**Є. І. Ткач, В. П. Сторожук**

# ЗАГАЛЬНА ТЕОРІЯ СТАТИСТИКИ

*3-тє видання*

ПІДРУЧНИК

*Затверджено*
*Міністерством освіти і науки України*
*як підручник для студентів*
*вищих навчальних закладів*

Київ
"Центр учбової літератури"
2009

ББК 65.051я73
УДК 330.101.52 (075.8)
Т 48

*Гриф надано*
*Міністерством освіти і науки України*
*(Лист № 14/18.2-688 від 07.04.2004)*

**Рецензенти:**

***Войнаренко М. П.*** – доктор економічних наук, професор Хмельницького національного технологічного університету “Поділля”;

***Цвєткова Л. О.*** – кандидат економічних наук, доцент Полтавського університету споживчої кооперації України.

Т 48 Ткач Є. І., Загальна теорія статистики: ***підручник [для студ. вищ. навч. закл.] /*** Ткач Є. І., Сторожук В. П. – [3-тє вид.] – К.: Центр учбової літератури, 2009. – 442 с. – ISBN 978-966-364-892-7

У підручнику “Загальна теорія статистики” розглянуто теоретичні і методичні основи побудови статистичних показників, які використовують для вивчення закономірностей суспільних явищ з урахуванням міжнародних стандартів статистики та обліку. Особливу увагу приділено статистичній методології, можливостям її використання в умовах суттєвих змін в економіці.

Для студентів економічних спеціальностей вищих закладів освіти, слухачів інститутів і факультетів післядипломної освіти, підприємців, менеджерів, економістів.

ББК 65.051я73
**ISBN 978-966-364-892-7**
УДК 330.101.52 (075.8)

# ПЕРЕДМОВА

Інтерес до статистики постійно зростає в усьому світі. Праця економіста любої спеціалізації неминуче зв’язана із збиранням, обробкою і аналізом статистичних матеріалів. Тому вивчення і оволодіння статистичною наукою при підготовці економістів високої кваліфікації має велике значення в системі вищої економічної освіти.

В нашій країні увага до статистичної науки надзвичайно загострена у зв’язку з проведенням економічних реформ, які зачіпають інтереси всіх людей.

Для підняття статистики до сучасного наукового рівня, задоволення потреб системи управління та інших соціально-економічних суб’єктів в якісній, повній, різноманітній і своєчасній інформації, вкрай необхідна докорінна її перебудова.

Важливою умовою правильного сприйняття і практичного використання статистичної інформації, кваліфікованих висновків і обґрунтованих прогнозів є знання статистичної методології вивчення кількісної сторони соціально-економічних явищ, природи масових статистичних сукупностей, пізнавальних властивостей статистичних показників, умов їх застосування в економічному дослідженні.

Одним із основних завдань статистики є оптимізація звітності, приведення об’єму інформації до сучасних потреб системи управління в умовах переходу до ринку. Потрібно впроваджувати замість суцільної звітності вибіркові обстеження, одноразові обліки чи опитування, що приведе до оперативного і поглибленого аналізу.

Забезпечення надійності і достовірності статистичної інформації можливе через підвищення наукового рівня всієї статистичної методології, наближення її до методології і стандартів світової статистичної практики.

В даний час перед статистикою стоять проблеми подальшого вдосконалення системи показників, прийомів і методів збору, обробки, зберігання і аналізу статистичної інформації. Це має важливе значення для розвитку і підвищення ефективності автоматизованих систем управління, створення автоматизованих банків даних, розподільчих банків даних, які в свою чергу могли б сприяти створенню Єдиної статистичної інформаційної системи (ЄСІС), що надасть можливість запровадити в практику сучасні статистичні методи аналізу, імітаційні та прогнозні методи.

В запропонованому підручнику міститься системний виклад загальних категорій, принципів і методів статистичної науки, теоретичних основ економіко-статистичних методів аналізу і прогнозування із застосуванням кореляційно-регресійного, табличного і графічного методів.

В підручнику “Загальна теорія статистики” послідовно розглядаються питання, які виникають на стадії статистичного спостереження, зведення і групування матеріалів спостереження та їх наступної обробки.

Даний підручник написаний для студентів економічних вузів і факультетів.

# Розділ 1. Предмет і метод статистики

## 1.1. Статистика як наука

Перехід до ринкової економіки ставить перед економістами всіх рівнів підвищені вимоги до їх статистичної підготовки. Оволодіння статистичною методологією є однією з основних і головних умов пізнання кон'юнктури ринку, вивчення тенденцій і прогнозування попиту і пропозиції, прийняття науково обґрунтованих рішень на різних рівнях комерційної діяльності на ринку товарів і послуг.

Термін "статистика" походить від латинського слова "статус" (status), що означає суму знань про державу.

В даний час статистика має три основних значення: 1) під статистикою розуміють практичну діяльність працівників статистичних органів, які збирають, обробляють і аналізують дані про соціально-економічний розвиток країни в цілому, а також окремих її регіонів, галузей економіки, конкретних підприємств і населення; 2) статистикою вважають статистичні дані подані в звітах підприємств, організацій і установ, а також опубліковані в статистичних збірниках, довідниках і періодичній пресі; 3) статистикою називають спеціальну науку, яка займається розробкою теоретичних положень і методів її практичного використання.

Між статистичною наукою і практичною діяльністю існує тісний зв'язок і взаємозалежність. Статистична наука використовує інформацію практичної діяльності господарських організацій, узагальнює її і розробляє методи проведення статистичних досліджень. В свою чергу, підприємства, організації і установи для практичної діяльності використовують теоретичні розробки і положення статистичної науки для розв'язання конкретних управлінських завдань.

Статистика являється складною і багатогранною наукою, яка вивчає суспільні явища і процеси з їх кількісної сторони. Як навчальна дисципліна статистика включає в себе цілий ряд розділів, таких як загальна теорія статистики, соціально-економічна статистика і серію галузевих статистик (статистику промисловості, статистику сільського господарства, статистику капітальних вкладень, статистику торгівлі, статистику матеріально-технічного постачання, статистику транспорту, статистику фінансів і кредиту, статистику права і ряд інших). Таким чином, курс "Загальна теорія статистики" відкриває цикл статистичних дисциплін.

Значення статистики в системі знань, які формують профіль економістів вищої кваліфікації, менеджерів, комерсантів надзвичайно велике. Кожний економіст вищої кваліфікації повинен вміти читати статистичні цифри, аналізувати їх і використовувати в своїй практичній роботі для обґрунтування своєї позиції і своїх висновків.

Статистика має свою багатовікову історію. З утворенням держав виникла потреба в статистичному зборі і підрахунку таких важливих для державного керівництва даних як чисельність населення, худоби, майна, земельних угідь і ін. Такий облік проводився ще до нашої ери в Китаї, Стародавньому Римі і Єгипті.

З розвитком суспільного виробництва, внутрішньої і зовнішньої торгівлі збільшувалась потреба в статистичній інформації, що розширювало сферу діяльності статистики і вимагало постійного вдосконалення її прийомів і методів. Різноманітна практика обліково-статистичних робіт почала піддаватись різним теоретичним узагальненням. Статистика стала формуватись як наука.

Вважається, що основи статистичної науки заклав англійський економіст В.Петті. Він замість словесних порівнянь використав новий спосіб доведення мовою чисел, ваги, мір тощо. Створений В.Петті новий напрямок отримав назву “політична арифметика”.

Представники політичної арифметики Д.Граут, П.Зюсмильх та ін. внесли вагомий вклад в розвиток демографії, ввели в науковий обіг таблиці і графіки. Дюто і Карлі при вивченні динаміки цін вперше використали індекси.

Основоположником іншого напрямку розвитку статистики вважається німецький вчений Г.Конрінг, який розробив систему опису державного устрою. Його послідовник професор Г.Ахенвань вперше в Марбургському університеті (1746 р.) почав читати нову дисципліну, яку назвав статистикою. Основним змістом цього курсу був опис політичного стану і достознаменностей держав.

М.В.Ломоносов в багатьох своїх працях розглядав питання населення, природних багатств, фінансів, торгівлі Росії, ілюстрував їх статистичними даними. Цей напрямок розвитку статистики називався описовим.

Професор Геттінгенського університету А.Шліцер вважав, що предметом статистики є не тільки опис політичного устрою держав, а все суспільство.

Бельгійський статистик А.Кетле вніс значний вклад в розробку теорії стійкості статистичних показників.

Математичний напрямок в статистиці розвинули такі вчені як Ф.Гальтон, К.Пірсон, В.Госсет, Р.Фішер, М.Мітчел і ін. Наприклад, К.Пірсон вніс значний вклад в розробку теорії кількісної оцінки зв’язку між явищами, В.Госсет (псевдонім Стьюдент) розробив теорію малої вибірки, Р.Фішер розвивав методи кількісного аналізу, М.Мітчел заснував ідею “економічного барометра”.

Подальший розвиток статистики вимагав вдосконалення методів збирання, обробки і узагальнення масових даних. Фундатором теорії статистики вважається А.Кетле, в працях якого простежується пошук філософських підвалин статистики. Він вважав, що предметом статистики є “людина в суспільстві”, а методологічною основою – принцип масовості, який пізніше отримав назву ”закону великих чисел”. Саме цей принцип для узагальнення характеристик сукупності зумовив необхідність обчислення середніх величин.

В Росії в кінці XIX століття статистику політико-економічного напрямку потіснив інтенсивний розвиток математичної статистики. Це привело до виділення двох концепцій наукового змісту статистики:

1) статистика як метод пізнання, яку підтримували А.А.Чупров, А.А.Кауфман, Н.А.Каблуков, Н.К.Дружинін та ін.;

2) статистика як наука, предметом якої є дослідження масових явищ і процесів. Представниками цієї концепції є: Ю.Є.Янсон, А.Ф.Фортунатов, В.С.Німчинов, Й.С.Пасхавер та ін.

Кожна з цих концепцій відображала лише одну сторону статистики, оскільки статистика одночасно є наукою із своїм методом. Дискусія вчених статистиків в кінці 50-х років XX століття завершилась визнанням статистики суспільною наукою.

Розвиток електронно-обчислювальної техніки сприяв для прогресу статистичної методології і поглибленого дослідження соціально-економічних процесів.

Історичний розвиток статистики узагальнений в працях вчених, таких як В.І.Хотимський, В.С.Німчинов, М.В.Птуха та ін. Вдосконалення методології статистичного вивчення соціально-економічних явищ і процесів висвітлене в працях вчених: А.І.Ротштейна, Д.В.Савінського, А.І.Газулова, П.П.Маслова, Т.В.Рябушкіна, Г.І.Бакланова, Г.С.Кільдішева, В.Є.Овсієнка, Н.Н.Ряузова, В.Є.Адамова та ін.

В даний час видаються підручники, монографії, наукові праці зі статистики, в тому числі найбільш цікаві праці іноземних статистиків, а також різні статистичні збірники. Вивчається все корисне, розроблене статистичною наукою в інших країнах.

### 1.2. Предмет статистики

Кожна наука володіє своїми суттєвими специфічними особливостями, які відрізняють її від інших наук і дають їй право на самостійне існування як осібної галузі знань. Головна особливість всякої науки заключається в предметі пізнання, в принципах і методах його вивчення, які в сукупності утворюють його методологію.

Статистика, як суспільна наука, вивчає кількісну сторону найрізноманітніших масових явищ і процесів суспільного життя.

Статистика встановлює кількісні характеристики різних показників суспільного життя, їх розвиток, визначає співвідношення між окремими показниками, дає цифрову оцінку закономірностям, які проявляються в них.

Статистика також вивчає вплив природних і технічних факторів на зміну кількісних характеристик соціально-економічних явищ і процесів.

Вивчаючи кількісну сторону явищ і процесів, статистика відображає її в своїх цифрових показниках, характеризуючи конкретну міру явищ, встановлює загальні властивості, виявляє схожість і відмінність окремих рис, об'єднує елементи в групи, виявляє певні типи явищ.

Вивчення кількісної сторони суспільних явищ і процесів нерозривно пов'язане з їх якісним змістом, оскільки кількісна розмірність не існує без якісної визначеності. Наприклад, при групуванні міст України за чисельністю жителів статистика виділяє наступні типи: малі міста (до 50 тис. чол.), середні (51 – 100 тис. чол.), великі (101 – 500 тис. чол.), дуже великі (501 – 1000 тис. чол.), міста мільйонери (понад 1000 тис. чол.). Однак, перш ніж проводити розрахунки, потрібно встановити якісні властивості і межі кожного типу міст.

Отже, предметом статистики є наука, яка вивчає кількісну сторону масових явищ соціально-економічного життя у нерозривному зв'язку з їх якісним змістом в конкретних умовах місця і часу.

Явища суспільного життя динамічні, протягом певного часу змінюються їх розміри, співвідношення і пропорції.

Явищам для окремих об'єктів і регіонів властиві безперервні зміни і розвиток, а тому їх кількісну сторону статистика повинна вивчати в залежності від конкретних умов простору і часу.

Іншою особливістю предмета статистики є можливість суспільних явищ які вона вивчає, які складаються з певної множини елементів, істотні властивості яких схожі. Так, студенти схожі тим, що вчаться, робітники – виконують фізичну працю, лікарі – лікують хворих та ін.

Наявність будь-якої властивості у окремого елемента є випадковістю, а при об'єднанні їх в єдине ціле вплив випадковостей нівелюється.

Розглядаючи кількісну сторону масових суспільних явищ статистика за допомогою чисел показує ступінь їх розвитку, напрям і швидкість змін, тісноту зв'язку і взаємозалежностей. Все це підтверджує, що статистика є одним із засобів пізнання суспільного життя.

Як суспільна наука статистика не може розвиватися поза розвитком теоретичних наук про суспільство, зокрема історичного матеріалізму, економічної теорії і соціології, які досліджують і формулюють закони розвитку соціально-економічних явищ, виясняють їх природу і значення в житті суспільства. Спираючись на знання положень економічної теорії, статистика формує статистичні сукупності, встановлює суттєві ознаки для виділення соціально-економічних типів, здійснює розробку відповідних методів для їх вивчення.

Використовуючи положення, розроблені економічною теорією, статистика збагачує економічні науки фактами, отриманими в статистичних дослідженнях, підтверджує або відхиляє їх теоретичні догми.

Опираючись на статистику економічна теорія розробляє і формулює закони розвитку соціально-економічних явищ і процесів.

Статистика проводить різноманітні прогнозні розрахунки, які використовуються для обґрунтування напрямків соціально-економічної політики.

Всі економічні науки використовують статистичну інформацію для перевірки, обґрунтування і ілюстрації своїх теоретичних розробок і положень.

## 1.3. Основні поняття в статистиці

Явища і процеси суспільного життя вивчаються за допомогою статистичних показників.

*Статистичний показник* – це кількісна оцінка властивості досліджуваного явища. В залежності від цільового призначення, статистичні показники ділять на обліково-оцінні і аналітичні.

*Обліково-оцінні показники* характеризують розмір якісно визначених соціально-економічних явищ і процесів в конкретних умовах місця і часу.

В залежності від специфіки досліджуваного явища, обліково-оцінні показники відображають або об'єми їх розповсюдження в просторі, або досягнені рівні розвитку на певні моменти часу.

За своєю архітектонікою статистичний показник складається з двох частин. Перша частина виражає суть поняття, а друга –

характеризує його величину. Наприклад, чисельність населення України на 1 січня 1996 р. становила 51,3 млн. осіб. Тут чисельність населення України є першою частиною, а 51,3 млн. осіб – другою частиною статистичного показника.

*Аналітичні показники* використовуються для аналізу статистичної інформації і характеризують особливості розвитку досліджуваного явища: типовість ознаки, співвідношення окремих його частин, ступінь розповсюдження в просторі, швидкість розвитку в часі і т.д. В якості аналітичних показників в статистиці застосовуються відносні і середні величини, показники варіації, динаміки і тісноти зв'язку, індекси і ін.

*Статистична ознака* – це характерна властивість досліджуваного явища, яка відрізняє його від інших явищ.

Статистичні показники виражаються смисловими поняттями і числовими значеннями. Показники виражені смисловими поняттями називаються *атрибутивними*. Наприклад, стать людини – чоловік і жінка; спеціальність працівника – лікар, вчитель, інженер, економіст і т.д. Якщо атрибутивні ознаки приймають тільки одне з двох протилежних значень, їх називають *альтернативними*. Наприклад, продукція якісна – неякісна, стандартна – нестандартна, товарна – нетоварна і т.д.

*Кількісні ознаки* завжди виражені числовими значеннями. Наприклад, вік людини, число членів сім'ї, розмір квартири, середня заробітна плата одного члена сім'ї і ін.

В конкретних статистичних дослідженнях ознаки діляться на *основні* (суттєві), які визначають головний зміст досліджуваного явища, і *другорядні*, не зв'язані безпосередньо з їх основним змістом. Наприклад, при вивченні прибутку від реалізації товарної продукції оптові ціни і собівартість будуть основними факторами, які впливають на зміну його об'єму.

Важливою особливістю статистики є те, що для вивчення свого предмету вона створює статистичну сукупність.

*Статистична сукупність* – це певна множина одиниць досліджуваного явища, об'єднаних у відповідності із завданням дослідження єдиною якісною основою (умовами існування і розвитку), але різняться між собою за деякими іншими ознаками. Наприклад, при визначенні об'єму роздрібного товарообороту всі торгові підприємства, які продають товари населенню, розглядаються як єдина статистична сукупність – роздрібна торгівля. Але за розміром обсягу реалізації товарів, торговою спеціалізацією, формами обслуговування покупців та іншими ознаками комерційної діяльності, одиниці даної статистичної сукупності будуть різнорідними.

Статистична сукупність називається однорідною, якщо суттєві ознаки для кожної її одиниці є в основному однакові, і різнорідною, якщо вона об'єднує різні типи явищ.

Ознака, яка приймає в межах однієї статистичної сукупності різні значення, називається варіаційною, а коливання значень ознаки – *варіацією*.

Статистичним сукупностям властиві закономірності масових соціально-економічних явищ і процесів.

*Статистична закономірність* – це певна впорядкованість, відносна постійність, повторюваність, послідовність і порядок в явищах.

Об'єктивною основою існування статистичних закономірностей є складне переплетення основних (тобто, загальних для всіх явищ) факторів, які формують масовий процес та індивідуальних для кожного з них окремо, але випадкових для маси. За великим числом подій випадкові причини взаємно урівноважуються (погашаються).

Статистичні закономірності масових соціально-економічних явищ і процесів відображають характер дії об'єктивних законів розвитку суспільства в конкретних умовах місця і часу. Отримуючи узагальнюючі характеристики масових суспільних явищ, статистика вияснює з їх допомогою певні закономірності, які проявляються в зміні статистичних даних, їх співвідношенні, розподілі і взаємозв'язку. Це можуть бути закономірності: а) розвитку (динаміки) явищ, що проявляються в збільшенні чисельності населення Землі, рості середньої тривалості життя, збільшенні виробництва продукції, зниженні собівартості виробництва продукції і т.д.; б) зміни структури явищ, що простежуються в збільшенні частки міського населення в його загальній чисельності, рості частки реалізації товарів культурно-побутового призначення в загальному обсязі товарообороту та ін.; в) розподілу одиниць у середині сукупності, про що свідчать розподіли населення за ростом, вагою, розміром одягу чи взуття, доходом на душу населення, кількістю членів сім'ї і т.д.; г) зв'язку між явищами, який існує між кількістю внесених добрив і урожайністю; між вартістю основних виробничих фондів і випуском продукції; між продуктивністю праці і собівартістю виробленої продукції і багато інших.

Характерною особливістю всіх названих закономірностей є те, що вони описуються узагальнюючими статистичними показниками, які формуються під дією складного комплексу причин. Статистичні закономірності відображають в різноманітних формах прояв економічних законів в цифрах в конкретних умовах місця і часу. Ці закономірності проявляються тільки в масових явищах і процесах, тільки у великих за чисельністю одиниць статистичних сукупностях. Тому статистичні закономірності безпосередньо зв'язані з дією закону великих чисел.

*Законом великих чисел* називається принцип, відповідно до якого закономірність масових явищ може проявлятися тільки при достатньо великому числі випадків.

Закон великих чисел створює умови для прояву статистичних закономірностей, виражає пряму залежність певного їх прояву від числа спостережень.

### 1.4. Метод статистики

Для своїх досліджень статистика розробляє і застосовує комплекс методів і засобів, сукупність яких формує *статистичну методологію*. Застосування в статистичному дослідженні конкретних методів визначається поставленими завданнями і залежить від характеру вихідної інформації.

Загальною основою розробки і застосування статистичної методології є положення соціально-економічної теорії і принцип діалектичного методу пізнання явищ суспільного життя. Вони складають теоретичну базу статистичної науки.

Теоретичний аналіз досліджуваного явища, заснований на соціально-економічних науках, завжди передує його статистичному вивченню і є необхідною умовою правильної організації статистичного дослідження та безпомилкових висновків за його результатами.

Необхідною умовою статистичного вивчення є також розуміння суті досліджуваного явища чи процесу, знання законів розвитку і особливостей конкретної обстановки.

Одночасно, керуючись положеннями економічної теорії, статистика збагачує соціально-економічні науки фактичними даними, отриманими в статистичному дослідженні.

У відповідності з діалектичним методом пізнання статистика вивчає всі явища і процеси в їх взаємозв'язку і взаємозалежності, в русі і зміні, виділяючи їх різні типи і форми, встановлює нове і прогресивне та визначає напрямки розвитку.

В процесі розвитку, поряд з кількісними змінами в досліджуваних явищах, відбуваються корінні якісні зміни, що потребує володіння методами, які дозволяють аналізувати такі зміни, фіксувати перехід кількісних змін в якісні.

В своїх дослідженнях статистика спирається на діалектичні категорії – випадкового і необхідного, одиничного і масового, індивідуального і загального.

Весь арсенал статистичних методів вивчення соціально-економічних явищ і процесів систематизується за їх цільовим призначенням і застосуванням, які знаходять свій вираз в трьох основних стадіях (етапах) економіко-статистичного дослідження.

На першому етапі статистичного дослідження збираються масові первинні дані шляхом реєстрації фактів чи опитування респондентів. Вимога масовості одиниць спостереження зумовлюється тим, що досліджувані статистикою закономірності проявляються лише в достатньо великому обсязі даних на основі дії закону великих чисел.

На другому етапі статистичного дослідження зібрані в процесі спостереження дані підлягають систематизації і групуванню. Це дозволяє виділити в досліджуваній сукупності соціально-економічні типи. Основна суть цього етапу статистичного дослідження полягає в переході від характеристик одиничного до загального. Якісна однорідність груп соціально-економічних явищ – одна з неодмінних умов узагальнюючих статистичних показників.

На третьому, завершальному етапі статистичного дослідження проводиться аналіз матеріалів зведення і групування за допомогою узагальнюючих статистичних показників таких, як абсолютні, відносні і середні величини, показники варіації, індекси, різні статистичні коефіцієнти і ін.

Аналіз отриманої статистичної інформації дозволяє розкрити причинні зв'язки досліджуваних явищ, визначити вплив різних факторів, оцінити ефективність прийнятих управлінських рішень і т.д.

Для аналізу і ілюстрації статистичної інформації широко застосовують статистичні таблиці і графіки.

Специфічний статистичний метод заснований на поєднанні аналізу і синтезу. Спочатку виділяються в складі досліджуваного явища окремі групи і підгрупи, вивчається суть відмінностей у величинах окремих ознак, виявляються причини відмінностей, а потім дається характеристика явища в цілому, у всій сукупності його сторін, тенденцій і форм розвитку.

Статистичні методи тісно пов'язані з математикою. В них спільні методи обробки і оцінки даних, але різні предмети пізнання. Математична статистика вивчає закономірності масових явищ і процесів в абстрактній формі, а статистика, як суспільна наука, характеризує суспільні явища в конкретних умовах їх існування і розвитку.

Статистика широко використовує такі математичні методи як аналіз варіаційних рядів, кореляційний і регресійний аналіз, теорію ймовірностей, методи оптимального програмування, теорію розпізнавання образів та ін.

Значення математики для розвитку статистичних методів різко підвищилось в сучасних умовах у зв'язку з широким використанням електронної обчислювальної техніки. Це дозволило значно скоротити обробку і передачу даних, впорядкувати зберігання інформації, полегшило і прискорило пошук необхідних даних у великих масивах.

В ході історичного розвитку статистичної науки в її складі відокремились ряд самостійних статистичних дисциплін: загальна теорія статистики, соціально-економічна статистика і ряд галузевих статистик.

### 1.5. Значення і основні завдання статистики

При переході України до ринкових відносин значення обліку і статистики значно зростає.

Велика роль статистики в управлінні народним господарством. Тільки через статистику органи управління отримують всесторонню характеристику об'єкта, яким управляють, незалежно чи це народне господарство в цілому, чи окремі його галузі або підприємства. Статистика сигналізує про неблагополучні моменти в окремих місцях механізму управління, показуючи таким чином потребу зворотного зв'язку – управлінських рішень.

Велике значення статистики для планування. Перш ніж планувати розвиток об'єкта, потрібно мати цифрові дані про цей об'єкт.

В даний час планування є багатоваріантним, тому його потреба в статистичних даних постійно зростає.

Статистика показує, як виконується план на окремому підприємстві, об'єднанні, галузі і народному господарству в цілому, виявляє можливі диспропорції в народному господарстві, що служить сигналом для керівних органів про необхідність їх усунення. За допомогою статистики чітко прослуховується пульс економіки, що дозволяє постійно контролювати хід виконання народногосподарських програм, виявляти резерви підвищення народногосподарського виробництва.

Корінним питанням здійснення радикальної економічної реформи в Україні є перехід від командно-адміністративних форм управління до ринково-економічних. Це ставить перед статистикою як складовою частиною системи управління народним господарством нові завдання, основними з яких на сучасному етапі її розвитку є:

1) всебічне дослідження глибоких перетворень економічних і соціальних процесів в суспільстві на основі використання науково обґрунтованої системи показників;

2) узагальнення, планування і прогнозування тенденцій розвитку народного господарства;

3) виявлення резервів росту ефективності суспільного виробництва;

4) своєчасне і повне забезпечення надійною інформацією законодавчих органів влади, управлінських, виконавчих і господарських органів та широкої громадськості.

Для статистики як фактора формування суспільної свідомості особливе значення має суттєве розширення гласності і доступності зведеної статистичної інформації при збереженні принципу конфіденційності індивідуальних даних. Це один з необхідних напрямків демократизації суспільства.

Головні напрямки розвитку статистики полягають у вдосконаленні аналізу статистичної інформації, впорядкуванні звітності та забезпеченні її вірогідності.

Засобом підвищення вірогідності статистичної інформації є подальше вдосконалення методології її формування.

Перед статистикою постають важливі проблеми теоретичного обґрунтування об’єму і структури статистичної інформації, яка відповідала б сучасним і перспективним умовам розвитку економіки.

Потрібно вирішити питання про перехід від суцільної звітності до несуцільних видів статистичного спостереження, що підвищить оперативність реагування на кон’юнктурні зміни, забезпечить управління інформацією для прийняття своєчасних рішень.

Особливе значення для статистичної науки має посилення прогнозованості аналітичної роботи. Вона повинна містити елементи передбаченості і виявлення критичних точок зміни явища, вказувати на можливі наслідки ситуації, що складається.

Перехід до ринкової економіки України зумовлює крайню необхідність впровадження в статистичний і бухгалтерський облік системи національних рахунків. У зв’язку з цим дуже важливо розвивати і поглиблювати прогресивні контакти з міжнародними статистичними службами ООН, з її Статистичною комісією.

Статистична комісія ООН здійснює розробку методології статистичних праць, порівняльності показників, готує рекомендації для Статистичного бюро Секретаріату ООН, координує роботу спеціалізованих статистичних органів ООН, здійснює консультації з питань збору, нагромадження, розробки, аналізу і розповсюдження статистичної інформації.

Статистичне бюро Секретаріату ООН, як виконавчий орган, збирає статистичну інформацію від держав-членів ООН, публікує ці дані, готує доповіді з різних питань статистики.

Регіональні економічні комісії розглядають питання статистики для Європи, Азії, Африки і Латинської Америки.

Міжнародним статистичним органом є Міжнародний статистичний інститут, який здійснює узагальнення наукових досліджень в області теорії і методології статистики.

## 1.6. Сучасна організація статистики в Україні

Вивченням економічного і соціального розвитку країни, окремих її регіонів, галузей, об'єднань, фірм, підприємств займаються спеціально створені з цією метою органи, які в сукупності називаються *статистичною службою*.

В Україні функції статистичної служби виконують органи державної статистики і органи відомчої статистики.

Організація державної статистики в країні та її завдання видозмінювались у відповідності із зміною органів державного управління, їх функцій, з врахуванням особливостей розвитку економічного і соціального життя суспільства у відповідні історичні періоди.

В даний час головним обліково-статистичним центром в країні є Державний комітет статистики України (Держкомстат України). Він здійснює керівництво статистикою України у відповідності із статтею 92 п. 12 Конституції України "Виключно законами України визначаються: організація і діяльність органів виконавчої влади, основи державної служби, організації державної статистики та інформатики". В його завдання входить подання офіційної статистичної інформації Президенту, Уряду, Парламенту, громадським і міжнародним організаціям, розробка науково обґрунтованої статистичної методології, координація статистичної діяльності регіональних органів виконавчої влади, аналіз економіко-статистичної інформації, складання національних рахунків і балансових розрахунків.

Система органів державної статистики утворена у відповідності з адміністративно-територіальним розподілом України. В Автономній Республіці Крим діє Державний комітет по статистиці, в областях –

обласні управління статистики, а в районах і містах – відділи статистики.

Державний Комітет Статистики України складається з управлінь, всередині яких є ряд відділів.

ТИПОВА СТРУКТУРА ДЕРЖАВНОГО КОМІТЕТУ СТАТИСТИКИ УКРАЇНИ СТАНОМ НА ЛИСТОПАД 1997 РОКУ.

I. Управління зведеної інформації та регіональної статистики:
- відділ регіональної статистики та ринкових перетворень.

II. Управління статистичної методології та планування:
- відділ державних реєстрів та статистичних класифікаторів;
- відділ організації статистичного обліку;
- відділ планування та координації методів статистичних спостережень.

III. Управління національних рахунків:
- відділ зведених національних рахунків;
- відділ короткострокових розрахунків валового внутрішнього продукту;
- відділ рахунків споживання та нагромадження;
- відділ рахунків виробництва та утворення доходу;
- відділ міжгалузевого балансу.

IV. Управління статистики умов життя населення:
- відділ зведених робіт та моніторингу соціальної сфери;
- відділ обстеження домашніх господарств;
- відділ статистики соціальної інфраструктури та житлових умов населення;
- відділ моральної статистики.

V. Управління статистики праці:
- відділ зведених робіт та одноразових обстежень;
- відділ статистики оплати праці;
- відділ статистики зайнятості.

VI. Управління статистики населення:
- відділ демографічної статистики;
- відділ перепису населення.

VII. Управління статистики промисловості:
- відділ зведених робіт;
- відділ статистики виробництва продукції;
- відділ статистики військово-промислового комплексу та конверсії;
- відділ статистики науково-технічного прогресу.

VIII. Управління статистики сільського господарства та навколишнього середовища:
- відділ зведених робіт;
- відділ статистики землеробства;
- відділ статистики тваринництва;
- відділ статистики сільськогосподарських підприємств;
- відділ статистики природних ресурсів та навколишнього середовища;
- відділ статистики заготівель.

IX. Управління статистики будівництва та основних фондів:
- відділ зведених робіт;
- відділ статистики інвестицій;
- відділ статистики будівництва;
- відділ статистики підрядних організацій;
- відділ статистики основних фондів.

X. Управління статистики послуг, транспорту та зв’язку:
- відділ статистики транспорту, зв’язку та інформатики;
- відділ статистики послуг внутрішнього ринку;
- відділ статистики послуг зовнішньоекономічної діяльності.

XI. Управління статистики торгівлі:
- відділ статистики товарних ринків;
- відділ статистики роздрібної торгівлі та громадського харчування;
- відділ статистики зовнішньої торгівлі.

XII. Управління статистики фінансів:
- відділ статистики фінансів підприємств та організацій;
- відділ статистики державних фінансів та грошово-кредитної системи.

XIII. Управління статистики цін:
- відділ зведених робіт та міжнародних співставлень;
- відділ статистики споживчих цін;
- відділ статистики цін виробників.

XIV. Управління статистики зарубіжних країн та міжнародного співробітництва:
- відділ статистики зарубіжних країн;
- відділ міжнародного співробітництва;
- відділ протоколу.

XV. Управління кадрів і навчальних закладів:
- відділ кадрів;
- відділ навчальних закладів.

XVI. Перший відділ.

XVII. Фінансово-економічне управління:
- фінансово-економічний відділ;
- головний бухгалтер.

XVIII. Управління інформатизації.

XIX. Управління справами:
- відділ діловодства та контролю виконання;
- відділ матеріально-технічного забезпечення, охорони праці та господарського обслуговування.

Структура місцевих органів державної статистики складається з ряду відділів, таких як соціальної статистики, статистики праці, демографічної статистики, статистики торгівлі, статистики промисловості та ін.

Поряд із загальнодержавною статистикою існує відомча статистика, яка обслуговує підприємства, об'єднання, відомства і міністерства. Вона виконує роботи, пов'язані з отриманням, обробкою і аналізом статистичної інформації, потрібної для керівництва і планування їх діяльності. Для ведення статистики на підприємствах, в об'єднаннях, концернах, асоціаціях, міністерствах утворені певні статистичні органи.

Значення відомчої статистики в даний час значно зросло у зв'язку з тим, що розвиток ринкової економіки, самостійність підприємств і повна відповідальність за результати виробничо-господарської діяльності вимагають більш глибокого аналізу економічних процесів, які відбуваються на підприємствах. Такий аналіз вимагає різноманітної статистичної інформації, яку отримують на основі первинного обліку і спеціально організованих обстежень.

Головним завданням відомчої статистики є забезпечення інформацією, яка характеризує виконання внутрівиробничих планів і програм, наявність внутрівиробничих резервів збільшення випуску продукції, покращення використання виробничого потенціалу.

Управляти складними соціальними і економічними системами можна лише володіючи оперативною, вірогідною і повною статистичною інформацією.

### *Контрольні запитання*

1. Що означає термін "статистика"?
2. Чим зумовлене виникнення і розвиток статистичної практики і науки?
3. Хто вважається основоположником статистичної науки?
4. Що є предметом статистичної науки?
5. Дайте визначення статистичного показника і вкажіть їх види.

6. Що характеризує статистична ознака?
7. Що ви розумієте під статистичною сукупністю?
8. Дайте визначення статистичної закономірності.
9. Поняття про закон великих чисел.
10. В чому заключається суть статистичної методології?
11. Перерахуйте стадії статистичного дослідження, розкрийте їх зміст.
12. Роль і значення математики в статистичному дослідженні.
13. Які основні завдання стоять перед статистикою на сучасному етапі розвитку суспільства?
14. Перерахуйте основні частини статистичної науки.
15. Які основні принципи організації статистики в України в даний час?
16. Які завдання стоять перед державною статистикою в умовах переходу до ринкової економіки?
17. Які завдання відомчої статистики та її роль в сучасних умовах?
18. Охарактеризуйте структуру органів Державної статистики в Україні.

# Розділ 2. Статистичне спостереження

## 2.1. Поняття про статистичне спостереження

Для отримання повних і точних відомостей про стан будь-якого явища на певний момент часу, або про результати його розвитку за відповідний період часу проводять статистичне дослідження, яке складається з трьох послідовних етапів: 1) статистичного спостереження; 2) зведення і групування матеріалів статистичного спостереження; 3) економічного аналізу даних, отриманих в результаті зведення і групування.

Етапи статистичного дослідження тісно пов’язані між собою, тому успіху можна досягти лише при добре підготовленій і організованій роботі на всіх його стадіях.

Всяке економічне дослідження починається з отримання вихідної інформації – первинного статистичного матеріалу, який формується в процесі статистичного спостереження і є фундаментом статистичного дослідження.

*Статистичне спостереження* – як перша стадія статистичного дослідження, являє собою планомірну, систематизовану, науково організовану роботу по збиранню і реєстрації масових первинних даних про явища і процеси суспільного життя.

Ці дані в залежності від мети статистичного дослідження можуть бути різні за своїм змістом і способами отримання. Вони пізніше систематизуються, групуються, обробляються, аналізуються і узагальнюються.

Саме статистичне спостереження також складається з трьох етапів: а) підготовки спостереження; б) збирання матеріалу; в) контролю зібраного матеріалу.

На підготовчому етапі статистичного спостереження, відповідно до його мети і завдань, розробляється програма і організаційний план проведення спостереження. Тут вирішують питання про зміст вихідної інформації, яким способом, якими засобами і в які терміни буде проведений облік фактів, як будуть організовані збирання і контроль отриманих первинних матеріалів. Повинні бути враховані також відповідні вимоги до оформлення цих матеріалів, яких вимагає техніка подальшої їх обробки на ЕОМ.

Від якості статистичного спостереження залежить успіх всього статистичного дослідження. Статистичне спостереження повинно бути організоване таким чином, щоб в результаті його проведення були отримані об’єктивні, вірогідні, повні дані про досліджуване явище і, по можливості, в короткий термін. Це дасть змогу зробити правильні узагальнення і висновки.

Соціально-економічні явища і процеси, які спостерігаються, повинні мати наукову і практичну цінність та виражати їх типи.

Одним з важливих завдань статистичного спостереження є ретельна і всебічна перевірка якості зібраних матеріалів для забезпечення їх вірогідності.

Наукова організація статистичного спостереження потрібна для створення найкращих умов для отримання об'єктивно правильних матеріалів, які б давали змогу передбачати майбутні ситуації і робити обґрунтовані прогнози.

Всі вище поставлені питання знаходять свій розв'язок в теорії статистичного спостереження.

## 2.2. Програмно-методологічні питання статистичного спостереження

Статистичне спостереження проводять за строго визначеним планом, який включає програмно-методологічні і організаційні питання.

До програмно-методологічних відносять питання, зв'язані з розробкою програми спостереження, визначенням мети, об'єкта і одиниці спостереження, проектуванням формулярів і текстів інструкцій, встановленням джерел і способів збирання інформації.

До організаційних відносять питання про органи спостереження, терміни і місце проведення спостереження, складання попередніх списків одиниць досліджуваної сукупності, розставлення і підготовка кадрів та деякі інші.

Програма статистичного спостереження визначається правильно встановленими і конкретно сформульованими завданнями дослідження. Тому, перш за все, потрібно чітко сформулювати мету всієї роботи, а потім вирішувати всі інші питання програми спостереження.

*Мета спостереження* являє собою основний результат статистичного дослідження. Чітке і конкретне формулювання мети спостереження потрібне для того, щоб не збирати зайвих, непотрібних і неповних даних.

Завдання статистичного дослідження необхідно відобразити в статистичних показниках, для чого розробляють і складають макети кінцевих статистичних таблиць, в які заносять результати всієї роботи.

При організації статистичного спостереження важливо точно визначити об'єкт спостереження.

Об’єктом статистичного спостереження називається сукупність одиниць досліджуваного явища, про які повинні бути зібрані потрібні статистичні дані. Об’єктом статистичного спостереження можуть бути, наприклад, чисельність населення України, кількість промислових підприємств, сукупність вищих навчальних закладів та ін.

Визначивши об’єкт статистичного спостереження потрібно вказати на його важливі ознаки і основні розпізнавальні риси, тобто встановити межі досліджуваної сукупності. Так, під час перепису населення недостатньо сказати, що об’єктом спостереження є “населення”, потрібно обов’язково вказати, яке саме населення: міське чи сільське, постійне чи наявне. При обстеженні промислового обладнання об’єктом спостереження виступає сукупність верстатів, машин, механізмів, двигунів і інше обладнання. Тут потрібно строго визначити межі між виробничим і енергетичним обладнанням, якщо підлягає перепису тільки виробниче обладнання, то яке саме: наявне, встановлене чи діюче і т.д.

При періодичному обстеженні потрібно слідкувати, щоб досліджувана сукупність була більш менш однорідною. Для цього статистика використовує *ценз* – обмежувальну ознаку, яку повинні задовольняти всі одиниці досліджуваної сукупності.

Поряд з визначенням об’єкта статистичного спостереження визначають також одиницю сукупності і одиницю спостереження.

*Одиницею спостереження* називають той первинний складовий елемент об’єкта статистичного спостереження, який є носієм ознак, що підлягають реєстрації.

*Одиницею сукупності* називається та первинна ланка, від якої отримують необхідні статистичні відомості.

Одиницю спостереження слід відрізняти від одиниці сукупності. Наприклад, при проведенні перепису промислового обладнання одиницею спостереження буде промислове підприємство, а одиницею сукупності – окремий верстат. Отже, одиниця сукупності – це те, що підлягає обстеженню, а одиниця спостереження – це джерело звідки отримують дані. Визначення одиниці сукупності важливе при розробці програми спостереження, а визначення одиниці спостереження – при вирішенні питань організації збору інформації.

Після того як визначені об’єкт, одиниця спостереження і одиниця сукупності, потрібно розробити зміст програми спостереження, що є основним питанням статистичного спостереження.

Програмою статистичного спостереження називається перелік чітко сформульованих питань, на які намічають отримати відповіді в процесі обстеження. Наприклад, в бланку перепису населення передбачені ознаки, які характеризують окрему людину і реєструються при проведенні перепису: стать, вік, освіта, професія, національність, сімейне положення і ін.

Розробка програми являється однією з важливих теоретичних і практичних проблем статистичного спостереження. Від якості її розробки залежать якість і цінність зібраного статистичного матеріалу.

Розробка програми спостереження справа досить складна і відповідальна, тому до їх розробки повинні залучатися практичні і наукові працівники з великим досвідом роботи, які безпосередньо заінтересовані в результатах даного статистичного спостереження.

Зміст програми статистичного спостереження визначається суттю і властивостями об’єкта спостереження, тому для успішної її розробки потрібно мати достатньо повну уяву про досліджуваний об’єкт і його складові елементи. Зміст програми залежить також від мети спостереження, потребою у відповідних статистичних даних для державного управління, господарського керівництва і наукових досліджень.

Обсяг програми спостереження залежить від засобів, якими розпоряджаються статистичні органи для проведення спостереження, від термінів збору інформації і інших чинників.

До програми спостереження і розробки її змісту ставиться ряд вимог, яким вона повинна відповідати при будь-якому статистичному дослідженні. Вона повинна містити тільки істотні ознаки, які безпосередньо характеризують досліджуване явище, його тип, основні риси і властивості.

В програму не потрібно включати другорядні питання, які перешкоджатимуть збору інформації, а в подальшому – її обробку і аналіз.

Розробляючи програму слід прагнути до повноти збору необхідної і доброякісної інформації за широкою програмою.

В програму спостереження повинні включатися тільки такі питання, на які дійсно можна буде отримати об’єктивні і достовірні відповіді.

Якщо при можливості і доцільності треба отримати інформацію за час, який передував назначеному обстеженню, то в програмі спостереження потрібно передбачити збір таких даних.

В програму спостереження включають також питання контрольного характеру з метою перевірки і уточнення матеріалів, які збираються.

Одночасно з програмою спостереження складають програму розроблення матеріалів спостереження, яка конкретизує завдання статистичного дослідження, тобто допомагає уточнити програму спостереження.

На розробку програми спостереження впливає вдосконалення автоматизації статистичних робіт, утворення автоматизованої системи державної статистики (АСДС), де зберігається банк даних.

Статистичним формуляром називається документ особливої форми куди збирають і записують відповіді на питання програми спостереження. Обов'язковим елементом статистичного формуляра є титульна і адресна частини, які необхідні для перевірки зібраних даних і їх наступного розроблення.

В титульній частині, як правило, проставляють назву статистичного спостереження, назву органу, який проводить спостереження, ким і коли затверджений формуляр, присвоєний йому номер.

В адресній частині записують точний адрес одиниці або сукупності одиниць спостереження і деякі інші відомості (кому підпорядковане підприємство відомості про яке записуються у формуляр спостереження, коли і куди треба висилати звітність, підписи відповідальних осіб, які стверджують правильність даних обстеження).

В практиці застосовують два види або дві системи статистичних формулярів: індивідуальну (карткову) і спискову.

*Індивідуальним* називається такий статистичний формуляр, в який заносять відомості про одну одиницю спостереження (листок обліку кадрів, одне підприємство, один робітник і т.д.).

*Списковим* називається такий статистичний формуляр, в якому реєструються відомості по декількох одиницях спостереження (відомість на заробітну плату, екзаменаційна відомість і ін.).

Обидві системи статистичних формулярів мають свої переваги і недоліки.

Індивідуальний формуляр, як правило, містить більше питань, так як в ньому характеризується тільки одна одиниця спостереження. При використанні ручного способу зведення і групування матеріалів картки легко групуються за потрібною нам ознакою. Недоліком цієї форми є те, що в ній кожний раз потрібно писати адресну частину, яка може бути загальною для кількох одиниць спостереження, наприклад, для членів однієї сім'ї і т.д. Карткова система більш трудомістка і вимагає зайвих матеріальних витрат (бланків, паперу і ін.).

Спискова система статистичних формулярів економніша, більш зручна для перевірки матеріалу і його машинної обробки. Однак, списковий формуляр може стати незручним або мало придатним, тому що із збільшенням числа питань він розростається до великих розмірів, що, в свою чергу, затушовує загальну тенденцію досліджуваного явища. Для його заповнення потрібно готувати спеціальних реєстраторів.

Застосування обчислювальних машин при обробці статистичного матеріалу висуває свої специфічні вимоги до формулярів спостереження. Відповіді у формулярах повинні бути розташовані таким чином, щоб він був одночасно і технічним носієм інформації, який безпосередньо сприймався б машиною.

Якими б розумними не здавались питання формуляра статистичного спостереження, до нього обов'язково складається інструкція, яка сприяє ефективному виконанню розробленої програми дослідження.

*Інструкцією* називається сукупність пояснень і вказівок, які стосуються головним чином програми статистичного спостереження. Інструкція може бути подана у вигляді окремого документа, або викладена на формулярі спостереження. В ній даються вказівки, як потрібно розуміти те чи інше питання опитувального листа, як потрібно записати відповіді на нього.

Часто в інструкції даються вказівки щодо змісту оргплану спостереження, зокрема про об'єкт і одиницю спостереження, про його завдання, про те, хто проводить спостереження, про час, терміни і критичний момент проведення спостереження і ін.

Інколи в інструкції наводяться різні приклади випадків, які можуть зустрітись в процесі проведення спостереження та приклади заповнення формулярів спостереження.

Інструкція повинна бути написана коротко, лаконічно і просто, вказівки мають бути ясними, конкретними і чіткими.

Якщо у формулярі зразу після питання даються деякі роз'яснення або перераховуються можливі відповіді на нього, то це називається *статистичною підказкою*. Вона може бути *повною*, коли наводиться вичерпний перелік можливих відповідей на питання, або ж *неповною* – якщо вказуються тільки деякі із можливих відповідей.

Об'єкт статистичного спостереження постійно змінюється з плином часу. В залежності від характеру досліджуваного об'єкта і показників, які визначаються, відомості можуть збиратись або за певний період часу (місяць, квартал, рік),або на певний момент часу (на початок і кінець року, на перше число кожного місяця і т.д.). Тобто, в залежності від часу спостереження всі показники діляться на інтервальні і моментні.

*Інтервальними* називаються показники, які характеризують виробництво продукції, затрати робочого часу на виконання якихось робіт, число народжених або померлих за певний період часу і т.п.

*Моментними* називаються показники, які характеризують чисельність одиниць досліджуваної сукупності і окремих її частин на якийсь момент часу. Наприклад, чисельність населення України на 1.01. кожного року, в тому числі чоловіків і жінок; чисельність працівників підприємства на початок і кінець кожного місяця; чисельність багаторічних насаджень, худоби, машин і обладнання на певну дату і т.д.

Від вдалого вибору часу спостереження залежить успіх всього дослідження в цілому. Це повинен бути, як правило, час, коли об’єкт спостереження знаходиться в найбільш характерному для нього стані. При цьому враховують наявність і стан комунікацій, які використовують для зв’язків з одиницями спостереження, а також між працівниками і органами, які проводять спостереження.

При спостереженні масових і рухомих об’єктів (перепис населення) вибирають час, коли об’єкт знаходиться в найменш рухомому стані (як правило, середину січня), що в значній мірі полегшує проведення спостереження і сприяє отриманню більш точних даних.

*Періодом спостереження* називається час, на протязі якого здійснюється реєстрація одиниць спостереження за встановленою програмою.

Період спостереження, як правило, невеликий проміжок часу, який залежить від багатьох факторів, таких як специфічні особливості об’єкта спостереження, його розміри, його стан в той чи інший час, наявність кадрів, їх кількість і якість, які можуть бути залучені до спостереження та ряд інших.

Термін проведення спостереження повинен бути максимально наближеним до критичного моменту.

*Критичним моментом* статистичного спостереження називається момент часу (день, година) до якого приурочується збір і реєстрація даних.

Термін спостереження обмежується датами, та часом і годинами початку і закінчення спостереження. Інколи вказується число днів, на протязі яких повинно бути проведене спостереження.

Для багатьох статистичних спостережень встановлюється строк (день тижня, число місяця чи кварталу), до якого інформація повинна бути зібрана і передана за призначенням.

### 2.3. Важливі організаційні питання статистичного спостереження

З метою успішного проведення спостереження складають організаційний план.

*Організаційний план* – це основний документ, в якому зосереджені розв'язки важливих питань організації і проведення статистичного спостереження.

До організаційних питань статистичного спостереження належать: визначення об'єкта, місця, часу і термінів спостереження; постановка мети і завдань спостереження; визначення органів спостереження; визначення прав і обов'язків окремих установ і організацій, які беруть участь у спостереженні; підготовчі роботи проведення спостереження; добір, навчання і інструктаж масових кадрів, потрібних для проведення спостереження; розмноження і розсилка формулярів спостереження; порядок здачі і приймання матеріалів спостереження; порядок отримання і подання попередніх і остаточних підсумків спостереження та інші практичні питання.

Організаційні плани складаються статистичними органами держави починаючи з вищих і закінчуючи нижчими ланками. В різних ланках ці плани дещо відрізняються як за обсягом, так і за змістом.

Вищі статистичні органи головну увагу приділяють розв'язку загальних організаційно-методологічних питань, таких як визначення завдань спостереження, його об'єкта, одиниці і термінів проведення і ін.

Нижчі ланки статистичних органів розв'язують в основному конкретні організаційні завдання на місцях.

Для прикладу візьмемо деякі положення організаційного плану підготовки і проведення перепису населення. В цьому плані завжди вказуються: терміни проведення і хто підлягає перепису; порядок проведення перепису в поїздах і автобусах далекого прямування, на морських і річкових суднах, на вокзалах, станціях, портах, причалах і закритих закладах; порядок проведення суцільного і вибіркового перепису; обов'язки обліковців напередодні перепису і порядок проведення контрольних обходів і ін.

В організаційному плані статистичного спостереження вказуються органи, які здійснюють підготовку і проведення спостереження, а також відповідають за його якість. В ньому чітко і конкретно визначають права і обов'язки кожного органу і їх взаємозв'язок в процесі проведення спостереження.

Статистичне спостереження в загальнодержавному масштабі організовує Державний комітет статистики України і їх місцеві органи.

Міністерства, відомства, наукові і інші установи проводять статистичні спостереження в основному локального характеру.

Інколи для проведення крупних статистичних спостережень утворюють тимчасові органи відповідних установ (різні відділи, сектори, інструкторські і лічильні ділянки і т.п.).

*Місце спостереження* – це місце де проводиться реєстрація фактів спостереження, які записуються у статистичних формулярах.

Питання про місце спостереження виникає лише при спеціально організованому спостереженні (як правило, при переписах населення) відносно об'єкта, одиниці якого можуть міняти місце свого перебування.

Вибір місця спостереження повинен забезпечити повноту охоплення об'єкта спостереження, високу якість фіксації даних і простоту проведення спостереження.

Питання вибору місця спостереження має суттєве значення в організації різних соціологічних досліджень.

Існують також особливі місця спостереження, куди особи повинні в обов'язковому порядку з'являтись для реєстрації фактів. Наприклад, відділи реєстрації актів громадянського стану. Хоча ці відділи не являються статистичними органами, однак тільки через них органи державної статистики ведуть спостереження за природним рухом населення.

## 2.4. Основні організаційні форми, види і способи статистичного спостереження

Вся різноманітність організаційних форм статистичного спостереження здійснюється в двох основних формах: 1) у формі звітності підприємств, організацій і установ; 2) у формі спеціально організованого спостереження (перепис населення, облік багаторічних насаджень, переоцінка основних фондів і т.п.).

*Звітністю* називають таку організаційну форму статистичного спостереження, за якою відомості поступають в статистичні органи від підприємств, організацій і установ у вигляді обов'язкових звітів про їх діяльність в точно встановлені терміни.

Керівництво підприємств, організацій і установ несе персональну відповідальність за своєчасність і достовірність поданої у звітах інформації.

Звітність являється важливою формою статистичного спостереження в Україні і має винятково велике значення як для практики, так і для науки. Вона охоплює все народне господарство і є обов’язковою для підприємств, організацій і установ всіх форм власності.

В звітах містяться основні обліково-статистичні дані про стан, розвиток і діяльність підприємств, організацій і установ, за допомогою яких контролюють виконання народногосподарських планів і програм, розробляють перспективні плани і прогнози та здійснюють оперативне керівництво.

Статистичну звітність складають на підставі даних первинного обліку.

*Первинним обліком* в статистиці називається ведення систематичних записів у формах первинних облікових документів про різні явища і процеси, які стосуються діяльності підприємств, організацій чи установ.

Звітність подають вищим організаціям і органам Державного комітету статистики в порядку, встановленому Державним комітетом статистики України щодо кожної форми.

В нашій державі розрізняють дві основні форми звітності: а) загальнодержавну; б) внутрівідомчу.

*Загальнодержавна* звітність обов’язкова для всіх підприємств, організацій і установ. Вона подається уряду міністерствами і відомствами у зведеному вигляді безпосередньо, або через Державний комітет статистики України.

*Внутрівідомча* – це звітність, розроблена міністерствами і відомствами для своїх оперативно-господарських потреб.

Підприємства, організації і установи для характеристики своєї фінансово-господарської діяльності складають і подають у певні інстанції декілька різних за змістом форм звітності і в неоднакові терміни. Тому, для кращої організації звітності, складають табель звітності, в якому проставляють номер і назву форми, спосіб відправлення, періодичність і терміни подання та установу, якій її подають.

Звітність в даний час є одним з основних джерел статистичної інформації про соціально-економічний розвиток держави.

Поряд із звітністю в практиці різних статистичних досліджень широко використовують спеціально організовані статистичні спостереження.

*Спеціально організованим статистичним спостереженням* називається таке спостереження, яке проводиться із спеціальною метою на якусь дату для отримання інформації, котру, в силу певних причин, не можна зібрати із звітів, або для перевірки і уточнення даних звітності.

Одним з основних видів спеціально організованого спостереження є переписи.

*Перепис* – це спеціально організоване статистичне спостереження великого масштабу, яке охоплює всю країну, або значну її частину і проводиться одночасно за єдиною програмою. Його метою є визначення чисельності, складу, стану і розміщення досліджуваного об'єкта на встановлений критичний момент. Прикладами такої форми статистичного спостереження можуть бути: переписи населення; обстеження бюджетів сімей; генеральні інвентаризації і переоцінка основних фондів; періодичний облік спеціалістів в народному господарстві і багато інших робіт органів статистики.

Спеціально організовані спостереження дають додатковий матеріал для народногосподарського планування, оперативних заходів і наукового пізнання закономірностей розвитку суспільства.

За повнотою охоплення спостереженням досліджуваного об'єкта розрізняють два його види: а) суцільне; б) несуцільне.

Суцільним називається таке спостереження, при якому обстеженню і реєстрації підлягають всі без винятку одиниці досліджуваної сукупності. Прикладом такого спостереження може бути повний перепис всього населення країни, облік всієї виробленої продукції підприємствами промисловості та ін.

Несуцільним називається таке спостереження, при якому обстеженню і реєстрації підлягають не всі одиниці досліджуваної сукупності, а лише певна їх частина. При проектуванні спостереження заздалегідь встановлюють, що воно буде несуцільним, яка саме частина сукупності буде спостерігатись і яким чином відбиратимуться ті чи інші одиниці для обстеження. Прикладами несуцільного спостереження є обстеження бюджетів сімей, реєстрація торгових оборотів і цін на продуктових ринках, визначення процента жирності молока та ін.

Несуцільні спостереження мають ту перевагу перед суцільними, що вони вимагають значно менше затрат сил і засобів, дозволяють застосовувати докладнішу програму і досконаліший спосіб обліку фактів, швидше підводити підсумки обстеження і, отже, підвищують оперативність статистичних матеріалів.

В багатьох випадках несуцільне спостереження є єдино можливим способом дослідження статистичної сукупності. Наприклад, аналіз якості продукції, перевірка загазованості повітря і забрудненості води і т.п.

Несуцільні спостереження в статистиці суттєво доповнюють основні матеріали, отримані в результаті суцільних спостережень.

В практиці статистичної роботи застосовують наступні види несуцільного спостереження: 1) вибіркове спостереження; 2) монографічне спостереження; 3) метод основного масиву; 4) анкетне.

*Вибірковим* називається таке спостереження, при якому вся сукупність фактів характеризується за деякою її частиною, відібраною випадково. Вибіркове спостереження в статистиці являється найбільш розповсюдженим видом несуцільного спостереження. В його основі лежить випадковий відбір одиниць для обстеження, що гарантує незалежність результатів вибірки від волі осіб, які її проводять і не допускає тенденційних помилок.

Вибіркове спостереження при правильній його організації і проведенні дає достатньо достовірні дані, які цілком придатні для характеристики всієї досліджуваної сукупності явищ. В багатьох випадках воно може повністю замінити суцільне спостереження. Оскільки вибірковому спостереженню відводиться окремий розділ, то докладніше про нього мова йтиме далі.

*Монографічне* спостереження являє собою детальне вивчення і опис окремого об’єкта, або невеликої їх кількості за розширеною програмою. Таке спостереження проводиться з метою виявлення певних тенденцій і закономірностей розвитку явища, або для вивчення і розповсюдження передового досвіду окремих підприємств, організацій і установ. Воно також використовується для виявлення недоліків в роботі окремих підприємств з метою їх усунення і недопущення в майбутньому.

*Метод основного масиву* заключається в тому, що з усієї сукупності одиниць спостереженню підлягає переважна їх частина, в яку, як правило, попадають найбільш суттєві і крупні одиниці досліджуваної сукупності. Взяті разом вони мають значну питому вагу в сукупності за однією чи декількома основними для даного дослідження ознаками. Одиниці сукупності, які не можуть істотно вплинути на характеристику сукупності в цілому із обстеження виключаються. Це дає значну економію матеріальних і трудових ресурсів. Прикладом такого методу спостереження є дослідження кон’юнктури товарообороту і динаміки цін на продуктових ринках великих міст з метою вивчення їх впливу на купівельну спроможність міського населення.

Так як організувати постійний облік оборотів і цін на ринках всіх міст України досить важко та й недоцільно, то для цього використовують несуцільне спостереження, де міста для обстеження відбирають в порядку основного масиву. В першу чергу беруть великі міста і промислові центри, в результаті чого обстежують відносно невелику кількість міст (біля 5 %), в яких проживає більше половини всього міського населення. Таке спостереження дозволяє отримати достатньо надійний матеріал для загальної характеристики кон'юнктури міських ринків.

*Анкетне спостереження* ґрунтується на принципі добровільного заповнення адресатами надісланих або розданих їм спеціальних анкет з метою отримання потрібної для дослідження інформації. Як правило, заповнених анкет повертається менше розданих або розісланих.

Анкетне спостереження досить просте і дешеве. Його використовують установи зв'язку для збирання відгуків від читачів, при проведенні різних соціологічних обстежень, для вивчення попиту на деякі товари народного споживання і в багатьох інших випадках.

Недоліком анкетного спостереження є те, що перевірити достовірність зібраного матеріалу досить складно або неможливо. Тому такий спосіб спостереження застосовують у випадках, коли не вимагається висока точність інформації, а лише наближені її характеристики.

За часом проведення статистичне спостереження поділяють на: а) поточне; б) періодичне; в) одноразове.

*Поточним* називається таке спостереження, яке ведеться систематично при безперервній реєстрації фактів в міру їх виникнення. Наприклад, реєстрація громадянських актів (народження, смерть, шлюб, розлучення), облік виходів працівників на роботу, облік виробленої продукції на підприємстві та ін.

*Періодичним* називається таке спостереження, яке повторюється через певні, заздалегідь установлені рівні проміжки часу. Такі спостереження, як правило, характеризують стан явищ на певний момент часу. Наприклад, щорічний перепис худоби станом на 1 січня, облік чисельності працівників, товарних запасів, залишків матеріальних цінностей на 1 число кожного місяця і т.д.

*Одноразовим* називається таке спостереження, яке проводиться в міру потреби один раз, або час від часу, без дотримання точної періодичності. Прикладами такого спостереження можуть бути переписи виробничого обладнання, переписи багаторічних плодовоягідних насаджень та ін.

За способами збирання статистичних даних розрізняють: а) безпосереднє спостереження; б) документальне спостереження; в) опитування.

*Безпосереднім* називається таке спостереження, при якому самі реєстратори збирають потрібні дані шляхом особистих замірювань, зважувань і підрахунків одиниць об'єкта і на цій основі проводять записи у формулярі спостереження. Наприклад, інвентаризація товарно-матеріальних цінностей, перепис виробничого обладнання, реєстрація залишків товарів в магазинах і т.д.

*Документальним* називається таке спостереження, при якому потрібні дані збирають і записують у формуляри на підставі використання різної документації. Цей спосіб спостереження ґрунтується на основі збирання необхідної інформації із статистичних звітів, які заповнюють підприємства і організації. При належному контролі за веденням первинного обліку і правильності заповнення форм статистичної звітності документальний спосіб спостереження дає найбільш точні результати.

*Опитування* – це таке спостереження, при якому відповіді на питання формуляра записують зі слів опитуваної особи. Так проводять перепис населення, збираючи в опитуваної особи і записуючи у формуляр відомості про її вік, сімейне положення, національність, освіту і інше.

В статистичній практиці використовують наступні три способи опитування: а) усне; б) самореєстрація; в) кореспондентський спосіб.

При *усному* опитуванні спеціально виділений працівник (реєстратор) розмовляє з опитуваною особою і з її слів сам заповнює формуляр. Такий спосіб гарантує однакове розуміння реєстратором питань і максимальну правильність відповідей на них.

При *самореєстрації* опитуваній особі вручають бланк обстеження, пояснюють питання і опитувана особа сама заповнює формуляр. В назначений час обліковець збирає заповнені формуляри і перевіряє повноту і правильність їх заповнення.

*Кореспондентський* спосіб полягає в тому, що інформацію в органи, які проводять спостереження, надсилають добровільні кореспонденти, які попередньо отримують від статистичних органів формуляри і інструкції щодо їх заповнення.

Кожний з наведених способів опитування має свої переваги і недоліки. Усне опитування вимагає значних затрат і великого числа реєстраторів, але воно гарантує високу якість зібраних матеріалів. Спосіб самореєстрації потребує менших затрат, однак його використовують тоді, коли програма спостереження повністю гарантує можливість правильного заповнення формулярів обстежуваними особами самостійно. Кореспондентський спосіб найменш затратний, але він не гарантує високої якості зібраних матеріалів, так як не завжди є можливість безпосередньо на місці перевірити правильність отриманих відповідей.

Вибір способу опитування визначається завданням і програмою статистичного спостереження та засобами, якими розпоряджається дослідник.

В зв'язку з утворенням автоматизованої системи державної статистики передача інформації може здійснюватись автоматично від однієї автоматизованої системи збору і обробки первинної статистичної інформації до іншої.

## 2.5. Помилки статистичного спостереження і способи контролю зібраних даних

Точність статистичного спостереження являється важливою і основною вимогою органів державної статистики. Однак, хоч як би старанно не було підготовлене статистичне спостереження, в процесі його проведення трапляються помилки, які призводять до зниження його точності.

*Точністю* статистичного спостереження називають ступінь відповідності значення будь-якої ознаки визначеної за допомогою статистичного спостереження її дійсному значенню. Чим ближчі значення ознак, отриманих в результаті статистичного спостереження до їх фактичних значень, тим точніше проведене спостереження.

Точність статистичного спостереження визначається як відношення даних спостереження до дійсних значень досліджуваних величин, або як різниця між ними.

*Помилками спостереження* називаються розходження між встановленими статистичним спостереженням і дійсними значеннями досліджуваних величин. Помилки спостереження виникають внаслідок неточностей при збиранні і реєстрації значень досліджуваних ознак.

Недопущення і попередження помилок є одним з важливих завдань організації і проведення статистичного спостереження. Невірні статистичні дані можуть призвести до прорахунків в державному управлінні народним господарством, серйозних помилок в науковому плануванні і прогнозуванні та інших негативних наслідків. Тому в Україні встановлена сувора відповідальність посадових осіб за навмисні викривлення статистичних даних.

В залежності від характеру, ступеня впливу на кінцеві результати, джерел і причин виникнення неточностей розрізняють наступні типи помилок статистичного спостереження: а) помилки реєстрації; б) помилки репрезентативності. Кожний з цих типів помилок ділиться на випадкові (ненавмисні) і систематичні (навмисні).

*Помилки реєстрації* виникають внаслідок неправильного встановлення фактів в процесі спостереження, помилкового запису їх значень, або обох причин разом.

*Випадковими* називаються помилки реєстрації, які можуть виникати внаслідок різних випадкових причин. Наприклад, опитувана особа може обмовитись, а реєстратор недочути чи випадково переставити місцями цифри, замість віку 23 роки записати 32 і навпаки. Такі неточності діють в протилежних напрямках і при достатньо великому числі спостережень взаємно погашаються.

*Систематичні* помилки реєстрації виникають внаслідок певних причин, діють в одному і тому ж напрямку і спричиняють серйозні викривлення загальних результатів статистичного спостереження. Наприклад, під час перепису населення опитувані особи часто округлюють свій вік, як правило, на цифрах, які закінчуються на “5” і “0”. Замість 34-36 років говорять 35, замість 49-51 говорять що їм 50 років і т.п. Внаслідок цього виходить, що 35, 40, 45, 50-річних громадян значно більше, ніж 34, 41, 46, 51-річних.

Систематичні помилки реєстрації можуть бути внаслідок свідомого викривлення фактів. Це навмисні приписки або приховування у звітах фактичних даних.

Помилки реєстрації виникають як при суцільному, так і при несуцільному спостереженні.

Точну величину помилок реєстрації визначити дуже важко, а тому за допомогою додаткових обстежень стараються знайти їх наближену оцінку.

На відміну від помилок реєстрації, помилки репрезентативності властиві тільки несуцільному спостереженню.

*Помилками репрезентативності* називаються відхилення значень ознак відібраної і обстеженої частини сукупності від значень ознак всієї досліджуваної сукупності. Якщо, наприклад, вибірковим спостереженням встановлено, що середня урожайність озимої пшениці, – 45,5 ц з 1 га, а її середня фактична - 50,0 ц з 1 га, то різниця ц з 1 га (50,0 – 45,5) і буде помилкою репрезентативності за умови, що помилка реєстрації дорівнює нулю.

*Випадкові помилки репрезентативності* виникають внаслідок того, що відібрана випадковим, неупередженим способом частина досліджуваної сукупності недостатньо повно відтворює всю сукупність в цілому. Статистичні методи дозволяють оцінити величину випадкової помилки репрезентативності.

*Систематичні помилки репрезентативності* виникають внаслідок порушення принципів неупередженого, випадкового відбору одиниць для обстеження. Кількісній оцінці розміри систематичної помилки репрезентативності, як правило, не піддаються.

Щоб уникнути помилок статистичного спостереження, зменшити розміри, виявити і виправити їх, потрібно в процесі підготовки і проведення спостереження передбачити і здійснити ряд заходів, таких як: виразно поставити завдання спостереження; детально розробити програму спостереження; чітко сформулювати питання у формулярах спостереження; грамотно розробити інструкцію спостереження, в якій доступно пояснити суть питань програми спостереження і методологію вирахування досліджуваних показників; вміло вибрати дату, встановити терміни і вибрати місце спостереження; правильно підібрати і підготувати кадри.

З метою отримання в процесі статистичного спостереження високоякісних матеріалів, статистичні органи здійснюють постійний контроль за ходом проведення спостереження, систематично перевіряють стан первинного обліку і звітності на підприємствах, організаціях і установах.

Важливе значення для підвищення точності матеріалів спостереження має роз'яснювальна робота серед посадових осіб і населення про мету, завдання, серйозність і порядок проведення спостереження, щоб не допустити зривів спостереження чи викривлень його даних.

Після закінчення спостереження матеріали, зібрані в процесі його проведення, старанно перевіряються за повнотою охоплення об'єкта спостереження, якістю заповнення формулярів і інших документів.

Статистика використовує два способи контролю матеріалів спостереження: а) арифметичний (лічильний); б) логічний.

*Арифметичний контроль* полягає в лічильній перевірці підсумкових даних звітів або формулярів і погодженні тих показників, які взаємозв'язані між собою і можуть бути виведені одні з одних. Наприклад, в шаховій таблиці любого зведення підсумки рядків і колонок повинні співпадати, а якщо такого співпадання немає, тоді шукають помилку в рядках чи колонках.

*Логічний контроль* полягає в співставленні взаємозв'язаних між собою відповідей на питання формуляра статистичного спостереження і виясненні їх логічної сумісності. Якщо виявляють логічно несумісні відповіді, шляхом подальшого співставлення з відповідями на інші питання встановлюють, яка з відповідей записана невірно. Наприклад, якщо у формулярі переписного листа перепису населення записано, що опитувана особа у віці 7 років має сім'ю, вищу освіту, працює лікарем, то зрозуміло, що неправильно записаний вік.

Логічний контроль здійснюють також шляхом порівняння даних статистичного спостереження з показниками плану, прогнозу або даними за аналогічні попередні періоди, чи з даними суміжних об'єктів, які зв'язані в якійсь мірі певними показниками і т.д.

Основною умовою успішного проведення любого статистичного дослідження на всіх його стадіях, в тому числі і на стадії спостереження є висока якість зібраного матеріалу.

### *Контрольні запитання*

1. Які ви знаєте стадії статистичного дослідження?
2. Статистичне спостереження – як перша стадія статистичного дослідження.
3. Завдання статистичного спостереження.
4. Мета статистичного спостереження.
5. Що є об'єктом статистичного спостереження?
6. Що називається одиницею статистичного спостереження?
7. Що називається одиницею сукупності?
8. Для чого складається програма статистичного спостереження?
9. Зміст програми статистичного спостереження.
10. Що називається статистичним формуляром?
11. Індивідуальні і спискові формуляри.
12. Інструкція статистичного спостереження.
13. Поняття про інтервальні та моментні показники.
14. Для чого встановлюють період спостереження?
15. Що ви розумієте під критичним моментом статистичного спостереження?
16. Організаційний план статистичного спостереження.
17. Для чого вибирають місце спостереження?
18. Основні організаційні форми статистичного спостереження.
19. Поняття про статистичну звітність.
20. Загальнодержавна і внутрівідомча звітність.
21. Спеціально організоване статистичне спостереження.
22. Поняття про переписи.
23. Суцільні і несуцільні спостереження.
24. Види несуцільного спостереження.
25. Поняття про вибіркове спостереження.

26. Поняття про монографічне спостереження.
27. Метод основного масиву.
28. Анкетне спостереження.
29. Що ви розумієте під поточним спостереженням?
30. Що називається періодичним спостереженням?
31. Для чого проводять одноразове спостереження?
32. Які ви знаєте способи збирання статистичних даних?
33. Поняття про безпосереднє спостереження.
34. Документальне спостереження.
35. Опитування – як вид статистичного спостереження.
36. Які ви знаєте способи опитування?
37. Поняття про усне опитування.
38. Самореєстрація – як вид опитування.
39. Кореспондентський спосіб опитування.
40. Що ви розумієте під точністю статистичного спостереження?
41. Поняття про помилки статистичного спостереження.
42. Помилки реєстрації.
43. Випадкові і систематичні помилки реєстрації.
44. Помилки репрезентативності.
45. Випадкові і систематичні помилки репрезентативності.
46. Статистичні способи контролю матеріалів спостереження.
47. Арифметичний контроль.
48. Логічний контроль.

# Розділ 3. Зведення і групування статистичних матеріалів

## 3.1. Зміст і завдання статистичного зведення

В результаті статистичного спостереження отримують велику кількість різноманітних відомостей про кожну одиницю досліджуваної сукупності. Проте, щоб на основі цих відомостей можна було зробити певні висновки, потрібно всю масу окремих даних привести до відповідного порядку, систематизувати, обробити і на цій основі дати зведену характеристику всієї сукупності фактів за допомогою узагальнюючих статистичних показників. Цього досягають на другому етапі статистичного дослідження, який називається зведенням і групуванням статистичних матеріалів.

Отже, *статистичним зведенням* називається наукова обробка первинних даних статистичного спостереження з метою отримання узагальнюючих характеристик досліджуваного явища чи процесу за рядом суттєвих для них ознак.

Перш ніж приступити до зведення зібраного первинного статистичного матеріалу його потрібно проконтролювати і прийняти. Попередній теоретичний аналіз повинен сприяти тому, щоб під час зведення не губились основні риси досліджуваних явищ в загальних підсумках. Опрацьовуваний матеріал необхідно перевірити за повнотою охоплення обстежуваних одиниць і якістю отриманих про них даних. Якість і повноту зібраної інформації перевіряють за допомогою логічного і лічильного контролю, виявлені дефекти виправляють. Важливою умовою своєчасного і правильного проведення статистичного зведення є суворе дотримання звітної дисципліни. І тільки після того, як весь первинний статистичний матеріал старанно проконтрольований і належним чином виправлений, можна приступати до його зведення.

Зведення може бути *просте* – як вузькотехнічна операція по підрахунку підсумків первинного статистичного матеріалу, а також *складне* – яке передбачає групування даних, розробку системи показників, підрахунок групових і загальних підсумків та виклад результатів зведення у вигляді статистичних таблиць чи графіків.

Статистичне зведення проводять за наперед розробленою програмою, яка відповідає завданням статистичного дослідження з врахуванням прийнятої форми організації зведення.

За формою організації зведення буває централізоване і децентралізоване. При централізованій формі організації зведення всі матеріали спостереження обробляють і систематизують в Державному комітеті статистики України. Суттєвою перевагою даної форми зведення є те, що вона дає можливість його автоматизації і використання єдиної методології обробки даних. При децентралізованій формі організації зведення матеріали спостереження обробляють і узагальнюють на місцях, а в центральні органи, тобто Державному комітету статистики України, надсилають зведену інформацію по регіонах. Децентралізована форма зведення дещо дешевша і оперативніша за централізовану.

На практиці, як правило, поєднують територіально-децентралізовану і централізовану форму зведення.

В залежності від завдань статистичного дослідження програма зведення встановлює групувальні ознаки, кількість груп та макети розроблюваних таблиць. Програма повинна бути складена таким чином, щоб в результаті зведення отримати матеріал, який характеризує досліджуване явище з різних його сторін.

Для успішного здійснення статистичного зведення складається план його проведення. План має містити розв'язок питань організації зведення куди входять: послідовність і терміни виконання окремих частин зведення, оформлення його результатів у вигляді таблиць, публікацій у вигляді статистичних збірників і ін. В плані зведення також вказується на яких машинах вона буде проводитись і яка техніка зведення.

### 3.2. Завдання і значення статистичних групувань

Під час обробки статистичних матеріалів виникає потреба виділення однорідних груп, типів, а вже потім їх описання за допомогою відповідних кількісних характеристик. Застосування таких узагальнюючих показників як відносні і середні величини, індекси та інші, можливе лише після того, як статистичний матеріал розподілений на однорідні групи.

*Групуванням* в статистиці називається розчленування усіх одиниць досліджуваної сукупності на групи за певними істотними для них ознаками. Серед багатьох методів, які роблять статистику одним з наймогутніших знарядь соціального пізнання, групування вважається найбільш ефективним. Воно є центральним моментом любого зведення, завдяки чому матеріал статистичного спостереження приймає систематизованого вигляду.

При статистичному вивченні соціально-економічних явищ і процесів групування є одним з основних методів аналізу і синтезу.

Ознаки, покладені в основу групування називаються групувальними. Групування одиниць досліджуваної сукупності за якою-небудь ознакою веде до рядів розподілу.

Групувальні ознаки можуть мати кількісний вираз (наприклад, вік працівника, стаж роботи, тарифний розряд, заробітна плата, урожайність і т.п.), тому вони називаються кількісними, а ряди їх розподілу – *варіаційними рядами*.

Якщо групувальні ознаки відображають певні властивості одиниць сукупності (наприклад, стать, національність, освіту, професію і т.п.), вони називаються якісними, а ряди розподілу побудовані за такими ознаками, називаються *атрибутивними*.

При групуванні одиниць сукупності за територіальною ознакою отримують *географічні* або *територіальні ряди* розподілу. Вони дають уяву про розміщення або ступінь розповсюдження тих чи інших явищ в просторі.

Особливим видом групувань в статистиці є класифікація. Класифікацією в статистиці називається стійке і фундаментальне групування одиниць сукупності за атрибутивною ознакою на подібні і відмінні групи і підгрупи. Перелік цих груп і підгруп розглядається як своєрідний статистичний стандарт, затверджений Державним комітетом статистики України. Наприклад, класифікація галузей народного господарства, класифікація основних фондів, класифікація професій і т.д. Статистичні класифікації ґрунтуються на таких суттєвих ознаках, які мало змінюються і існують тривалий час.

Метод статистичних групувань застосовується для розв’язання багатьох завдань, що виникають в ході наукового соціально-економічного дослідження. Всі вони зводяться до впорядкування первинного статистичного матеріалу, розподілу його за суттєвими варіаційними ознаками, для подальшого економічного аналізу. Групування, будучи першою сходинкою статистичного аналізу, одночасно є підготовчою стадією для більш глибокого аналізу досліджуваного статистичного матеріалу. В цьому заключається основне значення статистичного групування і визначається його провідна роль в зведенні первинного статистичного матеріалу.

Із багатьох завдань, які розв’язуються з допомогою статистичних групувань, можна виділити три основних:

1) розподіл всієї сукупності на якісно однорідні соціально-економічні типи;

2) вивчення структури явищ і структурних зрушень в них;

3) виявлення і характеристика взаємозв’язку між ознаками досліджуваного явища.

Ці завдання розв'язуються за допомогою типологічних, структурних і аналітичних групувань.

### 3.3. Основні правила утворення груп

Перед проведенням простих, а тим більше комбінованих групувань потрібно розв'язати питання про кількість груп, розмір інтервалу та ін.

При групуванні за атрибутивними ознаками число груп, на які ділять досліджувану сукупність, визначається кількістю різновидів цієї ознаки. Групування ж населення за національністю допускає утворення стількох груп, скільки різних національностей є серед населення досліджуваного регіону.

Окремим випадком атрибутивних групувань є альтернативне групування, коли є всього два варіанти атрибутивної ознаки, один з яких виключає інший. Наприклад, розподіл робітників які мають спеціальну освіту і які такої освіти не мають, розподіл студентів які отримують стипендію, і тих, що не отримують її, і т.п.

Інший характер має групування за кількісними ознаками, при якому виникає питання про кількість груп, числові межі окремих груп та інтервали групування. Наприклад, групування робітників за стажем роботи, тарифним розрядом чи заробітною платою; групування заводів за вартістю основних виробничих фондів, випуском або реалізацією продукції і т.д.

При розв'язанні питання про те, скільки доцільно утворити груп, беруть до уваги варіацію ознаки і число спостережень. Чим інтенсивніше змінюється ознака і чим більша сукупність одиниць, тим більше число груп можна утворити. Однак, як загальний принцип розв'язання питання про необхідну кількість груп виступає вимога, щоб вона була оптимальною, і щоб до кожної групи потрапила достатньо велика кількість одиниць. При великій кількості груп відбудеться розпорошення одиниць досліджуваної сукупності, однорідні одиниці попадуть в різні групи. А при малій кількості груп до однієї і тієї ж групи попадуть одиниці різних типів. Утворення великого число груп, як і малого, приведе до помилкових висновків.

Для отримання типових показників необхідно брати велику кількість одиниць досліджуваної сукупності, тоді до кожної групи попаде достатня кількість однорідних одиниць. Однак, в окремих випадках доцільно брати і малочисельні групи, зокрема коли мова йде про вловлювання нового, передового, яке тільки зароджується і ще не стало масовим, яке проявляється в незначному числі фактів. Завдання групування полягає в тому, щоб виділити ці факти, якими б малочисельними вони не були сьогодні, з метою їх подальшого глибокого вивчення.

Конкретне визначення достатньої чисельності груп потрібно погоджувати з суттю досліджуваного об'єкту. Таким чином, при розв'язанні питання про чисельність груп необхідно керуватись логікою і знанням суті досліджуваного питання.

При групуванні за кількісними ознаками виникає суттєве питання про вибір розміру інтервалів групування.

*Інтервалом* групування називається різниця між максимальним і мінімальним значеннями ознаки в кожній групі.

Інтервали в структурних і аналітичних групуваннях можуть бути рівними і нерівними в залежності від характеру розподілу одиниць сукупності за даною ознакою. Нерівні інтервали, в свою чергу, можуть бути прогресивно-зростаючими або прогресивно-спадаючими.

Якщо варіація досліджуваної ознаки знаходиться в порівняно вузьких межах і розподіл близький до нормального, то застосовують *рівні* інтервали.

Величину інтервалу, при групуванні з рівними інтервалами, визначають шляхом ділення розмаху варіації на число груп за формулою:

$$i = \frac{x_{max} - x_{min}}{n},$$

де $x_{max}$ – максимальна величина ознаки;
$x_{min}$ – мінімальна величина ознаки;
i – розмір інтервалу;
n – число груп.

Розмір інтервалу залежить від числа груп і варіації досліджуваної ознаки. Чим більшою буде варіація ознаки, тим більшим буде розмір інтервалу і чим більше число груп, тим менший розмір інтервалу.

Число груп орієнтовно можна визначити за формулою американського вченого Стерджеса:

$$n \approx 1 + 3{,}321 \lg N,$$

де N – число одиниць сукупності, яка підлягає групуванню.

Тоді формула для визначення розміру інтервалу матиме вигляд:

$$i = \frac{x_{max} - x_{min}}{1 + 3{,}321 \lg N}.$$

Якщо, наприклад, потрібно провести групування з рівними інтервалами 100 робітників підприємства за розміром заробітної плати, максимальне значення якої 280 грн., а мінімальне – 120 грн., то $n = 1 + 3{,}321 \lg N = 7{,}642 \approx 8$ груп, а $i = \dfrac{280 - 120}{8} = 20$ грн.

В даному випадку оптимальним розміром інтервалу може бути величина 20 грн., а число груп – 7 або 8.

Якщо в результаті ділення отримують не кругле число (наприклад 19,5 грн.) і виникає потреба в заокругленні, то заокруглюють в більшу сторону, в даному випадку до 20 грн.. В нашому прикладі ми будемо мати наступні вісім груп з рівними інтервалами:

| | | | |
|---|---|---|---|
| I | від 120 до 139 грн.; | V | від 200 до 219 грн.; |
| II | від 140 до 159 грн.; | VI | від 220 до 239 грн.; |
| III | від 160 до 179 грн.; | VII | від 240 до 259 грн.; |
| IV | від 180 до 199 грн.; | VIII | від 260 до 279 грн. |

Часто таке групування записують спрощено:

| | | | |
|---|---|---|---|
| I | 120 – 140 грн.; | V | 200 – 220 грн.; |
| II | 140 – 160 грн.; | VI | 220 – 240 грн.; |
| III | 160 – 180 грн.; | VII | 240 – 260 грн.; |
| IV | 180 – 200 грн.; | VIII | 260 – 280 грн. |

В цьому випадку використовують правило, що ліва цифра включає зазначене число, а права не включає. Це означає, що робітник із заробітною платою 140 грн. повинен бути віднесений до другої групи, 160 грн. – до третьої групи, 180 грн. – до четвертої і т.д.

Інколи перший і останній інтервали бувають відкритими:

| | | | |
|---|---|---|---|
| I | до 140 грн.; | V | 200 – 220 грн.; |
| II | 140 – 160 грн.; | VI | 220 – 240 грн.; |
| III | 160 – 180 грн.; | VII | 240 – 260 грн.; |
| IV | 180 – 200 грн.; | VIII | 260 грн. і більше, або понад 260 грн. |

Тоді стоїть питання, до якої групи віднести заробітну плату 140 грн., 160 грн., 180 грн. і т.д. В даному випадку розглядають як записані відкриті інтервали. Якщо в останньому інтервалі записано “260 грн. і більше”, це означає, що заробітна плата 260 грн. відноситься до останньої групи, отже, заробітну плату робітника в 140 грн. відносять до другої групи, а якщо в останньому інтервалі записано “понад 260 грн.”, то робітника із заробітною платою в 260 грн. відносять до передостанньої групи, а із заробітною платою 140 грн. – до першої групи.

Часто виникає потреба у визначенні серединних значень інтервалів. Для знаходження серединних значень інтервалів в групах з рівними інтервалами діють таким чином: спочатку визначають серединне значення для третього інтервалу як півсуму верхніх значень другого і третього інтервалів [(180 + 160) : 2] = 170 грн. Серединне значення другого інтервалу визначають вирахуванням із серединного значення третього інтервалу розміру інтервалу 170 – 20 = 150 грн., а серединне значення першого інтервалу – як різниця між серединним значенням другого інтервалу і розміром інтервалу 150 – 20 = 130 грн. Серединне значення останнього інтервалу визначають додаванням до серединного значення передостаннього інтервалу розміру інтервалу 250 + 20 = 270 грн.

В соціально-економічній статистиці часто застосовуються групування з *нерівними* інтервалами. Застосування нерівних, які прогресивно збільшуються або зменшуються, інтервалів зумовлено самою природою більшості соціально-економічних явищ, коли в нижчих групах велике значення мають навіть малі відмінності в показниках, а у вищих групах такі відмінності суттєвого значення не мають. Так, наприклад, для нижчих груп, при групуванні підприємств за чисельністю працюючих, різниця в 50 чи 100 чоловік має велике значення, а для вищих груп, в яких зосереджені великі підприємства, така різниця не суттєва. При групуванні промислових підприємств за чисельністю працюючих можна виділити такі групи:

| | | | |
|---|---|---|---|
| I | до 100 чол.; | V | 1001 – 3000 чол.; |
| II | 101 – 200 чол.; | VI | 3001 – 10000 чол.; |
| III | 201 – 500 чол.; | VII | 10001 – 20000 чол.; |
| IV | 501 – 1000 чол.; | VIII | понад 20000 чол. |

Вибір тих чи інших розмірів інтервалів залежить від числа одиниць, віднесених в кожну групу, або від ступеня заповнення кожного інтервалу. Ступінь заповнення отриманих інтервалів є одним з найголовніших критеріїв доцільності любого групування. Якщо групи з широкими інтервалами багаточисельні, то їх потрібно розбити на більш вузькі, а якщо групи з вузькими інтервалами заповнені слабо, або взагалі не заповнені, то виникає необхідність в розширенні інтервалів груп.

В статистичних групуваннях часто намагаються розмежувати дві якісно відмінні групи підприємств. Наприклад, підприємства, які не виконали план і ті, які виконали план на 100 % і більше. Групування підприємств за ступенем виконання плану, як правило, має такі інтервали:

| | | | |
|---|---|---|---|
| I | до 90 %; | V | від 99 до 100 %; |
| II | від 90 до 95 %; | VI | від 100 до 101 %; |
| III | від 95 до 98 %; | VII | від 101 до 103 %; |
| IV | від 98 до 99 %; | VIII | 103 % і більше. |

Інтервали, починаючи з другої групи до п’ятої включно, поступово зменшуються, в міру наближення підприємств до рівня виконання плану на 100 %, тому що, чим ближче до 100 %, тим більше підприємств в цих групах. А тому, дуже важливо виділити ті з них, які впритул наблизились до виконання плану. Аналогічно трактуються малі розміри інтервалів в шостій і сьомій групах, тому що тут кожний відсоток виконання і перевиконання плану має велике значення.

В групуваннях, метою яких є утворення якісно однорідних груп, застосовують *спеціалізовані* інтервали. В таких групуваннях межа інтервалу встановлюється там, де відбувається перехід від однієї якості до іншої. Кількість груп встановлюється у відповідності з теорією досліджуваного питання. Наприклад, при групуванні дітей за віком встановлюють наступні групи:

I діти до 1 року – немовлята;
II до 3 років – діти ясельного віку;
III від 3 до 7 років – дошкільнята;
IV від 7 до 17 років – діти шкільного віку.

Виділені таким чином групи дітей істотно відрізняються одна від одної в плані вивчення дитячої смертності (дітей віком до 1 року); в обслуговуванні і охороні дітей і материнства (друга група); в обслуговуванні і підготовці дітей до школи (третя група); в забезпеченні дітей школами, педагогічними кадрами, підручниками і шкільним приладдям та іншими необхідними для їх розвитку закладами (четверта група).

За характером відношення чоловічого населення до трудової діяльності виділяють наступні групи за віком:

I 0 – 15 років – непрацездатне населення;
II 16 – 18 років – особи напівпрацездатного віку;
III 19 – 59 років – працездатне населення;
IV 60 – 69 років – особи напівпрацездатного віку;
V 70 років і старше – непрацездатне населення.

Межі спеціалізованих інтервалів потрібно встановлювати в залежності від умов місця і часу проведення групування. Так, наприклад, підприємства різних галузей промисловості мають різну енергомісткість. Тому групування підприємств за рівнем енергоозброєності робітників з виділенням трьох груп (високої, середньої і низької енергоозброєності) потрібно проводити диференційовано за галузями промисловості. Адже, одне і те ж кількісне значення цього показника може по різному характеризувати ступінь енергоозброєності. Наприклад, на підприємствах кольорової металургії воно може бути віднесене до низької групи, а на підприємствах швейної або взуттєвої галузі – до високої. Отже, в різних галузях для віднесення підприємств до тієї чи іншої групи потрібно застосовувати різні кількісні міри та інтервали за ознакою енергоозброєності робітників.

Аналогічний приклад можна навести і з групуванням промислових підприємств за числом працівників. Так підприємство з 300 працівниками в машинобудуванні є порівняно мале, а в електроенергетиці – велике.

### 3.4. Типологічні групування

Аналізуючи розвиток суспільних явищ і процесів в часі, потрібно виділити соціально-економічні типи, так як в зародженні, розвитку, боротьбі і відмиранні різних соціально-економічних типів заключається суть історичного процесу розвитку любого суспільства. Виділення соціально-економічних типів при дослідженні певного явища є одним з головних і вирішальних завдань методу статистичних групувань.

*Типологічними* називаються групування, за допомогою яких проводять розподіл досліджуваного суспільного явища на класи або соціально-економічні типи.

Типологічні групування в статистичних дослідженнях займають одне з центральних місць. На основі всебічного теоретичного аналізу досліджуваної сукупності виділяють її головні і найхарактерніші типи або групи, вивчають істотні відмінності між ними, а також спільні ознаки для всіх груп.

За допомогою типологічних групувань вивчають класовий склад населення, розподіл підприємств за формами власності, поділ суспільного виробництва на I і II підрозділи, продукції промисловості на групи “А” і “Б”, поділ народного господарства на сферу матеріального виробництва і невиробничу сферу та ін.

В якості прикладу візьмемо розподіл населення України, зайнятого в народному господарстві за галузями матеріального виробництва і невиробничої сфери, % (табл. 3.1).

Таке групування дає можливість простежити тенденції зрушень в структурі зайнятості населення України в народному господарстві за ряд років.

Без типологічного групування важко зрозуміти і статистично правильно охарактеризувати процеси розвитку промисловості, сільського господарства, будівництва, торгівлі, питання виробництва і розподілу валового внутрішнього продукту і національного доходу в любому суспільстві без врахування класової структури, виділення соціально-економічних типів явищ і однорідних в класовому відношенні груп.

Таблиці 3.1*

| Зайняте населення | Роки | | | | |
|---|---|---|---|---|---|
| | 1960 | 1970 | 1980 | 1986 | 1993 |
| Всього зайнятого в народному господарстві | 100 | 100 | 100 | 100 | 100 |
| У галузях матеріального виробництва, в тому числі вантажному транспорті, зв'язку, обслуговуванні виробництва, торгівлі, особистому підсобному господарстві: | 85,9 | 79,9 | 76,0 | 74,4 | 74,9 |
| робітники і службовці; | 40,7 | 52,7 | 57,4 | 58,2 | 58,7 |
| колгоспники, зайняті в громадському та особистому підсобному господарстві; | 42,0 | 25,8 | 17,5 | 15,4 | 15,0 |
| члени сімей робітників і службовців, зайняті в особистому підсобному господарстві; | 3,0 | 1,4 | 0,8 | 0,7 | - |
| інше населення (кустарі, одноосібники). | 0,2 | 0,0 | 0,1 | 0,1 | - |
| У невиробничій сфері | 14,1 | 20,1 | 24,2 | 25,5 | 25,1 |

Користуючись методом статистичних групувань досліджують утворення і розвиток нових економічних типів явищ.

У випадках, коли статистика характеризує явища, які складаються з різних соціально-економічних типів і мають різні закони розвитку, зведені статистичні характеристики у вигляді середніх величин будуть правильно характеризувати розвиток явища тільки в тому випадку, якщо попередньо виділені за допомогою групувань якісно однорідні типи явищ.

Типологічні групування дають можливість аналізувати своєрідні особливості і розвиток окремих типів, зміну їх співвідношення в загальному об'ємі певного явища. Такі групування, як правило, містять цілу систему статистичних показників, які дозволяють глибоко і всебічно проаналізувати відмінності окремих типів і їх питому вагу за рядом показників в загальній сукупності.

Важким і важливим питанням методу любих групувань є правильний вибір групувальної ознаки, від якої залежать результати

* Народне господарство України у 1993 році. – К., 1994. – С. 275.

групування. Типологічні групування вимагають особливого підходу до вибору групувальної ознаки. Якщо такими ознаками виступають атрибутивні ознаки, наприклад, класовий склад населення, форма власності, галузь виробництва, то утворення числа груп і їх назви визначаються самою ознакою.

Однак, часто доводиться виділяти типи на основі групувань за кількісною ознакою. Наприклад, виділяючи із підприємств великі, середні і малі, або передові і відстаючі, ми групуємо їх відповідно за вартістю основних виробничих фондів (виробленої продукції), і рівнями виконання плану (продуктивності праці), тобто за кількісною ознакою. Тут важливо правильно встановити інтервал групування, щоб кількісно виділити одні типи від інших. Це питання розв'язується на основі визначення таких кількісних меж, які виділять нову якість.

Виділення типів зумовлене багатьма формами переходу одних типів в інші, багатьма ознаками, зв'язаними між собою. Тому, для виділення окремих типів, потрібно брати сукупність ознак, які характеризують різні сторони досліджуваного явища. Так, при виділенні соціально-економічних типів промислових підприємств, потрібно врахувати форму власності, вартість основних виробничих фондів, об'єм виробництва продукції, середньоспискову чисельність працівників та інші ознаки в залежності від умов місця і часу.

Прикладом такого типологічного групування можуть бути наступні дані (табл. 3.2).

Таблиця 3.2

Енергомісткість основних підгалузей хімічної промисловості, %*

| Підгалузі | Частка галузі в енерговитратах хімічної промисловості | Частка енерговитрат у собівартості продукції |
|---|---|---|
| Високоенергомісткі: | | |
| азотна; | 32,5 | 21,9 |
| хімічних волокон; | 13,9 | 9,7 |
| каустичної соди; | 9,7 | 18,5 |
| содова. | 5,7 | 23,4 |
| Середньоенергомісткі: | | |
| основна хімія; | 4,3 | 7,6 |
| гірнича хімія; | 3,7 | 14,5 |

* Розміщення продуктивних сил України: Підручник / Є.П.Качан, М.О.Ковтонюк, М.О.Петрига та ін.: За ред. Є.П.Качана. – К.: Вища школа, 1997. – С. 153.

Продовження табл. 3.2

| | | |
|---|---|---|
| пластмас і синтетичних смол; | 2,9 | 7,2 |
| калійна; | 2,9 | 13,9 |
| аніліно-фарбова. | 2,1 | 7,2 |
| Малоенергомісткі: | | |
| переробка пластмас; | 1,8 | 3,3 |
| хімічних засобів захисту рослин; | 1,7 | 8,3 |
| лакофарбова; | 1,6 | 2,4 |
| сірчана; | 1,4 | 15,1 |
| склопластиків; | 0,9 | 6,4 |
| реактивів; | 0,9 | 4,3 |
| фотохімічна; | 0,9 | 5,6 |
| киснева; | 0,6 | 27,0 |
| побутова хімія | 0,6 | 2,1 |

Вибір групувальних ознак вимагає всебічного теоретичного аналізу суті явища яке вивчається, щоб, керуючись завданнями дослідження, покласти в основу групування суттєві ознаки, які відповідають вимогам поставленого завдання.

Типологічні групування використовують всюди, де потрібно охарактеризувати якісні особливості окремих груп.

З розвитком народного господарства, переходом економіки до ринкових відносин значення типологічних групувань постійно зростає.

### 3.5. Структурні групування

Виділені в результаті типологічних групувань, окремі типи явищ вивчаються за їх складом. Групування дозволяють також охарактеризувати структуру і структурні зрушення досліджуваного явища. Так, за допомогою групувань вивчають склад населення за віком, статтю, освітою, національністю, сімейним положенням; склад працівників – за професіями, стажем роботи, віком та іншими ознаками.

Отже, *структурним групуванням* в статистиці називається розчленування якісно однорідної сукупності одиниць на групи, які характеризують її структуру і структурні зрушення за відповідний період часу.

Велике значення мають структурні групування при вивченні концентрації промислових, сільськогосподарських, будівельних, торгових, транспортних та інших підприємств.

Для вивчення процесу концентрації в промисловості, підприємства групують за вартістю основних виробничих фондів, за середньосписковою чисельністю працюючих, за обсягом виробленої чи реалізованої продукції, за рівнем механізації і автоматизації виробництва і т.д. Кожна з наведених групувальних ознак по-своєму відображають концентрацію. Для трудомістких галузей народного господарства концентрацію відображають через групування за числом робітників, для енергомістких галузей – з допомогою групувань за енергозатратами, а в деяких випадках групування за однією ознакою доповнюють групуваннями за іншими ознаками.

Структурні групування широко використовують в аналізі виконання плану. Групування підприємств певної галузі за процентом виконання плану дозволяє виявити які з них виконали і перевиконали план, а які його не виконали, встановити причини невиконання плану і показати шляхи ліквідації відставання від планових завдань. Це дозволить розрахувати певні резерви для виконання і перевиконання плану по галузі в цілому.

Метод структурних групувань дає можливість аналізувати статистичні сукупності по економічних і адміністративних регіонах, по галузях народного господарства, по географічних зонах і за багатьма іншими ознаками.

В зміні структури суспільних явищ відображаються важливі закономірності їх розвитку. Наприклад, індустріалізація народного господарства проявляється в рості частки промислової продукції у всій продукції народного господарства, а також в збільшенні питомої ваги виробництва засобів виробництва в складі продукції промисловості. Зростання культурного рівня населення простежується в підвищенні його грамотності і частки людей із середньою спеціальною і вищою освітою.

Якщо в структурних групуваннях порівняти дані в часі, то отримаємо інформацію, яка характеризуватиме закономірності в зміні структури, тобто про структурні зрушення в досліджуваному явищі. Тому структурні групування часто подаються в динамічних таблицях, наприклад (табл. 3.3).

Таблиця 3.3

Чисельність населення України (на 1 січня)*

| Показник | Роки | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1970 | 1990 | 1992 | 1993 | 1994 | 1995 | 1996 |
| Чисельність населення, млн. осіб | 47,1 | 51,8 | 52,1 | 52,2 | 52,1 | 51,7 | 51,3 |
| в тому числі, %: | | | | | | | |
| чоловіки; | 45,2 | 46,3 | 46,4 | 46,6 | 46,4 | 46,4 | 46,6 |
| жінки. | 54,8 | 53,7 | 53,6 | 53,4 | 53,6 | 53,6 | 53,4 |

* Розміщення продуктивних сил України: Підручник / Є.П.Качан, М.О.Ковтонюк, М.О.Петрига та ін.: За ред. Є.П.Качана. – К.: Вища школа, 1997. – С. 31.

Структурне групування показало певну стабілізацію за останні роки частки чоловіків і жінок в загальній чисельності населення України.

В наведеній таблиці структурне групування має лише один показник – питому вагу даної групи в загальному обсязі сукупності. Однак, в багатьох випадках структурне групування представляється цілою системою показників, що дозволяє вивчати розподіл підприємств за рядом показників, які характеризують їх роботу.

Багатопоказникове структурне групування покажемо на прикладі групування промислових підприємств одного з економічних регіонів України за обсягом виробництва валової продукції в 2007 р.

Таблиця 3.4

Групування промислових підприємств за обсягом валової продукції в 2007 р.

*(в відсотках до підсумку, дані умовні)*

| Групи підприємств за річним обсягом валової продукції, тис. грн. | Число підприємств | Валова продукція | Середньорічна чисельність промислово-виробничого персоналу | Середньорічна вартість промислово-виробничих основних фондів | Споживання електроенергії |
|---|---|---|---|---|---|
| До 100 | 5,8 | 0,1 | 0,2 | 0,1 | 0,1 |
| 101 – 500 | 13,3 | 0,2 | 1,5 | 0,7 | 0,2 |
| 501 – 1000 | 12,8 | 0,7 | 2,1 | 1,0 | 0,5 |
| 1001 – 5000 | 37,0 | 8,6 | 14,0 | 8,9 | 1,6 |
| 5001 – 10000 | 13,1 | 8,9 | 12,1 | 8,0 | 7,0 |
| 10001 – 50000 | 14,2 | 25,3 | 30,0 | 24,5 | 19,0 |
| 50001 – 100000 | 2,5 | 18,0 | 14,9 | 15,9 | 15,6 |
| 100001 і більше | 1,3 | 38,2 | 25,2 | 40,9 | 56,0 |
| Разом: | 100,0 | 100,0 | 100,0 | 100,0 | 100,0 |

За матеріалами даного групування простежується значна концентрація промисловості. Аналізуючи наведені показники бачимо, що великі підприємства мають більш високу продуктивність праці, більш високу ступінь використання основних промислово-виробничих фондів та більш високий рівень енергоозброєності праці.

Таке структурне групування можна одночасно розглядати і як типологічне (виділення малих, середніх і великих підприємств), і як аналітичне, яке показує зв’язок між рядом показників, що характеризують окремі групи.

### 3.6. Аналітичні групування

Одним з основних завдань, яке розв'язується за допомогою статистичних групувань, є дослідження взаємозв'язків варіаційних ознак в межах, як правило, однорідної сукупності.

*Аналітичними групуваннями* в статистиці називаються такі, за допомогою яких виявляють і вивчають взаємозв'язок між окремими ознаками досліджуваного соціально-економічного явища.

Всі явища суспільного життя і їх ознаки зв'язані між собою і залежать одне від одного. Взаємозв'язані ознаки діляться на факторні і результативні. *Факторною* називається ознака, під впливом якої змінюється залежна від неї інша ознака. *Результативною* називається ознака, яка змінюється під впливом факторної ознаки. Наприклад, внесення добрив – це фактор того, що в результаті буде вища врожайність сільськогосподарських культур. Внесення добрив впливає на урожайність, яка, в свою чергу, змінюється і залежить від кількості і якості цих добрив. Отже, внесення добрив – це факторна ознака, а вища врожайність сільськогосподарських культур – результативна ознака.

Розглянемо для прикладу аналітичне групування промислових підприємств однієї із галузей народного господарства за вартістю основних виробничих фондів і випуском валової продукції.

Таблиця 3.5

Залежність випуску валової продукції від вартості основних виробничих фондів.

| Групи заводів за вартістю основних виробничих фондів, млн. грн. | Кількість заводів | Вартість основних виробничих фондів в середньому на один завод, млн. грн. | Валова продукція за звітний період (у порівняльних цінах) в середньому на один завод, млн. грн. | Фондо-віддача, грн. |
|---|---|---|---|---|
| I 1,0 – 2,5 | 3 | 1,70 | 1,97 | 1,16 |
| II 2,5 – 4,0 | 10 | 3,16 | 3,79 | 1,20 |
| III 4,0 – 5,5 | 7 | 4,60 | 5,78 | 1,25 |
| IV 5,5 – 7,0 | 5 | 6,26 | 10,68 | 1,71 |
| Разом: | 25 | 4,01 | 5,51 | 1,37 |

Із таблиці чітко простежується тенденція до збільшення випуску валової продукції із зростанням вартості основних виробничих фондів. Групування показує залежність випуску валової продукції від вартості основних виробничих фондів, тобто від розміру підприємства. Це пояснюється тим, що на великих підприємствах підсилюється роль науково-технічного прогресу (вони мають більш високий рівень механізації і автоматизації виробництва), покращується організація праці, збільшується число висококваліфікованих спеціалістів та ін.

Аналітичне групування показує закономірність, яка заключається в тому, що із ростом обсягу виробництва підвищується ефективність роботи підприємств. В нашому прикладі про це свідчить тенденція росту фондовіддачі.

Взаємозв'язок між досліджуваними ознаками проявляється також в тому, що із збільшенням факторної ознаки результативна може не тільки збільшуватись, але й зменшуватись. Наприклад, із збільшенням рівня продуктивності праці зменшується собівартість одиниці продукції, із ростом обсягу виробництва знижується фондомісткість продукції і т.д.

Аналітичне групування може проводитись як за факторною, так і за результативною ознаками в залежності від того, що є основним при статистичному дослідженні. Якщо вивчається вплив якоїсь однієї причини на різні явища – групування проводять за факторною ознакою, а якщо вивчають вплив різних причин на яке-небудь одне явище, то групують сукупність за результативною ознакою. Групування в цих випадках допомагають виявити і вивчити різноманітні зв'язки і залежності між варіаційними ознаками, які відображають всебічні властивості досліджуваних статистичних сукупностей.

Характерною особливістю аналітичного групування є те, що кожна група, утворена за суттєвою факторною ознакою, характеризується середніми величинами результативних ознак.

Аналітичні групування дозволяють, при більш глибокому аналізі, знайти форму і виміряти тісноту зв'язку між варіаційними ознаками і на цій основі зробити важливі практичні висновки для планування і прогнозування.

### 3.7. Вторинні групування

Всі раніше наведені групування, зроблені на основі первинного статистичного матеріалу називаються *первинними* групуваннями.

Іноді, при статистичному дослідженні, раніше згрупований матеріал доводиться перегруповувати.

*Вторинним групуванням* в статистиці називається процес утворення нових груп на основі раніше проведеного групування первинних даних.

Метод вторинного групування використовується для утворення, на основі групування за кількісною ознакою, якісно однорідних груп або типів, приведення кількох групувань з різними інтервалами до одного виду, з метою порівняльності та утворення більш укрупнених груп, в яких чіткіше простежується характер розподілу. Його застосовують також при економічному аналізі роботи господарств для приведення до порівняльного виду їх вже згрупованих даних, але не зіставних за територією чи періодами часу.

Статистика використовує два способи утворення нових груп. Перший, найбільш простий і розповсюджений – це спосіб зміни (як правило укрупнення) інтервалів первинного групування. В більшості випадків тут виходять із передбачення про нормальний розподіл ознак в середині інтервалів.

Другий спосіб вторинного групування базується на закріпленні за кожною групою певної частки одиниць сукупності (спосіб дольового перегрупування).

Перший спосіб вторинного групування застосовується для зведення двох і більше групувань з неоднаковими інтервалами до одного виду, з метою зіставлення. Проілюструємо цей спосіб перегрупування на прикладі (табл. 3.6).

Таблиця 3.6

Розподіл робітників за виконанням норм виробітку

| Базисний період | | Звітний період | |
|---|---|---|---|
| Групи робітників за виконанням норм виробітку | Число робітників, % | Групи робітників за виконанням норм виробітку | Число робітників, % |
| I До 80 | 0,8 | I До 90 | 2,0 |
| II 80 – 85 | 1,2 | II 90 – 100 | 10,0 |
| III 85 – 90 | 2,3 | III 100 – 110 | 71,0 |
| IV 90 – 100 | 2,5 | IV 110 – 120 | 11,0 |
| V 100 – 110 | 70,2 | V 120 і більше | 6,0 |
| VI 110 – 115 | 16,0 | | |
| VII 115 – 120 | 2,8 | | |
| VIII 120 – 125 | 2,2 | | |
| IX 125 і більше | 2,0 | | |
| Разом: | 100,0 | Разом: | 100,0 |

Порівняти ці два періоди між собою не можна через те, що вони мають різні інтервали. Але їх можна порівняти, провівши вторинне групування.

Узявши за основу групи робітників звітного періоду, встановлюємо, що в першу групу увійдуть всі три групи робітників базисного періоду, інтервали від 90 % до 100 % і 100 % – 110 % обох періодів співпали, в четверту групу (110 – 120 %) – шоста і сьома групи робітників базисного періоду, і в останню групу (120 % і більше) – восьма і дев'ята групи.

Таблиця 3.7

Групування робітників за виконанням норм виробітку

| Групи робітників за виконанням норм виробітку | Число робітників, % до підсумку | |
|---|---|---|
| | базисний період | звітній період |
| I До 90 | 4,3 | 2,0 |
| II 90 – 100 | 2,5 | 10,0 |
| III 100 – 110 | 70,2 | 71,0 |
| IV 110 – 120 | 18,8 | 11,0 |
| V 120 і більше | 4,2 | 6,0 |
| Разом: | 100,0 | 100,0 |

Як видно з таблиці 3.7, найбільшою є частка робітників (понад 80 %) з виконанням норм виробітку від 100 до 120 %. Вторинне групування дало можливість зіставити дані обох періодів.

Суть другого способу вторинного групування полягає в подрібненні груп за принципом пропорційності. Покажемо проведення вторинного групування цим способом на прикладі.

Таблиця 3.8

Розподіл продуктових магазинів області за обсягом роздрібного товарообороту

| Групи магазинів за обсягом річного роздрібного товарообороту, тис. грн. | Кількість магазинів | | Роздрібний товарооборот за рік | |
|---|---|---|---|---|
| | шт. | % | тис. грн. | % |
| I До 10 | 15 | 5,8 | 130 | 1,1 |
| II 10 – 15 | 20 | 7,7 | 260 | 2,1 |
| III 15 – 20 | 22 | 8,5 | 396 | 3,2 |
| IV 20 – 25 | 12 | 4,6 | 276 | 2,2 |
| V 25 – 30 | 8 | 3,1 | 216 | 1,7 |
| VI 30 – 35 | 9 | 3,5 | 293 | 2,4 |
| VII 35 – 40 | 33 | 12,7 | 1221 | 9,9 |
| VIII 40 – 50 | 42 | 16,1 | 1806 | 14,6 |
| IX 50 – 70 | 39 | 15,0 | 2457 | 19,9 |
| X 70 – 90 | 30 | 11,5 | 2400 | 19,4 |
| XI 90 – 100 | 25 | 9,6 | 2375 | 19,2 |
| XII більше 100 | 5 | 1,9 | 525 | 4,3 |
| Разом: | 260 | 100,0 | 12355 | 100,0 |

Згрупуємо наведені магазини продторгу за річним обсягом роздрібного товарообороту в три групи: малі, середні і великі. До малої групи віднесемо 50 % загальної кількості магазинів, середньої – 30 % і великої – 20 %. У малу групу (50 %) увійдуть перші сім груп, які разом становлять 45,9 % магазинів. До 50 % не вистачає 4,1 % магазинів (50 – 45,9 ), які візьмемо з восьмої групи. Ця група становить 16,1 % магазинів. Питома вага магазинів, яких не вистачає в малій групі, становить 4,1 : 16,1 = 0,255.

Очевидно, для визначення питомої ваги роздрібного товарообороту, що належить до малої групи, треба взяти повністю відповідні дані по перших семи групах і 0,255 частки від даних восьмої групи. Отже, відсоток роздрібного товарообороту малої групи становитиме:

$$22{,}6 + 14{,}6 \cdot 0{,}255 = 26{,}3\ \%.$$

Для складання середньої групи беруть 30 % магазинів, включаючи 12 % із залишених у восьмій групі, повністю дев’яту групу (15 %) і 3 % із десятої групи (12 + 15 + 3 = 30 %). Відповідно частки роздрібного товарообороту будуть: 12 % магазинів восьмої групи, 15 % дев’ятої групи і 3 % десятої групи: 10,9 + 19,9 + 0,261 · 19,4 = 35,9 %, де 0,261 = 3 : 11,5.

В третю (велику) групу (20%) магазинів увійдуть 8,5 % магазинів десятої групи і повністю одинадцята і дванадцята групи (9,6 % і 1,9 %); відповідно частки роздрібного товарообороту:

$$14{,}3 + 19{,}2 + 4{,}3 = 37{,}8\ \%.$$

Вторинне групування магазинів за обсягом річного роздрібного товарообороту покажемо в таблиці 3.9.

Таблиця 3.9

Розподіл продуктових магазинів за обсягом річного товарообороту

| Групи магазинів за обсягом річного роздрібного товарообороту, тис. грн. | Кількість магазинів (в % до підсумку) | Роздрібний товарооборот за рік (в % до підсумку) |
|---|---|---|
| Малі | 50 | 26,3 |
| Середні | 30 | 35,9 |
| Великі | 20 | 37,8 |
| Разом: | 100 | 100,0 |

Вторинне групування показало, що п’ята частина (20 %) великих магазинів володіють обсягом річного роздрібного товарообороту в розмірі 37,8 %, а великі і середні магазини разом, які становлять 50 % від їх загальної кількості володіють майже ¾ (73,7 %) обсягу річного роздрібного товарообороту.

Вторинне групування, для приведення двох групувань з різними розмірами інтервалів до єдиного виду, з метою порівняльності, деколи проводять двома способами одночасно.

Покажемо проведення такого групування на прикладі.

Таблиця 3.10

Розподіл робітників двох цехів одного промислового підприємства за розміром місячної заробітної плати

| Цех 1 | | | Цех 2 | | |
|---|---|---|---|---|---|
| Групи робітників за розміром заробітної плати, грн. | | Питома вага робітників в групі, % до підсумку | Групи робітників за розміром заробітної плати, грн. | | Питома вага робітників в групі, % до підсумку |
| I | 110 – 120 | 5 | | - | - |
| II | 120 – 130 | 8 | I | 120 – 135 | 8 |
| III | 130 – 140 | 10 | II | 135 – 150 | 12 |
| IV | 140 – 150 | 12 | III | 150 – 165 | 14 |
| V | 150 – 160 | 20 | IV | 165 – 180 | 18 |
| VI | 160 – 170 | 26 | V | 180 – 195 | 19 |
| VII | 170 – 180 | 8 | VI | 195 – 210 | 15 |
| VIII | 180 – 190 | 6 | VII | 210 – 225 | 8 |
| IX | 190 – 200 | 4 | VIII | 225 – 240 | 6 |
| X | 200 – 210 | 1 | | - | - |
| Разом: | | 100 | Разом: | | 100 |

Безпосередньо порівняти розподіл робітників за місячною заробітною платою по двох цехах неможливо, тому що ці розподіли мають різні інтервали: 10 грн. в цеху 1 і 15 грн. в цеху 2.

За допомогою вторинного групування приведемо ці групування до порівняльного виду. З цією метою перегрупуємо заробітну плату в групи, єдині для обох цехів, вибравши інтервал 30 грн.

В нашому випадку показники по цеху 1 об’єднаємо в групи сумуванням, використавши перший спосіб вторинного групування, а по цеху 2 – другий спосіб (пропорційне подрібнення інтервалів).

По цеху 1 в першу групу (110 – 140 грн.) увійдуть дані трьох перших груп первинного групування: 23 (5 + 8 + 10); в другу групу (140 – 170 грн.) – три наступні групи: 58 (12 + 20 + 26); в третю групу (170 – 200 грн.) – дані сьомої, восьмої і дев’ятої груп: 18 (8 + 6 + 4); в четверту групу (200 – 230 грн.) дані десятої групи (1), а новостворена п’ята група (230 – 260 грн.) по цьому цеху даних не має.

По цеху 2 в першу групу (110 – 140 грн.) попадуть дані першої групи (8 %) і $^1/_3$ [(140 – 135 = 5), $\frac{5}{15}=\frac{1}{3}$] з другої групи (4 %): $12\left(8+\frac{1}{3}12\right)$; в другу групу (140 – 170 грн.) – $^2/_3$ даних [(150 – 140 = 10), $\frac{10}{15}=\frac{2}{3}$] з другої групи (8 %), повністю дані третьої групи (14 %) і $^1/_3$ даних з четвертої групи (5 %): $28\left(\frac{2}{3}12+14+\frac{1}{3}15\right)$; і т.д.

Дані вторинного групування покажемо в таблиці 3.11.

Таблиця 3.11

Розподіл робітників двох цехів одного промислового підприємства за розміром місячної заробітної плати

| Групи робітників за розміром місячної заробітної плати, грн. | | Питома вага робітників в відсотках до підсумку | |
|---|---|---|---|
| | | цех 1 | цех 2 |
| I | 110 – 140 | 23 | 12 |
| II | 140 – 170 | 58 | 28 |
| III | 170 – 200 | 18 | 36 |
| IV | 200 – 230 | 1 | 20 |
| V | 230 – 260 | - | 4 |
| Разом: | | 100 | 100 |

Проведене вторинне групування по обох цехах дало можливість порівняти їх дані. Вони показують, що в цеху 2 місячна заробітна плата робітників більш диференційована, ніж в цеху 1.

За даними аналогічного прикладу можна провести вторинне групування з розширеними інтервалами, з метою більш чіткого виявлення характеру розподілу досліджуваної сукупності. Способи проведення вторинного групування тут такі ж, як і в попередньому прикладі.

### 3.8. Складні групування

Зв'язки між явищами суспільного життя складні і різноманітні, вони залежать від багатьох причин, в яких переплетені різні, часто суперечливі, тенденції. Для повнішого і глибшого дослідження процесів розвитку цих явищ, потрібно групувати дані за двома і більше ознаками. Такі групування в статистиці називаються *складними*.

Найбільш розповсюдженим видом складних групувань є комбіновані групування, при яких розчленовані на групи сукупності піддаються подальшому дрібненню на підгрупи за іншими додатковими ознаками.

Отже, *комбінованими* в статистиці називаються групування, коли утворені групи за однією ознакою, діляться потім на підгрупи за однією і більше ознаками, взятими в комбінації.

Одночасне використання декількох групувальних ознак дозволяє виявити і порівняти такі відомості і зв'язки між досліджуваними ознаками, які не можна знайти через ізольовані групування за рядом групувальних ознак. Прикладом комбінованого групування може служити таблиця 3.12.

Таблиця 3.12

Групування 40-а підприємств машинобудування за вартістю промислово-виробничих фондів і чисельністю промислово-виробничого персоналу

| Групи підприємств за вартістю промислово-виробничих фондів, млн. грн. | Підгрупи за чисельністю промислово-виробничого персоналу, чол. | Число підпри-ємств | Середньо-річна вартість промислово-виробничих фондів, млн. грн. | Середньо-спискова чисельність промислово-виробничого персоналу, чол. | Товарна продукція в середньому на одне підпри-ємство, млн. грн. |
|---|---|---|---|---|---|
| I 10 – 160 | I до 15000 | 18 | 80,7 | 6540 | 103,5 |
| | II 15000 і більше | 7 | 102,7 | 16258 | 166,4 |
| Всього по групі | | 25 | 86,9 | 9262 | 121,2 |
| II 160 – 310 | I до 15000 | 3 | 184,7 | 12633 | 142,7 |
| | II 15000 і більше | 5 | 230,0 | 20764 | 223,4 |
| Всього по групі | | 8 | 213,0 | 17715 | 193,1 |
| III 310 – 460 | I до 15000 | 3 | 328,0 | 12850 | 359,3 |
| | II 15000 і більше | 4 | 380,5 | 22077 | 513,7 |
| Всього по групі | | 7 | 358,0 | 18123 | 447,6 |
| Разом: | | 40 | 159,5 | 12503 | 192,7 |

В таблиці чітко простежується залежність результативного показника (товарної продукції) від вартості промислово-виробничих фондів. Підсумкові дані по групах показують, що із збільшенням вартості промислово-виробничих фондів зростає середня величина товарної продукції (121,2; 193,1; 447,6).

За даними таблиці також спостерігається залежність обсягу товарної продукції від чисельності промислово-виробничого персоналу. Для цього потрібно розглядати зміну результативного показника по підгрупах. В підгрупах всіх груп товарна продукція в середньому на одне підприємство зростає.

Таким чином, комбіноване групування за двома ознаками дозволяє розглянути вплив обох групувальних ознак на результативний показник.

В даний час комбіноване групування відіграє важливу роль в комплексному статистичному дослідженні соціально-економічних явищ.

Разом з тим, комбінація групувальних ознак веде до різкого збільшення числа груп, що може привести до недостатньої чисельності одиниць в кожній групі, а це, в свою чергу – до випадкових і необґрунтованих висновків. Також, при вивченні впливу великої кількості ознак, застосування комбінованого групування стає неможливим через надмірне подрібнення інформації при побудові комбінованих таблиць, що затушовує прояв закономірностей і не дозволяє простежити одночасний вплив всього комплексу факторних ознак на досліджуваний результативний показник.

До комбінованих групувань вдаються при достатньо чисельному первинному матеріалі і якщо завдання заключається в дослідженні залежності результативних ознак від декількох факторних ознак. На практиці, при проведенні комбінованого групування, обмежуються трьома-чотирма ознаками. Все це вимагає пошуків нових принципів групування, які б зняли вказані вище обмеження.

Для розв’язання цього завдання математичною статистикою розроблений *метод багатовимірних групувань – метод розпізнавання образів*. В цьому випадку явища класифікуються за принципом подібності (близькості), які потім відносять до вже сформованих груп (образів) нових одиниць. На відміну від комбінованого групування, де віднесення до певних груп відбувається за окремою ознакою, при класифікації на основі розпізнавання образів віднесення до окремих груп проводять за комплексом ознак, які утворюють “ознаковий простір”, де кожна ознака розглядається як координата.

Якщо в наборі сукупності є ряд одиниць з фіксованими ознаками, то кожна з цих одиниць розглядається як точка в багатовимірному просторі. Всі точки, розташовані в багатовимірному просторі, утворюють скупчення, які складають групи (таксони, кластери). Завданням виділення цих скупчень точок в просторі і займається теорія розпізнавання образів.

Геометрична близькість декількох точок в багатовимірному просторі означає подібність становища відповідних об’єктів, тобто їх кількісна однорідність за досліджуваними ознаками.

Одним з розповсюджених критеріїв міри близькості між об’єктами є евклідова віддаль:

$$d_{ij} = \sqrt{\sum_{i=1}^{n}\left(x_{ki} - x_{kj}\right)^2} ,$$

де $x_{ki}$ – значення k-ої ознаки в i-му об’єкті;

$x_{kj}$ – значення k-ої ознаки в j-му об’єкті.

При зменшенні цієї віддалі близькість зростає, а при її збільшенні – зменшується.

В різних алгоритмах розпізнавання образів проводиться неоднаково, але загальним для всіх є те, що групи формуються на основі подібності об’єктів за великим числом ознак. Знаходження цих груп здійснюється одним із методів статистичної теорії розпізнавання образів – методом кластерного аналізу з допомогою сучасних ПЕОМ.

Багатовимірні групування дозволяють розв’язати цілий ряд важливих завдань економіко-статистичного дослідження, таких як побудова статистичних моделей економічних показників роботи підприємства і його частин, формування однорідних сукупностей, відбір суттєвих факторних ознак, виділення типових груп об’єктів та багато інших.

### 3.9. Необхідність створення системи групувань, та основні вимоги до них

Завдання всебічного аналізу соціально-економічних явищ і процесів не розв’язуються через складання якогось одного групування, яке б характеризувало типи, структуру і взаємозв’язок даного економічного явища. Такий аналіз вимагає складання системи групувань за багатьма ознаками, яка допоможе охарактеризувати розвиток певного явища в цілому.

Утворені системи групувань повинні відповідати загальним методологічним вимогам і підпорядковуватись логічним і формальним критеріям.

*Логічні критерії* вимагають наступного:

а) система повинна всебічно характеризувати досліджуваний об’єкт;

б) системою групувань повинні розв’язуватись типологічні, структурні і аналітичні завдання дослідження;

в) кожне окреме групування повинно бути однією з логічних ланок в загальній системі;

г) висновки за кожним групуванням не повинні суперечити одні одним;

д) система групувань повинна бути стабільною і не піддаватись частим змінам в часі.

Основними *формальними критеріями* є:

а) групування за якісними ознаками повинні передувати групуванням за кількісними ознаками;

б) результативні ознаки повинні бути виражені в однакових для всієї системи абсолютних, відносних і середніх показниках;

в) таблиці системи групувань повинні мати протягом тривалого часу стабільну нумерацію;

г) інтервали групувань також повинні бути стабільними і без потреби часто не змінюватись.

Дотримання вищенаведених вимог робить систему групувань більш гнучкою і одночасно стабільною, що є необхідною умовою всебічного економіко-статистичного аналізу.

### ***Контрольні запитання***

1. Поняття про статистичне зведення.
2. Завдання зведення і його основний зміст.
3. Групування як основа наукової обробки статистичних даних.
4. Завдання групувань і їх значення в статистичних дослідженнях.
5. Основні правила утворення груп.
6. Інтервали статистичних групувань.
7. Поняття про спеціалізовані інтервали.
8. Вибір групувальних ознак.
9. Поняття типологічних групувань.
10. Поняття структурних групувань.
11. Поняття аналітичних групувань.
12. Поняття про факторні і результативні ознаки.
13. Поняття про вторинні групування.
14. Прості і складні групування.
15. Поняття про комбіновані групування.
16. Багатовимірні групування.
17. Необхідність створення системи групувань.
18. Основні вимоги до системи групувань.

# Розділ 4 Статистичні таблиці

## 4.1. Завдання та значення табличного методу викладу статистичних даних

Результати статистичного зведення і групування викладаються, як правило, у формі таблиць. В науковій діяльності і практичній роботі досить широко застосовуються таблиці. Важливу роль вони відіграють в економічній роботі, а тому кожний економіст повинен вміти правильно скласти статистичну таблицю і зробити вірні висновки.

Важливою вимогою, яка пред’являється до статистичної таблиці, – це подання досліджуваного матеріалу в наглядній для читача формі.

Основною особливістю табличного викладу матеріалу є те, що показники, які характеризуються в статистичній таблиці можна об’єднати під одним загальним заголовком.

Проте не всяка таблиця являється статистичною. На відміну від допоміжних розрахункових таблиць (логарифмів, коефіцієнтів, множення та ін.) статистичними таблицями можуть вважатися тільки ті, що містять інформацію статистичного аналізу соціально-економічних явищ і процесів.

Отже, *статистична таблиця* – це форма раціонального, наочного та систематизованого викладення числових характеристик досліджуваних суспільних явищ і процесів.

Вперше статистичні таблиці були застосовані російським географом і статистиком І.К.Кириловим в 1727 р., за допомогою яких він описав окремі області держави і всю державу в цілому.

У статистичній таблиці за аналогією з граматичним реченням є підмет і присудок. *Підметом* таблиці являється статистична сукупність, тобто ті об’єкти або їх частини, які характеризуються рядом числових показників. *Присудком* статистичної таблиці називаються показники, що характеризують досліджувану сукупність і її частини.

Підмет таблиці переважно розташовується зліва і складає зміст рядків. Присудок таблиці, як правило, поміщається зверху і складає зміст колонок. Підмет і присудок таблиці інколи можуть мінятись місцями.

Кожна таблиця має ряд горизонтальних рядків і вертикальних граф (стовпчиків, колонок), які при перетині утворюють клітки, котрі заповнюються статистичними даними.

Розграфлена сітка (без слів і цифр) складає скелет таблиці. Отже, ззовні таблиці являють собою перетин рядків і колонок. Якщо записати назви підмета і присудка, то матимемо макет таблиці. Для отримання повної таблиці потрібно заповнити всі клітки відповідними статистичними даними.

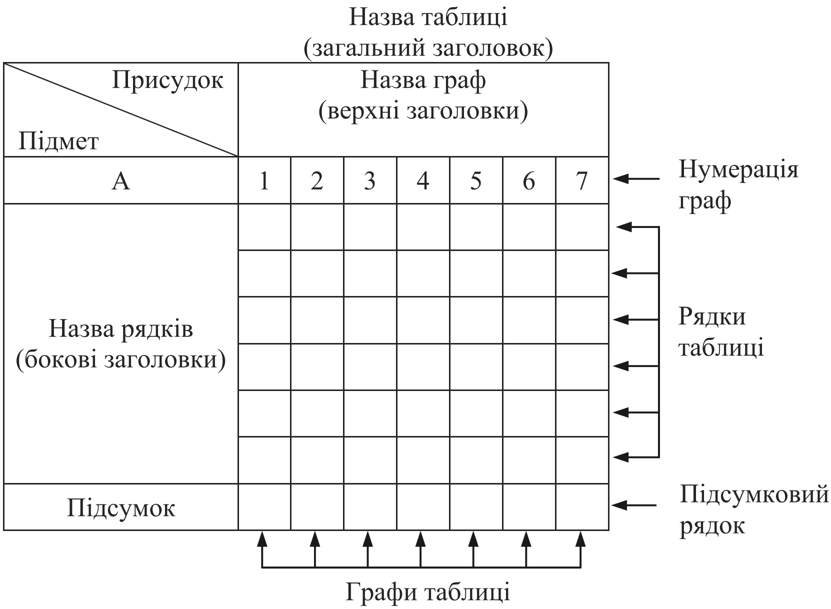
Обов'язковою складовою частиною статистичної таблиці є загальний і внутрішні заголовки. Загальний заголовок розташовується над таблицею, в якому коротко зазначається, про що йде мова в таблиці, до якого місця і часу вона відноситься і в яких одиницях наведені дані.

Внутрішні заголовки розміщаються усередині таблиці (збоку і зверху).

Приступаючи до складання статистичної таблиці, потрібно перш за все розробити її макет, що є важливою умовою планування опрацювання статистичних матеріалів.

Макет статистичної таблиці

Таблиця 4.1




До деяких таблиць подаються примітки, де роз'яснюється зміст окремих заголовків чи показників.

При складанні статистичних таблиць необхідно дотримуватись ряду правил їх побудови, які ми розглянемо нижче.

## 4.2. Види статистичних таблиць

В залежності від побудови підмета розрізняють три види статистичних таблиць: прості, групові і комбінаційні.

В простих таблицях в підметі перераховуються лише одиниці сукупності, в групових – цифрові дані об'єднуються в групи, а в комбінаційних – групи ще розбиваються на підгрупи. В цій класифікації відмінність між простою і груповою таблицями заключається в різному прийомі об'єднання цифрових даних. Відмінність же між груповою і комбінаційною таблицями заключається в різному числі ознак, за якими проходить розчленування одиниць сукупності. Завдання простої таблиці – дати лише перелік ознак, а групової і комбінаційної – у вивченні взаємозв'язків ознак досліджуваного явища.

В статистичній практиці найбільш широке розповсюдження мають прості таблиці.

*Простими* називаються статистичні таблиці в підметі яких немає груп.

До простих таблиць належить, наприклад, статистична характеристика міст і районів даної області з перерахуванням у підметі всіх районів, характеристика всіх підприємств даного міста, зведення про хід сільськогосподарських робіт, про виробництво і реалізацію м'яса, молока, яєць і т.д. по районах області або по сільськогосподарських колективах району, підсумки різних переписів в розрізі окремих регіонів, змалювання якого-небудь явища в часі. Прості таблиці ще називають *інформаційними*.

Вся різноманітність простих статистичних таблиць може бути зведена до перелікових, територіальних і хронологічних, або до деяких їх сполучень.

Простими *переліковими* таблицями називаються такі, в підметі яких наводиться перелік одиниць або показників, що вивчаються. Так, наприклад, за даними Державного Комітету Статистики України про виробництво основних видів промислової продукції може бути складена така таблиця:

Простими *територіальними* називаються таблиці, в підметі яких дається перелік територій (країн, областей, районів), кожна з яких характеризується відповідними показниками. Наприклад, характеристика територій і чисельності населення України на 1 січня 1996 р. (табл. 4.3).

Простими *хронологічними* називаються таблиці, в підметі яких наводяться певні відрізки часу, а в присудку – один або декілька статистичних показників. Таблиця залишається також простою і в тому випадку, якщо в ній одночасно поєднуються одиниці об'єкта або території, і відрізки часу.

Таблиця 4.2

Виробництво основних видів продукції промисловості в Україні за 1995 р.*

| Продукція | Одиниці виміру | 1995 р. |
|---|---|---|
| Електроенергія | млрд. кВт-год. | 194,0 |
| Нафта з газовим конденсатом | тис. т | 4100,0 |
| Газ | млрд. $м^3$ | 18,2 |
| Вугілля | млн. т | 83,8 |
| Сталь | млн. т | 22,3 |
| Готовий прокат | млн. т | 17,8 |
| Мінеральні добрива | тис. т | 2200,0 |
| Волокна та нитки | тис. т | 41,3 |
| Цемент | тис. т | 7,6 |
| Папір | тис. т | 98,0 |
| Тканини | млн. $м^2$ | 169,0 |
| Взуття | млн. пар | 20,6 |
| Трикотажні вироби | млн. шт. | 27,0 |
| Телевізори | тис. шт. | 315,0 |
| Холодильники побутові | тис. шт. | 562,0 |
| Пральні машини | тис. шт. | 213,0 |

Таблиця 4.3

Територія і чисельність населення а Автономній Республіці Крим і областях України (на 1 січня 1996 р.)**

| | Територія, тис. $км^2$ | Чисельність населення, тис. чол. | Міське населення, % | Густота населення, чол./$км^2$ |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| Україна | 603,7 | 51334,1 | 68 | 85,0 |
| Автономна Республіка Крим | 26,1 | 2205,6 | 63 | 84,5 |

* Розміщення продуктивних сил України: Підручник / Є.П.Качан, М.О.Ковтонюк, М.О.Петрига та ін.: За ред. Є.П.Качана. – К.: Вища школа, 1997. – С. 59.

** Розміщення продуктивних сил України: Підручник / Є.П.Качан, М.О.Ковтонюк, М.О.Петрига та ін.: За ред. Є.П.Качана. – К.: Вища школа, 1997. – С. 37.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Вінницька | 26,5 | 1876,0 | 47 | 70,8 |
| Волинська | 20,2 | 1075,2 | 52 | 53,2 |
| Дніпропетровська | 31,9 | 3852,2 | 84 | 120,8 |
| ............................. | ..... | ..... | ..... | ..... |
| Тернопільська | 13,8 | 1175,4 | 44 | 85,2 |
| .............................. | ..... | ..... | ..... | ..... |
| м. Київ | 0,8 | 2638,7 | 100 | 3298,4 |
| м. Севастополь | 0,9 | 406,9 | 94 | 452,1 |

Візьмемо для прикладу чисельність населення, зайнятого в народному господарстві Подільського економічного району (тис. чол., табл. 4.4)*.

Таблиця 4.4

| Область | Роки | | | | | |
|---|---|---|---|---|---|---|
| | 1985 | 1990 | 1991 | 1992 | 1993 | 1995 |
| Вінницька | 920,5 | 882,5 | 883,9 | 839,3 | 812,7 | 739,8 |
| Тернопільська | 528,7 | 533,9 | 527,5 | 508,2 | 496,6 | 445,1 |
| Хмельницька | 721,0 | 708,2 | 704,3 | 676,7 | 665,1 | 601,9 |
| Подільський економічний район | 2170,2 | 2124,6 | 2115,7 | 2024,2 | 1974,4 | 1786,8 |

Прості статистичні таблиці є лише підсумковими зведеннями, тому вони не дозволяють виявляти типів досліджуваного явища, його структури, а також взаємозв’язків між окремими ознаками. Значно більші можливості мають групові і тим більше комбінаційні таблиці.

*Груповими* називаються такі статистичні таблиці, підмет яких утворений в результаті групування одиниць досліджуваного об’єкта за тією чи іншою ознакою. Тобто, групові таблиці виникають в результаті проведення групування при зведенні статистичних матеріалів.

Найпростішим видом групових таблиць є ряди розподілу, утворені на основі кількісних або атрибутивних ознак. Наприклад:

Таблиця показує, що найбільше населення проживає в малих, великих містах та містах мільйонерах (відповідно 31,3; 24,9 і 21,6 %), а в середніх і дуже великих містах, при значній різниці в їх кількості (55 і 6) проживає приблизно однакова чисельність населення (11,0 і 11,2 %).

* Розміщення продуктивних сил України: Підручник / Є.П.Качан, М.О.Ковтонюк, М.О.Петрига та ін.: За ред. Є.П.Качана. – К.: Вища школа, 1997. – С. 312.

Таблиця 4.5

Структура міст і міського населення України (на 1 січня 1993 р.)*

| Міста і селища міського типу | Всього міст | | В них населення | |
|---|---|---|---|---|
| | к-сть | % | тис. чол. | % |
| Малі (до 50 тис. чол.) | 1242 | 92,0 | 11119,1 | 31,3 |
| Середні (50 – 100 тис. чол.) | 55 | 4,1 | 3884,1 | 11,0 |
| Великі (100 – 500 тис. чол.) | 40 | 3,0 | 8820,0 | 24,9 |
| Дуже великі (500 – 1000 тис. чол.) | 6 | 0,5 | 3993,3 | 11,2 |
| Міста мільйонери (понад 1000 тис.чол.) | 5 | 0,4 | 7654,5 | 21,6 |
| Всього: | 1356 | 100,0 | 35471,0 | 100,0 |

В більш складних групових таблицях присудок містить поряд з числом одиниць, що входять до тієї чи іншої групи, ряд інших показників, які належать до характеристики груп підмета. Наприклад, робота 100 заводів однієї із галузей промисловості за звітний рік характеризується наступними даними (табл. 4.6, дані умовні).

Таблиця 4.6

Групування заводів за вартістю основних виробничих фондів і випуском продукції

| Групи заводів за вартістю основних виробничих фондів, млн. грн. | | Кіль-кість заво-дів | Середні рівні, млн. грн. | | Фондовіддача, грн. (відношення вартості продукції до фондів) | Фондоємкість, грн. (відношення вартості фондів до продукції) |
|---|---|---|---|---|---|---|
| | | | основних виробничих фондів | випуску продукції | | |
| I | 1 – 3 | 5 | 2,25 | 2,90 | 1,29 | 0,77 |
| II | 3 – 5 | 19 | 4,13 | 5,55 | 1,34 | 0,75 |
| III | 5 – 7 | 53 | 6,08 | 8,26 | 1,36 | 0,74 |
| IV | 7 – 9 | 17 | 8,03 | 11,09 | 1,38 | 0,72 |
| V | 9 – 11 | 6 | 10,47 | 14,76 | 1,41 | 0,71 |

Підмет наведеного вище групування заводів показує структурний розподіл за кількісною ознакою. В таблиці чітко простежується зв’язок між вартістю основних виробничих фондів і випуском продукції. Тобто, із збільшенням вартості основних виробничих фондів в середньому на один завод випуск продукції зростає, росте фондовіддача і знижується фондоємкість.

Таблиця дозволяє провести глибокий аналіз зв’язку між вартістю основних виробничих фондів і випуском продукції, а саме:

* Населення України, 1992. – К., 1993. – С. 15.

виявити його напрямок, встановити силу, оцінити вплив факторної ознаки на зміну результативної і ін.

Наведемо ще один приклад групової таблиці, в якій показано групування населення за атрибутивною ознакою. За матеріалами анкетного обстеження населення Західного регіону України отримані наступні дані (табл. 4.7).

Таблиця 4.7

Розподіл сімей України за кількістю наявних в користуванні товарів культурно-побутового призначення в залежності від місця проживання.

<table>
<tr><th rowspan="2">Групи сімей за місцем проживання</th><th colspan="7">Доля сімей, які мають в користуванні 2-3 і більше товарів культурно-побутового призначення, в загальній кількості, %</th></tr>
<tr><th>радіо-приймачі і радіоли</th><th>телеві-зори різні</th><th>фото-апарати</th><th>швейні машини</th><th>пральні машини</th><th>електро-пилососи</th><th>велосипеди і мопеди</th></tr>
<tr><td>1. Місто</td><td>53,8</td><td>50,0</td><td>70,8</td><td>63,2</td><td>63,6</td><td>92,0</td><td>61,1</td></tr>
<tr><td>2. Село</td><td>46,2</td><td>50,0</td><td>29,2</td><td>36,8</td><td>36,4</td><td>8,0</td><td>38,9</td></tr>
<tr><td>Разом</td><td>100,0</td><td>100,0</td><td>100,0</td><td>100,0</td><td>100,0</td><td>100,0</td><td>100,0</td></tr>
</table>

Дані таблиці показують, що рівень забезпеченості товарами культурно-побутового призначення міських сімей значно перевищує сільських.

*Комбінаційними* називаються статистичні таблиці, у підметі яких групи утворені за однією ознакою діляться на підгрупи за іншими ознаками. Прикладом комбінаційної таблиці є табл. 4.8.

Таблиця 4.8

Вплив стажу роботи і тарифного розряду на заробітну плату робітників

<table>
<tr><th colspan="2" rowspan="2">Групи робітників за стажем роботи, р.</th><th rowspan="2">Підгрупи робітників за тарифним розрядом</th><th rowspan="2">Число робітників, чол.</th><th colspan="3">Середні рівні</th></tr>
<tr><th>стажу роботи, р.</th><th>тарифного розряду</th><th>заробітної плати, грн.</th></tr>
<tr><td rowspan="2">I</td><td rowspan="2">До 5 р. включно</td><td>I – III</td><td>5</td><td>2,2</td><td>2,4</td><td>120,00</td></tr>
<tr><td>IV – VI</td><td>3</td><td>4,3</td><td>4,7</td><td>153,33</td></tr>
<tr><td colspan="3">В середньому по групі</td><td>х</td><td>3,0</td><td>3,2</td><td>132,50</td></tr>
<tr><td rowspan="2">II</td><td rowspan="2">Більше 5 р.</td><td>I – III</td><td>4</td><td>7,0</td><td>2,8</td><td>134,50</td></tr>
<tr><td>IV – VI</td><td>8</td><td>8,9</td><td>4,6</td><td>162,75</td></tr>
<tr><td colspan="3">В середньому по групі</td><td>х</td><td>8,2</td><td>4,0</td><td>153,33</td></tr>
<tr><td colspan="4">В середньому по всіх робітниках</td><td>6,2</td><td>3,7</td><td>145,00</td></tr>
</table>

Як видно із комбінаційної таблиці із збільшенням стажу роботи і тарифного розряду заробітна плата в середньому на одного робітника зростає.

Пізнавальна властивість комбінаційних таблиць заключається в тому, що вони дозволяють простежити вплив на ознаки присудка не одного, а двох і більше факторів, закладених в основу комбінаційного групування.

Для того щоб вияснити, як змінюється, наприклад, заробітна плата в залежності від зміни стажу роботи робітника, достатньо простежити, як змінились середні дані по великих групах, а для визначення її залежності від тарифного розряду потрібно порівняти дані по підгрупах хоча б однієї з груп. Порівняння з підгрупами інших груп або підтверджує, або відкидає встановлений зв'язок.

Комбінаційна таблиця показує дію двох і більше факторів разом, виявляє взаємозв'язок між ними. Для виявлення дії одного фактора досить простежити зміну ознаки в однойменних групах, а для виявлення дії всіх факторів, потрібно проаналізувати дані всієї сукупності.

Групування за трьома, чотирма і більше ознаками, з побудовою відповідних таблиць, на практиці зустрічається дуже рідко, так як число утворюваних підгруп зростає в геометричній прогресії, що робить таблицю трудно оглядовою.

У випадку коли потрібно проаналізувати більше трьох взаємозв'язаних ознак в підметі таблиці вказують три ознаки в комбінації між собою, а четверту і наступні виносять в заголовок таблиці і замість однієї складають декілька таблиць.

Інколи в комбінаційних таблицях групи за однією ознакою розташовують в підметі, а за іншою – в присудку. Наприклад, забезпеченість сімей набором найбільш важливих товарів культурно-побутового призначення (телевізорами, холодильниками, пральними і швейними машинами та електропилососами) в залежності від віку сімей і середньодушевого грошового доходу наведена в табл. 4.9.

Таблиці 4.9

(в розрахунку на 100 сімей)

| Групи сімей за їх віком, р. | Групи сімей за середньодушовим грошовим доходом, гривень в місяць | | | | | |
|---|---|---|---|---|---|---|
| | до 50 | 51 – 60 | 61 – 80 | 81 – 100 | 101 – 130 | 130 і більше |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| До 5 | 14 | 13 | 24 | 37 | 64 | 73 |
| 6 – 10 | 15 | 14 | 29 | 47 | 54 | 97 |

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| 11 – 15 | 17 | 19 | 24 | 40 | 65 | 98 |
| 16 – 20 | 17 | 18 | 36 | 49 | 87 | 100 |
| 21 – 25 | 20 | 24 | 35 | 46 | 92 | 101 |
| 26 і більше | 29 | 33 | 47 | 63 | 96 | 112 |

В комбінаційних таблицях не можна довільно міняти місцями ознаки, які знаходяться у певній комбінації. Їх потрібно розставляти за вагомістю або послідовністю вивчення.

Групові і комбінаційні статистичні таблиці дають всебічну характеристику суспільних явищ і процесів для глибокого і всебічного економічного аналізу.

Крім розглянутих вище таблиць, статистика використовує також типові і балансові таблиці.

*Типовими* називаються статистичні таблиці, підмет яких утворений в результаті типологічного групування. Прикладом типової таблиці виступає табл. 3.1.

*Балансовими* називаються таблиці, які характеризують зв'язок між поступленням і витрачанням ресурсів. Прикладом балансової таблиці є табл. 4.10.

Таблиця 4.10

Схема
балансу грошових доходів і видатків населення

| Доходи | Всього | Видатки | Всього |
|---|---|---|---|
| А. Поступлення від державних, кооперативних, приватних та інших підприємств і організації – всього<br>в т.ч. за видами поступлень<br>Б. Поступлення від обміну товарами і послугами між групами населення | | А. Видатки здійснені через державні, кооперативні, приватні та інші підприємства і організації – всього<br>в т.ч. за видами видатків<br>Б. Видатки на оплату товарів і послуг в порядку обміну між групами населення | |
| Всього грошових доходів | | Всього грошових видатків | |
| Перевищення доходів над видатками | | Перевищення видатків над доходами | |
| Баланс | | Баланс | |

Балансові таблиці використовуються при складанні: балансу виробництва, споживання і нагромадження суспільного продукту (зведений матеріальний баланс); балансу виробництва, розподілу, перерозподілу і кінцевого використання суспільного продукту і національного доходу (зведений фінансовий баланс); звітного міжгалузевого балансу виробництва і розподілу продукції; балансу основних фондів народного господарства; балансу трудових ресурсів і ін.

Всі розглянуті вище статистичні таблиці називають ще результативними або публікаційними. Для їх розробки в процесі зведення матеріалів статистичного спостереження складають робочі або допоміжні розрахункові таблиці. Вони служать для розноски матеріалів спостереження по групах, різних допоміжних розрахунків, дають можливість швидше знайти помилки, тобто вони є першим етапом побудови результативних і публікаційних таблиць.

### 4.3. Розробка присудка таблиці

В присудку статистичної таблиці наводяться показники, які характеризують досліджуване явище. Вони можуть бути подані або в згорнутому або в розгорнутому вигляді. Згорнутий присудок містять вищенаведені таблиці 4.2, 4.3, 4.5. Розгорнутий присудок містить табл. 4.11, в якій всі спеціалісти діляться за статтю і освітою.

Таблиця 4.11

Чисельність спеціалістів з вищою і середньою спеціальною освітою, зайнятих в народному господарстві (дані умовні).

| Рік | Всього спеціалістів з вищою і середньою спеціальною освітою, тис. чол. | В тому числі | | Відсоток жінок в загальній чисельності спеціалістів з вищою і середньою спеціальною освітою |
|---|---|---|---|---|
| | | з вищою освітою | з середньою спеціальною освітою | |
| 1960 | 864 | 312 | 552 | 36 |
| 1970 | 5189 | 1865 | 3324 | 59 |
| 1980 | 9900 | 3568 | 6332 | 59 |
| 1990 | 16956 | 6410 | 10546 | 59 |
| 2000 | 20757 | 8115 | 12642 | 60 |

Розробка показників присудка таблиці може проводитися двома способами: а) способом простої розробки (табл. 4.11); б) способом складної (комбінованої розробки (табл. 4.12).

При простій розробці складові частини показника в присудку таблиці не діляться на підгрупи.

При складній розробці складові частини показника в присудку таблиці діляться на підгрупи. Якщо, наприклад, при простій розробці присудка ми дістали інформацію про розподіл спеціалістів за освітою, то при складній (комбінованій) – ми можемо додатково отримати дані про їх розподіл за статтю, профілем освіти, місцем проживання, галузями народного господарства і ін.

Таблиці 4.12

Чисельність спеціалістів з вищою і середньою спеціальною освітою зайнятих в народному господарстві (дані умовні)

| Рік | Всього спеціалістів зайнятих в народному господарстві, тис. чол. | В тому числі з освітою | | | | | |
|---|---|---|---|---|---|---|---|
| | | вищою | | | середньою спеціальною | | |
| | | чоловіки | жінки | обидві статі | чоловіки | жінки | обидві статі |
| 1960 | 989 | 200 | 112 | 312 | 353 | 199 | 552 |
| 1970 | 5624 | 765 | 1100 | 1865 | 1363 | 1961 | 3324 |
| 1980 | 10519 | 1463 | 2105 | 3568 | 2596 | 3736 | 6332 |
| 1990 | 17681 | 2628 | 3782 | 6410 | 4324 | 6222 | 10546 |
| 2000 | 21657 | 3246 | 4869 | 8115 | 5057 | 7585 | 12642 |

Більш глибоку статистичну характеристику спеціалістів, чисельність їх за освітою і статтю в поєднанні цих ознак дає комбінаційна розробка обох показників. Тут обидва показники присудка таблиці зв'язані між собою, що дає можливість простежити чисельність спеціалістів, не тільки за освітою, але й визначити число чоловіків і жінок в даних освітніх групах.

В таблицях з простою розробкою присудка кожна група може мати стільки даних, скільки ознак зареєстровано в статистичному спостереженні. При складній розробці присудка кожна група може бути охарактеризована різною комбінацією цих ознак, що залежать від завдання статистичного дослідження.

Складна розробка показників присудка при всіх інших однакових умовах дає більш повну характеристику об'єкту, поряд з цим, вона приводить до збільшення розмірів таблиць, що зменшує наочність і забруднення користування ними.

Комбінована розробка присудка таблиці дає великі можливості для статистичного аналізу. Із таблиць з комбінованою розробкою присудка можна отримати ряд таблиць з простою розробкою присудка.

Тому при проектуванні макетів статистичних таблиць потрібно детально проаналізувати, які ознаки піддати простій а які складній розробці.

Як при комбінованій, так і при простій розробці присудка, таблиці можуть бути простими, груповими і комбінаційними.

Отже, вид таблиці не залежить від способу розробки присудка, а виключно від розробки її підмета.

### 4.4. Оформлення статистичних таблиць

Глибоку і виразну характеристику досліджуваних явищ і процесів можуть дати лише грамотно і раціонально побудовані статистичні таблиці.

Наукою і практикою вироблені наступні основні правила складання і оформлення статистичних таблиць:

1. Перш за все необхідно розробити макет таблиці. При потребі форма статистичної таблиці повинна бути погоджена з раніше складеними таблицями для забезпечення можливості порівняння даних.
2. Таблиця повинна бути по можливості невеликою за розміром, включати тільки ті відомості, які необхідні для вивчення певного явища, таку таблицю легше аналізувати. Інколи доцільно замість однієї великої, побудувати декілька невеликих таблиць.
3. Загальний заголовок таблиці повинен коротко відображати основний зміст таблиці, а не перераховувати всього того, що в ній викладено. Якщо в таблиці приводяться дані, які охоплюють не всю сукупність, а тільки її частину (наприклад, продукція всієї промисловості і продукція легкої промисловості, або об’єм товарообороту по всіх не продуктових товарах і товарах культурно-побутового призначення), то в заголовку таблиці потрібно виділити слова, які окреслюють, яку частину сукупності охоплюють дані таблиці. В заголовку таблиці повинні бути вказані місце, період чи момент часу, на які припадають наведені в ній дані і загальна одиниця виміру (виноситься за межі таблиці і береться в дужки).

Якщо змістом таблиці є динамічні порівняння з будь-яким базисним роком за всіма елементами присудка, тоді в загальному заголовку в дужках проставляється рік порівняння, наприклад: (в порівнянні з 1990 р.).

Якщо в таблиці наводяться дані, не порівняльні за територіальною ознакою, то в її заголовку в лапках робиться застереження про межі досліджуваної сукупності, наприклад: “в межах відповідного року”.

На кінець, якщо для дослідження якогось явища складається ряд таблиць, то після заголовка в кожній з них проставляється порядковий номер.

4. Заголовки рядків підмета і граф присудка повинні формулюватись коротко, чітко і змістовно. Рядки в підметі і графи в присудку можуть нумеруватись, якщо число показників в них велике. При цьому графи підмета нумеруються великими літерами алфавіту, а графи з показниками присудка – арабськими цифрами.

Назви заголовків підмета і присудка пишуться з великої літери, підзаголовків – з малої, слова в них, по можливості, не скорочуються.

Якщо відсутня загальна одиниця виміру, то в кожній графі проставляється своя одиниця виміру, яка відділяється від речення комою, або береться в дужки і, як правило, пишеться скорочено (грн., шт., т, м, л, кВт-год., люд-дн., і т.д.). У випадку великої різноманітності одиниць виміру в наступній за підметом виділяється спеціальна графа "Одиниці виміру".

Адміністративний розподіл території в підметі таблиці розташовується за алфавітом. Інколи, з аналітичною метою їх розміщують в порядку зростання або спадання показника, який оцінює місце даної адміністративно-територіальної одиниці серед інших. Дані за багато років розташовують в хронологічному порядку.

Частини підмета і показники присудка, як правило, розташовуються за принципом від часткового до загального, наприклад, продукція рослинництва, продукція тваринництва, продукція всього сільського господарства. Якщо наводяться не всі складові, а виділяються найбільш важливі з них, то спочатку показують загальний підсумок, а потім пишуть "в тому числі" і виділяють найбільш важливі його складові частини, наприклад, всього населення зайнятого в народному господарстві, в тому числі: з вищою освітою; середньою спеціальною; і т.д.

Для зручності користування таблицями з абсолютними і відносними величинами за ряд років, спочатку наводять абсолютні величини, а потім – відповідні до них відносні дані, наприклад, магазин при плані 200 тис. грн., реалізував товарів на 215 тис. грн., тобто виконав план на 107,5 %.

У таблицях, які характеризують зв'язані між собою дані, для кращої наочності і порівнянності, групи підмета і графи присудка повинні розташовуватись в однаковій послідовності.

5. При заповненні кліток таблиці потрібно точно дотримуватись умовних позначень. Якщо дане явище відсутнє (при нульовому значенні ознаки) ставиться тире (–), якщо ж немає відомостей про його

розмір записується “немає відомостей” або проставляється три крапки (...). Коли величина показника не перевищує 0,05 ставиться 0,0. Якщо числове значення ознаки не має логічного змісту в клітці ставиться знак множення (х).

6. Кількісні показники в межах однієї графи повинні бути округлені з однаковим ступенем точності, тобто до 0,1 до 0,01, до 0,001 і т.д. При застосуванні багатозначних чисел потрібно кожні три цифри (починаючи справа) використовувати просвіти розбиті на класи мільйонів, тисяч і одиниць, наприклад: 8 021 948. Оперуючи числами із 6 – 7 знаків, інколи доцільно їх заокруглити до 2 – 3 знаків (8,02 млн.). Коли показники в процентах виражаються великими числами, то починаючи від 300 і більше їх замінюють відношенням в разах. Наприклад, пишуть не 948 %, а в 9,5 раз.

7. Якщо наводяться крім звітних ще й розрахункові дані, про це роблять застереження в таблиці або в примітці до неї. В примітці також вказуються джерела запозичених даних і інші необхідні пояснення. Якщо ж в таблиці наводяться взаємозв’язані дані, наприклад, абсолютні прирости, темпи росту, темпи приросту, то їх потрібно розташовувати в поряд стоячих графах.

8. Таблиці, як правило, повинні бути замкненими, тобто мати підсумковий рядок даних, за винятком аналітичних таблиць, в яких підсумки не обов’язкові.

Вимогами правил побудови статистичних таблиць нехтувати не можна. Правильно і грамотно складені статистичні таблиці є важливим засобом викладу опрацьованого матеріалу для економічного аналізу.

### 4.5. Аналіз статистичних таблиць

Щоб навчитись користуватися таблицями потрібно вміти їх читати і аналізувати. Читання і аналіз статистичних таблиць має велике науково-пізнавальне і практичне значення.

Читання включає: поступове ознайомлення з назвою таблиці, заголовками граф і рядків; встановлення, до якої території відносяться дані і на яку дату або на який період вони приведені; в яких одиницях вимірюються; вияснення взаємозв’язків процесів, що характеризуються середніми і відносними величинами. Слід пам’ятати, що аналіз є невід’ємною частиною статистичного дослідження і перед тим як аналізувати статистичну таблицю, потрібно піддати її логічному і арифметичному контролю.

Під аналізом розуміють метод наукового дослідження явища шляхом його фактичного, або умовного розчленування на складові частини. Аналіз взаємозв'язаний з синтезом – методом вивчення явища в цілому, в єдності і взаємозв'язку його частин.

Аналіз даних статистичної таблиці потрібно починати з підсумків, що дає загальну уяву про відомості таблиці. Потім потрібно перейти до вивчення окремих груп підмета, тобто окремих рядків таблиці, а також її граф. Вибирають спочатку часткові і найбільш характерні дані, а потім аналізують всі інші.

Зіставляючи окремі групи підмета з показниками присудка встановлюють зміну і залежність одних ознак від інших, а також характер цього зв'язку (прямий чи зворотній, сильний чи слабкий). За аналізом таблиці завжди іде синтез, заради чого і здійснюється аналіз.

Так, дані табл. 4.9 свідчать про зв'язок між віком сімей, середньодушевим доходом на одного члена сім'ї і забезпеченістю сімей набором найбільш важливих і необхідних товарів культурно-побутового призначення. Аналіз таблиці по рядках показує, що забезпеченість сімей цими товарами збільшується із ростом середньодушевого доходу на одного члена сім'ї. Аналіз граф свідчить, що із збільшенням віку сім'ї, її забезпеченість даними товарами також збільшується. Оцінити цей зв'язок кількісно, встановити його силу допоможе кореляційно-регресійний аналіз.

Аналіз даних в статистичних таблицях часто вимагає виконання додаткових розрахунків різних відносних, середніх величин, індексів та інших показників.

Наведені вище властивості статистичних таблиць, як форму компактного, виразного і рельєфного викладу даних, широко використовують вчені в наукових дослідженнях, а економісти в практичній діяльності.

### ***Контрольні запитання***

1. Поняття про статистичні таблиці і вимоги до їх побудови.
2. Призначення статистичних таблиць та переваги табличного методу.
3. Завдання табличного методу викладу статистичних даних.
4. Які таблиці не являються статистичними?
5. Що являється підметом статистичної таблиці?
6. Що називається присудком статистичної таблиці?
7. Що собою являють скелет і макет статистичної таблиці?
8. Що характеризує загальний заголовок статистичної таблиці?
9. Види статистичних таблиць.
10. Які таблиці називаються простими?
11. Поняття про прості перелікові таблиці.

12. Прості територіальні таблиці.
13. Прості хронологічні таблиці.
14. Які таблиці називаються груповими?
15. Використання комбінаційних таблиць.
16. Поняття про типові таблиці.
17. Використання балансових таблиць.
18. Розробка присудка таблиці.
19. Основні правила оформлення статистичних таблиць.
20. З чого починається оформлення таблиць?
21. Якою повинна бути таблиця за розміром і які відомості включати?
22. Якими повинні бути загальний заголовок таблиці, заголовки рядків підмета і граф присудка?
23. Яких умовних позначень потрібно дотримуватись при заповненні кліток таблиці?
24. Вимоги до округлень статистичних показників.
25. Аналіз статистичних таблиць.

# Розділ 5 Статистичні графіки

## 5.1. Поняття про статистичні графіки і правила їх побудови

В результаті опрацювання даних різного роду спостережень отримують велику кількість цифрового матеріалу, який розміщують у таблицях. Застосування табличного методу значно полегшує орієнтацію в зібраному і згрупованому матеріалі. Проте в багатьох випадках статистичних досліджень не можна обмежуватись одними таблицями.

Таблична форма викладу цифрового матеріалу не завжди дозволяє достатньо наглядно і чітко відображати загальну картину стану або розвитку якого-небудь явища, розкрити закономірності зв'язку статистичних показників між собою, або їх розподілу. А тому для розв'язку цих та інших завдань поряд із статистичними таблицями широко застосовується графічний спосіб зображення статистичних величин.

*Статистичний графік* – це особливий спосіб наочного зображення і узагальнення статистичних даних про соціально-економічні явища і процеси за допомогою геометричних образів, малюнків або схематичних географічних карт і пояснень до них. Графіки застосовуються, головним чином, для характеристики (порівняння) розвитку показників в часі і просторі, вивчення структури і структурних зрушень, контролю за виконанням планових завдань, характеристики просторового розміщення і просторового розповсюдження явищ. Графіки застосовуються також для аналізу зв'язків і залежності між різними показниками або між значеннями варіаційної ознаки і частотами чи частками.

При побудові статистичного графіка потрібно знати, з якою метою складається графік, вивчити вихідний матеріал і володіти методикою графічних зображень.

Основними елементами графіка є: поле графіка, графічні образи, масштабні орієнтири і експлікація графіка. Кожний елемент має своє призначення і виконує відповідну роль в побудові і інтерпретації графіка.

*Поле графіка* – це простір, на якому розміщуються геометричні та інші знаки, які створюють графік. Цей простір обмежується або аркушем чистого паперу, або географічною чи контурною картою.

Розмір поля залежить від призначення графіка. В статистичних дослідженнях найбільш часто зустрічаються графіки у вигляді прямокутників з нерівними сторонами по вертикалі і горизонталі, хоча також застосовуються графіки у виглядів квадратів. В практиці співвідношення нерівних сторін полів графіка береться від 1 : 1,33 до 1:1,50, якщо вертикальну сторону прийняти за 1.

Просторові орієнтири задаються у вигляді прямокутної системи координат, тобто координатної сітки. В картограмах засобами просторової орієнтації виступають географічні карти.

*Графічний образ* – це сукупність різноманітних геометричних знаків, за допомогою яких відображаються статистичні величини. В статистичних графіках використовуються такі геометричні знаки як, крапки, відрізки прямих ліній, квадрати, прямокутники, кола, півкола, сектори, а також негеометричні знаки – символи у вигляді силуетів або малюнків. Це і є основою графіка, його мовою.

*Масштабні орієнтири* статистичних графіків – це масштаб, масштабні шкали і масштабні знаки, які використовуються для визначення розмірів геометричних та інших графічних знаків.

*Масштаб* – це умовна міра переводу числової величини статистичного явища в графічну і навпаки. Тобто, це довжина відрізка шкали, прийнята за числову одиницю. Наприклад, 1 см на графіку відповідає 1000 одиницям виробленої продукції, або 1 см$^2$ дорівнює 100 км$^2$ на досліджуваній території.

При побудові графіка масштаб повинен бути таким, щоб ясно і чітко проявлялися відмінності зображення статистичних величин і разом з цим їх легко можна було б порівняти між собою. Найбільш розповсюдженою при побудові статистичних графіків є система прямокутних координат. При цьому найкраще співвідношення масштабу по осі абсцис і ординат 1,41 : 1, яке відоме під назвою "золотого перетину". На осі ординат графіка повинна бути нульова точка. У випадках, коли мінімальне значення ознаки значно вище нуля, доцільно робити розрив вертикальної шкали (див. мал. 5.12).

*Масштабна шкала* – це лінія, поділена на відрізки точками відповідно до прийнятого масштабу. Носієм шкали можуть виступати пряма або крива лінії. Залежно від цього масштабні шкали поділяють на прямолінійні і кругові.

Довжину відрізків між сусідніми поділками шкали називають графічним інтервалом, а різницю між числовими значеннями цих поділок – числовим інтервалом. Обидва інтервали можуть бути рівними і нерівними.

Шкалу, в якій рівним графічним інтервалом відповідають рівні числові інтервали називають рівномірною, або арифметичною.

Якщо рівним графічним інтервалам відповідають нерівні числові інтервали шкалу називають нерівномірною, або функціональною. Для побудови статистичних графіків з функціональною шкалою найчастіше застосовують логарифмічну функцію y = lg x.

*Масштабні знаки* – це еталони, які зображають на графіку статистичні величини у вигляді квадратів, кругів, силуетів тощо. Ними користуються для визначення розмірів і співвідношень статистичних величин, зображених на графіку, тобто для порівняння графічних знаків із знаком-еталоном.

*Експлікація графіка* – це словесні пояснення, які розкривають його зміст і основні елементи: заголовок графіка, одиниці виміру, умовні позначення.

Загальний заголовок повинен ясно, чітко і коротко розкрити основний його зміст і відповісти на три питання – що, де, коли?

На кожній масштабній шкалі графіка вказуються розміщені на них статистичні величини і одиниці їх вимірювання.

Пояснювальні надписи до окремих елементів графічного образу можуть знаходитись на полі графіка, або у формі умовних позначень виноситись за його межі.

*Класифікація статистичних графіків*. Класифікація графіків дає можливість визначати їх загальні риси, аналітичні можливості та техніку побудови. Графіки класифікуються за функціонально-цільовим призначенням, видами, формами і типами основних елементів.

За загальним призначенням графіки ділять на аналітичні, ілюстративні та інформаційні.

За функціонально-цільовим призначенням розрізняють графіки групувань і рядів розподілу, динаміки, взаємозв’язку і порівняння.

За формою графічних образів графіки поділяють на крапкові, лінійні, площинні, просторові і фігурні.

За типом системи координат розрізняють графіки у прямокутній і полярній системі координат, а за масштабними шкалами – графіки з рівномірними функціональними і змішаними шкалами.

Класифікація графіків за видом їх поля дає змогу виділити дві великі групи графіків: а) діаграми; б) статистичні карти.

З точки зору розв’язуваних завдань статистичні графіки поділяють на : 1) графіки порівняння статистичних величин; 2) графіки структури і структурних зрушень; 3) графіки зображення динаміки статистичних показників; 4) графіки контролю виконання плану; 5) графіки просторового розміщення і розповсюдження; 6) графіки варіаційних рядів (див. розділ 9); 7) графіки взаємозв’язку і взаємозалежності (див. розділ 10).

Графіки, які застосовуються для зображення статистичних даних надзвичайно різноманітні. В даній главі розглянемо графіки, які найбільш часто зустрічаються в статистичній практиці.

## 5.2. Графіки порівняння статистичних величин

В статистичній практиці для порівняння величин статистичного показника, які характеризують його зміну в просторі, застосовують діаграми.

*Діаграми* – це такий вид графіків, в якому цифрові дані зображаються з допомогою різних геометричних фігур і ліній. Діаграми є стовпчикові, стрічкові, секторні, лінійні та інші.

*Стовпчикові діаграми* являють собою найбільш простий, наочний і широко розповсюджений вид графіків в одному вимірі. В них статистичні дані зображаються у вигляді стовпчиків – прямокутників однакової ширини, розміщених вертикально по осі абсцис і однакової або різної висоти. Кожний окремий стовпчик характеризує окремий об'єкт. Загальне число стовпчиків дорівнює числу порівнюваних об'єктів. Віддаль між стовпчиками береться однакова, але інколи стовпчики розташовують упритул один до одного.

Покажемо побудову стовпчикової діаграми на прикладі за даними таблиці 5.1.

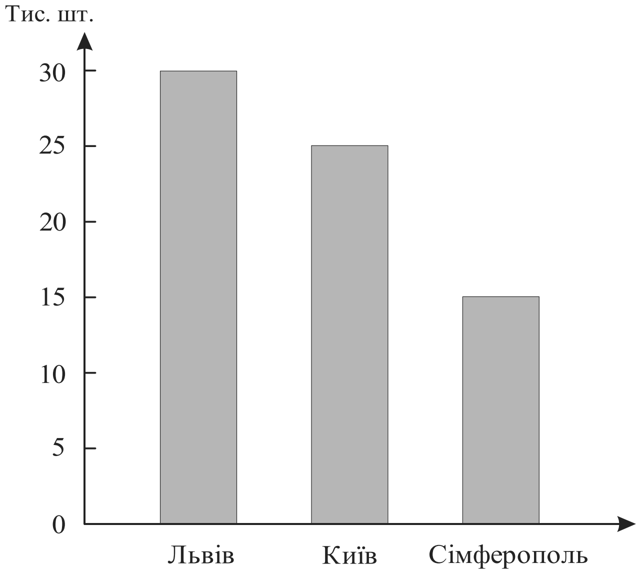
Таблиця 5.1

Виробництво телевізорів кольорового зображення заводами України в 2007 р. (дані умовні)

| Заводи | Вироблено телевізорів, тис. шт. | Середня ціна одного телевізора, грн. | Вартість вироблених телевізорів, тис. грн. |
|---|---|---|---|
| Львівський | 30 | 650 | 19500 |
| Київський | 25 | 450 | 11250 |
| Сімферопольський | 15 | 670 | 10050 |

Для побудови діаграми на осі абсцис на однаковій віддалі один від одного відкладемо три відрізки рівної довжини – основи стовпчиків. Заводи розмістимо на графіку ранжировано: в порядку зменшення чисельності виробництва телевізорів кольорового зображення. Масштаб на осі ординат – 5 тис. шт. телевізорів на 1 см.

Для наочності стовпчики заштриховують або замальовують. Наявність даної діаграми досягається шляхом порівняння висоти стовпчиків, котра відповідає чисельності телевізорів кольорового зображення. Внизу під стовпчиками вказують назви об’єктів порівняння, в даному випадку міст, де знаходяться телевізорні заводи.




Мал. 5.1 Виробництво телевізорів кольорового зображення заводами України в 2007 р.

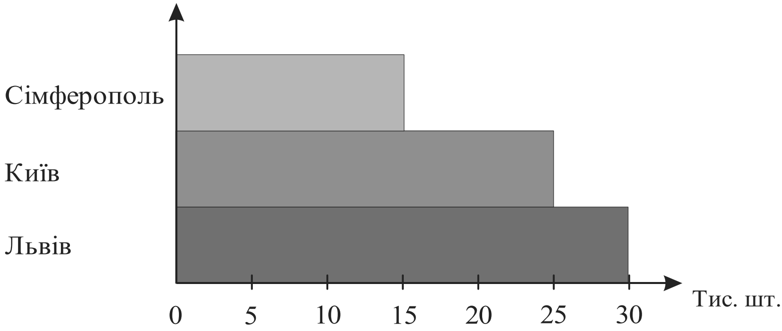
Якщо стовпчики-прямокутники, які зображають числа, розташувати не по вертикалі, а по горизонталі, тоді таку діаграму називають *стрічковою*.

Стовпчикові і стрічкові діаграми взаємозамінні, так як в обох випадках використовується один вимір – висота стовпчика або довжина стрічки.

Зображення діаграм у вигляді стрічок краще ніж у вигляді стовпчиків, так як при цьому вигідніше кожному прямокутнику дати відповідну горизонтальну назву.

Проілюструємо побудову стрічкової діаграми за даними табл. 5.1 на мал. 5.2.

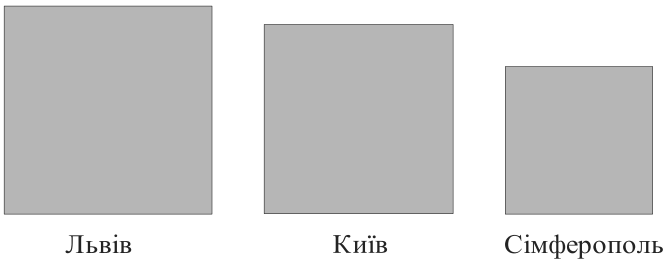
Стрічки на даній діаграмі розташовані (щільно) упритул одна до одної.




Мал. 5.2 Виробництво телевізорів кольорового зображення заводами України в 2007 р.

Для порівняння декількох абсолютних величин між собою використовують також *квадратні діаграми*. Для визначення сторони квадрату потрібно добути корінь квадратний із абсолютної величини явища, в даному випадку обсягу виробництва телевізорів кольорового зображення. Для Львівського заводу корінь квадратний із 30 дорівнює приблизно 5,5; Київського – корінь квадратний із 25 дорівнює 5 і Сімферопольського – корінь квадратний із 15 дорівнює 3,9. При виборі масштабу орієнтуються на найбільше число.

Прийнявши масштаб 2 = 1 см, визначаємо сторони квадратів: для Львова – 5,5 : 2 = 2,75 см; Києва – 5,0 : 2 = 2,5 см; Сімферополя – 3,9 : 2 = 1,95 см. Визначивши сторони квадратів будують діаграму (мал. 5.3).

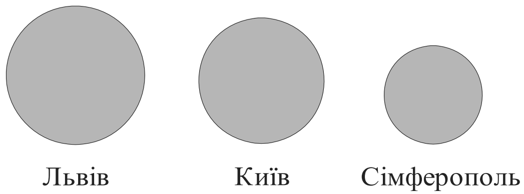



Мал. 5.3 Виробництво телевізорів кольорового зображення заводами України в 2007 р.

*Кругові діаграми* основані на використанні площ кругів для порівняння однорідних абсолютних величин між собою.

При побудові кругової діаграми, потрібно прийняти до уваги, що площі кругів відносяться між собою як квадрати їх радіусів. Отже, щоб знайти радіус, потрібно добути корінь квадратний із абсолютних величин і на цій основі визначити радіуси.

Довжина радіуса кола при масштабі 3 = 1 см буде становити: для Львова 1,83 см (5,5 : 3); Києва – 1,66 (5,0 : 3) і Сімферополя – 1,30 (3,9 : 3). Після того як визначені довжини радіусів описують кола і отримують наступну діаграму (мал. 5.4).




Мал. 5.4 Виробництво телевізорів кольорового зображення заводами України в 2007 р.

*Прямокутні діаграми* застосовують в тих випадках, коли потрібно порівняти величини, які являють собою добуток двох співмножників і показати роль кожного з них у формуванні цієї величини. Ці діаграми вперше запропонував російський статистик В.Є.Варзар (1851 – 1940 рр.), а тому прямокутні діаграми називають ще знаки Варзара.

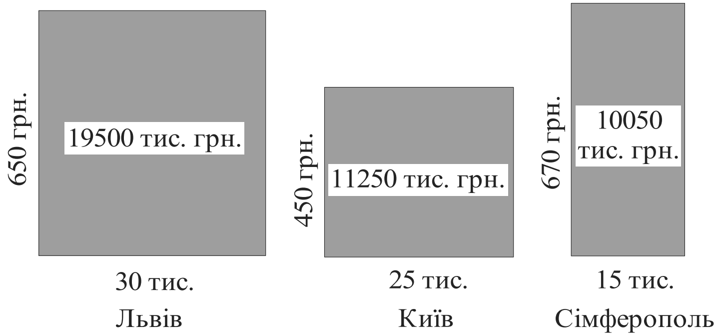
При побудові прямокутних діаграм встановлюють два масштаби: один для множника, який приймають за основу, а другий для множника, який приймають за висоту. В нашому прикладі (див. табл. 5.1) основа прямокутника буде характеризувати кількість телевізорів, висота – середню ціну одного телевізора, площа прямокутника – вартість всіх виготовлених телевізорів. Приймемо масштаб для основи прямокутника (10 тис. шт. = 1 см) і висоти (200 грн. = 1 см), тоді діаграма матиме вигляд (мал. 5.5).

Для більшої наочності зображення статистичних явищ можна замінити абстрактні геометричні фігури малюнками. Такого виду діаграми називаються *картинними або фігурними*.

Картинні діаграми будують двома способами:

- перший, коли малюють фігури розмір яких пропорційний величині зображуваного явища;

другий, коли встановлюють певний масштаб для фігур. В нашому прикладі 1 картинка може відповідати 5 тис. шт. телевізорів. Ці картинки розташовують у вигляді стрічки. Так, для Львівського заводу стрічка буде складатись з 6 картинок, Київського – 5 картинок, Сімферопольського – 3 картинки.




Мал. 5.5 Порівняння трьох заводів за кількістю виготовлених телевізорів, середньою ціною і вартістю.

Фігурні діаграми фіксують на собі увагу, достатньо зрозумілі і дохідливі, а тому вони часто використовуються як агітаційний інструмент.

### 5.3. Наочне зображення структури і структурних зрушень

Для статистичного дослідження складу сукупності використовують структурні діаграми. *Структурні діаграми* – це діаграми питомих ваг, які характеризують відношення окремих частин сукупності до її загального об'єму. За видами вони діляться на стовпчикові, стрічкові і секторні.

Стовпчикові і стрічкові діаграми застосовують не тільки для порівняння самих величин між собою, але й для одночасного порівняння частин цих величин. Звернемось до прикладу.

Таблиця 5.2

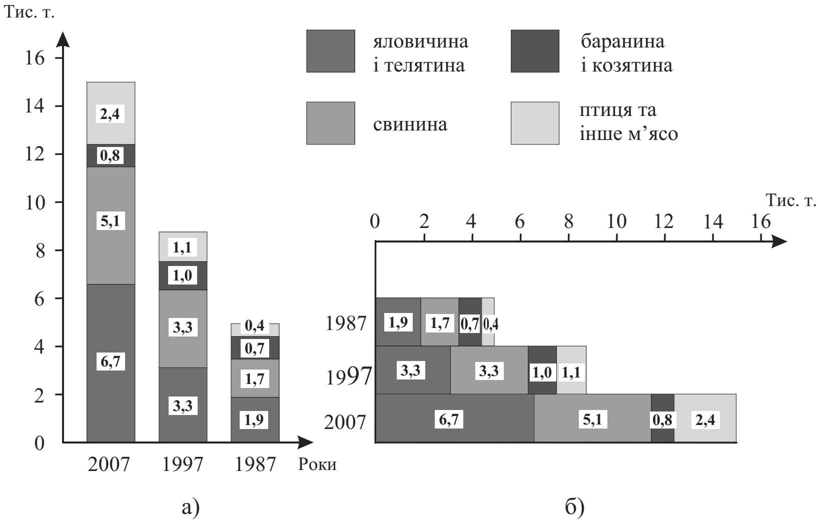
Виробництво м'яса всіма категоріями господарств регіону (дані умовні)

| Роки | М'ясо (в загальній вазі), тис. т. | в тому числі | | | |
|---|---|---|---|---|---|
| | | яловичина і телятина | свинина | баранина і козятина | м'ясо птиці та інше |
| 1987 | 4,7 | 1,9 | 1,7 | 0,7 | 0,4 |
| 1997 | 8,7 | 3,3 | 3,3 | 1,0 | 1,1 |
| 2007 | 15,0 | 6,7 | 5,1 | 0,8 | 2,4 |

З метою характеристики і ілюстрації об'єму і структури виробництва м'яса в регіоні побудуємо стовпчикову діаграму (мал. 5.6, а). Виберемо і відкладемо масштаб по осі ординат, в нашому прикладі

1 см. відповідає 2 тис. т. м'яса. По осі абсцис, на однаковій віддалі будуємо стовпчики, розбивши їх на частини, величини яких відповідають об'єму виробництва різних категорій м'яса.

Аналогічно будуємо і стрічкову діаграму, тільки в даному випадку масштабна шкала відкладається на осі абсцис, а перпендикулярно до осі ординат малюють смужки (стрічки), які відображають статистичне явище (мал. 5.6, б). Для кожної частини стовпчика встановлюємо відповідне штрихування.



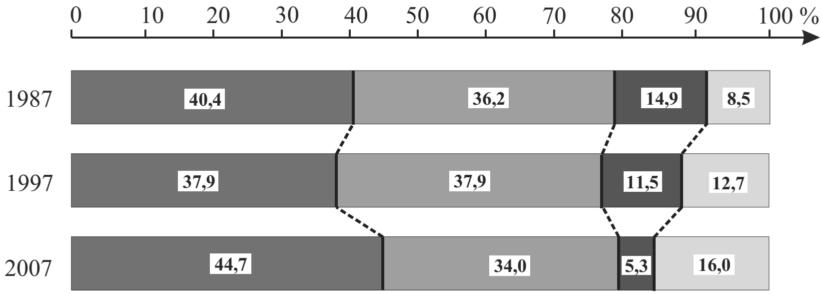

Мал. 5.6 Об'єм і структура виробництва м'яса в регіоні в 1987, 1997 і 2007 рр.

Для більш наочного зображення структури і структурних зрушень на графіку відкладають не самі абсолютні величини, а їх питомі ваги в загальному підсумку. Стовпчики або стрічки в цьому випадку мають однаковий розмір, який відповідає 100 %. В такій діаграмі стовпчики або стрічки розбивають на частини відповідно питомим вагам, котрі інколи для кращого порівняння структурних зрушень з'єднують пунктирними лініями.

Охарактеризуємо структуру виробництва м'яса господарствами регіону в процентах (табл. 5.3) і побудуємо графік (мал. 5.7).

Таблиця 5.3

| Роки | М'ясо (в загальній вазі), % | в тому числі | | | |
|---|---|---|---|---|---|
| | | яловичина і телятина | свинина | баранина і козлятина | м'ясо птиці та інше |
| 1987 | 100 | 40,4 | 36,2 | 14,9 | 8,5 |
| 1997 | 100 | 37,9 | 37,9 | 11,5 | 12,7 |
| 2007 | 100 | 44,7 | 34,0 | 5,3 | 16,0 |




Мал. 5.7 Структура виробництва м'яса в регіоні в 1987, 1997 і 2007 рр.

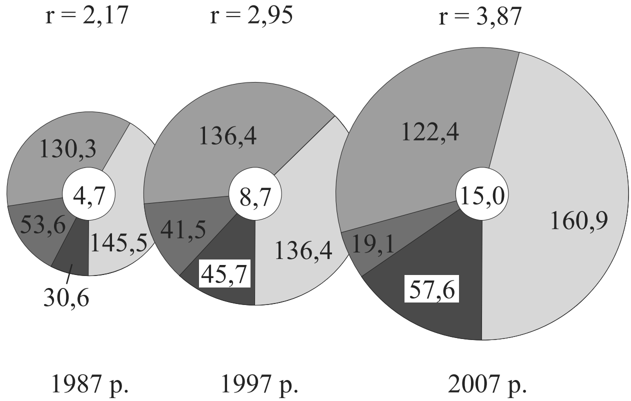
*Секторні діаграми* являють собою графічні зображення на площі круга, розділеного радіусами на окремі сектори за кількістю різновидів номінальних ознак. Ці діаграми застосовуються для наочної ілюстрації структури явища, для характеристики питомих ваг окремих частин цілого, для виявлення структурних зрушень.

На секторних діаграмах можуть зображуватись частини абсолютних величин явищ, або їх процентний вираз.

Для побудови секторної діаграми, яка характеризує абсолютні величини, спочатку потрібно знайти радіуси кругів, добувши квадратні корені з цих абсолютних величин. Наприклад, для 1987 р. $r = \sqrt{4{,}7} = 2{,}17$ см, 1997 р. – $r = \sqrt{8{,}7} = 2{,}95$ см, 2007 р. – $r = \sqrt{15{,}0} = 3{,}87$ см. При необхідності використовують масштаб.

Щоб розбити круг на сектори, які відповідають величинам частин цілого, потрібно 360° розділити на об'єм цілого (цим самим ми знайдемо скільки градусів припадає на одиницю явища) і отриманий результат перемножити на величину частин. Наприклад, для 2007 р. 360° : 15 = 24°, тобто 1 тонна м'яса дорівнює 24°, в такому випадку сектор яловичини і телятини становитиме 160,9° (6,7 x 24); свинини – 122,4° (5,1 x 24); баранини і козятини – 19,1° (0,8 x 24); м'яса птиці та іншого – 57,6° (2,4 x 24). Аналогічно визначають сектори виробництва різних категорій м'яса і за інші роки.

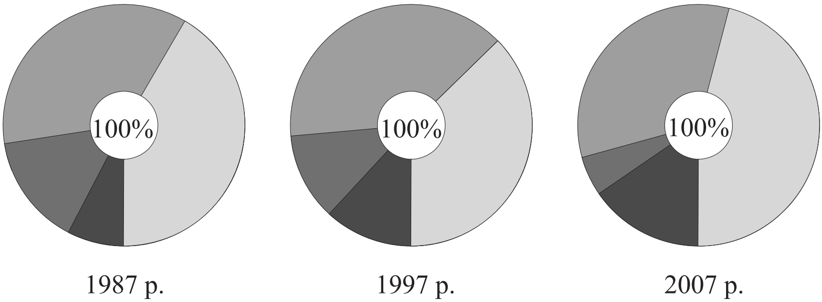
Кути секторів відкладаємо за допомогою транспортира, прийнявши який-небудь радіус за початок відліку (мал. 5.8).




Мал. 5.8 Об'єм і структура виробництва м'яса в регіоні в 1987, 1997 і 2007 рр.

Якщо секторна діаграма враховує лише питомі ваги частин явища, абстрагуючись від розмірів явища, креслять круги однакових діаметрів. Вся величина явища приймається за 100 %, розраховуються долі окремих його частин в відсотках (табл. 5.3). Круг розбивається на сектори пропорційно частинам зображуваного цілого. Таким чином на 1 % припадає 3,6°. Для отримання кутів секторів, які зображають долі частин цілого, потрібно їх процентний вираз перемножити на 3,6°. Наприклад, для 2007 р.: кут сектора виробництва яловичини і телятини становитиме 160,9° (44,7 x 3,6); свинини – 122,4° (34 x 3,6); баранини і козятини – 19,1° (3,3 x 3,6); м'яса птиці та іншого – 57,6° (16 x 3,6). Аналогічно визначаємо кути секторів виробництва різних категорій м'яса в 1987 і 1997 рр. Покажемо цю діаграму на мал. 5.9.

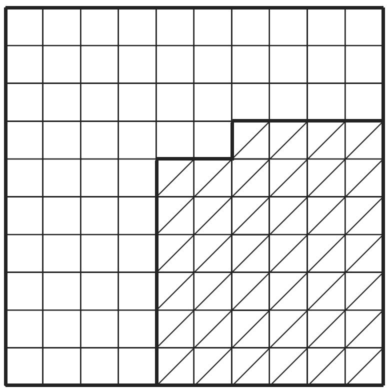
Секторні діаграми виразні в тих випадках, коли досліджувана сукупність ділиться не більше ніж на 4 – 5 частин і спостерігаються помітні структурні зрушення. Якщо ці структурні зрушення незначні, або сукупність ділиться на більше число секторів, тоді для графічного зображення структури доцільно використовувати стовпчикові або стрічкові діаграми.




Мал. 5.9 Структура виробництва м'яса в регіоні в 1987, 1997 і 2007 рр.

В деяких випадках для характеристики структури сукупності використовувати також квадратні і кругові діаграми.

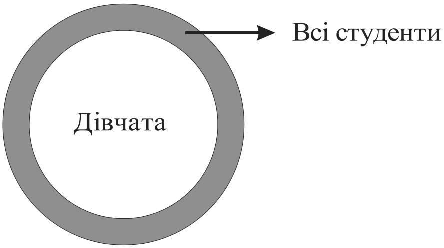
Для зображення структури сукупності, яка складається (в більшості випадків) з двох частин, беруть квадрат. Площу квадрата ділять на 100 рівних частин. Кожний маленький квадратик дорівнює одній сотій всієї площі великого квадрата. Потім ці квадратики заштриховують у відповідності із процентною структурою досліджуваної сукупності. Звернемось до прикладу. Нехай статева структура студентів економічного вузу в 2007 р. відповідно становила: дівчата – 60 %, хлопці – 40 %. Покажемо побудову такої діаграми на графіку (мал. 5.10).



Мал. 5.10 Питома вага дівчат і хлопців в економічному вузі в 2007 р.

В тому випадку, коли частину і ціле зображують при допомозі кругової діаграми, тоді круги креслять не окремо один від одного, а накладають один на другий.

Якщо для нашого прикладу візьмемо радіус R = 2 см, то площа круга становитиме $S = \pi r^2 = 3{,}1416 \cdot 4 = 12{,}5664$ см$^2$, яка відповідає 100 % всіх студентів. Для 60 % дівчат площа круга – $S_д = \pi r^2 = 3{,}1416 \cdot 4 \cdot 0{,}6 = 7{,}5398$ см$^2$. Радіус такої площі дорівнює R = 1,55 см. $\left(R = \sqrt{S : \pi} = \sqrt{7{,}5398 : 3{,}1416}\right)$. Після розрахунків будуємо кругову діаграму статевої структури студентів економічного вузу (мал. 5.11).




Мал. 5.11 Статева структура студентів економічного вузу в 2007 р.

Інколи доцільно показати три круги, один в другому, але навколо різних центрів.

Зустрічається також комбінування кругових діаграм із секторними, коли круги різної величини подають з розбивкою на сектори.

### 5.4. Графічне зображення динаміки статистичних показників

Графіки, які ілюструють зміну статистичних явищ в часі називаються *динамічними*. Для зображення динаміки явищ часто використовують розглянуті нами в п. 5.2 стовпчикові, стрічкові, квадратні, кругові і картинні діаграми, в яких кожний стовпчик, стрічка, квадрат і т.д. зображають величину статистичного явища на певну дату, або за відповідний проміжок часу.

Крім названих вище графіків нерідко застосовуються і лінійні графіки.

*Лінійні графіки* використовуються для характеристики зміни явищ в часі, виявлення залежності між двома показниками і деяких інших завдань. Вони будуються при допомозі прямокутної системи координат, на осі абсцис якої розміщують шкалу характеристик часу, а на осі ординат – рівні ряду динаміки.

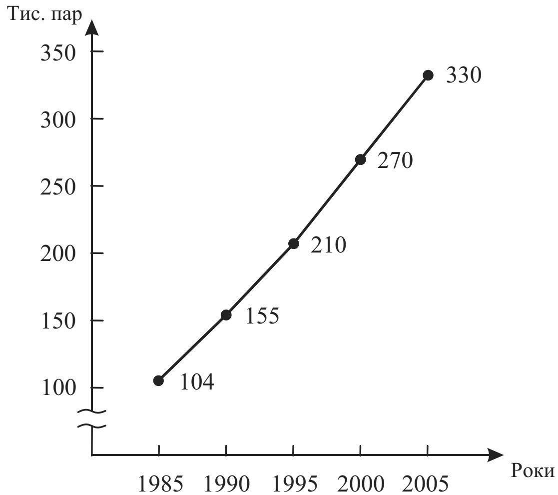
У лінійній діаграмі динаміки шкала на осі ординат повинна починатися з нуля, інакше діаграма буде не правильно відображати характер розвитку явища.

Оскільки при великих значеннях рівнів динамічного ряду діаграма з початковим нульовим рівнем ординати буде невиразною і некомпактною, тоді на осі ординат слід зробити розрив шкали (мал. 5.12). Для базисних характеристик швидкості зміни досліджуваного явища початковий рівень ординати може починатись із 100.

Лінійні діаграми дають можливість наочно визначити періоди часу, коли явища зростали (зменшувались) більш чи менш інтенсивно, або залишались без змін.

Особливістю лінійного графіка наочного зображення даних, які характеризують підсумки розвитку явища за певний період часу є те, що динаміка показується у вигляді неперервної лінії, котра характеризує неперервність процесу.

Покажемо побудову лінійного графіку на основі наступних даних: виробництво шкіряного взуття швейним об’єднанням і 1985 р. становило 104 тис. пар, в 1990 р. – 155 тис. пар, в 1995 р. – 210 тис. пар, в 2000 р. – 270 тис. пар і в 2005 р. – 330 тис. пар. Зобразимо ці дані графічно (мал. 5.12).




Мал. 5.12 Динаміка виробництва шкіряного взуття швейним об’єднанням в 1985 – 2005 рр.

Часто на одному лінійному графіку приводиться декілька кривих, котрі дають порівняльну характеристику динаміки різних показників або одного і того ж показника але різних об’єктів. В таких випадках спочатку потрібно показники рядів динаміки, які будемо наносити на графік, привести до однієї основи, тобто абсолютні показники кожного ряду замінити базисними темпами росту, прийнявши для всіх рядів один і той же період в якості бази порівняння. В цих графіках лінії всіх рядів розходяться із однієї точки, прийнятої за 1 або 100 %. Розглянемо побудову такого графіку (мал. 5.13) за даними таблиці 5.4.

Таблиця 5.4

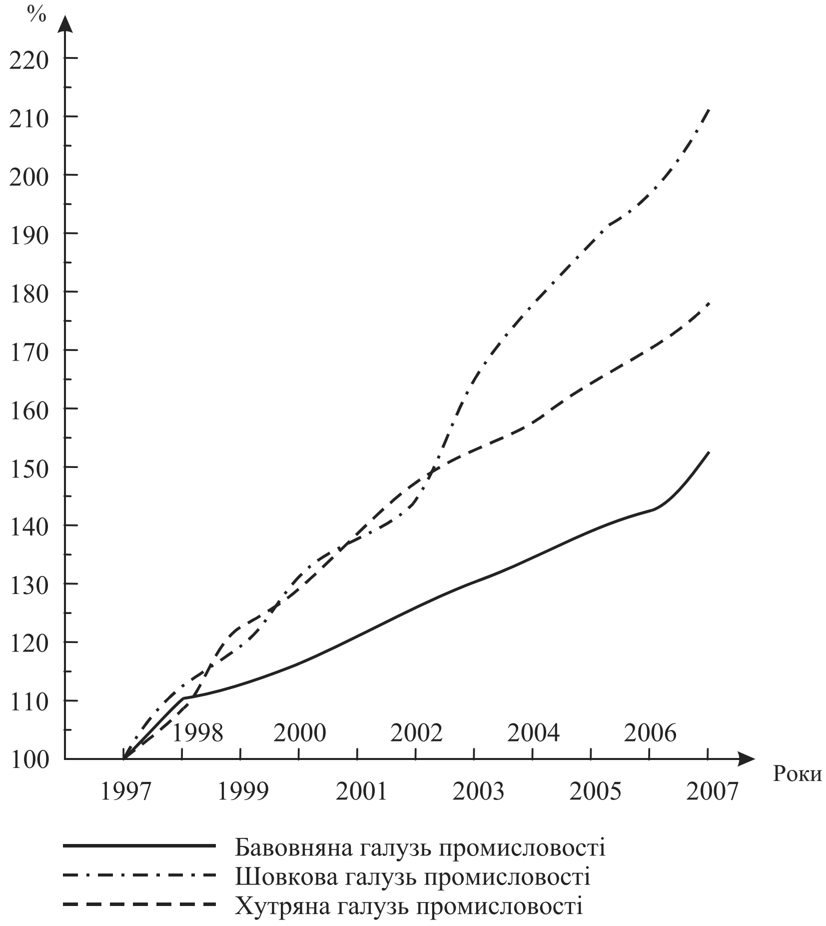
Темпи росту загального об’єму продукції деяких галузей легкої промисловості за період 1997 – 2007 рр.
(у відсотках до 1997 р., дані умовні)

| Роки | Галузі | | |
|---|---|---|---|
| | бавовняна | шовкова | хутряна |
| 1997 | 100 | 100 | 100 |
| 1998 | 114 | 115 | 110 |
| 1999 | 116 | 120 | 125 |
| 2000 | 118 | 130 | 129 |
| 2001 | 122 | 135 | 136 |
| 2002 | 124 | 142 | 146 |
| 2003 | 128 | 155 | 153 |
| 2004 | 133 | 178 | 159 |
| 2005 | 138 | 194 | 163 |
| 2006 | 140 | 201 | 167 |
| 2007 | 151 | 210 | 175 |

Лінійні графіки використовують в статистиці не тільки для ілюстрації динаміки якого-небудь явища, але і для наочного зображення рядів розподілу. В цьому випадку на осі абсцис відкладаються варіанти, а на осі ординат – частоти ряду розподілу.

Лінійними графіками користуються також для наглядного зображення залежності однієї варіаційної ознаки від іншої. Такі графіки наведені в розділі 10.

В статистичній практиці побудови графіків для аналізу темпів динаміки явища використовують лінійні графіки на напівлогарифмічній сітці. *Напівлогарифмічною* називається сітка, в котрій на осі абсцис нанесений звичайний масштаб, а на осі ординат – логарифмічний.




Мал.. 5.13 Динаміка виробництва продукції бавовняною, шовковою і хутряною галузями легкої промисловості за 1997 – 2007 рр.

Перевага напівлогарифмічної сітки в аналізі динаміки явища заключається в тому, що вона дає більш коректну уяву про темпи динаміки. Діаграму на напівлогарифмічний сітці ще називають *діаграмою темпів*.

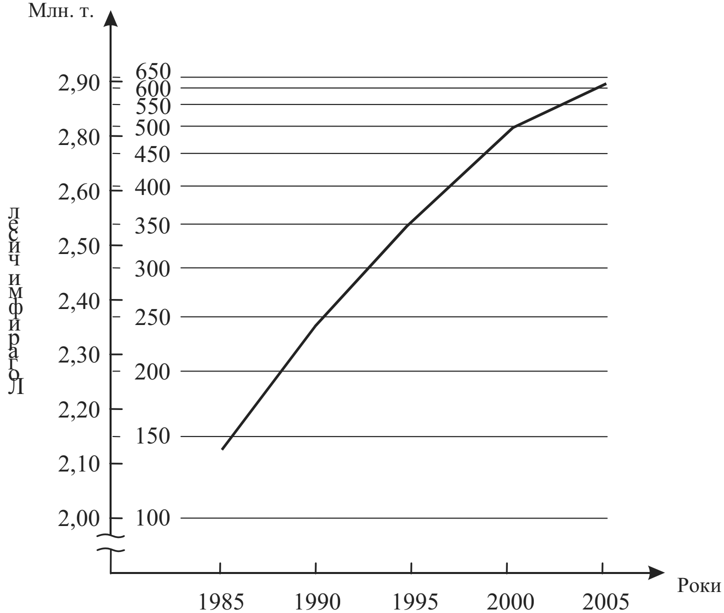
Для побудови лінійного графіка з напівлогарифмічною шкалою по осі ординат замість звичайної шкали відкладають логарифмічну з рівними інтервалами. Далі за таблицею логарифмів, знаходять логарифми для цілих чисел, які проставляють з правої сторони осі ординат для кращої наочності. За масштабом логарифмічної шкали знаходять відповідні точки, які проставляють на графіку і з'єднують їх лініями.

Побудову лінійного графіка на напівлогарифмічній сітці краще зрозуміти на прикладі.

Таблиця 5.5

Видобуток вугілля однією із шахт Львівсько-Волинського басейну за 1985 – 2005 рр. характеризуються такими даними (дані умовні)

| Роки | Видобуток вугілля, тис. т. | Логарифми числа добування вугілля |
|---|---|---|
| 1985 | 148 | 2,1703 |
| 1990 | 243 | 2,3856 |
| 1995 | 353 | 2,5478 |
| 2000 | 491 | 2,6911 |
| 2005 | 603 | 2,7803 |




Мал. 5.14 Динаміка видобутку вугілля за 1985 – 2005 рр.

Різновидністю лінійної діаграми є *радіальні діаграми*, побудовані в полярних координатах і призначені для відображення процесів і явищ, які періодично повторюються в часі (переважно сезонних коливань).

За вісь ординат в полярних координатах приймаються радіуси, а за вісь абсцис – коло. Пунктом відліку служить центр кола, або його окружність.

Радіальні діаграми бувають двох видів – замкнуті і спіральні.

*Замкнуті діаграми* відображують весь внутрішньо річний цикл зміни явища за один рік.

Для того щоб побудувати радіальну діаграму замкнутого виду, у якої пунктом відліку служитиме центр кола, креслять коло радіусом, рівним середньомісячному показнику. Усе коло ділять на стільки частин, скільки внутрішньорічних періодів і відповідно їм проводять радіуси. Періоди часу розміщують за годинниковою стрілкою, причому розміщення місяців (якщо коло розбите на 12) аналогічне циферблату годинника. На кожному радіусі відповідно до прийнятого масштабу відкладають від центра кола відрізки пропорційно рівням показників конкретного місяця. Дані, які перевищили середньомісячний рівень, відкладаються за межами кола на продовженні радіуса. Потім кінці відрізків на радіусах з’єднуються лініями, причому точка грудня з’єднується із точкою січня одного і того ж року. Побудову радіальної замкнутої діаграми розглянемо на прикладі.

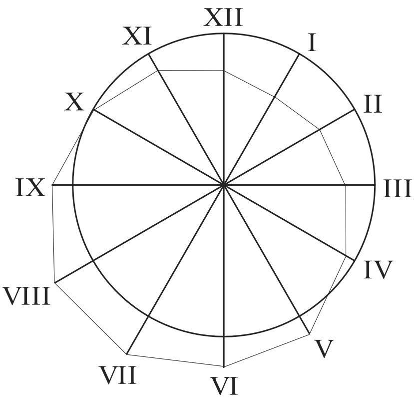
Маємо умовні дані про реалізацію молока на ринку одного міста по місяцях 2007 року:

| Місяць | I | II | III | IV | V | VI | VI | VIII | IX | X | XI | XII | Разом |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Продано молока, т. | 35; | 40; | 43; | 48; | 55; | 58; | 65; | 63; | 60; | 50; | 45; | 38 | 600 |

Середньомісячна реалізація молока за рік складає 50 т (600 : 12).

Креслимо коло з радіусом, який дорівнює середньомісячному показнику (R = 50 т). На горизонтальному діаметрі побудуємо шкалу, взявши довжину радіуса 2,5 см. Відповідно, 1 см – 20 т (50 : 2,5). Все коло ділимо на 12 радіусів, за кількістю місяців в році і відкладаємо наведені в умові прикладу за масштабом дані. Мітки на радіусах різних місяців з’єднуємо між собою (мал. 5.15).

*Спіральна радіальна діаграма* будується в тому випадку, коли є дані по місяцях за ряд років. Принцип її побудови той же, що і замкнутих, однак різниця лише в тому, що в спіральних діаграмах грудень одного року, а з січнем наступного року, в результаті чого виходить крива у вигляді спіралі.




Мал. 5.15 Сезонність продажу молока на ринку в 2007 р. по місяцях.

## 5.5. Контрольно-планові графіки

Графічний метод широко використовується для поточного контролю за ходом виконання плану. Форми графічного зображення для порівняння планових і фактичних показників досить різноманітні. В даному розділі розглянемо два основних види графіків:

а) лінійні графіки виконання плану;

б) обліково-планові графіки.

*Лінійні графіки* є зручними засобом контролю виконання плану по одному якому-небудь об’єкту або показнику. При цьому для аналізу на графіку доцільно показати наростаючим підсумком не тільки планові і фактичні показники у звітному періоді, але й фактичні – за минулий рік. Покажемо це на прикладі (табл. 5.6).

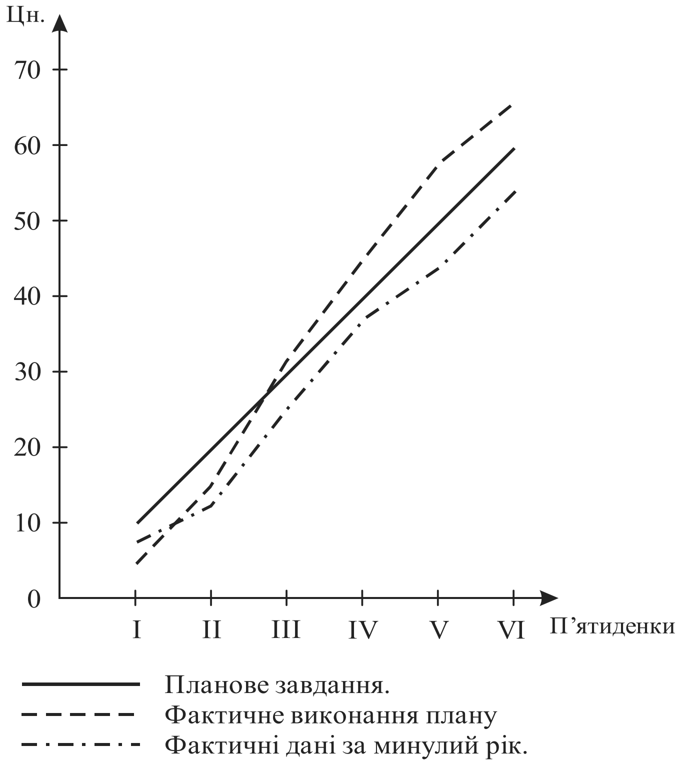
Таблиця 5.6

Виконання плану заготівлі лікарських рослин райспоживспілкою в серпні 2007 р.

| П’ятиденки | План | Фактично заготовлено лікарських рослин (тонн – наростаючим підсумком) | |
|---|---|---|---|
| | | в поточному році | в минулому році |
| **1** | **2** | **3** | **4** |
| I | 10 | 5 | 8 |
| II | 20 | 15 | 13 |
| III | 30 | 30 | 25 |

| 1 | 2 | 3 | 4 |
|---|---|---|---|
| IV | 40 | 43 | 38 |
| V | 50 | 58 | 45 |
| VI | 60 | 65 | 55 |

Побудуємо графік виконання плану (мал. 5.16).




Мал. 5.16 Виконання плану заготівлі лікарських рослин райспоживспілкою в серпні 2007 р.

На графіку наочно видно хід виконання плану заготівлі лікарських рослин по п'ятиденках: в першій і другій п'ятиденках фактичні заготівлі відстають від плану, третя п'ятиденка співпала з планом і в подальшому фактична лінія йде вище планової. Видно також, що об'єм заготівель поточного року (крім першої п'ятиденки) весь час вищий минулого.

В тих випадках, коли потрібно організувати наочний контроль виконання плану одночасно на декількох об'єктах, будують *обліково-планові графіки*.

Їх будують на спеціально розграфленій сітці, яка має форму таблиці, і на якій по горизонталі відкладають одиниці часу (день, п'ятиденку, декаду, місяць, квартал), а по вертикалі розміщують об'єкти дослідження.

Кожний відрізок по горизонталі відповідає 100 % виконання планового завдання, який, в свою чергу ділиться на п'ять рівних частин (кліток) по 20 % на кожну.

Ступінь виконання плану по кожному об'єкту зображається двома лініями: тонкою переривчастою, яка показує ступінь виконання плану за одиницю часу і жирною суцільною, яка характеризує виконання плану за звітний період в цілому.

Розглянемо порядок побудови обліково-планового графіка за такими даними (табл. 5.7).

Таблиця 5.7

Виконання плану по цехах заводу

| Число місяця | Цех 1 | | | | Цех 2 | | | |
|---|---|---|---|---|---|---|---|---|
| | план на день, шт. | вироб-лено за день, шт. | вико-нання плану, % | фактично вироблено з початку місяця, шт. | план на день, шт. | вироб-лено за день, шт. | вико-нання плану, % | Фактично вироблено з початку місяця, шт. |
| 1 | 500 | 450 | 90 | 450 | 500 | 500 | 100 | 500 |
| 2 | 500 | 500 | 100 | 950 | 510 | 561 | 110 | 1061 |
| 3 | 500 | 525 | 105 | 1470 | 520 | 546 | 105 | 1607 |
| 4 | 500 | 530 | 106 | 2000 | 530 | 535 | 101 | 2142 |
| 5 | 500 | 500 | 100 | 2500 | 540 | 590 | 109 | 2732 |
| і т.д. | | | | | | | | |

Цех 1 щоденно за планом повинен був виготовляти однакову кількість деталей – 200 шт., а цех 2 не однакову їх кількість: за перший день – 500 шт., другий – 510 шт., третій – 520 шт., і т.д.

Побудова графіка виконання плану за цими варіантами дещо відрізняється.

Нанесення тонких ліній на графік нескладне. Розглянемо спосіб їх нанесення в таких випадках: 1) планове завдання не виконано; 2) планове завдання виконано на 100 %; 3) планове завдання перевиконано.

Цех 1 планове завдання за перший день виконав на 90 % і тонка лінія займає 4 клітки (80 %) і половину п'ятої (10 %). За другий день планове завдання цех виконав на 100 %, а тому тонка лінія займає всі 5 кліток. Планове завдання третього числа виконано на 105 %, а тому тонка лінія займає всі п’ять кліток третього дня (100 %) і під лінією в першій клітці проведена маленька лінія, яка доходить до чверті (5 %). Ця додаткова лінія показує перевиконання денного планового завдання.

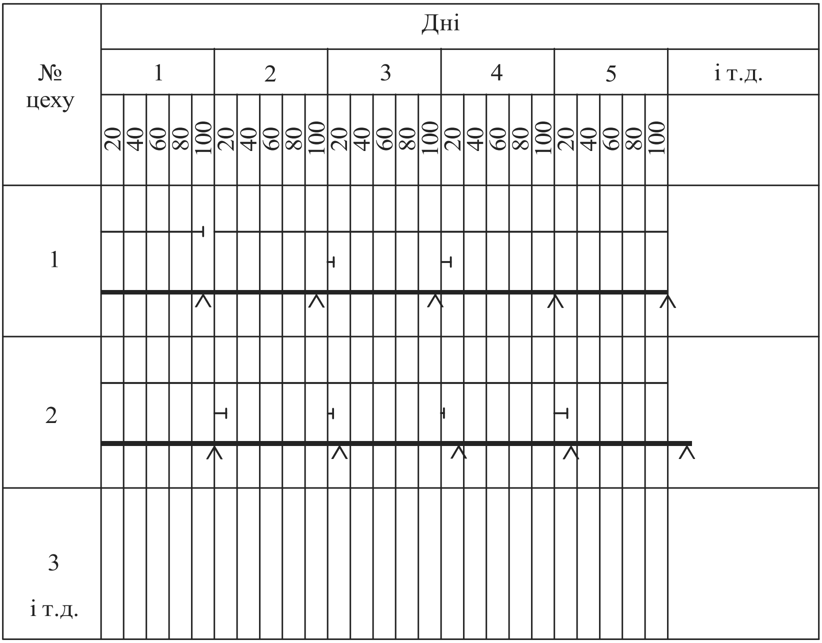
Техніка нанесення суцільної жирної лінії проста в тих випадках, коли планове завдання за кожний день не змінюється на протязі тривалого періоду, на котрий складається графік. Цех 1 щоденно планував виготовити по 500 деталей, і в перший день він виконав план на 90 %, другий – 100 %, третій – 105 %, четвертий – 106 % і в п’ятий – 100 %.

За перший день жирна лінія буде такої довжини, що й тонка. В другий день план був виконаний на 100 %, за два дні на 190 % (90 + 100), тобто суцільна жирна лінія доводиться за перший день до 100 %, а за другий день – до поділки котра відповідає 90 %. Це показує, що планове завдання за перші два дні не виконано.

За третій день планове завдання виконано на 105 %, а за три дні – на 295 % (90 + 100 + 105). Жирна лінія займе по 5 кліток першого і другого дня і 95 % третього. Всі подальші розрахунки проводять аналогічно до попередніх.

По цеху 2 принцип проведення тонкої лінії такий же, як і по цеху 1. складніше нанести на графік жирну лінію, тому що планові завдання за кожний день змінювались і не можна додавати проценти щоденного виконання завдання. В даному випадку ступінь виконання планового завдання за окремі періоди часу знаходять відношенням наростаючих підсумків абсолютних даних фактичного виконання до плану.

Суцільна жирна лінія повинна показати виконання планового завдання за всі попередні дні з початку звітного періоду. За перший день жирна лінія співпадає з тонкою. За другий день виготовлено продукції 561 шт., із яких 510 зараховують у виконання плану другого дня і 51 шт., або 9,8 % (51 : 520), – в рахунок виконання плану третього дня. За третій день виготовлено 546 деталей, до яких добавляються ще 51 шт. деталей виготовлених за другий день в рахунок третього, разом – 597 деталей, що складає 114,8 % (597 : 520) денного планового завдання третього дня. Жирна лінія займе п’ять кліток третього дня і ще майже ¾ першої клітки відведеної для четвертого дня (відповідає приблизно 15 %). Це означає, що за три дня другий цех повністю виконав триденне планове завдання і продовжував працювати в рахунок четвертого дня. Таким чином, жирна лінія графіка наочно демонструє хід виконання плану за кожний день з початку звітного періоду (мал. 5.17).




Мал. 5.17 Виконання плану виробництва деталей цехами заводу за першу п’ятиденку місяця.

Ступінь виконання плану за кожний день на жирній лінії позначається крапкою або штрихом.

Зображення на одному графіку декількох об’єктів дає можливість порівнювати підсумки їх роботи в цілому і оцінювати рівномірність виконання плану.

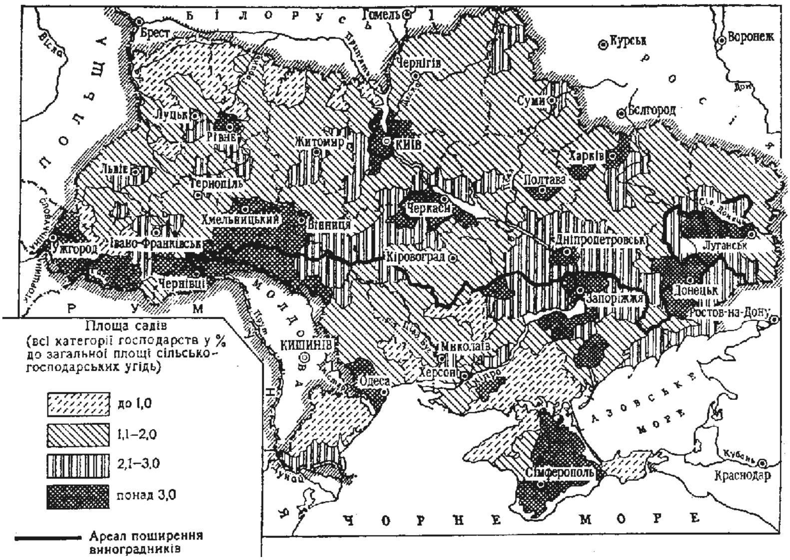
## 5.6 Графіки просторового розміщення і просторового розповсюдження

Для вивчення розміщення, рівня і ступеня розповсюдження якого-небудь явища в просторі використовується три види графіків: а) картограма; б) картодіаграма; в) центрограма.

*Картограма* – це схематична географічна карта, на якій розподіл зображуваних явищ по територіях дається за допомогою розмальовування, штрихування або крапок.

В залежності від використовуваних символів розрізняють фонові і крапкові картограми. Для побудови *фонових картограм* використовується вся поверхня карти в кордонах досліджуваної території. На цій карті повинні бути чітко позначені контури меж адміністративного поділу країни, області, району.

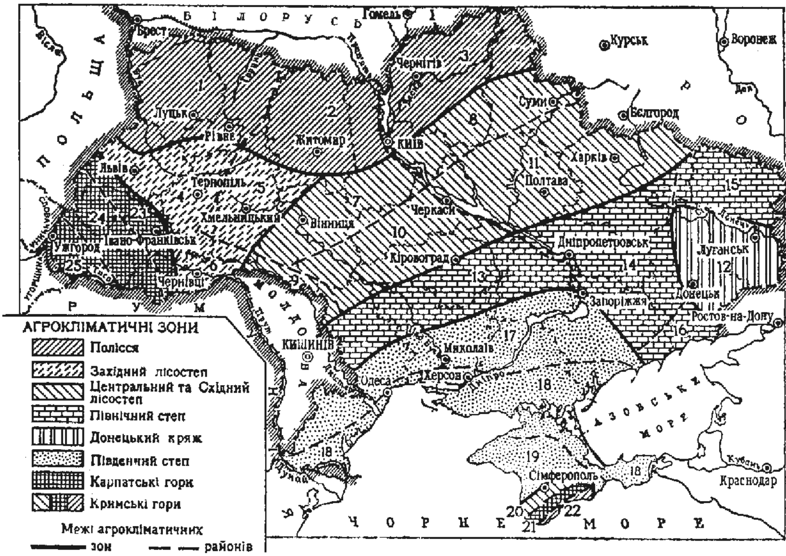
Географічний ряд, призначений для картограмування, потрібно оптимально розбити на групи (райони), що дозволить простежити закономірності розміщення зображуваного явища. Кожній групі (району) надається певний тип штрихування або колір, а потім їх наносять на карту. Так, наприклад, якщо ми хочемо дати картограму розміщення садів і виноградників в Україні, то ми повинні всі дані про це по окремих областях розбити, припустимо, на чотири групи і відповідним штрихуванням (мал. 5.18). В першу групу увійдуть всі категорії господарств з площею садів і виноградників, питома вага яких в загальній площі сільськогосподарських угідь до 1 %; другу – 1,1 – 2,0 %; третю – 2,1 – 3,0 % і четверту – понад 3,0 %. Після, на кожну область у відповідності з тією групою, до котрої вона попала, наноситься вказана штриховка. Інколи в якості умовного знаку замість штрихування користуються кольором, тільки при цьому вибирають кольори таким чином, щоб була витримана зростаюча інтенсивність по мірі переходу від нижчих груп до вищих.




Мал. 5.18 Розміщення садів і виноградників.*

* Заставний Ф.Д. Географія України: У 2-х книгах. – Львів: Світ, 1994. – С.296.

При зображенні деяких явищ, які вивчаються статистикою, розподіл за адміністративними районами не має великого значення, а тому в подібних випадках виділяються райони з однаковим показником досліджуваного явища за допомогою *ізолінійних картограм*. Такі картограми використовуються в метеорології і геодезії. В якості прикладу ізолінійної картограми візьмемо агро кліматичне районування за В.П.Поповим (мал. 5.19).




Мал. 5.19 Агрокліматичне районування (за В.П.Поповим).*

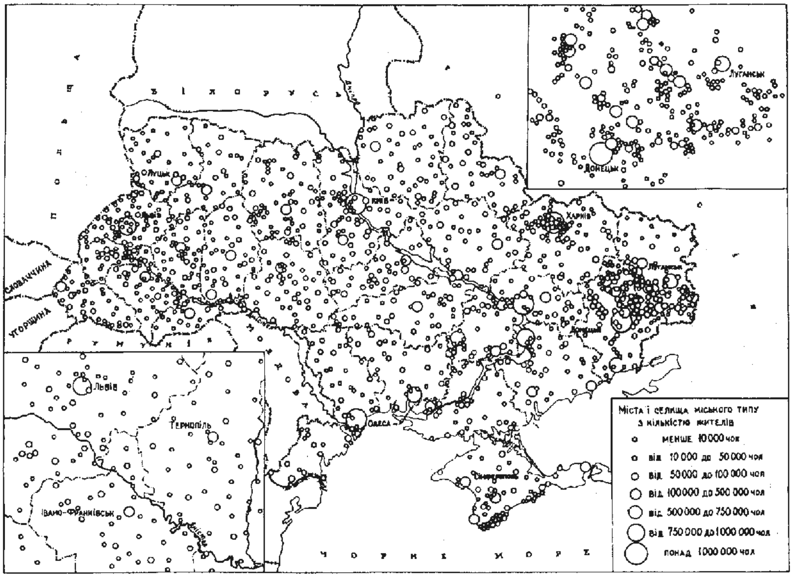
В економіці ізолінійні картограми застосовуються для встановлення часу виконання основних сільськогосподарських робіт (ізотопи), для зображення регіонів з однаковими цінами (ізопрайси) і т.д. на ізолінійних картограмах замкнутими плавними лініями зображуються контури приблизно рівних величин статистичного показника.

Недоліком штрихових картограм є те, що певний географічний регіон штрихується однаково, без переходу по густоті штрихів, хоч в дійсності розподіл любої ознаки на місцевості не завжди рівномірний.

* Там же, - С. 78.

А тому замість фарби і штрихування в якості графічних символів в картограмах використовують крапки.

В *крапкових картограмах* графічним знаком статистичних даних є крапки строго визначеного розміру, розміщені в заданих межах. Кожна крапка відповідає певній числовій величині і є носієм елементу обліку. Крапки на картограмі надають обліку наочність і природність. Крапкові картограми розподілу території України за густотою заселення міст і селищ міського типу в 1989 р. показують чітку характеристику їх розміщення в географічному розрізі (мал.. 5.20). Вони добре ілюструють ступінь концентрації об’єктів промисловості і сільського господарства в різних районах і можуть використовуватись в багатьох галузях статистики.

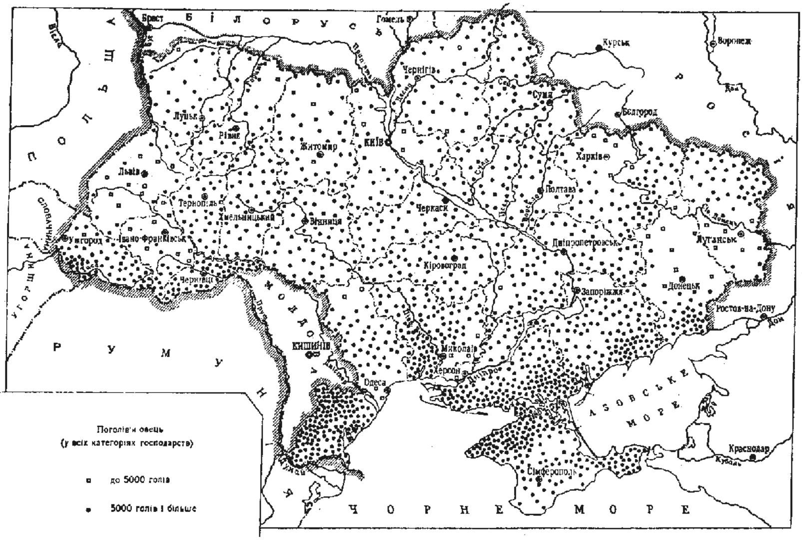



Мал. 5.20 Людність міст і селищ міського типу (1989 р.).*

Крапки на розрахунковій картограмі розміщують на контурах територій з врахуванням їх фактичного розподілу на окремих ділянках цієї території. Це дозволяє порівнювати щільність розміщення досліджуваних об’єктів на різних ділянках території за загущеністю крапок. Виразність крапкової діаграми залежить від розміру крапки.

* Заставний Ф.Д. Географія України: У 2-х книгах. – Львів: Світ, 1994. – С.184.

Якщо зменшити розмір крапки, тоді на цій же площі можна розмістити більше крапок, і не буде нашарування крапок однієї на другу. Покажемо вигляд такої діаграми на мал. 5.21.




Мал. 5.21 Розміщення поголів’я овець.*

Крапкові діаграми за своєю суттю близькі до фонових. Однак фонові картограми, як правило, використовуються для зображення середніх і відносних показників. Крапкові ж картограми використовуються для об’ємних (кількісних) показників. Їх застосовують в тих випадках, коли сума ваг статистичного розподілу по районах має економічний зміст. У фонових діаграмах сума ваг економічного змісту немає.

Якщо після заштрихування, фарбування або нанесення крапок на відповідні ділянки карти виявляється певна закономірність в *географічному* розміщенні території з однаковою величиною зображуваного показника, тоді можна судити про залежність даного показника від географічного фактора. Якщо ж райони з однаковим зображенням розміщені на карті в хаотичному порядку, це свідчить про відсутність певної закономірності в просторовому розміщенні

* Заставний Ф.Д. Географія України: У 2-х книгах. – Львів: Світ, 1994. – С.301.

даного показника, тобто розповсюдження або рівень не зв'язані з географічним положенням району.

*Картодіаграма* – це поєднання схематичної географічної карти із діаграмою. Основне завдання картодіаграм заключається в тому, щоб показати географічний розподіл зображуваного статистикою явища. Головна їх особливість заключається в розміщенні на контурній географічній карті спеціальних знаків-символів у вигляді стовпчиків, квадратів, кругів та інших. Величина геометричного знаку залежить від розміру даного явища в зображуваному районі. Знаки і символи на картодіаграмі розміщуються не в простій лінійній послідовності, а орієнтуються в географічному просторі.

Основна перевага картодіаграм перед звичайними діаграмами заключається в точній географічній орієнтації статистичних величин, у встановленні їх взаємної відповідності і просторовому розподілі.

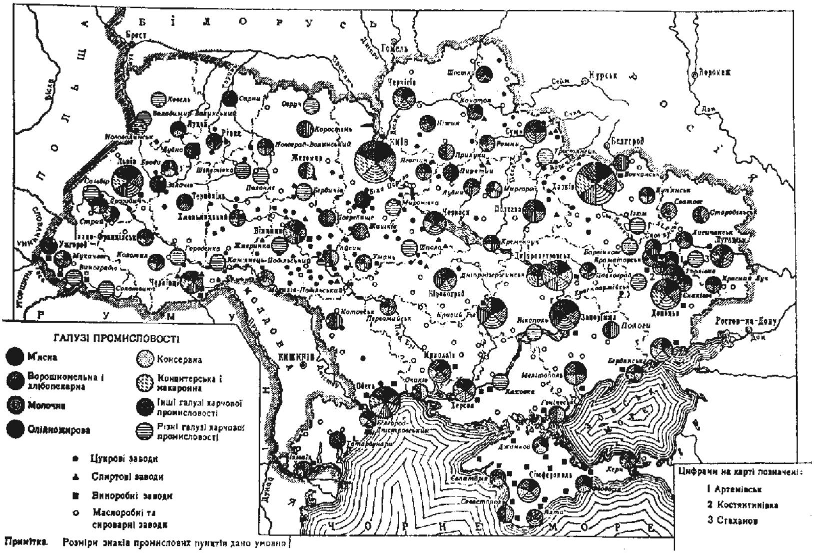
Певну перевагу картодіаграми мають також і перед картограмами. На картодіаграмі зображуються самі досліджувані величини, що сприяє більш точному їх відображенню. На картограмі зображуються головним чином середні, крайні значення, або значення інтервалів.

Основним видом знаків-символів при побудові картодіаграми є кругові і секторні діаграми. За допомогою цих символів на картодіаграмі зображують одночасно як об'єм так і структуру (склад) статистичного явища, розміщеного в просторі. Покажемо таку картодіаграму на прикладі географічного розміщення галузей харчової промисловості в Україні (мал. 5.22).

Якщо на картодіаграмі зображують лише структуру досліджуваного явища без врахування його об'єму, тоді будують круги однакового радіуса.

Для зображення розподілу по території абсолютних величин, на карту наносять прямокутники у вигляді стовпчиків або стрічок. Ці прямокутники або стрічки можна використати також для *графічного* зображення структури явища.

В економічних дослідженнях доводиться інколи поєднувати картодіаграми з картограмами. Діаграми якби накладаються на картограми. Картограми в поєднанні з картодіаграмами при вмілій їх побудові є важливим засобом наочного зображення і аналізу суспільно-економічних явищ і процесів. Наприклад, такий графік дає можливість проаналізувати територіальне розміщення міського населення за чисельністю в поєднанні із щільністю розселення сільського населення. Міста за чисельністю жителів на карті зображують за допомогою кругових діаграм, а щільність сільського населення – за вибраним штрихуванням відповідних територій.

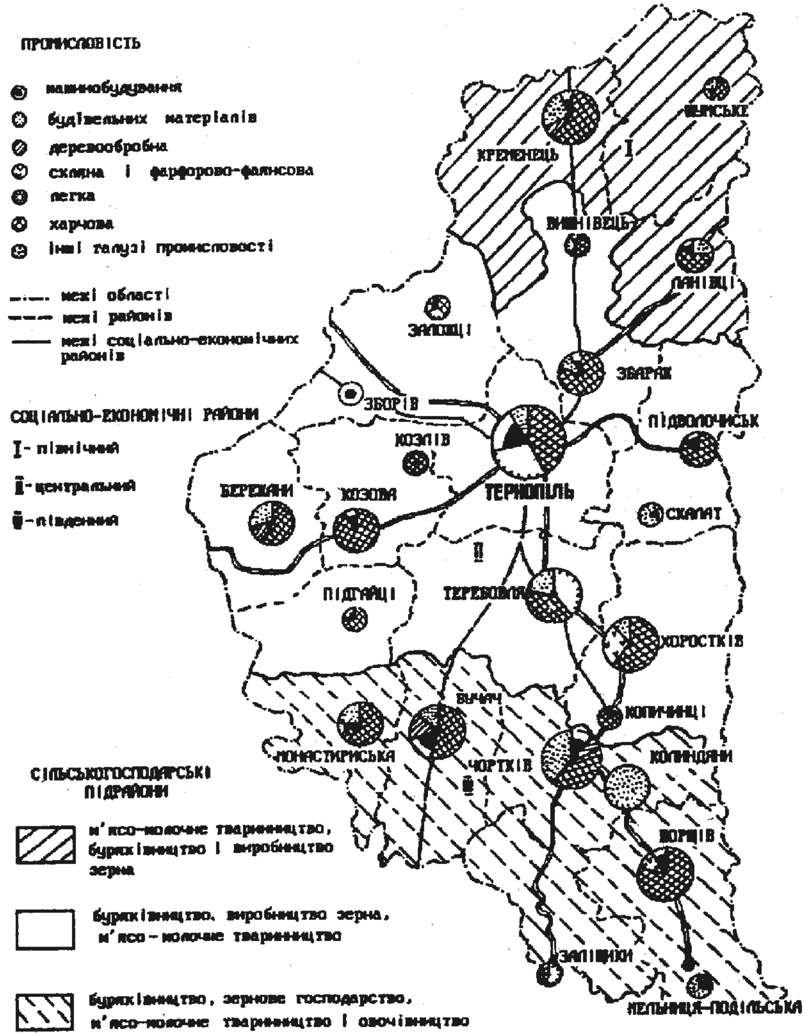



Мал. 5.22 Географія харчової промисловості.*

Соціально-економічне районування на географічній карті наочно зображують і аналізують також за допомогою поєднання цих двох графіків. Об'єм і структуру промисловості на карті показують за допомогою секторних діаграм, а напрямки сільськогосподарського виробництва – через заштриховку відповідних регіонів. Побудову такого графіка покажемо на прикладі соціально-економічного районування Тернопільської області (мал. 5.23).

*Центрограма* – це контурна карта на якій розміщуються короткі цифрові таблиці з інформацією про історико-географічний розвиток і розміщення досліджуваного явища чи процесу. Центрограми ще називають історико-географічними картами. Вони дозволяють скласти цілі статистико-географічні описи нанісши цифрові ряди на карті для різних територій, що дає можливість наочно уявити окремі сторони протікання досліджуваного процесу в динаміці.

* Заставний Ф.Д. Географія України: У 2-х книгах. – Львів: Світ, 1994. – С.304.




Мал. 5.23 Соціально-економічне районування.*

---

* Заставецька О.В., Заставецький Б.І., Ткач Д.В. Географія Тернопільської області: Навчальний посібник. – Тернопіль, 1994. – С. 62.

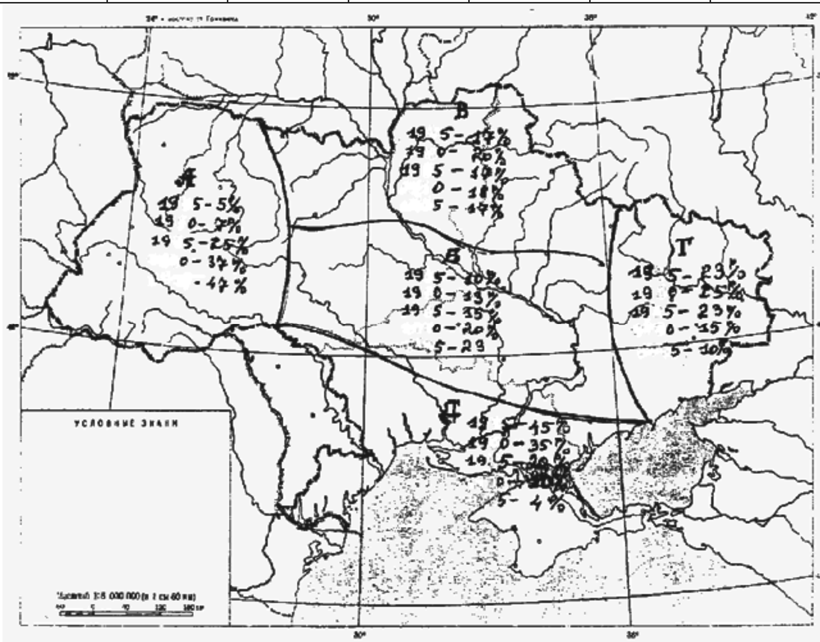
Центрограми дозволяють визначити питому вагу окремих регіонів, а також тенденцію переміщення центру розвитку в розташуванні окремих явищ.

Приклад побудови центрограми покажемо за даними табл. 5.8 на мал. 5.24.

Таблиця 5.8

Розміщення поголів'я великої рогатої худоби по території України в 1985 – 2005 рр.

| Роки | Поголів'я великої рогатої худоби в % | | | | | |
|---|---|---|---|---|---|---|
| | район А | район Б | район В | район Г | район Д | по Україні |
| 1985 | 5 | 10 | 17 | 23 | 45 | 100 |
| 1990 | 7 | 13 | 20 | 25 | 35 | 100 |
| 1995 | 25 | 15 | 17 | 23 | 20 | 100 |
| 2000 | 37 | 20 | 18 | 15 | 10 | 100 |
| 2005 | 47 | 23 | 17 | 10 | 4 | 100 |




Мал.. 5.24 Розподіл поголів'я великої рогатої худоби по районах України в 1985 – 2005 рр.

Дана центрограма достатньо виразна і показує помітне переміщення центру вирощування великої рогатої худоби з південно-східних районів через північні до центрально-західних за період з 1975 року по 1995 рік.

Центрограми широко використовуються при вивченні міграції народонаселення.

Такі дослідження проводяться центрографічним методом в багатьох країнах світу. Цей метод також використовується для вивчення переміщення центрів виробництва різних галузей народного господарства.

### *Контрольні запитання*

1. Поняття про статистичні графіки і вимоги до їх побудови.
2. Призначення графіків та переваги графічного методу.
3. Які основні елементи графіків?
4. Що собою являє шкала графіка і які її види?
5. Для чого в графіках застосовують штрихування, фарбування?
6. Загальні поняття про види графіків.
7. Застосування стовпчикових та стрічкових діаграм.
8. Квадратні і кругові діаграми та їх використання в статистиці.
9. Використання секторних діаграм.
10. Прямокутні діаграми (знаки Варзара).
11. Лінійні діаграми та їх застосування.
12. Контрольно-планові графіки.
13. Графіки просторового розміщення і просторового розповсюдження.

# Розділ 6 Абсолютні і відносні величини

## 6.1. Узагальнюючі статистичні показники

Інформація про суспільні явища і процеси створюється, передається і зберігається у вигляді *статистичних показників*. Вони є однією з основних економічних категорій, за допомогою яких відображають кількісну і якісну сторони явищ і процесів.

В результаті статистичного спостереження отримують інформацію де кожна одиниця досліджуваної сукупності у вигляді зібраних відповідних показників характеризується рядом ознак.

В назві показника (ціна товару, собівартість продукції, чисельність працівників і т.д.) відображається якісна визначеність явища зумовлена його суттю. Такий показник складається з моделі кількісної сторони і числового виразу явищ, що вивчається. Модель розкриває структуру показника і відповідає на питання: що, де, коли і яким чином підлягає вимірюванню і реєстрації. Вона обґрунтовує одиниці виміру, техніку збуту і обробку даних та ін.

Значення будь-якого статистичного показника, що досліджується, повинно володіти атрибутами місця і часу. Наприклад, чисельність населення за даними перепису на 12 січня 1989 р. в Україні становила 51,4 млн. чол.

Достовірність статистичної інформації залежить від об’єктивності відображення статистичним показникам суті досліджуваного соціально-економічного явища чи процесу. Вона характеризується адекватністю і точністю показника.

*Адекватність* або *відповідність* показника заключається в його спроможності відобразити ті властивості явищ, які намічені програмою статистичного дослідження. Для того, щоб показник відповідав призначенню і виконував свої функції, на стадії його проектування проводять логічне і статистичне обґрунтування.

Логічне – передбачає всебічний теоретичний аналіз досліджуваних економічних категорій (ціна, собівартість, рентабельність і ін).

Статистичне обґрунтування передбачає розробку методології і методики кількісної оцінки цих категорій.

*Точність* оцінки залежить від структури показника, організації статистичного спостереження та обробки отриманих даних.

Модель статистичного показника поєднує в собі як кількісну і якісну визначеність економічних категорій, так і адекватність і точність їх вимірювання.

За статистичною природою показники дуже різноманітні і поділяються за такими ознаками як спосіб вирахування, часова визначеність і адитивність.

За *способом вирахування* розрізняють первинні і вторинні (похідні) статистичні показники. Первинні отримують в результаті зведення матеріалів статистичного спостереження у формі абсолютних величин (кількість заводів галузі, вартість основних виробничих фондів і випущеної ними продукції). На основі первинних даних обчислюють вторинні, похідні від перших (середню вартість основних виробничих фондів і випуск продукції в середньому на один завод), а при визначенні інтенсивності використання основних виробничих фондів матимемо похідні показники другого порядку (фондовіддача, фондоємкість).

За *часовою ознакою* всі статистичні показники поділяють на інтервальні і моментні. Інтервальні показники характеризують явище за певний проміжок часу (тиждень, декада, місяць, квартал, рік). Наприклад, виробництво продукції за місяць, валовий збір зерна за сезон, чисельність народжених на протязі року та ін. Моментні показники дають кількісну характеристику явищ на певний момент часу (поголів'я великої рогатої худоби на початок кожного року, перепис багаторічних насаджень на початок року, чисельність працівників на перше число кожного місяця і т.д.). Інтервальні і моментні статистичні показники можуть бути первинними, а також похідними різних порядків від них.

Характерною особливістю первинних інтервальних показників є *адитивність*, тобто їх сумування має економічний зміст. Первинні інтервальні показники, а також похідні від них залежать від проміжку часу, за котрий вони вираховуються.

Моментні і в своїй більшості похідні показники неадитивні, тому що їх підсумування немає економічного змісту.

На другому етапі статистичного дослідження узагальнюють зібранні первинні матеріали і в результаті отримують зведені таблиці, де вся сукупність одиниць подається в цілому і по групах. Таким чином, *узагальнюючими показниками* в статистиці називаються показники, які характеризують сукупність одиниць в цілому або по групах.

Узагальнення показників є важливим завданням, а також характерним і специфічним для статистики методом.

Узагальнюючі статистичні показники можуть бути об'єднані в три великі групи: абсолютних, відносних і середніх величин.

Початковим видом узагальнюючих показників, як було відмічено вище, виступають абсолютні величини, отримані безпосередньо в результаті зведення первинного матеріалу. А вже на основі абсолютних величин вираховують відносні і середні величини.

В більшості випадків, для характеристики якого-небудь явища часто використовують всі три види узагальнюючих показників. Наприклад, при дослідженні роботи торгових підприємств облспоживспілки магазини групують за товарними групами, в результаті отримують зведену статистичну таблицю (табл. 6.1, макет).

Таблиці 6.1

Характеристика роботи ОСС за 199_ р.

| Групи товарів по торгах | Число магазинів | Товаро-оборот, млн. грн. | Питома вага, % | | Товарооборот в середньому на один магазин, млн. грн. |
|---|---|---|---|---|---|
| | | | за числом магазинів | за товаро-оборотом | |
| А | 1 | 2 | 3 | 4 | 5 |
| Продуктові<br>Непродуктові<br>Разом: | | | 100 | 100 | |

По кожному торгу зокрема і в цілому по всій торгівлі системи облспоживспілки підраховують число магазинів і об’єм товарообороту. Таким чином отримують узагальнюючі абсолютні величини, які характеризують загальну чисельність досліджуваної сукупності (число магазинів) і загальний обсяг ознаки, яка відображає результати торгової діяльності (об’єм товарообороту). Для з’ясування питомої ваги окремих торгів в системі споживчої кооперації області, або в загальному об’ємі товарообороту, вираховують відносні величини (процентне відношення частини до цілого). Користуючись середніми величинами (обчисливши об’єм товарообороту в середньому на один магазин) визначають характерні для кожного торгу розміри магазинів. Таким чином, комплексне використання всіх трьох узагальнюючих статистичних величин дозволяє різнобічно охарактеризувати досліджуване явище чи процес.

### 6.2. Абсолютні статистичні величини.

*Абсолютними величинами* в статистиці називаються первинні узагальнюючі показники, які характеризують суспільні явища і процеси в конкретних умовах місця і часу. Наприклад, чисельність населення Тернопільської області на 1 січня 1997 р. становила

1172,4 тисячі чоловік, у тому числі: міського – 512,1 тисячі чоловік і сільського – 660,3 тисячі чоловік.

Абсолютні статистичні величини отримують в результаті зведення шляхом безпосереднього підрахунку первинного статистичного матеріалу, або розрахунків, на основі інших показників досліджуваної сукупності (наприклад, сума припливу грошей в каси ощадбанку визначається як різниця між поступленням грошей і їх видачею з кас).

Характерною особливістю абсолютних величин є те, що вони безпосередньо зв'язані із соціальною, економічною і природною основою або речовою формою явищ і процесів, які вони представляють. Абсолютні величини відображають кількісну сторону певної суті явища тієї чи іншої його властивості.

Абсолютні величини як узагальнюючі показники характеризують сукупність за її *чисельністю* (число працівників, кількість магазинів, бібліотек, лікарень) і *об'ємом* (валовий випуск продукції, фонд заробітної плати, обсяг роздрібного товарообороту і т.д.).

Об'ємні показники, в свою чергу, часто використовують для характеристики сукупності в цілому або її частини.

Статистика виділяє три види абсолютних величин: індивідуальні, групові і загальні.

*Індивідуальними* називаються такі абсолютні величини, які виражають розміри кількісних ознак окремих одиниць досліджуваної сукупності. Індивідуальні абсолютні величини установлюються безпосередньо під час проведення статистичного спостереження і реєструються у формулярах спостереження.

На базі індивідуальних абсолютних величин утворюються загальні і групові абсолютні величини.

*Групові* і *загальні* абсолютні статистичні величини виражають величину ознаки у всіх одиниць даної сукупності, або окремих її груп.

Абсолютні статистичні величини виражають розміри явищ в таких одиницях міри як: вага, об'єм, площа, довжина, вартість і ін.

Абсолютні статистичні величини завжди числа іменовані. Вони мають певну розмірність і певні одиниці виміру.

Вибір одиниць виміру абсолютних величин залежить від фізичних і соціально-економічних властивостей явища, а також від суті і мети дослідження. В статистиці використовується велике число різноманітних одиниць виміру, які можна об'єднати в три групи: натуральні, вартісні і трудові.

*Натуральними* називаються одиниці виміру, які виражають розміри конкретних явищ у фізичних вимірниках (тоннах, кілограмах, метрах, гектарах, літрах, кубометрах і ін.).

Натуральні одиниці виміру можуть бути простими, складними і умовно-натуральними.

Складні натуральні одиниці виміру отримують шляхом перемноження двох величин різних розмірностей. Наприклад, потужність електродвигунів вимірюється в кіловатах, а спожита ними енергія в складних одиницях – кіловат-годинах, об’єм перевезених вантажів вимірюється в тоннах, а вантажооборот – в тонно-кілометрах, верстатний парк цеху обчислюється в штуках, а робота верстатів у верстато-днях, верстато-змінах і т.д.

Як показують наведені приклади, складні одиниці вимірюють сукупності, які характеризують результати функціонування продуктивних сил, і відображають два фактори: в першому випадку – потужність електродвигунів і час їх роботи, в другому – об’єм вантажів і віддаль перевезень, в третьому – кількість верстатів і тривалість їх роботи.

В ряді випадків статистика використовує умовно-натуральні одиниці виміру. Такі одиниці виміру використовуються для зведення докупи декількох різновидностей однакової споживної вартості. Одну з них приймають за еталон, а всі інші перераховують за допомогою спеціальних перевідних коефіцієнтів в одиниці виміру взятого еталону. Перерахунок в умовно-натуральні одиниці здійснюють за формулою:

$$у = е + k \cdot н,$$

де у – кількість умовно-натуральних одиниць;

е – кількість еталонних одиниць;

н – кількість одиниць сукупності, які відрізняються від еталонних;

k – коефіцієнт перерахунку не еталонних одиниць сукупності в еталонні.

Коефіцієнт перерахунку – це відношення величини якої-небудь властивості з різновидностей складових елементів явища до його величини, прийнятої за еталон. Наприклад, нехай маємо 150 т. мила із 40 % вмістом жирних кислот, 100 т. – із 50 % і 50 т. – із 60 % вмістом жирних кислот. Потрібно перерахувати все мило на 40 %-не. Спочатку визначимо коефіцієнти перерахунку: для 40 % мила $k_1 = \frac{40}{40} = 1$, для 50 % – $k_2 = \frac{50}{40} = 1{,}25$ і для 60 % – $k_3 = \frac{60}{40} = 1{,}5$. Перерахуємо все мило в умовно-натуральне (на 40 %-не, взяте за еталон):

$$150 \cdot 1 + 100 \cdot 1{,}25 + 50 \cdot 1{,}5 = 350 \text{ т.}$$

Отже, ми маємо 350 умовно-натуральних тонн мила.

Перерахунок в умовно-натуральні одиниці виміру отримав велике розповсюдження в практичній роботі і наукових дослідженнях. Використовуються найрізноманітніші перерахунки: паливо різних видів перераховують в еталонне кам'яне вугілля, теплотворна спроможність якого становить 29,309 Дж/кг (7000 ккал/кг); консервну продукцію вимірюють в тубах (тисячах умовних банок), за умовну банку приймається банка вагою нетто 400 г. – для одних видів консервів і банка об'ємом 353,4 $см^3$ – для інших; в сільському господарстві різні види робіт перераховують в умовну ріллю, різні види рогатої худоби – в умовні голови, різні види кормів – в кормові одиниці; деякі продукти харчування перераховують за засвоєністю частки білка, який вміщується в кожному продукті та ін.

Потрібно пам'ятати, що умовні абсолютні величини непридатні для обчислення різнорідної продукції.

*Вартісними* називаються одиниці виміру, які використовуються для характеристики в грошовому виразі багатьох різноманітних статистичних показників. Наприклад, собівартість і ціна одиниці продукції обліковується в гривнях і копійках, обсяг товарообороту продуктового магазину – в тисячах гривень, а національний дохід держави в мільйонах або мільярдах гривень. Вартісні вимірники дають можливість зіставляти найрізноманітніші види продукції. Недоліком цього виду одиниць виміру є те, що з плином часу ціни на окремі товари і послуги змінюються, а тому деякі сумарні величини стають не порівняльними. Цей недолік усувається шляхом переоцінки цих сум в ціни одного і того ж періоду (порівняльні ціни).

*Трудовими* називаються одиниці виміру, які використовуються для обліку затрат робочого часу, для визначення рівня продуктивності праці, величини трудових ресурсів і раціонального їх використання, та для деяких інших розрахунків. Трудові вимірники виражаються в людино-годинах, людино-днях, людино-роках, машино-днях, верстато-днях, коне-днях.

Поряд з абсолютними показниками, отриманими шляхом зведення даних спостереження, статистика використовує абсолютні показники, отримані розрахунковим шляхом. Так, об'єм національного доходу обчислюють на базі даних про валову продукцію галузей матеріального виробництва і даних про матеріальне споживання в цих галузях. Багато абсолютних величин розраховують балансовим методом. Розрахунок ґрунтується на тому, що в балансі всі показники зв'язані між собою так званою балансовою ув'язкою. Баланс складається з двох розділів: активу і пасиву. Баланс руху товарно-матеріальних цінностей в простій формі в активі має залишок на початок періоду і поступлення, а в пасиві – вибуття і залишок на кінець періоду. Сума активу завжди рівна сумі пасиву. А тому, якщо один з чотирьох показників невідомий, його легко визначити через балансову ув'язку. Якщо, наприклад, невідомий об'єм реалізації товарно-матеріальних цінностей за звітний період, то його розраховують як залишок на початок періоду плюс поступлення і мінус залишок на кінець періоду. Нехай залишок товарів на початок кварталу в одному з торгів становив 100 млн. грн., поступило за даний період товарів на суму 500 млн. грн., а залишок товарів на кінець цього кварталу склав 200 млн. грн. Об'єм реалізації за даний період становитиме:

$$100 + 500 - 200 = 400 \text{ млн. грн.}$$

Балансову ув'язку можна виразити рівністю:

$$a + b = c + d$$

де $a$ – залишок на початок періоду;
$b$ – поступлення за даний період;
$c$ – вибуття за даний період;
$d$ – залишок на кінець періоду.

Таким чином за балансовою ув'язкою:

$$100 + 500 = 400 + 200,$$

звідси $a = c + d - b = 400 + 200 - 500 = 100$ млн. грн.;
$b = c + d - a = 400 + 200 - 100 = 500$ млн. грн.;
$c = a + b - d = 100 + 500 - 200 = 400$ млн. грн.;
$d = a + b - c = 100 + 500 - 400 = 200$ млн. грн.

Розраховують також об'єм ознаки за даними про його середнє значення і чисельність одиниць сукупності. Наприклад, якщо відома середня вага помідорів в ящику (25 кг.) і кількість ящиків з помідорами, завезеними на продуктовий ринок (300 ящиків), то, перемноживши середню вагу одного ящика з помідорами на кількість ящиків, визначимо вагу всіх помідорів ($25 \cdot 300 = 7500$ кг., або 7,5 т.). Аналогічно визначають обсяг реалізації і інших видів товарів на продуктових ринках.

Абсолютні статистичні величини мають велике практичне і наукове значення. Вони характеризують наявність всіх ресурсів – матеріальних, грошових, трудових; розміри виробництва різних видів продукції і національного багатства країни та ін. Абсолютні величини використовують для розробки народногосподарських планів і програм та контролю за ходом їх виконання. Вони є основою для обчислення різних видів відносних і середніх величин, індексів та інших узагальнюючих показників.

### 6.3. Відносні величини

*Відносними величинами* називають статистичні показники, які виражають кількісні співвідношення між соціально-економічними явищами і процесами. Їх отримують шляхом порівняння (ділення) двох однойменних, або різнойменних величин.

Величина, з якою проводять порівняння, називається основою відносної величини, базою порівняння або базисною величиною. Величина, яку порівнюють, називається поточною, порівнюваною чи звітною величиною.

Відносні величини показують, у скільки разів порівнювана величина більша (менша) за базисну, або яку частку перша займає в другій, або скільки одиниць однієї величини припадає на одиницю другої.

В залежності від бази порівняння відносні величини можуть виражатися у формі:

а) коефіцієнтів – якщо база порівняння приймається за одиницю;

б) процентів (%) – якщо база порівняння береться за 100;

в) проміле (‰) – якщо за базу порівняння взято 1000;

г) продециміле ($^{0}/_{000}$) – якщо база порівняння становить 10000;

д) просантиміле ($^{0}/_{0000}$) – якщо база порівняння прийнята за 100000.

Вибір форми вираження відносної величини визначається розмірністю порівнювальних величин з метою надання даній відносній величині найбільшої виразності. Якщо порівнювальна величина більша за базу порівняння то виникає проблема вибору між коефіцієнтом і процентом. Коефіцієнт показує в скільки разів порівнювана величина більша за базу порівняння. При процентному виразі, якщо від нього відняти 100, судять на скільки процентів порівнювана величина більша за базу порівняння.

Якщо порівнювана величина менша за базу порівняння, доцільно застосовувати проценти, а якщо значно менша – проміле, продециміле або просантиміле.

Проміле широко застосовується в статистиці населення, де народжуваність, смертність, природний приріст та інші демографічні показники вираховуються на 1000 душ населення. Таку форму відображення відносних величин використовують для зручності їх сприйняття і тлумачення.

При обчисленні відносних величин найчастіше порівнюють дві абсолютні величини. В деяких випадках порівнюють також відносні і середні величини.

Відносні величини мають велике значення при соціально-економічному аналізі, тому що абсолютна величина, взята сама по собі, не завжди дає правильну оцінку явища. В багатьох випадках тільки в порівнянні з іншою абсолютною величиною дана величина проявляє свою істинну значущість. Якщо, наприклад, відомо, що в одній області протягом року народилось 15000 дітей, а в другій 18000 дітей, то це ще нічого не говорить про рівень народжуваності в цих областях. Тільки зіставивши ці величини із загальною чисельністю населення даних областей, ми матимемо уявлення про рівні народжуваності. Нехай рівень народжуваності в першій області – 12,8 ‰ [(15000 : 117000) · 1000], в другій – 10,0 ‰ [(18000 : 1800000) · 1000], то це свідчить про те, що в першій області народжуваність в 1,28 рази вища, ніж в другій.

В залежності від змісту і пізнавального значення відносні величини, що використовуються в статистиці, поділяються на наступні основні види: відносні величини планового завдання, виконання плану, динаміки, структури, координації, інтенсивності і порівняння.

*Відносні величини планового завдання* характеризують відношення запланованого рівня показника до його фактично досягнутого рівня в минулому (до планового) періоді. Так, наприклад, якщо в 2007 р. завод виготовив продукції на 10 млн. грн., а на 2008 р. планується 10,5 млн. грн., то відносна величина планового завдання складе 10,5 : 10 = 1,05, або 105 %. В народногосподарських планах планові завдання доводяться як в абсолютних показниках, так і у формі відносних величин.

Планові завдання у формі відносних величин досить часто встановлюються, особливо за такими показниками, як продуктивність праці, собівартість продукції, рентабельність та ін.

*Відносною величиною виконання плану* називається статистичний показник, який характеризує ступінь виконання планового завдання, встановленого на даний період. Іншими словами, це процентне відношення фактично досягнутого рівня до запланованого за відповідний період часу (місяць, квартал і т.д.)

Відносні величини виконання плану мають велике народногосподарське значення, тому що вони є знаряддям розв’язку статистикою одного з найбільш важливих її завдань – контролю за ходом виконання планів окремим підприємством, об'єднанням, галузями народного господарства і народним господарством в цілому.

Припустимо, що підприємство за планом повинно було випустити на протязі другого кварталу продукції на суму 300 тис. грн., а фактично випустило на 315 тис. грн., тоді процент виконання плану буде становити 105 % (315 : 300 · 100). Отже, план було виконано на 105 %, або перевиконано на 5 %.

У випадку, коли план доведений у формі відносної величини, для визначення ступеня його виконання замість абсолютних величин у чисельнику і знаменнику беруть відносні величини. Тому для визначення ступеня виконання плану спочатку визначають фактичні зміни цього показника за звітній період порівняно з базою, відносно якої обчислене планове завдання, а потім ділять фактичний показник на запланований і множать на 100.

Наприклад, за річним планом намічалось підвищити продуктивність праці робітників заводу на 6 %, фактично вона зросла, за цей же період, на 7 %. Тобто, план підвищення продуктивності праці робітників заводу виконаний на 116,7 % [(107 : 106) · 100], або перевиконаний на 16,7 %.

Якщо планом передбачалось зниження рівня показника, наприклад, собівартості випуску продукції за рік на 5 %, а фактично вона знизилась на 7 %, то процент виконання плану становитиме 97,9 % [(100 – 7) : (100 – 5) · 100] = (93 : 95 · 100). Отже, фактична собівартість продукції нижча від планової на 2,1 % (100 – 97,9).

*Відносні величини динаміки* характеризують зміну рівня того чи іншого явища в часі. В залежності від характеру бази порівняння, розрізняють відносні величини динаміки із змінною базою порівняння, або ланцюгові, і відносні величини з постійною базою порівняння, або базисні.

Розраховані нами базисні і ланцюгові відносні величини динаміки показані в табл. 6.2.

Відносні величини динаміки характеризують напрямок і швидкість зміни явищ в часі, тобто темп розвитку.

Якщо ланцюгові темпи розвитку показують, як змінюється величина показника від одного періоду до іншого, то базисні величини характеризують поступове віддалення від базисного періоду, тобто рух вперед від вихідної (початкової) точки.

Таблиця 6.2

Валовий збір картоплі в області за 2002-2007 рр.

| Рік | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|---|---|
| Валовий збір картоплі, млн. т. | 21 | 23 | 27 | 30 | 31 | 36 |
| В процентах до 2002 р. (базисні) | 100,0 | 109,5 | 128,6 | 142,8 | 147,6 | 171,4 |
| В процентах до попереднього року (ланцюгові) | - | 109,5 | 117,4 | 111,1 | 103,3 | 116,1 |

Вибір бази порівняння для базисних відносних величин має суттєве значення. В якості бази порівняння потрібно брати дані за ті роки, які мають важливе (віхове) значення в розвитку досліджуваного процесу.

Відносні величини планового завдання, виконання плану і динаміки взаємозв'язані між собою. Для розгляду цього зв'язку позначимо абсолютну величину базисного періоду через $У_О$, планового завдання – через $У_П$ і виконання плану (звітного періоду) – через $У_1$. Виразимо, через прийняті символи, досліджувані нами відносні величини:

$$\text{планового завдання} = \frac{У_П}{У_О};$$

$$\text{виконання плану} = \frac{У_1}{У_П};$$

$$\text{динаміки} = \frac{У_1}{У_О}.$$

Тепер неважко побачити, що відносна величина динаміки дорівнює добутку відносних величин планового завдання і виконання плану: $\frac{У_1}{У_О} = \frac{У_П}{У_О} \cdot \frac{У_1}{У_П}$. Цей взаємозв'язок дозволяє визначити і дві інші відносні величини. Так, відносну величину планового завдання визначають шляхом ділення відносної величини динаміки на відносну величину виконання плану: $\frac{У_1}{У_О} : \frac{У_1}{У_П} = \frac{У_1 \cdot У_П}{У_О \cdot У_1} = \frac{У_П}{У_О}$. Зрештою, відносна величина виконання плану – це частка від ділення відносної величини динаміки на відносну величину планового завдання:

$\frac{У_1}{У_О} : \frac{У_П}{У_О} = \frac{У_1 \cdot У_О}{У_О \cdot У_П} = \frac{У_1}{У_П}$. Таким чином, знаючи любі дві з цих трьох відносних величин, завжди можна обчислити третю. Наприклад, нехай маємо наступні дані про об'єм товарообороту магазину № 1:

Таблиця 6.3

| Номер магазину | Об'єм роздрібного товарообороту, тис. грн. | | |
|---|---|---|---|
| | 2006 р. (звітний) | 2007 р. | |
| | | за планом | фактичний |
| 1 | 4280 | 4300 | 4350 |

За даними табл. 6.3 відносні величини дорівнюють: планового завдання $- \frac{4300}{4280} \cdot 100 = 100{,}5$ %; виконання плану $- \frac{4350}{4300} \cdot 100 = 101{,}1$ % і динаміки $- \frac{4350}{4280} \cdot 100 = 101{,}6$ %. Взаємозв'язок:

відносна величина планового завдання = $\frac{101{,}6}{101{,}1} \cdot 100 = 100{,}5$ %;

відносна величина виконання плану = $\frac{101{,}6}{100{,}5} \cdot 100 = 101{,}1$ %;

відносна величина динаміки – 1,005 · 1,011 = 1, 016 або 101,6 %.

*Відносні величини структури* характеризують питому вагу окремих частин досліджуваної сукупності в загальному її об'ємі. Їх обчислюють шляхом відношення частини до цілого. Дані відносні величини виражаються в частках одиниці або процентах. Наприклад, як було сказано вище, загальна чисельність населення Тернопільської області на 1 січня 2007 р. становила 1172,4 тис. чол., в тому числі: міського населення – 512,1 тис. чол. І сільського – 660,3 тис. чол. Отже, питома вага міського населення в області становила 43,7 % (512,1 : 1172,4 · 100), сільського населення – 56,3 % (660,3 : 1172,4 · 100).

З допомогою відносних величин структури виявляють також структурні зрушення в досліджуваній сукупності, тобто зміну її складу, напрямок і тенденцію за різні періоди часу. Структурні зрушення дозволяють простежити і вивчити внутрішні зміни, що відбуваються усередині складного явища.

Для прикладу структурних зрушень наведемо зміни в розподілі чисельності багатодітних матерів, які отримують щомісячну державну допомогу (табл. 6.4, дані умовні):

Таблиця 6.4

Розподіл багатодітних матерів, які отримують державну допомогу (%)

| Показники | Рік | | | | |
|---|---|---|---|---|---|
| | 1970 | 1980 | 1990 | 2000 | 2007 |
| Всього багатодітних матерів, які отримують державну допомогу | 100,0 | 100,0 | 100,0 | 100,0 | 100,0 |
| В тому числі: | | | | | |
| з чотирма дітьми | 67,9 | 65,8 | 61,2 | 60,0 | 59,2 |
| з п'ятьма дітьми | 18,8 | 17,1 | 14,9 | 13,9 | 13,4 |
| з шістьма дітьми | 7,2 | 9,2 | 10,4 | 9,1 | 8,5 |
| з сімома дітьми і більше | 6,1 | 7,9 | 13,5 | 17,0 | 18,9 |

Дані табл. 6.4 показують суттєві зміни в структурі багатодітних матерів, які отримують державну допомогу за період 1969-1997 рр. Помітно знизилась доля матерів з чотирма і п'ятьма дітьми і дуже сильно, більше ніж в 3 рази, зросла питома вага матерів із сімома і більше дітьми.

Відносні величини структури широко використовуються при статистичному дослідженні розвитку народного господарства, змін у складі населення, розміщенні продуктивних сил і ін.

Вивчення структури і структурних зрушень суспільних явищ і процесів є одним з важливих завдань статистики.

*Відносні величини координації* характеризують співвідношення частин цілого між собою. Одну із складових частин цілого приймають за базу порівняння, а всі інші частини знаходять як відношення до неї. Ці відносні величини характеризують не структуру сукупності, а співвідношення частин між собою, і являються особливим видом відносних величин.

За допомогою відносних величин координації визначають, скільки одиниць даної частини цілого припадає на 1, на 100, на 1000, на 10000 одиниць іншої частини, взятої за базу порівняння. Наприклад, кількість інженерів, службовців на 100 робітників, число жінок що припадає на одного чоловіка і навпаки, співвідношення хлопчиків і дівчаток в сукупності народжених і таке ін.

За даними Статистичного щорічника України за 1995 р. (К., Техніка, 1996) чисельність населення України на 1 січня 1996 р. Становила 51,3 млн. чол., в тому числі: чоловіків – 23,9 млн. чол. (46,6 %) і жінок – 27,4 млн. чол. (53,4 %), тобто на 1000 жінок припадало 872 чоловіки ($\frac{23,9}{27,4}\cdot 1000$), або на 1000 чоловіків припадало 1146 жінок ($\frac{27,4}{23,9}\cdot 1000$).

Покажемо розрахунок відносних величин координації на іншому прикладі.

Таблиця 6.5

Чисельність чоловіків і жінок в економічному вузі за формами навчання:

| Групи студентів за статтю | Групи студентів за формою навчання | | |
|---|---|---|---|
| | денна | вечірня | заочна |
| Чоловіки | 730 | 250 | 940 |
| Жінки | 1770 | 450 | 1980 |

Відносні величини координації за формами навчання:

денна: $\frac{1770}{730} = 2,4$ жінки на одного чоловіка, або

$\frac{730}{1770}\cdot 100 = 41$ чоловіків на 100 жінок;

вечірня: $\frac{450}{250} = 1,8$ жінки на одного чоловіка, або

$\frac{250}{450}\cdot 100 = 55$ чоловіків на 100 жінок;

заочна: $\frac{1980}{940} = 2,1$ жінки на одного чоловіка, або

$\frac{940}{1980}\cdot 100 = 47$ чоловіків на 100 жінок;

Характеризуючи співвідношення між найважливішими складовими частинами цілого, відносні величини координації дозволяють контролювати і регулювати дотримання необхідних пропорцій між ними.

*Відносні величини інтенсивності* характеризують ступінь поширення або розвитку даного явища в певному середовищі. Їх отримують шляхом зіставлення двох різнойменних абсолютних величин, пов'язаних між собою, але які не являються складовими цілого. Чисельник відносної величини інтенсивності виражає розмір досліджуваного явища, а знаменник – розмір середовища, в якому воно поширюється чи розвивається. Відносні величини інтенсивності завжди числа іменовані і виражаються одиницями виміру тих абсолютних величин, на основі яких вони розраховуються.

До цього виду відносних величин належить показник густоти населення, який є результатом відношення чисельності населення певного регіону до відповідної території і характеризується числом жителів на 1 кв. км. площі. Наведемо для прикладу дані про густоту населення в Україні і деяких її регіонах станом на 1 січня 1996 р.*: в Україні – 85 чол./км$^2$, ($\frac{51334{,}1 \text{ тис. чол.}}{603{,}7 \text{ тис. км}^2} = 85{,}0$), в т.ч. в областях: Дніпропетровській – 120, 8 чол./км$^2$; Донецькій – 196,3; Закарпатській – 100,6; Івано-Франківській – 105,5; Луганській – 104,4; Львівській – 126,7; Чернівецькій – 116,5; Чернігівській – 42,8; Херсонській – 44,4; Житомирській – 49,5; Кіровоградській – 49,8.

Як показують приведені дані, розташування населення по території України дуже нерівномірне. Так, в Дніпропетровській області порівняно з Чернігівською густота населення щільніша більш ніж в 4,5 рази ($\frac{196{,}2}{42{,}8}$); Херсонською – в 4,4 рази ($\frac{196{,}2}{44{,}4}$); Житомирською – в 4 рази ($\frac{196{,}2}{49{,}5}$) і т.д.

Прикладами відносних величин інтенсивності можуть служити й інші статистичні показники, широко застосовувані в статистичному аналізі, як наприклад: виробництво продуктів на душу населення; вихід сільськогосподарської продукції на 100 га земельних угідь, фондовіддача, фондоємкість, фондоозброєність, продуктивність праці і т.п.

*Відносні величини порівняння* характеризують співвідношення однойменних величин, що стосуються одного й того ж періоду або моменту часу, але різних об’єктів чи територій.

Їх виражають в процентах, або коефіцієнтах, які показують, у скільки раз одна порівняльна величина більша від іншої. Так, знаючи, що в Києві на початок 1996 р. Проживало 2638,7 тис. чол., а в Севастополі – 406,9 тис. чол., знайдемо, що чисельність населення

---

* Розміщення продуктивних сил України: Підручник / Є.П. Качан, М.О. Ковтонюк, И.О. Петрига та ін.: За ред. Є.П. Качана. – К.: Вища школа, 1997. – С. 37.

Києва приблизно в 6,5 рази більша, ніж чисельність населення Севастополя.*

Відносні величини порівняння знаходять широке використання в порівняльній оцінці розмірів територій, величин посівних площ, об'ємів промислової продукції окремих областей, регіонів, міст і т.д.

Відносні величини порівняння застосовуються в міжнародних зіставленнях, таких як виробництво важливих видів продукції на душу населення і ін.

## 6.4. Комплексне використання абсолютних і відносних статистичних величин

Явища суспільного життя надзвичайно складні і багатогранні. Будь-який узагальнюючий показник спроможний відтворити лише одну грань предмета пізнання. Розмір соціально-економічних явищ і їх кількісні співвідношення змінюються залежно від часу і місця з неоднаковою швидкістю і в різних напрямках. Це зумовлює необхідність комплексного підходу до вивчення конкретних суспільних явищ, а отже, і диференційованого використання в економічному аналізі абсолютних і відносних величин. Така умова випливає безпосередньо з характеру взаємозв'язку абсолютних і відносних величин. Взаємозв'язок між даними величинами заключається в тому, що відносні величини є похідними від абсолютних величин, виражають співвідношення між ними, а тому змінюються в залежності від зміни абсолютних величин.

Кількісне вираження відносних величин залежить не тільки від числової відмінності між порівнюваними абсолютними величинами, але й від розміру бази порівняння, чим вона менша, тим більша відносна величина і навпаки. Тому одна і та ж абсолютна величина, в залежності від бази порівняння, може бути виражена різною відносною величиною. Або однаковому проценту можуть відповідати різні абсолютні значення ознак. Якщо брати у відриві абсолютні і відносні величини, то вони не дадуть ясного і чіткого уявлення про досліджувані явища і процеси.

Наприклад, в двох областях України "А" і "Б" із середньорічною чисельністю населення відповідно 943,6 і 3088,4 тис. чол. протягом року народилось однакове число дітей, по 12 тис. чол. в кожній. Але

---

* Розміщення продуктивних сил України: Підручник / Є.П. Качан, М.О. Ковтонюк, И.О. Петрига та ін.: За ред. Є.П. Качана. – К.: Вища школа, 1997. – С. 37.

без застосування спеціальної відносної величини – коефіцієнта народжуваності, судити в якій області народжуваність вища неможливо. З цією метою обчислимо коефіцієнти народжуваності:

для області "А" $К_Н = \frac{12}{943{,}6} \cdot 1000 = 12{,}7$ ‰;

для області "Б" $К_Н = \frac{12}{3088{,}4} \cdot 1000 = 3{,}9$ ‰.

Отже, в області "А" народжуваність вища ніж в області "Б" в 3,2 рази ($\frac{12{,}7}{3{,}9}$).

Візьмемо інший приклад. Нехай дві торгових точки виконали план товарообороту на 112 % кожна, тобто обидві його перевиконали на 12 %. Але якщо взяти до уваги, що в першій з них при плані товарообороту 12 млн. грн. Фактичне виконання становить 13,44 млн. грн., тобто зверх плану виручено 1,44 млн. грн., а в другій – при плані 500 тис. грн. Фактичне виконання склало 560 тис. грн., або перевиконане на 60 тис. грн., то різниця, як бачимо, дуже суттєва.

Особливу увагу при розрахунку відносних величин потрібно приділяти питанню порівнянності порівнюваних абсолютних величин. Обов'язково потрібно, щоб при розрахунку відносних величин виконання плану, динаміки і порівняння, їх абсолютні величини були порівняні між собою по періодах або моментах часу, до яких вони відносяться за одиницею виміру їх обсягу, за колом охоплюваних об'єктів, за методикою їх обчислення, за територією і деякими іншими ознаками.

Точне дотримання вимоги порівнянності особливо потрібне при порівнянні народногосподарських показників між різними країнами світу в зв'язку з неоднаковою методологією їх розрахунків.

Таким чином, тільки в поєднанні і взаємному доповненні одних одними, абсолютні і відносні величини дають можливість повніше і глибше проаналізувати явища суспільного життя, їх особливості й закономірності.

Відносні величини у взаємозв'язку з абсолютними величинами виступають як важливий засіб інформації і аналізу різних сторін соціально-економічних явищ і процесів, що є найважливішим принципом їх використання.

### *Контрольні запитання*

1. Поняття про узагальнюючі статистичні показники.
2. Способи вирахування статистичних показників.
3. Поняття про первинні і вторинні показники.
4. Як поділяються статистичні показники за часовою ознакою?
5. Поняття про абсолютні статистичні показники.
6. Індивідуальні, групові і загальні абсолютні статистичні величини.
7. Натуральні одиниці виміру.
8. Необхідність застосування складних натуральних одиниць виміру.
9. Поняття та необхідність вирахування умовно-натуральних одиниць виміру.
10. Для чого використовують вартісні одиниці виміру?
11. Трудові одиниці виміру і їх статистичне використання.
12. Отримання абсолютних показників розрахунковим шляхом.
13. Поняття про відносні величини.
14. Види відносних величин.
15. Форми вираження відносних величин.
16. Відносні величини планового завдання.
17. Відносні величини виконання плану.
18. Відносні величини динаміки.
19. Взаємозв’язок відносних величин планового завдання, виконання плану і динаміки.
20. Відносні величини структури.
21. Відносні величини координації.
22. Відносні величини інтенсивності.
23. Відносні величини порівняння.
24. Комплексне використання абсолютних і відносних величин.
25. Питання порівнянності порівнюваних абсолютних величин при розрахунку відносних величин.
26. Приклади використання абсолютних і відносних показників в економіко-статистичному аналізі.

# Розділ 7 Середні величини

## 7.1. Суть і значення середніх величин

Для зведеної кількісної характеристики багатьох явищ і процесів суспільного життя статистика широко використовує такий розповсюджений узагальнюючий показник як середня величина (середня врожайність, середня продуктивність праці, середня заробітна плата і т.п.). Вона дає загальну характеристику однорідних елементів масових явищ, які мають різне кількісне значення (варіацію) в залежності від конкретних умов. Варіація любої ознаки формується під впливом багатьох загальних і індивідуальних умов. Середня абстрагується від індивідуальних відмінностей і характеризує те загальне, типове, що притаманне всій сукупності в цілому.

В середній погашаються випадкові відхилення індивідуальних значень і відображаються ті загальні умови, під впливом яких формувалась вся сукупність. Це одна з основних умов наукового застосування середніх, де проявляється в самому загальному вигляді дія закону великих чисел.

*Середня величина* – це узагальнюючий показник, який характеризує однорідну сукупність явищ за якою-небудь кількісною варіаційною ознакою в даних умовах місця і часу.

Не можна змішувати середні величини з відносними величинами інтенсивності. Як ми вже відмічали, середня завжди узагальнює кількісні ознаки притаманні всім без винятку елементам сукупності. Відносним величинам інтенсивності властивий прояв зовсім іншого характеру. Наприклад, в розрахунку на душу населення щорічно споживається п'ять літрів абсолютного алкоголю. Однак, це не означає що всі без винятку (діти, люди похилого віку, хворі) споживають спиртні напої. В даному випадку споживання спиртного як ознака не є властивим для всього населення. Це і є відносна величина інтенсивності.

Середні величини мають велике значення в любих економіко-статистичних дослідженнях. Вони використовуються для узагальнюючої характеристики рівня розвитку суспільних явищ і процесів. Тільки за допомогою середньої можна охарактеризувати сукупність за кількісною варіаційною ознакою. Наприклад, заробітну плату, яку отримують робітники двох підприємств, можна порівняти за допомогою такого узагальнюючого показника як середня заробітна плата.

Середні величини використовують для порівняння показників двох і більше об'єктів (порівняння урожайності окремих культур по господарствах області, порівняння цін на деякі товари на ринках певного регіону і т.п.).

Середніми величинами користуються для характеристики зміни рівнів явищ в часі. Для дослідження динаміки середньої урожайності, середньої заробітної плати, середньої продуктивності праці і т.п.

До середніх звертаються при вивченні взаємозв'язків між явищами та їх ознаками. Залежність випуску продукції від вартості основних виробничих фондів по групах підприємств можна виявити лише на основі зміни їх середніх.

Середні величини застосовують для проведення факторного аналізу явищ з метою виявлення невикористаних резервів. Так, наприклад, для виявлення резервів росту об'єму відпрацьованого часу використовують трьохфакторну модель, яка складається з добутку трьох середніх: середньоспискової чисельності робітників, середнього фактичного числа днів роботи одним робітником і середньої тривалості робочого дня.

Велике значення мають середні величини в плануванні і прогнозуванні завдань для народного господарства в цілому і окремих його галузей, де в багатьох випадках вдаються до системи середніх величин.

Математичні прийоми, які використовуються в різних галузях економіки, безпосередньо зв'язані з обчисленням середніх величин. Варіація ознак в рядах розподілу визначається за допомогою відхилень від середньої. Всі формули агрегатних і середніх індексів характеризують зміну середнього рівня економічних показників для сукупності елементів, різнорідних за своєю натуральною формою. Алгебраїчне і графічне зображення загальної тенденції в рядах динаміки здійснюється за допомогою знаходження середньої лінії розвитку цього ряду.

Багатогранність суспільних явищ обумовлює виняткову важливість застосування середніх величин в економіко-статистичних дослідженнях. Вони є активним засобом управління, планування і прогнозування народного господарства держави.

## 7.2. Середня арифметична і її властивості

Найбільш поширеним видом середніх величин в статистиці є середня арифметична. Вона застосовується у формі простої середньої і зваженої середньої.

*Середня арифметична проста* застосовується в тих випадках, коли всі варіанти зустрічаються один раз, або мають однакові частоти в досліджуваній сукупності. Її отримують шляхом додавання окремих варіантів і діленням суми на число доданків.

Формула середньої арифметичної простої має вигляд:

$$\overline{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n} = \frac{\sum x}{n},$$

де $\overline{x}$ – середня величина ознак;

$x_1, x_2, x_3, \ldots, x_n$ – окремі варіанти ознаки;

$\Sigma$ – (велика грецька літера “сігма”) – знак суми;

n – кількість варіантів.

Розглянемо приклад. Маємо такі дані про добове видобування вугілля на шахті за першу декаду червня (табл. 7.1):

Таблиця 7.1

| Числа місяця | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Добове видобування вугілля, тис. т | 5,4 | 5,3 | 5,6 | 5,5 | 5,7 | 5,9 | 6,0 | 5,8 | 6,2 | 6,1 |

В нашому прикладі дані про добове видобування вугілля зустрічаються однакове число раз, а тому середньодобове значення цієї ознаки потрібно вираховувати за формулою середньої арифметичної простої:

$$\overline{x} = \frac{\sum x}{n} = \frac{5{,}4 + 5{,}3 + 5{,}6 + 5{,}5 + 5{,}7 + 5{,}9 + 6{,}0 + 5{,}8 + 6{,}2 + 6{,}1}{10} =$$

$$= \frac{57{,}5}{10} = 5{,}75 \text{ тис. т.}$$

Якщо в сукупності варіанти зустрічаються неоднакову кількість раз, то їх об’єднують в групи і, таким чином, переходять від середньої арифметичної простої до зваженої.

*Середня арифметична зважена* обчислюється як частка від ділення суми добутків варіантів і їх частот на суму частот. Вона визначається за формулою:

$$\overline{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \ldots + x_n f_n}{f_1 + f_2 + f_3 + \ldots + f_n} = \frac{\sum xf}{\sum f},$$

де f – частоти (ваги), які показують скільки раз зустрічаються значення ознаки в сукупності.

Отже, для обчислення середньої арифметичної зваженої потрібно: а) кожний варіант перемножити на його частоту; б) знайти суму їх добутків; в) суму добутків поділити на суму ваг.

Наведемо приклад. Маємо дані про заробітну плату робітників і число робітників, які отримують дану заробітну плату.

Таблиця 7.2

| Табельний номер робітника | Заробітна плата одного робітника, грн. (x) | Число робітників, чол. (f) |
|---|---|---|
| 1 | 150 | 10 |
| 2 | 180 | 17 |
| 3 | 200 | 40 |
| 4 | 210 | 18 |
| 5 | 250 | 15 |
| Разом | x | 100 |

Середня заробітна плата робітників за цими даними становитиме:

$$\overline{x} = \frac{\sum xf}{\sum f} = \frac{150 \cdot 10 + 180 \cdot 17 + 200 \cdot 40 + 210 \cdot 18 + 250 \cdot 15}{10 + 17 + 40 + 18 + 15} =$$

$$= \frac{20090}{100} = 200{,}9 \text{ грн.}$$

Часто доводиться обчислювати середню із варіантів, які самі по собі є середніми величинами, тобто середню з середніх. Така середня визначається так само, як і середні з початкових значень ознаки. В даному випадку середні, які служать основою для обчислення загальної середньої, беруться в якості варіантів.

Таблиця 7.3

Середня урожайність озимої пшениці по групі господарств району

| Номер господарства | Середня урожайність озимої пшениці, ц/га (x) | Посівна площа озимої пшениці, га (f) |
|---|---|---|
| 1 | 28,5 | 300 |
| 2 | 30,0 | 250 |
| 3 | 34,9 | 800 |
| 4 | 40,0 | 350 |
| 5 | 42,6 | 300 |
| Разом | x | 2000 |

Звідси, середня врожайність озимої пшениці по всіх господарствах складе:

$$\overline{x} = \frac{\sum xf}{\sum f} = \frac{28{,}5 \cdot 300 + 30{,}0 \cdot 250 + 34{,}9 \cdot 800 + 40{,}0 \cdot 350 + 42{,}6 \cdot 300}{300 + 250 + 800 + 350 + 300} =$$

$$= \frac{70750}{2000} = 35{,}4 \text{ ц/га.}$$

Іноді середні величини потрібно обчислити не з конкретних значень варіантів досліджуваної ознаки, а із значень величин, виражених у вигляді інтервалів. В таких випадках потрібно для кожного інтервалу знайти його середину за простою середньою між верхньою і нижньою межею кожного інтервалу і після цього проводити обчислення за формулою середньої арифметичної зваженої.

Покажемо обчислення середньої на основі інтервального ряду (табл. 7.4).

Таблиця 7.4

Розподіл виробів за їх вагою

| Вага виробів, г (x) | Число виробів, шт. (f) | Середина інтервалу (x) | Вага всіх деталей, г (x · f) |
|---|---|---|---|
| до 100 | 5 | 97,5 | 487,5 |
| 100-105 | 19 | 102,5 | 1947,5 |
| 105-110 | 52 | 107,5 | 5590,0 |
| 110-115 | 18 | 112,5 | 2025,0 |
| 115 і більше | 6 | 117,5 | 705,0 |
| Разом | 100 | х | 10755,0 |

В даному прикладі варіанти наведені у вигляді інтервалів (“від – до”), при цьому верхня і нижня межа інтервального ряду відкриті. Для знаходження нижньої межі першого інтервалу спочатку визначають розмір наступного інтервалу: 105 – 100 = 5. Потім, від верхньої межі першого інтервалу віднімаємо цей розмір 100 – 5 = 95, таким чином будемо мати нижню межу першого інтервалу. Для знаходження верхньої межі останнього відкритого інтервалу беруть розмір попереднього інтервалу, який додають до нижньої межі шуканого інтервалу: 115 + 5 = 120. Далі перетворюємо інтервальний ряд в конкретні числа за формулою середньої арифметичної простої: для першого інтервалу 95 + 100 : 2 = 97,5 г; для другого 100 + 105 : 2 = 102,5 г і т.д. (див. передостанню колонку).

Перемноживши вагу одного виробу на їх кількість, будемо мати вагу всіх виробів. Середню вагу одного виробу знайдемо шляхом ділення ваги всіх виробів на кількість виробів.

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{10755}{100} = 107{,}55 \text{ г.}$$

Середня арифметична має деякі математичні властивості, що мають практичне значення для спрощеного обчислення середньої за даними варіаційного ряду.

Найважливіші з них такі:

1. Якщо всі варіанти збільшити або зменшити на одне й теж число (А), то й середня арифметична збільшиться (зменшиться) на те ж число (А):

$$\frac{\sum (x - A)\cdot f}{\sum f} = \bar{x} - A \text{, звідки } \bar{x} = \frac{\sum (x - A)\cdot f}{\sum f} + A \,.$$

2. Якщо всі варіанти збільшити або зменшити в одне й те ж число (і) раз, то й середня арифметична відповідно збільшиться (зменшиться) в (і) раз:

$$\frac{\sum \frac{x}{i} f}{\sum f} = \frac{\bar{x}}{i} \text{, звідки } \bar{x} = \frac{\sum \frac{x}{i} f}{\sum f} \cdot i \,.$$

3. Якщо всі частоти (ваги) поділити або помножити на яке-небудь число (z), то середня арифметична від цього не зміниться:

$$\bar{x} = \frac{\sum x \cdot \frac{f}{z}}{\sum \frac{f}{z}} = \frac{\sum xf}{\sum f} \,.$$

4. Добуток середньої на суму частот завжди дорівнює сумі добутків варіантів на частоти:

$$\bar{x} \cdot \sum f = \sum x f \,.$$

5. Сума відхилень варіантів від їх середньої завжди дорівнює нулю:

для простої середньої $\sum (x - \bar{x}) = \sum x - n\bar{x} = 0$;

для зваженої середньої $\sum (x - \bar{x}) f = \sum xf - \bar{x} \sum f = 0$.

Використання першої і другої властивостей середньої арифметичної дозволяє значно спростити її обчислення. Цей метод в статистиці називається *метод моментів*, або *метод відліку від* *умовного нуля*. Розглянемо спрощений спосіб обчислення середньої арифметичної методом моментів за даними попереднього прикладу (табл. 7.5).

Таблиця 7.5

| Вага виробів, г<br>x | Число виробів, шт.<br>f | Середина інтервалу<br>x | Скорочені варіанти | | Зважені скорочені варіанти<br>$\left(\frac{x-A}{i}\right)f$ |
|---|---|---|---|---|---|
| | | | $x - A$<br>$A = 107{,}5$ | $\frac{x-A}{i}$<br>$i = 5$ | |
| до 100 | 5 | 97,5 | - 10 | - 2 | - 10 |
| 100-105 | 19 | 102,5 | - 5 | - 1 | - 19 |
| 105-110 | 52 | 107,5 | 0 | 0 | 0 |
| 110-115 | 18 | 112,5 | 1 | 1 | 18 |
| 115 і більше | 6 | 117,5 | 2 | 2 | 12 |
| Разом | 100 | x | x | x | + 1 |

Формула для знаходження середньої арифметичної способом моментів має вигляд:

$$\bar{x} = i \cdot m_1 + A\,,$$

де $m_1 = \dfrac{\sum\left(\dfrac{x-A}{i}\right)f}{\sum f}$ – момент першого порядку;

x – варіант;

f – частота;

A – умовно взяте число;

i – розмір інтервалу.

Визначаємо момент першого порядку:

$$m_1 = \frac{\sum\left(\dfrac{x-A}{i}\right)f}{\sum f} = \frac{1}{100} = 0{,}01\,.$$

Підставляємо значення в формулу:

$$\bar{x} = i \cdot m_1 + A = 5 \cdot 0{,}01 + 107{,}5 = 107{,}55 \text{ г.}$$

Отже, ми отримали той самий результат, що й при обчисленні за звичайною формулою середньої арифметичної зваженої.

Спосіб моментів використовують в тих випадках, коли вихідні дані наведені у вигляді інтервального ряду з рівними інтервалами.

Третю властивість середньої арифметичної використовують тоді, коли доводиться мати справу з великими частотами. В цьому випадку частоти скорочують і подальші розрахунки проводять з малими числами.

На цій же властивості середньої арифметичної ґрунтується перехід від абсолютних значень частот до відносних величин, які визначаються як процентне відношення кожної групи частот до їх загального підсумку.

Розглянемо застосування цієї властивості на прикладі. Нехай маємо розподіл п’яти груп господарств за врожайністю гречки (табл. 7.6).

Таблиця 7.6

| Групи господарств | Урожайність гречки, ц/га $x$ | Посівна площа, га $f$ | Скорочені ваги $f' = \frac{f}{10000}$ | Варіанти, зважені на скорочені ваги $xf'$ | Питомі ваги $P = \frac{f}{\sum f} \cdot 100$ | Варіанти, зважені на питомі ваги $x \cdot P$ |
|---|---|---|---|---|---|---|
| I | 10 | 10000 | 1 | 10 | 6,67 | 66,70 |
| II | 12 | 30000 | 3 | 36 | 20,00 | 240,00 |
| III | 14 | 50000 | 5 | 70 | 33,33 | 466,62 |
| IV | 16 | 40000 | 4 | 64 | 26,67 | 426,72 |
| V | 18 | 20000 | 2 | 36 | 13,33 | 239,94 |
| Разом | x | 150000 | 15 | 216 | 100 | 1439,98 |

$$f_1' = \frac{f_1}{10000} = \frac{10000}{10000} = 1; \; f_2' = \frac{f_2}{10000} = \frac{30000}{10000} = 3; \text{ і т.д.}$$

Середня врожайність $\bar{x} = \frac{\sum xf'}{\sum f'} = \frac{216}{15} = 14{,}4$ ц/га.

$$P_1 = \frac{f_1}{\sum f} \cdot 100 = \frac{10000}{150000} \cdot 100 = 6{,}67 \;\%;$$

$$P_2 = \frac{f_2}{\sum f} \cdot 100 = \frac{30000}{150000} \cdot 100 = 20{,}00; \text{ і т.д.}$$

Середня врожайність $\bar{x} = \frac{\sum xp}{100} = \frac{1439{,}98}{100} = 14{,}4$ ц/га.

### 7.3. Середня гармонічна

В статистичній практиці часто зустрічаються випадки, коли середню потрібно обчислювати за формулою *середньої гармонічної*. Це відбувається тоді, коли підсумовуванню підлягають не самі варіанти, а обернені їм числа. В цьому випадку, для знаходження середнього значення варіаційної ознаки, застосовують формулу *середньої гармонічної простої*, яка має вигляд:

$$\bar{x} = \frac{n}{\sum \frac{1}{x}},$$

де $n$ – число індивідуальних значень ознак;

$\sum \frac{1}{x}$ – сума обернених значень ознак.

Техніку обчислення середньої гармонічної простої покажемо на такому прикладі.

В ремонтній майстерні підприємства працювало п'ять робітників по 8 годин кожний. Перший витрачав на обробку однієї деталі 10 хв., другий – 12 хв., третій 13 хв., четвертий – 9 хв. і п'ятий – 15 хв. Потрібно визначити середні затрати часу на обробку однієї деталі.

Відомо, що між продуктивністю праці робітника і затратами часу на обробку однієї деталі існує обернена залежність. Тобто, чим вища продуктивність праці робітника, тим менше він витрачає часу на обробку однієї деталі. Так як кожен робітник працював однакову кількість часу, то середні затрати часу потрібно обчислювати за формулою середньої гармонічної простої.

$$\bar{x} = \frac{n}{\sum \frac{1}{x}} = \frac{5}{\frac{1}{10} + \frac{1}{12} + \frac{1}{13} + \frac{1}{9} + \frac{1}{15}} = \frac{5}{0{,}438} = 11{,}4 \text{ хв.}$$

Перевіримо правильність зробленого нами розрахунку. Перший робітник обробив за зміну 48 (480 : 10) деталей, другий – 40 (480 : 12) деталей, третій – 37 (480 : 13) деталей; четвертий – 53 (480 : 9) деталі і п'ятий – 32 (480 : 15) деталі.

Кількість деталей оброблених за зміну одним робітником буде становити в середньому:

$$\bar{x} = \frac{\sum x}{n} = \frac{48 + 40 + 37 + 53 + 32}{5} = \frac{270}{5} = 42 \text{ деталі.}$$

Кількість часу затраченого в середньому на обробіток однієї деталі робітником за зміну буде дорівнювати:

$$\bar{x} = \frac{8 \cdot 60}{42} = \frac{480}{42} = 11{,}4 \text{ хв.}$$

Отже, *середня гармонічна* являє собою обернену величину середньої арифметичної, обчисленої з обернених значень ознак.

*Середню гармонічну зважену* застосовують в тих випадках, коли є дані про індивідуальні значення ознаки в загальній сукупності і загальний обсяг сукупності, але в готовому виді немає частот.

$$\bar{x} = \frac{\sum W}{\sum \frac{W}{x}},$$

де $\sum \frac{W}{x}$ – сума добутку обернених ознак і частот, тобто $x \cdot f = W$, звідси $f = \frac{W}{x}$.

Розглянемо приклад. Маємо дані про середню заробітну плату робітників заводу в розрізі цехів і фонд заробітної плати (табл. 7.7).

Таблиця 7.7

| Номер цеху | Середня заробітна плата одного робітника, грн. (x) | Фонд заробітної плати, грн. (W) |
|---|---|---|
| 1 | 290 | 17400 |
| 2 | 320 | 16000 |
| 3 | 350 | 14000 |

Підставивши у формулу середньої гармонічної зваженої дані з нашого прикладу, отримаємо середню заробітну плату одного робітника по заводу в цілому:

$$\bar{x} = \frac{\sum W}{\sum \frac{W}{x}} = \frac{17400 + 16000 + 14000}{\frac{17400}{290} + \frac{16000}{320} + \frac{14000}{350}} = \frac{47400}{150} = 316 \text{ грн.}$$

Якщо сукупність складається з двох частин ($n_1$ і $n_2$) і є дані про їх середні гармонічні $(\bar{x}_1 \text{ і } \bar{x}_2)$, то загальну середню гармонічну для всієї сукупності визначають як зважену гармонічну середню з часткових гармонічних середніх.

$$\bar{x} = \frac{n_1 + n_2}{\frac{n_1}{\bar{x}_1} + \frac{n_2}{\bar{x}_2}}.$$

Звернемось до прикладу. Нехай маємо дані про чисельність робітників двох цехів: $n_1$ = 150 чол., $n_2$ = 140 чол. Середня заробітна плата робітників першого цеху $\bar{x}_1 = 300$ грн., другого $\bar{x}_2 = 350$ – грн. Потрібно визначити середню заробітну плату по обох цехах разом:

$$\bar{x} = \frac{n_1 + n_2}{\frac{n_1}{\bar{x}_1} + \frac{n_2}{\bar{x}_2}} = \frac{150 + 140}{\frac{150}{300} + \frac{140}{350}} = \frac{290}{0,9} = 322,22 \text{ грн.}$$

Інколи, для характеристики деяких явищ і процесів використовують формулу *середньої антигармонічної*, яка має вигляд:

$$\bar{x} = \frac{\sum x_i^2}{\sum x_i} - \text{ проста;}$$

$$\bar{x} = \frac{\sum x_i^2 f_i}{\sum x_i f_i} - \text{зважена.}$$

Наведемо три приклади її використання.

Приклад 1. Маємо n тренерів, і кожний з них підготував $x_i$ спортсменів. Якщо кожний спортсмен підготовить стільки ж нових спортсменів, скільки підготував його тренер, то середнє співвідношення спортсменів і тренерів (продуктивність роботи тренера) можна відобразити середньою антигармонічною.

Нехай п’ять тренерів підготували відповідно 10, 16, 20, 17 і 13 спортсменів.

Тоді $\bar{x} = \dfrac{\sum x_i^2}{\sum x_i} = \dfrac{10^2 + 16^2 + 20^2 + 17^2 + 13^2}{10 + 16 + 20 + 17 + 13} = \dfrac{1214}{76} = 16$ чол.

Тобто, продуктивність праці тренера в справі підготовки спортсменів складає в середньому 16 чоловік.

Приклад 2. Маємо п’ять галузей народного господарства і дані про ефективність грошових вкладень в кожну з них: $x_1 = 1,1$; $x_2 = 1,2$; $x_3 = 1,3$; $x_4 = 1,2$ і $x_5 = 1,1$. Одна вкладена гривня в поточному році дасть дохід $x_i$ гривень в наступному році. Якщо $x_i$ незмінні, то ефективність вкладень виразиться формулою середньої антигармонічної.

$$\bar{x} = \frac{\sum x_i^2}{\sum x_i} = \frac{1,1^2 + 1,2^2 + 1,3^2 + 1,2^2 + 1,1^2}{1,1 + 1,2 + 1,3 + 1,2 + 1,1} = \frac{6,99}{5,9} = 1,18 \text{ грн.}$$

Приклад 3. Розподіл обстежених сто жінок за народженням дітей представлений в табл. 8.7.

Таблиця 7.8

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Число жінок ($f_i$) | 5 | 8 | 22 | 46 | 14 | 4 | 1 |
| Число народжених дітей ($x_i$) | 6 | 4 | 1 | 2 | 3 | 5 | 8 |

Припустимо, що кожна дочка народила стільки ж дітей, скільки мала її мати, при цьому число народжених дівчаток і хлопчиків однакове. Потрібно встановити співвідношення чисельності між поколіннями.

$$\bar{x} = \frac{\sum x_i^2 f_i}{\sum x_i f_i} = \frac{6^2 \cdot 5 + 4^2 \cdot 8 + 1^2 \cdot 22 + 2^2 \cdot 46 + 3^2 \cdot 14 + 5^2 \cdot 4 + 8^2 \cdot 1}{6 \cdot 5 + 4 \cdot 8 + 1 \cdot 22 + 2 \cdot 46 + 3 \cdot 14 + 5 \cdot 4 + 8 \cdot 1} =$$

$$= \frac{804}{246} = 3{,}27 \text{ чол.}$$

Отже, на кожну людину першого покоління буде припадати в середньому 3,27 чол. наступного покоління.

### 7.4. Інші види середніх.

Особливе місце в статистиці відводиться різним модифікаціям *степеневої середньої*. Загальна формула степеневої середньої має такий вигляд:

$$\bar{x} = \sqrt[k]{\frac{\sum x^k}{n}} \text{ – степенева проста;}$$

$$\bar{x} = \sqrt[k]{\frac{\sum x^k f}{\sum f}} \text{ – степенева зважена,}$$

де $\bar{x}$ – степенева середня;
$x$ – варіанти;
$n$ – число варіантів;
$f$ – частоти;
$k$ – показник степеня, який визначає вид середньої.

Степінь степеневої середньої визначає вид середньої величини. Так, при $k = 1$ отримаємо середню арифметичну: $\bar{x}_a = \frac{\sum x}{n}$;

при $k = 2$ – середню квадратичну: $\bar{x}_q = \sqrt{\frac{\sum x^2}{n}}$;

при $k = 3$ – середню кубічну: $\bar{x}_к = \sqrt[3]{\frac{\sum x^3}{n}}$;

при $k = – 1$ – середню гармонічну: $\bar{x}_h = \frac{n}{\sum \frac{1}{x}}$;

при k = 0 – середню геометричну: $\bar{x}_g = \sqrt[n]{Dx}$,

де D – знак множення.

Чим вищий степінь у формулі степеневої середньої, тим буде більшою величина середньої для однієї і тієї ж системи індивідуальних значень ознак. Таким чином, наведені вище середні розмістяться в наступному порядку:

$$\bar{x}_h < \bar{x}_g < \bar{x}_a < \bar{x}_q < \bar{x}_k \text{ і т.д.}$$

Дане співвідношення називається в статистиці правилом *мажорантності* (від французького слова majeur – більший). Це правило буде достовірним для середніх, обчислених за різними формулами, але тільки з величин однієї і тієї ж сукупності. Різниця між середніми буде тим значніша, чим більша варіація осереднюваних величин.

В статистиці правильну характеристику сукупності за варіаційною ознакою в кожному окремому випадку дає тільки вірно вибраний відповідний вид середньої.

Із всіх видів степеневих середніх в статистиці найбільш часто використовують середню арифметичну, дещо рідше – середню гармонічну. Середню квадратичну використовують тільки при обчисленні показників варіації, а середню геометричну – при обчисленні середніх темпів динаміки. Більш детально останні дві середні будуть розглянуті в наступних темах.

В окремих випадках, для узагальнюючої характеристики двох і більше ознак, часто з різними одиницями виміру, використовують *багатомірну середню*. Її розрахунок базується на відношенні індивідуальних значень ознак $x_{ij}$ до середньої по сукупності в цілому:

$$P_{ij} = \frac{x_{ij}}{\bar{x}_i}, \text{ де } \bar{x}_i = \frac{\sum^{n} x_j}{n}.$$

Багатомірну середню обчислюють як середню арифметичну просту з відношень $P_{ij}$ за формулою:

$$\bar{P}_{ij} = \frac{\sum^{m} P_{ij}}{m},$$

де m – число ознак.

Вона характеризує місце j-го елементу сукупності в багатомірному полі дослідження. Методику обчислення багатомірних середніх покажемо на прикладі. Нехай забезпеченість трьох областей Галичини деякими основними товарами культурно-побутового призначенні характеризується наступними даними (табл. 7.9, дані умовні):

Таблиця 7.9

| Область | Забезпеченість в розрахунку на 100 сімей, шт. | | | Співвідношення $P_{ij} = \frac{x_{ij}}{\bar{x}_i}$ | | | Багатомірні середні $\bar{P}_{ij} = \frac{\sum^{m} P_{ij}}{m}$ |
|---|---|---|---|---|---|---|---|
| | Телевізорами кольорового зображення ($x_1$) | Холодильниками побутовими (x2) | Пральними машинами (x3) | $P_{1j}$ | $P_{2j}$ | $P_{3j}$ | |
| Ів.-Франківська | 92 | 85 | 73 | 0,979 | 1,000 | 1,000 | 0,993 |
| Львівська | 98 | 87 | 72 | 1,042 | 1,024 | 0,986 | 1,017 |
| Тернопільська | 96 | 84 | 74 | 1,021 | 0,988 | 1,014 | 1,008 |
| Україна в цілому | 94 | 85 | 73 | 1,000 | 1,000 | 1,000 | 1,000 |

Так, $P_{11} = \frac{x_{11}}{\bar{x}_1} = \frac{92}{94} = 0{,}979$; $P_{22} = \frac{x_{22}}{\bar{x}_2} = \frac{87}{85} = 1{,}024$;

$P_{33} = \frac{x_{33}}{\bar{x}_3} = \frac{74}{73} = 1{,}014$ і т.д. Базою зіставлення в даному випадку ми взяли середній рівень забезпеченості окремими товарами культурно-побутового призначення по країні в цілому.

Багатомірна середня із відносних величин для Івано-Франківської області становить:

$$\bar{P}_1 = \frac{\sum_{m} P_{ij}}{m} = \frac{0{,}979 + 1{,}000 + 1{,}000}{3} = 0{,}993;$$

Львівської – $\bar{P}_2 = 1{,}017$ і Тернопільської – $\bar{P}_3 = 1{,}008$.

У Львівській і Тернопільській областях забезпеченість сімей даними товарами культурно-побутового призначення вища за середню по Україні (значення багатомірних середніх більше за одиницю), а в Івано-Франківській області дещо нижча за середньодержавну (багатомірна середня менша за одиницю).

В деяких випадках для поглибленого економічного аналізу, обчислюють *прогресивну середню*. На відміну від загальної середньої, де в розрахунок беруть всі без винятку варіанти, прогресивна середня дає зведену характеристику не взагалі, а тільки кращих індивідуальних показників.

При обчисленні середньої прогресивної можливі два варіанти. Перший – коли кращими будуть вищі показники від загальної середньої (вища врожайність, вища продуктивність праці, більша заробітна плата і т.д.). В цьому випадку: а) з усіх наявних варіантів обчислюють загальну середню; б) відбирають всі ті варіанти, які перевищують загальну середню; в) з цих кращих варіантів обчислюють прогресивну середню.

$$\bar{x}' = \frac{\sum x'}{n},$$

де $x'$ – кращі (вищі) варіанти;

$n$ – число кращих варіантів.

Приклад 1. Маємо дані про заробітну плату десяти робітників бригади.

Таблиця 7.10

| Табельний номер робітника | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Заробітна плата, грн. | 150 | 262 | 173 | 280 | 200 | 240 | 185 | 300 | 192 | 168 |

Обчислимо загальну середню:

$$\bar{x} = \frac{\sum x}{n} = \frac{150+262+173+280+200+240+185+300+192+168}{10} =$$

$$= \frac{2150}{10} = 215 \text{ грн.}$$

Відберемо вищі варіанти за 215 грн. і визначимо прогресивну середню:

$$\bar{x}' = \frac{\sum x'}{n} = \frac{262+280+240+300}{4} = 270{,}5 \text{ грн.}$$

За другим варіантом кращими визнають показники нижчі загальної середньої (нижча собівартість одиниці продукції, менші затрати часу на виготовлення одиниці продукції, менша фондоємкість, матеріалоємність, енергоємність і т.д.). Тут також спочатку: а) обчислюють загальну середню; б) відбирають кращі (менші) варіанти; в) з кращих (менших) варіантів обчислюють прогресивну середню:

$$\bar{x}'' = \frac{\sum x''}{n},$$

де $x''$ – кращі (менші) варіанти;

$n$ – число цих варіантів.

Приклад 2. Нехай маємо дані про собівартість одиниці однойменної продукції, яка виробляється на десятьох підприємствах області (табл. 7.11, дані умовні):

Таблиця 7.11

| Номер підприємства | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Собівартість, грн. | 2,9 | 3,1 | 3,0 | 2,8 | 2,7 | 3,2 | 2,6 | 3,1 | 3,0 | 2,8 |

Середня собівартість одиниці продукції по десяти заводах буде дорівнювати:

$$\overline{x} = \frac{\sum x}{n} = \frac{2{,}9 + 3{,}1 + 3{,}0 + 2{,}8 + 2{,}7 + 3{,}2 + 2{,}6 + 3{,}1 + 3{,}0 + 2{,}8}{10} =$$

$$= \frac{29{,}2}{10} = 2{,}92 \text{ грн.}$$

Далі відбираємо кращі (менші) варіанти і обчислюємо прогресивну середню:

$$\overline{x}'' = \frac{\sum x''}{n} = \frac{2{,}9 + 2{,}8 + 2{,}7 + 2{,}6 + 2{,}8}{5} = \frac{13{,}8}{5} = 2{,}76 \text{ грн.}$$

Часто виникає необхідність в обчисленні середнього рівня варіантів моментного ряду з рівними інтервалами, розміщених в хронологічному порядку. З цією метою використовують формулу *середньої хронологічної* виду:

$$\overline{x} = \frac{\frac{1}{2}x_1 + x_2 + x_3 + \ldots + x_{n-1} + \frac{1}{2}x_n}{n-1},$$

де $x_1$, $x_2$, $x_3$, ..., $x_n$ – варіанти ряду;
n – кількість варіантів.

Покажемо застосування цієї середньої на прикладі. В одному відділенні Національного банку заборгованість по всіх планових позиках наступна (табл. 7.12, дані умовні):

Таблиця 7.12

| Дані на початок місяця | $^1/_{01}$ | $^1/_{02}$ | $^1/_{03}$ | $^1/_{04}$ | $^1/_{05}$ | $^1/_{06}$ | $^1/_{07}$ | $^1/_{08}$ | $^1/_{09}$ | $^1/_{10}$ | $^1/_{11}$ | $^1/_{12}$ | $^1/_{01}$ наступного року |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Заборгованість по позиках, тис. грн. | 20 | 30 | 80 | 70 | 60 | 50 | 60 | 90 | 80 | 60 | 40 | 30 | 30 |

Потрібно визначити середній залишок по позиках за I, II, III, IV квартали і за рік в цілому.

$$\overline{x}_I = \frac{\frac{1}{2}x_1 + x_2 + x_3 + \frac{1}{2}x_4}{n-1} = \frac{\frac{20}{2} + 30 + 80 + \frac{70}{2}}{3} = 51{,}7 \text{ тис. грн.}$$

$$\bar{x}_{II} = \frac{\frac{1}{2}x_4 + x_5 + x_6 + \frac{1}{2}x_7}{n-1} = \frac{\frac{70}{2} + 60 + 50 + \frac{60}{2}}{3} = 58{,}3 \text{ тис. грн.}$$

$$\bar{x}_{III} = \frac{\frac{1}{2}x_7 + x_8 + x_9 + \frac{1}{2}x_{10}}{n-1} = \frac{\frac{60}{2} + 90 + 80 + \frac{60}{2}}{3} = 76{,}7 \text{ тис. грн.}$$

$$\bar{x}_{IV} = \frac{\frac{1}{2}x_{10} + x_{11} + x_{12} + \frac{1}{2}x_1}{n-1} = \frac{\frac{60}{2} + 40 + 30 + \frac{30}{2}}{3} = 38{,}3 \text{ тис. грн.}$$

$$\bar{x}_p = \frac{\frac{20}{2} + 30 + 80 + 70 + 60 + 50 + 60 + 90 + 80 + 60 + 40 + 30 + \frac{30}{2}}{12} =$$

$$= \frac{675}{12} = 56{,}25 \text{ тис. грн., або}$$

$$\bar{x}_p = \frac{\bar{x}_I + \bar{x}_{II} + \bar{x}_{III} + \bar{x}_{IV}}{4} = \frac{51{,}7 + 58{,}3 + 76{,}7 + 38{,}3}{4} = \frac{225}{4} = 56{,}25 \text{ тис. грн.}$$

Середня хронологічна буде тим точніша, чим менші проміжки часу взяті між моментами. Більш детально застосування цієї середньої розглянемо в темі “Ряди динаміки”.

### 7.5. Структурні середні

Поряд з розглянутими вище середніми, для статистичної характеристики варіантних рядів обчислюють так звані структурні (порядкові) середні, до яких відносять моду і медіану.

*Модою* називається величина ознаки (варіанта), яка найчастіше зустрічається в даній сукупності.

Знаходження моди в дискретному варіаційному ряду не представляє великої складності. Розглянемо табл. 7.13 з розподілом студентів за їх ростом.

Таблиця 7.13

| Ріст студентів, см | 165 | 167 | 170 | 173 | 176 | 178 | 180 | 182 | 185 | 187 | 189 | 191 | 195 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Число студентів, чол. | 8 | 13 | 24 | 30 | 38 | 47 | 55 | 31 | 16 | 9 | 7 | 3 | 2 |

Очевидно, в цьому прикладі модою буде студент, який має ріст 180 см., так як цьому значенню варіанти відповідає найбільше число студентів (55 чол.).

В розподілах, де всі варіанти зустрічаються однакове число раз, моди немає взагалі, або говорять, що всі варіанти однаково модальні.

Якщо два варіанти мають однакові частоти, тоді буде дві моди, тобто розподіл буде бімодальним.

*Медіана* – це варіант, який займає середнє положення в рангованому варіаційному ряді.

Щоб знайти медіану в дискретному варіаційному ряді, потрібно спочатку розташувати всі варіанти в зростаючому або спадаючому порядку. Потім визначити номер медіани, який вкаже на її розташування в рангованому ряді, за формулою:

$$\text{№ } M_e = \frac{n+1}{2},$$

де $M_e$ – медіана;

$n$ – число варіантів.

Розглянемо приклад. Маємо дані про розподіл дев'яти деталей за їх масою:

Таблиця 7.14

| Номер деталі | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Маса, г | 2,6 | 3,4 | 3,3 | 2,7 | 3,0 | 2,9 | 2,8 | 3,1 | 3,2 |

Перегрупуємо деталі за їх масою в зростаючому порядку.

Таблиця 7.15

| Номер деталі | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Маса, г | 2,6 | 2,7 | 2,8 | 2,9 | 3,0 | 3,1 | 3,2 | 3,3 | 3,4 |

Визначимо номер медіани:

$$\text{№ } M_e = \frac{n+1}{2} = \frac{9+1}{2} = 5.$$

Тобто, під п'ятим номером від початку або від кінця ряду маса деталі буде медіаною. В нашому прикладі медіана $M_e$ = 3,0 г.

Коли варіаційний ряд має парну кількість членів, тоді медіана буде розраховуватись як півсума двох варіантів, які займають середнє положення в рангованому ряді. Припустимо, що в нас є ще десята деталь з масою 3,5 г. Номер медіани буде рівним 5,5 $\left(\frac{10+1}{2}\right)$. В даному випадку медіана буде розташована між п'ятим і шостим порядковим номером деталей.

$$M_e = \frac{3{,}0+3{,}1}{2} = 3{,}05 \text{ г.}$$

Для знаходження номера медіани в згрупованому дискретному варіаційному ряді, потрібно до суми нагромаджених частот добавити одиницю і поділити на два. Обчислимо медіану за даними таблиці 7.13.

Таблиці 7.16

Розрахункова таблиця

| Ріст студента, см | 165 | 167 | 170 | 173 | 176 | 178 | 180 | 182 | 185 | 187 | 189 | 191 | 195 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Нагромаджені частоти (Σf) | 8 | 21 | 45 | 75 | 113 | 160 | 215 | 246 | 262 | 271 | 278 | 281 | 283 |

Знаходимо номер медіани:

$$\text{№ } M_e = \frac{\sum f + 1}{2} = \frac{283 + 1}{2} = 142 .$$

Так, в розподілі 283 студентів за їх ростом медіаною буде 142-й варіант, який ділить впорядкований варіаційний ряд навпіл. Отже, 142-й варіант відповідає шостому значенню варіаційної ознаки, і медіаною буде студент з ростом 178 см.

Моду і медіану із інтервальних рядів визначають розрахунковим шляхом за наступними інтерполяційними формулами:

$$M_o = x_{mo} + i\frac{f_2 - f_1}{(f_2 - f_1) + (f_2 - f_3)},$$

де $M_o$ – мода;

$x_{mo}$ – нижня межа модального інтервалу;

i – розмір інтервалу;

$f_1$ – передмодальна частота;

$f_2$ – модальна частота;

$f_3$ – післямодальна частота.

$$M_e = x_{me} + i\frac{\frac{\sum f}{2} - S_{me-1}}{f_{me}},$$

де $M_e$ – медіана;

$x_{me}$ – нижня межа медіанного інтервалу;

i – розмір медіанного інтервалу;

$\sum f$ – сума частот;

$\frac{\sum f}{2}$ – порядковий номер медіани;

$S_{me-1}$ – сума нагромаджених частот до медіанного інтервалу;

$f_{me}$ – частота медіанного інтервалу.

Покажемо обчислення моди і медіани для інтервального варіаційного ряду на прикладі, наведеному в табл. 7.17.

Таблиця 7.17

Розподіл 500 робітників за заробітною платою

| Заробітна плата, грн. | Число робітників, чол. | Нагромаджені частоти |
|---|---|---|
| 150-160 | 5 | 5 |
| 160-170 | 10 | 15 |
| 170-180 | 61 | 76 |
| 180-190 | 105 | 181 |
| 190-200 | 130 | 311 |
| 200-210 | 109 | 420 |
| 210-220 | 62 | 482 |
| 220-230 | 11 | 493 |
| 230-240 | 7 | 500 |
| Разом: | 500 | х |

В нашому прикладі мода знаходиться в інтервалі від 190 до 200 грн., тому, що йому відповідає найбільша частота (найбільше число робітників – 130 чол.). Цей інтервал називається модальним.

Підставляючи числові значення із нашого прикладу в формулу моди, отримаємо:

$$M_o = 190 + 10\frac{130-105}{(130-105)+(130-109)} = 195{,}43 \text{ грн.}$$

Цей показник означає, що найбільше було робітників із заробітною платою 195 грн. 43 коп.

Щоб визначити медіану інтервального варіаційного ряду, спочатку, за допомогою нагромаджених частот, потрібно знайти інтервал, що містить медіану. Медіанному інтервалу відповідає перша з нагромаджених частот, яка перевищує півсуму частот всього обсягу сукупності $\left(\frac{\sum f}{2} = 250\right)$. Отже, медіана знаходиться в інтервалі від 190 до 200 грн., що в даному разі співпав з модальним інтервалом.

Підставляючи в формулу медіани значення з нашого прикладу, отримаємо:

$$M_e = 190 + 10\frac{250-181}{130} = 195{,}31 \text{ грн.}$$

Це означає, що половина робітників отримує заробітну плату меншу 195, 31 грн.; а друга половина – більшу.

В доповнення до медіани, для характеристики структури варіаційного ряду, вираховують *квартилі* і *децилі*, які відповідно ділять ряд за сумою частот на чотири і десять рівних частин.

Перший або нижній квартиль відсікає чверть сукупності знизу, другий – рівний медіані, а третій або верхній – відсікає чверть сукупності зверху.

В інтервальному варіаційному ряду квартилі у середині, визначеного за нагромадженими частотами, інтервала обчислюються за формулами:

а) перший (нижній) квартиль:

$$Q_1 = x_{Q_1} + i\frac{\frac{\sum f}{4} - S_{Q_1-1}}{f_{Q_1}},$$

де $x_{Q_1}$ – нижня межа першого квартильного інтервалу;

$i$ – величина інтервалу;

$\frac{\sum f}{4}$ – порядковий номер першого квартиля;

$S_{Q_1-1}$ – нагромаджена частота перед першим квартильним інтервалом;

$f_{Q_1}$ – частота першого квартильного інтервалу;

б) третій (верхній) квартиль:

$$Q_3 = x_{Q_3} + i\frac{3\frac{\sum f}{4} - S_{Q_3-1}}{f_{Q_3}},$$

де $x_{Q_3}$ – нижня межа третього квартильного інтервалу;

$i$ – величина інтервалу;

$3\frac{\sum f}{4}$ – порядковий номер третього квартиля;

$S_{Q_3-1}$ – нагромаджена частота перед третім квартильним інтервалом;

$f_{Q_3}$ – частота третього квартильного інтервалу.

Обчислення цих показників у варіаційному ряду абсолютно аналогічне знаходженню медіани і починається з розрахунку порядкового номера відповідного варіанта, а потім за нагромадженими частотами визначається інтервал, в якому міститься шуканий варіант.

Так, за даними табл. 7.17 знаходимо, що перший квартиль, порядковий номер якого 12,5 $\left(\frac{1}{4}\sum f = \frac{500}{4}\right)$, міститься в інтервалі від 180 до 190 грн.

Звідси,

$$Q_1 = 180 + 10\frac{125 - 76}{105} = 184{,}67 \text{ грн.}$$

Це означає, що у одної четверті всіх робітників заробітна плата не перевищує 184,67 грн., а в трьох четвертях – вона рівна або перевищує 184,67 грн.

Аналогічно третій (верхній) квартиль, порядковий номер якого 375 $\left(\frac{3}{4}\sum f = \frac{3 \cdot 500}{4}\right)$, шукаємо в інтервалі від 200 до 210 грн.:

$$Q_3 = 200 + 10\frac{375 - 311}{109} = 205{,}87 \text{ грн.}$$

Отже, заробітна плата кожного четвертого робітника перевищує 205 грн. 87 коп.

Формули для децилів в інтервальному варіаційному ряду мають вигляд:

а) дециль перший:

$$D_1 = x_{o1} + i\frac{\frac{1}{10}\sum f - S_{D_1 - 1}}{f_{D_1}};$$

де $D_1$ – перший дециль;

$x_{o1}$ – нижня межа першого дециля;

i – розмір інтервалу;

$\frac{1}{10}\sum f$ – порядковий номер першого дециля;

$S_{D_1 - 1}$ – нагромаджена частота перед першим децильним інтервалом;

$f_{D_1}$ – частота першого децильного інтервалу;

б) дециль другий:

$$D_2 = x_{o2} + i\frac{\frac{2}{10}\sum f - S_{D_2 - 1}}{f_{D_2}};$$

в) дециль третій:

$$D_3 = x_{o3} + i\frac{\frac{3}{10}\sum f - S_{D_3-1}}{f_{D_3}};$$

і т.д.

Вираховуються децилі по тій ж схемі, що медіани і квартилі.

Так, $D_1 = 170 + 10\frac{50-15}{61} = 170 + 10 \cdot 0{,}574 = 175{,}74$ грн.;

$$D_2 = 180 + 10\frac{100-76}{105} = 180 + 10 \cdot 0{,}228 = 182{,}28 \text{ грн.};$$

$$D_3 = 180 + 10\frac{150-76}{105} = 180 + 10 \cdot 0{,}705 = 187{,}05 \text{ грн.};$$

. . . . . . . . . . . . . . . . . . . . . . . . . .

$$D_{10} = 230 + 10\frac{500-493}{7} = 230 + 10 \cdot 1 = 240{,}00 \text{ грн.}$$

Мода, медіана, квартилі і децилі відносяться до так званих порядкових статистик, під якими розуміють варіант, який займає певне порядкове місце в рангованому варіаційному ряду.

Їх використання в статистичному аналізі варіаційних рядів дозволяє більш глибоко дослідити і детальніше охарактеризувати сукупність, яка вивчається.

## 7.6. Основні правила застосування середніх в статистиці

В статистичних дослідженнях вірну характеристику сукупності за варіаційною ознакою в кожному окремому випадку дає тільки правильно визначений вид середньої.

Для знаходження цього виду середньої проф. А.Я. Боярський в 1929 р. запропонував критерій у вигляді визначальної властивості середньої: *"Середня тільки тоді буде вірною узагальнюючою характеристикою сукупності за варіаційною ознакою, коли при заміні всіх варіантів середньою загальною загальний обсяг варіаційної ознаки залишиться незмінним"*. Отже, в залежності від утворення загального обсягу варіаційної ознаки визначається вид вибраної середньої.

Так, середня арифметична застосовується тоді, коли загальний обсяг варіаційної ознаки утворюється як *сума* окремих варіантів; середня квадратична – коли такий обсяг утворюється як *сума квадратів* окремих варіантів; середня гармонічна – коли загальний обсяг утворюється як *сума обернених значень* окремих варіантів; середня геометрична – коли обсяг варіаційної ознаки утворюється як *добуток окремих варіантів*.

Наукове використання середніх в статистиці базується на певних умовах.

Головна умова наукового використання середньої полягає в тому, що середні характеристики повинні вираховуватись на основі масового узагальнення фактів. Тільки тоді вони відображають суть явища, на значення якого не впливають випадкові одиничні фактори. Ця умова пов'язує статистичні середні із законом великих чисел.

Іншою важливою умовою застосування середніх в статистиці є якісна однорідність всіх одиниць сукупності. Вона заключається в тому, що не можна обчислювати середню з неоднорідної сукупності, окремі елементи якої підпорядковані різним законам розвитку по відношенню до осереднюваної ознаки.

Середня величина тільки тоді відобразить типовий розмір ознаки та її загальні риси, якщо це загальне реально існує, всі елементи якого якісно однорідні і типові.

Якісна однорідність досліджуваного явища, його однотипність встановлюється на основі всестороннього теоретичного аналізу суті цього явища. Статистики виділяють якісно однорідні сукупності з допомогою групувань. Таким чином, застосування методу середніх в статистиці тісно і нерозривно зв'язане з методом групувань.

Зв'язок методу середніх і методу групувань полягає також в тому, що загальні середні, обчислені для якісно однорідних одиниць, в багатьох випадках повинні доповнюватися груповими середніми, так як загальні середні можуть не розкрити повністю закономірності досліджуваних процесів.

Загальні середні потрібно доповнювати груповими середніми в тих випадках, коли варіаційна ознака суттєво відрізняється по окремих групах і в порівнюваних групах існує різне співвідношення груп. В таких випадках розмір загальної середньої визначається через розміри групових середніх і структуру досліджуваної сукупності.

Особливого значення набуває доповнення загальної середньої груповими середніми при вивченні взаємозв'язку і взаємозалежності одних показників і ознак від інших.

Пізнавальні можливості методу середніх значно зростуть, якщо їх характеристики доповнити рядами розподілу (див. розділ 9).

При використанні середніх потрібно також пам'ятати, що середні величини не можуть і не повинні підміняти індивідуальні показники, а доповнюватись вивченням кращих і гірших одиниць сукупності.

Із всього вищенаведеного випливає, що середні в статистиці потрібно застосовувати на основі і в органічній єдності з методом групувань, який, в свою чергу, дозволяє відмежувати якісно однорідні сукупності для доповнення загальної середньої груповими середніми, а також рядами розподілу, в яких розкриваються передові досягнення або недоліки. Іншими словами, наукове застосування середніх в статистиці повинно виходити з діалектичного поєднання категорій загального і індивідуального, масового і одиничного.

### *Контрольні запитання*

1. Дайте визначення суті середньої.
2. Значення середніх величин в системі узагальнюючих статистичних показників.
3. Які ви знаєте основні положення теорії середньої?
4. Покажіть взаємозв’язок методу середніх і методу групувань.
5. Назвіть види середніх і способи їх обчислення.
6. Застосування середньої арифметичної простої і зваженої.
7. Математичні властивості середньої арифметичної.
8. Визначення середньої арифметичної способом моментів.
9. Застосування середньої гармонічної простої і зваженої.
10. Використання формули середньої антигармонічної простої і зваженої.
11. Інші види середніх.
12. Степенева середня проста і зважена та її модифікації.
13. Поняття про правило мажорантності.
14. В який випадках використовують багатомірну середню?
15. Як визначають прогресивну середню?
16. Коли застосовують середню хронологічну?
17. Вибір форми середньої в залежності від економічного змісту осереднюваних показників.
18. Поняття про структурні середні.
19. Визначення моди в дискретному і інтервальному ряду.
20. Визначення медіани в дискретному і інтервальному ряду.
21. Поняття про квартилі і децилі.
22. Основні правила застосування середніх в статистиці.

# Розділ 8. Показники варіації

## 8.1. Поняття про показники варіації і способи їх обчислення

Середні величини мають велике теоретичне і практичне значення, вони дають узагальнюючу характеристику сукупності за варіаційними ознаками, виражають типовий, для даних умов, рівень цих ознак. Проте, для характеристики досліджуваних явищ одних тільки середніх величин недостатньо, оскільки, при однакових значеннях середньої величини, різні сукупності можуть істотно відрізнятися одна від одної за характером варіації величини досліджуваної ознаки.

Середні величини не виражають індивідуальних особливостей досліджуваної сукупності, які породжують варіацію ознаки її окремих елементів, а тому, їх потрібно доповнювати показниками, що характеризують коливання значень ознаки в сукупності.

*Варіацією* в статистиці називаються коливання ознаки в одиниць сукупності, а показники, що характеризують ці коливання називаються показниками варіації. Вони показують як розміщуються навколо середньої окремі значення осереднюваної ознаки.

Розглянемо такий приклад. Маємо наступні дані про продуктивність праці робітників-відрядників в двох бригадах:

Таблиця 8.1

| Табельний номер робітника | Вироблено деталей за зміну, шт. | |
|---|---|---|
| | I бригада | II бригада |
| 1 | 10 | 120 |
| 2 | 20 | 110 |
| 3 | 100 | 100 |
| 4 | 180 | 90 |
| 5 | 190 | 80 |
| Разом: | 500 | 500 |

Середня продуктивність праці в двох бригадах буде однакова і становитиме:

$$\overline{x}_1 = \frac{\sum x_1}{n} = \frac{500}{5} = 100 \text{ шт.}; \quad \overline{x}_2 = \frac{\sum x_2}{n} = \frac{500}{5} = 100 \text{ шт.}$$

Однак, коливання продуктивності праці робітників-відрядників в першій бригаді значно більше, ніж у другій. Отже, друга бригада працює ритмічніше, ніж перша.

Під час проведення статистичного дослідження важливо визначити варіацію ознаки, що ховається за середніми величинами, які характеризують певну сукупність.

Показники варіації характеризують також типовість самої середньої величини.

Для вимірювання варіації у статистиці використовують такі показники як: розмах варіації, середнє лінійне відхилення, середній квадрат відхилення (дисперсія), середнє квадратичне відхилення і коефіцієнт варіації.

Найбільш простим із цих показників є розмах варіації (R).

*Розмах варіації* являє собою різницю між найбільшим і найменшим значенням ознаки і визначається за формулою:

$$R = x_{max} - x_{min},$$

де $x_{max}$ – максимальне значення ознаки;
$x_{min}$ – мінімальне значення ознаки.

В нашому прикладі розмах варіації:

для першої бригади $R_1 = 190 - 10 = 180$ шт.,

для другої бригади $R_2 = 120 - 80 = 40$ шт.

Розмах варіації простий для обчислення, але він відображає лише крайні значення ознаки і не дає уяви про ступінь варіації в середині сукупності. Тому, він використовується для наближеної оцінки варіації. Цей недолік перекривають інші показники варіації.

*Середнє лінійне відхилення* ($\bar{l}$) являє собою середню арифметичну з абсолютних значень відхилень окремих варіантів від їх середньої арифметичної.

Середнє лінійне відхилення – величина іменована і визначається за формулами:

а) середнє лінійне відхилення просте

$$\bar{l} = \frac{\sum |x - \bar{x}|}{n};$$

б) середнє лінійне відхилення зважене

$$\bar{l} = \frac{\sum |x - \bar{x}| f}{\sum f}.$$

Обчислимо середнє лінійне відхилення для нашого прикладу:

Таблиця 8.2

Розрахунок середнього лінійного відхилення

| Табельний номер робітника (n) | I бригада | | II бригада | |
|---|---|---|---|---|
| | вироблено деталей за зміну ($x_1$) | $\lvert x_1 - \bar{x}_1 \rvert$ | Вироблено деталей за зміну ($x_2$) | $\lvert x_2 - \bar{x}_2 \rvert$ |
| 1 | 10 | 90 | 120 | 20 |
| 2 | 20 | 80 | 110 | 10 |
| 3 | 100 | 0 | 100 | 0 |
| 4 | 180 | 80 | 90 | 10 |
| 5 | 190 | 90 | 80 | 20 |
| Разом: | $\sum x_1 = 500$ | $\sum \lvert x_1 - \bar{x}_1 \rvert = 340$ | $\sum x_2 = 500$ | $\sum \lvert x_2 - \bar{x}_2 \rvert = 60$ |

$$\bar{l}_1 = \frac{\sum \lvert x_1 - \bar{x}_1 \rvert}{n} = \frac{340}{5} = 68 \text{ шт.}; \quad \bar{l}_2 = \frac{\sum \lvert x_2 - \bar{x}_2 \rvert}{n} = \frac{60}{5} = 12 \text{ шт.}$$

Отже, кількість вироблених деталей за зміну окремими робітниками відрізняється від середньої в першій бригаді в середньому на 68 шт., а в другій бригаді – в середньому на 12 шт. Таким чином, середнє лінійне відхилення по виробництву деталей за зміну в першій бригаді у 5,7 разів більше, ніж у другій.

Середнє лінійне відхилення буде мінімальним, якщо відхилення розраховані від медіани:

$$\bar{l} = \frac{\sum \lvert x - M_e \rvert f}{\sum f} - \min .$$

Основними узагальнюючими показниками варіації в статистиці є дисперсія і середнє квадратичне відхилення.

*Середній квадрат відхилення*, або *дисперсія* ($\sigma^2$) визначається як середня арифметична з квадратів відхилень окремих варіантів від їх середньої.

В залежності від вихідних даних, дисперсію обчислюють за формулами:

а) дисперсія незважена (проста)

$$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n};$$

б) дисперсія зважена

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f},$$

де $\sigma$ – (грецька буква “сигма”), $\sigma^2$ – дисперсія.

*Середнє квадратичне відхилення* ($\sigma$), являє собою корінь квадратний з дисперсії.

Воно визначається за формулами:

а) середнє квадратичне відхилення просте

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n}};$$

б) середнє квадратичне відхилення зважене

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2 f}{\sum f}}.$$

Середнє квадратичне відхилення часто називають стандартним відхиленням. Воно як і середнє лінійне відхилення, є іменованою величиною. Але, в порівнянні з середнім лінійним відхиленням, середнє квадратичне відхилення має ряд переваг, що випливають з його математичних властивостей. Його використовують при оцінці тісноти зв’язку між явищами, при обчисленні помилок вибіркового спостереження, дослідженні рядів розподілу та ін.

Середнє квадратичне відхилення дещо більше середнього лінійного відхилення. Для нормального або близького до нормального розподілу встановлено таке співвідношення між ними: $\sigma = 1{,}25\bar{l}$.

Середнє квадратичне відхилення не завжди зручне для використання, тому що воно не дозволяє порівнювати між собою середні квадратичні відхилення у варіаційних рядах, варіанти яких виражені у різних одиницях виміру. Крім цього, середнє квадратичне відхилення варіаційних рядів за своїм абсолютним значенням залежить не тільки від варіації ознаки, але й від абсолютних рівнів варіантів і середньої, що також не дає змоги порівнювати їх між собою.

Щоб мати можливість порівнювати середні квадратичні відхилення різних варіаційних рядів, потрібно перейти від абсолютних до відносних показників варіації.

До числа відносних показників варіації відносять коефіцієнт варіації ($V_\sigma$), відносний розмах варіації ($V_R$) і відносне лінійне відхилення ($V_{\bar{l}}$).

*Коефіцієнт варіації* являє собою відсоткове відношення середнього квадратичного відхилення до середньої арифметичної.

$$V_\sigma = \frac{\sigma}{\bar{x}} \cdot 100 .$$

Чим менше значення коефіцієнта варіації, тим однорідніша сукупність за досліджуваною ознакою і більш типова середня величина.

*Відносний розмах варіації* визначається шляхом ділення абсолютного розмаху варіації на середню арифметичну у відсотках.

$$V_R = \frac{R}{\bar{x}} \cdot 100 .$$

Якщо в якості абсолютного показника варіації використовують середнє лінійне відхилення, то відносним показником варіації буде *відносне лінійне відхилення*.

Воно визначається як відсоткове відношення середнього лінійного відхилення до середньої арифметичної або медіани.

$$V_{\bar{l}} = \frac{\bar{l}}{\bar{x}} \cdot 100 , \ V_{\bar{l}} = \frac{\bar{l}}{M_e} \cdot 100 .$$

Оскільки абсолютна величина середнього квадратичного відхилення завжди більша від абсолютної величини середнього лінійного відхилення, то й коефіцієнт варіації буде більшим відносного лінійного відхилення:

$$V_\sigma > V_{\bar{l}} .$$

Розглянемо обчислення вищенаведених показників варіації на прикладі. Нехай маємо дані про розподіл виробів за масою.

Таблиця 8.3

Розрахункова таблиця

| Маса виробу, г (x) | Число виробів, шт. (f) | x | xf | $\lvert x - \bar{x} \rvert$ | $\lvert x - \bar{x} \rvert f$ | $(x - \bar{x})^2$ | $(x - \bar{x})^2 f$ |
|---|---|---|---|---|---|---|---|
| до 100 | 5 | 97,5 | 487,5 | 10 | 50 | 100 | 500 |
| 100-105 | 19 | 102,5 | 1947,5 | 5 | 95 | 25 | 475 |
| 105-110 | 53 | 107,5 | 5697,5 | 0 | 0 | 0 | 0 |
| 110-115 | 17 | 112,5 | 1912,5 | 5 | 85 | 25 | 425 |
| 115-120 | 6 | 117,5 | 705,0 | 10 | 60 | 100 | 600 |
| Разом: | 100 | х | 10750,0 | х | 290 | х | 2000 |

Розмах варіації в нашому прикладі буде дорівнювати:

$$R = x_{max} - x_{min} = 117,5 - 97,5 = 20 \text{ г.}$$

Для обчислення наступних показників варіації потрібно визначити середню вагу виробу в даній сукупності:

$$\bar{x} = \frac{\sum x f}{\sum f} = \frac{10750}{100} = 107{,}5 \text{ г}$$

Середнє лінійне відхилення обчислимо за формулою:

$$\bar{l} = \frac{\sum |x - \bar{x}| f}{\sum f} = \frac{290}{100} = 2{,}9 \text{ г}$$

Отже, варіація маси виробів за середнім лінійним відхиленням – 2,9 г при загальній середній масі 107,5 г.

Обчислимо середній квадрат відхилення (дисперсію) звичайним методом за формулою:

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} = \frac{2000}{100} = 20 .$$

Середнє квадратичне відхилення дорівнює:

$$\sigma = \sqrt{\sigma^2} = \sqrt{20} = 4{,}472 \text{ г}$$

Коефіцієнт варіації середньої маси виробу становить:

$$V_\sigma = \frac{\sigma}{\bar{x}} \cdot 100 = \frac{4{,}472}{107{,}5} = 4{,}16 \text{ \%}.$$

Обчислимо відносне лінійне відхилення середньої маси виробу за формулою:

$$V_{\bar{l}} = \frac{\bar{l}}{\bar{x}} \cdot 100 = \frac{2{,}9}{107{,}5} = 2{,}70 \text{ \%}.$$

Як бачимо: $\sigma > \bar{l} \quad (4{,}472 > 2{,}900)$, а $V_\sigma > V_{\bar{l}} \quad (4{,}16 > 2{,}70)$.

Чим менші за величиною середнє квадратичне відхилення і коефіцієнт варіації, тим однорідніша сукупність досліджуваного явища і надійніше обчислена середня величина.

Величину відхилення якого-небудь варіанта від середньої величини визначають через, так зване, *нормоване відхилення* (t), яке визначається за формулою:

$$t = \frac{x - \bar{x}}{\sigma} .$$

Стосовно цього відхилення математична статистика для характеристики нормального розподілу варіантів широко використовує правило трьох сигм “3σ”, яке більш детально буде розглянуте в наступній темі.

## 8.2. Спрощені способи розрахунку дисперсії

Обчислення дисперсії, а отже, і середнього квадратичного відхилення в цілому ряді випадків пов'язано з великими і складними розрахунками, які потребують значних затрат часу і праці. Однак, їх можна значно спростити, якщо використати деякі математичні властивості дисперсії.

1. Якщо всі варіанти ознаки x зменшити на довільну величину A, то дисперсія від цього не зміниться:

$$\sigma^2{}_{(x-A)} = \sigma^2,$$

де $\sigma^2$ – дисперсія варіантів x;

$\sigma^2{}_{(x-A)}$ – дисперсія варіантів зменшена на довільну величину A.

2. Якщо всі значення варіантів x зменшити в i раз, то дисперсія зменшиться в $i^2$ раз, а середнє квадратичне відхилення – в i раз:

$$\sigma^2 = \sigma^2_{\frac{x}{i}} \cdot i^2, \text{ а } \sigma = \sigma_{\frac{x}{i}} \cdot i.$$

3. Якщо обчислити середній квадрат відхилень від любої величини A, яка в тій чи іншій мірі відрізняється від середньої арифметичної ($\bar{x}$), то він завжди буде більший за середній квадрат відхилень, обчислений від середньої арифметичної:

$$\sigma^2_A > \sigma^2.$$

При цьому більший на цілком певну величину – на квадрат різниці між середньою і цією умовно взятою величиною, тобто на $(\bar{x} - A)^2$. Все це можна записати рівнянням:

$$\sigma^2_A = \sigma^2 + (\bar{x} - A)^2, \text{ або } \sigma^2 = \sigma^2_A - (\bar{x} - A)^2,$$

де $\sigma^2$ – середній квадрат відхилень від середньої арифметичної ($\bar{x}$);

$\sigma^2_A$ – середній квадрат відхилень від довільної величини A.

Тобто, дисперсія від середньої арифметичної завжди буде менша від дисперсій вирахуваних від будь-яких інших величин, так як вона має властивість мінімальності.

Розглянемо методику такого розрахунку на нашому прикладі.

Таблиця 8.4

Розрахунок дисперсії $\sigma_A^2$ при А = 107,5

| Маса виробу, г (х) | Число виробів, шт. (f) | х – А<br>А = 107,5 | $(x-A)^2$ | $(x-A)^2 f$ |
|---|---|---|---|---|
| 97,5 | 5 | -10 | 100 | 500 |
| 102,5 | 19 | -5 | 25 | 475 |
| 107,5 | 53 | 0 | 0 | 0 |
| 112,5 | 17 | 5 | 25 | 425 |
| 117,5 | 6 | 10 | 100 | 600 |
| Разом: | 100 | х | х | 2000 |

$$\sigma_A^2 = \frac{\sum (x-A)^2 f}{\sum f} = \frac{2000}{100} = 20\ ;$$

$$\overline{x} = \frac{\sum x f}{\sum f} = \frac{10750}{100} = 107{,}5 \text{ (обчислена раніше).}$$

Обчислимо дисперсію за формулою:

$$\sigma^2 = \sigma_A^2 - (\overline{x} - A)^2 = 20 - (107{,}5 - 107{,}5)^2 = 20\ .$$

Ми отримали той самий результат, що й при звичайному способі розрахунку дисперсії. В даному випадку середня арифметична і довільна величина А співпали.

Техніку розрахунку дисперсії і середнього квадратичного відхилення можна значно спростити використавши математичні властивості дисперсії, тобто спосіб моментів або відліку від умовного нуля.

Покажемо методику цього розрахунку в табл. 8.5.

Таблиця 8.5

Розрахунок дисперсії способом моментів

| Маса виробу, г (х) | Число виробів, шт. (f) | х | х – А<br>А=107,5 | $\frac{x-A}{i}$<br>i = 5 | $\left(\frac{x-A}{i}\right) f$ | $\left(\frac{x-A}{i}\right)^2$ | $\left(\frac{x-A}{i}\right)^2 f$ |
|---|---|---|---|---|---|---|---|
| До 100 | 5 | 97,5 | - 10 | - 2 | - 10 | 4 | 20 |
| 100-105 | 19 | 102,5 | - 5 | - 1 | - 19 | 1 | 19 |
| 105-110 | 53 | 107,5 | 5 | 0 | 0 | 0 | 0 |
| 110-115 | 17 | 112,5 | 5 | 1 | 17 | 1 | 17 |
| 115-120 | 6 | 117,5 | 10 | 2 | 12 | 4 | 24 |
| Разом: | 100 | х | х | х | 0 | х | 80 |

Насамперед, обчислимо моменти першого і другого порядків:

$$m_1 = \frac{\sum\left(\frac{x-A}{i}\right)f}{\sum f} = \frac{0}{100} = 0\,;$$

$$m_2 = \frac{\sum\left(\frac{x-A}{i}\right)^2 f}{\sum f} = \frac{80}{100} = 0{,}8\,.$$

Дисперсію способом моментів, або відліку від умовного нуля, визначають за формулою:

$$\sigma^2 = i^2\left(m_2 - m_1^2\right) = 5^2\left(0{,}8 - 0^2\right) = 25 \cdot 0{,}8 = 20\,.$$

В тому випадку, коли довільна величина A = 0, а i = 1, дисперсію за способом відліку від умовного нуля визначають за формулою:

$$\sigma^2 = \overline{x^2} - (\bar{x})^2 = \frac{\sum x^2 f}{\sum f} - \left(\frac{\sum x f}{\sum f}\right)^2\,.$$

Обчислимо дисперсію за цією формулою за даними нашого прикладу.

Таблиця 8.6

Розрахунок дисперсії за формулою $\sigma^2 = \overline{x^2} - (\bar{x})^2$

| Маса виробу, г (x) | Число виробів, шт. (f) | x f | $x^2$ | $x^2f$ |
|---|---|---|---|---|
| 97,5 | 5 | 487,5 | 9506,25 | 47531,25 |
| 102,5 | 19 | 1947,5 | 10506,25 | 199618,75 |
| 107,5 | 53 | 5697,5 | 11556,25 | 612481,25 |
| 112,5 | 17 | 1912,5 | 12656,25 | 215156,25 |
| 117,5 | 6 | 705,0 | 13806,25 | 82837,50 |
| Разом: | 100 | 10750,0 | x | 1157625,00 |

Підставимо розраховані дані в табл. 8.6 у формулу дисперсії і отримаємо:

$$\sigma^2 = \frac{\sum x^2 f}{\sum f} - \left(\frac{\sum x f}{\sum f}\right)^2 = \frac{1157625}{100} - \left(\frac{10750}{100}\right)^2 = 115776{,}25 - 11556{,}25 = 20\,.$$

Як і слід було очікувати, результати обчислення дисперсії за всіма вищенаведеними способами однакові.

### 8.3. Дисперсія альтернативної ознаки

Поряд з вимірюванням варіації кількісних ознак, в статистичній практиці доводиться вивчати і обчислювати варіації якісних (альтернативних) ознак.

*Альтернативною* називається ознака, яка може набирати лише два взаємопротилежних значення. Наприклад, продукція на підприємстві може бути якісна і не якісна, товарна і нетоварна, стандартна і нестандартна і т.д.

Кількісно варіацію альтернативної ознаки виражають двома значеннями: наявність ознаки у одиниць сукупності позначають через 1, а її відсутність – через 0. Тоді, якщо частку одиниць, які володіють даною ознакою позначити через p, а частку одиниць, які не володіють ознакою, через q, то p + q = 1, звідси p = 1 – q, а q = 1 – p.

Спочатку обчислимо середнє значення альтернативної ознаки:

$$\bar{x} = \frac{\sum x f}{\sum f} = \frac{(1 \cdot p) + (0 \cdot q)}{p + q} = p .$$

Отже, середнє значення альтернативної ознаки дорівнює частці одиниць, які володіють даною ознакою.

Дисперсія альтернативної ознаки визначається за формулою:

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} = \frac{(1 - p)^2 p + (0 - q)^2 q}{p + q} =$$

$$= q^2 p + p^2 q = pq(q + p) = pq = p(1 - p)$$

Отже, дисперсія альтернативної ознаки визначається як добуток частки одиниць, які володіють даною ознакою, на частку одиниць, які нею не володіють.

Середнє квадратичне відхилення альтернативної ознаки визначається як корінь квадратний з дисперсії цієї ознаки:

$$\sigma = \sqrt{pq} = \sqrt{p(1 - p)} .$$

Покажемо хід обчислення середнього значення альтернативної ознаки, її дисперсії і середнього квадратичного відхилення на такому прикладі. При обстеженні 100 взірців готових виробів, відібраних у випадковому порядку, 20 % виявились неякісними. Потрібно визначити середню величину, дисперсію і середнє квадратичне відхилення.

Оскільки ознака якісних взірців готових виробів є альтернативною, то середня величина дорівнює частці, тобто

$$\bar{x} = p = \frac{80}{100} = 0{,}8 \text{, або } 80\ \%.$$

Дисперсію альтернативної ознаки обчислимо за формулою:

$$\sigma^2 = p(1 - p) = 0{,}8(1 - 0{,}8) = 0{,}16\ .$$

Звідси середнє квадратичне відхилення дорівнює:

$$\sigma = \sqrt{\sigma^2} = \sqrt{0{,}16} = 0{,}4 \text{, або } 40\ \%.$$

Граничне значення дисперсії альтернативної ознаки при p = 0,5 дорівнює 0,25.

## 8.4 Види дисперсій і правило їх додавання

Для більш детального вивчення варіації тієї чи іншої ознаки, статистика за допомогою правила додавання дисперсій виявляє і досліджує вплив окремих факторів і умов, які зумовили дану варіацію в цілому. Виявити частку варіації, зумовлену певними факторами, можна розбивши всю сукупність на групи за ознакою, вплив якої досліджується.

Якщо сукупність розбита на групи за одним фактором, то для неї можна обчислити такі види дисперсій: загальну, групові (часткові), середню з групових (часткових) і міжгрупову.

*Загальна дисперсія* визначається як середня арифметична з квадратів відхилень кожного значення ознаки від їх загальної середньої величини. Дана дисперсія характеризує варіацію досліджуваної ознаки за рахунок впливу всіх факторів. Її вираховують за формулами:

$$\sigma_з^2 = \frac{\sum(x - \bar{x}_з)^2}{n} \text{ – проста; } \sigma_з^2 = \frac{\sum(x - \bar{x}_з)^2 f}{\sum f} \text{ – зважена.}$$

*Групова (часткова) дисперсія* визначається як середня арифметична з квадратів відхилень кожного значення ознаки в групі від групової середньої. Групові дисперсії визначаються за формулами:

$$\sigma_i^2 = \frac{\sum(x_i - \bar{x}_i)^2}{n_i} \text{ – проста; } \sigma_i^2 = \frac{\sum(x_i - \bar{x}_i)^2 f_i}{\sum f_i} \text{ – зважена.}$$

де $x_i$ – індивідуальні значення групових ознак;

$\bar{x}_i$ – середнє значення ознак i-ї групи;

n – число одиниць сукупності в групі;

$f_i$ – частоти в групі.

Групова (часткова) дисперсія вимірює варіацію ознаки тільки за рахунок факторів, які діють в середині групи, тобто всіх інших факторів, крім покладеного в основу групування.

*Середню внутрішньогрупову (залишкову) дисперсію* визначають за формулами середньої арифметичної з групових дисперсій:

$$\overline{\sigma_i^2} = \frac{\sum \sigma_i{}^2}{n} \text{ – проста; } \overline{\sigma_i^2} = \frac{\sum \sigma_i{}^2 f_i}{\sum f_i} \text{ – зважена.}$$

*Міжгрупова дисперсія* визначається як середня арифметична з квадратів відхилень групових середніх від загальної середньої за формулами:

$$\delta^2 = \frac{\sum (\bar{x}_i - \bar{x}_з)^2}{n} \text{ – проста; } \delta^2 = \frac{\sum (\bar{x}_i - \bar{x}_з)^2 f_i}{\sum f_i} \text{ – зважена.}$$

де $\delta^2$ – міжгрупова дисперсія;
$\bar{x}_i$ – групові середні;
$\bar{x}_з$ – загальна середня;
n – число одиниць сукупності;
$f_i$ – ваги або частоти.

Міжгрупова дисперсія відображає варіацію досліджуваної ознаки під впливом фактора, покладеного в основу групування.

Математичною статистикою доведено, що між загальною дисперсією, середньою з групових дисперсій і міжгруповою дисперсією існує зв'язок, який описується рівністю:

$$\sigma_з^2 = \overline{\sigma_i^2} + \delta^2 .$$

Ця рівність в статистиці називається *правилом додавання дисперсій*. За допомогою даного правила, знаючи два види дисперсій, завжди можна визначити невідомий третій вид:

$$\overline{\sigma_i^2} = \sigma_з^2 - \delta^2 ;$$

$$\delta^2 = \sigma_з^2 - \overline{\sigma_i^2} .$$

Правило додавання дисперсій використовують при проведенні вибіркового спостереження, для спрощеного обчислення дисперсій громіздкого варіаційного ряду, вимірювання сили зв'язку між явищами та ін.

Проілюструємо застосування правила додавання дисперсій на прикладі.

Нехай маємо відомості про заробітну плату десяти робітників, розчленованих на дві групи за рівнем фахової підготовки.

Таблиця 8.7

Розрахункова таблиця

| Перша група | | | | Друга група | | | |
|---|---|---|---|---|---|---|---|
| Табель-ний номер робіт-ника ($n_1$) | Денна заробітна плата робітників-практиків, грн. ($x_1$) | $x_1-\overline{x}_1$ | $(x_1-\overline{x}_1)^2$ | Табель-ний номер робіт-ника ($n_2$) | Денна заробітна плата робітників, які закін-чили ПТУ, грн. ($x_2$) | $x_2-\overline{x}_2$ | $(x_2-\overline{x}_2)^2$ |
| 1 | 6,0 | -0,52 | 0,2704 | 6 | 8,2 | -0,42 | 0,1764 |
| 2 | 6,3 | -0,22 | 0,0484 | 7 | 8,5 | -0,12 | 0,0144 |
| 3 | 6,5 | -0,02 | 0,0004 | 8 | 8,6 | -0,02 | 0,0004 |
| 4 | 6,8 | 0,28 | 0,0784 | 9 | 8,8 | 0,18 | 0,0324 |
| 5 | 7,0 | 0,48 | 0,2304 | 10 | 9,0 | 0,38 | 0,1444 |
| Разом: | 32,6 | х | 0,6280 | Разом: | 43,1 | х | 0,3680 |

Спочатку обчислимо середньоденну заробітну плату робітників для кожної групи:

$$\overline{x}_1=\frac{\sum x_1}{n_1}=\frac{32,6}{5}=6,52\text{ грн.},\ \overline{x}_2=\frac{\sum x_2}{n_2}=\frac{43,1}{5}=8,62\text{ грн.}$$

За даними розрахункової таблиці визначимо групові дисперсії:

$$\sigma_1^2=\frac{\sum(x_1-\overline{x}_1)^2}{n_1}=\frac{0,6280}{5}=0,1256\ ;$$

$$\sigma_2^2=\frac{\sum(x_2-\overline{x}_2)^2}{n_2}=\frac{0,3680}{5}=0,0736\ .$$

Середня з групових дисперсій дорівнює:

$$\overline{\sigma_i^2}=\frac{\sum\sigma_i^2}{n}=\frac{0,1256+0,0736}{2}=\frac{0,1992}{2}=0,0996\ .$$

Для знаходження міжгрупової дисперсії потрібно обчислити загальну середню денну заробітну плату всіх робітників за формулою середньої арифметичної простої з групових середніх:

$$\overline{x}_з=\frac{\sum x_i}{n}=\frac{6,52+8,62}{2}=\frac{15,14}{2}=7,57\text{ грн.}$$

Тепер вже можемо обчислити міжгрупову дисперсію:

$$\delta^2 = \frac{\sum(\overline{x}_i - \overline{x}_з)^2}{n} = \frac{(6,52-7,57)^2 + (8,62-7,57)^2}{2} = \frac{2,205}{2} = 1,1025 .$$

Використавши правило додавання дисперсій, визначимо загальну дисперсію:

$$\sigma_з^2 = \overline{\sigma_i^2} + \delta^2 = 0,0996 + 1,1025 = 1,2021$$

Правильність наших розрахунків ми можемо перевірити, обчисливши загальну дисперсію звичайним способом:

$$\sigma_з^2 = \frac{\sum(x-\overline{x})^2}{n} = \frac{(6,0-7,57)^2 + (6,3-7,57)^2 + (6,5-7,57)^2 + (6,8-7,57)^2 +}{10}$$

$$\frac{+(7,0-7,57)^2 + (8,2-7,57)^2 + (8,5-7,57)^2 + (8,6-7,57)^2 +}{10}$$

$$\frac{+(8,8-7,57)^2 + (9,0-7,57)^2}{10} = \frac{12,021}{10} = 1,2021 .$$

Як показала перевірка, результати обчислення загальної дисперсії за обома способами такі самі, що свідчить про правильність проведених нами розрахунків.

Середнє квадратичне відхилення із загальної дисперсії дорівнює:

$$\sigma_з = \sqrt{\sigma_з^2} = \sqrt{1,2021} = 1,096 \text{ грн.}$$

Таким чином, ми визначили, що середня денна заробітна плата групи робітників в кількості десяти чоловік склала 7,57 грн. при середньому квадратичному відхиленні 1,096 грн.

При цьому, ми можемо стверджувати – якщо загальна дисперсія склала 1,2021, то міжгрупова дисперсія в розмірі 1,1025 викликана різницями кваліфікації в групах робітників, а середня внутрішньогрупова дисперсія в розмірі 0,0996 показує частку впливу інших, крім покладеного в основу групування, факторів.

Як бачимо, правило додавання дисперсій дозволяє визначити частку складових частин в загальній дисперсії.

Поділивши міжгрупову дисперсію на загальну дисперсію отримаємо показник, який називається *коефіцієнтом детермінації* і показує, яка частка всієї варіації ознаки, зумовлена ознакою, покладеною в основу групування. В нашому прикладі:

$$\eta^2 = \frac{\delta^2}{\sigma_з^2} = \frac{1,1025}{1,2021} = 0,917025 \text{, або } 91,7\ \%,$$

де $\eta^2$ – (грецька буква “ета” в квадраті) – коефіцієнт детермінації.

Значить, фактор кваліфікації робітників на 91,7 % зумовлює варіацію їхньої середньоденної заробітної плати.

Корінь квадратний із коефіцієнта детермінації називають *емпіричним кореляційним відношенням*, яке показує тісноту (силу) зв’язку між групувальною і результативною ознаками. Визначимо цей показник для нашого прикладу:

$$\eta = \sqrt{\frac{\delta^2}{\sigma_з^2}} = \sqrt{0,917} = 0,958\ .$$

Це говорить про те, що зв’язок між кваліфікацією робітників і їхньою середньоденною заробітною платою дуже сильний (тісний).

***Контрольні запитання***

1. Поняття про варіацію ознак та необхідність її статистичного вивчення.
2. Статистичні показники варіації та способи їх обчислення.
3. Абсолютний показник варіації.
4. Які ви знаєте середні показники варіації?
5. Дисперсія альтернативної ознаки.
6. Математичні властивості дисперсії.
7. Обчислення дисперсії способом моментів.
8. Необхідність обчислення і види відносних показників варіації.
9. Спрощені способи розрахунку дисперсії.
10. Види дисперсій.
11. Що характеризує загальна дисперсія?
12. Як визначається і що вимірює групова (часткова) дисперсія?
13. Як визначається середня внутрішньогрупова дисперсія?
14. Як визначається і що відображає міжгрупова дисперсія?
15. Правило додавання дисперсій.
16. Що показує коефіцієнт детермінації?
17. Як визначається і що характеризує емпіричне кореляційне відношення?
18. Використання правила додавання дисперсій в економіко-статистичному аналізі.

# Розділ 9. Ряди розподілу

## 9.1. Поняття про ряди розподілу, їх види

В результаті статистичного групування отримують ряди цифрових показників, які характеризують розподіл одиниць досліджуваної сукупності за варіюючою ознакою. Такі ряди називаються *рядами розподілу*.

Ряд розподілу складається з двох елементів – варіантів та частот. *Варіант* “x” – це окремі значення групувальної ознаки, які розташовані в певній послідовності. *Частоти* “f” – це числа, які показують, скільки разів певне значення ознаки зустрічається у сукупності, або скільки одиниць сукупності припадає на кожну групу.

Ряди розподілу відіграють важливу роль при вивченні складу сукупності за досліджуваною ознакою, закономірностей розподілу, використовуються при визначенні середніх величин, показників варіації, взаємозв’язку та ін.

В залежності від групувальної ознаки ряди розподілу поділяються на атрибутивні та варіаційні (кількісні).

В атрибутивних рядах розподілу варіанти не мають числового виразу (колір досліджуваного об’єкту, професія працівника, національність, стать та ін.). Альтернативна варіація – це варіація якісної ознаки, при якій ознака приймає тільки два взаємо протилежних значення (дешевий – дорогий, легкий – важкий, якісний – бракований і т.д.).

Варіаційні ряди, в яких варіанти мають числовий вираз, поділяються на дискретні та інтервальні. В дискретних рядах варіанти являють собою дискретні числа, в інтервальних – це інтервали групування.

В тому випадку, коли виконується групування за двома і більше ознаками отримують комбінаційний ряд розподілу.

При побудові атрибутивних рядів розподілу варіанти потрібно розташовувати в логічній послідовності. При використанні дискретних та інтервальних варіаційних рядів варіанти записують в порядку зростання або спадання. Для інтервальних рядів важливим є чітке розмежування варіантів.

Розрізняють ряди розподілу з абсолютними, відносними та нагромадженими частотами. Нагромаджені частоти ще називають кумулятивними. В першому випадку частоти являють абсолютні числа, в другому – питому вагу або частку кожної групи.

Ряди розподілу з абсолютними частотами характеризують склад сукупності, а з відносними – структуру сукупності.

Ряди розподілу з нагромадженими (кумулятивними) частотами показують, яка чисельність або питома вага одиниць має значення ознаки менше за дане. Кумулятивні частоти знаходять шляхом їх сумування по групах.

*Щільність розподілу* – це кількість одиниць сукупності, що припадає на одиницю ширини інтервалу групувальної ознаки. Розрізняють абсолютну ($f_d^a$) і відносну ($f_d^B$) щільність, яка розраховується за формулами:

$$f_d^a = \frac{f}{i};\; f_d^B = \frac{p}{i},$$

де f – частота;

p – частка (доля);

i – величина (розмір) інтервалу.

Звернемось до прикладу. Маємо такі дані про розподіл сімей за місячним доходом:

Таблиця 9.1

| Місячний доход на одного члена сім’ї, грн. | Число сімей | | Щільність розподілу | |
|---|---|---|---|---|
| | одиниць | % | число сімей | % |
| До 200 | 34 | 13,2 | 0,17 | 0,066 |
| 200-400 | 52 | 20,2 | 0,26 | 0,101 |
| 400-600 | 72 | 27,9 | 0,36 | 0,140 |
| 600-1000 | 70 | 27,1 | 0,18 | 0,068 |
| 1000-1500 | 30 | 11,6 | 0,06 | 0,023 |
| Разом: | 258 | 100,0 | x | x |

Розрахунки показують, що найбільшу щільність розподілу має третя група сімей з доходом 400-600 грн. на одного члена сім’ї.

*Інтерполяція* в рядах розподілу визначає, скільки одиниць сукупності (або процентів) мають значення ознаки менше за задане. Для інтерполяції використовують як абсолютні так і відносні нагромаджені частоти.

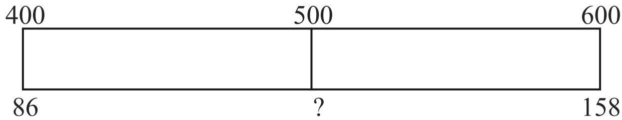
Методику розрахунку цього показника розглянемо за даними попереднього прикладу:

Таблиця 9.2

| Місячний доход на одного члена сім’ї, грн. | Нагромаджені частоти | |
|---|---|---|
| | число сімей | % |
| До 200 | 34 | 13,2 |
| 200-400 | 86 | 33,4 |
| 400-600 | 158 | 61,3 |
| 600-1000 | 228 | 88,4 |
| 1000-1500 | 258 | 100,0 |

Визначити: а) скільки сімей мають доход менше 500 грн. на одного члена сім'ї.

Розрахунок:




$$f_x < 500 = 86 + 0{,}5(158 - 86) = 122 \text{ сім'ї, або}$$

$$f_x < 500 = 33{,}4 + 0{,}5(61{,}3 - 33{,}4) = 47{,}4 \ \%;$$

б) скільки сімей мають доход менше 730 грн. на одного члена сім'ї.

$$f_x < 730 = 158 + \frac{730 - 600}{400} \cdot (228 - 158) = 181 \text{ сім'я, або}$$

$$f_x < 730 = 61{,}3 + 0{,}325 \cdot (88{,}4 - 61{,}3) = 70{,}1 \ \%.$$

З допомогою інтерполяції визначають менше якого значення мають ознаку певна кількість (питома вага) одиниць сукупності. Наприклад, менше якого мають доход 50 %, або 129 сімей.

Розрахунок:

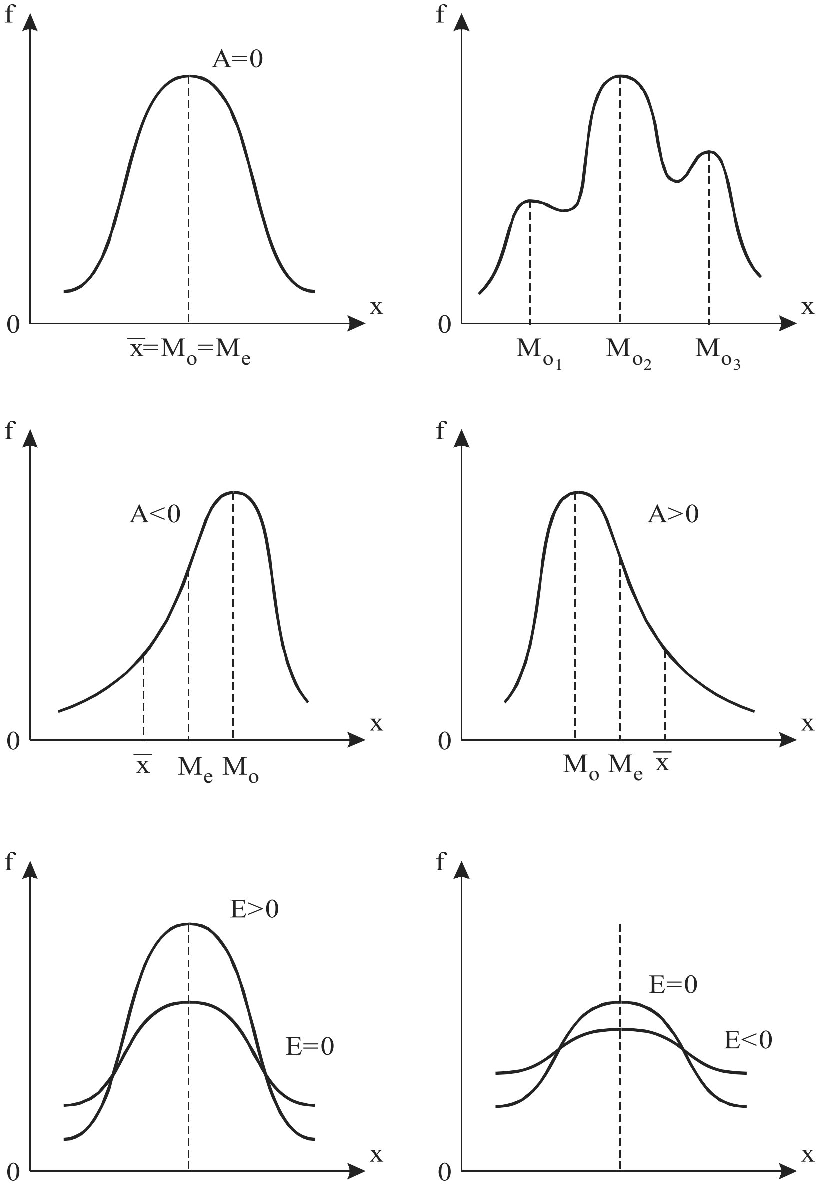
$$x_f = 50\,\% = 400 + \frac{50{,}0 - 33{,}4}{61{,}3 - 33{,}4} \cdot (600 - 400) = 520 \text{ грн.}$$

Отже, половина сімей має доход менше 520 грн. на одного члена сім'ї.

## 9.2. Форми рядів розподілу, та їх характеристика

Різноманітність статистичних сукупностей – передумова різних форм співвідношення частот і значень варіюючої ознаки. За своєю формою розподіли поділяють на такі види: одно-, дво- і багатовершинні. Наявність двох і більше вершин свідчить про неоднорідність сукупності, про поєднання в ній груп з різними рівнями ознаки. Розподіли якісно однорідних сукупностей, як правило, одновершинні. Серед одновершинних розподілів є симетричні і асиметричні (скошені), гостро- і плосковершинні. Різні види форм розподілу покажемо на графіках (мал. 9.1).

Якщо частоти варіантів рівновіддалені від центру значення ознаки, такий варіаційний ряд називається *симетричним*, а якщо ж вершина розподілу зміщена, тобто частоти по обидві сторони від центру змінюються неоднаково, тоді варіаційний ряд називається *асиметричним*, або *скошеним*.




Мал. 9.1 Різні види форм розподілу

Розрізняють правосторонню і лівосторонню асиметрію. Напрям асиметрії протилежний напряму зміщення вершини розподілу. При правосторонній асиметрії вершина розподілу зміщена уліво, при лівосторонній – управо. Асиметрія появляється із-за обмеженої варіації ознак в одному напрямку, або під впливом переважаючої причини розвитку явища, яка веде до зміщення центру його розподілу.

Відхилення між середньою арифметичною і медіаною або модою покажуть міру асиметрії. В симетричному розподілі необхідною умовою є рівність трьох характеристик: середньої арифметичної, моди і медіани:

$$\overline{x} = M_o = M_e .$$

В помірно асиметричних розподілах між середньою арифметичною, медіаною і модою існує певна наближена рівність:

$$M_o \approx \overline{x} - 3\left(\overline{x} - M_e\right).$$

У випадку, коли чітко простежується асиметрія варіаційного ряду для більш глибокого вивчення економічних явищ, середня величина ознаки повинна доповнюватися модою і медіаною.

В асиметричному розподілі між середньою арифметичною, медіаною і модою існують певні розбіжності. При правосторонній асиметрії $\overline{x} > M_e > M_o$, а при лівосторонній – $\overline{x} < M_e < M_o$.

Стандартизовані відхилення характеризують напрям і міру скошеності розподілу. Коефіцієнт асиметрії (А) визначається як відношення різниці між середньою арифметичною і модою чи медіаною до середнього квадратичного відхилення:

$$A = \frac{\overline{x} - M_o}{\sigma} \text{ або } A = \frac{\overline{x} - M_e}{\sigma} .$$

Коефіцієнт асиметрії коливається в межах від – 3 до + 3.

В симетричному розподілі А = 0, при правосторонній асиметрії А > 0, при лівосторонній А < 0. Використовуючи стандартизоване відхилення середньої і моди, оцінимо симетричність розподілу посівних площ за урожайністю гречки.

Таблиця 9.3

Розрахунок середньої і середнього квадратичного відхилення.

| Урожайність гречки, ц/га | Посівна площа, га (f) | x | xf | $x-\bar{x}$ | $(x-\bar{x})^2$ | $(x-\bar{x})^2 f$ |
|---|---|---|---|---|---|---|
| 10-12 | 17 | 11 | 187 | - 3,9 | 15,21 | 258,57 |
| 12-14 | 20 | 13 | 260 | - 1,9 | 3,61 | 72,20 |
| 14-16 | 27 | 15 | 405 | 0,1 | 0,01 | 0,27 |
| 16-18 | 23 | 17 | 391 | 2,1 | 4,41 | 101,43 |
| 18-20 | 13 | 19 | 247 | 4,1 | 16,81 | 218,53 |
| Разом: | 100 | x | 1490 | x | x | 651,00 |

$$\bar{x}=\frac{\sum xf}{\sum f}=\frac{1490}{100}=14{,}9 \text{ ц/га}; \ \sigma^2=\frac{\sum(x-\bar{x})^2 f}{\sum f}=\frac{651}{100}=6{,}51;$$

$$\sigma=\sqrt{\sigma^2}=\sqrt{6{,}51}=2{,}55 \text{ ц/га};$$

$$M_o=x_{mo}+i\frac{f_2-f_1}{(f_2-f_1)+(f_2-f_3)}=14+2\frac{27-20}{(27-20)+(27-23)}=15{,}27 \text{ ц/га};$$

$$A=\frac{\bar{x}-M_o}{\sigma}=\frac{14{,}9-15{,}27}{2{,}55}=-0{,}145.$$

Стандартизоване відхилення свідчить про незначну лівосторонню асиметрію, а тому розподіл посівних площ гречки за урожайністю можна вважати симетричним.

Крутість варіаційного ряду, тобто його високовершинність (гостровершинність) або низьковершинність (плосковершинність) називається *ексцесом*. Розподіли більш гостровершинні, аніж нормальні, володіють позитивним ексцесом, а більш плоско вершинні – від'ємним ексцесом. На практиці в одному розподілі часто поєднуються всі особливості: одновершинний розподіл може бути симетричним і високовершинним або скошеним і низьковершинним.

Як узагальнюючі характеристики в якості міри крутості розподілу використовують моменти. За їх допомогою можна описати будь-який розподіл. *Моментом розподілу* називається середня арифметична k-го ступеня відхилень варіантів "x" від деякої постійної величини "A".

В залежності від вибору величини "A" моменти поділяються на початкові, початкові відносно "$x_o$" (умовні) і центральні. Ступінь "k" визначає моменти різних порядків.

На практиці використовуються:

*Початкові моменти* (Mk) отримують у випадку, якщо постійна величина, від якої знаходять відхилення при вирахуванні моментів, дорівнює нулю A = 0.

*Початкові моменти:*

- нульового порядку (k = 0) $M_0 = \dfrac{\sum x^0 f}{\sum f} = 1$;
- першого порядку (k = 1) $M_1 = \dfrac{\sum x^1 f}{\sum f} = \bar{x}$;
- другого порядку (k = 2) $M_2 = \dfrac{\sum x^2 f}{\sum f} = \overline{x^2}$;
- третього порядку (k = 3) $M_3 = \dfrac{\sum x^3 f}{\sum f} = \overline{x^3}$;
- четвертого порядку (k = 4) $M_4 = \dfrac{\sum x^4 f}{\sum f} = \overline{x^4}$.

*Початкові моменти відносно “$x_o$”* ($M'_k$) отримують у випадку, якщо постійна величина, від якої знаходять відхилення при вирахуванні моментів, дорівнює довільно вибраній величині A = $x_o$.

*Початкові моменти відносно “$x_o$”* :

- нульового порядку $M'_0 = \dfrac{\sum (x - x_0)^0 f}{\sum f} = 1$;
- першого порядку $M'_1 = \dfrac{\sum (x - x_0) f}{\sum f} = \bar{x} - x_0$;
- другого порядку $M'_2 = \dfrac{\sum (x - x_0)^2 f}{\sum f}$;
- третього порядку $M'_3 = \dfrac{\sum (x - x_0)^3 f}{\sum f}$;
- четвертого порядку $M'_4 = \dfrac{\sum (x - x_0)^4 f}{\sum f}$.

*Центральні моменти* ($\mu_k$) отримують у випадку, якщо постійна величина, від якої знаходять відхилення при вирахуванні моментів дорівнює середній величині $A = \bar{x}$.

*Центральні моменти:*

- нульового порядку $\mu_0 = \dfrac{\sum (x - \bar{x})^0 f}{\sum f} = 1$;
- першого порядку $\mu_1 = \dfrac{\sum (x - \bar{x}) f}{\sum f} = 0$;
- другого порядку $\mu_2 = \dfrac{\sum (x - \bar{x})^2 f}{\sum f} = \sigma^2$;
- третього порядку $\mu_3 = \dfrac{\sum (x - \bar{x})^3 f}{\sum f}$;
- четвертого порядку $\mu_4 = \dfrac{\sum (x - \bar{x})^4 f}{\sum f}$.

Початковий момент першого порядку – це середня арифметична "$\bar{x}$", другого – середній квадрат значень ознаки $\overline{x^2}$. Центральний момент другого порядку характеризує варіацію $\mu_2 = \sigma^2$, третього – асиметрію, четвертого – ексцес.

Для порівняння ступеня асиметрії різних розподілів використовують стандартний момент:

$$A = \frac{\mu_3}{\sigma^3}.$$

Вважають, коли A < 0,25 асиметрія низька, якщо A не перевищує 0,5 – середня і при A більшому за 0,5 – висока.

Для вимірювання ексцесу використовують стандартизований момент четвертого порядку:

$$E = \frac{\mu_4}{\sigma^4}.$$

У симетричному розподілі E = 3, при гостровершинному – E >3, плосковершинному – E < 3.

Для розрахунку коефіцієнтів асиметрії і ексцесу використаємо дані таблиці 9.3.

Розрахунок центральних моментів третього і четвертого порядків, за даними центрального ряду розподілу, доцільно провести за формулами:

$$\mu_3 = \frac{\sum\left(\frac{x-\bar{x}}{i}\right)^3 f}{\sum f}\cdot i^3;\ \mu_4 = \frac{\sum\left(\frac{x-\bar{x}}{i}\right)^4 f}{\sum f}\cdot i^4,$$

де i – ширина інтервалу (може бути будь-яким числом);
f – частота або частка інтервалу.

Таблиця 9.4

Розрахунок коефіцієнтів асиметрії та ексцесу

| Урожайність гречки, ц/га | Посівна площа, га (f) | x | xf | $\frac{x-\bar{x}}{i}$ i = 2 | $\left(\frac{x-\bar{x}}{i}\right)^2$ | $\left(\frac{x-\bar{x}}{i}\right)^2 f$ | $\left(\frac{x-\bar{x}}{i}\right)^3 f$ | $\left(\frac{x-\bar{x}}{i}\right)^4 f$ |
|---|---|---|---|---|---|---|---|---|
| 10-12 | 0,17 | 11 | 1,87 | -1,95 | 3,8025 | 0,6464 | -1,2605 | 2,4580 |
| 12-14 | 0,20 | 13 | 2,60 | -0,95 | 0,9025 | 0,1805 | -0,1715 | 0,1629 |
| 14-16 | 0,27 | 15 | 4,05 | 0,05 | 0,0025 | 0,0007 | 0,0000 | 0,0000 |
| 16-18 | 0,23 | 17 | 3,91 | 1,05 | 1,1025 | 0,2536 | 0,2662 | 0,2796 |
| 18-20 | 0,13 | 19 | 2,47 | 2,05 | 4,2025 | 0,5463 | 1,1200 | 2,2959 |
| Разом: | 1,00 | x | 14,90 | x | x | 1,6275 | -0,0458 | 5,1964 |

$$\sigma = \sqrt{2^2\cdot 1{,}6275} = 1{,}2757 \text{ ц/га};$$

$$\mu_3 = \frac{\sum\left(\frac{x-\bar{x}}{i}\right)^3 f}{\sum f}\cdot i^3 = \frac{-0{,}0458}{1{,}000}\cdot 8 = -0{,}3664;$$

$$A = \frac{\mu_3}{\sigma^3} = \frac{-0{,}3664}{2{,}0761} = -0{,}176;$$

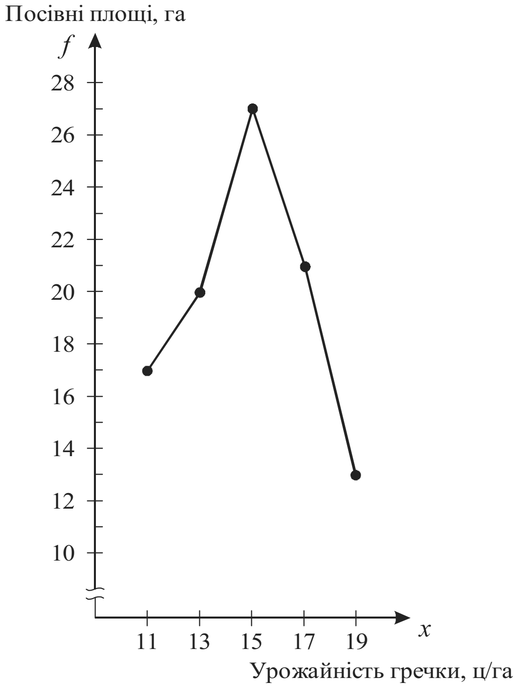
$$\mu_4 = \frac{\sum\left(\frac{x-\bar{x}}{i}\right)^4 f}{\sum f}\cdot i^4 = \frac{5{,}1964}{1{,}000}\cdot 16 = 83{,}1424;$$

$$E = \frac{\mu_4}{\sigma^4} = \frac{83{,}1424}{2{,}6485} = 31{,}392.$$

Отже, розподіл посівних площ за урожайністю гречки характеризується незначною лівосторонньою асиметрією.

Концентрація посівних площ (70 %) в другому, третьому і четвертому інтервалах відображена в коефіцієнті ексцесу. В нашому прикладі коефіцієнт ексцесу значно більший за 3 (Е = 31,392), що свідчить про гостровершинність розподілу.

Криву розподілу посівних площ гречки проілюструємо на мал. 9.2.




Мал. 9.2 Розподіл посівних площ за урожайністю гречки.

На мал. 9.2 чітко видно, що даний розподіл має позитивний ексцес, тобто характеризується нагромадженням членів ряду в центрі розподілу.

При безпосередній оцінці концентрації розподілу порівнюють частки розподілу елементів, сукупності ($d_j$) з обсягом ознаки ($xd_j$). У випадку рівномірного розподілу $d_j = xd_j$. Суттєві відхилення $xd_j - d_j$ вказують на певну концентрацію елементів сукупності. Для визначення міри концентрації вираховують коефіцієнт концентрації (К) за формулою:

$$К = \frac{\sum \left| xd_j - d_j \right|}{2}.$$

За умови рівномірного розподілу К = 0, при повній концентрації К = 1, а у всіх інших випадках цей коефіцієнт буде тим більший, чим більший ступінь концентрації. Розглянемо приклад (табл. 9.5).

Таблиця 9.5

Розрахунок коефіцієнта концентрації споживання електроенергії промисловими підприємствами Західного регіону (дані умовні)

| Групи підприємств за середньорічною вартістю основних виробничих фондів, тис. грн. | Доля підприємств $d_j$ | Доля спожитої електроенергії $xd_j$ | $\lvert xd_j - d_j \rvert$ |
|---|---|---|---|
| До 300 | 0,215 | 0,002 | 0,213 |
| 301-900 | 0,380 | 0,018 | 0,362 |
| 901-5000 | 0,217 | 0,050 | 0,167 |
| 5001-10000 | 0,141 | 0,131 | 0,010 |
| 10001-25000 | 0,023 | 0,102 | 0,079 |
| 25001-50000 | 0,013 | 0,117 | 0,104 |
| 50001-200000 | 0,011 | 0,580 | 0,569 |
| Разом: | 1,000 | 1,000 | 1,504 |

Дані таблиці 9.5 свідчать, що більш як половина спожитої електроенергії (56,9 %) припадає на підприємства, які становлять 1,1 % загального їх числа, і, навпаки, 21,5 % промислових підприємств першої групи споживають лише 0,2 % електроенергії. Коефіцієнт концентрації дорівнює:

$$К = \frac{\sum \left| xd_j - d_j \right|}{2} = \frac{1{,}504}{2} = 0{,}752 .$$

Він свідчить про високий рівень концентрації споживання електроенергії промисловими підприємствами Західного регіону.

## 9.3. Криві розподілу та способи перевірки гіпотез

Закони розподілу являють собою узагальнюючу характеристику варіації в однорідній сукупності. Фактичні розподіли можуть бути графічно зображені емпіричною кривою розподілу. Поглиблюючи аналіз варіації рядів розподілу можна описати закономірність співвідношення варіантів і частот функцією теоретичної кривої. Серед найбільш розповсюджених є *крива нормального розподілу*. Вона використовується як стандарт для порівняння інших розподілів, а також застосовується при проведенні вибіркового, кореляційно-регресійного, факторного та інших статистичних методів дослідження.

Нормальний розподіл близький до інших одновершинних розподілів, а тому його часто використовують як перше наближення при статистичному моделюванні. Деякі розподіли приводять до нормального виду через заміну значень “x” їх логарифмами “lg x”. Логарифмічною нормальною кривою можна описати ряд асиметричних розподілів.

Частоти теоретичної кривої нормального розподілу визначають за формулою:

$$f' = \sum f \frac{i}{\sigma} f(t), \text{ де } f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}.$$

Інтервальна функція розподілу має вигляд:

$$F(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_{-x}^{+x} e^{-\frac{t^2}{2}} dt.$$

Функція “F(x)” табульована, її значення знаходять по спеціальній таблиці.

Аналітично нормальний розподіл описується таким рівнянням:

$$\hat{y}_t = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2}},$$

де $\hat{y}_t$ – ордината кривої нормального розподілу;

t – нормоване відхилення, що дорівнює $\frac{x - \bar{x}}{\sigma}$;

$\sigma$ – середнє квадратичне відхилення;

$\pi$ – величина відношення довжини кола до діаметру, $\pi$ = 3,1415;

e – основа натуральних логарифмів, e ≈ 2,7182.

Для перевірки відповідності фактичного розподілу нормальному, частоти фактичного розподілу порівнюють з теоретичними, які характерні для нормального розподілу. З цією метою за фактичними даними вираховують теоретичні частоти кривої нормального розподілу. Тобто, фактичну криву розподілу вирівнюють кривою нормального розподілу.

Після розрахунку теоретичних частот виникає потреба перевірки висунутої гіпотези про відповідність або невідповідність того чи іншого теоретичного закону розподілу, прийнятого в якості математичної моделі для емпіричного розрахунку. При цьому виходять з того, що якщо відхилення між емпіричними і теоретичними частотами можна вважати випадковими, тоді гіпотеза про те, що прийнятий теоретичний розподіл відповідає даному емпіричному, не відхиляється.

Математична статистика дає декілька показників, за якими судять наскільки фактичний розподіл узгоджується з нормальним. Такі показники називають *критеріями узгодження*. Критерії узгодження виступають у вигляді деякої величини, яка оцінює досліджуване явище з певною ймовірністю.

Статистика використовує критерії узгодження К.Пірсона ($\chi^2$), А.Н.Колмогорова ($\lambda$), Б.С.Ястремського (L), В.І.Романовського (R), Р.Фішера (Z), Вілконсона та ін.

Одним із основних і найбільш розповсюджених показників є критерії Пірсона $\chi^2$ і Колмогорова $\lambda$.

*Критерій* $\chi^2$ запропонував англійський вчений К.Пірсон, статистичну характеристику якого обчислюють за формулою:

$$\chi^2 = \sum \frac{(f - f')^2}{f'},$$

де f і $f'$ – відповідно фактичні і теоретичні частоти.

За спеціальними таблицями визначають ймовірність досліджуваного значення $\chi^2$ в залежності від числа ступенів вільності. Число ступенів вільності визначається за формулою: k = m – r, m – число груп; r – число обмежених зв'язків. Якщо фактичне $\chi^2$ менше табличного ($\chi^2_ф < \chi^2_{табл}$), тоді при прийнятому рівні значимості розходження між фактичними і теоретичними частотами вважаються випадковими, гіпотеза про закон розподілу приймається.

Розглянемо на прикладі застосування критерію $\chi^2$ для доведення гіпотези про правильність вибору типу розподілу господарств за врожайність озимої пшениці.

Таблиця 9.6

Урожайність озимої пшениці господарств району

| Урожайність озимої пшениці, ц/га | Кількість господарств f | x | $\lvert x-\bar{x}\rvert$ | $t=\frac{x-\bar{x}}{\sigma}$ | $f(t)$ | $f'$ | $f-f'$ | $(f-f')^2$ | $\frac{(f-f')^2}{f'}$ |
|---|---|---|---|---|---|---|---|---|---|
| 40-42 | 4 | 41 | 5,72 | 2,22 | 0,0339 | 2 | 2 | 4 | 2,00 |
| 42-44 | 7 | 43 | 3,72 | 1,44 | 0,1415 | 11 | -4 | 16 | 1,45 |
| 44-46 | 28 | 45 | 1,72 | 0,67 | 0,3187 | 25 | 3 | 9 | 0,36 |
| 46-48 | 35 | 47 | 0,28 | 0,11 | 0,3965 | 31 | 4 | 16 | 0,52 |
| 48-50 | 16 | 49 | 2,28 | 0,88 | 0,2709 | 21 | -5 | 25 | 1,19 |
| 50-52 | 6 | 51 | 4,28 | 1,66 | 0,1006 | 8 | -2 | 4 | 0,50 |
| 52-54 | 4 | 53 | 6,28 | 2,43 | 0,0208 | 2 | 2 | 4 | 2,00 |
| Разом: | 100 | x | x | x | x | 100 | x | x | 8,02 |

$$\overline{x} = \frac{\sum xf}{\sum f} = \frac{4672}{100} = 46{,}72 \text{ ц/га}; \ \sigma = \sqrt{6{,}6414} = 2{,}58 \text{ ц/га};$$

$$\sum f \frac{i}{\sigma} = 100 \frac{2}{2{,}58} = 77{,}5 \ ; \ f' = \sum f \frac{i}{\sigma} \cdot f(t) = 77{,}5 \cdot 0{,}0339 = 2 \ ;$$

$$\chi^2_{ф} = \sum \frac{(f - f')^2}{f'} = 8{,}02 \ ; \ k = m - r = 7 - 3 = 4 \ ; \ \chi^2_{табл}(p = 0{,}99) = 13{,}28 \ .$$

Висновок: оскільки фактичний критерій $\chi^2_{ф}$ значно менший $\chi^2_{табл}$ (8,02 < 13,28), то із ймовірністю 0,99 можна вважати доведеним, що тип розподілу вибраний правильно, тобто, що розподіл господарств за врожайністю озимої пшениці є нормальним.

Покажемо на графіку вирівнювання фактичного розподілу господарств району за врожайністю озимої пшениці теоретичною кривою нормального розподілу (мал. 9.3).


Мал. 9.3. Крива розподілу господарств регіону за урожайністю озимої пшениці

На графіку, поданому на мал. 9.3, простежується досить велика близькість фактичних частот розподілу до теоретичних, що також свідчить про правильність вибраного типу розподілу.

Такого ж висновку можна дійти за допомогою і інших критеріїв узгодження.

*Критерій узгодження А.Н.Колмогорова* ($\lambda$) оцінює близькість фактичного розподілу до теоретичного шляхом знаходження величини (D), тобто максимальної різниці нагромаджених (кумулятивних) часток (частот) фактичного і теоретичного розподілів.

Критерій Колмогорова визначається за формулою:

$$\lambda = D \cdot \sqrt{n},$$

де D – абсолютна максимальна різниця кумулятивних часток $D = \max|S_d - S_{d'}|$, або частот $D = \max|S_f - S_{f'}|$ емпіричного і теоретичного розподілів;

n – число спостереження (чисельність одиниць сукупності).

Якщо розподіл задано в частотах, тоді формула матиме вигляд:

$$\lambda = \frac{D}{\sqrt{n}}.$$

Методику розрахунку цього показника розглянемо на прикладі даних табл. 9.6.

Таблиця 9.7

Розрахунок критерію “$\lambda$”

| Номер груп | Нагромаджені частоти | | Відхилення $\lvert S_f - S_{f'}\rvert$ |
|---|---|---|---|
| | емпіричні $S_f$ | теоретичні $S_{f'}$ | |
| 1 | 4 | 2 | 2 |
| 2 | 11 | 13 | 2 |
| 3 | 39 | 38 | 1 |
| 4 | 74 | 69 | 5 |
| 5 | 90 | 90 | 0 |
| 6 | 96 | 98 | 2 |
| 7 | 100 | 100 | 0 |

$$\lambda = \frac{D}{\sqrt{100}} = \frac{5}{10} 0{,}5 .$$

По спеціальній таблиці ймовірностей для критерію узгодження “$\lambda$” знаходимо, що значенню $\lambda = 0{,}5$ відповідає ймовірність 0,9639, це означає, що з ймовірністю 0,9639 можна стверджувати про нормальний розподіл господарств за врожайністю озимої пшениці.

Критерій узгодження А.Н.Колмогорова застосовують ще й для визначення належності двох вибіркових сукупностей “$n_1$” і “$n_2$” до однієї генеральної сукупності, а також для перевірки незалежності двох і більше ознак сукупності.

*Критерій узгодженості В.І.Романовського* (R) також використовується для оцінки наближення фактичного розподілу до теоретичного. Він визначається за формулою:

$$R = \frac{\chi^2 - k}{\sqrt{2k}}.$$

Скориставшись розрахунками прикладу за даними табл. 9.6, визначимо критерій узгодження Романовського.

$$R = \frac{\chi^2 - k}{\sqrt{2k}} = \frac{8,02 - 4}{\sqrt{2 \cdot 4}} = \frac{4,02}{2,8284} = 1,42 < 3.$$

Якщо при дослідженні наближення фактичного розподілу до теоретичного величина цього виразу менша трьох (1,42 < 3), це дає підставу для стверджування про можливість розподілу за законом даного розподілу. Тобто розподіл господарств за врожайністю озимої пшениці є нормальним.

Критерій узгодження Б.С. Ястремського (L) використовується для прямої відповіді на питання про міру розбіжності між фактичним і теоретичним розподілом. Для визначення критерію Ястремського використовується критерій Пірсона. В загальному вигляді критерій Ястремського визначають за наступною формулою:

$$L = \frac{\chi^2 - n}{\sqrt{2n + 4Q}},$$

де $Q = \frac{(f_i - f_i')^2}{f_i(1 - P_i)}$ – при кількості груп, менше 20 ($n < 20$), дорівнює 0,6.

За даними попередніх прикладів покажемо розрахунок цього показника.

$$L = \frac{\chi^2 - n}{\sqrt{2n + 4Q}} = \frac{8,02 - 7}{\sqrt{2 \cdot 7 + 4 \cdot 0,6}} = \frac{1,02}{4,02} = 0,25.$$

Так як величина $L < 3$ ($0,25 < 3$), то із ймовірністю 0,997 можна стверджувати, що розподіл господарств за врожайністю озимої пшениці є нормальним.

Всі розглянуті нами критерії узгодження дають загальну оцінку ступеня відхилення емпіричного розподілу від нормального, але не визначають його характеру, а тому при суттєвих їх відхиленнях аналіз розподілу доцільно доповнювати характеристиками асиметрії і ексцесу.

Таким чином, для перевірки висунутої гіпотези про відповідність чи невідповідність теоретичного закону розподілу емпіричному можна використати любий із наведених критеріїв, які забезпечують дослідження законів розподілу з різною точністю, надійністю і трудоємкістю.

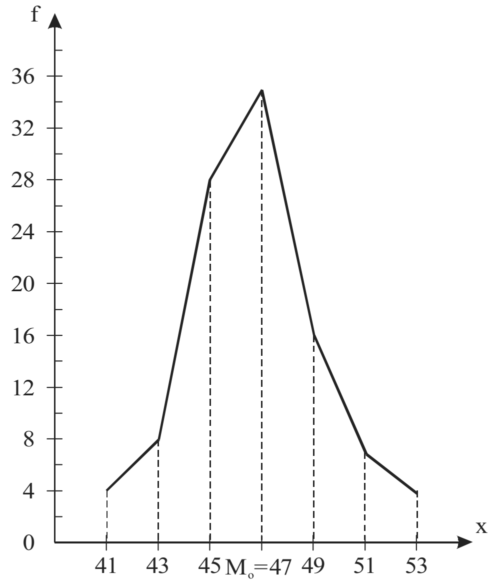
## 9.4. Графічне зображення рядів розподілу

Для наглядного зображення варіаційного ряду використовують такі графіки як: полігон, гістограма, кумулята, огіва, крива концентрації (Лоренца), крива Парето, показникова крива, антимода та ін.

Полігон – це графічне зображення варіаційного ряду в прямокутній системі координат, при котрому величина ознаки відкладається на осі абсцис, частоти або частки (щільність розподілу) – на осі ординат.

Частіше полігон застосовується для зображення дискретного варіаційного ряду, однак може використовуватись і для інтервального ряду.

З метою ілюстрації деяких графіків рядів розподілу використовуємо дані табл. 9.6. Так, за цими даними графік полігону матиме вигляд:

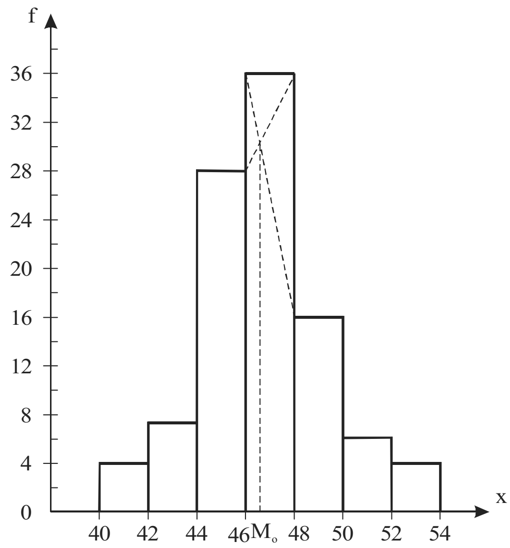



Мал. 9.4. Полігон розподілу господарств регіону за урожайністю озимої пшениці

За допомогою полігона можна визначити також моду, для цього з його вершини опускають перпендикуляр на вісь абсцис, а точка їх перетину і буде модальною величиною. В нашому прикладі мода відповідає врожайності 47 центнерів з гектара озимої пшениці.

*Гістограма* – це графічне зображення інтервального варіаційного ряду у вигляді прямокутників відповідної висоти, основи яких знаходяться, як певні відрізки, на осі абсцис, і, які відповідають інтервалам зміни ознаки. Висоти прямокутників пропорційні при рівності інтервалів частотам або часткам інтервалів, а при нерівності – абсолютним або відносним щільностям, що робить гістограму незалежною від ширини інтервалів.

Гістограму ряду розподілу за даними табл. 9.6 наведемо на мал. 9.5.




Мал. 9.5. Гістограма розподілу господарств регіону за урожайністю озимої пшениці

Для графічного визначення моди за допомогою гістограми пунктирними лініями з'єднують верхні кути модального інтервалу і стовпчиків, що прилягають до нього. Модою буде перетин осі абсцис перпендикуляром, опущеним з точки зіткнення цих прямих, як показано на мал. 9.5. Цей метод більш коректно оцінює модальну величину, так як враховує передмодальну і післямодальну частоти, тобто за гістограмою мода дорівнює 46,67 центнери з гектара. Графічне визначення моди можливе тільки у варіаційних рядах розподілу з рівними інтервалами.

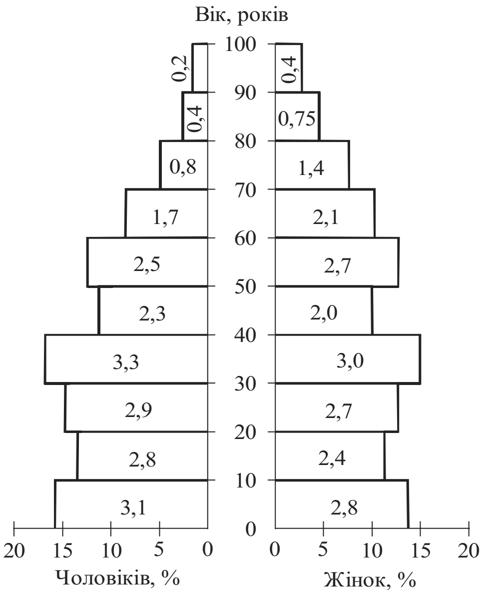
*Двобічну гістограму* використовують для графічного зображення комбінаційних розподілів, де одна групувальна ознака приймає два значення, а друга – кількість груп з рівними інтервалами.

Для її побудови по осі ординат відкладають межі інтервалів за першою ознакою, а на осі абсцис по обидві сторони від осі ординат – однакові відрізки для часток або частот розподілу за другою групувальною ознакою.

Двобічні діаграми широко використовуються при дослідженні особливостей розподілу демографічних явищ, де за характерною формою графічного зображення їх називають пірамідою. Відхилення від піраміди свідчить про наявність певних особливостей у розподілі сукупності.

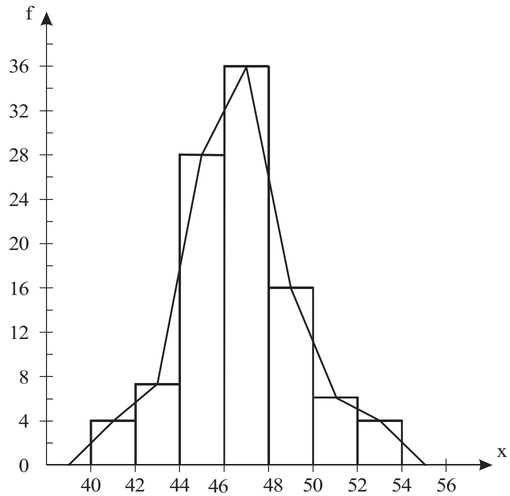
Покажемо побудову двобічної діаграми за даними чисельності населення одного з регіонів України (дані умовні) за віком і статтю (мал. 9.6).

Для графічного порівняння двох і більше інтервальних рядів розподілу на гістограму накладають полігон розподілу, з'єднавши середини вершин прямокутників гістограми ламаною лінією.




Мал. 9.6. Статево-вікова піраміда розподілу населення регіону України

Для забезпечення відповідності площ полігона і гістограми, потрібно по обидві сторони останньої відкласти додаткові відрізки завширшки крайнім інтервалам, і з'єднати їх середини з серединами прямокутників гістограми (мал. 9.7).



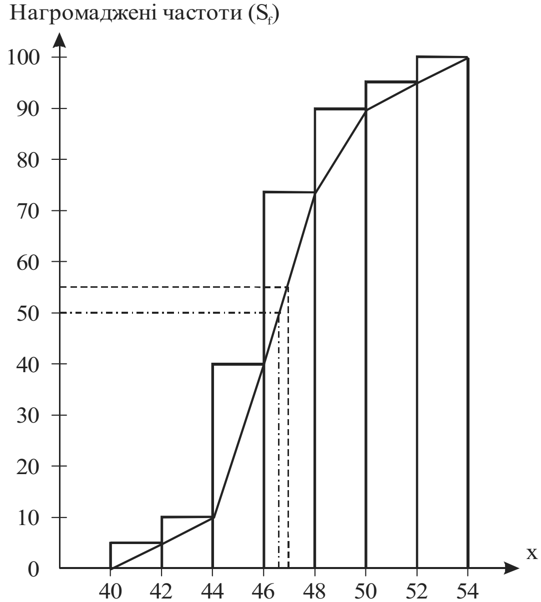

Мал.. 9.7. Полігон розподілу господарств регіону за урожайністю озимої пшениці

*Кумулятивні діаграми* використовують для графічного порівняння двох або більше варіаційних розподілів з рівними чи нерівними інтервалами. Для їх побудови використовують прямокутну систему координат де на осі абсцис відкладають відрізки інтервалів групувань, а на осі ординат – нагромаджені частоти або частки. Висота прямокутників відповідає кумулятивним частотам або часткам конкретних інтервалів ряду розподілу.

При побудові кумуляти інтервальної ознаки нижньої межі першого інтервалу відповідає частота, яка дорівнює нулю, а верхній – частота першого інтервалу. Верхній границі другого і наступних інтервалів відповідають їх нагромаджені частоти, а останньою інтервалу – сума всіх частот (графік кумуляти побудований за даними табл. 9.7, приведений на мал. 9.8).

На основі кумулятивної кривої розподілу визначають, скільки одиниць сукупності, або яка частка не перевищує певного значення групувальної ознаки. На мал. 9.8 штрихові лінії показують, що із 100 господарств регіону 55 мали врожайність озимої пшениці менше 47 центнерів з гектара.

Для графічного визначення медіани з точки на осі ординат, яка відповідає півсумі нагромаджених частот, проводять пряму (штрих-пунктирну) лінію (мал. 9.8), паралельну осі абсцис. Перпендикуляр опущений з точки перетину цієї прямої з кумулятою на вісь абсцис вкаже на медіану. В нашому прикладі $M_e$ = 46,67 центнерів з гектара.

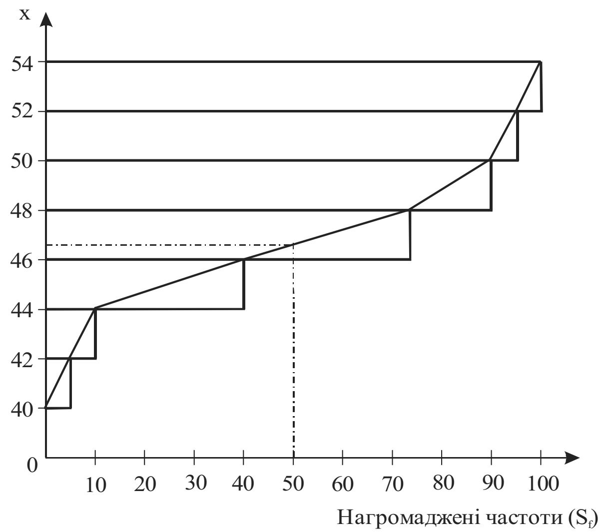



Мал. 9.8. Кумулятивна гістограма розподілу господарств регіону за урожайністю озимої пшениці

При побудові кумуляти дискретної ознаки на вісь абсцис наносять значення варіантів, а на ординату – нагромаджені підсумки частот або часток. З'єднавши вершини ординат прямими лініями отримуємо кумулятивну криву, яку ще називають кумулятивним полігоном.

Різновидом кумулятивного розподілу варіаційного ряду є *огіва*. Вона являє собою дзеркальне відображення кумуляти розподілу. При її побудові на осі абсцис відкладають нагромаджені частоти або частки, а на осі ординат – межі інтервалів варіаційного ряду розподілу. Побудову кривої огіви розподілу регіону за врожайністю озимої пшениці покажемо на мал. 9.9. Штрих-пунктирною лінією визначена медіана.

*Крива концентрації Лоренца*, як різновид кумулятивної діаграми, використовується в якості показника ступеня оцінки концентрації розподілу частот варіаційного ряду. Вона відображає також ступінь рівномірності розподілу частот.




Мал. 9.9. Огіва розподілу господарств регіону за урожайністю озимої пшениці

Криву Лоренца будують у прямокутній системі координат, де на осі абсцис і ординат наносять однакові масштаби шкали від нуля до ста. Така діаграма має форму квадрата, на її осі абсцис відкладають значення нагромаджених часток, які характеризують розподіл сукупності за групувальною ознакою, а на осі ординат – нагромаджені значення обсягу даної ознаки.

Першою на осі "х" відкладається точка (одиниця або група одиниць), яка має найвищий показник за обсягом досліджуваного явища, а потім кумулятивні підсумки в порядку рангів їх одиниць. Із кожної кумулятивної точки на осі абсцис відкладається перпендикуляр, висота якого відповідає кумулятивному підсумку обсягу досліджуваного явища. Послідовно з'єднавши вершини цих перпендикулярів отримаємо криву концентрації Лоренца.

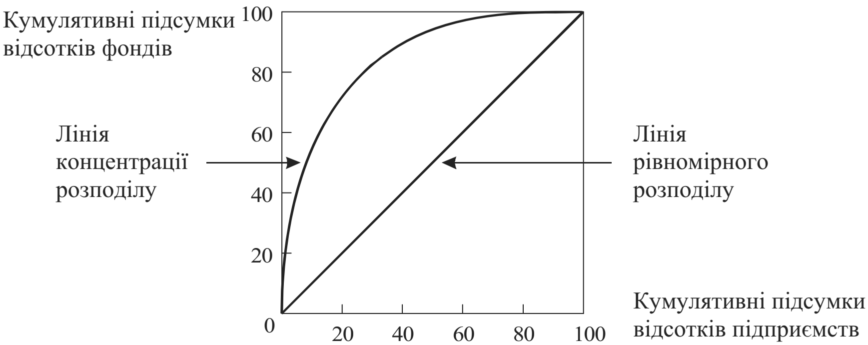
При рівномірному розподілі елементів сукупності за групувальною ознакою повинна зберегтись рівність $x = y$ і на графіку ця залежність виразиться прямою лінією, яка проходить через початок координат під кутом в $45^{o}$. Тобто це буде діагональ квадрата, яка з'єднує нижній лівий кут з верхнім правим на якій будується крива Лоренца. Ця діагональ розглядається як лінія рівномірного розподілу. Всякі відхилення від неї свідчать про нерівномірність розподілу. Чим більше крива Лоренца буде відхиляться від діагоналі, тим більша нерівномірність розподілу і вища концентрація обсягу досліджуваного явища.

Розглянемо методику побудови кривої концентрації Лоренца на прикладі розподілу підприємств за вартістю основних виробничих фондів (табл. 9.8, дані умовні).

Таблиця 9.8.

| Групи підприємств за середньорічною вартістю основних виробничих фондів, тис. грн. | В відсотках | | | |
|---|---|---|---|---|
| | число підпри-ємств, x | основні виробничі фонди, y | нагромаджені підсумки | |
| | | | число підпри-ємств, x | основні виробничі фонди, y |
| 100000 і більше | 1,7 | 27,0 | 1,7 | 27,0 |
| 50000 – 10000 | 2,0 | 19,0 | 3,7 | 46,0 |
| 10000 – 50000 | 13,0 | 20,0 | 16,7 | 66,0 |
| 5000 – 10000 | 12,8 | 17,0 | 29,5 | 83,0 |
| 1000 – 5000 | 37,0 | 8,0 | 66,5 | 91,0 |
| 500 – 1000 | 12,0 | 7,9 | 78,5 | 98,6 |
| 300 – 500 | 12,6 | 0,6 | 91,1 | 99,5 |
| 100 – 300 | 5,0 | 0,3 | 96,1 | 99,8 |
| менше 100 | 3,9 | 0,2 | 100,0 | 100,0 |
| Разом: | 100,0 | 100,0 | x | x |

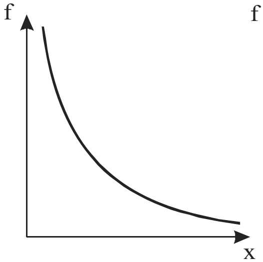
Для побудови кривої концентрації Лоренца на квадраті 100x100 (мал. 9.10) знаходимо точки з координатами, які відповідають нагромадженим підсумкам проценту числа підприємств і обсягу основних фондів в них. Послідовно з’єднуємо всі точки плавною лінією, ми отримуємо криву концентрації основних виробничих фондів на підприємствах.



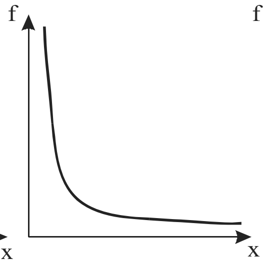

Мал. 9.10.Крива концентрації основних виробничих фондів на підприємствах

Графік Лоренца можна будувати не тільки за спадаючими, але й зростаючими значеннями досліджуваного показника, але тоді крива концентрації буде не випуклою, а ввігнутою, розміщеною під діагоналлю.

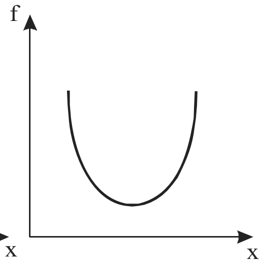
До інших видів графічного зображення рядів розподілу відносять: показникові криву, криву Парето і антимоду. Ці графіки мають вигляд:




Мал. 9.11.
Показникова крива




Мал. 9.12.
Крива Парето




Мал. 9.13.
Модель (Портера) Антимода

Прикладом показникового розподілу є термін служби товарів народного споживання, які вибувають із-за аварійності – довговічність посуду в підприємствах громадського харчування, тривалість телефонних викликів, час між простоями верстатів і т.д.

Розподіл Парето показує розподіл доходів у відповідності з їх величиною.

Антимода – це абсциса нижньої точки, розміщеної в центральній частині U– подібного розподілу досліджуваної ознаки. Як видно із мал. 9.13, обидві сторони антимоди частоти зростають.

### *Контрольні запитання*

1. Поняття про ряди розподілу та їх використання у варіаційному аналізі.
2. Види рядів розподілу.
3. Щільність розподілу та інтерполяція в рядах розподілу.
4. Форми рядів розподілу.
5. Характеристика форм розподілу.
6. Симетричні і асиметричні розподіли, статистичні методи їх оцінки.
7. Поняття про ексцес.
8. Моменти розподілу та їх використання в статистичному дослідженні.

9. Показники концентрації рядів розподілу.
10. Криві розподілу та їх типи.
11. Крива нормального розподілу і її використання у варіаційному аналізі.
12. Способи перевірки гіпотез.
13. Критерій узгодження Пірсона.
14. Критерій узгодження Колмогорова.
15. Критерій узгодження Романовського.
16. Критерій узгодження Ямстремського.
17. Графічне зображення рядів розподілу.
18. Поняття про полігон і гістограму та способи їх побудови.
19. Використання двобічної гістограми.
20. Кумулятивні діаграми розподілу.
21. Крива концентрації Лоренца.
22. Інші види графічного зображення рядів розподілу.

# Розділ 10. Статистичні методи вивчення взаємозв'язків

## 10.1. Зв'язки суспільних явищ і завдання їх статистичного вивчення

Одним з найбільш загальних законів об'єктивного світу є закон зв'язку і залежності між явищами суспільного життя. Ці явища найбільш складні, оскільки вони формуються під дією багаточисельних, різноманітних і взаємозв'язаних факторів.

Усі явища суспільного життя існують не ізольовано, вони органічно зв'язані між собою, залежать одні від одних, обумовлюють одні одних і знаходяться в постійному русі і розвитку. Розкриваючи взаємозв'язки і взаємозалежності між явищами можна пізнати їх суть і закони розвитку. Тому вивчення взаємозв'язків є основним завданням всякого статистичного аналізу.

Причинна залежність являє собою головну форму закономірних зв'язків, які діють на наслідок в певних умовах місця і часу. Отже, для виникнення наслідку потрібні і причини, і умови, тобто відповідні фактори.

Суспільні явища або окремі їх ознаки, які впливають на інші і обумовлюють їх зміну, називаються *факторними*, а суспільні явища або окремі їх ознаки, які змінюються під впливом факторних, називаються *результативними*.

За характером залежності явищ розрізняють функціональні і кореляційні зв'язки.

*Функціональним* називається зв'язок, при якому певному значенню факторної ознаки завжди відповідає одне значення результативної ознаки. Функціональні зв'язки характеризуються повною відповідністю між причиною і наслідком. Тому функціональна залежність виражається точною математичною формулою. Наприклад, зв'язок між довжиною кола і радіусом круга описується формулою: $l = 2\pi r$ . Існують функціональні зв'язки, в яких результативна ознака є функцією кількох факторних ознак. Для прикладу можна взяти зв'язок між площею трикутника, його основою і висотою.

Функціональні залежності вивчаються точними науками, такими як математика, фізика, хімія та інші. Вони досить рідко використовуються для дослідження суспільних явищ.

*Кореляційним* називається зв'язок, при якому кожному значенню факторної ознаки, відповідає декілька значень результативної ознаки. В кореляційних зв'язках між причиною і наслідком немає повної відповідності, а спостерігається лише певне співвідношення. На відміну від функціонального зв'язку кореляційний зв'язок проявляється не в кожному окремому випадку, а в середньому при великому числі спостережень. Кореляційні зв'язки найбільше використовуються при дослідженні суспільних явищ. Прикладами кореляційних зв'язків можуть бути: зв'язок між заробітною платою робітників і стажем їх роботи, собівартістю продукції і продуктивністю праці, врожайністю і розміром внесених в ґрунт добрив і т.д.

За напрямком розрізняють зв'язки прямі і обернені.

*Прямий зв'язок* – це такий зв'язок, коли із зростанням факторної ознаки, результативна ознака також зростає.

При *оберненому зв'язку* із збільшенням факторної ознаки результативна зменшується або, навпаки, із зменшенням факторної ознаки, результативна зростає. Наприклад, ріст продуктивності праці приводить до зменшення собівартості одиниці продукції, скорочення термінів збирання зернових приводить до зростання врожайності і т.д.

За своїм аналітичним виразом (за формою) зв'язок ділиться на прямолінійний і криволінійний.

При прямолінійній кореляційній залежності рівним змінам середніх значень факторної ознаки відповідають приблизно рівні зміни середніх значень результативної ознаки.

При криволінійній кореляційній залежності рівним змінам середніх значень факторної ознаки відповідають нерівні зміни середніх значень результативної ознаки.

Статистичне вивчення взаємозв'язків розв'язує наступні завдання:

1) визначаються форми зв'язку;
2) вимірюється тіснота (сила) зв'язку;
3) виявляється вплив окремих факторів на результативну ознаку.

## 10.2 Загальні методи вивчення зв'язків

Зв'язки і залежності суспільних явищ вивчаються різними методами, які дають уявлення про їх наявність і характер. До цих методів відносять: балансовий метод, метод порівняння паралельних рядів, графічний метод, метод аналітичних групувань, індексний метод, кореляційно-регресійний метод та інші методи математичної статистики.

Одним із поширених методів статистичного вивчення зв'язків суспільних явищ є балансовий метод, як прийом аналізу зв'язків і пропорцій в народному господарстві.

*Статистичний баланс* являє собою систему показників, яка складається із двох сум абсолютних величин, зв'язаних між собою знаком рівності:

$$а + б = в + г.$$

Співставляючи баланси зв'язують в єдину систему абсолютні рівні, які характеризують рух ресурсів. Одним з таких балансів є баланс матеріальних ресурсів на якому-небудь підприємстві. Цю балансову ув'язку можна зобразити через балансове рівняння:

залишок на початок + поступлення = видатки + залишок на кінець.

Наведена балансова рівність характеризує єдиний процес руху матеріальних ресурсів і показує взаємозв'язок і пропорції окремих елементів цього процесу. Між поступленнями і видатками повинно витримуватись певне співвідношення, а якщо воно порушується, тоді різко змінюється питома вага запасів на кінець періоду в порівнянні з початковим періодом. Отже, нормальний хід процесу вимагає дотримання певної пропорціональності між всіма елементами балансу.

Можливості в аналізі взаємозв'язків і пропорцій балансу значно розширяться, якщо поступлення розбити за джерелами (постачальниками), а видатки – за призначенням (покупцями). В цьому випадку баланс покаже взаємозв'язок як між поступленнями, видатками і залишками в межах підприємства, так і між даним підприємством з іншими підприємствами. Тобто, з підприємствами які постачають дане підприємство матеріальними ресурсами, а також з підприємствами-споживачами його ресурсів. З допомогою таких балансів вивчають рух (оборот) робочої сили, грошових засобів, основних фондів і т.д.

Розглянемо найпростіший баланс руху окремих товарів в торговому підприємстві (табл. 10.1).

Таблиця 10.1

Баланс руху товарів за I квартал, тис. грн.

| Товар | Залишок на 1.01 | Поступило в I кварталі | Разом в приході (гр.1+гр.2) | Продано оптом | Продано в роздріб | Залишок на 1.04 | Разом в розході (гр.4+гр.5 +гр.6) |
|---|---|---|---|---|---|---|---|
| А | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| а | 8 | 24 | 32 | 5 | 20 | 7 | 32 |
| б | 11 | 35 | 46 | 10 | 30 | 6 | 46 |
| в | 6 | 16 | 22 | 2 | 13 | 7 | 22 |
| Разом | 25 | 75 | 100 | 17 | 63 | 20 | 100 |

В наведеній таблиці по кожному товару окремо і по всіх товарах разом показано їх рух за квартал. Дані показники в балансі розглядаються у взаємозв'язку і взаємозалежності. Так, товару "а" поступило за квартал менше ніж продано, в результаті зменшився залишок на 1.04 порівняно з 1.01. Аналогічне співвідношення склалося по товару "б", і зовсім інше – по товару "в", тут поступлення перевищило реалізацію, внаслідок чого залишок на кінець кварталу більший залишку на початок..

З балансу видно пропорції розподілу реалізації товарів оптом і в роздріб, а також співвідношення залишків до товарообороту. Все це дає можливість аналізувати процес руху товарів в цілому, а також пропорції його окремих частин.

Балансова форма дає можливість здійснювати взаємний контроль даних, а також розраховувати відсутні показники. Балансова ув'язка дозволяє виявити неточності окремих показників, уточнити їх.

Більш широкі можливості аналізу взаємозв'язків відкриваються при складанні балансів руху матеріальних, трудових і фінансових ресурсів для сукупності підприємств, галузі, економічного району і країни в цілому. В даному випадку баланси дають можливість виявити взаємозв'язки в утворенні і розподілі ресурсів між підприємствами, районами і галузями народного господарства, проаналізувати пропорції руху ресурсів, міжгалузеві і міжрайонні зв'язки.

Аналіз таких балансів дозволяє глибоко дослідити закономірності процесу відтворення для підвищення наукового рівня планування і прогнозування соціально-економічних явищ і процесів.

До числа поширених методів аналізу взаємозв'язків відноситься і метод порівняння паралельних рядів.

Суть *методу порівняння паралельних рядів* полягає в тому, що отримані в результаті групування і лічильної обробки матеріали статистичного спостереження розташовуються рангованими паралельними рядами за факторною ознакою. Паралельно записуються значення результативної ознаки. Це дає можливість, порівнюючи їх, простежити співвідношення, виявити існування зв'язку і його напрямок.

Покажемо застосування цього методу на прикладі. Нехай маємо такі дані про роботу десяти однотипних підприємств (табл. 10.2).

Таблиця 10.2

| Номер підприємства | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Вартість основних виробничих фондів, млн. грн. | 5,3 | 6,4 | 7,9 | 8,3 | 9,2 | 10,1 | 12,5 | 13,0 | 14,6 | 15,7 |
| Випуск продукції, млн. грн. | 5,8 | 7,6 | 8,7 | 9,1 | 11,9 | 12,3 | 13,8 | 14,0 | 15,2 | 17,6 |

З таблиці видно, що із збільшенням вартості основних виробничих фондів випуск продукції зростає.

На основі порівняння паралельних рядів вираховують напрямок і силу зв'язку за допомогою коефіцієнта Фехнера і кореляції рангів.

Використаємо дані табл. 10.2 для ілюстрації розрахунку коефіцієнта Фехнера.

Таблиця 10.3

Взаємозв'язок між вартістю основних виробничих фондів і випуском однорідної продукції по 10 підприємствах.

| Номер підприємства | Вартість основних виробничих фондів, млн. грн. x | Випуск продукції, млн. грн. y | $x-\bar{x}$ | $y-\bar{y}$ | С або Н |
|---|---|---|---|---|---|
| 1 | 5,3 | 5,8 | - | - | С |
| 2 | 6,4 | 7,6 | - | - | С |
| 3 | 7,9 | 8,7 | - | - | С |
| 4 | 8,3 | 9,1 | - | - | С |
| 5 | 9,2 | 11,9 | - | + | Н |
| 6 | 10,1 | 12,3 | - | + | Н |
| 7 | 12,5 | 13,8 | + | + | С |
| 8 | 13,0 | 14,0 | + | + | С |
| 9 | 14,6 | 15,2 | + | + | С |
| 10 | 15,7 | 17,6 | + | + | С |
| Разом | 103,0 | 116,0 | x | x | x |

*Коефіцієнт Фехнера* оцінює силу зв'язку на основі порівняння знаків відхилень значень варіантів від їх середньої по кожній ознаці.

Визначимо середні: $\bar{x} = \frac{\sum x}{n} = \frac{103}{10} = 10,3$ млн. грн.

$$\bar{y} = \frac{\sum y}{n} = \frac{116}{10} = 11,6 \text{ млн. грн.}$$

Знак мінус означає, що значення ознаки менше середньої, а знак плюс – що більше середньої. Співпадання знаків за обома ознаками означає узгоджену варіацію, неспівпадання – порушення такої узгодженості. За цим принципом побудований коефіцієнт Фехнера, який має вигляд:

$$K_Ф = \frac{\sum C - \sum H}{\sum C + \sum H} = \frac{8-2}{8+2} = \frac{6}{10} = 0{,}60\,,$$

де $\sum C$ – сума знаків, які співпали по обох рядах;

$\sum H$ – сума знаків, які не співпали.

В даному прикладі із 10 спостережень, у 8 випадках знаки співпали і в 2 випадках не співпали.

Коефіцієнт Фехнера коливається в межах від + 1 до – 1. При наближенні цього коефіцієнта до + 1 спостерігається пряма і сильна узгодженість, при – 1 будемо мати сильну але обернену узгодженість. При нулю, узгодженість між досліджуваними ознаками відсутня.

В нашому прикладі $К_Ф$ показує, що між вартістю основних виробничих фондів і випуском продукції існує прямий і досить тісний зв'язок.

Більш точно оцінює силу зв'язку *коефіцієнт кореляції рангів*. Цей коефіцієнт враховує узгодженість рангів, які займають окремі одиниці сукупності по кожній із двох досліджуваних ознак.

Сукупність рангується за факторною ознакою в порядку зростання і проставляються відповідні ранги. Паралельно проставляються ранги тих же одиниць сукупності, які вони б зайняли в рангованому ряду за результативною ознакою.

Коефіцієнт кореляції рангів запропонований американським вченим К. Спірменом має вигляд:

$$\rho = 1 - \frac{6\sum d^2}{n\left(n^2 - 1\right)},$$

де $\rho$ (грецька буква "ро") – коефіцієнт кореляції рангів К. Спірмена;

$d^2$ – квадрат різниці між величинами рангів в порівняльних рядах;

n – число рангів.

Існує правило, що для варіантів, які повторюються, ранг визначається як середня арифметична відповідних рангів, наприклад, ранг однакових величин, які займають 4 і 5 місця, дорівнює 4,5.

Коефіцієнт рангової кореляції може приймати значення в межах: $-1 \le \rho \le +1$.

Коли ранги факторної ознаки $R_x$ повністю співпадають з рангами результативної ознаки $R_y$, тоді кожне значення $R_x = R_y$ і $\sum d^2 = 0$. В цьому випадку можна судити про майже повний прямий зв'язок, $\rho = 1$.

Якщо ранги розташувались строго в протилежному напрямку, тобто першому рангу факторної ознаки відповідає останній ранг результативної ознаки і т.д., тоді спостерігається повна обернена кореляція рангів. В даному випадку $\frac{6\sum d^2}{n(n^2-1)} = 2$, а коефіцієнт ρ = – 1.

Коли $\frac{6\sum d^2}{n(n^2-1)} = 1$, кореляція рангів відсутня і ρ = 0.

Потрібно мати на увазі, що цей емпіричний показник менш точний в порівнянні з лінійним коефіцієнтом кореляції і емпіричним кореляційним відношенням, а тому коли він отримує крайні значення ± 1 або 0, то це не означає, що існує функціональний зв'язок, або залежність абсолютно відсутня. В усіх інших випадках, коли коефіцієнт рангової кореляції не приймає крайніх значень він інтерпретується так же, як і коефіцієнт лінійної кореляції і володіє такими ж самими властивостями.

Для розрахунку коефіцієнта кореляції рангів можна не знати кількісних значень варіантів ознаки, а достатньо лише знати їх ранги.

Розглянемо приклад. Потрібно визначити силу зв'язку між вартістю основних виробничих фондів і виробітком продукції на одного робітника за такими даними:

Таблиця 10.4

Показники роботи 10 підприємств і розрахунок зв'язку між ними

| Номер підприємства | Вартість основних виробничих фондів, тис. грн. x | Виробіток одного робітника, тис. грн. y | Ранги | | Різниця рангів | |
|---|---|---|---|---|---|---|
| | | | $R_x$ | $R_y$ | $d=R_x-R_y$ | $d^2$ |
| 1 | 2348 | 20 | 1 | 2 | -1 | 1 |
| 2 | 2654 | 32 | 2 | 4 | -2 | 4 |
| 3 | 2780 | 41 | 3 | 7 | -4 | 16 |
| 4 | 2891 | 43 | 4 | 8 | -4 | 16 |
| 5 | 3125 | 18 | 5 | 1 | +4 | 16 |
| 6 | 3240 | 24 | 6 | 3 | +3 | 9 |
| 7 | 3915 | 37 | 7 | 5 | +2 | 4 |
| 8 | 4000 | 39 | 8 | 6 | +2 | 4 |
| 9 | 4137 | 43 | 9 | 9 | 0 | 0 |
| 10 | 5199 | 45 | 10 | 10 | 0 | 0 |
| Разом | x | x | x | x | x | 70 |

$$\rho = 1 - \frac{6\sum d^2}{n\left(n^2 - 1\right)} = 1 - \frac{6 \cdot 70}{10\left(100 - 1\right)} = 1 - 0{,}424 = 0{,}576\,.$$

Коефіцієнт кореляції рангів К. Спірмена вказує на помітний прямий зв'язок між вартістю основних виробничих фондів і виробітком продукції на одного робітника.

Англійський статистик Кендел для визначення тісноти зв'язку між корельованими ознаками запропонував свою формулу коефіцієнта кореляції рангів:

$$\tau = \frac{S}{\frac{1}{2}n\left(n-1\right)}, \text{ або } \tau = \frac{2S}{n\left(n-1\right)},$$

де $\tau$ (грецька буква "тау") – коефіцієнт кореляції рангів Кендела;
$S$ – фактична сума балів;
$n$ – число рангів.

Величина S являє собою різницю двох складових:

$$S = \sum S_1 - \sum S_2\,,$$

де $S_1$ – число рангів, які перевищують номер рангу, записаного в рангах за результативною ознакою $R_y$;
$S_2$ – число рангів менших $R_y$ в наступних записах.

Для визначення цього показника складемо додаткову таблицю і розрахуємо величини $S_1$ і $S_2$ за даними попереднього прикладу.

Таблиця 10.5

| Підприємства | $R_y$ | $S_1$ | $S_2$ |
|---|---|---|---|
| 1 | 2 | 8 | 1 |
| 2 | 4 | 6 | 2 |
| 3 | 7 | 3 | 4 |
| 4 | 8 | 2 | 4 |
| 5 | 1 | 5 | 0 |
| 6 | 9 | 4 | 0 |
| 7 | 5 | 3 | 0 |
| 8 | 6 | 2 | 0 |
| 9 | 9 | 1 | 0 |
| 10 | 10 | 0 | 0 |
| Разом | x | 34 | 11 |

Скориставшись даними таблиці 10.5 знаходимо величину $S = \sum S_1 - \sum S_2 = 34 - 11 = 23$.

Звідси $\tau = \dfrac{S}{\frac{1}{2}n(n-1)} = \dfrac{23}{\frac{1}{2}10(10-1)} = \dfrac{23}{45} = 0{,}511$.

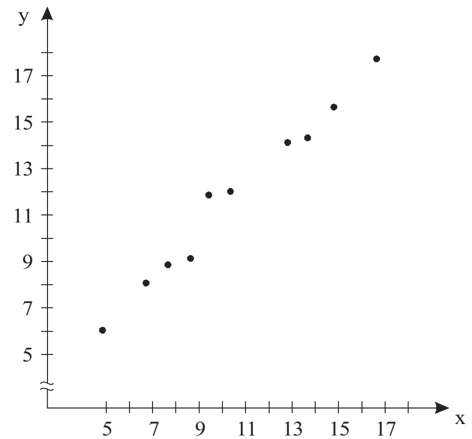
Таким чином коефіцієнт кореляції рангів Кендела оцінює зв'язок між даними ознаками більш обережно, ніж коефіцієнт Спірмена.

Між коефіцієнтами кореляції рангів існує співвідношення:

$$\rho = \frac{3}{2}\tau .$$

Достоїнство коефіцієнтів рангової кореляції заключається в простоті методики розрахунку і в швидкості оцінки взаємозв'язку між явищами.

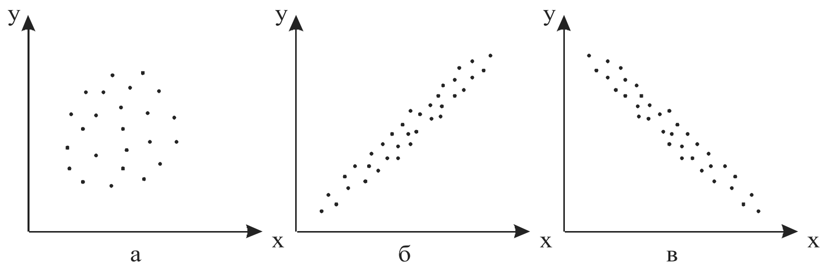
*Графічний метод виявлення кореляційної залежності* заключається в зображенні статистичних характеристик, отриманих в результаті зведення і обробки вихідної інформації на графіку, яке наочно покаже форму зв'язку між досліджуваними ознаками, та його напрямок. Так, виявлений в попередньому прикладі (табл. 10.3) за допомогою порівняння паралельних рядів зв'язок між вартістю основних виробничих фондів (x) і випуском продукції (y) можна наочно уявити, якщо побудувати графік, відклавши на осі абсцис значення ознаки X, а на осі ординат – значення ознаки Y. Нанісши на графіку точки, які відповідають значенням X і Y, отримаємо кореляційне поле, де за характером розміщення точок можемо судити про напрямок і силу зв'язку (мал. 10.1).




Мал. 10.1 Кореляційне поле зв'язку між вартістю основних виробничих фондів і випуском продукції.

Якщо точки розташовані хаотично по всьому полю, це говорить про відсутність залежності між двома ознаками (мал. 10.2, а); якщо вони сконцентровані навколо осі, яка йде від нижнього лівого кута до верхнього правого (мал. 10.2, б) – це пряма залежність між досліджуваними ознаками; якщо точки будуть сконцентровані навколо осі, яка проляже від верхнього лівого кута до нижнього правого (мал. 10.2, в) – маємо обернену залежність.

В нашому прикладі кореляційне поле ясно показує, що між вартістю основних виробничих фондів і випуском продукції існує пряма кореляційна залежність.




Мал. 10.2 Розподіл кореляційного поля при різних видах залежності.

*Метод статистичних групувань*, як прийом виявлення кореляційної залежності, відноситься до числа найважливіших прийомів дослідження взаємозв'язків. Для виявлення залежності між ознаками за допомогою цього методу матеріал статистичного спостереження групується за факторною ознакою, і для кожної групи вираховуються середні значення як факторної так і результативної ознак. Порівнюючи зміни середніх значень результативної ознаки в міру зміни середніх значень факторної ознаки, виявляють характер зв'язку між ними.

Статистичні групування, проведені з метою виявлення і аналізу взаємозв'язків між ознаками суспільних явищ, називаються *аналітичними*. За їх допомогою розкривається вплив одних ознак на інші.

Проблема групувань зводиться до утворення оптимального числа груп для кожного конкретного випадку, щоб групові середні перестали носити випадковий характер і в той же час групувальна ознака проявила себе в повній мірі.

Звернемось до прикладу. Нехай ми провели аналітичне групування 20 робітників-відрядників за стажем роботи з метою виявлення його впливу на місячну заробітну плату, утворивши за факторною ознакою 5 груп з рівними інтервалами. Дані групувань викладені в аналітичній таблиці 10.6.

Таблиця 10.6

Залежність місячної заробітної плати робітників-відрядників від стажу роботи.

| Групи робітників-відрядників за стажем роботи, років | Число робітників-відрядників | Середні рівні | |
|---|---|---|---|
| | | стажу роботи, років | місячної заробітної плати, грн. |
| I 1 – 4 | 3 | 2,07 | 154,67 |
| II 4 – 7 | 6 | 5,40 | 171,83 |
| III 7 – 10 | 5 | 8,44 | 191,00 |
| IV 10 – 13 | 4 | 10,92 | 193,50 |
| V 13 – 16 | 2 | 15,00 | 214,00 |
| Разом | 20 | 7,72 | 182,60 |

Як видно із таблиці, середній місячний заробіток робітників-відрядників збільшується разом із ростом стажу їх роботи. Це свідчить про пряму залежність заробітної плати робітників-відрядників від стажу їх роботи.

Групування дозволяють також виявити одночасний вплив декількох факторів на результативну ознаку. Для цього проводять комбінаційні групування, дані якого викладають в комбінаційних таблицях. Наприклад, фонд заробітної плати залежить від трьох факторів: числа робітників, середньої заробітної плати і структури робітників; валовий збір зернових – від середньої врожайності, розміру посівних площ і їх структури і т.д.

На практиці при проведенні комбінаційного групування, як правило, обмежуються трьома-чотирма факторними ознаками, оскільки при більшому числі ознак таке групування стає неможливим, тому що не дозволяє виявити одночасно вплив всього комплексу факторних ознак на досліджуваний показник.

Аналітичні групування характеризують лише загальні риси зв’язку, який вивчається, його тенденцію, але не дають кількісної оцінки його сили. На основі аналітичних групувань це завдання розв’язується за допомогою розрахунку *емпіричного кореляційного відношення*.

Вимірювання тісноти зв’язку за допомогою емпіричного кореляційного відношення ґрунтується на правилі складання дисперсій. Основною метою дисперсійного аналізу є виявлення на основі величини загальної дисперсії впливу окремих факторів чи умов, які визначають варіацію ознаки. Для оцінки частки варіації, зумовленої тією чи іншою ознакою, сукупність розподіляють на групи за ознакою, вплив якої досліджується. Це дозволяє розкласти загальну варіацію ознаки на дві дисперсії, з яких одна частина варіації визначається впливом фактора, закладеного в основу групування, а друга – варіацією, зумовленою впливом усіх інших факторів, крім того, що вивчається. Отже, згідно з правилом складання дисперсій для розрахунку використовують загальну, міжгрупову і внутрішньогрупову (залишкову) дисперсії.

Загальна дисперсія характеризує варіацію ознаки у статистичній сукупності в результаті впливу всіх факторів, міжгрупова – вплив фактора, покладеного в основу групування, залишкова – впливом усіх інших факторів.

Таким чином, загальна дисперсія результативної ознаки складається з двох частин: міжгрупової та середньої з групових (залишкової).

Для кількісної оцінки зв’язку між явищами на базі матеріалів аналітичного групування вираховують коефіцієнт детермінації і емпіричне кореляційне відношення.

*Коефіцієнт детермінації* показує ступінь варіації ознаки під впливом фактора покладеного в основу групування. Він визначається як відношення міжгрупової дисперсії до загальної.

$$\eta^2 = \frac{\sigma^2_{\bar{y}}}{\sigma^2_з},$$

де $\eta^2$ – коефіцієнт детермінації ($\eta$ – грецька буква “ета”);

$\sigma^2_{\bar{y}}$ – міжгрупова дисперсія;

$\sigma^2_з$ – загальна дисперсія.

Критерієм суттєвості і сили зв’язку між факторною і результативною ознаками виступає *емпіричне кореляційне відношення*, яке визначається за формулою:

$$\eta = \sqrt{\frac{\sigma^2_{\bar{y}}}{\sigma^2_з}}.$$

Для якісної оцінки сили зв’язку між досліджуваними ознаками на основі емпіричного кореляційного відношення використовують таку шкалу (табл. 10.7).

Таблиця 10.7

| Величина $\eta$ | 0,1 – 0,3 | 0,3 – 0,5 | 0,5 – 0,7 | 0,7 – 0,9 | 0,9 – 0,99 |
|---|---|---|---|---|---|
| Сила зв'язку | слабкий | помірний | помітний | сильний | дуже сильний |

Скористаємось матеріалами аналітичного групування (табл. 10.6) і вирахуємо вищенаведені показники.

Таблиця 10.8

Розрахунок міжгрупової дисперсії

| Групи робітників-відрядників за стажем роботи, років | Число робітників, чол. $f_i$ | Середня місячна заробітна плата одного робітника, грн. $\bar{y}_i$ | $\bar{y}_i - \bar{y}_з$ | $\left(\bar{y}_i - \bar{y}_з\right)^2$ | $\left(\bar{y}_i - \bar{y}_з\right)^2 f_i$ |
|---|---|---|---|---|---|
| I 1 – 4 | 3 | 154,67 | -27,93 | 780,08 | 2340,25 |
| II 4 – 7 | 6 | 171,83 | -10,77 | 115,99 | 695,96 |
| III 7 – 10 | 5 | 191,00 | 8,40 | 70,56 | 352,80 |
| IV 10 – 13 | 4 | 193,50 | 10,90 | 118,81 | 475,24 |
| V 13 – 16 | 2 | 214,00 | 31,40 | 985,96 | 1971,92 |
| Разом | 20 | $\bar{Y}_з = 182{,}60$ | x | x | 5936,17 |

Міжгрупову дисперсію визначаємо за формулою:

$$\sigma_{\bar{y}}^2 = \frac{\sum\left(\bar{y}_i - \bar{y}_з\right)^2 f_i}{\sum f_i} = \frac{5936{,}17}{20} = 291{,}81 .$$

Розрахуємо загальну дисперсію за формулою:

$$\sigma_з^2 = \overline{y^2} - (\bar{y})^2 = 33666{,}40 - 33342{,}76 = 323{,}64 .$$

Обчислимо коефіцієнт детермінації і емпіричне кореляційне відношення:

$$\eta^2 = \frac{\sigma_{\bar{y}}^2}{\sigma_з^2} = \frac{291{,}81}{323{,}64} = 0{,}9 \text{ або } 90\ \%;$$

$$\eta = \sqrt{\frac{\sigma_{\bar{y}}^2}{\sigma_з^2}} = \sqrt{0{,}9} = 0{,}95 .$$

Коефіцієнт детермінації показує, що заробітна плата робітників-відрядників на 90 % залежить від стажу їх роботи і на 10 % від інших факторів.

Емпіричне кореляційне відношення свідчить про те, що зв'язок між стажем роботи і середньою місячною заробітною платою робітників-відрядників дуже сильний.

Кореляційне відношення змінюється в межах від 0 до 1. Якщо зв'язок відсутній, тоді емпіричне кореляційне відношення $\eta = 0$, всі групові середні рівні між собою, міжгрупова дисперсія також дорівнює $\sigma^2_{\bar{y}} = 0$, так як міжгрупової варіації немає.

Якщо зв'язок функціональний, то кореляційне відношення $\eta = 1$, міжгрупова дисперсія дорівнює загальній $\sigma^2_{\bar{y}} = \sigma^2_{з}$, а середня з групових $\overline{\sigma^2_i} = 0$. В цьому випадку кожному значенню факторної ознаки відповідає єдине значення результативної ознаки.

Чим більше значення кореляційного відношення наближається до одиниці, тим кореляційний зв'язок ближчий до функціональної залежності між ознаками.

Слід відмітити, якщо кореляційне відношення $\eta > 0$, то це не може бути доказом наявності кореляційного зв'язку між ознаками. Відмінне від нуля кореляційне відношення може з'явитись при неправильному розподілі досліджуваної сукупності на групи.

Перевірка істотності відхилень групових середніх здійснюється за допомогою критеріїв, розроблених математичною статистикою.

Емпіричне кореляційне відношення повинно мати високий рівень надійності. Для оцінки надійності кореляційних характеристик використовують критерій Фішера (F-критерій) або Стьюдента (t-критерій).

Критерій Фішера (F-критерій) визначається за формулою:

$$F_Ф = \frac{\sigma^2_{\bar{y}}}{\overline{\sigma^2_i}} \cdot \frac{k_2}{k_1}, \text{ або } F_Ф = \frac{\sigma^2_{\bar{y}} \cdot k_2}{\overline{\sigma^2_i} \cdot k_1},$$

де $\sigma^2_{\bar{y}}$ – міжгрупова дисперсія;

$\overline{\sigma^2_i}$ – середня групова (залишкова) дисперсія;

$k_1$, $k_2$ – ступені вільності для великої і малої дисперсій.

Фішер встановив розподіл відношень дисперсій і розробив відповідні математичні таблиці. В них наводиться F-критерій теоретичний ($F_T$) при двох ймовірностях 0,95 і 0,99. Якщо $F_Ф > F_T$ то з прийнятою ступінню ймовірності можна стверджувати про наявність впливу фактора, який вивчається. Коли ж $F_Ф \leq F_T$, можна стверджувати, що різниця між дисперсіями обумовлена впливом випадкових факторів.

Розподіл в таблицях Фішера для знаходження $F_T$ залежить від ступенів вільності міжгрупової ($k_1$) і середньої з групових ($k_2$) дисперсій. В аналітичному групуванні вони обчислюються за формулами:

$$k_1 = m - 1; \qquad k_2 = n - m,$$

де n – кількість елементів досліджуваної сукупності;

m – число груп.

За даними нашого прикладу

$$\overline{\sigma_i^2} = \sigma_з^2 - \sigma_{\bar{y}}^2 = 323{,}64 - 291{,}81 = 31{,}83\,.$$

Тоді $F_Ф = \dfrac{\sigma_{\bar{y}}^2 \cdot k_2}{\overline{\sigma_i^2} \cdot k_1} = \dfrac{291{,}81 \cdot 15}{31{,}83 \cdot 4} = 34{,}38$.

Для оцінки отриманого відношення, його порівнюють з табличним.

Ступінь вільності для великої (міжгрупової) дисперсії $k_1 = m - 1 = 5 - 1 = 4$, для малої (внутрішньогрупової) – $k_2 = n - m = 20 - 5 = 15$.

Знаходимо $F_T$ при ймовірності 0,95 і даних ступенях вільності по математичній таблиці. Воно становить $F_{T\,(0,95)} = 3{,}06$.

Таким чином $F_Ф > F_T$ (34,38 > 3,06), що свідчить про суттєвість впливу стажу роботи на середню місячну заробітну плату робітників-відрядників.

До аналогічного висновку можна прийти за оцінкою надійності кореляційного відношення по критерію Стьюдента (t-критерію), який визначається за формулою:

$$t_\eta = \frac{\eta}{\mu_\eta}\,,$$

де $\mu_\eta$ – середня помилка кореляційного відношення.

Вона визначається за формулою:

$$\mu_\eta = \frac{1 - \eta^2}{\sqrt{\eta}}\,.$$

Якщо критерій Стьюдента дорівнює або більший 3 ($t_\eta \geq 3$) показник кореляційного відношення вважають вірогідним (тобто зв’язок між досліджуваними явищами є доведеним). Якщо ж цей критерій менший 3 ($t_\eta < 3$), то не можна зробити висновок про вірогідність зв’язку між досліджуваними явищами.

Для нашого прикладу $\mu_\eta = \dfrac{1-\eta^2}{\sqrt{\eta}} = \dfrac{1-0{,}9}{\sqrt{0{,}95}} = 1{,}1053$ ,

$$t_\eta = \frac{\eta}{\mu_\eta} = \frac{0{,}9487}{0{,}1054} = 9{,}0 .$$

Так як критерій Стьюдента значно більший за 3, то кореляційне відношення вважається вірогідним, а зв'язок між стажем роботи і середньою місячною заробітною платою робітників-відрядників доведеним.

Переконавшись за допомогою аналітичного групування і розрахунку показника емпіричного кореляційного відношення, що сила зв'язку між досліджуваними явищами достатньо тісна, можна перейти до кореляційно-регресійного аналізу.

## 10.3 Кореляційний і регресійний методи аналізу зв'язку

Основне завдання кореляційного і регресійного методів заключається в аналізі статистичних даних з метою виявлення залежності між досліджуваними ознаками у вигляді певної математичної формули і встановлення за допомогою коефіцієнта кореляції порівняльної оцінки тісноти взаємозв'язку.

Після того як через економічний аналіз встановлено наявність зв'язку між явищами і загальний характер цього зв'язку, статистика з допомогою кореляційного і регресійного методів надає цим зв'язкам числового виразу.

Кореляційний і регресійний методи аналізу розв'язують два основних завдання:

1) визначають з допомогою рівнянь регресії аналітичну форму зв'язку між варіацією ознак X і Y;

2) встановлюють міру тісноти зв'язку між ознаками.

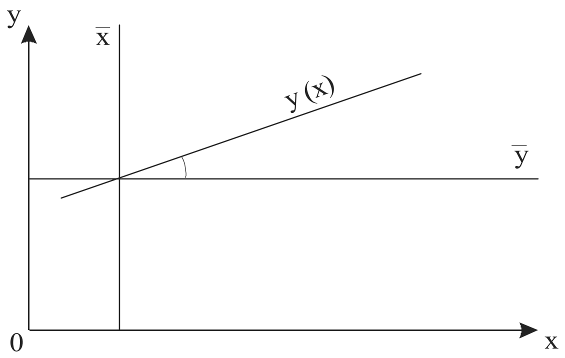
Найбільш часто зустрічаються наступні типи кореляційних зв'язків:

1) факторна ознака безпосередньо зв'язана з результативною;

2) результативна ознака визначається комплексом діючих факторів;

3) дві результативні ознаки викликані дією однієї загальної причини.

В практиці економіко-статистичних досліджень часто доводиться мати справу з прямолінійною формою зв'язку, яка виражається за допомогою рівняння регресії (мал. 10.3).




Мал. 10.3. Теоретична лінія регресії.

На графіку (мал. 10.3) середню арифметичну результативної ознаки $\bar{y}$ відображає пряма, паралельна осі абсцис, лінійне кореляційне рівняння $y(x)$ зображується похилою прямою, а кут нахилу між ними характеризує тісноту зв'язку.

Рівняння регресії характеризує зміну середнього рівня результативної ознаки (у) в залежності від зміни факторної ознаки (х). Воно визначає математичне сподівання групових середніх результативної ознаки під впливом різних значень факторної ознаки.

У випадку лінійної форми зв'язку, результативна ознаки змінюється під впливом факторної ознаки рівномірно. Така форма зв'язку виражається рівнянням прямої:

$$\hat{y}_x = a_0 + a_1 x,$$

де $\hat{y}_x$ – вирівняне середнє значення результативної ознаки;

х – значення факторної ознаки;

$a_0$ і $a_1$ – параметри рівняння;

$a_0$ – значення у при х = 0;

$a_1$ – коефіцієнт регресії.

Коефіцієнт регресії $a_1$ показує на скільки зміниться результативна ознака (у) при зміні факторної ознаки (х) на одиницю.

Якщо $a_1$ має позитивний знак, то зв'язок прямий, якщо від'ємний – то зв'язок обернений.

Параметри рівняння зв'язку визначаються способом найменших квадратів складеної і розв'язаної системи двох рівнянь з двома невідомими:

$$\begin{cases} \sum y = na_0 + a_1 \sum x; \\ \sum yx = a_0 \sum x + a_1 \sum x^2, \end{cases}$$

де $n$ – число членів в кожному з двох порівнювальних рядів;

$\sum x$ – сума значень факторної ознаки;

$\sum x^2$ – сума квадратів значень факторної ознаки;

$\sum y$ – сума значень результативної ознаки;

$\sum yx$ – сума добутків значень факторної ознаки на значення результативної ознаки.

Розв'язавши дану систему рівнянь, отримаємо такі значення параметрів:

$$a_0 = \frac{\sum x^2 \sum y - \sum x \sum yx}{n \sum x^2 - \sum x \sum x}; \qquad a_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - \sum x \sum x}.$$

Вирахувавши за фактичними даними всі записані вище суми і підставивши їх у наведені формули, знайдемо параметри шуканої прямої.

Покажемо розрахунок параметрів лінійного рівняння зв'язку між вартістю основних виробничих фондів і випуском продукції за даними десяти однорідних підприємств (табл. 10.9).

Таблиця 10.9

| Номер заводу n | Вартість основних виробничих фондів, млн. грн. x | Випуск продукції, млн. грн. y | $x^2$ | xy | $y^2$ | $\hat{y}_x = 0{,}167 + 0{,}421x$ |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | 12 | 5,6 | 144 | 67,2 | 31,36 | 5,2 |
| 2 | 8 | 4,0 | 64 | 32,0 | 16,00 | 3,5 |
| 3 | 10 | 4,0 | 100 | 40,0 | 16,00 | 4,4 |
| 4 | 6 | 2,4 | 36 | 14,4 | 5,76 | 2,7 |
| 5 | 9 | 3,6 | 81 | 32,4 | 12,96 | 4,0 |
| 6 | 15 | 5,0 | 225 | 75,0 | 25,00 | 6,5 |
| 7 | 11 | 4,6 | 121 | 50,6 | 21,16 | 4,8 |
| 8 | 13 | 6,5 | 169 | 84,5 | 42,25 | 5,6 |
| 9 | 14 | 7,0 | 196 | 98,0 | 49,00 | 6,1 |
| 10 | 10 | 4,5 | 100 | 45,0 | 20,25 | 4,4 |

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Разом | 108 | 47,2 | 1236 | 539,1 | 239,74 | 47,2 |
| В середньому на один завод | 10,8 | 4,72 | 123,6 | 53,91 | 23,974 | x |

За способом найменших квадратів визначимо параметри:

$$a_0 = \frac{1236 \cdot 47{,}2 - 108 \cdot 539{,}1}{10 \cdot 1236 - 108 \cdot 108} = \frac{58339{,}2 - 58222{,}8}{12360 - 11664} = \frac{116{,}4}{696{,}0} = 0{,}167\ ;$$

$$a_1 = \frac{10 \cdot 539{,}1 - 108 \cdot 47{,}2}{696{,}0} = \frac{5391{,}0 - 5097{,}6}{696{,}0} = \frac{293{,}4}{696{,}0} = 0{,}421\ .$$

Тоді лінійне рівняння регресії зв’язку між вартістю основних виробничих фондів і випуском продукції матиме вигляд:

$$\hat{y}_x = 0{,}167 + 0{,}421x\ .$$

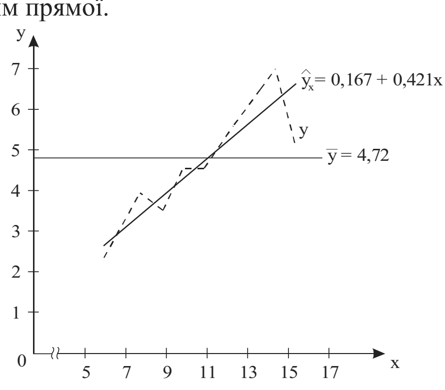
Таким чином, при збільшенні вартості основних виробничих фондів на 1 млн. грн. випуск продукції зросте на 0,421 млн. грн.

Підставляючи в дане рівняння послідовно значення факторної ознаки (x), отримаємо вирівняні значення результативної ознаки ($\hat{y}_x$), які показують, яким теоретично повинен бути середній розмір випуску продукції при даному розмірі основних виробничих фондів (за інших рівних умов).

Вирівняні (теоретичні) значення (із заокругленням до десятих) наведені в останній графі таблиці 10.9. Якщо параметри рівняння визначені правильно, то

$$\sum y = \sum \hat{y}_x = 47{,}2\ .$$

Побудуємо графік, який покаже вирівнювання емпіричних даних рівнянням прямої.




Мал. 10.4. Емпіричні і вирівняні рівні ряду.

Для економічної інтерпретації лінійних і нелінійних зв'язків між двома досліджуваними явищами часто використовують розраховані на основі рівнянь регресії коефіцієнти еластичності.

*Коефіцієнт еластичності* показує, на скільки відсотків зміниться в середньому результативна ознака (у) при зміні факторної ознаки (х) на 1 %.

Для лінійної залежності коефіцієнт еластичності визначається за формулою:

$$\varepsilon = a_1 \frac{x}{\hat{y}_x}, \text{ або } \varepsilon = a_1 \frac{\overline{x}}{\overline{y}},$$

де $\varepsilon$ – коефіцієнт еластичності.

Підставляючи в дану формулу різні значення х, будемо отримувати різні значення $\varepsilon$. В нашому прикладі коефіцієнт еластичності на першому підприємстві при х = 12 буде дорівнювати:

$\varepsilon_1 = a_1 \frac{x}{\hat{y}_x} = 0{,}421 \cdot \frac{12}{5{,}2} = 0{,}97$ %. Отже, на 1 % приросту вартості основних виробничих фондів, випуск продукції зросте на 0,97 %. На п'ятому підприємстві при х = 9 $\varepsilon_5 = 0{,}421 \cdot \frac{9}{4} = 0{,}95$ %, на десятому – при х = 10 $\varepsilon_{10} = 0{,}96$ %.

Для всіх підприємств разом коефіцієнт еластичності буде дорівнювати:

$$\varepsilon = a_1 \frac{\overline{x}}{\overline{y}} = 0{,}421 \cdot \frac{10{,}8}{4{,}72} = 0{,}963 \text{ \%}.$$

Це означає, що при збільшенні середньої вартості основних виробничих фондів на 1 % випуск продукції зросте в середньому на 0,963 %.

Якщо залежність між ознаками представлена за даними, вирівняними за параболою другого порядку, то коефіцієнт еластичності матиме вигляд:

$$\varepsilon = (a_1 + a_2 x) \frac{\overline{x}}{\overline{y}}.$$

Визначення тісноти зв'язку в кореляційно-регресійному аналізі ґрунтується на правилі складання дисперсій як і в методі аналітичного групування. Але на відміну від нього, де для оцінки лінії регресії застосовують групові середні результативної ознаки, в кореляційно-регресійному аналізі для цієї мети використовують теоретичні значення результативної ознаки.

Для наочної уяви і обґрунтування кореляційно-регресійного аналізу звернемось до графіка (мал. 10.4). На цьому графіку нами проведені три лінії:

1) y – ламана лінія фактичних даних;
2) $\hat{y}_x$ – пряма похила лінія теоретичних значень (y) при абстрагуванні від впливу всіх факторів, крім фактора (x) (змінна середня);
3) $\bar{y}$ – пряма горизонтальна лінія, з середнього значення якої виключено вплив на (y) всіх без винятку факторів (постійна середня).

Неспівпадання лінії змінної середньої $\hat{y}_x$ з лінією постійної середньої $\bar{y}$ пояснюється впливом факторної ознаки x, що, в свою чергу, свідчить про існування між ознаками y і x неповного, нефункціонального зв'язку. Для визначення тісноти цього зв'язку потрібно вирахувати дисперсію відхилень y і $\hat{y}_x$, тобто залишкову дисперсію, яка зумовлена впливом всіх факторів, крім фактора x. Різниця між загальною і залишковою дисперсіями дає нам теоретичну (факторну) дисперсію, яка вимірює варіацію, зумовлену фактором x. На порівнянні цієї різниці із загальною дисперсією побудований *індекс кореляції*, або *теоретичне кореляційне відношення*, які визначаються за такими формулами:

$$R = \sqrt{\frac{\sigma_з^2 - \sigma_e^2}{\sigma_з^2}} = \sqrt{1 - \frac{\sigma_e^2}{\sigma_з^2}}, \text{ або } R = \sqrt{\frac{\sigma_{\hat{y}}^2}{\sigma_з^2}},$$

де R – індекс кореляції (теоретичне кореляційне відношення);

$\sigma_з^2$ – загальна дисперсія;

$\sigma_e^2$ – залишкова дисперсія;

$\sigma_{\hat{y}}^2$ – факторна (теоретична) дисперсія.

Факторну дисперсію з теоретичних значень обчислюють за формулою:

$$\sigma_{\hat{y}}^2 = \frac{\sum (\hat{y}_x - \bar{y})^2}{n},$$

або за формулою без теоретичних значень

$$\sigma_{\hat{y}}^2 = \frac{\left(a_0 \sum y + a_1 \sum xy\right) - (\bar{y})^2}{n}.$$

Залишкову дисперсію визначають або за формулою $\sigma_e^2 = \frac{\sum (y - \hat{y}_x)^2}{n}$, або за правилом складання дисперсій $\sigma_e^2 = \sigma_з^2 - \sigma_{\hat{y}}^2$.

У наведеному прикладі (за даними розрахунків в табл. 10.9) факторна дисперсія дорівнює:

$$\sigma_{\hat{y}}^2 = \frac{(0{,}167 \cdot 47{,}2 + 0{,}421 \cdot 539{,}1) - 4{,}72^2}{10} = 1{,}206 .$$

Загальну дисперсію обчислимо за формулою:

$$\sigma_з^2 = \overline{y^2} - (\overline{y})^2 = 23{,}974 - 22{,}278 = 1{,}696 .$$

Залишкову дисперсію визначимо як різницю між загальною і факторною дисперсіями:

$$\sigma_e^2 = \sigma_з^2 - \sigma_{\hat{y}}^2 = 1{,}696 - 1{,}206 = 0{,}490 .$$

Таким чином, індекс кореляції за вище наведеними формулами буде становити:

$$R = \sqrt{\frac{\sigma_з^2 - \sigma_e^2}{\sigma_з^2}} = \sqrt{\frac{1{,}696 - 0{,}490}{1{,}696}} = 0{,}843 \text{, або}$$

$$R = \sqrt{1 - \frac{\sigma_e^2}{\sigma_з^2}} = \sqrt{1 - \frac{0{,}490}{1{,}696}} = 0{,}843 \text{, або}$$

$$R = \sqrt{\frac{\sigma_{\hat{y}}^2}{\sigma_з^2}} = \sqrt{\frac{1{,}206}{1{,}696}} = \sqrt{0{,}711} = 0{,}843 .$$

Індекс кореляції показує тісну залежність випуску продукції від вартості основних виробничих фондів.

*Коефіцієнт детермінації* ($R^2$) характеризує ту частину варіації результативної ознаки (y), яка відповідає лінійному рівнянню регресії:

$$R^2 = \frac{\sigma_{\hat{y}}^2}{\sigma_з^2} = \frac{1{,}206}{1{,}696} = 0{,}711 .$$

Отже, в обстеженій сукупності заводів 71,1 % варіації випуску продукції пояснюється різними рівнями оснащеності заводів основними виробничими фондами.

Індекс кореляції приймає значення від 0 до 1. Коли R = 0, то зв'язку між варіацією ознак y і x немає. Залишкова дисперсія дорівнює загальній ($\sigma_e^2 = \sigma_з^2$), а теоретична дисперсія дорівнює нулю ($\sigma_{\hat{y}}^2 = 0$). Всі теоретичні значення $\hat{y}_x$ збігаються з середніми значеннями $\overline{y}$, лінія $\hat{y}_x$ на графіку співпадає з лінією $\bar{y}$, тобто приймає горизонтальне положення.

При R = 1 теоретична дисперсія дорівнює загальній ($\sigma^2_{\hat{y}} = \sigma^2_з$), а залишкова – $\sigma^2_e = 0$. Фактичні значення у збігаються з теоретичними $\hat{y}_x$, зв’язок між досліджуваними ознаками лінійно-функціональний.

Індекс кореляції придатний для вимірювання тісноти зв’язку при любій її формі. Він, як і емпіричне кореляційне відношення вимірює лише тісноту зв’язку і не показує її напрямок.

Для вимірювання тісноти зв’язку і визначення його напрямку при лінійній залежності використовують *лінійний коефіцієнт кореляції*, який визначається за формулою:

$$r = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\sigma_x \cdot \sigma_y}.$$

Значення r коливається в межах від –1 до +1. Додатнє значення r означає прямий зв’язок між ознаками, а від’ємне – зворотній.

Оцінка тісноти зв’язку проводиться за наступною приблизною схемою (табл. 10.10).

Таблиця 10.10

Для оцінки тісноти зв’язку

| Сила зв’язку | Величина лінійного коефіцієнта кореляції при наявності: | |
|---|---|---|
| | прямого зв’язку | оберненого зв’язку |
| Слабка | 0,1 – 0,30 | (-0,1) – (-0,30) |
| Середня | 0,3 – 0,70 | (-0,3) – (-0,70) |
| Тісна | 0,7 – 0,99 | (-0,7) – (-0,99) |

Всі дані для обчислення лінійного коефіцієнта кореляції в нашому прикладі є в таблиці 10.9.

$$\sigma_x = \sqrt{\overline{x^2} - (\bar{x})^2} = \sqrt{123{,}6 - 10{,}8^2} = \sqrt{6{,}96} = 2{,}638;$$

$$\sigma_y = \sqrt{\overline{y^2} - (\bar{y})^2} = \sqrt{23{,}974 - 4{,}72^2} = 1{,}302;$$

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y} = \frac{53{,}91 - 10{,}8 \cdot 4{,}72}{2{,}638 \cdot 1{,}302} = \frac{2{,}9340}{3{,}4349} = 0{,}854.$$

Скористаємось для знаходження лінійного коефіцієнта кореляції іншою формулою:

$$r = a_1 \frac{\sigma_x}{\sigma_y} = 0{,}421 \frac{2{,}638}{1{,}302} = 0{,}853,$$

тобто відповідь отримана така сама. Це означає, що зв'язок між вартістю основних виробничих фондів і випуском продукції сильний (тісний) і прямий.

Абсолютна величина лінійного коефіцієнта кореляції збігається з індексом кореляції (відхилення в 0,01).

Із наведених формул коефіцієнта кореляції можна визначити коефіцієнт регресії, не вираховуючи рівняння зв'язку.

$$a_1 = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x^2} = \frac{2,934}{6,960} = 0,421 \text{, або}$$

$$a_1 = r \cdot \frac{\sigma_y}{\sigma_x} = 0,853 \cdot \frac{1,302}{2,638} = 0,421 \,.$$

Перевірку істотності зв'язку в кореляційно-регресійному аналізі здійснюють за допомогою тих самих критеріїв і процедур, що і в аналітичному групуванні. Ступені вільності залежать від числа параметрів рівняння регресії $k_1 = m - 1$ і кількості одиниць досліджуваної сукупності $k_2 = n - m$.

Істотність зв'язку коефіцієнта детермінації $R^2$ перевіряють за допомогою таблиці критерію F для 5 %-ного рівня значимості. Так, при $k_1 = m - 1 = 2 - 1 = 1$ (для лінійної моделі) і $k_2 = n - m = 10 - 2 = 8$. Фактичне значення F-критерію для нашого прикладу визначають за формулою:

$$F_Ф = \frac{R^2}{1 - R^2} \cdot \frac{k_2}{k_1} = \frac{0,711}{1 - 0,711} \cdot \frac{8}{1} = 19,68 \,.$$

Критичне значення $F_{T(0,95)} = 5,32$ значно менше від фактичного $F_{T(0,95)} < F_Ф$ (5,32 < 19,68), що підтверджує істотність кореляційного зв'язку між досліджуваними ознаками.

Для встановлення достовірності обчисленого нами лінійного коефіцієнта кореляції використовують критерій Стьюдента (t-критерій):

$$t_r = \frac{|r|}{\mu_r} \,,$$

де $\mu_r$ – середня помилка коефіцієнта кореляції, яку визначають за формулою:

$$\mu_r = \frac{1 - r^2}{\sqrt{n - 1}} \,.$$

При достатньо великому числі спостережень (n > 50) коефіцієнт кореляції можна вважати достовірним, якщо він перевищує свою помилку в 3 і більше раз, а якщо він менше 3, то зв'язок між досліджуваними ознаками у і х не доведений.

В нашому прикладі середня помилка коефіцієнта кореляції дорівнює:

$$\mu_r = \frac{1-0{,}853^2}{\sqrt{9}} = \frac{1-0{,}723}{3} = \frac{0{,}277}{3} = 0{,}092 \,.$$

Відношення коефіцієнта кореляції до його середньої помилки становить:

$$t_r = \frac{0{,}853}{0{,}092} = 9{,}27 \,.$$

Це дає нам право вважати, що обчислений лінійний коефіцієнт кореляції достатньо точно характеризує силу зв'язку між досліджуваними ознаками.

### 10.4. Нелінійні залежності

Якщо попередній аналіз явищ, зв'язок яких досліджується, показує, що рівним змінам середніх значень факторної ознаки відповідають нерівні зміни середніх значень результативної ознаки, то для вираження загального характеру зв'язку застосовують криволінійні форми кореляційних рівнянь. В практиці економічного аналізу найбільш часто використовують наступні нелінійні функції залежності: гіперболічну, параболічну другого порядку, напівлогарифмічну, логістичну та деякі інші.

Якщо результативна ознака із збільшенням факторної ознаки зростає або спадає не безкінечно, а прямує до кінцевої мети, то для її аналізу застосовується *рівняння гіперболи*:

$$\hat{y}_x = a_0 + a_1 \frac{1}{x} \,.$$

Для знаходження параметрів цього рівняння способом найменших квадратів використовується система нормальних рівнянь:

$$\begin{cases} \sum y = na_0 + a_1 \sum \frac{1}{x} \,; \\ \sum y \frac{1}{x} = a_0 \sum \frac{1}{x} + a_1 \sum \frac{1}{x^2}. \end{cases}$$

Покажемо побудову гіперболічного кореляційного рівняння на прикладі.

Маємо дані про денний виробіток п'яти робітників і собівартість одиниці виробленої ними продукції (табл. 10.11).

Таблиця 10.11

Вирівнювання по гіперболі.

| № робіт-ника n | Денний виробіток робітника, тис. грн. x | Собівартість одиниці продукції, грн. y | $\frac{1}{x}$ | $\frac{1}{x^2}$ | $\frac{y}{x}$ | $\hat{y}_x = 1{,}014 + 6{,}305\frac{1}{x}$ |
|---|---|---|---|---|---|---|
| 1 | 1 | 7,0 | 1,00 | 1,00 | 7,00 | 7,22 |
| 2 | 2 | 5,0 | 0,50 | 0,25 | 2,50 | 4,17 |
| 3 | 3 | 3,0 | 0,33 | 0,11 | 1,00 | 3,09 |
| 4 | 5 | 2,0 | 0,20 | 0,04 | 0,40 | 2,28 |
| 5 | 10 | 1,5 | 0,10 | 0,01 | 0,15 | 1,64 |
| Разом: | 21 | 18,5 | 2,13 | 1,41 | 11,05 | 18,50 |

За способом найменших квадратів знайдемо параметри гіперболи:

$$a_0 = \frac{\sum\frac{1}{x^2}\cdot\sum y - \sum\frac{1}{x}\cdot\sum\frac{y}{x}}{n\sum\frac{1}{x^2} - \sum\frac{1}{x}\cdot\sum\frac{1}{x}} = \frac{1{,}41\cdot 18{,}5 - 2{,}13\cdot 11{,}05}{5\cdot 1{,}41 - 2{,}13\cdot 2{,}13} = 1{,}014\ ;$$

$$a_1 = \frac{n\sum\frac{y}{x} - \sum\frac{1}{x}\cdot\sum y}{n\sum\frac{1}{x^2} - \sum\frac{1}{x}\cdot\sum\frac{1}{x}} = \frac{5\cdot 11{,}05 - 2{,}13\cdot 18{,}5}{2{,}5131} = 6{,}305\ .$$

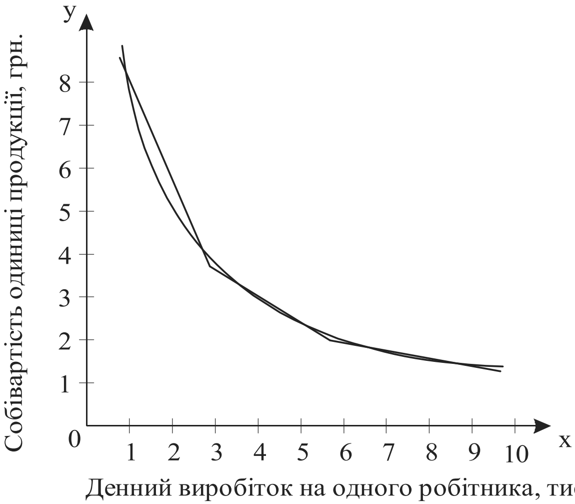
Отже, шукане кореляційне рівняння гіперболи матиме вигляд:

$$\hat{y}_x = 1{,}014 + 6{,}305\cdot\frac{1}{x}\ .$$

Підставляючи в це рівняння відповідні значення x, отримаємо теоретичні значення собівартості одиниці продукції $\hat{y}_x$, наведені в останній графі таблиці 10.11.

Нанесемо фактичні і теоретичні значення досліджуваних ознак на графік (мал. 10.5).

Проставлені на графіку значення змінної середньої $\hat{y}_x$ досить близько відтворюють фактичні значення y. Наочно видно вирівнювання емпіричної ламаної лінії зв’язку гіперболою. Вона зображає обернений зв’язок, де із збільшенням денного виробітку продукції на одного робітника собівартість одиниці продукції знижується. При цьому зменшення результативної ознаки відбувається повільніше, ніж зростання факторної.




Мал. 10.5. Графік кореляційної залежності собівартості одиниці продукції від денного виробітку на одного робітника.

Для визначення тісноти зв'язку між результативною і факторною ознаками обчислюємо кореляційне відношення за формулою:

$$\eta_{\frac{y}{x}} = \sqrt{1 - \frac{\sum\left(y - \hat{y}_x\right)^2}{\sum\left(y - \bar{y}\right)^2}}\ .$$

Складаємо розрахункову таблицю 10.12.

Таблиця 10.12

| № робіт-ника n | Денний виробіток робітника, тис. грн. x | Собівартість одиниці продукції, грн. y | $y - \bar{y}$ | $(y - \bar{y})^2$ | $\hat{y}_x$ | $y - \hat{y}_x$ | $(y - \hat{y}_x)^2$ |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 7,0 | 3,3 | 10,89 | 7,22 | -0,22 | 0,0484 |
| 2 | 2 | 5,0 | 1,3 | 1,69 | 4,17 | 0,83 | 0,6889 |
| 3 | 3 | 3,0 | -0,7 | 0,49 | 3,09 | -0,09 | 0,0081 |
| 4 | 5 | 2,0 | -1,7 | 2,89 | 2,28 | -0,28 | 0,0784 |
| 5 | 10 | 1,5 | -2,2 | 4,84 | 1,64 | -0,14 | 0,0196 |
| Разом: | 21 | 18,5 | x | 20,80 | 18,50 | x | 0,8434 |

Середнє значення результативної ознаки:

$$\bar{y} = \frac{\sum y}{n} = \frac{18{,}5}{5} = 3{,}7 \text{ грн.}$$

Підставивши підсумкові дані таблиці 10.12 у вищенаведену формулу кореляційного відношення, отримаємо:

$$\eta_{\frac{y}{x}} = \sqrt{1 - \frac{0{,}8434}{20{,}8000}} = \sqrt{1 - 0{,}0405} = \sqrt{0{,}9595} = 0{,}979 \; .$$

Кореляційне відношення показує, що між собівартістю і виробітком продукції одним робітником існує тісна обернена залежність.

Надійність показника кореляційного відношення перевіримо по t-критерію Стьюдента. Для цього спочатку визначимо середню помилку кореляційного відношення:

$$\mu_\eta = \frac{1-\eta^2}{\sqrt{n-1}} = \frac{1-0{,}979^2}{\sqrt{4}} = \frac{0{,}041}{2} = 0{,}0208 \; .$$

t-критерій дорівнює:

$$t = \frac{\eta}{\mu_\eta} = \frac{0{,}9790}{0{,}0208} = 47{,}07 \; .$$

Це свідчить про високу надійність обчисленого показника кореляційного відношення.

*Парабола другого порядку* як форма математичного вираження зв'язків між досліджуваними явищами застосовується в тих випадках, коли із зростанням факторної ознаки відбувається нерівномірне зростання або спадання результативної ознаки.

При знаходженні рівняння зв'язку між такого роду показниками x і y в якості апроксимативної функції застосовується тип кривої, вираженої у вигляді параболи другого порядку:

$$\hat{y}_x = a_0 + a_1 x + a_2 x^2 \; .$$

Параметри цього рівняння знаходять способом найменших квадратів шляхом складання і розв'язку системи нормальних рівнянь:

$$\begin{cases} \sum y = na_0 + a_1 \sum x + a_2 \sum x^2; \\ \sum xy = a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3; \\ \sum x^2 y = a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4. \end{cases}$$

Розглянемо в якості прикладу залежність між урожайністю насіння багаторічних трав і глибиною зрошення за даними спостереження на 10 ділянках.

Всі необхідні дані для розв'язку системи рівнянь розраховані та наведені в табл. 10.13.

Таблиця 10.13.

| Номер спостереження n | Глибина зрошення, см. x | Урожайність насіння трав, ц/га y | $x^2$ | $x^3$ | $x^4$ | xy | $x^2y$ | $\hat{y}_x$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1,2 | 0 | 0 | 0 | 0 | 0 | 3,0 |
| 2 | 5 | 5,0 | 25 | 125 | 625 | 25,0 | 125,0 | 4,7 |
| 3 | 10 | 7,0 | 100 | 1000 | 10000 | 70,0 | 700,0 | 6,2 |
| 4 | 15 | 8,0 | 225 | 3375 | 50625 | 120,0 | 1800,0 | 7,4 |
| 5 | 20 | 9,2 | 400 | 8000 | 160000 | 184,0 | 3680,0 | 8,4 |
| 6 | 25 | 9,5 | 625 | 15625 | 390625 | 237,5 | 5937,5 | 9,2 |
| 7 | 30 | 9,7 | 900 | 27000 | 810000 | 291,0 | 8730,0 | 9,6 |
| 8 | 35 | 9,9 | 1225 | 42875 | 1500625 | 346,5 | 12127,5 | 10,0 |
| 9 | 40 | 9,8 | 1600 | 64000 | 2560000 | 392,0 | 15680,0 | 9,9 |
| 10 | 45 | 8,7 | 2025 | 91125 | 4100625 | 391,5 | 17617,5 | 9,6 |
| Разом: | 225 | 78,0 | 7125 | 253125 | 9583125 | 2057,5 | 66697,5 | 78,0 |

Підставляючи дані таблиці у систему рівнянь, матимемо:

$$78 = 10\ a_0 + 225\ a_1 + 7125\ a_2;$$
$$2057{,}5 = 225\ a_0 + 7125\ a_1 + 253125\ a_2;$$
$$66697{,}5 = 7125\ a_0 + 253125\ a_1 + 9583125\ a_2.$$

Звільняємося від коефіцієнта при $a_0$, для чого кожний член рівняння ділимо на відповідний коефіцієнт при $a_0$. Отримуємо:

$$7{,}800 = a_0 + 22{,}5\ a_1 + 712{,}5\ a_2;$$
$$9{,}144 = a_0 + 31{,}667\ a_1 + 1125\ a_2;$$
$$9{,}361 = a_0 + 35{,}265\ a_1 + 1345\ a_2.$$

Віднімаємо із другого рівняння перше, а з третього – друге, внаслідок чого отримаємо два рівняння з двома невідомими:

$$1{,}344 = 9{,}167\ a_1 + 412{,}5\ a_2;$$
$$0{,}301 = 3{,}589\ a_1 + 220{,}0\ a_2.$$

Ділимо кожний член цих рівнянь на відповідний коефіцієнт при $a_1$ і віднімаємо від першого рівняння друге:

$$-\begin{cases} 0{,}147 = a_1 + 44{,}998\ a_2; \\ 0{,}060 = a_1 + 61{,}298\ a_2; \end{cases}$$
$$\overline{0{,}087 = -\ 16{,}3\ a_2.}$$

Звідси $a_2 = \frac{0{,}087}{-16{,}3} = -0{,}005$ .

Методом підстановки отримуємо параметри $a_1$ і $a_0$:

$a_1$ = 0,147 – 44,998 · (-0,005) = 0,372;

$a_0$ = 7,8 – 22,5 · 0,372 – 712,5 · (-0,005) = 2,9925.

Таким чином, рівняння параболи другого порядку матиме вигляд:

$$\hat{y}_x = 2{,}9925 + 0{,}372x + 0{,}005x^2 .$$

Значення коефіцієнта регресії $a_1$ в рівнянні свідчить, що при збільшенні глибини зрошення на 1 см урожайність насіння багаторічних трав зростає в середньому на 37,2 кг, а значення коефіцієнта $a_2$ вказує на те, що цей процес відбувається із сповільненням на 0,05 кг.

Далі знаходимо теоретичну лінію регресії рівняння параболи другого порядку $\hat{y}_x$, підставляючи в рівняння регресії значення x і $x^2$:

$$\hat{y}_1 = 2{,}9925 + 0{,}372 \cdot 0 - 0{,}005 \cdot 0 = 2{,}9925 \approx 3{,}0 ;$$

$$\hat{y}_2 = 2{,}9925 + 0{,}372 \cdot 5 - 0{,}005 \cdot 25 = 4{,}7275 \approx 4{,}7 ;$$

$$\hat{y}_3 = 2{,}9925 + 0{,}372 \cdot 10 - 0{,}005 \cdot 100 = 6{,}2125 \approx 6{,}2 ; \text{ і т.д.}$$

Визначаємо середню урожайність насіння багаторічних трав:

$$\bar{y} = \frac{\sum y}{n} = \frac{78}{10} = 7{,}8 \text{ ц/га.}$$

З метою оцінки тісноти зв'язку між урожайністю насіння багаторічних трав і глибиною зрошення визначимо кореляційне відношення, для чого проведемо додаткові розрахунки в табл. 10.14.

Таблиця 10.14

| Номер спостереження n | y | $y-\bar{y}$ | $(y-\bar{y})^2$ | $\hat{y}_x$ | $\hat{y}_x-\bar{y}$ | $(\hat{y}_x-\bar{y})^2$ |
|---|---|---|---|---|---|---|
| 1 | 1,2 | -6,6 | 43,56 | 3,0 | - 4,8 | 23,04 |
| 2 | 5,0 | -2,2 | 4,84 | 4,7 | - 3,1 | 9,61 |
| 3 | 7,0 | -0,8 | 0,64 | 6,2 | - 1,6 | 2,56 |
| 4 | 8,0 | 0,2 | 0,04 | 7,4 | - 0,4 | 0,16 |
| 5 | 9,2 | 1,4 | 1,96 | 8,4 | 0,6 | 0,36 |
| 6 | 9,5 | 1,7 | 2,89 | 9,2 | 1,4 | 1,96 |
| 7 | 9,7 | 1,9 | 3,61 | 9,6 | 1,8 | 3,24 |
| 8 | 9,9 | 2,1 | 4,41 | 10,0 | 2,2 | 4,84 |
| 9 | 9,8 | 2,0 | 4,00 | 9,9 | 2,1 | 4,41 |
| 10 | 8,7 | 0,9 | 0,81 | 9,6 | 1,8 | 3,24 |
| Разом: | 78,0 | x | 66,76 | 78,0 | x | 53,42 |

Кореляційне відношення за даними наших розрахунків дорівнює:

$$\eta_{\frac{y}{x}} = \sqrt{\frac{\sum(\hat{y}_x - \bar{y})^2}{\sum(y - \bar{y})^2}} = \sqrt{\frac{53,42}{66,76}} = \sqrt{0,800} = 0,895\ .$$

Зв'язок між урожайністю насіння багаторічних трав і глибиною зрошення тісний.

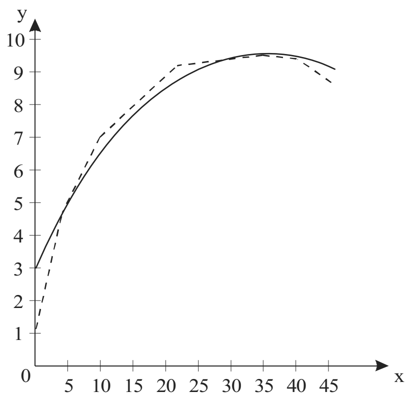
Для встановлення достовірності обчисленого кореляційного відношення скористаємось t-критерієм Стьюдента.

$$\mu_\eta = \frac{1-\eta^2}{\sqrt{n-1}} = \frac{1-0,800}{3} = 0,067\ ;$$

$$t_\eta = \frac{\eta}{\mu_\eta} = \frac{0,895}{0,067} = 13,36\ .$$

Так як $t_\eta > t_T$ (13,36 > 3) залежність урожайності насіння багаторічних трав від глибини зрошення є доведеною.

Покажемо вирівнювання емпіричних даних за параболою другого порядку на графіку (мал. 10.6).




Мал. 10.6. Графік залежності між глибиною зрошення (x) і урожайністю насіння багаторічних трав (y).

Вирівнювання за напівлогарифмічною кривою проводять в тих випадках, коли із зростанням факторної ознаки, середня результативна ознака спочатку до певних меж зростає досить швидко, але пізніше темпи її зростання поступово сповільнюються.

Прикладом вирівнювання такого взаємозв'язку може бути залежність між виробітком одного продавця (y) і зростанням об'єму товарообороту магазину (x). З укрупненням магазину розмір товарообороту в середньому на одного продавця зростає нерівномірно. В невеликих магазинах виробіток росте високими темпами, але в міру збільшення розмірів торгових точок темпи його зростання спадають. Якраз таку залежність найкраще описує напівлогарифмічна функція виду:

$$\hat{y}_x = a_0 + a_1 \log x .$$

Покажемо застосування цієї функції на прикладі роботи 25 магазинів (табл. 10.15).

Таблиця 10.15

Розрахункова таблиця для вирівнювання за напівлогарифмічною кривою.

| Номер магазину n | Об'єм товаро-обороту магазину, млн. грн. x | Виробіток на одного продавця, тис. грн. y | log x | $(\log x)^2$ | y log x | $\hat{y}_x$ |
|---|---|---|---|---|---|---|
| 1 | 10,0 | 5,2 | 1,0000 | 1,000 | 5,200 | 4,989 |
| 2 | 10,2 | 5,3 | 1,0086 | 1,017 | 5,346 | 5,019 |
| 3 | 10,4 | 5,2 | 1,0170 | 1,034 | 5,288 | 5,048 |
| 4 | 12,0 | 5,4 | 1,0792 | 1,165 | 5,828 | 5,264 |
| 5 | 15,3 | 5,1 | 1,1847 | 1,404 | 6,042 | 5,631 |
| 6 | 16,2 | 5,3 | 1,2095 | 1,463 | 6,410 | 5,718 |
| 7 | 16,6 | 5,5 | 1,2201 | 1,489 | 6,710 | 5,755 |
| 8 | 18,8 | 5,4 | 1,2742 | 1,624 | 6,881 | 5,943 |
| 9 | 20,0 | 5,9 | 1,3010 | 1,693 | 7,676 | 6,036 |
| 10 | 20,3 | 6,5 | 1,3075 | 1,709 | 8,499 | 6,059 |
| 11 | 20,8 | 5,8 | 1,3181 | 1,737 | 7,645 | 6,096 |
| 12 | 22,0 | 6,8 | 1,3424 | 1,802 | 9,262 | 6,180 |
| 13 | 22,5 | 5,8 | 1,3522 | 1,828 | 7,843 | 6,214 |
| 14 | 26,0 | 5,7 | 1,4150 | 2,002 | 8,066 | 6,433 |
| 15 | 28,0 | 7,0 | 1,4472 | 2,094 | 10,130 | 6,545 |
| 16 | 29,0 | 6,6 | 1,4624 | 2,139 | 9,652 | 6,598 |
| 17 | 33,2 | 7,1 | 1,5211 | 2,314 | 10,800 | 6,802 |
| 18 | 34,0 | 8,0 | 1,5315 | 2,345 | 12,252 | 6,838 |
| 19 | 39,0 | 7,5 | 1,5911 | 2,531 | 11,933 | 7,045 |
| 20 | 40,0 | 6,9 | 1,6021 | 2,567 | 11,054 | 7,084 |
| 21 | 45,5 | 6,8 | 1,6580 | 2,749 | 11,274 | 7,248 |
| 22 | 56,0 | 7,3 | 1,7482 | 3,056 | 12,762 | 7,592 |
| 23 | 64,0 | 7,7 | 1,8062 | 3,262 | 13,908 | 7,794 |
| 24 | 67,0 | 7,9 | 1,8261 | 3,335 | 14,426 | 7,863 |
| 25 | 72,0 | 8,0 | 1,8573 | 3,450 | 14,858 | 7,971 |
| Разом: | 748,8 | 159,8 | 35,0807 | 50,809 | 229,746 | 159,765 |

Для знаходження параметрів напівлогарифмічної функції способом найменших квадратів, потрібно розв'язати систему двох рівнянь:

$$\begin{cases} \sum y = na_0 + a_1 \sum \log x; \\ \sum y \log x = a_0 \sum \log x + a_1 \sum (\log x)^2. \end{cases}$$

Підставляючи у рівняння для знаходження параметрів розрахункові дані із таблиці, отримаємо:

$$159{,}8 = 25\ a_0 + 35{,}08\ a_1;$$
$$229{,}746 = 35{,}08\ a_0 + 50{,}809\ a_1.$$

Вирівнюємо коефіцієнти при $a_0$, перемножуючи перше рівняння на 1,4032 (35,08 : 25) і віднімаємо заново створене рівняння від другого:

$$-\ \begin{array}{l} 229{,}746 = 35{,}08\ a_0 + 50{,}809\ a_1; \\ \underline{224{,}231 = 35{,}08\ a_0 + 49{,}224\ a_1;} \end{array}$$
$$5{,}515 = 1{,}585\ a_1.$$

Звідси $a_1 = \dfrac{5{,}515}{1{,}585} = 3{,}479$ .

Підставляючи це значення в перше рівняння визначимо $a_0$:

$$25\ a_0 = 159{,}8 – 122{,}04 = 37{,}76,$$
$$a_0 = \frac{37{,}76}{25} = 1{,}51 .$$

Отже, напівлогарифмічне рівняння зв'язку прийме вигляд:

$$\hat{y}_x = 1{,}51 + 3{,}479 \log x .$$

На основі цього рівняння розрахуємо теоретичні значення $\hat{y}_x$ (див. останню колонку табл. 10.15) і вирівняємо за їх допомогою фактичні значення результативної ознаки у на графіку (мал. 10.7).

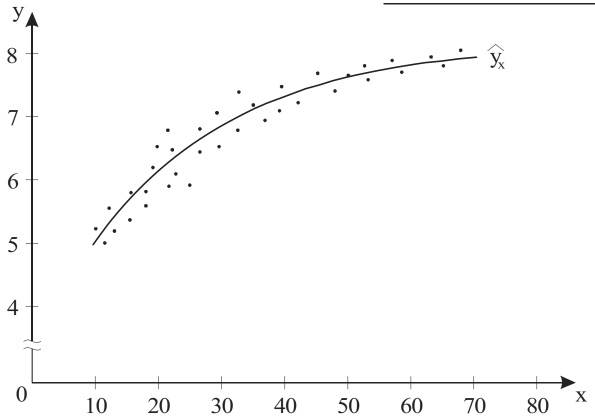
Оцінимо ступінь тісноти зв'язку між обсягом товарообороту і виробітком одного продавця за допомогою кореляційного відношення, а надійність цієї залежності перевіримо за t-критерієм Стьюдента, для чого побудуємо розрахункову таблицю 10.16.

Знаходимо середній розмір виробітку:

$$\bar{y} = \frac{\sum y}{n} = \frac{159{,}8}{25} = 6{,}392 \approx 6{,}4 \text{ тис. грн.}$$

Визначаємо кореляційне відношення:

$$\eta_{\frac{y}{x}} = \sqrt{\frac{\sum (\hat{y}_x - \bar{y})^2}{\sum (y - \bar{y})^2}} = \sqrt{\frac{19{,}48}{23{,}90}} = 0{,}903 .$$




Мал. 10.7. Графік кореляційної залежності між розміром товарообороту магазину (x) і виробітком одного продавця (y).

Таблиця 10.16

| Номер магазину n | y | $y-\bar{y}$ | $(y-\bar{y})^2$ | $\hat{y}_x$ | $\hat{y}_x-\bar{y}$ | $(\hat{y}_x-\bar{y})^2$ |
|---|---|---|---|---|---|---|
| 1 | 5,2 | -1,2 | 1,44 | 5,0 | -1,4 | 1,96 |
| 2 | 5,3 | -1,1 | 1,21 | 5,0 | -1,4 | 1,96 |
| 3 | 5,2 | -1,2 | 1,44 | 5,0 | -1,4 | 1,96 |
| 4 | 5,4 | -1,0 | 1,00 | 5,3 | -1,1 | 1,21 |
| 5 | 5,1 | -1,3 | 1,69 | 5,6 | -0,8 | 0,64 |
| 6 | 5,3 | -1,1 | 1,21 | 5,7 | -0,7 | 0,49 |
| 7 | 5,5 | -0,9 | 0,81 | 5,8 | -0,6 | 0,36 |
| 8 | 5,4 | -1,0 | 1,00 | 5,9 | -0,5 | 0,25 |
| 9 | 5,9 | -0,5 | 0,25 | 6,0 | -0,4 | 0,16 |
| 10 | 6,5 | 0,1 | 0,01 | 6,1 | -0,3 | 0,09 |
| 11 | 5,8 | -0,6 | 0,36 | 6,1 | -0,3 | 0,09 |
| 12 | 6,9 | 0,5 | 0,25 | 6,2 | -0,2 | 0,04 |
| 13 | 5,8 | -0,6 | 0,36 | 6,2 | -0,2 | 0,04 |
| 14 | 5,7 | -0,7 | 0,49 | 6,4 | 0,0 | 0,00 |
| 15 | 7,0 | 0,6 | 0,36 | 6,6 | 0,2 | 0,04 |
| 16 | 6,6 | 0,2 | 0,04 | 6,6 | 0,2 | 0,04 |
| 17 | 7,1 | 0,7 | 0,49 | 6,8 | 0,4 | 0,16 |
| 18 | 8,0 | 1,6 | 2,56 | 6,8 | 0,4 | 0,16 |
| 19 | 7,5 | 1,1 | 1,21 | 7,1 | 0,7 | 0,49 |
| 20 | 6,9 | 0,5 | 0,25 | 7,1 | 0,7 | 0,49 |
| 21 | 6,8 | 0,4 | 0,16 | 7,2 | 0,8 | 0,64 |
| 22 | 7,3 | 0,9 | 0,81 | 7,6 | 1,2 | 1,44 |
| 23 | 7,7 | 1,3 | 1,69 | 7,8 | 1,4 | 1,96 |
| 24 | 7,9 | 1,5 | 2,25 | 7,9 | 1,5 | 2,25 |
| 25 | 8,0 | 1,6 | 2,56 | 8,0 | 1,6 | 2,56 |
| Разом: | 159,8 | x | 23,90 | 159,8 | x | 19,48 |

Значення кореляційного відношення досить близьке до одиниці, що свідчить про значну силу кореляційного зв’язку між досліджуваними ознаками.

В нашому прикладі середня помилка кореляційного відношення дорівнює:

$$\mu_\eta = \frac{1-\eta^2}{\sqrt{n-1}} = \frac{1-0{,}815}{\sqrt{24}} = 0{,}038\ .$$

Відношення значення кореляційного відношення до його помилки дорівнює:

$$t_\eta = \frac{\eta}{\mu_\eta} = \frac{0{,}903}{0{,}038} = 23{,}76\ .$$

Так як критерій Стьюдента більший за три, то можна вважати, що обчислене кореляційне відношення досить точно характеризує силу зв’язку між обсягом товарообороту і виробітком на одного продавця.

## 10.5. Багатофакторний кореляційно-регресійний аналіз

В багатьох випадках на результативну ознаку впливає не один, а декілька факторів. Між факторами існують складні взаємозв’язки, тому їх вплив на результативну ознаку комплексний і його не можна розглядати як просту суму ізольованих впливів.

*Багатофакторний кореляційно-регресійний аналіз* дозволяє оцінити міру впливу на досліджуваний результативний показник кожного із введених в модель факторів при зафіксованому на середньому рівні інших факторів. При цьому важливою умовою тут є відсутність функціонального зв’язку між факторами.

Математично завдання зводиться до знаходження аналітичного виразу, який найкращим чином відображав би зв’язок факторних ознак з результативною, тобто знайти функцію:

$$\hat{y} = f(x_1, x_2, \ldots, x_n).$$

Найбільш складною проблемою є вибір форми зв’язку, яка виражається аналітичним рівнянням, на основі котрого за існуючими факторами визначаються значення результативної ознаки – функції. Ця функція повинна краще за інші відображати реально існуючі зв’язки між досліджуваним показником і факторами. Емпіричне обґрунтування типу функції за допомогою графічного аналізу зв’язків для багатофакторних моделей практично не придатне.

Форму зв’язку можна визначати шляхом перебору функцій різних типів, але це зв’язане з великою кількістю зайвих розрахунків. Але, беручи до уваги, що любу функцію багатьох змінних шляхом логарифмування або заміни змінних можна звести до лінійного виду, рівняння множинної регресії можна будувати в лінійній формі:

$$\hat{y}_x = a_0 + a_1x_1 + a_2x_2 + \ldots + a_nx_n\,.$$

Параметри рівняння знаходять за способом найменших квадратів.

Так, для розрахунку параметрів рівняння лінійної двофакторної регресії

$$\hat{y}_x = a_0 + a_1x_1 + a_2x_2\,,$$

де $\hat{y}_x$ – розраховані значення результативної ознаки – функції;

$x_1$ і $x_2$ – факторні ознаки;

$a_0$, $a_1$ і $a_2$ – параметри рівняння,

будується система нормальних рівнянь:

$$\begin{cases} \sum y = na_0 + a_1\sum x_1 + a_2\sum x_2\,; \\ \sum yx_1 = a_0\sum x_1 + a_1\sum x_1^2 + a_2\sum x_1x_2\,; \\ \sum yx_2 = a_0\sum x_2 + a_1\sum x_1x_2 + a_2\sum x_2^2. \end{cases}$$

Кожний коефіцієнт рівняння показує ступінь впливу відповідного фактора на результативний показник при фіксованому положенні решти факторів, тобто, як із зміною окремого фактора на одиницю змінюється результативний показник. Вільний член рівняння множинної регресії економічного змісту не має.

Звернемось до прикладу. Нехай маємо дані по 10 робітниках підприємства про їхній стаж роботи, тарифний розряд і денну заробітну плату. Потрібно встановити залежність заробітної плати (y) від двох факторів: стажу роботи робітників ($x_1$) і тарифного розряду ($x_2$), для чого складемо розрахункову таблицю 10.17.

Підставляємо отримані дані в систему нормальних рівнянь:

$$106 = 10\,a_0 + 87\,a_1 + 41\,a_2;$$
$$1183 = 87\,a_0 + 1009\,a_1 + 416\,a_2;$$
$$502 = 41\,a_0 + 416\,a_1 + 189\,a_2.$$

Для розв'язку системи нормальних рівнянь поділимо всі члени рівнянь на коефіцієнти при $a_0$:

$$10{,}6 = a_0 + 8{,}7\,a_1 + 4{,}1\,a_2;$$
$$13{,}6 = a_0 + 11{,}6\,a_1 + 4{,}78\,a_2;$$
$$12{,}244 = a_0 + 10{,}146\,a_1 + 4{,}61\,a_2.$$

Віднімемо від другого рівняння перше, а від третього рівняння друге:

$$3 = 2{,}9\,a_1 + 0{,}68\,a_2;$$
$$-1{,}356 = -1{,}454\,a_1 - 0{,}17\,a_2.$$

Таблиця 10.17

| Номер робітника n | Стаж роботи, років $x_1$ | Тариф-ний розряд $x_2$ | Денна заробітна плата, грн. y | $yx_1$ | $yx_2$ | $x_1^2$ | $x_2^2$ | $y^2$ | $x_1x_2$ | $\hat{y}_x$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 3 | 3 | 6 | 1 | 4 | 9 | 2 | 2,3 |
| 2 | 3 | 3 | 6 | 18 | 18 | 9 | 9 | 36 | 9 | 5,0 |
| 3 | 6 | 3 | 5 | 30 | 15 | 36 | 9 | 25 | 18 | 7,4 |
| 4 | 5 | 2 | 7 | 35 | 14 | 25 | 4 | 49 | 10 | 5,7 |
| 5 | 8 | 5 | 10 | 80 | 50 | 64 | 25 | 100 | 40 | 10,8 |
| 6 | 10 | 4 | 9 | 90 | 36 | 100 | 16 | 81 | 40 | 11,6 |
| 7 | 9 | 6 | 13 | 117 | 78 | 81 | 36 | 169 | 54 | 12,5 |
| 8 | 15 | 5 | 18 | 270 | 90 | 225 | 25 | 324 | 75 | 16,6 |
| 9 | 12 | 5 | 15 | 180 | 75 | 144 | 25 | 225 | 60 | 14,1 |
| 10 | 18 | 6 | 20 | 360 | 120 | 324 | 36 | 400 | 108 | 20,0 |
| Разом | 87 | 41 | 106 | 1183 | 502 | 1009 | 189 | 1418 | 416 | 106,0 |
| В серед-ньому | 8,7 | 4,1 | 10,6 | 118,3 | 50,2 | 100,9 | 18,9 | 141,8 | 41,6 | 10,6 |

Поділимо кожний член обох рівнянь на коефіцієнти при $a_1$ і віднімемо від першого рівняння друге:

$$\begin{array}{r} 1{,}034 = a_1 + 0{,}234\ a_2 \\ - \\ 0{,}932 = a_1 + 0{,}117\ a_2 \\ \hline 0{,}102 = 0{,}117\ a_2 \end{array}$$

Звідси $a_2 = \dfrac{0{,}102}{0{,}117} = 0{,}872$ .

Підставляючи значення параметра $a_2$ в рівняння, отримаємо параметри $a_1$ і $a_0$:

$$a_1 = 1{,}034 - 0{,}234 \cdot 0{,}872 = 0{,}83;$$
$$a_0 = 10{,}6 - 8{,}7 \cdot 0{,}83 - 4{,}1 \cdot 0{,}872 = -0{,}196.$$

Рівняння зв'язку, яке визначає залежність результативної ознаки від двох факторних матиме вигляд:

$$\hat{y} = -0{,}196 + 0{,}83\,x_1 + 0{,}872\,x_2\,.$$

Таким чином, із збільшенням стажу роботи робітника на 1 рік, денна заробітна плата зросте на 0,83 грн., а збільшення тарифного розряду на 1 дає ріст в заробітній платі в розмірі 0,872 грн.

Підставляючи в це рівняння значення $x_1$ і $x_2$, отримаємо відповідні значення змінної середньої (остання графа табл. 10.17), які досить близько відтворюють значення фактичних рівнів заробітної плати. Це свідчить про правильний вибір форми математичного вираження кореляційного зв'язку між трьома досліджуваними ознаками.

Однак на основі коефіцієнтів регресії не можна судити, яка із факторних ознак найбільше впливає на результативну ознаку, так як коефіцієнти регресії між собою не порівняльні, оскільки вони володіють різними одиницями виміру. З метою виявлення порівняльної сили впливу окремих факторів і резервів, які закладені в них, статистика вираховує часткові коефіцієнти еластичності ($\varepsilon_i$), а також бета-коефіцієнти ($\beta_i$) за формулами:

$$\varepsilon_i = a_i \frac{\bar{x}_i}{\bar{y}}; \ \beta_i = a_i \frac{\sigma_{x_i}}{\sigma_y},$$

де $a_i$ – коефіцієнт регресії при і-му факторі;

$\bar{x}_i$ – середнє значення і-го фактора;

$\bar{y}$ – середнє значення результативної ознаки;

$\sigma_{x_i}$ – середнє квадратичне відхилення і-го фактора;

$\sigma_y$ – середнє квадратичне відхилення результативної ознаки.

*Часткові коефіцієнти еластичності* показують, на скільки відсотків в середньому зміниться результативна ознака із зміною на 1 % кожного фактора при фіксованому положенні інших факторів.

Для визначення факторів, в розвитку котрих закладені найбільші резерви покращення досліджуваної ознаки, з врахуванням ступеня варіації факторів рівняння множинної регресії, вираховують *часткові β-коефіцієнти*, які показують на яку частину середнього квадратичного відхилення змінюється результативна ознака із зміною відповідної факторної ознаки на величину її середнього квадратичного відхилення.

Розрахуємо коефіцієнти еластичності і β-коефіцієнти для нашого прикладу:

$$\varepsilon_1 = a_1 \frac{\bar{x}_1}{\bar{y}} = 0{,}83 \cdot \frac{8{,}7}{10{,}6} = 0{,}681;$$

$$\varepsilon_2 = a_2 \frac{\bar{x}_2}{\bar{y}} = 0{,}872 \cdot \frac{4{,}1}{10{,}6} = 0{,}337.$$

Аналіз часткових коефіцієнтів еластичності показує, що за абсолютним приростом найбільший вплив на заробітну плату робітників має стаж роботи – фактор $x_1$, із збільшенням якого на 1 % заробіток приросте на 0,68 %, а при збільшенні тарифного розряду на 1 % заробітна плата підвищиться на 0,34 %.

Для розрахунку β-коефіцієнтів потрібно вирахувати відповідні середні квадратичні відхилення:

$$\sigma_{x_1} = \sqrt{\overline{x_1^2} - \left(\overline{x_1}\right)^2} = \sqrt{100{,}9 - 8{,}7^2} = \sqrt{25{,}21} = 5{,}02\ ;$$

$$\sigma_{x_2} = \sqrt{\overline{x_2^2} - \left(\overline{x_2}\right)^2} = \sqrt{18{,}9 - 4{,}1^2} = \sqrt{2{,}09} = 1{,}44\ ;$$

$$\sigma_{y} = \sqrt{\overline{y^2} - \left(\overline{y}\right)^2} = \sqrt{141{,}8 - 10{,}6^2} = \sqrt{29{,}44} = 5{,}42\ .$$

Тоді:

$$\beta_1 = a_1 \frac{\sigma_{x_1}}{\sigma_y} = 0{,}83 \cdot \frac{5{,}02}{5{,}42} = 0{,}769\ ;$$

$$\beta_2 = a_2 \frac{\sigma_{x_2}}{\sigma_y} = 0{,}872 \cdot \frac{1{,}44}{5{,}42} = 0{,}232\ .$$

Аналіз β-коефіцієнтів показує, що на заробітну плату робітників найбільший вплив з двох досліджуваних факторів має фактор $x_1$ – стаж роботи, так як йому відповідає найбільше значення β-коефіцієнта.

Для характеристики ступеня тісноти зв’язку в множинній прямолінійній кореляції використовують *множинний коефіцієнт кореляції*, формула якого має вигляд:

$$R_{yx_1x_2} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1} \cdot r_{yx_2} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2}}\ ,$$

де $r_{yx_1}; r_{yx_2}$ і $r_{x_1x_2}$ – парні коефіцієнти лінійної кореляції, які визначаються за формулами:

$$r_{yx_1} = \frac{\overline{yx_1} - \overline{y} \cdot \overline{x}_1}{\sigma_y \cdot \sigma_{x_1}}\ ;$$

$$r_{yx_2} = \frac{\overline{yx_2} - \overline{y} \cdot \overline{x}_2}{\sigma_y \cdot \sigma_{x_2}}\ ;$$

$$r_{x_1x_2} = \frac{\overline{x_1x_2} - \overline{x}_1 \cdot \overline{x}_2}{\sigma_{x_1} \cdot \sigma_{x_2}}\ .$$

Множинний коефіцієнт кореляції показує, яку частину загальної кореляції складають коливання, під впливом факторів $x_1$, $x_2$, ..., $x_n$ закладених в багатофакторну модель для дослідження.

Множинний коефіцієнт кореляції коливається в межах від 0 до ±1. При R = 0 зв’язок між досліджуваними ознаками відсутній, при R=1 функціональний.

Перш ніж розрахувати множинний коефіцієнт кореляції, потрібно вирахувати парні коефіцієнти кореляції

$$r_{yx_1} = \frac{\overline{yx}_1 - \bar{y}\cdot\bar{x}_1}{\sigma_y \cdot \sigma_{x_1}} = \frac{118{,}3 - 10{,}6\cdot 8{,}7}{5{,}42\cdot 5{,}02} = 0{,}995\ ;$$

$$r_{yx_2} = \frac{\overline{yx}_2 - \bar{y}\cdot\bar{x}_2}{\sigma_y \cdot \sigma_{x_2}} = \frac{50{,}2 - 10{,}6\cdot 4{,}1}{5{,}42\cdot 1{,}44} = 0{,}864\ ;$$

$$r_{x_1x_2} = \frac{\overline{x_1x_2} - \bar{x}_1\cdot\bar{x}_2}{\sigma_{x_1} \cdot \sigma_{x_2}} = \frac{41{,}6 - 8{,}7\cdot 4{,}1}{5{,}02\cdot 1{,}44} = 0{,}820\ .$$

Високі значення парних коефіцієнтів кореляції свідчать про сильний вплив (окремо) стажу роботи і тарифного розряду на заробітну плату робітників.

На основі парних коефіцієнтів кореляції можна розрахувати часткові коефіцієнти першого порядку:

$$r_{yx_1(x_2)} = \frac{r_{yx_1} - r_{yx_2}\cdot r_{x_1x_2}}{\sqrt{\left(1 - r^2_{yx_2}\right)\cdot\left(1 - r^2_{x_1x_2}\right)}} = \frac{0{,}995 - 0{,}864\cdot 0{,}820}{\sqrt{\left(1 - 0{,}864^2\right)\cdot\left(1 - 0{,}820^2\right)}} = \frac{0{,}286}{0{,}293} = 0{,}976\ ;$$

$$r_{yx_2(x_1)} = \frac{r_{yx_2} - r_{yx_1}\cdot r_{x_1x_2}}{\sqrt{\left(1 - r^2_{yx_1}\right)\cdot\left(1 - r^2_{x_1x_2}\right)}} = \frac{0{,}864 - 0{,}995\cdot 0{,}820}{\sqrt{\left(1 - 0{,}995^2\right)\cdot\left(1 - 0{,}820^2\right)}} = \frac{0{,}0481}{0{,}0572} = 0{,}841\ .$$

Як бачимо з розрахунків часткових коефіцієнтів кореляції, зв'язок кожного фактора з досліджуваним показником за умови комплексної взаємодії факторів дещо слабший але достатньо тісний.

Для виявлення тісноти зв'язку між результативною ознакою і обома факторними ознаками одночасно вирахуємо сукупний коефіцієнт множинної кореляції:

$$R_{yx_1x_2} = \sqrt{\frac{r^2_{yx_1} + r^2_{yx_2} - 2r_{yx_1}\cdot r_{yx_2}\cdot r_{x_1x_2}}{1 - r^2_{x_1x_2}}} =$$

$$= \sqrt{\frac{0{,}995^2 + 0{,}864^2 - 2\cdot 0{,}995\cdot 0{,}864\cdot 0{,}820}{1 - 0{,}820^2}} = \sqrt{0{,}995} = 0{,}998,\ \text{або}$$

$$R_{yx_1x_2} = \sqrt{1 - \left[\left(1 - r^2_{yx_1}\right)\cdot\left(1 - r^2_{x_1x_2}\right)\right]} = \sqrt{1 - \left[\left(1 - 0{,}995^2\right)\cdot\left(1 - 0{,}820^2\right)\right]} =$$

$$= \sqrt{1 - \left(0{,}009975\cdot 0{,}3276\right)} = \sqrt{0{,}9967322} = 0{,}998\ .$$

Вирахуваний коефіцієнт множинної кореляції (R = 0,998) показує, що між двома факторними і результативною ознаками існує достатньо тісний зв'язок.

Сукупний коефіцієнт множинної детермінації ($R^2 = 0,995$) свідчить про те, що варіація заробітної плати робітників на 99,5 % обумовлюється двома факторами (стажем роботи і тарифним розрядом) введеними в кореляційну модель. Це означає, що вибрані фактори суттєво впливають на досліджуваний показник.

З метою більш глибокого економічного аналізу збільшують число суттєвих факторів, які включають в модель досліджуваного показника і будують багатофакторні рівняння регресії. Їх розраховують з допомогою персональних комп'ютерів.

### 10.6. Непараметричні показники тісноти зв'язку

Поряд з вивченням кореляційної залежності між кількісними показниками статистика встановлює також зв'язки і між якісними ознаками.

При вивченні залежності між якісними ознаками встановлюють наявність зв'язку і вимірюють його тісноту.

Для вимірювання тісноти зв'язку між двома ознаками, які мають альтернативний вираз застосовується *коефіцієнт асоціації*, запропонований статистиком Юлом.

З метою розрахунку коефіцієнта асоціації використовують таблицю, яка складається з чотирьох комірок позначених латинськими літерами a, b, c, d. Кожна з комірок відповідає відомій альтернативі тієї чи іншої ознаки. Така таблиця має вигляд:

Таблиця 10.18

| Ознака | A | не A | $\sum$ B |
|---|---|---|---|
| B | a | b | a + b |
| не B | c | d | c + d |
| $\sum$ A | a + c | b + d | a + b + c + d |

Коефіцієнт асоціації визначається за формулою:

$$Q = \frac{ad - bc}{ad + bc}.$$

Розглянемо конкретний приклад. Досліджується вплив мінеральних добрив на урожайність озимої пшениці. З цією метою нами побудована таблиця, в якій розставлені розміри посівів в кожній групі (га).

Таблиця 10.19

Залежність урожайності озимої пшениці від внесення мінеральних добрив в ґрунт

| Ділянки за урожайністю | Внесені добрива | Не внесені добрива | Всього |
|---|---|---|---|
| Підвищили урожайність | 110 | 10 | 120 |
| Не підвищили урожайність | 10 | 70 | 80 |
| Всього | 120 | 80 | 200 |

Коефіцієнт асоціацій дорівнює:

$$Q = \frac{ad - bc}{ad + bc} = \frac{110 \cdot 70 - 10 \cdot 10}{110 \cdot 70 + 10 \cdot 10} = 0{,}974 .$$

Отже, спостерігається досить тісний зв'язок між удобренням ділянок і урожайністю озимої пшениці.

Для дослідження кореляції альтернативних ознак Юл запропонував також і другий коефіцієнт – *коефіцієнт колігації*, який визначається за формулою:

$$W = \frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}} .$$

За даними таблиці 10.19 коефіцієнт колігації дорівнює:

$$W = \frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}} = \frac{\sqrt{7700} - \sqrt{100}}{\sqrt{7700} + \sqrt{100}} = \frac{77{,}7}{97{,}7} = 0{,}795 .$$

Чотири кліткова таблиця дозволила К.Пірсону винайти ще інший показник, який називається *коефіцієнтом контингенції*. Він розраховується за формулою:

$$K = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} = \frac{7600}{\sqrt{120 \cdot 80 \cdot 120 \cdot 80}} = \frac{7600}{9600} = 0{,}792 .$$

Коефіцієнт колігації як і коефіцієнт контингенції оцінюють зв'язок між внесенням мінеральних добрив в ґрунт і урожайністю озимої пшениці більш обережно, однак показують достатньо сильний зв'язок між цими ознаками.

Величина цих коефіцієнтів як показників зв'язку тлумачиться так само, як і коефіцієнт кореляції.

Коефіцієнти асоціації, колігації і контингенції є свого роду коефіцієнтами кореляції для якісних ознак. Причому завжди: Q>W>K.

В тому випадку, коли обидві взаємозв'язані ознаки розбиті більш ніж на дві групи, для вимірювання тісноти зв'язку застосовуються *показники взаємного спряження*, запропоновані К.Пірсоном і А.Чупровим.

Коефіцієнт взаємного спряження К.Пірсона вираховується за формулою:

$$C_1 = \sqrt{\frac{\varphi^2}{\varphi^2 + 1}} ,$$

де $\varphi$ (грецька буква “фі”) – сума квадратів частот кожної стрічки, розділеної на суму частот по графах, і в свою чергу, на суму частот по стрічці без одиниці.

Розглянемо приклад. Потрібно дослідити вплив термінів посівів ярої пшениці на врожайність. За обома ознаками сукупність із 500 дослідних ділянок розбиваємо на три групи (табл. 10.20).

Таблиця 10.20

| Врожай ($f_i$) | Терміни посівів ($f_j$) | | | Разом ($F_j$) |
|---|---|---|---|---|
| | пізній | середній | ранній | |
| Низький | 40 | 55 | 5 | 100 |
| Середній | 55 | 215 | 30 | 300 |
| Високий | 5 | 30 | 65 | 100 |
| Разом ($F_i$) | 100 | 300 | 100 | 500 |

Визначаємо $\varphi^2$ за формулою:

$$\varphi^2 = \sum \frac{f_{ij}^2}{F_i \cdot F_j} - 1 = \left(\frac{40^2}{100 \cdot 100}\right) + \left(\frac{55^2}{300 \cdot 100}\right) + \left(\frac{5^2}{100 \cdot 100}\right) + \left(\frac{55^5}{100 \cdot 300}\right) +$$

$$+ \left(\frac{215^2}{300 \cdot 300}\right) + \left(\frac{30^2}{100 \cdot 300}\right) + \left(\frac{5^2}{100 \cdot 100}\right) + \left(\frac{30^2}{300 \cdot 100}\right) + \left(\frac{65^2}{100 \cdot 100}\right) - 1 =$$

$$= 1{,}3627 - 1 = 0{,}3627 .$$

Отже, коефіцієнт взаємного спряження К.Пірсона буде дорівнювати:

$$C_1 = \sqrt{\frac{\varphi^2}{\varphi^2 + 1}} = \sqrt{\frac{0{,}3627}{1{,}3627}} = \sqrt{0{,}2662} = 0{,}516 .$$

Даний коефіцієнт свідчить про досить тісний зв’язок між урожайністю ярої пшениці і термінами посіву.

Коефіцієнт взаємного спряження А.Чупрова визначається за формулою:

$$C_2 = \sqrt{\frac{\varphi^2}{(K_1 - 1) \cdot (K_2 - 1)}} ,$$

де $K_1$ – число груп по графах;

$K_2$ – число груп по стрічках.

За даними попереднього прикладу цей коефіцієнт становитиме:

$$C_2 = \sqrt{\frac{\varphi^2}{(K_1 - 1) \cdot (K_2 - 1)}} = \sqrt{\frac{0{,}363}{2 \cdot 2}} = 0{,}301 .$$

Результат отриманий за коефіцієнтом взаємного спряження А.Чупрова більш точний, оскільки він враховує число груп по кожній з досліджуваних ознак. Його вигідно використовувати і при більшому розподілі одиниць сукупності на групи по взаємозв'язаних ознаках. Коефіцієнт взаємного спряження К.Пірсона застосовується в основному у випадку квадратної таблиці, тоді як А.Чупрова – придатний для вимірювання зв'язків в прямокутних таблицях.

Вважається, що вже при значенні коефіцієнтів взаємного спряження 0,3 можна говорити про тісний зв'язок між варіацією досліджуваних ознак.

Якщо одна із взаємозв'язаних ознак має кількісний вираз, а друга – альтернативний, то показником тісноти зв'язку виступає *бісеріальний коефіцієнт кореляції* (бі-серія – дві серії). Цей коефіцієнт визначається за формулою:

$$r = \frac{\overline{y}_1 - \overline{y}_2}{\sigma_y} \cdot \frac{p \cdot q}{z},$$

де $\overline{y}_1$ – середня ознака по першій альтернативній групі;

$\overline{y}_2$ – середня ознака по другій альтернативній групі;

$\sigma_y$ – середнє квадратичне відхилення по обох групах;

p – доля першої групи;

q – доля другої групи;

z – ордината нормальної кривої, яка ділить її площу у відношенні p : q (значення z для різних p наведені в табл. 10.21).

Таблиця 10.21

Значення z для бісеріального коефіцієнта кореляції при різних p. (Для $p \le 0{,}50$ знаходять z для $q = 1 - p$)

| P | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0,5 | 0,3889 | 3988 | 3984 | 3978 | 3969 | 3958 | 3944 | 3928 | 3909 | 3888 |
| 0,6 | 3863 | 3837 | 3808 | 3776 | 3741 | 3704 | 3664 | 3622 | 3576 | 3528 |
| 0,7 | 3477 | 3429 | 3366 | 3307 | 3244 | 3178 | 3109 | 3037 | 2961 | 2882 |
| 0,8 | 2800 | 2714 | 2624 | 2531 | 2433 | 2332 | 2226 | 2116 | 2000 | 1880 |
| 0,9 | 1755 | 1624 | 1487 | 1343 | 1191 | 1031 | 0862 | 0680 | 0484 | 0267 |

Прикладом такого взаємозв'язку може служити наступна дворядна таблиця.

Таблиця 10.22

Взаємозв’язок між процентом виконання норм виробіток і кваліфікацією робітників.

| Кваліфікація робітників | Процент виконання норм (y) | | | | | Всього робітників |
|---|---|---|---|---|---|---|
| | до 100 | 100–110 | 110–120 | 120–130 | понад 130 | |
| Закінчили ПТУ ($f_1$) | 8 | 36 | 34 | 50 | 72 | 200 |
| Не закінчили ПТУ ($f_2$) | 20 | 18 | 30 | 26 | 6 | 100 |
| Разом (f) | 28 | 54 | 64 | 76 | 78 | 300 |

Розрахуємо потрібні дані для нашого прикладу:

$$\bar{y}_1 = \frac{\sum y f_1}{\sum f_1} = \frac{24420}{200} = 122{,}1\ \%;\ \bar{y}_2 = \frac{\sum y f_2}{\sum f_2} = \frac{11300}{100} = 113{,}0\ \%;$$

$$\bar{y}_з = \frac{\sum y f}{\sum f} = \frac{35720}{300} = 119{,}1\ \%;\ \sigma_y^2 = \frac{\sum (y - \bar{y}_з)^2 f}{\sum f} = \frac{50439}{300} = 168{,}13;$$

$$\sigma_y = \sqrt{168{,}13} = 12{,}966\ \%;\ p = \frac{200}{300} = 0{,}67;\ q = \frac{100}{300} = 0{,}33;\ z = 0{,}3622;$$

$$r = \frac{\bar{y}_1 - \bar{y}_2}{\sigma_y} \cdot \frac{p \cdot q}{z} = \frac{122{,}1 - 113{,}0}{12{,}966} \cdot \frac{0{,}2211}{0{,}3622} = 0{,}428.$$

Бісеріальний коефіцієнт кореляції показує достатньо тісний зв’язок між кваліфікацією робітників і виконанням норм виробіток.

Цей вплив може бути оцінений також через бісеріальний коефіцієнт, розрахований за іншою формулою:

$$r_1 = \frac{\bar{y}_1 - \bar{y}_з}{\sigma_y \sqrt{\frac{\sum f}{\sum f_1} - 1}} = \frac{122{,}0 - 119{,}1}{12{,}966\sqrt{1{,}5 - 1}} = 0{,}316;$$

$$r_2 = \frac{\bar{y}_2 - \bar{y}_з}{\sigma_y \sqrt{\frac{\sum f}{\sum f_2} - 1}} = \frac{113{,}0 - 119{,}1}{12{,}966\sqrt{3 - 1}} = -0{,}333,$$

де $r_1$ – показує прямий зв’язок виконання норм виробіток робітниками які закінчили ПТУ;

$r_2$ – обернений зв’язок виконання норм виробіток робітниками які не мають профтехосвіти.

Оцінка емпіричних мір тісноти зв'язку для всіх непараметричних показників здійснюється через t-критерій Стьюдента.

### *Контрольні запитання*

1. Поняття про взаємозв'язки суспільних явищ.
2. Які ознаки називаються факторними, а які результативними?
3. Яка різниця між функціональним і кореляційним зв'язком?
4. Поняття про прямі і обернені зв'язки.
5. Поняття про прямолінійні і криволінійні залежності.
6. Поняття про статистичний баланс.
7. Суть методу порівняння паралельних рядів.
8. Оцінка зв'язку за коефіцієнтом Фехнера.
9. Оцінка зв'язку за допомогою коефіцієнтів кореляції рангів.
10. Графічний метод виявлення кореляційної залежності.
11. Поняття про метод статистичних групувань.
12. Що показують емпіричне кореляційне відношення і коефіцієнт детермінації?
13. Для чого використовують F-критерій Фішера та t-критерій Стьюдента?
14. Поняття про кореляційно-регресійний аналіз.
15. В яких випадках використовують рівняння прямої?
16. Що показує коефіцієнт регресії?
17. Що показують коефіцієнти еластичності?
18. Поняття про індекс кореляції та теоретичне кореляційне відношення.
19. Для чого використовують лінійний коефіцієнт кореляції?
20. Поняття про нелінійні залежності.
21. Коли застосовують рівняння гіперболи?
22. В якому випадку використовують рівняння параболи другого порядку?
23. Вирівнювання за напівлогарифмічною кривою.
24. Поняття про багатофакторний кореляційно-регресійний аналіз.
25. Що показують β-коефіцієнти?
26. Що показує множинний коефіцієнт кореляції?
27. Що показують парні та часткові коефіцієнти кореляції?
28. Коли використовують непараметричні показники тісноти зв'язку?
29. Як визначається коефіцієнт асоціації, запропонований Юлом?
30. Що показує коефіцієнт контингенції Пірсона?
31. Поняття про показники взаємного спряження К.Пірсона і А.Чупрова.
32. В яких випадках застосовують бісеріальний коефіцієнт кореляції?

# Розділ 11. Ряди динаміки

## 11.1. Поняття про ряди динаміки, їх види та правила побудови

В статистичній практиці доводиться мати справу з великою кількістю чисел, що характеризують розвиток явищ в часі. Для кращого розуміння і аналізу досліджуваних статистичних даних, їх потрібно систематизувати, побудувавши хронологічні ряди, які називаються *рядами динаміки* або *часовими рядами*. Отже, *рядами динаміки* в статистиці називаються ряди чисел, що характеризують закономірності і особливості зміни суспільних явищ і процесів в часі.

Кожний ряд динаміки складається з двох елементів:

1) періодів або моментів часу, до яких відносяться рівні ряду (t);
2) статистичних показників, які характеризують рівні ряду (y).

В залежності від характеру рівнів ряду розрізняють два види рядів динаміки: моментні і інтервальні (періодичні).

*Моментним* називається ряд динаміки, величини якого характеризують стан явищ на певний момент часу, прикладом моментного ряду динаміки можуть служити дані, наведені в таблиці 11.1.

Таблиця 11.1

Парк тракторів в сільських спілках району.

| Дані на початок місяця | 1.01 | 1.02 | 1.03 | 1.04 | 1.05 | 1.06 | 1.07 | 1.08 | 1.09 | 1.10 | 1.11 | 1.12 | 1.01 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Число тракторів, шт. | 622 | 640 | 643 | 640 | 664 | 670 | 682 | 733 | 753 | 768 | 800 | 826 | 888 |

Рівні моментного ряду сумувати не має змісту, так як це приведе до повторного рахунку, але різниця рівнів має певний економічний зміст.

*Інтервальним* називається такий ряд динаміки, величини якого характеризують розміри суспільних явищ за певні періоди часу (день, місяць, квартал і т.д.) Прикладом інтервального ряду динаміки можуть служити наступні дані:

Таблиця 11.2

Валовий збір картоплі в області за 2002 – 2007 рр.

| Рік | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|---|---|
| Валовий збір картоплі, млн. т. | 20 | 23 | 27 | 28 | 30 | 33 |

Важливе економічне значення має сумування рівнів інтервального ряду. Сума рівнів інтервального ряду динаміки характеризує рівень даних явища за більш тривалий проміжок часу.

Ряди динаміки бувають одномірні і багатомірні.

*Одномірні* ряди динаміки характеризують зміну одного показника (валовий збір картоплі).

*Багатомірні* ряди динаміки характеризують зміну двох, трьох і більше показників.

В свою чергу, багатомірні динамічні ряди поділяються на паралельні і ряди взаємозв'язаних показників.

*Паралельні* ряди динаміки відображають зміну або одного і того самого показника щодо різних об'єктів (чисельність населення по країнах), або різних показників щодо одного і того самого об'єкта (валовий збір пшениці, цукрових буряків і картоплі в районі).

Ряди взаємозв'язаних показників характеризують залежність одного явища від іншого (залежність заробітної плати робітників від їхнього тарифного розряду).

За повнотою часу динамічні ряди поділяються на повні і неповні.

В *повних* динамічних рядах дати або періоди ідуть один за одним з рівними інтервалами.

В *неповних* динамічних рядах в послідовності часу спостерігаються нерівні інтервали.

За способом вираження рівнів динамічного ряду вони поділяються на ряди *абсолютних*, *середніх* і *відносних* величин. Наведені ряди в таблицях 11.1 і 11.2 показують зміну в часі абсолютних показників.

Ряд динаміки середніх величин показаний в таблиці 11.3.

Таблиця 11.3

Середня урожайність озимої пшениці в господарствах району.

| Рік | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|---|---|
| Урожайність озимої пшениці, ц/га. | 45,8 | 48,3 | 52,5 | 55,8 | 57,4 | 60,0 |

Дані таблиці показують безперервний ріст урожайності озимої пшениці в господарствах району.

Ряд динаміки відносних величин наведемо в таблиці 11.4.

Таблиця 11.4

Темпи росту роздрібного товарообороту в торгових підприємствах споживчої кооперації області.

| Рік | 2002 | 2003 | 2004 | 2005 | 2006 | 2002 |
|---|---|---|---|---|---|---|
| Товарооборот в % до 2002 р. | 100 | 105 | 108 | 115 | 123 | 134 |

При формуванні динамічних рядів для наукового дослідження розвитку суспільних явищ в часі потрібно дотримуватись правил їх побудови. Важливим правилом побудови динамічних рядів є вимога порівняльності всіх рівнів ряду між собою. Показники ряду динаміки повинні бути порівняльні за територією, колом охоплюваних об'єктів, способами розрахунків, періодами часу, одиницями виміру.

Важливою вимогою любих динамічних порівнянь є вимога *порівняльності території*, до котрої відносяться рівні динамічного ряду. Межі територіальних одиниць держав, областей, районів на протязі досліджуваного періоду змінюються внаслідок приєднання до них нових територій, або відокремлення певних частин їх територій. В кожному окремому випадку питання порівняльності розв'язується в залежності від мети дослідження. Для приведення даних динамічного ряду до порівняльного виду проводиться перерахунок попередніх даних з врахуванням нових меж (кордонів).

Статистичні дані, які необхідні для побудови ряду динаміки повинні бути порівняльні за *колом охоплюваних об'єктів*. Непорівняльність може виникнути внаслідок переходу деяких об'єктів із одного підпорядкування в інше. Проте порівняльність зберігається, якщо появилось побудоване нове підприємство, або якесь вибуле підприємство зупинило роботу.

Порівняльність за колом охоплюваних об'єктів забезпечується *зімкненням динамічних рядів* шляхом заміни абсолютних рівнів відносними.

Наведемо приклад зімкнення динамічного ряду продукції фірми "Юмакс", в котру за 1999 р. входило 8 підприємств, а в 2003 р. добавилось ще 2 підприємства, які перейшли із другої фірми.

Таблиця 11.5

Реалізована продукція фірми "Юмакс", млн. грн.

| Рік | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|---|---|---|---|---|
| Продукція 8-и підприємств | 20 | 25 | 30 | 40 | 50 | | | | |
| Продукція 10-и підприємств | | | | | 70 | 75 | 80 | 93 | 100 |

Для побудови зімкненого ряду відносних величин, дані 2003 р. приймаються за 100 % як для попередніх, так і для наступних років. При цьому для попередніх років за 100 % береться 50 млн. грн., а для наступного – 70 млн. грн. Провівши відповідні розрахунки, отримаємо зімкнений ряд динаміки реалізації продукції даної фірми у відсотках до 2003 р.

Таблиця 11.6

Динаміка реалізації продукції фірми "Юмакс".

| Рік | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|---|---|---|---|---|
| Відносний рівень в % до 2003 р. | 40,0 | 50,0 | 60,0 | 80,0 | 100,0 | 107,1 | 114,3 | 132,8 | 142,8 |

Таким чином, із двох непорівняльних рядів отримали один зімкнений ряд реалізованої продукції даної фірми, рівні якого придатні для аналізу.

В моментних рядах динаміки виникає непорівняльність за *критичним моментом реєстрації* рівнів явищ, які піддаються сезонним коливанням.

Для прикладу візьмемо чисельність великої рогатої худоби, яка літом, як правило, завжди більша ніж зимою. А тому, не можна вважати порівняльний ряд динаміки, рівні якого відносяться до різних дат реєстрації.

Рівні динамічного ряду повинні бути порівняльні за *методикою їх розрахунку*. Наприклад, за попередні роки чисельність робітників заводу була визначена на початок кожного місяця, тобто на певну дату, а в наступні роки – як середньомісячна чисельність. Для забезпечення порівняльності показників ряду, потрібно дані про чисельність робітників заводу на початок місяця перерахувати в середньомісячну чисельність.

Статистичні дані динамічного ряду можуть бути непорівняльними за *різними періодами або тривалістю часу*. Для того щоб виявити закономірності розвитку явищ із значними сезонними коливаннями потрібно порівнювати між собою дані за одні і ті ж періоди часу і за однаковою тривалістю періоду. Так, наприклад, обсяг виробництва молока за різні роки потрібно порівнювати тільки таким чином: січень з січнем, квітень з квітнем, липень з липнем і т.д. Також не можна порівнювати виробництво продукції за місяць і за квартал, за квартал і за півріччя і т.д. тобто, інтервали часу, за які наведені дані динамічного ряду, повинні бути рівні.

Важливе значення має правильне визначення інтервалів і віддалі між моментами в рядах динаміки, що залежить від характеру досліджуваних явищ. Для явищ, які змінюються повільно, інтервал беруть більш широкий, а для динамічних явищ, навпаки, потрібно брати менший розмір інтервалу.

Часто статистичні дані виражені в різних одиницях виміру, що веде до непорівняльності таких динамічних рядів. Непорівняльність через різні одиниці виміру виникає внаслідок того, що ряд явищ обліковується паралельно в двох одиницях виміру, або за однією із них. Наприклад, сталеві труби обліковуються в тоннах і метрах, електромотори – в штуках і кіловатах потужності і т.д. Будуючи ряд за одним вимірником, отримаємо одну динаміку розвитку явища, а за другим – іншу, які можуть бути протилежними. Вага сталевих труб може зростати, а метраж зменшуватись. Порівняльність за одиницями виміру вимагає, щоб рівні динамічного ряду завжди були виражені в одних і тих самих одиницях виміру. Однак статистичний аналіз такого ряду доповнюють паралельним аналізом за іншим способом вимірювання.

Непорівняльність рядів динаміки через одиниці виміру виникає і внаслідок непорівняльності грошової оцінки (міняється грошова одиниця, інфляція, змінюється курс валюти та ін.). Для приведення до порівняльного виду таких рядів динаміки всі попередні рівні досліджуваних ознак перераховуються за діючою грошовою оцінкою.

Непорівняльність статистичних показників динаміки може бути зумовлена також різною структурою сукупності за ряд років. Для приведення даних таких рядів до порівняльного виду використовують так звану стандартизацію структури (стандартизовані коефіцієнти народжуваності, смертності, природного приросту і т.д.), що в значній мірі зв’язано із віковою структурою населення в різні роки. В якості стандартної структури використовують структуру якогось одного періоду часу, а всі показники інших періодів розраховуються за цією ж структурою, що робить такі показники порівняльними.

## 11.2 Основні характеристики рядів динаміки

Завдання статистики полягає в тому, щоб шляхом аналізу рядів динаміки розкрити і охарактеризувати закономірності, що проявляються на різних етапах розвитку того чи іншого явища, виявити тенденції розвитку та їх особливості.

В процесі аналізу динаміки розраховують і використовують наступні аналітичні показники динаміки: абсолютний приріст, темп росту, темп приросту і абсолютне значення одного процента приросту.

Розрахунок цих показників ґрунтується на абсолютному або відносному порівнянні між собою рівнів ряду динаміки. При цьому порівнюваний рівень називається *поточним*, а рівень, з яким роблять порівняння – *базисним*. За базу порівняння часто приймають або попередній рівень, або початковий (перший) рівень ряду динаміки.

Якщо кожний рівень порівнюється з попереднім, то отримують *ланцюгові показники динаміки*, а якщо кожний рівень порівнюють з одним і тим же рівнем, взятим за базу порівняння, то такі показники називаються *базисними*.

За постійну базу порівняння можна прийняти не тільки початковий, але й будь-який інший рівень ряду динаміки. Інколи за базу порівняння приймають середній рівень любого попереднього періоду. Вибір бази порівняння повинен бути обґрунтований історичними та економічними особливостями розвитку досліджуваного явища.

Розглянемо порядок вирахування цих показників за такими даними.

Таблиця 11.7

Валовий збір картоплі в області за 1999 – 2007 рр.

| Рік | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|---|---|
| Валовий збір картоплі, млн. т. | 20 | 23 | 27 | 28 | 30 | 33 |

*Абсолютний приріст* (Δу) обчислюється як різниця між поточним та базисним рівнями і показує, на скільки одиниць підвищився або зменшився рівень порівняно з базисним, за певний період часу. Він виражається в тих же одиницях виміру, що й рівні ряду динаміки і визначається за формулою:

$$\Delta_{у}^{б} = у_{і} - у_{1}\text{, або } \Delta_{у}^{л} = у_{і} - у_{і-1}\text{,}$$

де Δу – абсолютний приріст;

$у_і$ – поточний рівень ряду динаміки;

$у_1$ – початковий (перший) рівень ряду динаміки;

$у_{і-1}$ – попередній рівень ряду динаміки.

В нашому прикладі абсолютний приріст складає (млн.. т.):

| $\Delta_{у}^{б} = у_{і} - у_{1}$ | $\Delta_{у}^{л} = у_{і} - у_{і-1}$ |
|---|---|
| В 2003 р. 23 – 20 = 3 млн. т.; | 23 – 20 = 3 млн. т.; |
| В 2004 р. 27 – 20 = 7 млн. т.; | 27 – 23 = 4 млн. т.; |
| В 2005 р. 28 – 20 = 8 млн. т.; | 28 – 27 = 1 млн. т.; |
| В 2006 р. 30 – 20 = 10 млн. т.; | 30 – 28 = 2 млн. т.; |
| В 2007 р. 33 – 20 = 13 млн. т.; | 33 – 30 = 3 млн. т. |

Ланцюгові і базисні абсолютні прирости взаємозв’язані між собою: сума і послідовних ланцюгових приростів, починаючи з першого, дорівнює і-му базисному приросту. Так, в нашому прикладі сума приростів за 2003 та 2004 рр. дорівнює базисному приросту 2004 р. (3 + 4 = 7 млн. т. і т.д.). Різниця між базисними приростами 2004 р. і 2003 р. буде дорівнювати ланцюговому приросту 2004 р. (7 – 3 = 4 млн. т. і т.д.).

Абсолютний приріст виражає абсолютну швидкість зміни рівнів ряду динаміки. Для більш вичерпної і глибокої характеристики явища або процесу абсолютні величини доповнюють відносними.

*Темп росту* ($Т_р$) вираховується як відношення порівнюваного рівня з рівнем, прийнятим за базу зіставлення, і показує, в скільки разів (відсотків) порівнюваний рівень більший або менший базисного. Темп росту обчислюється за формулою:

$$Т_р^б = \frac{у_i}{у_1}, \text{ або } Т_р^л = \frac{у_i}{у_{i-1}}.$$

В 2003 р. 23 : 20 = 1,150; 23 : 20 = 1,150;
В 2004 р. 27 : 20 = 1,350; 27 : 23 = 1,174;
В 2005 р. 28 : 20 = 1,400; 28 : 27 = 1,037;
В 2006 р. 30 : 20 = 1,500; 30 : 28 = 1,071;
В 2007 р. 33 : 20 = 1,650; 33 : 30 = 1,100.

Темп росту може бути виражений у вигляді коефіцієнтів, коли визначається безпосереднє відношення абсолютних розмірів рівнів, і у відсотках, коли він показує, скільки відсотків складає поточний рівень відносно до базисного, прийнятого за 100 %.

Між ланцюговими і базисними темпами росту існує певний взаємозв'язок. Добуток кількох послідовних ланцюгових темпів росту дорівнює базисному темпу росту за відповідний період і, навпаки, поділивши наступний базисний темп росту на попередній, отримаємо відповідний ланцюговий темп росту. Наприклад:

$$Т^л_{р(2003\,р.)} \cdot Т^л_{р(2004\,р.)} = Т^б_{р(2004\,р.)} = 1{,}150 \cdot 1{,}174 = 1{,}350;$$

$$Т^б_{р(2004\,р.)} : Т^б_{р(2003\,р.)} = Т^л_{р(2004\,р.)} = 1{,}350 \cdot 1{,}150 = 1{,}174$$

і т.д.

Взаємозв'язок ланцюгових та базисних темпів росту використовують для переходу від одних темпів росту до других в тих випадках, коли невідомі абсолютні рівні ряду динаміки.

*Темп приросту* ($Т_{пр}$) визначається як відношення абсолютного приросту до абсолютного попереднього або початкового рівня і показує на скільки відсотків порівнюваний рівень більший або менший рівня, прийнятого за базу порівняння.

$$Т_{пр}^б = \frac{\Delta_у^б}{у_1}, \text{ або } Т_{пр}^л = \frac{\Delta_у^л}{у_{i-1}}.$$

В 2003 р. 3 : 20 = 0,150; 3 : 20 = 0,150;
В 2004 р. 2 : 20 = 0,350; 4 : 23 = 0,174;
В 2005 р. 8 : 20 = 0,400; 1 : 27 = 0,037;

В 2006 р. 10 : 20 = 0,500; 2 : 28 = 0,071;
В 2007 р. 13 : 20 = 0,650; 3 : 30 = 0,100.

Між темпами росту і приросту існує безпосередній взаємозв'язок:

$$T^{б}_{пр} = T^{б}_{р} - 1\text{, або } T^{б}_{пр}(\%) = T^{б}_{р}(\%) - 100\ (\%)\ ;$$

$$T^{л}_{пр} = T^{л}_{р} - 1\text{, або } T^{л}_{пр}(\%) = T^{л}_{р}(\%) - 100\ (\%)\ .$$

$$T^{б}_{пр(2003\,р.)} = 1{,}150 - 1 = 0{,}150\ ;$$

і т.д.

$$T^{л}_{пр(2004\,р.)} = 1{,}174 - 1 = 0{,}174\ ;$$

і т.д.

*Абсолютне значення одного процента приросту* (А) визначається шляхом ділення абсолютного приросту на темп приросту за один і той самий період. Абсолютне значення одного відсотка приросту можна вирахувати технічно більш легким шляхом, діленням початкового рівня на 100, оскільки за 100 % завжди приймається базисний рівень, то 1 % буде в 100 раз менший від базисного рівня:

$$A = \frac{\Delta_{у}}{T_{пр(\%)}}\text{, або } A = \frac{у_{б}}{100}\ .$$

В 2003 р. 3 : 15,0 = 0,20; 20 : 100 = 0,20;
В 2004 р. 4 : 17,4 = 0,23; 23 : 100 = 0,23;
В 2005 р. 1 : 3,7 = 0,27; 27 : 100 = 0,27;
В 2006 р. 2 : 7,1 = 0,28; 28 : 100 = 0,28;
В 2007 р. 3 : 10,0 = 0,30; 30 : 100 = 0,30.

Обчислений показник має важливе значення в економічному аналізі, оскільки темпи росту можуть мати тенденцію до зменшення, або залишатись на одному рівні, а абсолютне значення одного процента приросту зростати.

Подамо всі вирахувані показники в таблиці.

Таблиця 11.8

Розрахунок показників динаміки валового збору картоплі в області за 2002 – 2007 рр.

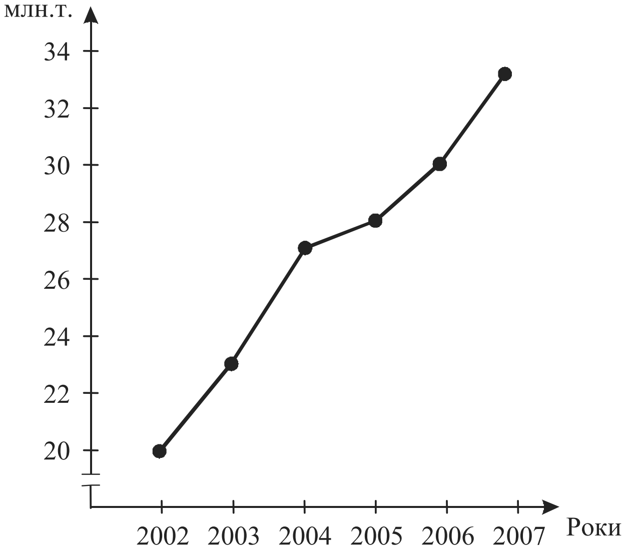
| Роки | Валовий збір картоплі, млн. т. | Абсолютний приріст, млн. т. | | Темп росту | | Темп приросту | | Абсолютне значення 1 % приросту, млн.. т. |
|---|---|---|---|---|---|---|---|---|
| | | базис-ний | ланцю-говий | базис-ний | ланцю-говий | базис-ний | ланцю-говий | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 2002 | 20 | - | - | 1,000 | - | - | - | - |

Продовження табл. 11.8

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|
| 2003 | 23 | 3 | 3 | 1,150 | 1,150 | 0,150 | 0,150 | 0,20 |
| 2004 | 27 | 7 | 4 | 1,350 | 1,174 | 0,350 | 0,174 | 0,23 |
| 2005 | 28 | 8 | 1 | 1,400 | 1,037 | 0,400 | 0,037 | 0,27 |
| 2006 | 30 | 10 | 2 | 1,500 | 1,074 | 0,500 | 0,071 | 0,28 |
| 2007 | 33 | 13 | 3 | 1,650 | 1,100 | 0,650 | 0,100 | 0,30 |

Ряди динаміки, як правило, подаються не тільки в таблицях, але й показуються на графіках. При графічному зображенні динамічного ряду на осі абсцис відкладається шкала часу, а на осі ординат – шкала рівнів ряду.

В якості ілюстрації покажемо побудову графіка за даними таблиці 11.7.




Мал. 11.1 Динаміка валового збору картоплі в області за 2002 – 2007 рр.

### 11.3. Середні показники динаміки

Динамічні ряди складаються з багатьох варіаційних рівнів, а тому, як люба статистична сукупність, вони потребують деяких узагальнюючих характеристик. Для цього вираховують середні показники: середні рівні ряду, середні абсолютні прирости, середні темпи росту і приросту.

При обчисленні середніх показників динаміки необхідно користуватись загальними положеннями теорії середніх.

Методи розрахунку середнього рівня інтервального і моментного рядів динаміки залежать від їх виду.

В інтервальному ряду з рівними інтервалами середній рівень ряду вираховується за формулою *середньої арифметичної простої*:

$$\bar{y} = \frac{\sum y}{n},$$

де $\bar{y}$ – середній рівень ряду;

$\sum y$ – сума рівнів ряду;

n – число рівнів.

В табл. 11.2 були наведені дані про валовий збір картоплі в області за 1999 – 2004 рр. за цими даними зробимо розрахунок середньорічного рівня валового збору картоплі за шість років:

$$\bar{y} = \frac{\sum y}{n} = \frac{20+23+27+28+30+32}{6} = \frac{161}{6} = 26{,}8 \text{ млн. т.}$$

Якщо окремі періоди інтервального ряду динаміки мають різну довжину, то для визначення середнього рівня використовують *середню арифметичну зважену*:

$$\bar{y} = \frac{\sum yt}{\sum t},$$

де y – рівні ряду;

t – проміжки часу.

Дану формулу використовують для знаходження середнього рівня в моментних рядах з нерівними інтервалами.

Нехай, наприклад, на 1 січня 2007 р. на підприємстві перебувало в списку працівників: 1210 чол.; 6 січня було прийнято 33 чол.; 15 січня звільнилось 7 чол.; 21 січня прийнято 12 чол.; 29 січня звільнилось 10 чол. В даному випадку моментний ряд буде мати вигляд:

Таблиця 11.9

Чисельність працівників на підприємстві в січні 2007 р.

| на 1.01 | на 6.01 | на 15.01 | на 21.01 | на 29.01 | на 1.02 |
|---|---|---|---|---|---|
| 1210 | 1243 | 1236 | 1248 | 1238 | 1238 |

В цьому ряду проміжки між окремими датами не рівні між собою: з 1 січня до 6 січня – 5 днів, з 6 до 15 січня – 9 днів, з 15 до 21 січня – 6 днів, з 21 до 29 січня – 8 днів, і з 29 по 31 січня включно – 3 дні.

Середньоспискова чисельність працівників підприємства в січні місяці буде дорівнювати:

$$\overline{y} = \frac{\sum yt}{\sum t} = \frac{1210 \cdot 5 + 1243 \cdot 9 + 1236 \cdot 6 + 1248 \cdot 8 + 1238 \cdot 3}{31} =$$

$$= \frac{38351}{31} = 1237 \text{ чол.}$$

Для наближеної оцінки інколи визначають півсуму рівнів на початок і кінець періоду і приймають її за характеристику середнього рівня всього періоду. Однак цей середній рівень є грубою оцінкою і застосовується рідко. За даними попереднього прикладу середньоспискове число працівників підприємства становитиме:

$$\overline{y} = \frac{y_1 + y_n}{2} = \frac{1210 + 1238}{2} = 1224 \text{ чол.},$$

де $y_n$ – кінцевий рівень ряду динаміки.

Як бачимо, отримана середня – неточна, тому, що в ній не враховані проміжні значення ряду динаміки.

В економічних дослідженнях для аналізу динаміки середніх рівнів часто практикують укрупнення інтервалу ряду. Візьмемо інтервальний ряд динаміки щоденного роздрібного товарообороту по одному із торгів області.

Таблиця 11.10

Роздрібний товарооборот торгу області в червні.

| Число місяця | Товарооборот, млн. грн. | Число місяця | Товарооборот, млн. грн. | Число місяця | Товарооборот, млн. грн. |
|---|---|---|---|---|---|
| 1 | 201 | 11 | 208 | 21 | 236 |
| 2 | 202 | 12 | 219 | 22 | 235 |
| 3 | 204 | 13 | 220 | 23 | 238 |
| 4 | 191 | 14 | 223 | 24 | 238 |
| 5 | 196 | 15 | 197 | 25 | 239 |
| 6 | 210 | 16 | 190 | 26 | 245 |
| 7 | 205 | 17 | 228 | 27 | 242 |
| 8 | 213 | 18 | 230 | 28 | 247 |
| 9 | 215 | 19 | 234 | 29 | 250 |
| 10 | 211 | 20 | 234 | 30 | 251 |

Для переходу від ряду динаміки щоденного товарообороту до ряду даних по п’ятиденках складемо роздрібний товарооборот за кожну п’ятиденку. Після цього перетворимо інтервальний ряд по п’ятиденках в ряд середньоденних рівнів роздрібного товарообороту шляхом ділення загального обсягу товарообороту за кожні п’ять днів на величину інтервалу (5 днів). Середньоденний товарооборот за місяць в цілому отримаємо шляхом ділення загального обсягу роздрібного товарообороту за місяць на 30 днів. Його можна також визначити шляхом додавання середньоденних даних по п'ятиденках і діленням отриманої суми на шість.

Наші розрахунки подамо в табл.. 11.11.

Таблиця 11.11

| Показники | П'ятиденки | | | | | | Разом за місяць |
|---|---|---|---|---|---|---|---|
| | I | II | III | IV | V | VI | |
| Роздрібний товарооборот, млн. грн. | 994 | 1054 | 1067 | 1116 | 1186 | 1235 | 6652 |
| Середньоденний роздрібний товарооборот, млн. грн. | 198,8 | 210,8 | 213,4 | 223,2 | 237,2 | 247,0 | 221,7 |

Укрупнення інтервалів в інтервальних рядах динаміки використовують для виявлення тенденції розвитку явища. В нашому прикладі простежується послідовне зростання роздрібного товарообороту як за укрупненими інтервалами, так і за середніми рівнями по п'ятиденках.

Якщо інтервальний ряд динаміки має різні інтервали, тоді потрібно застосовувати середню арифметичну зважену. Наприклад. Роздрібний товарооборот торгу за перше півріччя характеризується наступними даними (табл.. 11.12):

Таблиця 11.12

| | Січень | Лютий | Березень | II квартал |
|---|---|---|---|---|
| Роздрібний товарооборот, млн. грн.. | 5200 | 5160 | 5340 | 18000 |

За даними таблиці видно, що ми маємо три рівні ряду з інтервалом в один місяць, а четвертий рівень – квартал. Для визначення середньомісячного товарообороту за перше півріччя, потрібно взяти об'єм товарообороту за півроку і поділити цю суму на шість місяців (квартальний товарооборот береться з трійними вагами).

$$\bar{y}_{м} = \frac{(5200+5160+5340)+18000}{3+3} = \frac{33700}{6} = 5616{,}7 \text{ млн. грн.}$$

Якщо за вищенаведеними даними потрібно вирахувати середньоквартальний товарооборот, тоді переходять до рівноінтервального ряду шляхом укрупнення інтервалів.

$$\overline{y}_{к} = \frac{y_{Iк} + y_{IIк}}{2} = \frac{15700 + 18000}{2} = \frac{33700}{2} = 16850 \text{ млн. грн.}$$

Для визначення середнього рівня в моментному динамічному ряду з рівними або приблизно рівними проміжками часу між сусідніми датами, застосовують формулу *середньої хронологічної*, яка має вигляд:

$$\overline{y} = \frac{\frac{1}{2}y_1 + y_2 + y_3 + \ldots + y_{n-1} + \frac{1}{2}y_n}{n-1}.$$

В знаменнику береться число рівнів за мінусом одиниці тому, що в чисельнику серед доданків перший і останній беруться в половинному розмірі.

Розглянемо для прикладу моментний ряд з рівними інтервалами, який відображає зміну тракторного парку в сільських спілках району (табл..11.1).

Визначимо середнє число тракторів за кожний квартал, перше і друге півріччя і за рік в цілому.

$$\overline{y}_{Iк} = \frac{\frac{622}{2} + 640 + 643 + \frac{640}{2}}{3} = \frac{1914}{3} = 638 \text{ шт.};$$

$$\overline{y}_{IIк} = \frac{\frac{640}{2} + 664 + 670 + \frac{682}{2}}{3} = \frac{1995}{3} = 665 \text{ шт.};$$

$$\overline{y}_{IIIк} = \frac{\frac{682}{2} + 733 + 753 + \frac{768}{2}}{3} = \frac{2211}{3} = 737 \text{ шт.};$$

$$\overline{y}_{IVк} = \frac{\frac{768}{2} + 800 + 826 + \frac{888}{2}}{3} = \frac{2454}{3} = 818 \text{ шт.}$$

За кожне півріччя середні рівні можна шукати за формулою середньої хронологічної, або за середньою арифметичною простою з середньоквартальних рівнів.

$$\overline{y}_{Iп} = \frac{\frac{622}{2} + 640 + 643 + 640 + 664 + 670 + \frac{682}{2}}{6} = 651{,}5 \text{ шт.};$$

або $\overline{y}_{Iп} = \frac{\overline{y}_{Iк} + \overline{y}_{IIк}}{2} = \frac{638 + 665}{2} = 651{,}5$ шт.;

$$\bar{y}_{IIп} = \frac{\frac{682}{2} + 733 + 753 + 768 + 800 + 826 + \frac{888}{2}}{13-1} = 777{,}5 \text{ шт.};$$

$$\text{або } \bar{y}_{IIп} = \frac{\bar{y}_{IIIк} + \bar{y}_{IVк}}{2} = \frac{737 + 818}{2} = 777{,}5 \text{ шт.};$$

Як бачимо, відповідь, не залежно від способу визначення середніх рівнів, в обох випадках однакові.

В цілому за рік середньомісячне число тракторів в селянських спілках можна визначати за формулами середньої хронологічної або середньої арифметичної із середніх квартальних чи піврічних рівнів:

$$\bar{y}_{р} = \frac{\frac{622}{2} + 640 + 643 + 640 + 664 + 670 + 682 + 733 + 753 + 768 + 800 + 826 + \frac{888}{2}}{13-1} = \frac{8574}{12} = 714{,}5 \text{ шт.}$$

$$\text{або } \bar{y}_{р} = \frac{638 + 665 + 737 + 818}{4} = \frac{2858}{4} = 714{,}5 \text{ шт.};$$

$$\text{або } \bar{y}_{р} = \frac{\bar{y}_{Iп} + \bar{y}_{IIп}}{2} = \frac{651{,}5 + 777{,}5}{2} = 714{,}5 \text{ шт.}$$

Дещо по-іншому вираховують середні рівні в моментному ряду з нерівними інтервалами. Припустимо, що ми маємо такі дані про спискову чисельність працівників промислового підприємства за перше півріччя 2007 р. На 1.01 на підприємстві перебувало в списку 2000 чол., на 1.02 – 1950 чол., на 1.03 – 2010 чол., на 1.04 – 2020 чол. і на 1.07 2040 чол. Потрібно вирахувати середньоспискове число працівників за перший і другий квартал і за перше півріччя.

Знаходимо середньоспискове число працівників за перший квартал.

Хід розв’язку можна подати так:

$$\bar{y}_{Iк} = \frac{\frac{2000 + 1950}{2} + \frac{1950 + 2021}{2} + \frac{2021 + 2020}{2}}{3} =$$

$$= \frac{1975 + 1980 + 2015}{3} = \frac{5970}{3} = 1990 \text{ чол.},$$

або перетворивши чисельник, отримаємо:

$$\bar{y}_{Iк} = \frac{\frac{2000}{2} + \frac{1950+1950}{2} + \frac{2010+2010}{2} + \frac{2020}{2}}{4-1} =$$

$$= \frac{\frac{2000}{2} + 1950 + 2010 + \frac{2020}{2}}{4-1} = \frac{5970}{3} = 1990 \text{ чол.},$$

Тобто, середньоспискову чисельність працівників підприємства отримуємо за формулою середньої хронологічної.

Вираховуємо середньоспискове число працівників за другий квартал за даними на 1 квітня і 1 липня:

$$\bar{y}_{IIк} = \frac{2020+2040}{2} = 2030 \text{ чол.}$$

Знаходимо середньоспискове число працівників за перше півріччя:

$$\bar{y}_{Iп} = \frac{\bar{y}_{Iк} + \bar{y}_{IIк}}{2} = \frac{1990+2030}{2} = \frac{4020}{2} = 2010 \text{ чол.},$$

або $\bar{y}_{Iп} = \frac{1975+1980+2015+(2030 \cdot 3)}{6} = \frac{12060}{6} = 2010$ чол.

В даному випадку використовуємо формулу середньої арифметичної зваженої, взявши середньоспискове число працівників за другий квартал з потрійними вагами.

*Середній абсолютний приріст* визначається як середня арифметична проста з ланцюгових абсолютних приростів за певні періоди і показує на скільки одиниць в середньому змінився рівень у порівнянні з попереднім.

$$\bar{\Delta}_y = \frac{\sum \Delta_y^л}{n},$$

де $\bar{\Delta}_y$ – середній абсолютний приріст;

n – кількість приростів.

Оскільки сума ланцюгових приростів ($\sum \Delta_y^л$) дорівнює приросту за весь досліджуваний період ($y_n$ – $y_1$), то формула середнього абсолютного приросту може мати такий вигляд:

$$\bar{\Delta}_y = \frac{y_n - y_1}{n-1}.$$

Число абсолютних приростів менше числа рівнів динамічного ряду на одиницю. За таблицею 11.8 середньорічний абсолютний приріст валового збору картоплі складе:

$$\overline{\Delta}_у = \frac{\sum \Delta_у^л}{n} = \frac{3+4+1+2+3}{5} = \frac{13}{5} = 2{,}6 \text{ млн. т.}$$

$$\text{або } \overline{\Delta}_у = \frac{y_n - y_1}{n-1} = \frac{33-20}{6-1} = \frac{13}{5} = 2{,}6 \text{ млн. т.}$$

*Середній темп росту* вираховується за формулою *середньої геометричної*:

$$\overline{T}_р = \sqrt[n]{T_{р1}^л \cdot T_{р2}^л \cdot T_{р3}^л \cdot \ldots \cdot T_{рn}^л} \ ,$$

де $\overline{T}_р$ – середній темп росту;

$T_{р1}^л, T_{р2}^л, T_{р3}^л, \ldots, T_{рn}^л$ – ланцюгові темпи росту;

n – число темпів.

Для обчислення середнього темпу росту використовують також іншу формулу:

$$\overline{T}_р = \sqrt[n-1]{\frac{y_n}{y_1}} \ ,$$

де $y_n$ – кінцевий рівень ряду;

$y_1$ – початковий рівень ряду;

n – кількість рівнів динамічного ряду.

Обчислення середніх темпів росту проводять у випадках, коли відомі початковий і кінцевий рівні ряду, або коли відомі темпи росту чи приросту за весь період.

Підставивши значення ланцюгових темпів росту, обчислених в таблиці 11.8 у формулу середньої геометричної, знайдемо середньорічний темп росту за 2002 – 2007 рр.:

$$\overline{T}_р = \sqrt[5]{1{,}150 \cdot 1{,}174 \cdot 1{,}037 \cdot 1{,}071 \cdot 1{,}100} = \sqrt[5]{1{,}650} = 1{,}105 \ .$$

При обчисленні середньої геометричної з коренями високих ступенів використовують логарифми, тому що середня геометрична в загальному вигляді може бути визначена як антилогарифм від середньої арифметичної із логарифмів чисел. В нашому прикладі, логарифмуючи ліву і праву частини, отримаємо:

$$\lg \overline{T} = \frac{1}{5}\left(\lg 1{,}15 + \lg 1{,}174 + \lg 1{,}037 + \lg 1{,}071 + \lg 1{,}100\right) =$$

$$= \frac{1}{5}\left(0{,}0607 + 0{,}0697 + 0{,}0158 + 0{,}0298 + 0{,}0414\right) = \frac{0{,}2174}{5} = 0{,}04348$$

Антилогарифм для 0,0435 дорівнює 1,105. Отже, середній річний темп росту валового збору картоплі в області за 2002 – 2007 рр. склав 1,105 або 110,5 %.

Користуючись другою формулою, обчислимо середньорічний темп росту валового збору картоплі в області:

$$\overline{T}_p = \sqrt[n-1]{\frac{y_n}{y_1}} = \sqrt[5]{\frac{33}{20}} = \sqrt[5]{1{,}650} = 1{,}105\,;$$

$$\lg \overline{T}_p = \frac{1}{5}\lg 1{,}65 = 0{,}0435\,;\ \overline{T}_p = 1{,}105 \text{ або } 110{,}5\ \%.$$

З метою перевірки точності обчислення середньорічного темпу росту перемножимо початковий рівень ряду 1992 р. ($y_1$ = 20 млн. т.) послідовно 5 раз на 1,105, в результаті чого ми повинні отримати рівень валового збору картоплі в області за 1997 р. ($y_n$ = 33 млн. т.).

$$y_n = y_1 \cdot \overline{T}_p = 20 \cdot 1{,}105 \cdot 1{,}105 \cdot 1{,}105 \cdot 1{,}105 \cdot 1{,}105 = 20 \cdot 1{,}105^5 =$$

$$= 32{,}9489 \approx 33 \text{ млн.т.}$$

Обчислюючи середні темпи росту за формулою середньої геометричної доводиться вдаватись до таблиць логарифмів. На практиці дослідники користуються готовими таблицями А.М. Айрапетова (видавництво “Статистика”, М., 1979 р.), за допомогою яких легко знаходять середньорічні темпи росту, приросту або зниження.

Середні темпи росту обчислюють крім середньої геометричної простої, ще й за формулою *середньої геометричної зваженої*, яка має вигляд:

$$\overline{T}_p = \sqrt[\sum t]{T_{p1}^{t_1} \cdot T_{p2}^{t_2} \cdot T_{p3}^{t_3} \cdot \ldots \cdot T_{pn}^{t_n}}\,,$$

де t – інтервал часу, на протязі якого зберігається даний темп росту;

$\sum t$ – сума відрізків часу періоду.

Розглянемо приклад. Середній річний темп росту продуктивності праці робітників заводу за перших два роки п’ятиріччя склав 1,04, а за наступні три роки – 1,05. Визначити середньомісячний темп росту продуктивності праці робітників заводу за 5 років:

$$\overline{T}_p = \sqrt[5]{1{,}04^2 \cdot 1{,}05^3}\,;$$

$$\lg \overline{T}_p = \frac{1}{5}\left(2\lg 1{,}04 + 3\lg 1{,}05\right) = \frac{1}{5}\left(2 \cdot 0{,}0170 + 3 \cdot 0{,}0212\right) =$$

$$= \frac{1}{5}\left(0{,}0340 + 0{,}0636\right) = \frac{1}{5} \cdot 0{,}0976 = 0{,}0195;$$

$$\lg \overline{T}_p = 0{,}0195\,;\ \overline{T}_p = 1{,}046\,, \text{ або } 104{,}6\ \%.$$

При використанні логарифмів для визначення середніх темпів росту потрібно пам'ятати, що логарифми величин менших за одиницю, але позитивних, мають від'ємну характеристику.

Потрібно також мати на увазі, що середній темп можна вираховувати тільки в тих випадках, коли темпи динамічного ряду безперервно зростають або спадають. Коли темпи коливаються в різних напрямках, вирахуваний за ними середній темп є величиною дуже умовною.

*Середній темп приросту* визначається як різниця між середнім темпом росту і одиницею (якщо середній темп росту у вигляді коефіцієнта), або 100 (якщо він у відсотках):

$$\overline{T}_{пр} = \overline{T}_{р} - 1 \text{ (у вигляді коефіцієнтів);}$$

$$\overline{T}_{пр} = \overline{T}_{р} - 100 \text{ (у вигляді відсотків).}$$

Середній темп приросту показує, на скільки відсотків збільшився, або зменшився рівень у порівнянні з попереднім в середньому за одиницю часу.

Вище обчислено, що середньорічний темп росту валового збору картоплі в області склав 1,105, або 110,5 %. Значить середньорічний темп приросту валового збору картоплі складе у вигляді коефіцієнтів:

$$\overline{T}_{пр} = 1{,}105 - 1 = 0{,}105$$

у вигляді відсотків:

$$\overline{T}_{пр} = 110{,}5 - 100 = 10{,}5 \text{ \%.}$$

Застосування наведених показників динаміки є лише першим етапом аналізу динамічних рядів, який дозволяє виявити швидкість і інтенсивність розвитку явищ.

Подальший аналіз рядів динаміки соціально-економічних показників зв'язаний з більш складними узагальненнями, з визначенням основної тенденції, вивченням сезонних коливань рівнів і дослідженням зв'язку між рядами.

## 11.4 Виявлення тенденцій розвитку явищ

Виявлення основної тенденції (тренду) ряду, є одним з головних методів аналізу і узагальнення динамічних рядів. Зображена на графіку лінія тренду динамічного ряду покаже плавну зміну досліджуваного явища в часі, яке звільнене від короткочасних відхилень, викликаних різними причинами. В статистичній практиці виявлення основної тенденції розвитку явищ в часі проводиться методами укрупнення інтервалів, рухомої середньої і аналітичним вирівнюванням.

Одним з найпростіших способів обробки ряду з метою виявлення закономірності зміни його рівнів є *укрупнення інтервалів* (періодів) часу. Суть цього методу полягає в тому, що дані динамічного ряду об'єднуються в групи по періодах і розраховується середній показник на період – трирічча, п'ятиріччя і т.д. Такій обробці доцільно піддавати динамічний ряд з більш-менш систематичними коливаннями рівня, що дозволяє чіткіше виявити загальну тенденцію розвитку явища.

Укрупнення інтервалів проілюструємо за даними наступного прикладу (табл. 11.13).

Таблиця 11.13

Динаміка врожайності озимої пшениці в селянській спілці "Колос".

| Роки | Урожайність озимої пшениці, ц/га | Сумарна врожайність, ц/га (за трирічча) | Середня врожайність, ц/га (за трирічча) |
|---|---|---|---|
| 1993 | 15,6 | | |
| 1994 | 16,0 | 50,5 | 16,8 |
| 1995 | 18,9 | | |
| 1996 | 15,7 | | |
| 1997 | 20,0 | 55,3 | 18,4 |
| 1998 | 19,6 | | |
| 1999 | 19,8 | | |
| 2000 | 21,5 | 61,3 | 20,4 |
| 2001 | 20,0 | | |
| 2002 | 27,3 | | |
| 2003 | 24,4 | 79,9 | 26,6 |
| 2004 | 28,2 | | |
| 2005 | 27,9 | | |
| 2006 | 33,1 | 93,7 | 31,2 |
| 2007 | 32,7 | | |

Якщо розглядати рівні урожайності озимої пшениці за короткі проміжки часу, то в силу впливу багатьох факторів, головним чином, погодних умов різних років, які діють в різних напрямках, в рядах динаміки спостерігається зниження або підвищення їх рівнів. Через це не видно основної тенденції розвитку досліджуваного явища.

В результаті проведеного укрупнення інтервалів, взявши дані за трирічча, ми отримали новий ряд динаміки сумарної урожайності за три роки, який показує її збільшення.

Середня річна врожайність за трирічча також показує тенденцію її росту.

Важливим способом виявлення загальної тенденції ряду динаміки є *згладжування за допомогою рухомої середньої*. Тут також вдаються до укрупнення періодів, але воно проводиться шляхом послідовних зміщень на одну дату при збереженні постійного інтервалу періоду.

Порядок розрахунку рухомої середньої за даними попереднього прикладу покажемо в таблиці 11.14.

Таблиця 11.14

Розрахунок трирічної рухомої середньої врожайності озимої пшениці.

| Роки | Урожайність озимої пшениці, ц/га | Сумарна врожайність, ц/га (за трирічча) | Середня врожайність, ц/га (за трирічча) |
|---|---|---|---|
| 1993 | 15,6 | - | - |
| 1994 | 16,0 | 50,5 | 16,8 |
| 1995 | 18,9 | 50,6 | 16,9 |
| 1996 | 15,7 | 54,6 | 18,2 |
| 1997 | 20,0 | 55,3 | 18,4 |
| 1998 | 19,6 | 59,4 | 19,8 |
| 1999 | 19,8 | 60,9 | 20,3 |
| 2000 | 21,5 | 61,3 | 20,4 |
| 2001 | 20,0 | 68,8 | 22,9 |
| 2002 | 27,3 | 71,7 | 23,9 |
| 2003 | 24,4 | 79,9 | 26,6 |
| 2004 | 28,2 | 80,5 | 26,8 |
| 2005 | 27,9 | 89,2 | 29,7 |
| 2006 | 33,1 | 93,7 | 31,2 |
| 2007 | 32,7 | - | - |

Як видно із таблиці 11.14, згладжений ряд, який складається з рухомих середніх показує більш плавне підвищення урожайності озимої пшениці.

Згладжування способом рухомої середньої можна проводити також і за парним числом членів ряду. Однак таке згладжування трохи складніше, так як середня мусить бути віднесена тільки до середини між двома датами, котрі знаходяться в середині інтервалу. Щоб ліквідувати такий зсув застосовують спосіб перетворення рівнів, або центрування. При *перетворенні рівнів* рівень першого інтервалу ділять на два і його половина входить в суму, за якою вираховується рухома середня. Потім беруться всі наступні рівні в повному розмірі і до них додається половина рівня, який виходить за межі згладжування.

*Центрування* заключається в тому, що із кожної пари згладжених рухомих середніх розраховується середня арифметична, яка і відноситься до відповідної дати.

При застосуванні методу рухомої середньої важливе значення має вибір періоду або інтервалу згладжування, який повинен відповідати періоду коливань в даному динамічному ряду. Якщо, наприклад, періодичність коливань встановлення в 3 місяці, то береться 3-членна рухома середня, в 6 місяців – 6-членна рухома середня і т.д.

Найбільш ефективним і складним способом виявлення основної тенденції є *аналітичне вирівнювання*. При цьому рівні ряду динаміки розглядаються як функція часу $\left[\hat{y}_t = f(t)\right]$, а завдання вирівнювання зводиться до знаходження того виду функції, ординати точок якої були б найбільш близькі до значень фактичного динамічного ряду.

На практиці найбільш поширеними формулами, які виражають тенденцію розвитку (тренд) явищ є: пряма, показникова функція, парабола другого і третього порядків, гіпербола, логістична функція, експонента, ряд Фур'є та деякі інші.

*Вирівнювання за прямою* використовують в тих випадках, коли абсолютні прирости більш-менш постійні, тобто коли рівні динамічного ряду змінюються в арифметичній прогресії, або близькі до неї.

Рівняння прямої має вигляд:

$$\hat{y}_t = a_0 + a_1 t\,,$$

де $\hat{y}_t$ – вирівняні значення динамічного ряду;

$a_0$ і $a_1$ – параметри шуканої прямої (початковий рівень і щорічний приріст);

t – умовне позначення часу.

Для знаходження параметрів $a_0$ і $a_1$, потрібно розв'язати за способом найменших квадратів систему нормальних рівнянь:

$$\begin{cases} \sum y = na_0 + a_1 \sum t; \\ \sum yt = a_0 \sum t + a_1 \sum t^2, \end{cases}$$

де y – фактичні рівні динамічного ряду;

n – число членів ряду динаміки.

Дану систему нормальних рівнянь можна легко спростити, якщо відлік часу брати з середини ряду таким чином, щоб сума часу дорівнювала нулю ($\sum t = 0$). При непарному числі рівнів серединна точка приймається за 0, тоді попередні періоди позначаються відповідно через –1, –2, –3 і т.д., а наступні за серединним періоди –

відповідно через +1, +2, +3 і т.д. При парному числі рівнів динамічного ряду два серединних моменти часу позначаються через –1 і +1, а всі решта через два інтервали, тобто попередні періоди до середини через –3, –5, –7 і т.д., а наступні – відповідно через +3, +5, +7 і т.д.

При відліку часу від середини ряду коли $\sum t = 0$, тоді система рівнянь для знаходження параметрів $a_0$ і $a_1$ матиме вигляд:

$$\begin{cases} \sum y = na_0; \\ \sum yt = a_1 \sum t^2, \end{cases}$$

Звідки

$$a_0 = \frac{\sum y}{n}, \ a_1 = \frac{\sum yt}{\sum t^2}.$$

Методику вирівнювання врожайності озимої пшениці за рівнянням прямої покажемо на прикладі (табл. 11.15).

Таблиця 11.15

Розрахункова таблиця для вирівнювання ряду динаміки за прямою.

| Роки | Урожайність озимої пшениці, ц/га y | Умовне позначення часу t | $t^2$ | yt | Вирівняна урожайність $\hat{y}_t = a_0 + a_1 t$ |
|---|---|---|---|---|---|
| 1993 | 15,6 | -7 | 49 | -109,2 | 14,1 |
| 1994 | 16,0 | -6 | 36 | -96,0 | 15,3 |
| 1995 | 18,9 | -5 | 25 | -94,5 | 16,5 |
| 1996 | 15,7 | -4 | 16 | -62,8 | 17,8 |
| 1997 | 20,0 | -3 | 9 | -60,0 | 19,0 |
| 1998 | 19,6 | -2 | 4 | -39,2 | 20,2 |
| 1999 | 19,8 | -1 | 1 | -19,8 | 21,5 |
| 2000 | 21,5 | 0 | 0 | 0 | 22,7 |
| 2001 | 20,0 | 1 | 1 | 20,0 | 23,9 |
| 2002 | 27,3 | 2 | 4 | 54,6 | 25,2 |
| 2003 | 24,4 | 3 | 9 | 73,2 | 26,4 |
| 2004 | 28,2 | 4 | 16 | 112,8 | 27,7 |
| 2005 | 27,9 | 5 | 25 | 139,5 | 28,9 |
| 2006 | 33,1 | 6 | 36 | 198,6 | 30,1 |
| 2007 | 32,7 | 7 | 49 | 228,9 | 31,4 |
| n = 15 | $\sum y = 340{,}7$ | $\sum t = 0$ | $\sum t^2 = 280$ | $\sum yt = 346{,}1$ | $\sum \hat{y}_t = 340{,}7$ |

Використовуючи розрахункові підсумки, отримуємо:

$$a_0 = \frac{\sum y}{n} = \frac{340,7}{15} = 22,713;$$

$$a_1 = \frac{\sum yt}{\sum t^2} = \frac{346,1}{280} = 1,236.$$

Звідси рівняння прямої буде мати наступний вигляд:

$$\hat{y}_t = 22,713 + 1,236\,t.$$

Коефіцієнт регресії в даному рівнянні ($a_1$ = 1,236) характеризує середній приріст урожайності озимої пшениці за рік. Величина 22,713 буде показувати теоретичну врожайність 2000 р., для якого ми взяли 0 номер року. Підставляючи у рівняння $\hat{y}_t = 22,713 + 1,236\,t$ послідовно значення t (–7, –6, –5 і т.д.), отримаємо вирівняний (теоретичний) ряд динаміки урожайності озимої пшениці в селянській спілці "Колос", який абстрагований від випадкових коливань і характеризується систематичним ростом (див. останню графу табл. 11.15).

Якщо розрахунки виконані правильно, то $\sum y = \sum \hat{y}_t$. В нашому прикладі $\sum y = 340,7 = \sum \hat{y}_t$. Отже, значення рівнів вирівняного динамічного ряду знайдені правильно.

Результати проведеного аналітичного вирівнювання ряду динаміки урожайності озимої пшениці за 1993 – 2007 рр. і фактичні дані покажемо на графіку (мал. 11.2).

До основних кривих, які використовуються для аналітичного вирівнювання динамічних рядів відносять: гіперболу, параболу другого і третього порядків, показникову функцію, та деякі інші.

*Вирівнювання за гіперболою* проводиться тоді, коли із плином часу ряд динаміки зростає або спадає до певної межі.

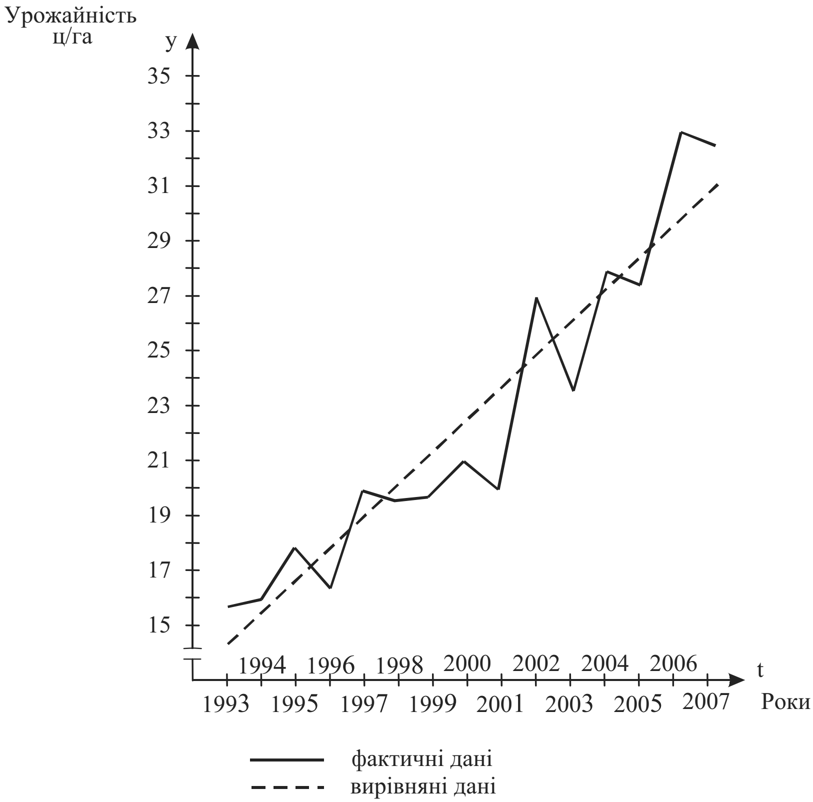
Рівняння гіперболи визначається за формулою:

$$\hat{y}_t = a_0 + a_1 \frac{1}{t}.$$

Для знаходження параметрів $a_0$ і $a_1$ в даному рівнянні способом найменших квадратів застосовують систему нормальних рівнянь:

$$\begin{cases} \sum y = na_0 + a_1 \sum \frac{1}{t}; \\ \sum y\frac{1}{t} = a_0 \sum \frac{1}{t} + a_1 \sum \frac{1}{t^2}. \end{cases}$$

Якщо добитись, щоб $\sum t = 0$, тоді параметри $a_0$ і $a_1$ знаходять за новою системою рівнянь:




Мал. 11.2 Урожайність озимої пшениці в селянській спілці “Колос”.

$$\begin{cases} \sum y = na_0; \\ \sum y\frac{1}{t} = a_1 \sum \frac{1}{t^2}. \end{cases}$$

Перетворивши цю систему рівнянь в логарифмічну, будемо мати:

$$\begin{cases} \sum \lg y = n \lg a_0; \\ \sum t \lg y = \lg a_1 \sum t^2, \end{cases}$$

Звідки

$$\lg a_0 = \frac{\sum \lg y}{n} \text{ і } \lg a_1 = \frac{\sum t \lg y}{\sum t^2}.$$

Розглянемо вирівнювання за гіперболою на умовному прикладі, який характеризує динаміку зниження собівартості одиниці продукції за 2003-2007 рр. (табл. 11.16).

Таблиця 11.16

Динаміка зниження собівартості продукції за 2003-2007 рр. і розрахункові дані для вирівнювання за гіперболою.

| Роки | Собівартість одиниці продукції, грн. y | lg y | t | $t^2$ | t lg y | $\lg \hat{y}_t$ | Вирівняні рівні $\hat{y}_t = a_0 + a_1 \frac{1}{t}$ |
|---|---|---|---|---|---|---|---|
| 2003 | 70 | 1,8451 | -2 | 4 | -3,6902 | 1,8484 | 70,53 |
| 2004 | 50 | 1,6990 | -1 | 1 | -1,6990 | 1,6749 | 47,31 |
| 2005 | 30 | 1,4914 | 0 | 0 | 0 | 1,5013 | 31,72 |
| 2006 | 20 | 1,3010 | 1 | 1 | 1,3010 | 1,3278 | 21,27 |
| 2007 | 15 | 1,1761 | 2 | 4 | 2,3522 | 1,1542 | 14,27 |
| n = 5 | $\sum y = 185$ | $\sum \lg y = 7{,}5066$ | $\sum t = 0$ | $\sum t^2 = 10$ | $\sum t \lg y = -1{,}7356$ | $\sum \lg \hat{t}_t = 7{,}5066$ | $\sum \hat{y}_t = 185{,}10$ |

Прийнявши в якості вирівнювання рівняння гіперболи $\hat{y}_t = a_0 + a_1 \frac{1}{t}$ і прологарифмувавши його $\lg \hat{y}_t = \lg a_0 + t \lg a_1$ параметри $a_0$ і $a_1$ знайдемо за наведеними вище формулами:

$$\lg a_0 = \frac{\sum \lg y}{n} = \frac{7{,}5066}{5} = 1{,}50132 \text{, звідси } a_0 = 31{,}72;$$

$$\lg a_1 = \frac{\sum t \lg y}{\sum t^2} = \frac{-1{,}7356}{10} = -0{,}17356 \text{, звідси } a_1 = -1{,}491.$$

Отже, $\lg \hat{y}_t = 1{,}50132 - 0{,}17356\, t$, або $\hat{y}_t = 31{,}72 - 1{,}491 \frac{1}{t}$.

Для розрахунку вирівняних рівнів зручніше користуватися рівнянням логарифмів.

Підставляючи в формулу значення t = –2, –1, 0, +1, +2, знайдемо логарифми ($\lg \hat{y}_t$), а потім за таблицями антилогарифмів – $\hat{y}_t$.

Так, для 2003 р.:

$$\lg \hat{y}_t = 1{,}50132 - 0{,}17356\,(-2) = 1{,}8484 \text{, звідси } \hat{y}_t = 70{,}53.$$

Для 2004 р.:

$$\lg \hat{y}_t = 1{,}50132 - 0{,}17356\,(-1) = 1{,}6749 \text{, звідси } \hat{y}_t = 47{,}31.$$

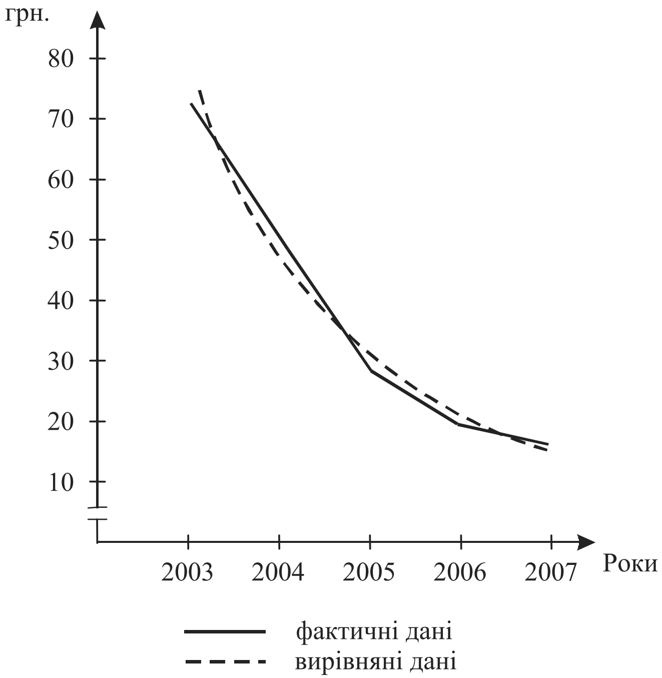
Для 2005 р.:

$\lg \hat{y}_t = 1{,}50132 - 0{,}17356\,(0) = 1{,}5013$, звідси $\hat{y}_t = 31{,}72$ і т.д.

Логарифми вирівняних рівнів і самі рівні показані в двох останніх колонках табл. 11.16.

Як видно із таблиці, вирівняні рівні дуже близькі до емпіричних, що в свою чергу свідчить про відповідність рівняння гіперболи для відображення тренду.

Результати проведеного вирівнювання ряду динаміки зниження собівартості продукції за 2002-2007 рр. покажемо на графіку (мал. 11.3).




Мал. 11.3. Динаміка зниження собівартості продукції за 2003-2007 рр.

На графіку чітко видно вирівнювання емпіричної лінії гіперболою. Вона зображає обернений зв'язок, де з плином часу собівартість зменшується.

При *вирівнюванні за параболою другого порядку* $\hat{y}_t = a_0 + a_1 t + a_2 t^2$ параметри $a_0$, $a_1$ і $a_2$ визначаються способом найменших квадратів, для чого складають і розв'язують систему нормальних рівнянь:

$$\begin{cases} \sum y = na_0 + a_1 \sum t + a_2 \sum t^2; \\ \sum yt = a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3; \\ \sum yt^2 = a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4. \end{cases}$$

Якщо добитись щоб $\sum t = 0$, тоді і $\sum t^3 = 0$, а, отже, система рівнянь спроститься:

$$\begin{cases} \sum y = na_0 + a_2 \sum t^2; \\ \sum yt = a_1 \sum t^2; \\ \sum yt^2 = a_0 \sum t^2 + a_2 \sum t^4. \end{cases}$$

Із цієї системи $a_1 = \frac{\sum yt}{\sum t^2}$, а параметри $a_0$ і $a_2$ визначаються розв'язком системи двох рівнянь з двома невідомими.

Покажемо на прикладі вирівнювання динамічного ряду рівнянням параболи другого порядку. Для отримання необхідних даних побудуємо розрахункову таблицю 11.17.

Таблиця 11.17

Виробництво взуття на одній із фабрик області за 1998-2007 рр.

| Роки | Виробництво взуття, тис. пар y | t | $t^2$ | $t^4$ | yt | $yt^2$ | Вирівняні рівні $\hat{y}_t = a_0 + a_1 t + a_2 t^2$ |
|---|---|---|---|---|---|---|---|
| 1998 | 12 | -9 | 81 | 6561 | -108 | 972 | 19,6 |
| 1999 | 50 | -7 | 49 | 2401 | -350 | 2450 | 43,9 |
| 2000 | 70 | -5 | 25 | 625 | -350 | 1750 | 63,9 |
| 2001 | 80 | -3 | 9 | 81 | -240 | 720 | 79,7 |
| 2002 | 92 | -1 | 1 | 1 | -92 | 92 | 91,2 |
| 2003 | 95 | 1 | 1 | 1 | 95 | 95 | 98,6 |
| 2004 | 97 | 3 | 9 | 81 | 291 | 873 | 101,7 |
| 2005 | 99 | 5 | 25 | 625 | 495 | 2475 | 100,6 |
| 2006 | 98 | 7 | 49 | 2401 | 686 | 4802 | 95,2 |
| 2007 | 87 | 9 | 81 | 6561 | 783 | 7047 | 85,6 |
| n=10 | $\sum y = 780$ | $\sum t = 0$ | $\sum t^2 = 330$ | $\sum t^4 = 19338$ | $\sum yt = 1210$ | $\sum yt^2 = 21276$ | $\sum \hat{y}_t = 780$ |

В нашому прикладі $a_1 = \frac{\sum yt}{\sum t^2} = \frac{1210}{330} = 3{,}667$.

Система рівнянь для знаходження параметрів $a_2$ і $a_0$:

$$\begin{cases} \sum y = na_0 + a_2 \sum t^2; \\ \sum yt^2 = a_0 \sum t^2 + a_2 \sum t^4. \end{cases}$$

Підставляємо дані із таблиці в рівняння цієї системи:

$$\begin{cases} 780 = 10\, a_0 + 330\, a_2; \\ 21276 = 330\, a_0 + 19338\, a_2. \end{cases}$$

Поділивши члени першого рівняння на 10, а члени другого рівняння на 330, отримаємо систему рівнянь:

$$\begin{cases} 78{,}000 = a_0 + 33\, a_2; \\ 64{,}473 = a_0 + 58{,}6\, a_2. \end{cases}$$

Віднімаємо від другого рівняння перше:

$$-13{,}527 = 25{,}6a_2, \text{ звідси } a_2 = \frac{-13{,}527}{25{,}6} = -0{,}528;$$

$$78 = a_0 + 33 \cdot (-0{,}528) = a_0 - 17{,}424;$$

$$a_0 = 95{,}424.$$

Отже, рівняння параболи другого порядку матиме вигляд:

$$\hat{y}_t = 95{,}424 + 3{,}667\, t - 0{,}528\, t^2.$$

Розрахуємо вирівняні рівні:

для 1998 р. $\hat{y}_t = 95{,}424 + 3{,}667 \cdot (-9) - 0{,}528 \cdot 81 = 19{,}653$;

для 1999 р. $\hat{y}_t = 95{,}424 + 3{,}667 \cdot (-7) - 0{,}528 \cdot 49 = 43{,}883$;

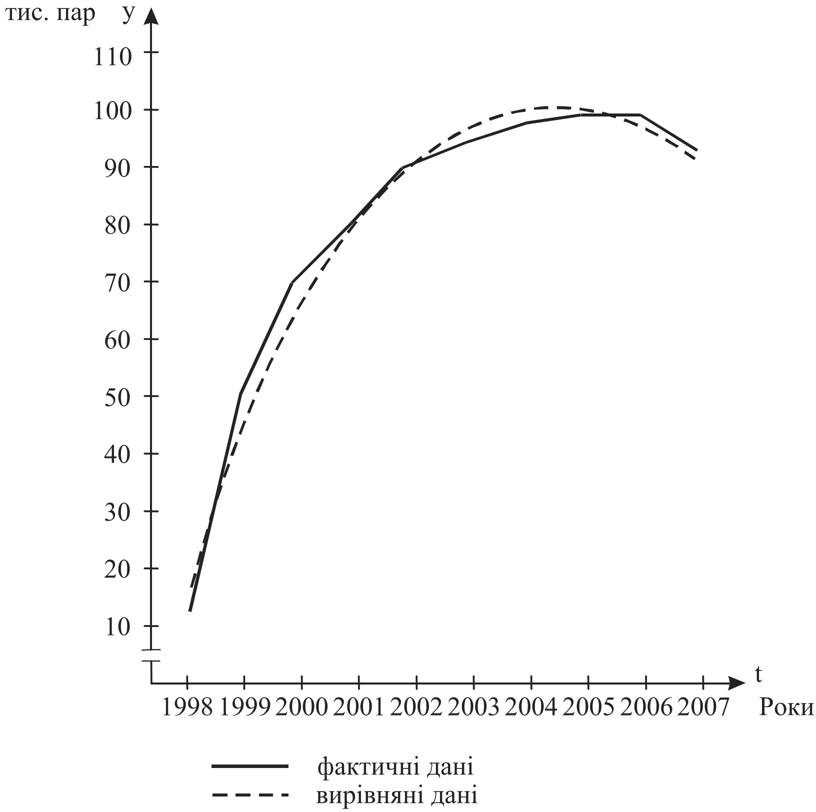
для 2000 р. $\hat{y}_t = 95{,}424 + 3{,}667 \cdot (-5) - 0{,}528 \cdot 25 = 69{,}889$;

і т.д.

Результати вирівнювання записані в останній графі табл. 11.17. Як видно із розрахунків, і в даному прикладі вирівняні рівні динамічного ряду досить близькі до даних емпіричного ряду. Отже, парабола другого порядку досить точно відображає тренд на даному відрізку часу.

Параметри рівняння параболи другого порядку потрібно інтерпретувати наступним чином: $a_0$ – це величина, яка виражає середні умови утворення рівнів ряду; $a_1$ – швидкість розвитку даних динамічного ряду; $a_2$ – характеризує прискорення цього розвитку.

Проілюструємо вирівнювання за параболою другого порядку на графіку (мал. 11.4).




Мал. 11.4. Виробництво взуття на одній із фабрик в області.

Як показує графік, розраховані нами дані досить добре відтворюють емпіричні рівні ряду.

В наукових дослідженнях аналітичне вирівнювання проводять також і за багаточленами більш високих ступенів, таких як парабола третього порядку, котра описується рівнянням:

$$\hat{y}_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3 .$$

Потрібно мати на увазі, що чим вищий порядок параболи, тим вона більш точно відтворює фактичні дані.

*Вирівнювання за показниковою функцією* проводиться в тих випадках, коли динамічний ряд розвивається в геометричній прогресії, тобто тоді, коли ланцюгові темпи росту більш-менш постійні.

Показникова функція, яка використовується для вирівнювання динамічного ряду описується рівнянням:

$$\hat{y}_t = a_0 \cdot a_1^t .$$

Для визначення параметрів $a_0$ і $a_1$ цього рівняння методом найменших квадратів попередньо логарифмують рівні, тоді логарифми показникової функції описують лінійною функцією:

$$\lg \hat{y}_t = \lg a_0 + t \lg a_1 .$$

Система нормальних рівнянь в даному випадку має такий вигляд:

$$\begin{cases} \sum \lg y = n \lg a_0 + \lg a_1 \sum t; \\ \sum t \lg y = \lg a_0 \sum t + \lg a_1 \sum t^2, \end{cases}$$

коли добитись, щоб $\sum t = 0$, тоді

$$\begin{cases} \sum \lg y = n \lg a_0; \\ \sum t \lg y = \lg a_1 \sum t^2, \end{cases}$$

звідки $\lg a_0 = \dfrac{\sum \lg y}{n}$ і $\lg a_1 = \dfrac{\sum t \lg y}{\sum t^2}$.

Розглянемо вирівнювання за показниковою функцією на прикладі динаміки росту чисельності населення одного з міст області України за 2001-2007 рр. (табл. 11.18, дані умовні):

Таблиця 11.18

| Роки n | Чисельність населення міста на початок року, тис. чол. y | lg y | t | $t^2$ | t lg y | $\lg \hat{y}_t$ | Вирівняні рівні $\hat{y}_t = a_0 a_1^t$ |
|---|---|---|---|---|---|---|---|
| 2001 | 200,0 | 2,3010 | -3 | 9 | -6,9030 | 2,3006 | 199,8 |
| 2002 | 205,8 | 2,3135 | -2 | 4 | -4,6270 | 2,3138 | 206,0 |
| 2003 | 212,3 | 2,3269 | -1 | 1 | -2,3269 | 2,3270 | 212,3 |
| 2004 | 218,8 | 2,3401 | 0 | 0 | 0 | 2,3402 | 218,9 |
| 2005 | 225,5 | 2,3532 | 1 | 1 | 2,3532 | 2,3534 | 225,6 |
| 2006 | 232,6 | 2,3666 | 2 | 4 | 4,7332 | 2,3666 | 232,6 |
| 2007 | 240,0 | 2,3802 | 3 | 9 | 7,1406 | 2,3798 | 239,7 |
| n=7 | $\sum y = 1535$ | $\sum \lg y = 16,3815$ | $\sum t = 0$ | $\sum t^2 = 28$ | $\sum t \lg y = 0,3701$ | $\sum \lg \hat{y}_t = 16,3814$ | $\sum \hat{y}t = 1534,9$ |

Використавши розраховані в таблиці дані, визначимо коефіцієнти:

$$\lg a_0 = \frac{\sum \lg y}{n} = \frac{16{,}3815}{7} = 2{,}3402 \text{ , } a_0 = 218{,}9;$$

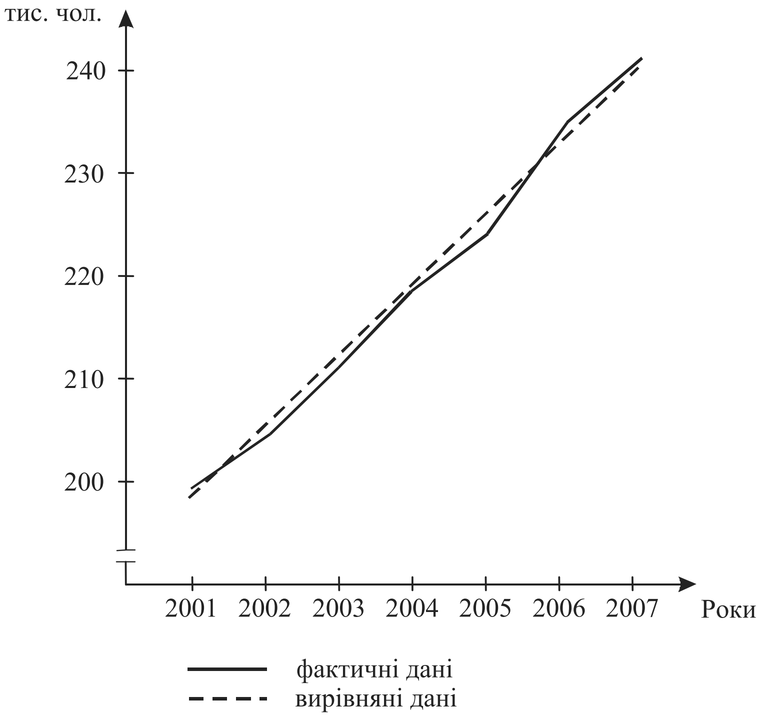
$$\lg a_1 = \frac{\sum t \lg y}{\sum t^2} = \frac{0{,}3701}{28} = 0{,}0132 \text{ , } a_1 = 1{,}030.$$

Отже $\lg \hat{y}_t = 2{,}3402 + 0{,}0132\ t$ , або $\hat{y}_t = 218{,}9 \cdot 1{,}03^t$ .

Для розрахунку вирівняних рівнів користуються логарифмічною функцією підставляючи послідовно значення t = -3, -2, -1, 0, +1, +2, +3, а знайшовши ($\lg \hat{y}_t$) за таблицями десяткових антилогарифмів визначають теоретичні (вирівняні) значення ($\hat{y}_t$).

Логарифми вирівняних рівнів і самі теоретичні рівні наведені в двох останніх колонках табл. 11.18.

Результати проведеного вирівнювання ряду динаміки чисельності населення одного з міст України за 2001-2007 рр. покажемо на графіку (мал. 11.5).




Мал. 11.5 Динаміка росту чисельності міста.

Як за даними таблиці 11.18, так і на графіку (мал. 11.5) чітко простежується близькість емпіричних і теоретичних рівнів, що свідчить про правильність вибраної нами функції для вирівнювання.

Коефіцієнт $a_1$ в показниковій функції характеризує середній темп росту досліджуваної ознаки. В нашому прикладі $a_1 = 1{,}03$ означає, що чисельність населення міста щорічно збільшується в 1,03 рази, або на 3 %.

Особливе місце в аналітичному вирівнюванні рядів динаміки займає *вирівнювання за допомогою ряду Фур'є*, який описується рівнянням:

$$\hat{y}_t = a_0 + \sum_{n=1}^{m}\left(a_n \cos kt + a_{n+1} \sin kt\right),$$

де k – ступінь точності гармонік (найчастіше від 1 до 4);
t – час, виражений в радіанній мірі або градусах.

Вирівнювання за наведеною формулою проводять в тих випадках, коли в емпіричному ряду спостерігається певна періодичність змін його рівнів, яка виступає у вигляді синусоїдних коливань. Останні являють собою гармонійні коливання, а синусоїди, отримані при вирівнюванні рядом Фур'є, називають гармоніками відповідних порядків.

При вирівнюванні по ряду Фур'є періодичні коливання рівнів динамічного ряду виступають у вигляді суми декількох гармонік, нашарованих одна на одну. Так, наприклад, при k = 1 рівняння Фур'є матиме вигляд:

$$\hat{y}_t = a_0 + a_1 \cos t + a_2 \sin t\,.$$

При k = 2 відповідно:

$$\hat{y}_t = a_0 + a_1 \cos t + a_2 \sin t + a_3 \cos 2t + a_4 \sin 2t\text{ , і т.д.}$$

В залежності від величини часу (t) знаходять за певною таблицею (11.19) відповідні значення cos і sin.

Таблиця 11.19

Час в радіанній мірі і в градусах.

| Число рівнів (n) | I | II | III | IV | V | VI | VII | VIII | IX | X | XI | XII |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Радіанна міра (t) | 0 | $\frac{\pi}{6}$ | $\frac{\pi}{3}$ | $\frac{\pi}{2}$ | $\frac{2\pi}{3}$ | $\frac{5\pi}{6}$ | $\pi$ | $\frac{7\pi}{6}$ | $\frac{4\pi}{3}$ | $\frac{3\pi}{2}$ | $\frac{5\pi}{3}$ | $\frac{11\pi}{6}$ |
| Градуси | 0 | 30 | 60 | 90 | 120 | 150 | 180 | 210 | 240 | 270 | 300 | 330 |

Вирівнювання по ряду Фур'є, як правило, використовують для дослідження сезонності різних соціально-економічних явищ і процесів.

Параметри рівняння теоретичних рівнів визначаються за способом найменших квадратів.

Знайшовши часткові похідні функції ряду Фур'є і прирівнявши їх до нуля, отримаємо систему нормальних рівнянь, за якими можна вирахувати параметри:

$$a_0 = \frac{\sum y}{n},$$

$$a_1 = \frac{2\sum y \cos t}{n},$$

$$a_2 = \frac{2\sum y \sin t}{n}.$$

Покажемо вирівнювання за періодичною функцією ряду Фур'є на умовному прикладі про об'єм реалізації побутових холодильників торговельними підприємствами однієї із облспоживспілок України за 2005-2007 рр. (табл. 11.20). В цій же таблиці наведемо значення cos і sin для відповідних t.

Таблиця 11.20

| Рік, квартал (n) | Радіанна міра (t) | Реалізація холодильників, шт. (y) | cos t | sin t | $y\cos t$ | $y\sin t$ | Вирівняні рівні $\hat{y}_t = a_0 + a_1\cos kt + a_2\sin kt$ |
|---|---|---|---|---|---|---|---|
| 2005 I | 0 | 1942 | 1,000 | 0 | 1942,00 | 0 | 2733,03 |
| II | π/6 | 2957 | 0,866 | 0,500 | 2560,76 | 1478,50 | 2533,65 |
| III | π/3 | 2504 | 0,500 | 0,866 | 1250,00 | 2168,46 | 2370,34 |
| IV | π/2 | 2194 | 0 | 1,000 | 0 | 2194,00 | 2286,84 |
| 2006 I | 2π/3 | 2126 | -0,500 | 0,866 | -1063,00 | 1841,12 | 2305,56 |
| II | 5π/6 | 2704 | -0,866 | 0,500 | -2341,66 | 1352,00 | 2421,45 |
| III | π | 3291 | -1,000 | 0 | -3291,00 | 0 | 2603,47 |
| IV | 7π/6 | 1745 | -0,866 | -0,500 | -1511,17 | -872,50 | 2802,85 |
| 2007 I | 4π/3 | 2505 | -0,500 | -0,866 | -1252,50 | -2169,33 | 2966,16 |
| II | 3π/2 | 3704 | 0 | -1,000 | 0 | -3704,00 | 3049,66 |
| III | 5π/3 | 3834 | 0,500 | -0,866 | 1917,00 | -3320,24 | 3030,94 |
| IV | 11π/6 | 2513 | 0,866 | -0,500 | 2176,26 | -1256,50 | 2915,05 |
| Разом | × | 32019 | × | × | +388,69 | -2288,49 | 32019,00 |

В таблиці 11.20 показано розрахунок необхідних даних для знаходження параметрів рівняння ряду Фур'є за першою гармонікою:

$$a_0 = \frac{\sum y}{n} = \frac{32019}{12} = 2668{,}25;$$

$$a_1 = \frac{2\sum y \cos t}{n} = \frac{\sum y \cos t}{n/2} = \frac{388{,}69}{6} = 64{,}78;$$

$$a_2 = \frac{2\sum y \sin t}{n} = \frac{\sum y \sin t}{n/2} = \frac{-2288{,}49}{6} = -381{,}41.$$

Звідси $\hat{y}_t = 2668{,}25 + 64{,}78\cos t - 381{,}41\sin t$.

Підставляючи в дане рівняння значення cos t і sin t (із табл. 11.20), отримаємо теоретичні значення обсягу реалізації холодильників по кварталах року:

$$\hat{y}_I = 2668{,}25 + 64{,}78 \cdot 1{,}000 - 381{,}41 \cdot 0 = 2733{,}03;$$

$$\hat{y}_t = 2668{,}25 + 64{,}78 \cdot 0{,}866 - 381{,}41 \cdot 0{,}500 = 2533{,}65;$$

$$\hat{y}_t = 2668{,}25 + 64{,}78 \cdot 0{,}500 - 381{,}41 \cdot 0{,}866 = 2370{,}34;$$

і т.д.

Як показують дані таблиці 11.20, вже перша гармоніка ряду Фур'є досить добре апроксимує емпіричний ряд динаміки.

Порівняння емпіричних і теоретичних ліній реалізації побутових холодильників в облспоживспілці (мал. 11.6) показує, що вирівнювання за рівнянням першої гармоніки ряду Фур'є рівнів динамічного ряду досить точно відтворюють фактичні дані реалізації цієї групи товарів.

Аналогічно розраховують рівняння ряду Фур'є із застосуванням другої, третьої і четвертої гармонік з перевіркою близькості їх теоретичних і фактичних значень.

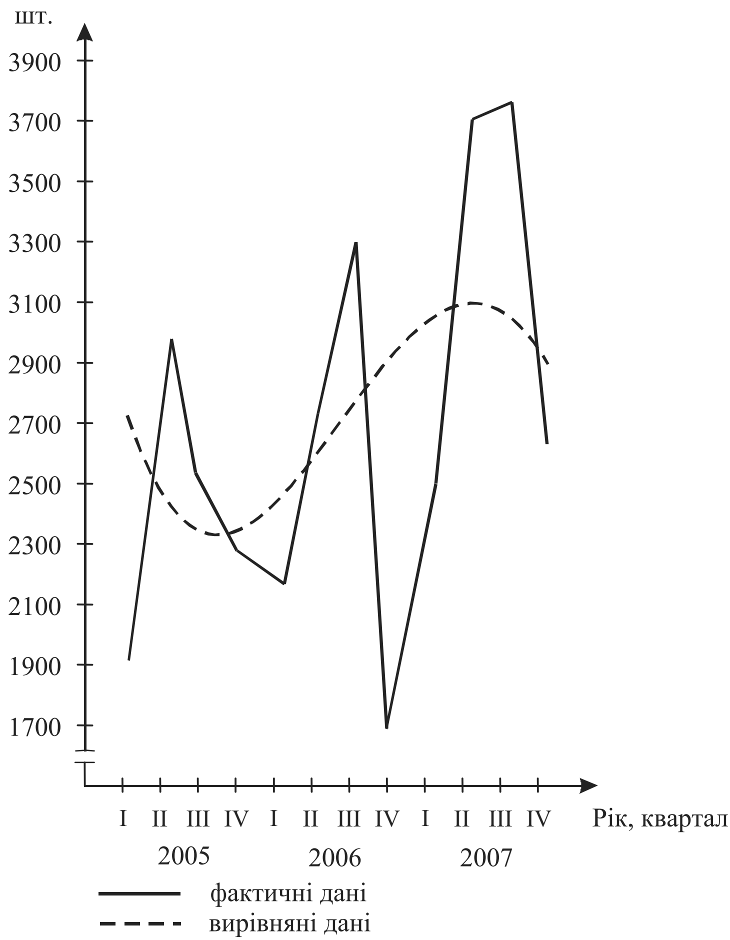
Вирівнювання рядів динаміки відіграє важливу роль в аналізі соціально-економічних процесів, які змінюються в часі. Правильний підбір типу прямої або кривої для визначення тренду має велике теоретичне і практичне значення, особливо при прогнозуванні.

Вирівнювання рядів динаміки використовують також для знаходження відсутніх членів ряду за допомогою інтерполяції і екстраполяції.

*Інтерполяцією* називається в статистиці знаходження відсутнього показника всередині ряду. Вона ґрунтується на припущенні, що за наявними даними можна визначити характер розвитку явища в цілому.

Найбільш простим є знаходження відсутнього рівня всередині ряду у випадку зміни явища по прямій. Наприклад, на початок 2007 р.

в селянській спілці числилось 10 тис. голів овець, а на кінець року – 12,4 тис. голів. Потрібно визначити ймовірну чисельність овець на 1.04.2007 р.




Мал. 11.6 Поквартальна реалізація побутових холодильників торговими підприємствами споживчої кооперації області.

Визначимо річний абсолютний приріст овець:

$$\Delta_y = y_{12} - y_1 = 12,4 - 10,0 = 2,4 \text{ тис. голів.}$$

Знаходимо середньомісячний абсолютний приріст:

$$\overline{\Delta_y} = \frac{\sum \Delta y}{n} = \frac{2,4}{12} = 0,2 \text{ тис. голів.}$$

Якщо припустити, що кожного місяця абсолютний приріст овець був приблизно однаковий, тоді на 1.04.1996 р. в спілці числилось 10,6 тис. голів:

$$\hat{y}_{1.04.2007\,р.} = y_1 + \overline{\Delta_y} \cdot t = 10 + 0,2 \cdot 3 = 10,6 \text{ тис. голів.}$$

*Екстраполяцією* в статистиці називається знаходження невідомих рівнів в кінці або на початку динамічного ряду. Цей прийом заключається в тому, що за знайденими математичними рівняннями передбачають попередній або майбутній розвиток явищ. Звернемось до прикладу.

Нехай в місті на 1.01.2007 р. проживало 200 тис. чол. Середньорічний темп приросту за попередні п’ять років склав 2 %. Потрібно визначити ймовірну чисельність населення міста на 1.01.2010 р.

Для знаходження перспективної чисельності населення станом на 1.01.2010 р. використовуємо формули:

$$\hat{y}_t = y_1 \cdot \overline{T}^t \text{, або } \hat{y}_t = y_1 + \overline{\Delta_y} \cdot t \text{ ;}$$

$$\hat{y}_{1.01.2010\,р.} = 200 \cdot 1,02^3 = 212 \text{ тис. чол., або}$$

$$\hat{y}_{1.01.2010\,р.} = 200 + 4 \cdot 3 = 212 \text{ тис. чол.}$$

Як інтерполяція, так і екстраполяція ґрунтуються на припущенні, що наявні величини цілком достатньо визначають темп розвитку досліджуваного явища і, отже, його можна поширювати на відсутні рівні динамічного ряду.

Іноді виникає потреба порівняти між собою зміну динамічних рядів за ряд років декількох споріднених або взаємозв’язаних явищ (наприклад, виплавку чавуну і сталі, видобуток вугілля, нафти і газу і т.п.). З цією метою переводять абсолютні показники рядів динаміки у відносні, прийнявши рівні якого-небудь періоду за одиницю або сто. Таке перетворення динамічних рядів називається *приведенням їх до однієї основи*. Розглянемо приклад (табл. 11.21).

За основу для порівняння ми взяли абсолютний рівень 2003 р. В даному прикладі темпи росту грошових заощаджень населення в області “А” значно вищі, ніж в області “Б”.

В процесі аналізу динамічних рядів часом доводиться обчислювати коефіцієнт випередження, який являє собою відношення більшого середньорічного темпу приросту до меншого і показує, у скільки разів швидше зростає рівень одного ряду динаміки порівняно з іншими. Так, наприклад, за п’ятиріччя середньорічний темп приросту грошових надходжень у відділеннях національного банку становив 19,5 %, а у відділення Промінвестбанку – 6,5 %. Отже коефіцієнт випередження дорівнює: 19,5 : 6,5 = 3. Таким чином, грошові надходження у відділення Національного банку зростали в три рази швидше, ніж у відділення Промінвестбанку.

Таблиця 11.21

Грошові заощадження населення в двох областях України (дані умовні).

| Роки | Заощадження населення | | | |
|---|---|---|---|---|
| | область “А” | | область “Б” | |
| | тис. грн. | в % до 2000 р. | тис. грн. | в % до 2000 р. |
| 2003 | 815 | 100,0 | 631 | 100,0 |
| 2004 | 849 | 104,2 | 637 | 100,9 |
| 2005 | 893 | 109,6 | 642 | 101,7 |
| 2006 | 900 | 110,4 | 658 | 104,3 |
| 2007 | 915 | 112,3 | 660 | 104,6 |

В економічній практиці при контролі за виконанням планових завдань використовують *динамічні ряди наростаючих підсумків*. Такий контроль здійснюється шляхом процентного відношення кожного окремого підсумку до планового завдання на весь період. Внаслідок цього отримуємо ряд чисел, які характеризують хід виконання планового завдання в процентному відношенні. Наприклад, підприємство планувало реалізувати готової продукції за рік в обсязі 24 млн. грн., фактично реалізувало в січні продукції на суму 1,3 млн. грн., в лютому – 1,8 млн. грн. і в березні – 2,4 млн. грн.

Визначимо відсотки виконання планового завдання реалізації продукції наростаючим підсумком з початку року:

за січень $1{,}3 : 24 \cdot 100 = 5{,}4$ %;

за січень і лютий $(1{,}3 + 1{,}8) : 24 \cdot 100 = 12{,}9$ %;

за січень, лютий і березень $(1{,}3 + 1{,}8 + 2{,}4) : 24 \cdot 100 = 22{,}9$ %;

і т.д.

## 11.5. Вимірювання сезонних коливань

*Сезонними коливаннями* називаються більш-менш стійкі внутрішньорічні коливання в рядах динаміки, обумовлені специфічними умовами виробництва чи споживання певного виду продукції.

Для дослідження внутрішньорічних коливань можна використати цілий ряд методів (простої середньої, Персонса, рухомої середньої, аналітичного вирівнювання, рядів Фур'є), які забезпечують їх оцінку з різною точністю, надійністю і трудоємкістю.

Аналіз сезонності покажемо на прикладі реалізації товарів культурно-побутового призначення за допомогою різних методів, з поступовим переходом від простих способів дослідження до більш складних.

Сезонні коливання характеризуються спеціальним показником, який називається індексом сезонності ($I_s$). В сукупності ці індекси утворюють сезонну хвилю.

*Індекс сезонності* – це процентне відношення однойменних місячних (квартальних) фактичних рівнів динамічних рядів до їх середньорічних або вирівняних рівнів.

Наочну уяву про зміну попиту населення на товари культурно-побутового призначення в окремі періоди року дають графіки. Наприклад, поквартальна реалізація побутових холодильників торговими підприємствами споживчої кооперації області за 2002-2004 рр. показана на графіку (мал. 11.6).

Індекси сезонності (сезонну хвилю) реалізації цих товарів розрахуємо *методом простих середніх*.

Індекс сезонності за методом простої середньої визначається за формулою:

$$I_s = \frac{\overline{y}_i}{\overline{y}_з} \cdot 100 ,$$

де $\overline{y}_i$ – середні місячні або квартальні рівні;

$\overline{y}_з$ – загальна середня (місячна або квартальна).

Сезонну хвилю, або індекс сезонності реалізації побутових холодильників розрахуємо в три етапи, для чого складемо розрахункову таблицю 11.22.

Таблиця 11.22

Розрахунок сезонної хвилі реалізації побутових холодильників торговими підприємствами споживчої кооперації області за 2005-2007 рр., шт.

| Квартал | Роки | | | Разом | В середньому $\overline{y}_i$ | Сезонна хвиля $I_s = \frac{\overline{y}_i}{\overline{y}_з} \cdot 100$ |
|---|---|---|---|---|---|---|
| | 2005 | 2006 | 2007 | | | |
| I | 1942 | 2126 | 2505 | 6573 | 2191,00 | 82,1 |
| II | 2957 | 2704 | 3704 | 9365 | 3121,67 | 117,0 |
| III | 2504 | 3291 | 3834 | 9629 | 3209,67 | 120,3 |
| IV | 2194 | 1745 | 2513 | 6452 | 2150,67 | 80,6 |
| Разом | 9597 | 9866 | 12556 | 32019 | $\overline{y}_з = 2668,25$ | х |

На першому етапі розрахуємо середню реалізацію холодильників в кожному кварталі за три роки. Цей розрахунок ліквідовує вплив випадкових причин і проводиться за середньою арифметичною простою. Наприклад, середнє сезонне коливання реалізації холодильників в І кварталі за три роки:

$$\overline{y}_I = \frac{6573}{3} = 2191 \text{ шт.}$$

Аналогічно розраховуються сезонні коливання для II, III і IV кварталів.

На другому етапі визначаємо середню реалізацію холодильників за весь досліджуваний період.

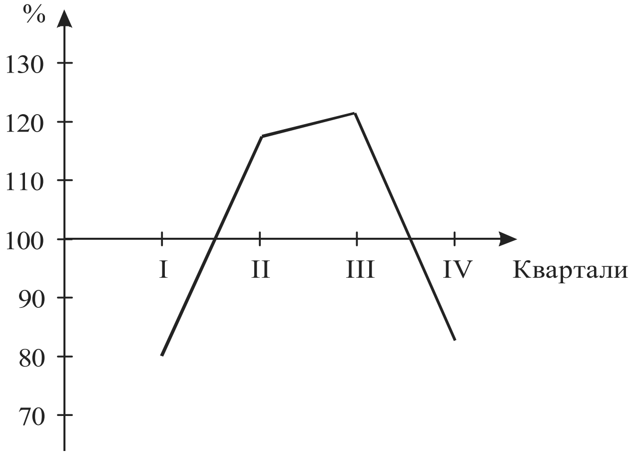
$$\overline{y}_з = \frac{32019}{12} = 2668,25 \text{ шт.}$$

На третьому етапі обчислимо сезонну хвилю, або індекс сезонності за весь період:

$$I_{s_I} = \frac{\overline{y}_1}{\overline{y}_з} \cdot 100 = \frac{2191,00}{2668,25} \cdot 100 = 82,1 \text{ \%;}$$

$$I_{s_{II}} = \frac{\overline{y}_2}{\overline{y}_з} \cdot 100 = \frac{3121,67}{2668,25} \cdot 100 = 117,0 \text{ \%, і т.д.}$$

Покажемо сезонну хвилю реалізації холодильників на графіку (мал. 11.7).




Мал. 11.7. Сезонна хвиля реалізації побутових холодильників системою споживчої кооперації області за 2005-2007 рр.

Судячи з таблиці 11.22 і графіку (мал. 11.7), реалізація побутових холодильників суттєво падає в першому і четвертому кварталах і різко зростає в другому і третьому кварталах року. В середньому за досліджуваний період цих товарів продавалось в першому кварталі на 17,9 пунктів (82,1 – 100), а в четвертому кварталі на 19,4 пунктів (80,6 – 100) менше середньо квартальної реалізації, тоді як в другому кварталі на 17,0 (117,0 – 100) і третьому на 20,3 (120,3 – 100) пунктів більше цього рівня.

Більшість динамічних рядів досліджуваних явищ мають тенденцію росту, тому для більш точного визначення сезонної хвилі в таких рядах необхідна нейтралізація еволюції тренду.

З цією метою використовуємо *метод ланцюгових індексів (метод Персонса)*, спосіб розрахунку якого покажемо в табл. 11.23.

Таблиця 11.23

Розрахунок сезонної хвилі реалізації холодильників побутових системою споживчої кооперації області методом У. Персонса.

| Роки / Квартали | Реалізація побутових холодильників, шт. | | | Ланцюгові індекси по роках | | | Середні ланцюгові індекси | Базисні індекси $y_i$ | Сезонна хвиля, % $I_s = \frac{\overline{y}_i}{\overline{y}_з} \cdot 100$ |
|---|---|---|---|---|---|---|---|---|---|
| | 2005 | 2006 | 2007 | 2005 | 2006 | 2007 | | | |
| I | 1942 | 2126 | 2505 | - | 0,969 | 1,436 | 1,202 | 1,000 | 82,0 |
| II | 2957 | 2704 | 3704 | 1,523 | 1,272 | 1,479 | 1,425 | 1,425 | 116,9 |
| III | 2504 | 3291 | 3834 | 0,847 | 1,179 | 1,035 | 1,020 | 1,453 | 119,2 |
| IV | 2194 | 1745 | 2513 | 0,876 | 0,530 | 0,655 | 0,687 | 0,998 | 81,9 |
| Разом | 9597 | 9866 | 12556 | 3,246 | 3,950 | 4,605 | × | 4,876 | 400,0 |
| В середньому | × | × | × | 1,082 | 0,988 | 1,151 | × | $\overline{y}_з = 1{,}219$ | 100,0 |

Визначаємо ланцюгові індекси як відношення реалізації холодильників за кожний наступний квартал до попереднього, тобто 2955 : 1942 = 1,523; 2504 : 2957 = 0,847 і т.д.

Потім знаходимо середні ланцюгові індекси за кожний квартал: (0,969 + 1,436) : 2 = 1,202; (1,523 + 1,272 + 1,479) : 3 = 1,425 і т.д.

Отриманий ряд середніх ланцюгових індексів служить основою для визначення базисних індексів. Прийнявши середню за перший квартал за 100 % і залишивши середню за другий квартал без змін (1,425), перераховуємо всі інші середні по відношенню до першого кварталу. Тоді відношення рівня третього кварталу до першого складе: 1,425 · 1,020 = 1,453; 1,453 · 0,687 = 0,998.

Процентне відношення базисних індексів за кожний квартал до середньорічного індексу дасть нам сезонну хвилю реалізації холодильників побутових:

$$I_{s_I} = \frac{y_I}{\bar{y}_з} \cdot 100 = \frac{1{,}000}{1{,}219} \cdot 100 = 82{,}0\ \%;$$

$$I_{s_{II}} = \frac{y_{II}}{\bar{y}_з} \cdot 100 = \frac{1{,}425}{1{,}219} \cdot 100 = 116{,}9\ \%\ \text{і т.д.}$$

Метод Персонса дозволяє визначити зсув сезонної хвилі під впливом загальної тенденції розвитку реалізованого попиту на холодильники (1,202 · 0,998 = 1,199) або 19,9 %. Цей зсув визначається як добуток середнього ланцюгового індексу за перший квартал на значення базисного індексу за четвертий квартал.

Сезонну хвилю методом Персонса можна також визначити і за медіанними значеннями ланцюгових індексів. При вирахуванні індексів сезонності враховують зсув сезонної хвилі під впливом тренду. Виходячи з гіпотези, що загальна тенденція ряду динаміки розвивається по прямій лінії, виключення тренду із сезонної хвилі проводять шляхом рівномірного розподілу зсуву по квартальних значеннях базисних індексів.

Скореговані базисні індекси і розрахунок хвилі показані в табл. 11.24.

Таблиця 11.24

Розрахунок виправленої сезонної хвилі реалізації холодильників побутових за рахунок виключення тренду.

| Квартал | Медіанний ланцюговий індекси | Базисні індекси | Базисні індекси після виключення зсуву загальної тенденції ряду | Сезонна хвиля $I_s = \frac{\bar{y}_i}{\bar{y}_з} \cdot 100$ |
|---|---|---|---|---|
| I | 0,969 | 1,000 | 1,000 | 81,4 |
| II | 1,272 | 1,272 | 1,330 | 108,3 |
| III | 1,179 | 1,499 | 1,614 | 131,5 |
| IV | 0,530 | 0,794 | 0,967 | 78,8 |
| Разом | × | × | 4,911 | 400,0 |
| В середньому | × | × | 1,228 | 100,0 |

Медіанні ланцюгові індекси беруться із середнього ряду ланцюгових індексів, тобто ряду ланцюгових індексів другого року порівняння. Цей ряд знову перетворюємо в ряд базисних індексів, а потім визначаємо зсув сезонної хвилі, як було показано в попередньому прикладі. В цьому прикладі він становить:

$$0{,}969 \cdot 0{,}794 = 0{,}769 \text{ або } -23{,}1\ \%.$$

В результаті поступового збільшення реалізації холодильників побутових при переході від одного кварталу даного року до такого ж кварталу наступного року зсув сезонної хвилі склав 23,1 %.

Виключаємо із сезонної хвилі тренд:

I квартал – 1,000;

II квартал – $1{,}272 + \dfrac{0{,}231 \cdot 1}{4} = 1{,}330$ ;

III квартал – $1{,}499 + \dfrac{0{,}231 \cdot 2}{4} = 1{,}614$ ;

IV квартал – $0{,}794 + \dfrac{0{,}231 \cdot 3}{4} = 0{,}967$ .

Поділивши скореговані базисні індекси за кожний квартал на середньо квартальний рівень за всі роки отримаємо виправлену сезонну хвилю реалізації холодильників побутових торговими підприємствами споживчої кооперації області.

Розвиток загальної тенденції різних динамічних рядів по прямій лінії зустрічається в реальній дійсності далеко не завжди. Він може приймати самі різноманітні форми, а тому для розрахунку сезонної хвилі доцільно використовувати і інші методи елімінування тренду, такі як рухома середня, аналітичне вирівнювання і ряд Фур'є.

Для вивчення сезонності часто доводиться вираховувати *рухому середню* з парним числом членів ряду, тому що характер динамічного ряду визначає тривалість періоду рухомої середньої, який повинен співпадати з періодом коливання, або бути кратним йому.

Наприклад, при розрахунку загальної тенденції реалізованого попиту на мотоцикли і велосипеди державної і кооперативної торгівлі України (табл. 11.25, дані умовні), маючи справу з квартальними даними, періодом рухомої середньої доцільно взяти чотири квартали (рік), так як коливання в досліджуваному емпіричному ряду повторюються щорічно.

Згладжування за парним числом членів ряду незручне тим, що середня мусить бути віднесена тільки до середини між двома датами, тобто проходить зсув періоду, до якого відноситься рівень.

Статистика для ліквідації такого зсуву використовує, розглянуті нами раніше, способи перетворення рівнів і центрування.

Усунення зсуву періоду проведемо способом центрування рівнів ряду динаміки.

Таблиця 11.25

Розрахунок відносних сезонних коливань реалізації мотоциклів і велосипедів державної і кооперативної торгівлі України після визначення і виключення загальної тенденції розвитку попиту методом рухомої середньої за парним числом членів ряду.

| Рік, квартал | | Реалізація мотоциклів і велосипедів, млн. грн. (y) | Сума чотирьох рівнів ряду | Центрування суміжних сум | Згладжений ряд за парними рівнями ряду ($\hat{y}_t$) | Сезонна хвиля $I_s = \frac{y_i}{\hat{y}_з} \cdot 100$ |
|---|---|---|---|---|---|---|
| А | | 1 | 2 | 3 | 4 | 5 |
| 2003 | I | 79,4 | - | - | - | - |
| | II | 102,4 | - | - | 90,4 | 113,3 |
| | | - | 330,2 | - | - | - |
| | III | 89,5 | - | 658,6 | 82,3 | 108,7 |
| | | - | 328,4 | - | - | - |
| | IV | 58,9 | - | 660,0 | 82,5 | 71,4 |
| 2004 | | - | 331,6 | - | - | - |
| | I | 77,6 | - | 661,5 | 82,7 | 93,8 |
| | | - | 329,9 | - | - | - |
| | II | 105,6 | - | 664,8 | 83,1 | 127,1 |
| | | - | 334,9 | - | - | - |
| | III | 87,8 | - | 671,3 | 83,9 | 104,6 |
| | | - | 336,4 | - | - | - |
| | IV | 63,9 | - | 656,3 | 82,0 | 77,9 |
| 2005 | | - | 319,9 | - | - | - |
| | I | 79,1 | - | 645,3 | 80,7 | 98,0 |
| | | - | 325,4 | - | - | - |
| | II | 89,1 | - | 674,1 | 84,3 | 105,7 |
| | | - | 348,7 | - | - | - |
| | III | 93,3 | - | 707,4 | 88,4 | 105,5 |
| | | - | 358,7 | - | - | - |
| | IV | 87,2 | - | 749,1 | 93,6 | 93,2 |
| 2006 | | - | 390,4 | - | - | - |
| | I | 89,1 | - | 788,0 | 98,5 | 90,4 |
| | | - | 397,6 | - | - | - |
| | II | 102,8 | - | 772,1 | 96,5 | 125,2 |
| | | - | 374,5 | - | - | - |
| | III | 100,5 | - | 755,0 | 94,4 | 106,5 |
| | | - | 380,5 | - | - | - |
| | IV | 64,1 | - | 771,1 | 96,4 | 66,5 |
| 2007 | | - | 390,6 | - | - | - |
| | I | 95,1 | - | 788,5 | 98,6 | 96,4 |
| | | - | 397,9 | - | - | - |
| | II | 130,9 | - | 800,3 | 100,0 | 130,9 |
| | | - | 402,4 | - | - | - |
| | III | 107,8 | - | - | 102,4 | 105,3 |
| | | - | - | - | - | - |
| | IV | 68,6 | - | - | - | - |

Перший і останній рівні згладженого ряду вираховуються як середня рухома із відповідних трьох рівнів: (79,4 + 102,4 + 89,2) : 3 = 90,4; (130,9 + 107,8 + 68,6) : 3 = 102,4.

Потім з першого кварталу по четвертий, а з другого кварталу першого року по перший квартал другого року включно і т.д. визначаємо суми чотирьох рівнів ряду:

(79,4 + 102,4 + 89,5 + 58,9) = 330,2 млн. грн.;

(102,4 + 89,5 + 58,9 + 77,6) = 328,4 млн. грн.;

і т.д.

Отримані дві суміжні суми центруємо:

330,2 + 328,4 = 658,6 млн. грн.

Вираховуємо рухому середню для другого рівня ряду:

658,6 : 8 = 82,3 млн. грн., і т.д.

Згладжений ряд реалізації мотоциклів і велосипедів за чотирьохчленною рухомою середньою в найбільшій мірі усуває випадкові коливання і відображає загальну тенденцію розвитку реалізованого попиту на ці товари.

Для розрахунку сезонних коливань реалізації мотоциклів і велосипедів державної і кооперативної торгівлі України використовуємо також метод *аналітичного вирівнювання* за рівнянням прямої (табл. 11.26).

Таблиця 11.26

| Рік, квартал | | Реалізація мотоциклів і велосипедів, млн. грн. (y) | t | $t^2$ | yt | $\hat{y}_t$ | Сезонна хвиля $I_s = \frac{y_i}{\hat{y}_з} \cdot 100$ |
|---|---|---|---|---|---|---|---|
| А | | 1 | 2 | 3 | 4 | 5 | 6 |
| 2003 | I | 79,4 | -19 | 361 | -1508,6 | 81,2 | 97,8 |
| | II | 102,4 | -17 | 289 | -1740,8 | 82,1 | 124,2 |
| | III | 89,5 | -15 | 225 | -1342,5 | 82,9 | 108,0 |
| | IV | 58,9 | -13 | 169 | -765,7 | 83,8 | 70,3 |
| 2004 | I | 77,6 | -11 | 121 | -853,6 | 84,7 | 91,6 |
| | II | 105,6 | -9 | 81 | -950,4 | 85,6 | 123,4 |
| | III | 87,8 | -7 | 49 | -614,6 | 86,5 | 101,4 |
| | IV | 63,9 | -5 | 25 | -319,5 | 87,3 | 73,2 |
| 2005 | I | 79,1 | -3 | 9 | -237,3 | 88,2 | 89,7 |
| | II | 89,1 | -1 | 1 | -89,1 | 89,1 | 100,0 |
| | III | 93,3 | 1 | 1 | 93,3 | 90,0 | 103,7 |
| | IV | 87,2 | 3 | 9 | 261,6 | 90,8 | 96,0 |

Продовження табл. 11.26

| А | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 2006 | I | 89,1 | 5 | 25 | 445,5 | 91,7 | 97,2 |
| | II | 120,8 | 7 | 49 | 845,6 | 92,6 | 130,4 |
| | III | 100,5 | 9 | 81 | 904,5 | 93,5 | 107,5 |
| | IV | 64,1 | 11 | 121 | 705,1 | 94,4 | 67,9 |
| 2007 | I | 95,1 | 13 | 169 | 1236,3 | 95,2 | 99,9 |
| | II | 130,9 | 15 | 225 | 1963,5 | 96,1 | 136,2 |
| | III | 107,8 | 17 | 289 | 1832,6 | 97,0 | 111,1 |
| | IV | 68,6 | 19 | 361 | 1303,4 | 97,8 | 70,1 |
| Разом | | 1790,7 | × | 2660 | 1169,3 | 1790,6 | × |

Параметри $a_0$ і $a_1$ для рівняння прямої $\hat{y}_t = a_0 + a_1 t$ знаходимо методом найменших квадратів розв'язавши систему нормальних рівнянь:

$$\sum y = na_0 + a_1 \sum t;$$

$$\sum yt = a_0 \sum t + a_1 \sum t^2;$$

$$a_0 = \frac{\sum y}{n} = \frac{1790,7}{20} = 89,535; \quad a_1 = \frac{\sum yt}{\sum t^2} = \frac{1169,3}{2660,0} = 0,439,$$

Звідси $\hat{y}_t = 89,535 + 0,439\ t$.

Підставляючи в дане рівняння послідовно значення (t) матимемо вирівняний ряд:

$$\hat{y}_{t_I} = 89,535 + 0,439 \cdot (-19) = 81,2;$$

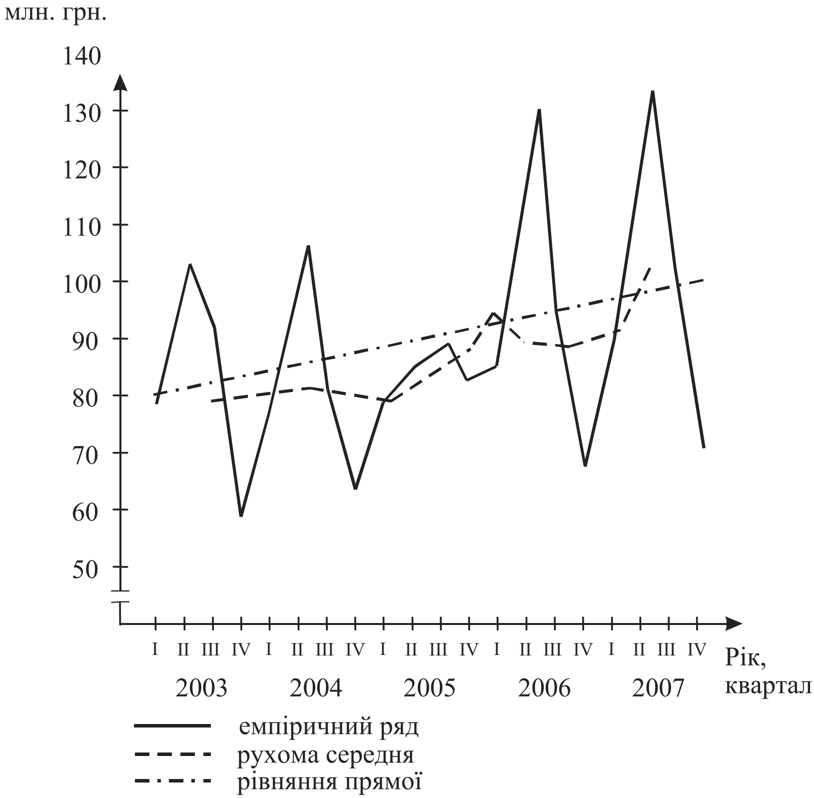
$$\hat{y}_{t_{II}} = 89,535 + 0,439 \cdot (-17) = 82,1; \text{ і т.д.}$$

Сезонну хвилю (індекси сезонності) реалізації мотоциклів і велосипедів визначаємо як процентне відношення емпіричних рівнів до теоретичних:

$$I_{s_I} = \frac{y_I}{\hat{y}_{t_I}} \cdot 100 = \frac{79,4}{81,2} \cdot 100 = 97,8\ \%;$$

$$I_{s_{II}} = \frac{y_{II}}{\hat{y}_{t_{II}}} \cdot 100 = \frac{102,4}{82,1} \cdot 100 = 124,2\ \%; \text{ і т.д.}$$

Результати згладжування внутрішньорічних коливань за методом рухомої середньої і визначення загальної тенденції реалізації мотоциклів і велосипедів населенню України за рівнянням прямої наочно показані на графіку (мал. 11.8).




Мал. 11.8. Згладжування внутрішньорічних коливань реалізованого попиту населення України на мотоцикли і велосипеди методом чотирьохчленної рухомої середньої і рівнянням прямої.

Судячи з графіка, пряма лінія цілком об’єктивно відображає тенденцію розвитку досліджуваного явища.

Моделювання сезонних коливань різних явищ можна проводити і з допомогою *ряду Фур’є*, аналітичний вираз якого стосовно динаміки має наступний вигляд:

$$\hat{y}_t = a_0 + a_1 \cos kt + a_2 \sin kt \ .$$

Таке моделювання сезонної хвилі покажемо на прикладі реалізації холодильників побутових, скориставшись попередньо проведеними розрахунками в табл. 11.20.

Визначимо за даними цього прикладу показник сили коливання динамічного ряду із-за сезонного характеру реалізації холодильників за формулою середнього квадратичного відхилення індексів сезонності:

$$\sigma = \sqrt{\frac{\sum(i-100)^2}{4}}\ .$$

Таблиця 11.27

Розрахунок сезонної хвилі реалізації холодильників побутових за допомогою ряду Фур'є ( $\hat{y}_t = 2668{,}25 + 64{,}78\cos t - 381{,}41\sin t$ ) і середнього квадратичного відхилення індексів сезонності.

| Рік, квартал | | Реалізація холодильників побутових, шт. (y) | Вирівняний ряд динаміки ( $\hat{y}_t$ ) | Сезонна хвиля $I_s = \frac{y}{\hat{y}_t}\cdot 100$ | i – 100 | (i - 100)$^2$ |
|---|---|---|---|---|---|---|
| 2005 | I | 1942 | 2733,03 | 71,0 | -29,0 | 841,00 |
| | II | 2957 | 2533,65 | 116,7 | 16,7 | 278,89 |
| | III | 2504 | 2370,34 | 105,6 | 5,6 | 31,36 |
| | IV | 2194 | 2286,84 | 95,9 | -4,1 | 16,81 |
| 2006 | I | 2126 | 2305,56 | 92,2 | -7,8 | 60,84 |
| | II | 2704 | 2421,45 | 111,7 | 11,7 | 136,89 |
| | III | 3291 | 2603,47 | 126,4 | 26,4 | 696,96 |
| | IV | 1745 | 2802,85 | 62,2 | -37,8 | 1428,84 |
| 2007 | I | 2505 | 2966,16 | 84,4 | -15,6 | 243,36 |
| | II | 3704 | 3049,66 | 121,4 | 21,4 | 457,96 |
| | III | 3834 | 3030,94 | 126,5 | 26,5 | 702,25 |
| | IV | 2513 | 2915,05 | 86,2 | -13,8 | 190,44 |
| Разом | | 32019 | 32019,00 | × | × | 5085,60 |

Середнє квадратичне відхилення для індексів сезонності вирахуване за допомогою рівнянь ряду Фур'є дорівнює:

$$\sigma = \sqrt{\frac{\sum(i-100)^2}{4}} = \sqrt{\frac{5085{,}6}{4}} = \sqrt{1271{,}4} = 35{,}6\ \%.$$

Даний показник свідчить про достатньо великий вплив на реалізацію холодильників побутових сезонного фактора.

Як показують вище наведені дані, розраховані індекси сезонності реалізованого попиту на товари культурно-побутового призначення достатньо надійні і точні.

Це дозволяє використати їх для екстраполяції реалізованого попиту на дані товари культурно-побутового призначення.

Для розрахунку показників реалізованого попиту на перспективу використовують модель прогнозу наступного виду:

$$y_t = I_k \cdot \hat{y}_t + \varepsilon_t ,$$

де $y_t$ – розмір реалізованого попиту в момент часу t;

$I_k$ – середній індекс сезонності k-того кварталу;

$\hat{y}_t$ – оцінка величини товарообороту в момент часу T;

$\varepsilon_t$ – випадковий компонент.

В якості прикладу покажемо прогнозування реалізації холодильників державною і кооперативною торгівлею однієї з областей України (розрахункова табл. 11.28, дані умовні).

Таблиця 11.28

Розрахунок сезонної хвилі реалізації холодильників державною і кооперативною торгівлею області за 2005-2007 рр., після виключення лінійного тренду $\hat{y}_t = 4041{,}5 + 51{,}91\,t$ .

| Рік, квартал | | Реалізація холодильників, шт. ($y_t$) | t | $t^2$ | yt | $\hat{y}_t$ | Сезонна хвиля $I_s = \frac{y}{\hat{y}_t} \cdot 100$ | $\varepsilon_t = y_t - \hat{y}_t$ | $\varepsilon^2 = (y_t - \hat{y}_t)^2$ |
|---|---|---|---|---|---|---|---|---|---|
| А | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 2005 | I | 2790 | -11 | 121 | -30690 | 3470,49 | 80,4 | -680,49 | 463066,6 |
| | II | 4584 | -9 | 81 | -41256 | 3574,31 | 128,2 | 1009,69 | 1019473,9 |
| | III | 4356 | -7 | 49 | -30492 | 3678,13 | 118,4 | 677,87 | 459507,7 |
| | IV | 3083 | -5 | 25 | -15425 | 3781,95 | 81,5 | -698,95 | 488531,1 |
| 2006 | I | 3052 | -3 | 9 | -9156 | 3885,77 | 78,5 | -833,77 | 695172,4 |
| | II | 4271 | -1 | 1 | -4271 | 3989,59 | 107,0 | 281,41 | 79191,6 |
| | III | 5086 | 1 | 1 | 5086 | 4093,41 | 124,2 | 992,59 | 985234,9 |
| | IV | 2789 | 3 | 9 | 8367 | 4197,23 | 66,4 | -1408,23 | 1983111,7 |
| 2007 | I | 3714 | 5 | 25 | 18570 | 4301,05 | 86,4 | -587,05 | 344627,7 |
| | II | 5495 | 7 | 49 | 38465 | 4404,87 | 124,7 | 1090,13 | 1188383,4 |
| | III | 5781 | 9 | 81 | 52029 | 4508,69 | 128,2 | 1272,31 | 1618772,7 |
| | IV | 3497 | 11 | 121 | 38467 | 4612,51 | 75,8 | -1115,51 | 1244362,6 |

Випадкову величину εt практично визначити дуже важко, а тому можна лише з певною ймовірністю стверджувати, що вирахувані за даними моделями показники прогнозу відрізняються від істинної реалізації холодильників на величину:

$$t \cdot \frac{\sigma_{\varepsilon_t}}{\sqrt{n}} ,$$

де t – число, яке гарантує з певною ймовірністю межі прогнозу;

$\sigma_{\varepsilon_t}$ – середньоквадратичне відхилення випадкового компоненту.

Величину $t \cdot \frac{\sigma_{\varepsilon_t}}{\sqrt{n}}$ для кожного кварталу визначаємо з ймовірністю 0,9545 і t = 2 (табл. 11.29).

Таблиця 11.29

Розрахунок величини $t \cdot \frac{\sigma_{\varepsilon_t}}{\sqrt{n}}$

| Квартал | Виправлена сезонна хвиля, % | Виправлений індекс сезонності $I_k$ | $\varepsilon_t^2 = \left(y_t - \hat{y}_t\right)^2$ | $\sigma_{\varepsilon_t} = \sqrt{\frac{\sum \varepsilon_t^2}{n-1}}$ | $t \cdot \frac{\sigma_{\varepsilon_t}}{\sqrt{n}}$ |
|---|---|---|---|---|---|
| I | 81,8 | 0,818 | 1502866,7 | 867,0 | 1002 |
| II | 120,0 | 1,200 | 2287048,9 | 1069,0 | 1236 |
| III | 123,6 | 1,236 | 3063515,3 | 1238,0 | 1431 |
| IV | 74,6 | 0,746 | 3716005,4 | 1363,0 | 1576 |

$$\sigma_{\varepsilon_t} = \sqrt{\frac{\sum \varepsilon_I^2}{n-1}} = \sqrt{\frac{1502866{,}7}{3-1}} = 867\,,\ t \cdot \frac{\sigma_{\varepsilon_I}}{\sqrt{n}} = 2 \cdot \frac{867}{1{,}73} = 1002\text{ ; і т.д.}$$

Визначивши значення $I_k$ і $\hat{y}_t$ отримаємо по кварталах:

$$\hat{y}_{t_I} = 0{,}818 \cdot \left(4041{,}5 + 51{,}91\,t\right) + \varepsilon_t\,;$$

$$\hat{y}_{t_{II}} = 1{,}200 \cdot \left(4041{,}5 + 51{,}91\,t\right) + \varepsilon_t\,;$$

$$\hat{y}_{t_{III}} = 1{,}236 \cdot \left(4041{,}5 + 51{,}91\,t\right) + \varepsilon_t\,;$$

$$\hat{y}_{t_{IV}} = 0{,}746 \cdot \left(4041{,}5 + 51{,}91\,t\right) + \varepsilon_t\,.$$

Вирахуємо прогноз (його верхні та нижні межі) реалізації холодильників до 2010 р.

Прогноз на 2008 р. по кварталах:

$$I_I = 0{,}818\ [4041 + 51{,}91 \cdot (+\ 13)] = 3858 \text{ шт.}$$

Знаходимо межі прогнозу за формулою:

$$\hat{y}_{t+k} - t \cdot \frac{\sigma_{\varepsilon_t}}{\sqrt{n}} \le y_{i+k} \le \hat{y}_{t+k} + t \cdot \frac{\sigma_{\varepsilon_t}}{\sqrt{n}}\,;$$

$$3850 - 1002 \le y_I \le 3850 + 1002\,;$$

$$2856 \le y_I \le 4860\,;$$

$$y_{II} = 1{,}200[4041 + 51{,}91 \cdot (+15)] = 5784 \text{ шт.};$$

$$5784 - 1236 \le y_{II} \le 5784 + 1236\,;$$

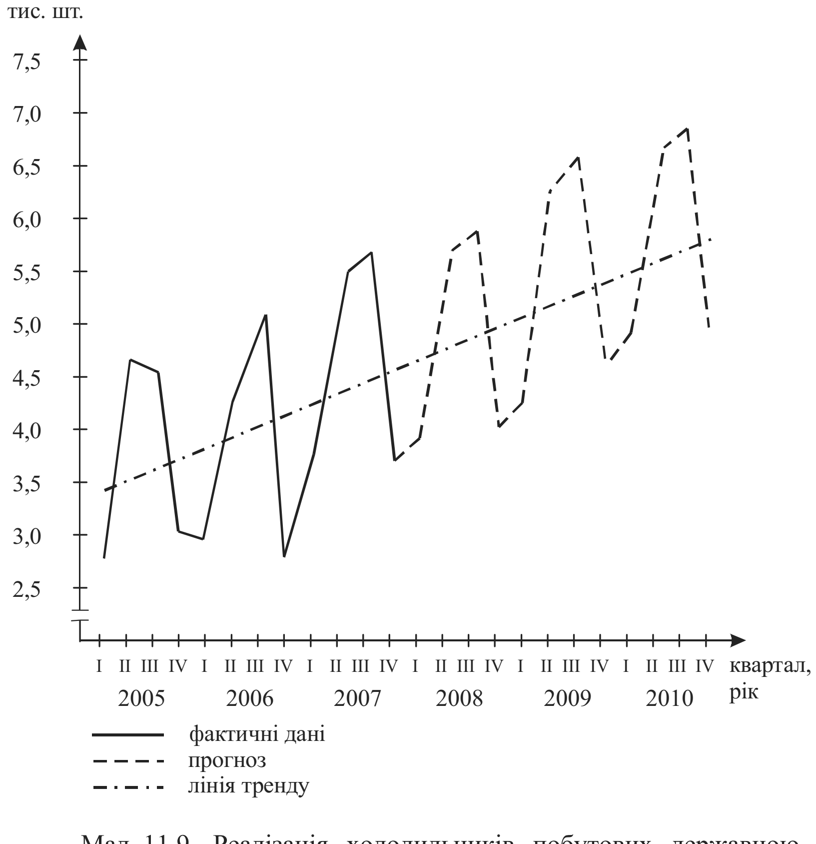
$$4548 \le y_{II} \le 7020\,;$$

і т.д.

Таблиця 11.30

Прогноз реалізації холодильників державною і кооперативною торгівлею області на 2008-2010 рр., в шт.

| Квартал | 2008 р. | | | 2009 р. | | | 2010 р. | | |
|---|---|---|---|---|---|---|---|---|---|
| | прогноз | нижня межа | верхня межа | прогноз | нижня межа | верхня межа | прогноз | нижня межа | верхня межа |
| I | 3858 | 2856 | 4860 | 4198 | 3196 | 5200 | 4622 | 3620 | 5624 |
| II | 5784 | 4548 | 7020 | 6282 | 5046 | 7518 | 6905 | 5669 | 8141 |
| III | 6086 | 4655 | 7517 | 6728 | 5297 | 8159 | 7241 | 5810 | 8672 |
| IV | 3751 | 2175 | 5327 | 4138 | 2562 | 5714 | 4448 | 2872 | 6024 |




Мал. 11.9. Реалізація холодильників побутових державною і кооперативною торгівлею області за 2005-2007 рр. (2008-2010 рр. прогноз).

За підсумками розрахунків, можна зробити висновок про те, що незалежно від методу вирівнювання емпіричних даних і способу визначення сезонної хвилі "піки" і "ями" реалізації однойменних товарів однаково розподіляються по кварталах, а квартальні індекси сезонності близькі за величиною.

Таким чином, внутрішньорічні коливання більшості соціально-економічних явищ можна визначати любим із наведених методів вирівнювання емпіричних даних і любим із способів вирахування сезонної хвилі. Тому з метою поточного аналізу сезонності будь-якого явища цілком прийнятне використання менш трудоємкого методу простої середньої.

## 11.6 Особливості вимірювання взаємозв'язків в рядах динаміки

При вивченні кореляційних зв'язків в багатомірних рядах динаміки спостерігається певна залежність рівнів даного періоду від попереднього, внаслідок чого виникають певні методологічні особливості. В таких динамічних рядах фактором зміни рівнів виступає, крім інших, також час.

Вплив даного рівня динамічного ряду на зміну наступного з плином часу приводить до так званої *автокореляції*. Тому кореляційно-регресійний метод правильно покаже зв'язок між явищами динамічних рядів лише в тому випадку, якщо в кожному з цих рядів відсутня автокореляція.

В практиці економіко-статистичного аналізу рядів динаміки застосовують різні способи усунення автокореляції, такі як спосіб різницевих перетворень (при лінійному тренді), спосіб відхилень тенденції (при нелінійній залежності), або введення змінної величини t в рівняння регресії $\hat{y}_t = f\left(x_1, x_2, x_3, \ldots, t\right)$, де вона відіграє роль фактора часу.

При застосуванні методу *регресії* для дослідження динамічних рядів виникає особливість, яка заключається в тому, що в рівнях динамічних рядів присутня *авторегресія*, яка проявляється так же, як і автокореляція.

Авторегресія виражає залежність величини рівня динамічного ряду від попередніх значень рівня в певні моменти часу.

Методику побудови рівня регресії з введенням фактора часу розглянемо на прикладі двох вхаємозв'язаних рядів динаміки: глибини зрошення багаторічних трав з добавленням органічних компонентів $y_t$, та урожайності насіння цих трав $x_t$ (табл. 11.31).

Таблиця 11.31

Розрахункова таблиця для обчислення коефіцієнтів регресії.

| Роки | Глибина зрошення з добавленням органічних компонентів, см. (x) | Урожайність насіння багаторічних трав, ц/га (y) | t | $t^2$ | $x^2$ | xy | xt | yt | $\hat{y}_t$ |
|---|---|---|---|---|---|---|---|---|---|
| 1997 | 12 | 5,6 | -5 | 25 | 144 | 67,2 | -60 | -28,0 | 4,9 |
| 1998 | 8 | 4,0 | -4 | 16 | 64 | 32,0 | -32 | -16,0 | 3,4 |
| 1999 | 10 | 4,0 | -3 | 9 | 100 | 40,0 | -30 | -12,0 | 4,3 |
| 2000 | 6 | 2,4 | -2 | 4 | 36 | 14,0 | -12 | -4,8 | 2,8 |
| 2001 | 9 | 3,6 | -1 | 1 | 81 | 32,4 | -9 | -3,6 | 4,0 |
| 2002 | 15 | 5,0 | 0 | 0 | 225 | 75,0 | 0 | 0 | 6,4 |
| 2003 | 11 | 4,6 | 1 | 1 | 121 | 50,6 | 11 | 4,6 | 4,9 |
| 2004 | 13 | 6,5 | 2 | 4 | 169 | 84,5 | 26 | 13,0 | 5,8 |
| 2005 | 14 | 7,0 | 3 | 9 | 196 | 98,0 | 42 | 21,0 | 6,3 |
| 2006 | 10 | 4,5 | 4 | 16 | 100 | 45,0 | 40 | 18,0 | 4,8 |
| 2007 | 12 | 6,0 | 5 | 25 | 144 | 72,0 | 60 | 30,0 | 5,6 |
| Разом | 120 | 53,2 | 0 | 110 | 1380 | 611,1 | 36 | 22,2 | 53,2 |

Зв’язок між глибиною зрошення багаторічних трав з добавлення органічних компонентів та врожайності насіння цих трав можна відобразити лінійною функцією:

$$\hat{y}_t = a_0 + a_1 x + a_2 t \,,$$

де $a_1$ – параметр, який характеризує середній приріст результативної ознаки на одиницю приросту факторної ознаки (x);

$a_2$ – середній щорічний приріст (y) під впливом зміни комплексу факторів крім (x);

t – час (роки).

Параметри цього рівняння визначаються методом найменших квадратів, склавши і розв’язавши систему нормальних рівнянь з трьома невідомими:

$$\begin{cases} \sum y = na_0 + a_1\sum x + a_2\sum t; \\ \sum xy = a_0\sum x + a_1\sum x^2 + a_2\sum xt; \\ \sum yt = a_0\sum t + a_1\sum xt + a_2\sum t^2. \end{cases}$$

Якщо добитись що $\sum t = 0$, то за даними табл. 11.31 система нормальних рівнянь буде мати вигляд:

$$\begin{cases} 53{,}2 = 11\,a_0 + 120\,a_1; \\ 611{,}1 = 120\,a_0 + 1380\,a_1 + 36\,a_2; \\ 22{,}2 = 36\,a_1 + 110\,a_2. \end{cases}$$

Розв’язавши цю систему, отримаємо наступні значення параметрів:

$$a_0 = 0{,}549; \qquad a_1 = 0{,}393; \qquad a_2 = 0{,}073.$$

Лінійне рівняння зв’язку буде мати вигляд:

$$\hat{y}_t = 0{,}549 + 0{,}393\,x + 0{,}073\,t\ .$$

Параметри рівняння регресії потрібно тлумачити так: якщо при інших рівних умовах глибина зрошення багаторічних трав з добавленням органічних компонентів збільшиться на 1 см., то врожайність насіння цих трав зросте на 0,393 ц/га. За рахунок впливу інших факторів, які рівномірно змінюються протягом часу, урожайність насіння багаторічних трав щорічно зростатиме в середньому на 0,073 ц/га.

Підставляючи в отримане рівняння регресії значення t і x визначимо теоретичні рівні врожайності багаторічних трав (остання колонка табл. 11.31):

$$\hat{y}_{1997\,р.} = 0,549 + 0,393 \cdot 12 + 0,073 \cdot (-5) = 4,9\ ;$$

$$\hat{y}_{1998\,р.} = 0,549 + 0,393 \cdot 8 + 0,073 \cdot (-4) = 3,4\ ;$$

$$\hat{y}_{1999\,р.} = 0,549 + 0,393 \cdot 10 + 0,073 \cdot (-3) = 4,3\ ;$$

і т.д.

Приведене рівняння регресії повинно виключити авторегресію. Для переконання в цьому знайдемо автокореляцію різниць між фактичними даними і вирівняними даними за цим рівнянням, тобто кореляцію величин $y - \hat{y}_t = \varepsilon_t$ .

Коефіцієнт автокореляції відхилень приймає значення в межах від –1 до +1 і визначається за формулою:

$$r_a = \frac{\sum \varepsilon_t \cdot \varepsilon_{t+1}}{\sum \varepsilon_t^2}\ .$$

Таблиця 11.32

Розрахунок коефіцієнта автокореляції.

| Роки | $y$ | $\hat{y}_t$ | $\varepsilon_t = (y - \hat{y}_t)$ | $\varepsilon_{t+1}$ | $\varepsilon_t \cdot \varepsilon_{t+1}$ | $\varepsilon_t^2 = (y - \hat{y}_t)^2$ |
|---|---|---|---|---|---|---|
| 1997 | 5,6 | 4,9 | 0,7 | 0,6 | 0,42 | 0,49 |
| 1998 | 4,0 | 3,4 | 0,6 | -0,3 | -0,18 | 0,36 |
| 1999 | 4,0 | 4,3 | -0,3 | -0,4 | 0,12 | 0,09 |
| 2000 | 2,4 | 2,8 | -0,4 | -0,4 | 0,16 | 0,16 |
| 2001 | 3,6 | 4,0 | -0,4 | -1,4 | 0,56 | 0,16 |
| 2002 | 5,0 | 6,4 | -1,4 | -0,3 | 0,42 | 1,96 |
| 2003 | 4,6 | 4,9 | -0,3 | 0,7 | -0,21 | 0,09 |
| 2004 | 6,5 | 5,8 | 0,7 | 0,7 | 0,49 | 0,49 |
| 2005 | 7,0 | 6,3 | 0,7 | -0,3 | -0,21 | 0,49 |
| 2006 | 4,5 | 4,8 | -0,3 | 0,4 | -0,12 | 0,09 |
| 2007 | 6,0 | 5,6 | 0,4 | 0,7 | 0,28 | 0,16 |
| Разом | 53,2 | 53,2 | 0 | 0 | 1,73 | 4,54 |

У нашому прикладі коефіцієнт автокореляції з *часовим зсувом – лагом* р = 1 дорівнює:

$$r_a = \frac{\sum \varepsilon_t \cdot \varepsilon_{t+1}}{\sum \varepsilon_t^2} = \frac{1,73}{4,54} = 0,381 .$$

Значення коефіцієнта додатнє $r_a = 0,381$ свідчить про незначний ступінь кореляції залишкових величин.

Висновок щодо наявності або відсутності автокореляції в залишкових величинах роблять порівнюючи фактичне значення $r_a$ з табличним для даного числа спостережень (n) і прийнятого рівня значимості.

Для нашого прикладу: $r_{a_Ф} = 0,381$, при n = 10 і 5 %-ному рівні ймовірності $r_{a_Т} = 0,360$. Тобто $r_{a_Ф} > r_{a_Т}$ на 0,021 пункти, що і засвічує про незначну автокореляцію.

В багатьох економіко-статистичних дослідженнях доводиться вивчати паралельно декілька динамічних рядів, в яких коливання рівнів взаємообумовлені, наприклад, динаміка цін на які-небудь овочі на ринку в значній мірі зв’язана з їх урожайністю; в свою чергу динаміка урожайності або валовий збір залежать від динаміки кількості опадів, агрохімобробітку; попит населення на певні товари народного споживання залежить від пропозиції, тобто від об’єму їх виробництва і т.д.

Для вимірювання залежності між такими рядами динаміки використовують методи кореляції, тобто розраховують різні коефіцієнти кореляції.

Розглянемо приклад. Нехай потрібно виміряти силу зв'язку між рядами динаміки, застосувавши метод корелювання рівнів умовних даних про зміну вартості основних виробничих фондів і випуском продукції за 1998-2007 рр.

З цією метою скористаємось лінійним коефіцієнтом кореляції для розрахунку якого побудуємо табл. 11.33.

Таблиця 11.33

| Роки | Вартість основних виробничих фондів, млн. грн. (x) | Випуск продукції, млн. грн. (y) | $x^2$ | xy | $y^2$ |
|---|---|---|---|---|---|
| 1998 | 5,3 | 5,8 | 28,09 | 30,74 | 33,64 |
| 1999 | 6,4 | 7,0 | 40,96 | 44,80 | 49,00 |
| 2000 | 7,9 | 8,7 | 62,41 | 68,73 | 75,69 |
| 2001 | 8,3 | 9,1 | 68,89 | 75,53 | 82,81 |
| 2002 | 9,2 | 10,9 | 84,64 | 100,28 | 118,81 |
| 2003 | 10,1 | 11,3 | 102,01 | 114,13 | 127,69 |
| 2004 | 12,5 | 13,8 | 156,25 | 172,50 | 190,44 |
| 2005 | 13,0 | 14,0 | 169,00 | 182,00 | 196,00 |
| 2006 | 14,6 | 15,9 | 213,16 | 232,14 | 252,81 |
| 2007 | 15,7 | 18,8 | 246,49 | 295,16 | 353,44 |
| Разом | 103,0 | 115,3 | 1171,90 | 1316,01 | 1480,33 |
| В середньому | 10,3 | 11,53 | 117,19 | 131,601 | 148,033 |

Визначаємо середні квадратичні відхилення для обох рівнів:

$$\sigma_x = \sqrt{\overline{x^2} - (\overline{x})^2} = \sqrt{117{,}19 - 10{,}3^2} = \sqrt{11{,}1} = 3{,}332 ;$$

$$\sigma_y = \sqrt{\overline{y^2} - (\overline{y})^2} = \sqrt{148{,}033 - 11{,}53^2} = \sqrt{15{,}092} = 3{,}885 .$$

Звідси лінійний коефіцієнт кореляції дорівнює:

$$r = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x \cdot \sigma_y} = \frac{131{,}601 - 10{,}3 \cdot 11{,}53}{3{,}332 \cdot 3{,}885} = 0{,}992 .$$

Коефіцієнт кореляції між рівнями цих динамічних рядів показує прямий і дуже сильний зв'язок. Однак корельовано дані розвиваються в часі, а тому, перш ніж робити висновок про тісноту зв'язку між досліджуваними явищами, потрібно перевірити обидва ряди на автокореляцію. Наявність автокореляції встановлюється за допомогою коефіцієнта кореляції.

Для розрахунку коефіцієнта кореляції по першому ряду (x – вартість основних виробничих фондів) побудуємо табл. 11.34.

Таблиця 11.34

| Роки | Вартість основних виробничих фондів, млн. грн. ($x_t$) | $x_{t+1}$ | $x_t \cdot x_{t+1}$ | $x_t^2$ |
|---|---|---|---|---|
| 1998 | 5,3 | 6,4 | 33,92 | 28,09 |
| 1999 | 6,4 | 7,9 | 50,56 | 40,96 |
| 2000 | 7,9 | 8,3 | 65,57 | 62,41 |
| 2001 | 8,3 | 9,2 | 76,36 | 68,89 |
| 2002 | 9,2 | 10,1 | 92,92 | 84,64 |
| 2003 | 10,1 | 12,5 | 126,25 | 102,01 |
| 2004 | 12,5 | 13,0 | 162,50 | 165,75 |
| 2005 | 13,0 | 14,6 | 189,80 | 169,00 |
| 2006 | 14,6 | 15,7 | 229,22 | 213,16 |
| 2007 | 15,7 | 5,3 | 83,21 | 246,49 |
| Разом | 103,0 | 103,0 | 1110,31 | 1171,90 |

Вирахуємо за підсумковими даними таблиці необхідні величини для знаходження коефіцієнта автокореляції ( $r_{a_1}$ ):

$$\overline{x_t \cdot x_{t+1}} = \frac{\sum x_t \cdot x_{t+1}}{n} = \frac{1110{,}31}{10} = 111{,}031\,;$$

$$\overline{x_t} = \frac{\sum x_t}{n} = \frac{103}{10} = 1{,}03\,;$$

$$\sigma^2_{x_t} = \frac{\sum x_t^2}{n} - \overline{x}_t^2 = \frac{1171{,}9}{10} - 10{,}3^2 = 117{,}19 - 106{,}09 = 14{,}1\,;$$

$$r_{a_1} = \frac{\overline{x_t \cdot x_{t+1}} - (\overline{x}_t)^2}{\sigma^2_{x_t}} = \frac{111{,}031 - 106{,}09}{14{,}1} = \frac{5{,}22}{14{,}1} = 0{,}370\,.$$

Потім проводимо аналогічні вирахування для другого ряду (y – випуск продукції) і отримуємо $r_{a_2} = 0{,}353$ .

Порівнюємо отримані коефіцієнти автокореляції з їх табличною величиною при чисельності спостережень n = 10, і рівні значимості p = 0,05 (5 %-ний рівень). Перший коефіцієнт автокореляції перевищив табличне значення $r_{a_T} = 0{,}360$, а другий – близький до нього, тому можна зробити висновок про існування автокореляції.

Отже, для того, щоб отримати правильне значення зв'язку між рядами динаміки, потрібно усунути певну тенденцію розвитку, а вже потім розраховувати коефіцієнт кореляції.

Одним із способів усунення автокореляції є *корелювання відхилень фактичних рівнів від вирівняних*, які відображають тренд. Для цього потрібно: а) провести аналітичне вирівнювання порівнювальних рядів; б) визначити величини відхилень кожного фактичного рівня динаміки від їх вирівняних значень; в) провести корелювання отриманих відхилень.

Розрахуємо за даними попереднього прикладу лінійні тренди для обох рядів:

$$\hat{x}_t = a_0 + a_1 t = 10{,}3 + 0{,}575\, t\ ;$$

$$\hat{y}_t = 11{,}53 + 0{,}6676\, t\ .$$

Фактичні рівні рядів динаміки і їх тренди наведені в табл. 11.35.

Таблиця 11.35

Розрахункова таблиця для визначення лінійного коефіцієнта кореляції.

| Роки | Вартість основних виробничих фондів, млн. грн. (x) | Випуск продукції, млн. грн. (y) | Вирівняні значення | | $d_x = x - \hat{x}_t$ | $d_y = y - \hat{y}_t$ | $d_x^2$ | $d_y^2$ | $d_x \cdot d_y$ |
|---|---|---|---|---|---|---|---|---|---|
| | | | $\hat{x}_t$ | $\hat{y}_t$ | | | | | |
| 1998 | 5,3 | 5,8 | 5,1 | 5,5 | 0,2 | 0,3 | 0,04 | 0,09 | 0,06 |
| 1999 | 6,4 | 7,0 | 6,3 | 6,9 | 0,1 | 0,1 | 0,01 | 0,01 | 0,01 |
| 2000 | 7,9 | 8,7 | 7,4 | 8,2 | 0,5 | 0,5 | 0,25 | 0,25 | 0,25 |
| 2001 | 8,3 | 9,1 | 8,6 | 9,5 | -0,3 | -0,4 | 0,09 | 0,16 | 0,12 |
| 2002 | 9,2 | 10,9 | 9,7 | 10,9 | -0,5 | 0 | 0,25 | 0 | 0 |
| 2003 | 10,1 | 11,3 | 10,9 | 12,2 | -0,8 | -0,9 | 0,64 | 0,81 | 0,72 |
| 2004 | 12,5 | 13,8 | 12,0 | 13,5 | 0,5 | 0,3 | 0,25 | 0,09 | 0,15 |
| 2005 | 13,0 | 14,0 | 13,2 | 14,9 | -0,2 | -0,9 | 0,04 | 0,81 | 0,18 |
| 2006 | 14,6 | 15,9 | 14,3 | 16,2 | 0,3 | -0,3 | 0,09 | 0,09 | -0,09 |
| 2007 | 15,7 | 18,8 | 15,5 | 17,5 | 0,2 | 1,3 | 0,04 | 1,69 | 0,26 |
| Разом | 103,0 | 115,3 | 103,0 | 115,3 | 0 | 0 | 1,70 | 4,00 | 1,66 |

Коефіцієнт кореляції відхилень розраховують за формулою:

$$r = \frac{\sum d_x \cdot d_y}{\sqrt{\sum d_x^2 \cdot \sum d_y^2}} = \frac{1{,}66}{\sqrt{1{,}7 \cdot 4{,}0}} = 0{,}636\ .$$

Отже, автокореляція завищила показник тісноти зв'язку, так як вирахуваний із безпосередніх рівнів ряду він дорівнює 0,992, а між залишковими величинами він значно менший (0,636).

До аналогічних результатів можна прийти, якщо знайти *кореляцію різниць між наступними і попередніми рівнями* обох рядів: $\Delta x = x_i - x_{i-1}$; $\Delta y = y_i - y_{i-1}$. При заміні рівнів динамічних рядів різницями між ними усувається вплив автокореляції в кожному динамічному ряду. Однак, потрібно мати на увазі, що різниці першого порядку виключають автокореляцію в рядах динаміки з прямолінійним трендом. Якщо зміна динамічних рядів в часі проходить по параболі другого порядку, тоді усунення впливу автокореляції проводять за допомогою кореляції других різниць (різниць між першими різницями).

Коефіцієнт кореляції перших різниць визначається за формулою:

$$r = \frac{\sum \Delta x \cdot \Delta y}{\sqrt{\sum \Delta_x^2 \cdot \sum \Delta_y^2}} .$$

Визначення цього коефіцієнта покажемо на даних попереднього прикладу (табл. 11.36).

Таблиця 11.36

| Роки | Вартість основних виробничих фондів, млн. грн. (x) | Випуск продукції, млн. грн. (y) | Різниця між наступними рівнями | | | | |
|---|---|---|---|---|---|---|---|
| | | | $\Delta_x$ | $\Delta_y$ | $\Delta_x^2$ | $\Delta_y^2$ | $\Delta_x \cdot \Delta_y$ |
| 1998 | 5,3 | 5,8 | - | - | - | - | - |
| 1999 | 6,4 | 7,0 | 1,1 | 1,2 | 1,21 | 1,44 | 1,32 |
| 2000 | 7,9 | 8,7 | 1,5 | 1,7 | 2,25 | 2,89 | 2,55 |
| 2001 | 8,3 | 9,1 | 0,4 | 0,4 | 0,16 | 0,16 | 0,16 |
| 2002 | 9,2 | 10,9 | 0,9 | 1,8 | 0,81 | 3,24 | 1,62 |
| 2003 | 10,1 | 11,3 | 0,9 | 0,4 | 0,81 | 0,16 | 0,36 |
| 2004 | 12,5 | 13,8 | 2,4 | 2,5 | 5,76 | 6,25 | 6,00 |
| 2005 | 13,0 | 14,0 | 0,5 | 0,2 | 0,25 | 0,04 | 0,10 |
| 2006 | 14,6 | 15,9 | 1,6 | 1,9 | 2,56 | 3,61 | 3,04 |
| 2007 | 15,7 | 18,8 | 1,1 | 2,9 | 1,21 | 8,41 | 3,19 |
| Разом | × | × | × | × | 15,02 | 26,20 | 18,34 |

Вирахуємо коефіцієнти кореляції різниць:

$$r = \frac{\sum \Delta x \cdot \Delta y}{\sqrt{\sum \Delta_x^2 \cdot \sum \Delta_y^2}} = \frac{18,34}{\sqrt{15,02 \cdot 26,20}} = 0,924\,.$$

Отриманий коефіцієнт кореляції різниць свідчить про прямий і тісний зв'язок між вартістю основних виробничих фондів і випуском продукції.

Досліджуючи кореляцію між рядами динаміки, можна зустріти багато випадків коли зміна рівнів одного ряду викликає зміну рівнів другого ряду тільки через певний проміжок часу (місяць, квартал, рік і т.д.).

Наприклад, зміна виробництва деяких товарів народного споживання, в даному періоді впливає на зміну об'єму товарообігу в майбутньому періоді; зміна числа шлюбів в даному році може вплинути на народжуваність в наступному році і т.д.

Тому, провівши попередній якісний і логічний аналіз з метою оцінки зв'язку між рядами динаміки, потрібно зсунути один ряд відносно другого на визначений проміжок часу (лаг) і *корелювати ряди з лагом*.

Для тривалого періоду залежність між рівнями рядів динаміки може змінюватись в часі. Тому показник тісноти зв'язку в таких динамічних рядах визначають як серію коефіцієнтів кореляції, розрахованих за аналогією рухомої середньої. Це дає можливість виявляти періоди в яких залежність збільшувалась або зменшувалась, і, знаючи такі періоди, легше пояснити зміну цієї залежності в конкретних умовах місця і часу.

Таким чином ми розглянули три основні особливості кореляції динамічних рядів:

- перша особливість заключається в усуненні автокореляції при корелюванні динамічних рядів;
- друга особливість проявляється в можливості корелювання динамічних рядів з часовим лагом;
- третьою особливістю кореляції динамічних рядів є можливість застосування змінної (рухомої) кореляції.

### *Контрольні запитання*

1. Поняття про ряди динаміки.
2. Види рядів динаміки.
3. Правила побудови рядів динаміки.
4. Основні характеристики рядів динаміки.
5. Як визначаються ланцюгові і базисні показники динаміки?

6. Як визначається абсолютний приріст, темп зростання, темп приросту і абсолютне значення одного процента приросту?
7. Коли використовують середню арифметичну просту, а коли зважену при визначенні середнього рівня динамічного ряду?
8. За якою формулою визначається середній рівень в момент них динамічних рядах з рівними інтервалами?
9. Визначення середнього абсолютного приросту, темпу зростання і приросту.
10. Які ви знаєте основні прийоми аналізу рядів динаміки?
11. Виявлення закономірності зміни рівнів ряду способом укрупнення інтервалів.
12. Для чого і як проводять згладжування рядів динаміки за допомогою рухомої середньої?
13. Які ви знаєте методи аналітичного вирівнювання рядів динаміки?
14. Вирівнювання за прямою.
15. Вирівнювання за гіперболою.
16. Вирівнювання за параболою другого порядку.
17. Вирівнювання за показниковою функцією.
18. Вирівнювання за допомогою ряду Фур’є.
19. Поняття про інтерпретацію і екстраполяцію.
20. Поняття про сезонні коливання в рядах динаміки.
21. Як визначається індекс сезонності?
22. Які ви знаєте методи дослідження внутрішньорічних коливань?
23. Визначення сезонної хвилі методом простої середньої.
24. Визначення сезонності методом ланцюгових індексів (методом Персонса).
25. Дослідження сезонності за допомогою рухомої середньої.
26. Використання методів аналітичного вирівнювання для вивчення сезонних коливань.
27. Які ви знаєте особливості вимірювання взаємозв’язків в рядах динаміки?
28. Поняття про автокореляцію і авторегресію.
29. Як визначається коефіцієнт автокореляції?
30. Які ви знаєте методи корелювання рядів динаміки?

# Розділ 12. Індекси

## 12.1. Поняття про індекси, їх види

Для характеристики соціально-економічних явищ і процесів статистика широко використовує узагальнюючі показники у вигляді середніх, відносних величин та коефіцієнтів. Одними з таких узагальнюючих показників і є індекси. В широкому розумінні слово Index в перекладі з латинського означає показник.

*Індексом* в статистиці називається відносний показник, що характеризує зміну рівня соціально-економічного явища в часі, порівняно з планом, базисним періодом або в просторі. Іншими словами індекс – це узагальнюючий показник порівняння двох сукупностей, які складаються з елементів безпосередньо несумарних. Так, наприклад, статистичне управління повідомляє, що продуктивність праці в промисловості підвищилась на 7,0 %, основні виробничі фонди збільшились в 1,3 рази, ціни на товари і послуги знизились на 5 %. Зрозуміло, що безпосередньо додавати всякого роду продукцію, ціни на різні товари і т.п. не можна. А тому для узагальнюючої характеристики зміни саме таких складних явищ використовують індекси.

В статистичних дослідженнях складних соціально-економічних явищ і процесів виділяють три великі сфери застосування економічних індексів.

До *першої сфери* застосування індексів відносять порівняльну характеристику несумарних сукупностей в часі. Сюди входять синтетичні індекси динаміки, виконання плану і територіальні індекси.

*Індекси динаміки* показують зміну якого-небудь складного явища в звітному періоді порівняно з базисним.

Для порівняння фактично досягнутого рівня з плановим завданням використовують *індекс виконання плану.*

Індекси можуть застосовуватись для просторового порівняння рівнів продуктивності площі, урожайності, цін в різних регіонах, в таких випадках їх називають *територіальними індексами*.

Для всіх цих індексів основною теоретичною проблемою є розробка принципів і методів їх обчислення як зведених синтетичних показників.

*Друга сфера* застосування індексів заключається в їх використанні для факторного аналізу складного явища через систему взаємозв'язаних індексів. До таких складних явищ можуть бути віднесені вартість виробленої чи реалізованої продукції, фонд заробітної плати, валовий збір зерна та ін.

Так, вартість виробленої продукції дорівнює добутку цін на кількість продукції, валовий збір зерна – добутку урожайності на посівну площу, фонд заробітної плати – добутку заробітної плати одного працівника на їх чисельність і т.д. Аналогічним чином зв'язані і індекси. Якщо індекс цін помножити на індекс фізичного об'єму отримаємо індекс вартості (товарообороту) продукції.

З допомогою індексних систем можна виміряти, яку роль в динаміці складного явища складають інтенсивний і екстенсивний фактори. На цій же основі визначають абсолютні зміни важливих елементів економіки, зумовлені дією інтенсивних і екстенсивних факторів розвитку.

З допомогою *третьої сфери*, через застосування економічних індексних систем, проводять аналіз динаміки середніх показників, зміна яких піддається впливу структурних зрушень у середині досліджуваної сукупності.

Динамічні процеси, як правило, в більшості випадків супроводжуються зміною в структурі досліджуваної сукупності. Змінюється частка окремих груп з різними рівнями варіаційних ознак в середині сукупності. А тому, структурні зрушення впливають на динаміку середніх показників. Вони свідчать про якісні зміни в розвитку повної сукупності, яка вивчається. В зв'язку з цим, велике значення має визначення впливу структурних зрушень на динаміку середніх показників через застосування системи взаємопов'язаних індексів змінного складу, постійного (фіксованого) складу і структурних зрушень.

Всі економічні індекси статистика класифікує за трьома основними ознаками:

а) за характером досліджуваних об'єктів;
б) за ступенем охоплення елементів сукупності;
в) за методикою розрахунку загальних індексів.

За *характером досліджуваних об'єктів* індекси ділять на індекси об'ємних (кількісних) і якісних показників.

До першої групи відносяться індекси фізичного обсягу продукції промисловості, сільського господарства, будівництва, роздрібного товарообороту, національного доходу, споживання і ін.

До другої групи якісних показників відносяться індекси цін, собівартості, урожайності, продуктивності праці і деякі інші індекси.

За *ступенем охоплення* елементів сукупності індекси ділять на:

а) індивідуальні;
б) загальні;
в) групові.

*Індивідуальні індекси* характеризують зміну окремих елементів складного явища. Так, наприклад, кількість реалізованої картоплі на ринку в 2006 р. склала 53000 кг., по ціні 25 коп. за один кілограм, в 2007. – 59360 кг., по 20 коп. за кілограм. Отже, реалізація картоплі в 2007 р. порівняно з 2006 р. збільшилась в 1,12 рази (59360:53000), або на 12 % (112-100), а ціни на картоплю за цей же період змінились на 20 %

$$\left[\left(\frac{20}{25}\cdot 100\right)-100\right].$$

В теорії індексів показник, зміну якого характеризує індекс, називається *індексованою* величиною.

Для обчислення індексів, що характеризують зміни явищ в часі, необхідно мати дані не менше ніж за два періоди. Період який порівнюють називається звітним, а період з яким порівнюють – називається базисним.

Біля індексованої величини звітного періоду внизу ставиться "1", а біля індексованої величини базисного періоду – "0".

Індекс, обчислений як відношення даних звітного періоду до даних базисного періоду, може бути виражений у вигляді коефіцієнтів (коли базисний рівень приймають за 1) або у відсотках (якщо базу порівняння беруть за 100). Якщо індекс більший за 1 або 100%, рівень досліджуваного явища зростає, а якщо менший 1 або 100% - знижується.

Індивідуальні індекси позначають малою латинською буквою "і", продукцію в натуральному виразі – через "q", ціну одиниці товару – через "р", собівартість одиниці продукції – через "z". Індивідуальні індекси цих ознак мають вигляд:

а) фізичного обсягу

$$i_q = \frac{q_1}{q_0};$$

б) ціни одиниці товару

$$i_p = \frac{p_1}{p_0};$$

в) собівартості одиниці продукції

$$i_z = \frac{z_1}{z_0};$$

де $i_q, i_p, i_z$ – індивідуальні індекси фізичного обсягу, ціни і собівартості одиниці продукції;

$q_1, q_0, p_1, p_0, z_1, z_0$ – фізичний обсяг, ціна, собівартість в звітному і базисному періодах.

*Загальні індекси* характеризують зміну сукупності в цілому і являють собою відносні числа, що визначають зміни в часі порівняно з плановим, базисним періодами або в просторі складного явища, яке складається з несумарних окремих елементів. Так, наприклад, виробництво продукції промисловості в області (у порівняльних цінах) в 2007 році порівняно з 2006 р. збільшилось 1,3 рази, тобто склало 130 %, виробництво товарів народного споживання (у відпускних цінах) за цей же період склало 141%, в т.ч.: продовольчих – 171 %; алкогольних напоїв – 89%; непродовольчих товарів – 102 %. Показники 130 % (1,3), 141 % (1,41), 89 % (0,89), 102 % (1,02) являються загальними індексами продукції промисловості, а також наведених товарів народного споживання.

*Груповими або субіндексами* називаються такі індекси, які охоплюють не всі елементи сукупності, а тільки яку-небудь частку або їх групу. Так, індекс виробництва всіх товарів народного споживання в 2004 р. в порівняні з 2003 р. склав 1,41, а індекси по окремих групах товарів відповідно склали: продовольчих – 1,71, алкогольних напоях – 0,89 і непродовольчих – 1,02. Дані індекси будуть груповими по відношенню до всіх товарів народного споживання.

Групові індекси мають велике значення, тому що вони розкривають, за допомогою методу групувань, закономірності в розвитку окремих частин досліджуваного явища. В них поєднується безпосередній зв’язок індексів з методом групувань.

В залежності від методології обчислення, загальні і групові індекси діляться на агрегатні і середні з індивідуальних індексів.

Агрегатні індекси є основною формою економічних індексів, а середні з індивідуальних індексів – похідними, отриманими в результаті перетворення агрегатних індексів.

*Базисні і ланцюгові індекси* обчислюють в тих випадках, коли доводиться вивчати яке-небудь явище суспільного життя за ряд послідовних років (трирічча, п’ятиріччя і т.д.).

Економічні індекси поєднують в собі властивості середніх і відносних величин. Так, для обчислення середньої арифметичної величини використовують підсумовування, а відносні величини знаходять шляхом ділення. Статистичні індекси синтезують обидві ці операції.

Всі види відносних величин, крім структури і інтенсивності, є індексами. Отже, любий індекс – це відносна величина, але не будь-яка відносна величина може бути індексом.

## 12.2. Агрегатні індекси як вихідна форма індексів

Агрегатний індекс являється основною формою економічного індекса. Його назва пішла від латинського слова *aggrego* – приєдную. Чисельник і знаменник цього індекса являє собою агрегат, набір різнорідних елементів.

Отже, *агрегатним індексом* в статистиці називається загальний індекс, який є відношенням сум добутків індексованих (зіставлюваних) величин порівнюваних періодів на ваги (співвимірники, за допомогою яких сумуються різнорідні елементи).

Основна трудність побудови формули для обчислення загального індексу полягає в знаходженні таких ваг, за допомогою яких можна перейти від сукупності безпосередньо несумарних елементів, до іншої сукупності, елементи якої можна підсумовувати.

При побудові формул агрегатних індексів використовують наступне правило: "*якщо індексована величина – якісний показник, який знаходять шляхом ділення (ціна, собівартість, урожайність і т.д.) ваги беруться звітного періоду, а якщо індексована величина – кількісний показник, який можна підсумовувати (фізичний об'єм продукції, чисельність працівників, посівна площа) ваги беруться базисного періоду*".

Покажемо застосування цього правила при побудові формул агрегатних індексів.

*Загальний індекс цін* визначається за формулою:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1},$$

де $I_p$ – загальний індекс цін;
$p$ – індексована величина;
$q$ – вага.

Цей індекс показує як змінились ціни на всі досліджувані товари в звітному періоді порівняно з базисним.

*Загальний індекс фізичного обсягу* продукції визначається за такою формулою:

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0},$$

де $I_q$ - загальний індекс фізичного обсягу продукції;
$q$ - індексована величина;
$p$ – вага.

Даний індекс показує зміну кількості виробленої або реалізованої продукції в звітному періоді порівняно з базисним.

*Загальний індекс обсягу* товарообороту показує зміну виробництва або реалізованої продукції в звітному періоду порівняно з базисним у фактичних цінах і визначається за наступною формулою:

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0}.$$

Вищенаведені індекси взаємозв'язані. Якщо загальний індекс цін перемножити на загальний індекс фізичного обсягу, отримаємо загальний індекс товарообороту у фактичних цінах.

$$I_p \cdot I_q = I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0} \cdot \frac{\sum q_1 p_0}{\sum q_0 p_0} = \frac{\sum p_1 q_1}{\sum p_0 q_0}.$$

Звідси $I_p = \frac{I_{pq}}{I_q}$, а $I_q = \frac{I_{pq}}{I_p}$.

Абсолютна сума економії або перевитрат від зміни цін визначається як різниця між чисельником і знаменником загального індекса цін:

$$\Delta pq(p) = \sum p_1 q_1 - \sum p_0 q_1.$$

Розглянемо приклад. Маємо дані про реалізацію деяких продовольчих товарів на ринках міста в січні місяці (табл. 12.1)

Таблиця 12.1

| Назва товару | Продано товарів тис. кг | | Середня ціна одного кілограма, грн. | |
|---|---|---|---|---|
| | 2006 р. | 2007 р. | 2006 р. | 2007 р. |
| Морква, кг. | 15,3 | 16,2 | 0,22 | 0,20 |
| Яблука, кг. | 49,8 | 51,6 | 0,90 | 0,85 |

Обчислимо спочатку індивідуальні індекси цін і фізичного обсягу.

$$i_p = \frac{p_1}{p_0} = \frac{0,20}{0,22} = 0,909; \qquad i_p = \frac{p_1}{p_0} = \frac{0,85}{0,90} = 0,944;$$

$$i_q = \frac{q_1}{q_0} = \frac{16,2}{15,3} = 1,059; \qquad i_q = \frac{q_1}{q_0} = \frac{51,6}{49,8} = 1,036.$$

Ціна кілограма моркви в січні 2004 року порівняно з січнем 2003 р. знизилась на 9,1 % (100 – 90,9), яблук – на 5,6 (100 – 94,4), а кількість реалізованої моркви збільшилася за цей же період в 1,059 рази, або на 5,9% (105,9 – 100), яблук – в 1,036 рази або на 3,6 %.

Визначимо загальний індекс цін на дані продукти:

$$I_p = \frac{\sum p_1q_1}{\sum p_0q_1} = \frac{0{,}20 \cdot 16{,}2 + 0{,}85 \cdot 51{,}6}{0{,}22 \cdot 16{,}2 + 0{,}90 \cdot 51{,}6} = \frac{47{,}100}{50{,}004} = 0{,}942\ .$$

Абсолютна сума виграшу населення від зниження цін:

$$\Delta pq(p) = \sum p_1q_1 - \sum p_0q_1 = 47{,}100 - 50{,}004 = -2{,}904 \text{ тис. грн.}$$

Індекс фізичного обсягу продуктів дорівнює:

$$I_q = \frac{\sum q_1p_0}{\sum q_0p_0} = \frac{16{,}2 \cdot 0{,}22 + 51{,}6 \cdot 0{,}90}{15{,}3 \cdot 0{,}22 + 49{,}8 \cdot 0{,}90} = \frac{50{,}004}{48{,}196} = 1{,}037\ .$$

Загальний індекс обсягу товарообороту обчислюємо за формулою:

$$I_{pq} = \frac{\sum p_1q_1}{\sum p_0q_0} = \frac{47{,}100}{48{,}196} = 0{,}977\ .$$

Абсолютна зміна товарообороту становить:

$$\Delta pq = \sum p_1q_1 - \sum p_0q_0 = 47{,}100 - 48{,}196 = -1{,}096 \text{ тис. грн.}$$

в т.ч. за рахунок зміни кількості проданих товарів:

$$\Delta pq(q) = \sum q_1p_0 - \sum q_0p_0 = 50{,}004 - 48{,}196 = 1{,}808 \text{ тис. грн.}$$

Перевіримо правильність наших розрахунків:

$$I_{pq} = I_p \cdot I_q = 0{,}977 = 0{,}942 \cdot 1{,}037;$$

$$\Delta pq = \Delta pq(p) + \Delta pq(q) = -1{,}096 = -2{,}904 + 1{,}808\ .$$

Отже, ціни на продукти на ринку в січні 2007 р. порівняно з січнем 2006 р. знизились на 5,8 %, внаслідок чого населення зекономило 2,904 тис. грн., кількість реалізованих продуктів за цей же період збільшилась в 1,037 рази або на 3,7%, а товарооборот в фактичних цінах – зменшився на 2,3 % або на 1,096 тис. грн., в т.ч. за рахунок зниження цін на 2,904 тис. грн., а за рахунок збільшення кількості проданих товарів він зріс на 1,808 тис. грн.

Аналогічно розраховують систему індексів, зв'язаних із собівартістю і кількістю виготовленої продукції.

Агрегатний індекс собівартості продукції визначають за формулою:

$$I_z = \frac{\sum z_1q_1}{\sum z_0q_1}\ .$$

Загальний індекс фізичного обсягу в даному випадку має вигляд:

$$I_q = \frac{\sum q_1z_0}{\sum q_0z_0}\ .$$

Зміна загального обсягу затрат та виробництво продукції визначається через індекс затрат виду:

$$I_{zq} = \frac{\sum z_1 q_1}{\sum z_0 q_0}.$$

Ці індекси взаємозв'язані:

$$I_z \cdot I_q = I_{zq} = \frac{\sum z_1 q_1}{\sum z_0 q_1} \cdot \frac{\sum q_1 z_0}{\sum q_0 z_0} = \frac{\sum z_1 q_1}{\sum z_0 q_0},$$

$$I_z = \frac{I_{zq}}{I_q}, \text{ а } \quad I_q = \frac{I_{zq}}{I_z}.$$

Загальна зміна затрат на виробництво продукції в звітному періоді порівняно з базисним дорівнює:

$$\Delta zq = \sum z_1 q_1 - \sum z_0 q_0,$$

в т.ч. за рахунок:

а) зміни собівартості одиниці продукції

$$\Delta zq(z) = \sum z_1 q_1 - \sum z_0 q_1,$$

б) зміни кількості виготовленої продукції

$$\Delta zq(q) = \sum q_1 z_0 - \sum z_0 q_0 .$$

### 12.3. Середньозважені індекси

В деяких випадках загальні індекси обчислюють як середні арифметичні, перетворені з відповідних агрегатних індексів.

Перетворюють агрегатний індекс в середній з індивідуальних індексів, підставляючи у його чисельник або знаменник замість індексованого показника його вираз, виведений з формули індивідуального індекса. Якщо таку заміну роблять у чисельнику, агрегатний індекс перетворюється в середній арифметичний, а якщо у знаменнику – в середній гармонічний.

Перетворимо агрегатний індекс фізичного обсягу в середній арифметичний. Як відомо, формула загального індексу фізичного обсягу така: $I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}$, а індивідуальний індекс фізичного обсягу $i_q = \frac{q_1}{q_0}$, звідки $q_1 = i_q \cdot q_0$. Замінивши в формулі агрегатного індекса фізичного обсягу продукції індексовану величину $q_1$ на $i_q q_0$, отримаємо формулу *середнього арифметичного індекса* фізичного обсягу продукції:

$$\bar{I}_q = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0}.$$

Таким чином, даний індекс виступає як середня арифметична величина із індивідуальних індексів, зважених на вартість продукції базисного періоду ($q_0 p_0$).

Наведемо приклад. Нехай маємо дані про продаж товарів в універмазі міста:

Таблиця 12.2

| Товарні групи | Продано в 2006 р., тис. грн., ($q_0 p_0$) | Індекси кількості проданих товарів ($i_q$) |
|---|---|---|
| Трикотажні вироби | 150 | 0,98 |
| Тканини | 200 | 1,05 |
| Галантерея | 30 | 1,20 |

За цими даними середній арифметичний індекс фізичного обсягу реалізованих товарів дорівнює:

$$\bar{I}_q = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0} = \frac{0{,}98 \cdot 150 + 1{,}05 \cdot 200 + 1{,}2 \cdot 30}{150 + 200 + 30} = \frac{393}{380} = 1{,}034.$$

Щоб перетворити агрегатний індекс цін у середній гармонічний, потрібно в знаменнику агрегатного індекса цін замінити індексовану величину $P_0$ її виразом із індивідуального індекса $\frac{P_1}{i_p}$, ( $i_p = \frac{P_1}{P_0}$ звідки $P_0 = \frac{P_1}{i_p}$ ), отримаємо формулу *середнього гармонічного індекса* цін виду:

$$\bar{I}_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}}.$$

Цей індекс являє собою середню гармонічну із величин, обернених індивідуальним індексам цін, зважених на вартість продукції в звітному періоді ($p_1 q_1$).

Покажемо обчислення такого індекса на прикладі. Маємо наступні дані про продаж товарів продуктовими магазинами споживчої кооперації за два квартали:

Таблиця 12.3

| Товари | Товарооборот в діючих цінах, тис. грн. | | Зміна середніх цін в ІІ кварталі порівняно з І кварталом, % |
|---|---|---|---|
| | І квартал | ІІ квартал | |
| Овочі | 60 | 64 | - 20 |
| М'ясо | 42 | 44 | + 10 |
| Зерно | 35 | 38 | без змін |

Спочатку визначимо індивідуальні індекси цін:

$$i_p = \frac{p_1}{p_0} = \frac{100-20}{100} = 0{,}8\,; \qquad i_p = \frac{p_1}{p_0} = \frac{100+10}{100} = 1{,}1\,;$$

$$i_p = \frac{p_1}{p_0} = \frac{100}{100} = 1{,}0\,.$$

Тепер маємо всі дані для обчислення середнього гармонічного індекса цін:

$$\bar{I}_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}} = \frac{64+44+38}{\frac{64}{0{,}8}+\frac{44}{1{,}1}+\frac{38}{1{,}0}} = \frac{146}{158} = 0{,}924\,.$$

Аналогічно перетворюють агрегатний індекс собівартості в середній гармонічний індекс:

$$I_z = \frac{\sum z_1 q_1}{\sum z_0 q_1}\,; \qquad i_z = \frac{z_1}{z_0}\,; \qquad z_0 = \frac{z_1}{i_z}\,;$$

$$\bar{I}_z = \frac{\sum z_1 q_1}{\sum \frac{z_1 q_1}{i_z}}\,.$$

Середній арифметичний і середній гармонічний індекси обчислюють в тих випадках, коли для визначення агрегатного індекса немає потрібних даних.

Середній арифметичний індекс цін і середній гармонічний індекс фізичного обсягу мають вигляд:

$$\bar{I}_p = \frac{\sum i_p p_0 q_1}{\sum p_0 q_1}\,; \qquad \bar{I}_q = \frac{\sum q_1 p_0}{\sum \frac{q_1 p_0}{ip}}\,.$$

Однак, дані індекси в статистичній практиці майже не використовуються, тому що в них вагами індивідуальних індексів є умовний товарооборот ($p_0q_1$), який не передбачено в звітності.

Середні арифметичні індекси і середні гармонічні індекси повинні співпадати за своєю величиною з відповідними агрегатними індексами.

Вибір форми індекса залежить від поставленого завдання дослідження і від наявності даних, необхідних для обчислення того чи іншого індекса.

Наведемо загальну схему перетворення основних агрегатних індексів в середні арифметичні і середні гармонічні індекси (табл. 12.4.)

Таблиця 12.4

Схема перетворення агрегатних індексів

| Назва індекса | Індивідуальний індекс | Перетворення індивідуального індекса | Агрегатний індекс | Середній арифметичний індекс | Середній гармонічний індекс |
|---|---|---|---|---|---|
| Фізичного обсягу | $i_q = \frac{q_1}{q_0}$ | $q_1 = i_q \cdot q_0$<br>$q_0 = q_1 : i_q$ | $I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}$ | $\bar{I}_q = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0}$ | $\bar{I}_q^* = \frac{\sum q_1 p_0}{\sum \frac{q_1 p_0}{i_q}}$ |
| Цін | $i_p = \frac{p_1}{p_0}$ | $p_1 = i_p \cdot p_0$<br>$p_0 = p_1 : i_p$ | $I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}$ | $\bar{I}_p^* = \frac{\sum i_p p_0 q_1}{\sum p_0 q_1}$ | $\bar{I}_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}}$ |
| Собівартості | $i_z = \frac{z_1}{z_0}$ | $z_1 = i_z \cdot z_0$<br>$z_0 = z_1 : i_z$ | $I_z = \frac{\sum z_1 q_1}{\sum z_0 q_1}$ | $\bar{I}_z^* = \frac{\sum i_z z_0 q_1}{\sum z_0 q_1}$ | $\bar{I}_z = \frac{\sum z_1 q_1}{\sum \frac{z_1 q_1}{i_z}}$ |
| Урожайності | $i_y = \frac{y_1}{y_0}$ | $y_1 = i_y \cdot y_0$<br>$y_0 = y_1 : i_y$ | $I = \frac{\sum y_1 \Pi_1}{\sum y_0 \Pi_1}$ | $\bar{I}_y^* = \frac{\sum i_y y_0 \Pi_1}{\sum y_0 \Pi_1}$ | $\bar{I}_y = \frac{\sum y_1 \Pi_1}{\sum \frac{y_1 \Pi_1}{i_y}}$ |
| Заробітної плати | $i_f = \frac{f_1}{f_0}$ | $f_1 = i_f \cdot f_0$<br>$f_0 = f_1 : i_f$ | $I_f = \frac{\sum f_1 T_1}{\sum f_0 T_1}$ | $\bar{I}_f^* = \frac{\sum i_f f_0 T_1}{\sum f_0 T_1}$ | $\bar{I}_f = \frac{\sum f_1 T_1}{\sum \frac{f_1 T_1}{i_f}}$ |
| Продуктивності праці | $i_w = \frac{w_1}{w_0}$ | $w_1 = i_w \cdot w_0$<br>$w_0 = w_1 : i_w$ | $I_w = \frac{\sum w_1 T_1}{\sum w_0 T_1}$ | $\bar{I}_w^* = \frac{\sum i_w w_0 T_1}{\sum w_0 T_1}$ | $\bar{I}_w = \frac{\sum w_1 T_1}{\sum \frac{w_1 T_1}{i_w}}$ |

* Такі формули теоретично можливі, але на практиці майже не використовуються.

## 12.4. Базисні і ланцюгові індекси з постійними і змінними вагами

В ряді випадків доводиться аналізувати явища суспільного життя не за два, а за три і більше послідовних періодів. В такому разі, в залежності від бази порівняння, обчислюють індекси з постійною базою порівняння (базисні) і змінною базою порівняння (ланцюгові).

*Базисними* називаються індекси, які вираховуються шляхом порівняння даних кожного періоду з даними будь-якого одного періоду, прийнятого за базу порівняння. Наприклад:

$$i_{q1} = \frac{q_1}{q_0}, \quad i_{q2} = \frac{q_2}{q_0}, ..., \; i_{qn} = \frac{q_n}{q_0}.$$

*Ланцюговими* називаються індекси, обчислені шляхом порівняння даних кожного періоду з даними попереднього періоду. Наприклад:

$$i_{q1} = \frac{q_1}{q_0}, \;\; i_{q2} = \frac{q_2}{q_1}, \quad i_{q3} = \frac{q_3}{q_2}, \; ..., \; i_{qn} = \frac{q_n}{q_{n-1}}.$$

Між базисними і ланцюговими індексами існує взаємозв’язок, що дозволяє переходити від одного виду індексів до іншого.

Базисні індекси можна визначити через ланцюгові, послідовно перемноживши останні:

$$i_q = \frac{q_1}{q_0} \cdot \frac{q_2}{q_1} \cdot \frac{q_3}{q_2} = \frac{q_3}{q_0}.$$

Ланцюгові індекси визначають через базисні шляхом ділення відповідного базисного індекса до попереднього базисного індекса.

$$i_q = \frac{q_2}{q_0} : \frac{q_1}{q_0} = \frac{q_2 q_0}{q_0 q_1} = \frac{q_2}{q_1}.$$

Базисні і ланцюгові індекси є індивідуальні і загальні.

Якщо порівнюваних періодів три і більше, то загальні (базисні і ланцюгові) індекси обчислюють з постійними і змінними вагами.

Якщо для всього індексованого ряду беруть ваги якогось одного періоду, то отримують базисні і ланцюгові індекси з постійними вагами. Якщо ваги змінюються від одного індекса до іншого, то матимемо базисні і ланцюгові індекси із змінними вагами.

Розглянемо вирахування базисних і ланцюгових індексів з постійними і змінними вагами на прикладі.

Таблиця 12.5

Динаміка реалізації овочів і цін на ринку за три літні місяці

| Товар | Кількість проданих товарів, тис. кг | | | Середня ціна за 1 кг, грн. | | |
|---|---|---|---|---|---|---|
| | червень | липень | серпень | червень | липень | серпень |
| Помідори | 15 | 48 | 62 | 1,00 | 0,80 | 0,60 |
| Огірки | 12 | 50 | 55 | 1,20 | 1,00 | 0,70 |
| Капуста | 10 | 58 | 70 | 0,90 | 0,50 | 0,20 |

Визначимо базисні і ланцюгові індекси цін з постійними вагами.

Базисні індекси цін з постійними вагами:

$$I_{p1}=\frac{\sum p_1q_0}{\sum p_0q_0}=\frac{0,80\cdot 15+1,00\cdot 12+0,50\cdot 10}{1,00\cdot 15+1,20\cdot 12+0,90\cdot 10}=\frac{29,0}{38,4}=0,755\ ;$$

$$I_{p2}=\frac{\sum p_2q_0}{\sum p_0q_0}=\frac{0,60\cdot 15+0,70\cdot 12+0,20\cdot 10}{1,00\cdot 15+1,20\cdot 12+0,90\cdot 10}=\frac{19,4}{38,4}=0,505\ .$$

Ланцюгові індекси цін з постійними вагами:

$$I_{p1}=\frac{\sum p_1q_0}{\sum p_0q_0}=\frac{0,80\cdot 15+1,00\cdot 12+0,50\cdot 10}{1,00\cdot 15+1,20\cdot 12+0,90\cdot 10}=\frac{29,0}{38,4}=0,755\ ;$$

$$I_{p2}=\frac{\sum p_2q_0}{\sum p_1q_0}=\frac{0,60\cdot 15+0,70\cdot 12+0,20\cdot 10}{0,80\cdot 15+1,00\cdot 12+0,50\cdot 10}=\frac{19,4}{29,0}=0,669\ .$$

Обчислимо базисні і ланцюгові індекси цін із змінними вагами.

Базисні індекси цін із змінними вагами:

$$I_{p1}=\frac{\sum p_1q_1}{\sum p_0q_1}=\frac{0,80\cdot 48+1,00\cdot 50+0,50\cdot 58}{1,00\cdot 48+1,20\cdot 50+0,90\cdot 58}=\frac{117,4}{160,2}=0,733\ ;$$

$$I_{p2}=\frac{\sum p_2q_2}{\sum p_0q_2}=\frac{0,60\cdot 62+0,70\cdot 55+0,20\cdot 70}{1,00\cdot 62+1,20\cdot 55+0,90\cdot 70}=\frac{89,7}{191,0}=0,470\ .$$

Ланцюгові індекси цін із змінними вагами:

$$I_{p1}=\frac{\sum p_1q_1}{\sum p_0q_1}=\frac{0,80\cdot 48+1,00\cdot 50+0,50\cdot 58}{1,00\cdot 48+1,20\cdot 50+0,90\cdot 58}=\frac{117,4}{160,2}=0,733\ ;$$

$$I_{p2}=\frac{\sum p_2q_2}{\sum p_1q_2}=\frac{0,60\cdot 62+0,70\cdot 55+0,20\cdot 70}{0,80\cdot 62+1,00\cdot 55+0,50\cdot 70}=\frac{89,7}{139,6}=0,642\ .$$

Між базисними і ланцюговими індексами з постійними вагами існує таке ж співвідношення, яке встановлене для індивідуальних індексів:

$$\frac{\sum p_1q_0}{\sum p_0q_0}\cdot\frac{\sum p_2q_0}{\sum p_1q_0}=\frac{\sum p_2q_0}{\sum p_0q_0}=0,755\cdot 0,669=0,505;$$

$$\frac{\sum p_2q_0}{\sum p_0q_0}:\frac{\sum p_1q_0}{\sum p_0q_0}=\frac{\sum p_2q_0}{\sum p_1q_0}=0,505:0,755=0,669.$$

Як видно із нашого прикладу, добуток послідовних ланцюгових індексів цін з постійними вагами дорівнює базисному індексу цін з постійними вагами наступного періоду, а відношення наступного базисного індексу до попереднього дорівнює ланцюговому індексу цін наступного періоду.

Для ланцюгових індексів із змінними вагами такого взаємозв'язку не існує.

В статистичній практиці з постійними вагами обчислюють, в основному, індекси фізичного обсягу продукції, а із змінними вагами – індекси цін, собівартості, урожайності та інших якісних показників.

Для визначення індексів якісних показників за тривалий період обчислюють, як правило, ланцюгові індекси із змінними вагами, а для індекса фізичного обсягу та інших індексів об'ємних показників ефективніше вираховувати базисні індекси з постійними вагами.

Таким чином, в залежності від мети статистичного дослідження і змісту соціально-економічного явища, застосовують базисні і ланцюгові індекси з постійними і змінними вагами.

## 12.5. Індекси змінного, постійного складу і структурних зрушень

Для якісних показників, таких як середня ціна, собівартість, урожайність та інших по однойменній продукції, але віднесеній до різних об'єктів, обчислюють загальні індекси змінного, постійного (фіксованого) складу і структурних зрушень.

Зміна середньої величини досліджуваного якісного показника залежить від варіації значення кожної окремої величини явища, яке вивчається і від зміни структури явища. Наприклад, зміна середньої заробітної плати працівників залежить від зміни заробітної плати окремих категорій працівників і від їх питомої ваги в загальній чисельності.

Індекс, який характеризує спільний вплив обох факторів, називається *індексом змінного складу* і визначається за формулою:

$$I_{\bar{z}} = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} = \frac{\bar{z}_1}{\bar{z}_0},$$

де $I_{\bar{z}}$ – загальний індекс собівартості продукції змінного складу;

$z_1$, $z_0$ – собівартість одиниці продукції в звітному і базисному періодах;

$q_1$, $q_0$ – кількість виробленої продукції в натуральному виразі в звітному і базисному періодах;

$\bar{z}_1, \bar{z}_0$ – осереднені ознаки.

На величину індекса собівартості змінного складу впливають зміни рівнів собівартості і зміни в структурі продукції (її складі). Щоб виявити роль кожного фактора в загальній динаміці середньої, потрібно індекс змінного складу розкласти на два індекси-співмножники, кожний з яких відображає вплив тільки одного фактора.

Перший індекс, який характеризує вплив тільки індексованої величини (в якому змінюється тільки собівартість), називається *індексом постійного (фіксованого) складу*. Він обчислюється за формулою:

$$I_z = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_1}{\sum q_1} = \frac{\sum z_1 q_1}{\sum z_0 q_1}.$$

Другий індекс показує, як змінюється середній рівень (середня собівартість) тільки за рахунок зміни структури явища (структури продукції). Він називається *індексом структурних зрушень* і визначається за наступною формулою:

$$I_{zdq} = \frac{\sum z_0 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} = \frac{\sum z_0 q_1}{\bar{z}_0 \sum q_1}.$$

Взаємозв'язок між перерахованими індексами можна подати у вигляді наступної системи:

$$I_{\bar{z}} = I_z \cdot I_{zdq} = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} =$$

$$= \left( \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_1}{\sum q_1} \right) \cdot \left( \frac{\sum z_0 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} \right).$$

Розглянемо застосування даної системи індексів на прикладі аналізу середньої собівартості продукції “А” на двох заводах (табл. 12.6).

Таблиця 12.6

| № заводу | Базисний період | | | Звітний період | | |
|---|---|---|---|---|---|---|
| | Вироблено продукції, тис. шт. | Собівартість одиниці, грн. | Питома вага продукції, % | Вироблено продукції, тис. шт. | Собівартість одиниці, грн. | Питома вага продукції, % |
| 1 | 70 | 25 | 50 | 80 | 22 | 40 |
| 2 | 70 | 23 | 50 | 120 | 20 | 60 |
| Разом | 140 | x | 100 | 200 | x | 100 |

Обчислимо індекс собівартості продукції змінного складу:

$$I_{\bar{z}} = \frac{\bar{z}_1}{\bar{z}_0} = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} = \frac{22 \cdot 80 + 20 \cdot 120}{200} : \frac{25 \cdot 70 + 23 \cdot 70}{140} =$$

$$= \frac{4160}{200} : \frac{3360}{140} = \frac{20{,}8}{24{,}0} = 0{,}867, \quad \text{або } 86{,}7\ \%.$$

Отже, середня собівартість знизилась на 13,3% (100,0 – 86,7).

Економія на одиницю продукції становить – 3,2 грн. (20,8 – 24,0), а на весь обсяг продукції звітного періоду – 640 тис. грн. (3,2 · 200).

Зниження середньої собівартості одиниці продукції зумовлене зміною собівартості продукції на кожному заводі і зміною структури продукції. Для виявлення впливу кожного з цих факторів на динаміку середньої собівартості, обчислимо індекси собівартості:

а) постійного (фіксованого) складу:

$$I_z = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_1}{\sum q_1} = \frac{22 \cdot 80 + 20 \cdot 120}{200} : \frac{25 \cdot 80 + 23 \cdot 120}{200} =$$

$$= \frac{4160}{200} : \frac{4760}{200} = \frac{20{,}8}{23{,}8} = 0{,}874, \quad \text{або } 87{,}4\ \%.$$

Собівартість продукції по двох заводах разом в середньому знизилась на 12,6 % (100,0 – 87,4). Економія затрат на виробництво продукції в звітному періоді складає 600 тис. грн. [(20,8 – 23,8) · 200];

б) структурних зрушень:

$$I_{zdq} = \frac{\sum z_0 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} = \frac{23{,}8}{24{,}0} = 0{,}992, \quad \text{або } 99{,}2\ \%.$$

Це означає, що середня собівартість виробу “А” в звітному періоді додатково знизилась на 0,8 % (100,0 – 99,2) за рахунок зміни структури, тобто за рахунок збільшення частки продукції другого заводу з 50 % до 60 %, на якому рівень собівартості був дещо нижчим в порівнянні з першим заводом. За рахунок цієї зміни економія затрат виробництва досягла в звітному періоді 40 тис. грн. [(23,8 – 24,0) · 200].

Проведемо перевірку наших розрахунків:

$$I_{\bar{z}} = I_z \cdot I_{zdq} = 0{,}867 = 0{,}874 \cdot 0{,}992;$$

$$\Delta_{zq} = \Delta zq(z) + \Delta zq(dq) = 640 = 600 + 40\,.$$

Отже, всі індекси обчислені правильно.

Вирахувані нами індекси можна обчислити іншим способом, за питомими вагами продукції заводів, вираженими в коефіцієнтах:

а) індекс собівартості змінного складу:

$$I_{\bar{z}} = \frac{\sum z_1 d_1}{\sum z_0 d_0} = \frac{22 \cdot 0{,}4 + 20 \cdot 0{,}6}{25 \cdot 0{,}5 + 23 \cdot 0{,}5} = \frac{20{,}8}{24{,}0} = 0{,}867, \text{ або } 86{,}7\ \%\,;$$

де $d_1$, $d_0$ – коефіцієнти частки продукції заводів в звітному і базисному періодах;

б) індекс собівартості продукції постійного складу:

$$I_z = \frac{\sum z_1 d_1}{\sum z_0 d_1} = \frac{22 \cdot 0{,}4 + 20 \cdot 0{,}6}{25 \cdot 0{,}4 + 23 \cdot 0{,}5} = \frac{20{,}8}{23{,}8} = 0{,}874, \text{ або } 87{,}4\ \%\,;$$

в) індекс структурних зрушень:

$$I_{zdq} = \frac{\sum z_0 d_1}{\sum z_0 d_0} = \frac{22 \cdot 0{,}4 + 23 \cdot 0{,}6}{25 \cdot 0{,}5 + 23 \cdot 0{,}5} = \frac{23{,}8}{24{,}0} = 0{,}992, \text{ або } 99{,}2\ \%.$$

Індекс структурних зрушень може бути вирахуваний також через взаємозв'язок вищенаведених індексів:

$$I_{zdq} = \frac{I_{\bar{z}}}{I_z} = \frac{0{,}867}{0{,}874} = 0{,}992$$

Наведемо формули найбільш розповсюджених індексів змінного складу, постійного складу і структурних зрушень.

Таблиця 12.7

| Назва індекса | ІНДЕКСИ | | |
|---|---|---|---|
| | Змінного складу | Постійного складу | Структурних зрушень |
| Цін | $\frac{\sum p_1 q_1}{\sum q_1} : \frac{\sum p_0 q_0}{\sum q_0}$ | $\frac{\sum p_1 q_1}{\sum q_1} : \frac{\sum p_0 q_1}{\sum q_1} = \frac{\sum p_1 q_1}{\sum p_0 q_1}$ | $\frac{\sum p_0 q_1}{\sum q_1} : \frac{\sum p_0 q_0}{\sum q_0} = \frac{\sum p_0 q_1}{\bar{p}_0 \sum q_1}$ |
| Собівартості | $\frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0}$ | $\frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_1}{\sum q_1} = \frac{\sum z_1 q_1}{\sum z_0 q_1}$ | $\frac{\sum z_0 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} = \frac{\sum z_0 q_1}{\bar{z}_0 \sum q_1}$ |
| Урожайності | $\frac{\sum y_1 \Pi_1}{\sum \Pi_1} : \frac{\sum y_0 \Pi_0}{\sum \Pi_0}$ | $\frac{\sum y_1 \Pi_1}{\sum \Pi_1} : \frac{\sum y_0 \Pi_1}{\sum \Pi_1} = \frac{\sum y_1 \Pi_1}{\sum y_0 \Pi_1}$ | $\frac{\sum y_0 \Pi_1}{\sum \Pi_1} : \frac{\sum y_0 \Pi_0}{\sum \Pi_0} = \frac{\sum y_0 \Pi_1}{\bar{y}_0 \sum \Pi_1}$ |
| Заробітної плати | $\frac{\sum f_1 T_1}{\sum T_1} : \frac{\sum f_0 T_0}{\sum T_0}$ | $\frac{\sum f_1 T_1}{\sum T_1} : \frac{\sum f_0 T_1}{\sum T_1} = \frac{\sum f_1 T_1}{\sum f_0 T_1}$ | $\frac{\sum f_0 T_1}{\sum T_1} : \frac{\sum f_0 T_0}{\sum T_0} = \frac{\sum f_0 T_1}{\bar{f}_0 \sum T_1}$ |
| Продуктивності праці | $\frac{\sum w_1 T_1}{\sum T_1} : \frac{\sum w_0 T_0}{\sum T_0}$ | $\frac{\sum w_1 T_1}{\sum T_1} : \frac{\sum w_0 T_1}{\sum T_1} = \frac{\sum w_1 T_1}{\sum w_0 T_1}$ | $\frac{\sum w_0 T_1}{\sum T_1} : \frac{\sum w_0 T_0}{\sum T_0} = \frac{\sum w_0 T_1}{\bar{w} \sum T_1}$ |

де $p_1$, $p_0$ – ціна одиниці продукції в звітному і базисному періодах;

$z_1$, $z_0$ – собівартість одиниці продукції в звітному і базисному періодах;

$q_1$, $q_0$ – кількість продукції в звітному і базисному періодах;

$y_1$, $y_0$ – урожайність сільськогосподарських культур в звітному і базисному періодах;

$n_1$, $n_0$ – посівна площа в звітному і базисному періодах;

$f_1$, $f_0$ – заробітна плата працівників в звітному і базисному періодах;

$T_1$, $T_0$ – чисельність працівників в звітному і базисному періодах;

$w_1$, $w_0$ – продуктивність праці в звітному і базисному періодах.

## 12.6. Територіальні індекси

В практиці статистичних досліджень часто виникає потреба зіставлення рівнів економічних явищ в просторі, для чого використовують територіальні індекси.

*Територіальні індекси* – це узагальнюючі відносні величини, що дають порівняльну характеристику в розрізі територій або об'єктів.

Індивідуальні територіальні індекси визначають як звичайні відносні величини порівняння.

При побудові загальних територіальних індексів якісних показників виникає питання вибору бази порівняння і ваг.

Якщо порівнювання здійснюється по двох територіях (об'єктах), то базою порівняння може бути показник будь-якої території (об'єкта). У випадку, коли порівняння проводять по багатьох територіях (об'єктах) – база порівняння повинна бути економічно обґрунтованою. Так, наприклад, якщо порівнюється собівартість продукції по групі однотипних підприємств з приблизно однаковими техніко-економічними умовами виробництва, то за базу порівняння доцільно взяти передове підприємство, на якому зафіксована найнижча собівартість одиниці продукції.

При побудові територіальних індексів якісних показників вагами можуть виступати:

а) кількісний (екстенсивний) показник тієї території, на якій якісний (інтенсивний) показник найбільш економічно кращий;

б) кількісний показник однієї з двох порівнюваних територій (об'єктів);

в) середній кількісний показник з багатьох порівнюваних територій (об'єктів);

г) об'ємний кількісний показник (сума екстенсивних показників декількох територій або об'єктів);

д) кількісний показник, прийнятий за стандарт.

Проілюструємо обчислення територіальних індексів цими способами на умовному прикладі.

Таблиця 12.8

Урожайність і посівні площі зернових двох районів

<table>
<tr><th rowspan="3">Куль-тура</th><th colspan="2">Середня урожайність, ц/га</th><th colspan="4">Посівна площа, га</th></tr>
<tr><th rowspan="2">по району "А"</th><th rowspan="2">по району "В"</th><th rowspan="2">по району "А"</th><th rowspan="2">по району "В"</th><th colspan="2">по області</th></tr>
<tr><th>в га</th><th>в %</th></tr>
<tr><td>Пшениця</td><td>36</td><td>42</td><td>190</td><td>210</td><td>2850</td><td>50</td></tr>
<tr><td>Жито</td><td>19</td><td>23</td><td>80</td><td>100</td><td>1200</td><td>21</td></tr>
<tr><td>Ячмінь</td><td>25</td><td>30</td><td>110</td><td>150</td><td>1650</td><td>29</td></tr>
</table>

Обчислимо, спочатку, територіальні індекси урожайності зернових, прийнявши в якості бази порівняння район "В", в якому вища урожайність:

а) з вагами (посівними площами) цього району:

$$I_у = \frac{\sum у_А \cdot П_В}{\sum у_В \cdot П_В} = \frac{36 \cdot 210 + 19 \cdot 100 + 25 \cdot 150}{42 \cdot 210 + 23 \cdot 100 + 30 \cdot 150} = \frac{13210}{15620} = 0,846\ ;$$

б) з вагами району “А”:

$$I_у = \frac{\sum у_А \cdot П_А}{\sum у_В \cdot П_А} = \frac{36 \cdot 190 + 19 \cdot 80 + 25 \cdot 110}{42 \cdot 190 + 23 \cdot 80 + 30 \cdot 110} = \frac{11110}{13120} = 0,847\ .$$

Замінемо базу порівняння і визначимо територіальні індекси:

а) з вагами бази порівняння:

$$I_у = \frac{\sum у_В \cdot П_А}{\sum у_А \cdot П_А} = \frac{42 \cdot 190 + 23 \cdot 80 + 30 \cdot 110}{36 \cdot 190 + 19 \cdot 80 + 25 \cdot 110} = \frac{13120}{11110} = 1,181;$$

б) з вагами району “В”:

$$I_у = \frac{\sum у_В \cdot П_В}{\sum у_А \cdot П_В} = \frac{42 \cdot 219 + 23 \cdot 100 + 30 \cdot 150}{36 \cdot 210 + 19 \cdot 100 + 25 \cdot 150} = \frac{15620}{13210} = 1,182\ .$$

Визначимо тепер територіальні індекси урожайності зернових з об’ємними вагами:

а) для району “А” порівняно з районом “В”:

$$I_у = \frac{\sum у_А \cdot (П_А + П_В)}{\sum у_В \cdot (П_А + П_В)} =$$

$$= \frac{36 \cdot (190 + 210) + 19 \cdot (80 + 100) + 25 \cdot (110 + 150)}{42 \cdot (190 + 210) + 23 \cdot (80 + 100) + 30 \cdot (110 + 150)} =$$

$$= \frac{36 \cdot 400 + 19 \cdot 180 + 25 \cdot 260}{42 \cdot 400 + 23 \cdot 180 + 30 \cdot 260} = \frac{24320}{28740} = 0,846\ ;$$

б) для району “В” порівняно з районом “А”:

$$I_у = \frac{\sum у_А \cdot (П_А + П_В)}{\sum у_В \cdot (П_А + П_В)} = \frac{28740}{24320} = 1,182\ .$$

Аналогічні результати отримаємо, використавши в якості ваг середні посівні площі для окремих культур зернових:

а) для району “А” порівняно з районом “В”:

$$I_у = \frac{\sum у_А \overline{П}_{(А+В)}}{\sum у_В \overline{П}_{(А+В)}} = \frac{36 \cdot 200,8 + 19 \cdot 90,9 + 25 \cdot 131,8}{42 \cdot 200,8 + 23 \cdot 90,9 + 30 \cdot 131,8} = \frac{12250,9}{14478,3} = 0,846\ ;$$

б) для району “В” порівняно з районом “А”:

$$I_у = \frac{\sum у_В \overline{П}_{(А+В)}}{\sum у_А \overline{П}_{(А+В)}} = \frac{14478,3}{12250,9} = 1,182\ .$$

І, зрештою, знайдемо територіальні індекси урожайності зернових із стандартними вагами:

а) для району “А” порівняно з районом “В”:

$$I_у = \frac{\sum у_А П_{ст}}{\sum у_В П_{ст}} = \frac{36 \cdot 50 + 19 \cdot 21 + 25 \cdot 29}{42 \cdot 50 + 23 \cdot 21 + 30 \cdot 29} = \frac{2924}{3453} = 0,847 ;$$

б) для району “В” порівняно з районом “А”:

$$I_у = \frac{\sum у_В П_{ст}}{\sum у_А П_{ст}} = \frac{3453}{2924} = 1,181 .$$

Отже, урожайність зернових в районі “А” нижча, ніж у районі “В” на 15,3% (100 – 84,7), а в районі “В” порівняно з районом “А” вона вища в 1,181 рази, або на 18,1 % (181,1 – 100).

Як показали наші розрахунки, всі способи обчислення територіальних індексів дали приблизно однакові результати, що свідчить про те, що любий з них може бути використаний для порівняльного аналізу рівнів економічних явищ в просторі.

Будуючи територіальні індекси для кількісних показників (в основному фізичного об’єму продукції), вагами виступає середній якісний показник по території, для якої здійснюється порівняння, або використовують стандартні ваги.

Наведемо приклад. Нехай, маємо дані про реалізацію овочів на ринках двох районів.

Таблиця 12.9

| Товари | Район “А” | | Район ”В” | |
|---|---|---|---|---|
| | Ціна за 1 кг., грн. | Продано, т | Ціна за 1 кг., грн. | Продано, т |
| Картопля | 0,20 | 700 | 0,25 | 900 |
| Капуста | 0,25 | 300 | 0,20 | 500 |
| Помідори | 0,40 | 100 | 0,35 | 200 |

Спочатку визначимо середні ціни на овочі:

$$\overline{P}_1 = \frac{\sum pq}{\sum q} = \frac{0,20 \cdot 700 + 0,25 \cdot 900}{700 + 900} = \frac{365}{1600} = 0,23 \text{ грн. };$$

$$\overline{P}_2 = \frac{\sum pq}{\sum q} = \frac{0,25 \cdot 300 + 0,20 \cdot 500}{300 + 500} = \frac{175}{800} = 0,22 \text{ грн. };$$

$$\overline{P}_3 = \frac{\sum pq}{\sum q} = \frac{0,40 \cdot 100 + 0,35 \cdot 200}{100 + 200} = \frac{110}{300} = 0,37 \text{ грн. }.$$

Обчислимо територіальні індекси реалізації овочів за формулою:

а) для району “А” порівняно з районом “В”:

$$I_q = \frac{\sum q_A \overline{P}}{\sum q_B \overline{P}} = \frac{700 \cdot 0{,}23 + 300 \cdot 0{,}22 + 100 \cdot 0{,}37}{900 \cdot 0{,}23 + 500 \cdot 0{,}22 + 200 \cdot 0{,}37} = \frac{264}{391} = 0{,}675;$$

б) для району “В” порівняно з районом “А”:

$$I_q = \frac{\sum q_B \overline{P}}{\sum q_A \overline{P}} = \frac{900 \cdot 0{,}23 + 500 \cdot 0{,}22 + 200 \cdot 0{,}37}{700 \cdot 0{,}23 + 300 \cdot 0{,}22 + 100 \cdot 0{,}37} = \frac{391}{264} = 1{,}481.$$

Отже, кількість проданих овочів в районі “А” нижча, ніж в районі “В” на 32,5 %, відповідно в районі “В” реалізація вища за район “А” в 1,481 рази, або на 48,1 %.

Вибір бази порівняння і ваг в кожному конкретному випадку потрібно вирішувати, виходячи з мети самого дослідження.

## 12.7. Використання системи взаємозв’язаних індексів в аналізі факторів динаміки

Соціально-економічні явища і процеси взаємозв’язані між собою, що виражається у взаємозв’язку відповідних показників. Одна з форм взаємозв’язку між економічними показниками заключається в тому, що їх можна виразити як добуток кількох інших показників. Так, фонд заробітної плати може бути поданий у вигляді добутку заробітної плати одного працівника на загальне число працівників.

$$F = f \cdot T,$$

де $F$ – фонд заробітної плати;

$T$ – число працівників;

$f$ – заробітна плата одного працівника $\left(f = \dfrac{F}{T}\right)$.

Товарооборот у фактичних цінах можна виразити як добуток ціни на кількість проданих товарів, валовий збір любої культури – як добуток врожайності на посівну площу і т.д.

Показники-співмножники виступають тут як фактори, від величини яких залежить результативна ознака. Так, збільшення фонду заробітної плати може бути в результаті підвищення середньої заробітної плати працівників або збільшення числа працівників. Збільшення валового збору сільськогосподарських культур відбувається або за рахунок підвищення врожайності, або збільшення розміру посівних площ.

У зв’язку з цим, при економічному аналізі динаміки потрібно виявити і оцінити роль кожного окремого фактора в зміні результативного показника. Наприклад, як вплинули на зміну товарообороту зниження цін і збільшення реалізації кількості товарів, або скільки надпланової продукції отримано за рахунок росту врожайності і скільки – за рахунок розширення посівних площ і т.п.

Виявлення ролі окремих факторів має велике значення для визначення, на які з них потрібно вплинути, щоб добитись бажаного економічного ефекту.

Завдання виявлення і оцінки ролі окремих факторів у зміні економічного явища статистика вирішує шляхом побудови системи взаємозв'язаних індексів.

Якщо має місце взаємозв'язок між економічними показниками, то, безперечно, такий же взаємозв'язок існує і між індексами цих показників.

Для індивідуальних індексів існує правило: *індекс добутку дорівнює добутку індексів співмножників*. Наприклад, індекс затрат на виробництво будь-якого товару дорівнює добутку індексу собівартості на індекс кількості товару (його фізичного обсягу):

$$i_{zq} = \frac{z_1 q_1}{z_0 q_0} = \frac{z_1}{z_0} \cdot \frac{q_1}{q_0} = i_z \cdot i_q .$$

Отже, знаючи взаємозв'язок економічних показників, легко будується система взаємозв'язаних індивідуальних індексів.

При побудові системи взаємозв'язаних загальних індексів, індекс результативного показника буде дорівнювати добутку факторних індексів при дотриманні певних умов. В таких системах кожний фактор в одному з індексів факторів-співмножників є індексованим показником, а в другому виступає як вага, причому ваги в індексах – співмножниках повинні бути зафіксовані на різних рівнях.

Індекс результативного показника завжди виступає як добуток індекс якісного показника на індекс об'ємного показника. Як правило, в індексах якісних показників ваги фіксуються на рівні звітного періоду, а в індексах кількісних показників – на рівні базисного періоду.

Так, наприклад, загальний індекс товарообороту може бути виражений як добуток індекса цін на індекс фізичного обсягу продукції:

$$\dot{I}_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \cdot \frac{\sum q_1 p_0}{\sum q_0 p_0} = \dot{I}_p \cdot \dot{I}_q .$$

Як видно з цієї системи фактор цін є індексованим показником загального індекса цін і вагою в другому індексі – співмножнику – фізичному обсязі продукції і, навпаки, фактор фізичного обсягу є індексованим показником загального індекса фізичного обсягу і вагою в індексі цін. В загальному (якісному) індексі цін ваги взяті за звітній період, а в індексі фізичного обсягу вони зафіксовані на рівні базисного періоду.

Аналогічно можна виразити взаємозв'язок між іншими показниками, наприклад, собівартістю одиниці продукції, обсягом продукції і затратами на її виробництво. Цей взаємозв'язок має вигляд:

$$\dot{I}_{zq} = \frac{\sum z_1 q_1}{\sum z_0 q_0} = \frac{\sum z_1 q_1}{\sum z_0 q_1} \cdot \frac{\sum q_1 z_0}{\sum q_0 z_0} = \dot{I}_z \cdot \dot{I}_q .$$

Якщо агрегатний індекс продуктивності праці помножити на індекс затрат часу, то отримаємо загальний індекс фізичного обсягу:

$$\dot{I}_{\overline{w}} \cdot \dot{I}_T = \dot{I}_q = \left( \frac{\sum q_1 p_0}{\sum T_1} : \frac{\sum q_0 p_0}{\sum T_0} \right) \cdot \frac{\sum T_1}{\sum T_0} = \frac{\sum q_1 p_0}{\sum q_0 p_0} .$$

Валовий збір якої-небудь сільськогосподарської культури в динаміці отримаємо, якщо перемножимо загальний індекс урожайності на індекс посівної площі:

$$\dot{I}_{\overline{y}} \cdot \dot{I}_{П} = \dot{I}_{yП} = \left( \frac{\sum y_1 П_1}{\sum П_1} : \frac{\sum y_0 П_0}{\sum П_0} \right) \cdot \frac{\sum П_1}{\sum П_0} = \frac{\sum y_1 П_1}{\sum y_0 П_0} .$$

Взаємозв'язок індексів використовують для знаходження за двома відомими індексами третього, величина якого невідома. Якщо, наприклад, валовий збір зернових збільшився в 1,25 рази, а посівна площа не змінилась, то середня урожайність зернових в звітному періоді порівняно з базисним періодом зросла в 1,25 рази:

$I_{yП} : I_{П} = I_{\overline{y}} = 1{,}25 : 1 = 1{,}25$, або на 25 % (125 – 100).

Візьмемо інший приклад. Нехай відомо, що собівартість одиниці продукції в звітному періоді порівняно з базисним періодом зменшилась на 5%, а фізичний обсяг виробленої продукції збільшився в 1,13 рази. Визначимо, через взаємозв'язок індексів, як змінились затрати на виробництво продукції.

$$\dot{I}_{zq} = \dot{I}_z \cdot \dot{I}_q = 0{,}95 \cdot 1{,}13 = 1{,}073 .$$

Таким чином, загальні затрати на виробництво продукції в звітному періоді порівняно з базисним збільшились в 1,073 рази або на 7,3% (107,3 – 100).

Інколи для обчислення невідомого індекса за зв'язаними з ним відомими індексами не обов'язкова побудова розгорнутої системи індексів. Так, якщо відомо, що обсяг національного доходу за звітній період збільшився в 1,5 рази, а чисельність населення за цей же період зросла на 2 %, то національний доход в розрахунку на душу населення збільшиться в 1,47 рази (1,5 : 1,02 = 1,47). Тобто, обсяг національного доходу на душу населення обчислюється діленням його загального обсягу на чисельність населення. Аналогічно зв'язані між собою і відповідні індекси.

$$\dot{I}_{N/S} = \dot{I}_N : \dot{I}_S = 1{,}5 : 1{,}02 = 1{,}47\ ,$$

де $\dot{I}_{N/S}$ – загальний індекс національного доходу в розрахунку на душу населення;

$\dot{I}_N$ – індекс загального обсягу національного доходу;

$\dot{I}_S$ – загальний індекс чисельності населення.

Таким чином, залежно від мети і характеру дослідження економічного явища, питання про те, якими саме факторами визначається величина цього явища і на які фактори-співмножники слід розкласти результативний показник, вирішується на основі всебічного аналізу, виходячи з конкретних завдань статистичного дослідження.

## 12.8. Розклад абсолютного приросту за факторами

Система взаємопов'язаних індексів використовується не тільки для характеристики відносної ролі факторів у зміні складного економічного явища. Вона дає можливість виявити і оцінити роль цих факторів також в абсолютному вираженні.

Розглянемо приклад. Нехай маємо дані про реалізацію продуктових товарів на ринку:

Таблиця 12.10

| Назва продуктів | Базисний рік | | | Звітний рік | | | Вартість проданого товару у звітному році за цінами базисного року, тис. грн. $p_0q_1$ |
|---|---|---|---|---|---|---|---|
| | Ціна одиниці продукції, грн. $p_0$ | Кількість проданого товару, тис. од. $q_0$ | Вартість проданого товару, тис. грн. $p_0q_0$ | Ціна одиниці продукції, грн. $p_1$ | Кількість проданого товару, тис. од. $q_1$ | Вартість проданого товару, тис. грн. $p_1q_1$ | |
| Картопля, кг. | 0,35 | 45 | 15,75 | 0,30 | 53 | 15,9 | 18,55 |
| Молоко, л. | 0,45 | 13 | 5,85 | 0,40 | 20 | 8,0 | 9,00 |
| М'ясо, кг. | 4,50 | 10 | 45,00 | 4,00 | 14 | 56,0 | 63,00 |
| Разом: | х | х | 66,60 | х | х | 79,9 | 90,55 |

За наведеними даними визначимо спочатку відносну зміну товарообороту (вартості проданих товарів) в фактичних цінах у звітному році порівняно з базисним роком:

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{79{,}9}{66{,}6} = 1{,}2 .$$

Різниця між чисельником і знаменником даного індекса дасть нам абсолютний приріст товарообороту в звітному році порівняно з базисним за рахунок зміни двох факторів – ціни одиниці товару і обсягу проданих товарів кожного виду:

$$\Delta_{pq} = \sum p_1 q_1 - \sum p_0 q_0 = 79{,}9 - 66{,}6 = 13{,}3 \text{ тис. грн.}$$

Отже, в цілому по всіх видах товарів, ріст обсягу товарообороту склав 13,3 тис. грн., що на 20 % більше ніж в базисному році.

Завдання полягає в тому, щоб визначити, як змінився обсяг товарообороту за рахунок кожного фактора зокрема. Тобто, загальний абсолютний приріст товарообороту (+13,3 тис. грн.) потрібно розкласти за факторами.

Розглянемо зміну товарообороту по кожному виду товарів в цілому і за рахунок обох факторів окремо:

по картоплі:

$$i_{pq} = \frac{p_1 q_1}{p_0 q_0} = \frac{15{,}90}{15{,}75} = 1{,}009 .$$

$$\Delta_{pq} = p_1 q_1 - p_0 q_0 = 15{,}90 - 15{,}75 = 0{,}15 \text{ тис. грн.}$$

в т.ч. за рахунок:

а) ціна 1 кг картоплі:

$$i_p = \frac{p_1}{p_0} = \frac{0{,}30}{0{,}35} = 0{,}857 .$$

$$\Delta_{pq}(P) = p_1 q_1 - p_0 q_1 = 15{,}9 - 18{,}55 = -2{,}65 \text{ тис. грн.}$$

або $\Delta_{pq}(p) = (p_1 - p_0) q_1 = (0{,}30 - 0{,}35) \cdot 53 = -2{,}65$ тис. грн.

б) кількості проданої картоплі:

$$i_q = \frac{q_1}{q_0} = \frac{53}{45} = 1{,}178 .$$

$$\Delta_{pq}(q) = q_1 p_0 - q_1 p_0 = 18{,}55 - 15{,}75 = 2{,}80 \text{ тис. грн.}$$

$$\Delta_{pq}(q) = (q_1 - q_0) p_1 = (53 - 45) \cdot 0{,}35 = 2{,}80 \text{ тис. грн.}$$

Перевірка:

$$i_{pq} = i_p \cdot i_q = 0{,}857 \cdot 1{,}178 = 1{,}009 .$$

$$\Delta_{pq} = \Delta_{pq}(p) + \Delta_{pq}(q) = -2{,}65 + 2{,}80 = 0{,}15 \text{ тис. грн.}$$

В результаті збільшення реалізації картоплі в фактичних цінах в 1,009 рази, або на 0,9 % отримано приріст товарообороту в сумі 0,15 тис. грн., в т.ч. за рахунок зниження цін на 14,3 % товарооборот зменшився на 2,65 тис. грн., а за рахунок росту фізичного обсягу реалізації на 17,8% - він збільшився на 2,80 тис. грн.

Аналогічні розрахунки по інших товарах для наочності показані в табл. 12.11.

Таблиця 12.11

Розклад приросту товарообороту за факторами (тис. грн.)

| Назва товару | Збільшення (+) або зменшення (-) товарообороту за рахунок зміни | | |
|---|---|---|---|
| | Ціни одиниці товару | Кількості проданих товарів | Обох факторів |
| Картопля | 15,9 – 18,55 = – 2,65 | 18,55 – 15,75 = 2,80 | 15,90 – 15,75 = 0,15 |
| Молоко | 8,0 – 9,00 = – 1,00 | 9,00 – 5,85 = 3,15 | 8,00 – 5,85 = 2,15 |
| М'ясо | 56,0 – 63,00 = – 7,00 | 63,00 – 45,00 = 18,00 | 56,00 – 45,00 = 11,00 |
| Разом: | 79,9 – 90,55 = – 10,65 | 90,55 – 66,60 = 23,95 | 79,9 – 66,60 = 13,30 |

Тепер обчислимо зміну товарообороту по всіх товарах разом в відносних і абсолютних показниках за рахунок зміни:

а) цін:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{79{,}90}{90{,}55} = 0{,}882\ ;$$

$$\Delta_{pq}(p) = \sum p_1 q_1 - \sum p_0 q_1 = 79{,}90 - 90{,}55 = -10{,}65 \text{ тис. грн.}$$

б) фізичного обсягу реалізації:

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0} = \frac{90{,}55}{66{,}60} = 1{,}360\ ;$$

$$\Delta_{pq}(q) = \sum q_1 p_0 - \sum q_0 p_0 = 90{,}55 - 66{,}60 = 23{,}95 \text{ тис. грн.}$$

Перевіримо через взаємозв'язок індексів проведені розрахунки:

$$I_{pq} = I_p \cdot I_q = 0{,}882 \cdot 1{,}360 = 1{,}200\ ;$$

$$\Delta_{pq} = \Delta_{pq}(p) + \Delta_{pq}(q) = -10{,}65 + 23{,}95 = 13{,}30 \text{ тис. грн.}$$

Таким чином, в цілому по всіх видах товарів загальний обсяг товарообороту за рахунок зниження цін на 11,8 % зменшився на 10,65 тис. грн., а за рахунок збільшення обсягу реалізації всіх видів товарів в 1,36 рази, або на 36 % він зріс на 23,95 тис. грн.

Як доводять наші розрахунки, відносний вплив кожного конкретного фактора на результативний показник визначається через обчислення відповідних індексів, а абсолютні зміни – як різниця між чисельником і знаменником цих індексів.

Розклад абсолютних приростів за двома факторами – це найбільш простий спосіб розподілу абсолютного приросту. В дійсності, на практиці, зустрічається дуже багато результативних показників, які залежать від трьох і більше факторів. Наведемо приклад. Нехай маємо дані по одній з галузей промисловості (дані умовні, табл. 12.12).

Таблиця 12.12

| Показники | Період | |
|---|---|---|
| | базисний | звітний |
| Валова продукція в порівняльних цінах, млн. грн. | 40000 | 46000 |
| Матеріальні затрати в порівняльних цінах, млн. грн. | 20000 | 21160 |
| Відпрацьований час працівниками галузі, млн. люд.-дн. | 4000 | 4100 |

Визначимо спочатку приріст валової продукції за двома факторами:

а) за рахунок збільшення відпрацьованого часу:

$$\dot{I}_T = \frac{\sum T_1 W_0}{\sum T_0 W_0},$$

$$\Delta_q(T) = \sum T_1 W_0 - \sum T_0 W_0 = \sum (T_1 - T_0) \cdot W_0 = \Delta T \cdot W_0 =$$

$$= (4100 - 4000) \cdot \frac{40000}{4000} = 100 \cdot 10 = 1000 \text{ млн. грн.}$$

б) за рахунок росту продуктивності праці:

$$\dot{I}_W = \frac{\sum W_1 T_1}{\sum W_0 T_1},$$

$$\Delta_q(W) = \sum W_1 T_1 - \sum W_0 T_1 = \sum (W_1 - W_0) \cdot T_1 = \Delta W \cdot T_1 =$$

$$= (11{,}219512 - 10) \cdot 4100 = 5000 \text{ млн. грн.}$$

Загальний приріст валової продукції дорівнює:

$$\dot{I}_q = \frac{\sum W_1 T_1}{\sum W_0 T_0},$$

$$\Delta_q = \sum W_1 T_1 - \sum W_0 T_0 = \sum q_1 - \sum q_0 = 46000 - 40000 =$$

$$= 6000 \text{ млн. грн.}$$

$$\Delta_q = \Delta_q(T) + \Delta_q(W) = 1000 + 5000 = 6000 \text{ млн. грн.}$$

Далі визначимо чисту продукції і її зміну в звітному періоді порівняно з базисним за рахунок трьох факторів.

$$ЧП_0 = ВП_0 - МЗ_0 = 40000 - 20000 = 20000 \text{ млн. грн.}$$

$$ЧП_1 = ЧП_1 - ЧП_0 = 24840 - 20000 = 4840 \text{ млн. грн.}$$

Зміна чистої продукції за рахунок:

а) відпрацьованого часу:

$$\dot{I}_T = \frac{T_1}{T_0} = \frac{4100}{4100} = 1{,}025,$$

$$\Delta_{ЧП}(T) = ЧП_0(\dot{I}_T - 1) = 2000 \cdot (1{,}025 - 1) = 500 \text{ млн. грн.}$$

б) продуктивності праці:

$$\dot{I}_W = \frac{ВП_1}{T_1} : \frac{ВП_0}{T_0} = \frac{46000}{4100} = \frac{40000}{4000} = \frac{11{,}219512}{10{,}000000} = 1{,}122,$$

$$\Delta_{ЧП}(W) = ЧП_0 \cdot \dot{I}_T \cdot (\dot{I}_W - 1) =$$
$$= 20000 \cdot 1{,}025(1{,}122 - 1) = 2500 \text{ млн. грн.}$$

в) частки чистої продукції у валовій продукції:

$$\dot{I}_d = \frac{ЧП_1}{ВП_1} : \frac{ЧП_0}{ВП_0} = \frac{24840}{46000} : \frac{20000}{40000} = \frac{0{,}54}{0{,}50} = 1{,}08,$$

$$\Delta_{ЧП}(d) = ЧП_0 \cdot I_T \cdot \dot{I}_W(\dot{I}_d - 1) =$$
$$= 20000 \cdot 1{,}025 \cdot 1{,}122 \cdot (1{,}08 - 1) = 1840,$$

або $\Delta_{ЧП}(d) = ВП_1 \cdot \Delta_{ЧП} = 46000 \cdot 0{,}04 = 1840$ млн. грн.

Проведемо перевірку зроблених нами розрахунків.

$$I_{ЧП} = \frac{ЧП_1}{ЧП_0} = \frac{24840}{20000} = 1{,}242.$$

$$I_{ЧП} = I_T \cdot I_W \cdot I_d = 1{,}025 \cdot 1{,}122 \cdot 1{,}08 = 1{,}242.$$

$$\Delta_{ЧП} = \Delta_{ЧП}(T) + \Delta_{ЧП}(W) + \Delta_{ЧП}(d) = 500 + 2500 + 1840 = 4840 \text{ млн. грн.}$$

Отже, приріст чистої продукції за факторами розкладений правильно.

Є і інший метод визначення впливу факторів на приріст чистої продукції, наприклад:

а) $\Delta_{ЧП}(T) = (T_1 - T_0) \cdot W_0 d_0 = (4100 - 4000) \cdot 10 \cdot 0{,}5 = 500$ млн. грн.;

б) $\Delta_{ЧП}(W) = T_1(W_1 - W_0) \cdot d_0 = 4100(11{,}219512 - 10) \cdot 0{,}5 = 2500$ млн. грн.;

в) $\Delta_{ЧП}(d) = T_1 \cdot W_1 \cdot (d_1 - d_0) = 4100 \cdot 11{,}219512(0{,}54 - 0{,}5) = 1840$ млн. грн..

Розглянемо ще один приклад розкладу складного результативного показника за факторами. Вартість матеріальних затрат, зв’язаних з виробництвом любої продукції на підприємстві, залежить від кількості виробленої продукції, питомих витрат сировини і цін на цю сировину. Отже, досліджуючи динаміку цього складного показника, важливо правильно побудувати систему взаємозв'язаних індексів і розкласти абсолютний приріст за факторами.

При визначенні впливу одного із трьох факторів на величину результативного показника, доводиться фіксувати в якості незмінних ваг два фактори, кожний з яких може фіксуватися на різному рівні.

В даному випадку факторні індекси повинні бути побудовані таким чином, щоб вони були взаємозв'язані в єдиній системі.

Якщо кількість виготовленої продукції позначити через "q, питомі витрати через "m", а ціни на матеріали через "p", то динаміку результативного показника – індекс загальної вартості матеріальних витрат обчислюють за формулою:

$$\dot{I}_{qmp} = \frac{\sum q_1 m_1 p_1}{\sum q_0 m_0 p_0}.$$

Різниця між чисельником і знаменником покаже загальну абсолютну зміну витрат в звітному періоді порівняно з базисним:

$$\Delta_{qmp} = \sum q_1 m_1 p_1 - \sum q_0 m_0 p_0$$

Система взаємопов'язаних індексів для розкладу загальної суми абсолютного приросту витрат матиме вигляд:

$$\dot{I}_{qmp} = \dot{I}_q \cdot \dot{I}_m \cdot \dot{I}_p = \frac{\sum q_1 m_1 p_1}{\sum q_0 m_0 p_0} =$$

$$= \frac{\sum q_1 m_0 p_0}{\sum q_0 m_0 p_0} \cdot \frac{\sum q_1 m_1 p_0}{\sum q_1 m_0 p_0} \cdot \frac{\sum q_1 m_1 p_1}{\sum q_1 m_1 p_0},$$

де $\dot{I}_q$ – індекс фізичного обсягу побудований при зафіксованих на базисному рівні питомих витрат матеріалів і цінах;

$\dot{I}_m$ – індекс питомих витрат матеріалів, при зафіксованих цінах на рівні базисного періоду, а фізичного обсягу – на рівні звітного періоду;

$\dot{I}_p$ – індекс цін при зафіксованих обсягах продукції і питомих витратах на рівні звітного періоду.

Віднімаючи від чисельника кожної формули її знаменник, отримаємо суму абсолютної зміни вартості матеріальних витрат за рахунок відповідного фактора.

Проілюструємо це на конкретному прикладі.

Таблиця 12.13

| Вид сировини | Витрати різних видів сировини на виробництво різнорідної продукції | | | | | |
|---|---|---|---|---|---|---|
| | Базисний період | | | Звітний період | | |
| | вироблено виробів, тис. шт. $q_0$ | витрати сировини на 1 виріб, кг $m_0$ | ціна 1 кг сировини, грн. $p_0$ | вироблено виробів, тис. шт. $q_1$ | витрати сировини на 1 виріб, кг $m_1$ | ціна 1 кг сировини, грн. $p_1$ |
| А | 45 | 30 | 10 | 50 | 28 | 10 |
| Б | 70 | 40 | 19 | 80 | 39 | 18 |
| В | 35 | 55 | 25 | 40 | 55 | 24 |

Визначимо загальний індекс витрат на виробництво різнорідної продукції та абсолютну суму приросту вартості матеріальних витрат:

$$\dot{I}_{qmp} = \frac{\sum q_1 m_1 p_1}{\sum q_0 m_0 p_0} = \frac{50 \cdot 28 \cdot 10 + 80 \cdot 39 \cdot 18 + 40 \cdot 55 \cdot 24}{45 \cdot 30 \cdot 10 + 70 \cdot 40 \cdot 19 + 35 \cdot 55 \cdot 25} = \frac{122960}{114825} = 1{,}0708\,,$$

$$\Delta_{qmp} = \sum q_1 m_1 p_1 - \sum q_0 m_0 p_0 = 122960 - 114825 = 8135 \text{ тис. грн.}$$

Розкладемо через систему взаємозв'язаних індексів цей приріст за факторами:

а) фізичним обсягом виробленої продукції:

$$\dot{I}_q = \frac{\sum q_1 m_0 p_0}{\sum q_0 m_0 p_0} = \frac{50 \cdot 30 \cdot 10 + 80 \cdot 40 \cdot 19 + 40 \cdot 55 \cdot 25}{114825} = \frac{130800}{114825} = 1{,}1391\,,$$

$$\Delta_{qmp}(q) = \sum q_1 m_0 p_0 - \sum q_0 m_0 p_0 = 130800 - 114825 = 15975 \text{ тис. грн.}$$

б) питомими витратами матеріалів:

$$\dot{I}_m = \frac{\sum q_1 m_1 p_0}{\sum q_1 m_0 p_0} = \frac{50 \cdot 28 \cdot 10 + 80 \cdot 39 \cdot 19 + 40 \cdot 55 \cdot 25}{130800} = \frac{128280}{130800} = 0{,}9807\,,$$

$$\Delta_{qmp}(m) = \sum q_1 m_1 p_0 - \sum q_1 m_0 p_0 = 128280 - 130800 = -2520 \text{ тис. грн.}$$

в) змінами цін:

$$\dot{I}_p = \frac{\sum q_1 m_1 p_1}{\sum q_1 m_1 p_0} = \frac{122960}{128280} = 0{,}9585\,,$$

$$\Delta_{qmp}(p) = \sum q_1 m_1 p_1 - \sum q_1 m_1 p_0 = 122960 - 128280 = -5320 \text{ тис. грн.}$$

Перевіримо проведені розрахунки через взаємозв'язок:

$$\dot{I}_{qmp} = I_q \cdot I_m \cdot I_p = 1{,}0708 = 1{,}1391 \cdot 0{,}9807 \cdot 0{,}9585\,,$$

$$\Delta_{qmp} = \Delta_{qmp}(q) + \Delta_{qmp}(m) + \Delta_{qmp}(p) = 15975 - 2520 - 5320 = 8123 \text{ тис. грн.}$$

Таким чином, загальні витрати на виробництво продукції в звідному періоді порівняно з базисним зросли в 1,0708 рази, або на

7,08 %, що дало в абсолютній сумі 8135 тис. грн. приросту, в тому числі за рахунок: збільшення виробництва продукції в 1,1391 рази (на 13,91 %) приріст склав 15975 тис. грн., зменшення питомих витрат матеріалів на 1,93 % приріст знизився на 2520 тис. грн. і зниження цін на матеріали на 4,15 % – приріст витрат зменшився на 5320 тис. грн.

В цілому прийом абсолютних приростів складного результативного показника за факторами через систему взаємопов'язаних індексів є дуже важливим в статистичному аналізі.

### *Контрольні запитання*

1. Поняття про статистичні індекси.
2. Види статистичних індексів.
3. Які ви знаєте сфери застосування індексів?
4. Як класифікуються індекси?
5. Як діляться індекси за характером досліджуваних об'єктів?
6. Класифікація індексів за ступенем охоплення елементів сукупності?
7. Що показують індивідуальні індекси?
8. Що називають індексованою величиною?
9. Що характеризують загальні індекси?
10. Поняття про групові індекси.
11. Базисні і ланцюгові індекси.
12. Що являє собою агрегатний індекс?
13. Правило побудови агрегатних індексів.
14. Як визначається загальний індекс цін?
15. Як визначається загальний індекс фізичного обсягу продукції?
16. Що характеризує загальний індекс товарообороту?
17. Взаємозв'язок індексів.
18. Загальний індекс собівартості, фізичного обсягу та затрат на виробництво продукції, їх взаємозв'язок.
19. Поняття про середньозважені індекси.
20. Середній арифметичний індекс фізичного обсягу.
21. Середній гармонічний індекс цін і собівартості.
22. Які індекси називаються базисними, а які ланцюговими?
23. Базисні і ланцюгові індекси з постійними і змінними вагами.
24. Що характеризують індекси змінного складу?
25. Що характеризують індекси постійного (фіксованого) складу?
26. Для чого обчислюють індекси структурних зрушень?
27. Як зв'язані між собою індекси змінного складу, постійного складу і структурних зрушень?
28. Поняття про територіальні індекси.

29. Використання системи взаємозв’язаних індексів в факторному аналізі.
30. Методи розкладу абсолютного приросту за факторами.
31. Які фактори впливають на загальну зміну товарообороту в фактичних цінах?
32. Які фактори впливають на загальну зміну обсягу валової продукції?
33. Як розкласти абсолютний приріст валової продукції за факторами?
34. Як розкладається абсолютний приріст чистої продукції за факторами?
35. Які індекси входять в систему загального індекса матеріальних витрат на виробництво продукції?
36. Розкладіть абсолютну суму приросту матеріальних витрат за факторами.

# Розділ 13. Вибіркове спостереження

## 13.1. Поняття про вибіркове спостереження та його основні завдання

Із всіх видів несуцільного спостереження в практиці статистичних досліджень найбільше визнання і застосування отримало вибіркове спостереження.

*Вибірковим спостереженням* називається такий вид несуцільного спостереження, за характеристикою відібраної частини одиниць якого судять про всю сукупність.

Розрізняють генеральну і вибіркову сукупності.

*Генеральною сукупністю* називається така маса одиниць, з якої проводиться відбір для дослідження. Частина генеральної сукупності відібрана для обстеження називається *вибірковою*. Обсяг генеральної сукупності позначають через (N), а вибіркової – через (n). Узагальнюючими показниками генеральної сукупності є: середній розмір ознаки ($\bar{x}$), доля (P), генеральна дисперсія ($\sigma^2$). Розрізняють також наступні узагальнюючі показники вибіркової сукупності: середню вибіркову ($\tilde{x}$), вибіркову долю (w) і дисперсію ($\sigma_в^2$).

Вибірковий метод відрізняється від інших видів несуцільного спостереження за двома ознаками: 1) заздалегідь встановлюється, яка частина одиниць генеральної сукупності буде обстежуватись, і 2) наперед визначається порядок відбору одиниць, який забезпечить в достатній мірі репрезентативність генеральної сукупності.

До вибіркового спостереження статистика вдається у випадках, коли потрібно зекономити сили і засоби при проведенні дослідження, тобто коли недоцільно або неможливо проводити суцільне спостереження.

Існує ряд причин, в силу яких у багатьох випадках вибірковому спостереженню надається перевага перед суцільним. Серед них найбільш суттєві це: економія часу і засобів в результаті скорочення обсягу робіт статистичного дослідження; зведення до мінімуму псуття або знищення досліджуваних об’єктів; можливість більш детального вивчення кожної одиниці спостереження при неможливості охоплення всіх одиниць; досягнення високої точності результатів обстеження за рахунок скорочення помилок результатів.

Вибіркове спостереження застосовують також у поєднанні із суцільним для поглиблення дослідження, або для уточнення і контролю результатів суцільного спостереження.

Вибіркове спостереження складається з таких етапів:

1) постановка мети спостереження;

2) складання програми спостереження і розробка відповідних даних;
3) вирішення організаційних питань проведення спостереження;
4) визначення відсотка і способу відбору одиниць;
5) проведення відбору;
6) реєстрація відповідних ознак у відібраних для дослідження одиниць;
7) узагальнення даних спостереження та розрахунок їх вибіркових характеристик;
8) знаходження помилок вибірки;
9) перерахунок характеристик вибіркового спостереження на всю сукупність.

Вибіркове спостереження проводиться для вирішення наступних основних завдань:

1) визначення середнього розміру досліджуваної ознаки;
2) визначення питомої ваги (частки);
3) визначення середньої і граничної помилки вибірки;
4) знаходження меж для середньої і частки при повторному і безповторному відборі;
5) визначення потрібної чисельності вибірки;
6) поширення даних вибіркового спостереження на всю сукупність.

### 13.2. Основні умови наукової організації вибіркового спостереження

Науковим обґрунтуванням можливості застосування вибіркового спостереження виступає діалектична єдність одиничного, особливого і загального, згідно якої в кожному одиничному є риси особливого і загального, а загальне володіє рисами одиничного і особливого. Ця єдність дозволяє за одиничним і особливим судити про загальне, за частиною – про ціле, якщо правильно встановлений зв'язок між ними.

Особливістю вибіркового спостереження в порівнянні з іншими видами несуцільного спостереження є те, що при відборі одиниць у вибіркову сукупність забезпечується рівна можливість попадання кожної одиниці у вибірку. Це досягається шляхом неупередженого строгого випадкового відбору за схемою, розробленою математичною статистикою.

Відповідь на питання про те, яка за розміром різниця між генеральними і вибірковими узагальнюючими показниками, з якою ймовірністю можна судити про цю різницю, дає теорія вибіркового методу, на основі закону великих чисел. За допомогою цього закону розв'язують два взаємозв'язаних завдання: 1) розраховують, із заданою ймовірністю, межі можливих відхилень вибіркового показника від відповідного показника в генеральній сукупності; 2) визначають ймовірність того, що розмір можливих відхилень вибіркового показника від генерального не перевищить встановленої межі.

Масові соціально-економічні явища піддаються впливу великого числа випадкових чинників, тому, вивчаючи їх, статистика використовує основний висновок граничних теорем теорії ймовірностей, який полягає в тому, що сукупна дія великої кількості випадкових факторів призводить, за деякої умови, до результату, майже незалежного від випадку. Так як вибіркове спостереження зв'язане із випадковими відхиленнями даних вибіркової і генеральної сукупностей, основне положення граничних теорем дозволяє стверджувати, що результати вибіркового спостереження достовірні, при достатньо великій кількості відібраних одиниць. Тоді вибіркові характеристики надійно відтворять генеральні характеристики.

При масовому спостереженні, розподіл емпіричних частот більшості явищ підпорядковується закону нормального розподілу. Любий розподіл частот в генеральній сукупності буде характеризуватись вибірковими середніми за їх розподілом, близькими до нормального.

Доведено, що за нормальним розподілом більша частина величин зосереджена навколо генеральної середньої. Біля 68,3% чисельності вибіркової середньої буде знаходитись в межах $\pm\ \sigma$ генеральної середньої; 95,4% цієї чисельності знаходитиметься в межах $\pm\ 2\sigma$ і 99,7 % - не вийде за межі $\pm\ 3\sigma$. Нормальний розподіл показує частоту виникнення помилок даного розміру середньої.

Випадкові помилки реєстрації при великому числі спостережень не впливають суттєво на результат дослідження, так як вони взаємно погашаються, а тому від них можна абстрагуватись і в подальшому розглядати тільки помилки вибірки.

Принцип строгої випадковості, який покладений в основу вибірки, забезпечує його об'єктивність, дозволяє встановити межі можливих помилок і отримати практично достовірні дані для характеристики всієї сукупності явищ. Така вибіркова сукупність називається *представницькою або репрезентативною*. В її склад входять представники всіх груп, з яких складається генеральна сукупність.

Точність результатів вибіркового спостереження, в кінцевому підсумку, буде залежати від способу відбору одиниць, ступеня коливання ознаки в сукупності та від числа одиниць, що їх спостерігатимуть.

### 13.3. Методи і способи відбору одиниць у вибіркову сукупність

Відбір одиниць із генеральної сукупності у вибіркову можна проводити по-різному, в залежності від багатьох умов. *Способом відбору* називається система організації відбору одиниць із генеральної сукупності.

Розрізняють два методи відбору одиниць у вибіркову сукупність: повторний і безповторний.

*Повторним* називається такий метод відбору, при якому кожна раніше відібрана одиниця повертається у генеральну сукупність і може знову брати участь у вибірці.

*Безповторним* називається такий метод відбору, при якому кожна раніше відібрана одиниця не повертається в генеральну сукупність і в подальшій вибірці участі не бере.

Оскільки безповторний відбір охоплює постійно нові одиниці сукупності, а повторний – одну і ту ж сукупність, тому безповторний відбір дає більш точні результати.

Повторний і безповторний методи відбору, в залежності від характеру одиниці відбору, застосовуються в поєднанні з іншими видами відбору. В практиці статистичних досліджень використовуються три види відбору:

1) індивідуальний – відбір окремих одиниць сукупності;
2) груповий (серійний) – відбір груп (серій) одиниць;
3) комбінований – комбінація індивідуального і групового.

Різні види відбору здійснюється різними способами проведення вибірки. За способом відбору одиниць для обстеження розрізняють такі види вибіркового спостереження:

1) власне випадкова вибірка;
2) механічна вибірка;
3) типова (районована) вибірка;
4) серійна (гніздова) вибірка;
5) комбінована вибірка;
6) одноступінчаста і багатоступінчаста вибірка;
7) однофазна і багатофазна вибірка;
8) інші види вибірок.

*Власне випадковою* називається така вибірка, при якій відбір одиниць з усієї генеральної сукупності є випадковим. Часто для цього застосовують жеребкування або таблицю випадкових чисел. Ця вибірка дає добрі результати за умови, коли між одиницями досліджуваної сукупності немає значних відхилень, тобто при однорідній сукупності.

Проведення власне випадкової вибірки вимагає достовірної інформації про всі одиниці генеральної сукупності, збирання якої досить важке. Під час організації і проведення жеребкування одиниці сукупності можуть заново виникати або ліквідовуватись, а при дослідженні якості продукції, яке проводиться безперервно протягом робочого дня, статистик ніколи не має вичерпної інформації про кількість одиниць. Незручність власне випадкової вибірки ще й в тому, що на всі одиниці генеральної сукупності потрібно заготовити жеребки, що також вимагає додаткових затрат коштів і часу.

При всій своїй простоті власне випадкова вибірка, в силу викладених обставин, на практиці використовується рідко.

В практиці вибіркового спостереження найбільше розповсюдження має, як різновидність власне випадкової, механічна вибірка.

*Механічна вибірка* – це послідовний відбір одиниць через рівні проміжки в порядку визначеного розположення їх в генеральній сукупності, або в якому-небудь переліку. Інтервали відбору визначаються у відповідності з питомою вагою відбору одиниць (кожна п'ята, десята, сота і т.д.).

Механічний відбір здійснюють за списками або безпосередньо на місці за природнім розміщенням одиниць генеральної сукупності.

Розміщення одиниць генеральної сукупності в списку або на місці спостереження буває впорядкованим або не впорядкованим відносно досліджуваної ознаки. Спосіб розміщення одиниць генеральної сукупності впливає на порядок їх відбору у вибіркову сукупність. Якщо в генеральній сукупності одиниці розміщені не впорядковано, тоді із першого інтервалу (наприклад, в десять одиниць) береться люба, а далі послідовно відбирається кожна десята одиниця. У випадку коли розположення впорядковане (ранговане), у вибіркову сукупність потрібно відбирати одиниці, розміщені посередині кожного інтервалу, тому що може виникнути систематична помилка вибірки.

Механічний відбір із впорядкованої сукупності називають ще систематичним відбором.

Механічна вибірка завжди безповторна.

В статистичній практиці найбільш часто використовують типову або районовану вибірку. При *типовому відборі* генеральну сукупність спочатку поділяють на однорідні групи за певною ознакою, райони, зони. Потім з кожної групи випадковим або механічним способом відбирають певну кількість одиниць, пропорційно питомій вазі групи в загальній сукупності.

Типова вибірка має ті переваги перед попередніми, що до вибірки потрапляють представники всіх типових груп, внаслідок чого вибірка стає більш достовірною. Використання цієї вибірки найбільш придатне при неоднорідності одиниць досліджувальної сукупності.

Досить часто в практиці вибіркового спостереження застосовують *серійну (гніздову) вибірку*. При серійній (гніздовій) вибірці відбір одиниць проводять цілими групами (серіями, гніздами) сукупності, в межах яких обстежують всі одиниці без винятку. Серії для спостереження відбирають випадково, частіше безповторним способом механічної вибірки.

Серійний відбір значно простіший в організаційному відношенні і дешевший за інші. Однак, отримана в процесі цього відбору випадкова помилка вибірки може бути більша, ніж при інших способах відбору. Для зменшення випадкової помилки серійної вибірки збільшують її обсяг.

В статистичній практиці вибіркового спостереження часто комбінують два або кілька видів вибірок. Таку вибірку називають *комбінованою*. Перш за все, можна комбінувати суцільне і вибіркове спостереження. В даному випадку, за основною програмою обстежується вся генеральна сукупність, а за додатковою – вибіркова сукупність. Наприклад, перепис всього населення проводять за програмою, яка складається з 11 питань, а 25% із нього – по бланку з 18 питань (11 основних і 7 додаткових). Цінність такого спостереження заключається в тому, що дані вибіркового спостереження тут тісно пов'язані з даними суцільного спостереження, що робить їх більш доказовими і значущими.

Можна комбінувати також серійну вибірку з власне випадковою. В цьому випадку спочатку розбивають генеральну сукупність на серії, відбирають потрібну кількість серій і проводять випадковий відбір одиниць з кожної серії.

Середня помилка вибірки при різних комбінаціях її видів і способів визначається по-різному, в залежності від ступінчастості вибірки.

В залежності від того, як змінюється одиниця відбору, при послідовному проведенні ряду вибірок розрізняють одноступінчастий і багатоступінчастий відбір одиниць в досліджувану сукупність.

Вибірка, коли із досліджуваної сукупності зразу відбираються одиниці або серії одиниць для безпосереднього обстеження, називається *одноступінчастою*.

*Багатоступінчаста* вибірка передбачає поступове вилучення із генеральної сукупності спочатку укрупнених груп одиниць, потім груп, менших за об'ємом, і так до тих пір, поки не відберуть відповідні групи або окремі одиниці, які і будуть досліджуватись. Вибірка може бути двох-, трьох і більше ступінчастою. Однак, потрібно уникати великого числа ступенів, детально продумуючи саму організацію вибіркового спостереження.

В багатоступінчастому відборі поєднуються різні способи відбору. Наприклад, при бюджетних обстеженнях сімей може застосовуватись чотирьохступінчаста вибірка. Так, відбір населених пунктів може проводитись типовим способом, вулиці в населених пунктах – випадковим, будинки – механічним, а конкретна сім'я – знову випадковим. Як видно із прикладу, в багатоступінчастій вибірці одиниця відбору на кожній ступені вибірки різні, а обстежують тільки одиниці, відібрані на останній ступені.

Якщо необхідні дані можна отримати на основі вивчення всіх первинно відібраних одиниць, застосовують *однофазну вибірку*, а якщо тільки на основі деякої їх частини, відібраної так, що вона складає підвибірку із початково проведеної вибірки – *багатофазну*.

*Багатофазною* називається така вибірка, коли одні відомості збираються від всіх одиниць відбору, потім відбираються ще деякі одиниці і обстежуються за більш широкою програмою. Вибірка може бути двохфазна, трьохфазна, чотирьохфазна і т.д. Вона може служити основою для дослідження сукупності за різними програмами.

Багатофазна вибірка відрізняється від багатоступінчастої тим, що при багатофазній вибірці на кожній фазі проводиться обстеження відібраних одиниць і завжди зберігається одна і таж одиниця відбору. Прикладом може служити відбір кращих студентів для участі в міжвузівській олімпіаді по статистиці. Спочатку з потоку відбирають кращих студентів і дають їм контрольну роботу з предмету. Студенти, які виконали першу контрольну роботу на відмінно, потрапляють в наступну фазу і виконують другу, більш складнішу контрольну роботу, і так аж до конкурсних завдань. Ці студенти, які пройшли всі фази відбору потрапляють на олімпіаду. Як бачимо на всіх фазах відбору і обстеження залишається одна і таж одиниця відбору – студент.

Важливою властивістю багатофазної вибірки є те, що вона дає можливість використовувати дані, зібрані на першій фазі, для додаткової характеристики і уточнення розрахунків, отриманих в наступних фазах дослідження.

Багатофазна вибірка може поєднуватись з багатоступінчастою, де кожна окрема ступінь багатоступінчастої вибірки може бути багатофазною.

Розрахунок помилок репрезентативності багатоступінчастої і багатофазної вибірок проводиться для кожної ступені і фази окремо. Визначення цих помилок є одним із складних завдань статистики.

Простота і одноманітність методики відбору досить бажані при проведені вибіркового спостереження. Але бувають випадки, коли необхідно застосувати інші способи відбору, такі як взаємопроникаючі і квантильні вибірки, направлений відбір, моментні спостереження, або скористатись малою вибіркою.

*Взаємопроникаючою* називається така вибірка, коли із однієї генеральної сукупності проводять одним і тим же способом декілька незалежних вибірок.

Взаємопроникаючі вибірки завжди проводять різні, незалежні один від одного дослідники, що дозволяє порівнювати підсумки по всіх частинах і забезпечити взаємну перевірку їх роботи. Взаємопроникаючі вибірки дають незалежні одна від одної оцінки значень досліджуваної сукупності, і, якщо результати різних вибірок близькі між собою, то такі оцінки дуже переконливі.

Переваги взаємопроникаючих вибірок стають більш рельєфними, якщо кожна з них є репрезентативною вибіркою із генеральної сукупності. Взаємопроникаюча вибірка зручна для порівняння результатів дослідження по окремих географічних районах.

До недоліків взаємопроникаючої вибірки можна віднести порівняно великі видатки на їх проведення, тому що однаковою за величиною частиною генеральної сукупності будуть займатись декілька дослідників, часто в різних районах.

Помилки взаємопроникаючих вибірок визначаються за формулами типової пропорційної вибірки.

*Квантильні вибірки* застосовують тоді, коли виникає потреба дослідження даних суцільного спостереження за додатковою програмою.

Для проведення квантильної вибірки рангують потрібну варіаційну ознаку і за її нагромадженими частотами будують огіву. За огівою механічним способом відбирають потрібну частину одиниць для дослідження цієї ж ознаки. Якщо огіва вибіркової сукупності добре відтворює огіву генеральної сукупності, то помилка репрезентативності буде мінімальною.

При застосуванні квантильних вибірок, помилки вибірки не вираховуються, а визначаються відхилення огів, побудованих за даними вибіркової і генеральної сукупностей.

Якщо середня квадратична помилка, визначена з різниці групових нагромаджених частот вибіркового спостереження, за величиною мала, то вибіркове спостереження добре відтворює дані генеральної сукупності.

Коли при оцінці відповідності між вибірковими даними і даними суцільного спостереження лінійний коефіцієнт кореляції наближається до одиниці, вибірку вважають репрезентативною.

Направлений відбір використовують тоді, коли за відомим середнім значенням ознаки в генеральній сукупності вибіркова сукупність повинна характеризувати її структуру за іншими ознаками.

*Направлений відбір* передбачає проведення відбору таким чином, щоб середній розмір відібраних одиниць дорівнював середньому розміру одиниць всієї сукупності.

Формування вибіркової сукупності проводять шляхом заміни одиниць сукупності, які значно відхиляються за своїми розмірами від середньої, іншими, відібраними з генеральної сукупності випадковим або типовим способами. При відсутності рівності середніх генеральних і вибіркових сукупностей процес заміни повторюється до тих пір, доки не буде досягнено врівноваженості. В тому випадку, коли заміна однієї одиниці іншою призводить до наближеної рівності середніх генеральної і вибіркової сукупностей, вибірку вважають врівноваженою і репрезентативною за всіма іншими ознаками сукупності.

Якщо врівноваження застосовується в типовому відборі, тоді врівноважується або вся сукупність, або кожна типова група зокрема.

Врівноваження можна проводити також за якісною ознакою. Наприклад, коли вивчають населення за місцем проживання (місто чи село), де одиницею відбору є адміністративний район з певним числом жителів, вибірку врівноважують за відсотковим співвідношенням цих двох груп населення.

Таким чином, *направленим відбором* називається врівноваження за однією ознакою для вибіркового дослідження інших ознак.

Помилку вибірки направленого відбору визначають в залежності від способу проведення відбору одиниць до врівноваження.

Особливим видом вибіркового спостереження є моментне спостереження, суть якого в тому, що на визначені моменти часу фіксується наявність окремих елементів досліджуваного явища. Цей вид вибіркового спостереження отримав широке розповсюдження в усіх галузях народного господарства.

*Моментне спостереження* використовується для вивчення використання робочого часу робітниками або часу роботи устаткування. В кожний момент спостереження фіксують, чи знаходився робітник або станок в роботі, а якщо ні, то з яких причин. Моментне спостереження охоплює роботу всіх робітників або станків цеху, а тому в цьому відношенні, воно є суцільним. Вибірковим його вважають через те, що охоплює не весь час роботи цеху, а лише визначені моменти часу, коли здійснюється контроль за роботою робітників або обладнання.

За допомогою моментного спостереження отримують потрібну інформацію скоріше і з меншими затратами, як при суцільному спостереженні.

Оцінка репрезентативності моментного спостереження здійснюється за формулами середньої і граничної помилки, вибіркової частки. Число записів моментного обстеження виступає в якості вибіркової сукупності.

Вибіркова сукупність, яка складається з порівняно невеликої кількості одиниць (20-30) називається *малою вибіркою*. Як відомо, із збільшенням вибіркової сукупності підвищується точність вибіркових даних. Проте, іноді доводиться обмежуватись малою кількістю спостережень (при перевірці якості продукції, зв'язаної із знищенням продукції, яку перевіряють). Математичною статистикою доведено, що і при малих вибірках характеристики вибіркової сукупності можна поширити на генеральну. Але розрахунок середньої і граничної помилок тут мають свої особливості.

При малих вибірках дисперсію обчислюють з врахуванням кількості ступенів вільності варіації.

Ступенем вільності варіації називають кількість одиниць, здатних змінюватись після того, як по всіх одиницях вибірки була визначена їх загальна характеристика.

Англійський вчений Стьюдент винайшов закон розподілу відхилень вибіркових середніх від генеральної середньої для малих вибірок. Опираючись на цей закон, він склав спеціальні таблиці, в яких наводяться значення критерію t для малих вибірок.

Згідно розподілу Стьюдента, ймовірна оцінка того, що гранична помилка не перевищить t-кратну середню помилку в малих вибірках, залежить і від величини t, і від чисельності вибірки.

На практиці малі вибірки застосовуються, головним чином, для оцінки суттєвості t відмінностей двох вибіркових середніх.

## 13.4. Знаходження середньої і граничної помилок та необхідної чисельності для різних видів вибірок

Для вибіркового спостереження властиві помилки реєстрацій і помилки репрезентативності. Помилки реєстрацій можуть виникати при усякому статистичному спостереженні. Помилки репрезентативності властиві тільки вибірковому спостереженню. Ці помилки виникають при дотриманні всіх правил проведення вибіркового спостереження, але при належній організації вибірки їх можна довести до зовсім малих розмірів.

*Помилки репрезентативності* становлять різницю між середніми і відносними показниками вибіркової сукупності та відповідними показниками генеральної сукупності. Вони поділяються на систематичні і випадкові.

*Систематичні помилки* репрезентативності виникають внаслідок порушення принципів проведення вибіркового спостереження.

Вони мають тенденційний характер викривлення величини досліджуваної ознаки в сторону її збільшення або зменшення.

*Випадкові помилки* репрезентативності зумовлені тим, що вибіркова сукупність не відображає точно середні і відносні показники генеральної сукупності.

Визначення величини випадкових помилок репрезентативності є одним з головних завдань теорії вибіркового методу.

Для узагальнюючої характеристики помилки вибірки вираховують середню помилку репрезентативності, яку позначають через грецьку букву “мю” ($\mu$) і називають ще стандартом.

Для визначення середньої помилки репрезентативності власне випадкової і механічної вибірки застосовують чотири формули для повторного і безповторного відбору.

Таблиця 13.1

| Спосіб відбору | При визначенні середньої | При визначенні частки |
|---|---|---|
| Повторний | $\mu = \sqrt{\frac{\sigma^2}{n}}$ | $\mu = \sqrt{\frac{W(1-W)}{n}}$ |
| Безповторний | $\mu = \sqrt{\frac{\sigma^2}{n}(1-\frac{n}{N})}$ | $\mu = \sqrt{\frac{W(1-W)}{n}(1-\frac{n}{N})}$ |

де $\mu$ - середня помилка репрезентативності;
$\sigma^2$ – середній квадрат відхилень у вибірці;

n – чисельність вибіркової сукупності;

N – чисельність генеральної сукупності;

n/N – частка обстеженої частини вибіркової сукупності;

1 – n/N – необстежена частка генеральної сукупності;

W – частка даної ознаки у вибірці;

1 – W – частка протилежної ознаки у вибірці.

На практиці частіше використовують безповторний відбір, який гарантує більш точні результати, оскільки при цьому відборі виключається можливість повторного обстеження одних і тих же одиниць генеральної сукупності.

Для узагальнюючої характеристики помилки вибірки поряд із середньою розраховують ще і граничну помилку вибірки. Стверджувати, що дана генеральна середня не вийде за межі середньої помилки вибірки можна тільки з визначеною ступінню ймовірності.

При вибірковому спостереженні розмір граничної помилки репрезентативності (Δ) може бути більший, дорівнювати, або менший від середньої помилки репрезентативності (μ). Тому величину граничної помилки репрезентативності обчислюють з певною ймовірністю (P), якій відповідає t-разове значення μ. З введенням показника кратності помилки t, формула граничної помилки репрезентативності матиме вигляд:

$$\Delta = t\mu ; \qquad t = \frac{\Delta}{\mu} ,$$

де μ – середня помилка вибірки;

t – коефіцієнт довір'я, який залежить від ймовірності визначення граничної помилки.

Формула граничної помилки вибірки випливає із основних положень теорії вибіркового методу, сформульованих в деяких теоремах ймовірностей, які відображають закон великих чисел.

Однією з основних теорем, які лягли в основу теорії вибіркового методу, є теорема П.Л. Чебишева, в якій він довів, що з ймовірністю скільки завгодно близькою до одиниці, можна стверджувати, що при достатньо великому числі незалежних спостережень, вибіркова середня буде скільки завгодно мало відрізнятись від генеральної середньої, тобто при проведенні повторної вибірки.

Академік А.А. Марков довів збереження цієї умови для залежних спостережень (безповторної вибірки).

Пізніше академік А.М. Ляпунов довів, що ймовірність відхилень вибіркової середньої від генеральної середньої при достатньо великому об'ємі вибірки і обмеженій дисперсії генеральної сукупності підпорядковується закону нормального розподілу. Ймовірність цих відхилень при різних значеннях t визначається за формулою

$$F(t) = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-\frac{t^2}{2}} dt \; .$$

Значення цього інтеграла при різних значеннях t табульовані і приводяться в спеціальних таблицях, наприклад:

для $t = 1$ $P(\Delta \le \mu) = 0{,}683$;

для $t = 2$ $P(\Delta \le \mu) = 0{,}954$;

для $t = 3$ $P(\Delta \le \mu) = 0{,}997$;

для $t = 4$ $P(\Delta \le \mu) = 0{,}999$.

Ці показники означають, що з ймовірністю 0,683 можна стверджувати, що гранична помилка вибірки не перевищить $\mu$, тобто в 68,3% випадків помилка репрезентативності не вийде за межі $\pm\ \mu$. Іншими словами, в 683 випадках із 1000 помилка репрезентативності не перевищить однієї величини середньої помилки. З ймовірністю 0,954 можна стверджувати, що помилка репрезентативності не перевищить $\pm\ 2\mu$, з ймовірністю 0,997 – не перевищить $\pm\ 3\mu$. З ймовірністю 0,999, тобто дуже близькою до одиниці, можна очікувати, що різниця між вибірковою і генеральною середніми не перевищить чотирьохкратної середньої помилки вибірки.

Гранична помилка вибірки розраховується при проведені вибіркового спостереження по-різному, в залежності від видів і способів відбору. Вона дає можливість встановити, в яких межах знаходиться величина генеральної середньої або частки.

Із теореми Чебишева знаходять, що

$$\overline{x} - \tilde{x} = \pm\Delta_x \;\; і \;\; \tilde{x} - \Delta_x \le \overline{x} \le \tilde{x} + \Delta_x \; .$$

Теорема Бернулі розглядає помилку вибірки для альтернативної ознаки. Вона стверджує, що при достатньо великому обсязі вибірки в міру її збільшення ймовірність відхилення між долями ознак у вибірковій і генеральній сукупностях буде наближатися до одиниці. Тобто, з ймовірністю, скільки завгодно близькою до одиниці, можна стверджувати, що при достатньо великому обсязі вибірки вибіркова частка скільки завгодно мало відрізняється від її ймовірності (частки в генеральній сукупності).

Додаючи граничну помилку вибірки до вибіркової частки і віднімаючи її від неї, знаходять межі генеральної долі (частки).

$$P - W = \pm\Delta_p \quad і \quad W - \Delta_p \le P \le W + \Delta_p \; .$$

Формули для розрахунку граничної помилки власне випадкової і механічної вибірки можна записати таким чином (табл. 13.2.).

Таблиця 13.2.

| Спосіб відбору | При визначенні середньої | При визначенні долі |
|---|---|---|
| Повторний | $\Delta_x = t\sqrt{\frac{\sigma^2}{n}}$ | $\Delta_P = t\sqrt{\frac{W(1-W)}{n}}$ |
| Безповторний | $\Delta_x = t\sqrt{\frac{\sigma^2}{n}\left(1-\frac{n}{N}\right)}$ | $\Delta_P = t\sqrt{\frac{W(1-W)}{n}\left(1-\frac{n}{N}\right)}$ |

де $\Delta_x$ – гранична помилка вибірки для середньої;

$\Delta_P$ – гранична помилка вибірки для долі.

На основі формул граничної помилки вибірки розв'язують наступні завдання:

1. Визначають довірчі межі генеральної середньої і частки з прийнятою ймовірністю.

2. Визначають ймовірність того, що відхилення між вибірковими і генеральними характеристиками не перевищать визначену величину.

3. Визначають необхідну чисельність вибірки, яка із заданою ймовірністю забезпечить прийняту точність вибіркових показників.

Розглянемо розв'язок цих завдань для різних видів вибірок на конкретних прикладах.

Для визначення довірчих меж генеральної середньої розраховують вибіркову середню ($\tilde{x}$) і з прийнятою ймовірністю (P) граничну помилку вибірки $\Delta_x = t\mu$. Припустимо, що при 2 % власне випадковому відборі у відібраних для обстеження 100 деталей встановлено, що середня вага однієї деталі 2500 г., дисперсія 900, із 100 деталей 10 виявились бракованими. З ймовірністю 0,954 встановити межі середньої ваги однієї деталі в генеральній сукупності, а з ймовірністю 0,997 – межі частки якісних деталей в генеральній сукупності.

Граничну помилку визначаємо $\Delta_x = t\mu$ за формулою безповторного відбору, так як чисельність генеральної сукупності можна знайти $N = \frac{100 \cdot 100}{2} = 5000$ шт. В спеціальній таблиці знаходимо, що для ймовірності 0,954 $t = 2$, а для ймовірності 0,997 $t = 3$.

Таким чином, отримаємо:

$$\Delta_x = t\sqrt{\frac{\sigma^2}{n}(1-\frac{n}{N})} = 2\sqrt{\frac{900}{100}(1-\frac{100}{5000})} = 2\sqrt{9\cdot 0{,}99} = 2\cdot 3 = 6 \text{ г.}$$

Звідси довірчі межі генеральної середньої будуть наступні:

$$\tilde{x} - \Delta_x \le \bar{x} \le \tilde{x} + \Delta_x,$$

$$2500 - 6 \le \bar{x} \le 2500 + 6,$$

$$2494 \le \bar{x} \le 2506,$$

тобто з ймовірністю 0,954 можна стверджувати, що середня вага однієї деталі в генеральній сукупності знаходиться в межах від 2494 г до 2506 г.

Аналогічно вирішується завдання і при визначенні інтервалів (меж) для генеральної частки. Вибіркова частка якісних деталей:

W = 90 : 100 = 0,9

Гранична помилка для частки буде дорівнювати:

$$\Delta_P = t\sqrt{\frac{w(1-w)}{n}(1-\frac{n}{N})} = 3\sqrt{\frac{0{,}9(1-0{,}9)}{100}(1-\frac{100}{5000})} =$$

$$= 3\sqrt{0{,}0009\cdot 0{,}98} = 3\cdot 0{,}03 = 0{,}09\,.$$

Генеральна частка дорівнюватиме:

$$P = W \pm \Delta_P = W \pm t\mu\,.$$

Довірчі межі генеральної частки будуть виражені наступним чином:

$$W - t\mu \le P \le W + t\mu\,,$$

$$0{,}9 - 0{,}09 \le P \le 0{,}9 + 0{,}09\,,$$

$$0{,}81 \le P \le 0{,}99\,.$$

Таким чином, з ймовірністю 0,997 можна стверджувати, що частка якісних деталей в генеральній сукупності знаходиться в межах від 81 % до 99 %.

При розрахунках вибіркових характеристик інколи ставиться завдання визначення ймовірності допуску певної помилки, тобто відхилення від відповідних характеристик генеральної сукупності не більше ніж на певну завдану величину, яку знаходять за формулою граничної помилки.

Використавши розрахунки попередньої задачі знаходимо для середньої:

$$t = \frac{\Delta_x}{\mu} = \frac{6}{3} = 2\,.$$

Для даного значення t = 2 відповідає ймовірність P = 0,954. Це дає право стверджувати, що при визначенні за вибірковими даними середньої ваги деталей ($\tilde{x} = 2500$ г) допущена помилка, яка не перевищує 6 г.

Аналогічно розраховують ймовірність допуску помилки для частки:

$$t = \frac{\Delta_P}{\mu} = \frac{0{,}09}{0{,}03} = 3\,.$$

Для цього значення t = 3 відповідає ймовірність P = 0,997. Таким чином, майже достовірно можна стверджувати, що при визначенні за вибірковими даними частки якісних деталей (W = 0,9) допущена помилка, яка не перевищує 9 %.

При організації проведення вибіркового спостереження важливе значення має правильне визначення необхідної чисельності (обсягу) вибірки, яка з відповідною ймовірністю забезпечить встановлену точність результатів спостереження. Надмірна чисельність вибірки призводить до затягнення строків дослідження, зайвої втрати сил і коштів, а недостатня – дає результати з великою помилкою репрезентативності. Чисельність вибірки залежить від наступних факторів:

1) від варіації досліджуваної ознаки. Чим більша варіація, тим більшою повинна бути чисельність вибірки, і навпаки;

2) від розміру можливої граничної помилки вибірки. Чим менший розмір можливої помилки, тим більша повинна бути чисельність вибірки. Існує правило, якщо помилку потрібно зменшити в три рази, то чисельність вибірки збільшують в дев'ять раз;

3) від розміру ймовірності, з якою гарантуватимуть результати вибірки. Чим більша ймовірність, тим більша повинна бути чисельність вибірки;

4) від способу відбору одиниць у вибіркову сукупність для обстеження.

Визначення необхідної чисельності вибірки будується на основі алгебраїчного перетворення формул граничної помилки вибірки при різних способах відбору.

Для власне випадкової і механічної вибірки виведення формул необхідної чисельності вибірки проводиться таким чином. З формули граничної помилки вибірки для середньої при повторному відборі $\Delta_x = t\sqrt{\frac{\sigma^2}{n}}$ потрібно визначити чисельність вибірки, тобто n. Для цього обидві частини даного рівняння підносимо до квадрату і отримуємо $\Delta_x^2 = \frac{t^2\sigma^2}{n}$, звідси необхідна чисельність вибірки дорівнює $n = \frac{t^2\sigma^2}{\Delta_x^2}$.

Дана формула є математичним підтвердженням залежності чисельності вибірки від розміру граничної помилки, величини коефіцієнта довір'я t і величини варіації (дисперсії).

Так само виводять формули необхідної чисельності вибірки при вивченні частки ознаки при повторному і безповторному відборі.

В готовому вигляді їх можна записати таким чином (табл. 13.3).

Табл. 13.3

| Спосіб відбору | Чисельність вибірки | |
|---|---|---|
| | при визначенні середньої | при визначенні частки |
| Повторний | $n = \frac{t^2\sigma^2}{\Delta_x^2}$ | $n = \frac{t^2 W(1-W)}{\Delta_P^2}$ |
| Безповторний | $n = \frac{t^2\sigma^2 N}{\Delta_x^2 N + t^2\sigma^2}$ | $n = \frac{t^2 W(1-W)N}{\Delta_P^2 N + t^2 W(1-W)}$ |

Покажемо застосування цих формул на прикладах. Припустимо, що для району, в якому є 8000 корів, необхідно організувати вибіркове обстеження з метою встановлення середньої річної удійності корів. Якою повинна бути чисельність вибірки?

Якщо орієнтуватись на повторний вибір, то при граничній помилці в 30 кг з ймовірністю Р = 0,954 і при середньому квадратичному відхиленні 300 кг, визначеному за результатами аналогічних обстежень, необхідна чисельність вибірки буде

$$n = \frac{t^2\sigma^2}{\Delta_x^2} = \frac{2^2 300^2}{30^2} = \frac{4 \cdot 90000}{900} = 400 \text{ корів.}$$

При безповторному відборі, за тих же умов, необхідна чисельність вибірки буде дорівнювати:

$$n = \frac{t^2\sigma^2 N}{\Delta_x^2 N + t^2\sigma^2} = \frac{4 \cdot 90000 \cdot 8000}{900 \cdot 8000 + 4 \cdot 90000} = 380 \text{ корів.}$$

Цей розрахунок підтверджує, що при інших рівних умовах, обсяг вибірки при безповторному відборі завжди буде менший, ніж при повторному.

Для визначення частки породних корів з помилкою не більше 5%, при ймовірності 0,954 і дисперсії альтернативної ознаки 0,25, знаходимо при повторному відборі:

$$n = \frac{t^2 W(1-W)}{\Delta_P^2} = \frac{4 \cdot 0{,}25}{0{,}0025} = 400 \text{ корів},$$

і при безповторному відборі:

$$n = \frac{t^2 W(1-W)N}{\Delta_P^2 N + t^2 W(1-W)} = \frac{8000}{0{,}0025 \cdot 8000 + 1} = \frac{8000}{21} = 380 \text{ корів}.$$

Таким чином, для забезпечення прийнятної точності при повторному відборі необхідно обслідувати 400 корів, а при безповторному відборі – 380.

При типовому відборі, коли відбір одиниць проводиться з окремих типово однорідних груп, виділених за відповідною ознакою, варіація групових середніх відпадає, і помилка типової вибірки залежить від середньої величини з групових дисперсій. А тому, при типовому відборі в формулах помилок вибірки замість загальної дисперсії слід використати середню з групових ( $\overline{\sigma_i^2}$ ) – для середньої, і $\overline{W(1-W)}$ – для частки.

Таким чином, розрахунок граничної помилки вибірки при типовому відборі проводиться за допомогою таких формул:

Таблиця 13.4.

| Спосіб відбору | Гранична помилка вибірки | |
|---|---|---|
| | Для середньої | Для частки |
| Повторний | $\Delta_x = t\sqrt{\frac{\overline{\sigma_i^2}}{n}}$ | $\Delta_p = t\sqrt{\frac{\overline{W(1-W)}}{n}}$ |
| Безповторний | $\Delta_x = t\sqrt{\overline{\sigma_i^2}\left(1-\frac{n}{N}\right)}$ | $\Delta_p = t\sqrt{\frac{\overline{W(1-W)}}{n}\left(1-\frac{n}{N}\right)}$ |

Розглянемо приклад. Припустимо, що в одному з цехів підприємства працює 1000 робітників трьох професійних груп, і з кожної групи проведена 10 %-на безповторна вибірка для визначення середнього відсотка виконання норм виробітку.

Результати дослідження наведені в табл. 13.5.

З ймовірністю Р = 0,954 потрібно встановити межі відсотка виконання норм виробітки всіма робітниками цеху.

Таблиця 13.5.

| Професії | Обсяг вибірки $f_i$ | Середній відсоток виконання норм $\tilde{x}_i$ | Середнє квадратичне відхилення $\sigma_i$ |
|---|---|---|---|
| Слюсарі | 25 | 102 | 2 |
| Токарі | 45 | 107 | 4 |
| Фрезерувальники | 30 | 104 | 3 |

Розрахуємо загальну вибіркову середню:

$$\tilde{x} = \frac{\sum \tilde{x}_i f_i}{\sum f_i} = \frac{102 \cdot 25 + 107 \cdot 45 + 104 \cdot 30}{25 + 45 + 30} = 104{,}85\ \%.$$

Визначимо середню з групових дисперсій:

$$\overline{\sigma_i^2} = \frac{\sum \sigma_i^2 f_i}{\sum f_i} = \frac{4 \cdot 25 + 16 \cdot 45 + 90 \cdot 30}{25 + 45 + 30} = 10{,}9\ .$$

Тоді гранична помилка вибірки буде дорівнювати:

$$\Delta_x = t\sqrt{\frac{\overline{\sigma_i^2}}{n}\left(1 - \frac{n}{N}\right)} = 2\sqrt{\frac{10{,}9}{100}\left(1 - \frac{100}{1000}\right)} = 0{,}62\ \%.$$

Відповідно, $\bar{x} = \tilde{x} \pm \Delta_x$ ,

$$104{,}85 - 0{,}62 \le \bar{x} \le 104{,}85 + 0{,}62\ ,$$

$$104{,}23 \le \bar{x} \le 105{,}47\ .$$

Висновок: з ймовірністю 0,954 можна стверджувати, що середній відсоток виконання норм виробітку всіма робітниками цеху знаходиться в межах від 104,23 % до 105, 47 %.

Покажемо розрахунок граничної помилки вибірки і визначення узагальнюючих показників генеральної сукупності за даними типової вибірки для частки на такому умовному прикладі. Припустимо, що з метою визначення частки браку у всій партії виготовлених деталей, була проведена 10 %-е типова вибірка з відбором одиниць пропорційно чисельності одиниць типових груп. В середині типових груп використовувався метод випадкового безповторного відбору. Результати вибірки наведені в табл. 13.6.

Таблиця 13.6.

| Типи станків | Виробіток одного станка, шт. | Відсоток браку за даними вибірки |
|---|---|---|
| 1 | 1750 | 1,8 |
| 2 | 2250 | 2,5 |
| 3 | 4320 | 3,0 |
| 4 | 5680 | 1,5 |
| 5 | 2500 | 1,0 |

З ймовірністю 0,997 визначити межі, в яких знаходиться доля браку у всій партії деталей, виготовлених на всіх станках.

Розрахуємо вибіркову долю браку:

$$\overline{w} = \frac{\sum w_i f_i}{\sum f_i} = \frac{1,8\cdot 1750 + 2,5\cdot 2250 + 3,0\cdot 4320 + 1,5\cdot 5680 + 1,0\cdot 2500}{1750+2250+4320+5680+2500} =$$

$$= \frac{32755}{16500} = 1,98\% .$$

Розрахуємо дисперсії типових груп:

для I групи: $\sigma_1^2 = w_1(1-w_1) = 1,8\cdot 98,2 = 176,76$ ;

для II групи: $\sigma_2^2 = w_2(1-w_2) = 2,5\cdot 97,5 = 243,75$ ;

для III групи: $\sigma_3^2 = w_3(1-w_3) = 3,0\cdot 97,0 = 291,00$ ;

для IV групи: $\sigma_4^2 = w_4(1-w_4) = 1,5\cdot 98,5 = 147,75$ ;

для V групи: $\sigma_5^2 = w_5(1-w_5) = 1,0\cdot 99,0 = 99,00$ .

Середня з внутрігрупових дисперсій визначається за формулою:

$$\overrightarrow{\sigma^2} = \overline{w(1-w)} = \frac{\sum w_i(1-w_i)f_i}{\sum f_i} =$$

$$= \frac{176,76\cdot 1750 + 243,75\cdot 2250 + 291,00\cdot 4320 + 147,75\cdot 5680 + 99,00\cdot 2500}{16500} =$$

$$= \frac{3201607,5}{16500} = 194 .$$

Визначаємо граничну помилку вибіркової частки:

$$\Delta_p = t\sqrt{\frac{W(1-W)}{n}(1-\frac{n}{N})} = 3\cdot\sqrt{\frac{194}{16500}(1-\frac{16500}{165000})} = 0{,}3,$$

$$1{,}98 - 0{,}3 \le P \le 1{,}98 + 0{,}3,$$

$$1{,}68 \le P \le 2{,}28.$$

З ймовірністю 0,997 можна стверджувати, що частка браку у всій партії деталей, виготовлених на всіх станках, коливається в межах від 1,68 % до 2,28 %.

Для визначення необхідної чисельності вибірки при типовому відборі спочатку визначається загальна чисельність вибірки за формулою $n = \frac{t^2\overline{\sigma}^2}{\Delta^2}$ – для повторного відбору і $n = \frac{t^2\overline{\sigma^2}N}{\Delta^2N + t^2\overline{\sigma^2}}$ – для безповторного відбору, а вже потім визначається потрібний відбір із кожної групи, в залежності від виду відбору: пропорційного, непропорційного, оптимального, який враховує і чисельність одиниць в кожній групі, і варіацію досліджуваної ознаки.

Звернемось до прикладу. В академії навчається 10 тисяч студентів. Із них 5 тисяч – студенти денної форми навчання, 4 тисячі – студенти заочної форми навчання і 1 тисяча – студенти вечірньої форми навчання.

Для визначення середнього балу за зданими екзаменами планується типова вибірка із випадковим безповторним відбором внутрітипових груп. Скільки студентів потрібно відібрати, щоб з ймовірністю 0,954 помилка вибірки не перевищила 0,5 бала, якщо на основі попередніх обстежень відомо, що дисперсія дорівнює 9?

Розрахуємо необхідну чисельність типової вибірки:

$$n = \frac{t^2\overline{\sigma^2}N}{\Delta_x^2N + t^2\overline{\sigma^2}} = \frac{4\cdot 9\cdot 10000}{0{,}5^2\cdot 10000 + 4\cdot 9} = 142 \text{ студенти.}$$

Потрібно відібрати 142 студенти.

Найбільш розповсюдженим способом серійного відбору є такий, при якому утворені в генеральній сукупності і відображені при вибірці серії (гнізда) є однакові за обсягом. Очевидно, що при серійній вибірці, яка передбачає суцільне спостереження одиниць у відібраних серіях, помилка вибірки буде залежати не від числа обстежених одиниць сукупності, а від кількості відібраних серій. Помилка вибірки залежатиме не від варіації ознаки у всій сукупності, а від варіації серійних середніх, яка вимірюється міжсерійною (міжгруповою) дисперсією.

Гранична помилка серійної вибірки визначається за формулами:

Таблиця 13.7.

| Спосіб відбору | Гранична помилка вибірки | |
|---|---|---|
| | для середньої | для частки |
| Повторний | $\Delta = t\sqrt{\frac{\delta_{\overline{x}}^2}{s}}$ | $\Delta = t\sqrt{\frac{\delta_{P}^2}{s}}$ |
| Безповторний | $\Delta = t\sqrt{\frac{\delta_{\overline{x}}^2}{s}(1-\frac{s}{S})}$ | $\Delta = t\sqrt{\frac{\delta_{P}^2}{s}(1-\frac{s}{S})}$ |

де $\delta^2$ – міжсерійна (міжгрупова) дисперсія середніх;

S – число серій в генеральній сукупності;

s – число відібраних серій.

Розглянемо приклад. Посівна площа під кукурудзою в селянських спілках регіону становить 5000 га. Посіви розміщені на 50 стогектарних ділянках. Для визначення врожайності в порядку серійної безповторної вибірки відібрано п’ять стогектарних ділянок, на яких проведено суцільний облік зібраного врожаю кукурудзи. Були отримані такі дані:

Таблиця 13.8.

| № ділянки | Середній врожай на ділянці, ц/га, ($\widetilde{x}_s$) | $(\widetilde{x}_s - \overline{x}_s)$ | $(\widetilde{x}_s - \overline{x}_s)^2$ |
|---|---|---|---|
| 1 | 48 | - 2 | 4 |
| 2 | 51 | 1 | 1 |
| 3 | 50 | 0 | 0 |
| 4 | 49 | -1 | 1 |
| 5 | 52 | 2 | 4 |
| Разом: | 250 | 0 | 10 |

Визначаємо середню врожайність кукурудзи по всіх господарствах:

$$\overline{x}_s = \frac{\sum \widetilde{x}_S}{S} = \frac{250}{5} = 50 \text{ ц/га.}$$

Міжсерійну дисперсію визначаємо за формулою:

$$\delta_{\overline{x}}^2 = \frac{\sum (\widetilde{x}_S - \overline{x}_S)^2}{S} = \frac{10}{5} = 2 \,.$$

Тоді гранична помилка середньої вибірки при ймовірності P = 0,954 дорівнює:

$$\Delta_x = t\sqrt{\frac{\delta_{\bar{x}}^2}{S}(1-\frac{s}{S})} = 2\sqrt{\frac{2}{5}(1-\frac{5}{50})} = 1{,}2 \text{ ц/га.};$$

Звідси, $\bar{x} = \tilde{x} \pm \Delta_x$,

$$50 - 1{,}2 \le \bar{x} \le 50 + 1{,}2 \ ,$$

$$48{,}8 \le \bar{x} \le 51{,}2 \ .$$

Отже, з ймовірністю 0,954 можна стверджувати, що середня врожайність кукурудзи в цілому по регіону коливається в межах від 48,8 ц/га до 51,2 ц/га.

Необхідна чисельність вибірки при серійному відборі визначається як відбір певної кількості серій, які забезпечать з відповідною ймовірністю потрібну точність результатів дослідження.

Для повторного відбору необхідна чисельність визначається за формулою $s = \frac{t^2\delta^2}{\Delta_x^2}$, а для безповторного – $s = \frac{t^2\delta^2 S}{\Delta_x^2 S + t^2\delta^2}$.

В статистичній практиці вибіркове спостереження із великих масивів генеральної сукупності часто проводиться у вигляді *комбінованої, ступінчастої або декілька фазної вибірки*. Вибіркова сукупність при комбінованій вибірці формується в результаті ступінчастого відбору.

Загальна помилка при комбінованій вибірці складається із помилок, які можливі на кожній ступені, і визначається як корінь квадратний із квадратів помилок відповідних вибірок. Якщо серійну вибірку скомбінувати з власне випадковою або механічною, то гранична помилка вибірки визначається за формулою:

$$\Delta_x = t\sqrt{\frac{\delta^2}{s}(1-\frac{s}{S}) + \frac{\sigma^2}{n}(1-\frac{n}{N})} \ .$$

При застосуванні комбінованої вибірки обов'язково потрібно знати склад генеральної сукупності, а також необхідно скласти обґрунтовану схему відбору одиниць по ступенях.

*При моментному методі спостереження* гранична помилка частки визначається як при звичайній повторній власне випадковій вибірці. Наприклад, дослідник провів 120 спостережень і отримав шість відміток про те, що обладнання простоювало. Вибіркова питома вага

$$w = \frac{m}{n} = \frac{6}{120} = 0{,}05 \ .$$

Помилка вибірки з довірчою ймовірністю Р = 0,954 визначається за формулою:

$$\Delta_p = t\sqrt{\frac{W(1-W)}{n}} = 2\sqrt{\frac{0{,}05 \cdot 0{,}95}{120}} = 0{,}04\,;$$

$$W - \Delta_p \leq P \leq W + \Delta_p,$$

$$0,05 - 0,04 \leq P \leq 0,05 + 0,04,$$

$$0,01 \leq P \leq 0,09.$$

Висновок: з ймовірністю 0,954 можна стверджувати, що частка простоїв обладнання за весь робочий час знаходиться в межах від 1 % до 9 %.

Відбір моментів здійснюється за схемою механічної вибірки, або за схемою власне випадкової вибірки за таблицею випадкових чисел. Другий спосіб доцільно застосовувати в тих випадках, коли спостереження повинно бути для об’єкта несподіваним, щоб не порушувати його звичайного трудового ритму.

Для визначення чисельності моментних спостережень використовують формулу помилки власне випадкової повторної вибірки. Відбір в моментних спостереженнях завжди безповторний, однак формулу безповторного відбору тут застосувати не можна, тому що чисельність генеральної сукупності моментів роботи визначити неможливо, практично вона нескінчена, якщо момент спостереження досить короткий. А тому, для визначення необхідної чисельності моментів спостереження застосовується формула:

$$n = \frac{0{,}25t^2}{\Delta^2}\,,$$

або якщо прийняти довірчу ймовірність Р = 0,954, тобто коефіцієнт довір’я t = 2, тоді

$$n = \frac{0{,}25 \cdot 2^2}{\Delta^2} = \frac{0{,}25 \cdot 4}{\Delta_2}\,, \quad n = \frac{1}{\Delta^2}\,.$$

Наприклад, нехай нам потрібно визначити число моментних спостережень за часом роботи обладнання, щоб гранична помилка не перевищила 1 %, при ймовірності Р = 0,954.

$$n = \frac{1}{\Delta^2} = \frac{1}{0{,}0001} = 10000 \text{ спостережень.}$$

Якщо спостереження буде проводитись протягом 25 днів і буде фіксуватися стан обладнання в моменти спостереження у 50 його одиниць, тоді в день повинно проводитись $\frac{10000}{50 \cdot 25} = 8$ спостережень за кожною одиницею.

При *малих вибірках* розподіл вибіркових середніх і помилок вибірки відрізняється від нормального. Тому для оцінки результатів малої вибірки використовують дещо видозмінені формули. Середня помилка малої вибірки розраховується за формулою:

$$\mu_{\text{м.в.}} = \sqrt{\frac{S^2_{\text{м.в.}}}{n-1}} = \frac{S_{\text{м.в.}}}{\sqrt{n-1}},$$

де $S_{\text{м.в.}} = \sqrt{\frac{\sum (x-\overline{x})^2}{n-1}}$;

n – 1 – число ступенів вільності варіацій, які вказують на кількість різних можливих значень варіантів і їх середньою арифметичною.

Для ув'язки середньої помилки малої вибірки з граничною, враховується те, що при недостатньо великому обсязі вибірки стандартизована різниця між вибірковою і генеральною середньою має розподіл Стьюдента, а не нормальний. Стьюдент винайшов закон розподілу відхилень вибіркових середніх від генеральної середньої для малих вибірок і склав спеціальні таблиці в яких наведені значення t при невеликому обсязі вибірки.

Звернемось до прикладу. Нехай на ділянці лісу вибірковим методом обстежено 10 дерев з метою визначення ділової деревини в одному дереві. Вибіркова середня склала 1,75 м$^3$ при середньому квадратичному відхиленні $S_{\text{м.в.}} = 0{,}25\text{м}^3$. Потрібно з ймовірністю Р = 0,95 визначити довірчі межі складу ділової деревини для генеральної середньої.

Знаходимо середню помилку вибіркової середньої:

$$\mu_{\text{м.в.}} = \frac{S_{\text{м.в.}}}{\sqrt{n-1}} = \frac{0{,}24}{\sqrt{10-1}} = 0{,}08.$$

За таблицею Стьюдента встановлюємо, що при 9(10-1) ступенях вільності варіації ймовірності Р = 0,05 відповідає значення

t = 2,265.

Отже, гранична помилка вибірки складає:

$$\Delta = t\mu_{\text{м.в.}} = 2{,}2665 \cdot 0{,}08 = 0{,}1812 \text{ м}^3,$$

Звідси

$$\tilde{x} - \Delta \le \overline{x} \le \Delta + \tilde{x},$$

$$1{,}75 - 0{,}18 \le \overline{x} \le 1{,}75 + 0{,}18,$$

$$1{,}57 \le \overline{x} \le 1{,}93.$$

Таким чином, з ймовірністю 0,95 можна стверджувати, що в досліджуваній сукупності в одному дереві вміщається ділової деревини від 1,57 до 1,93 м$^3$.

Як відмічалось, малі вибірки використовуються головним чином для оцінки суттєвості (достовірності) різниць двох вибіркових середніх.

Розглянемо приклад оцінки достовірності різниці середніх при малій вибірці (вибірки незалежні).

Маємо дані про живу вагу телят дослідної і контрольної групи в 3-місячному віці.

Таблиця 13.9.

| Спостереження | Жива вага телят, кг | |
|---|---|---|
| | дослідна група $\chi_1$ | контрольна група $\chi_2$ |
| 1 | 94 | 66 |
| 2 | 99 | 74 |
| 3 | 108 | 78 |
| 4 | 115 | 76 |
| 5 | 99 | 69 |
| 6 | 100 | 75 |
| 7 | 128 | 73 |
| 8 | 96 | 77 |
| 9 | 110 | 70 |
| 10 | 111 | 72 |
| Разом: | 1060 | 730 |

Потрібно встановити різницю між двома середніми живої ваги телят в дослідній і контрольній групах і в якій мірі ця різниця викликана згодовуванням в дослідній групі крім незбираного молока ще й концентратів.

Визначаємо середню вагу телят по групах:

$$\tilde{x}_1 = \frac{\sum x_1}{n} = \frac{1060}{10} = 106 \text{ кг.}$$

$$\tilde{x}_2 = \frac{\sum x_2}{n} = \frac{730}{10} = 73 \text{ кг.}$$

Знаходимо різницю між середніми двох вибірок:

$$\Delta_{ф} = \tilde{x}_1 - \tilde{x}_2 = 106 - 73 = 33 \text{ кг.}$$

Вираховуємо середнє квадратичне відхилення ваги для кожної групи телят:

$$S_1 = \sqrt{\frac{\sum (x_1 - \tilde{x}_1)^2}{n_1 - 1}} = \sqrt{\frac{988}{9}} = 10{,}45 \text{ кг.}$$

$$S_2 = \sqrt{\frac{\sum (x_2 - \tilde{x}_2)^2}{n_2 - 1}} = \sqrt{\frac{130}{9}} = 3{,}8 \text{ кг.}$$

Обчислимо середні помилки вибіркових середніх по кожній групі:

$$\mu_1 = \frac{S_1}{\sqrt{n_1}} = \frac{10{,}45}{3{,}162} = 3{,}3 \text{ кг.}$$

$$\mu_2 = \frac{S_2}{\sqrt{n_2}} = \frac{3{,}8}{3{,}162} = 1{,}2 \text{ кг.}$$

Середню помилку різниці двох вибіркових середніх визначимо за формулою:

$$\overline{\mu}_{1-2} = \sqrt{\mu_1^2 + \mu_2^2} = \sqrt{3{,}3^2 + 1{,}2^2} = \sqrt{12{,}33} = 3{,}51 .$$

Число ступенів вільності двох вибірок буде становити:

$$(n_1 - 1) + (n_2 - 1) = 9 + 9 = 18 .$$

При рівні значимості 0,05 і 18 ступенях вільності

$$t_T = 2{,}1009 \approx 2{,}1 .$$

Гранична помилка для двох вибіркових середніх буде дорівнювати:

$$\Delta_{0,05} = t \cdot \overline{\mu}_{1-2} = 2{,}10 \cdot 3{,}51 = 7{,}37 \text{кг} .$$

Порівнявши фактичну різницю між обома середніми $\Delta_Ф = 33$ кг., з граничною помилкою $\Delta_{0,05} = 7{,}37$ кг., бачимо, що перша значно перевищує другу. Це свідчить про те, що різниця в середній вазі телят в дослідній і контрольній групах зумовлена дією досліджуваного фактора.

До такого самого висновку можна прийти зіставивши фактичне нормоване відхилення з табличним. В нашому прикладі фактичне нормоване відхилення становить:

$$t_ф = \frac{\Delta_ф}{\overline{\mu}_{1-2}} = \frac{33{,}0}{3{,}51} = 9{,}4 .$$

Табличне нормоване відхилення $t_T = 2{,}1$ показує максимальну величину відношень граничних помилок вибіркових середніх до їх середньої помилки. В нашому прикладі фактичне нормоване відхилення значно перевищує табличне, а тому можна зробити висновок, що різниця ваги двох середніх є не випадковою, а цілком достовірна.

Для оцінки відмінності двох залежних вибіркових середніх застосовується середня різниця.

Розглянемо приклад. При порівнянні врожайності двох сортів ярого ячменю отримано такі дані (табл. 13.10):

Таблиця 13.10

| № повторностей | Урожайність, ц/га | | Різниця врожайності d = А - Б | Квадрат різниці врожайності $d^2$ |
|---|---|---|---|---|
| | Сорт А | Сорт Б | | |
| 1 | 37,5 | 32,2 | 5,3 | 28,09 |
| 2 | 38,9 | 33,8 | 5,1 | 26,01 |
| 3 | 45,0 | 40,0 | 5,0 | 25,00 |
| 4 | 43,5 | 38,6 | 4,9 | 24,01 |
| 5 | 40,1 | 30,4 | 9,7 | 94,09 |
| Суми | 205,0 | 175,0 | 30,0 | 197,20 |
| Середні | 41,0 | 35,0 | 6,0 | × |

Потрібно дати статистичну оцінку результатам досліду, тобто встановити істотність різниці врожайності ярого ячменю між середніми двох сортів, при рівні значимості 0,05. Врожайність обох сортів можна порівнювати попарно в межах кожної повторності.

Вираховуємо середню різницю врожайності обох сортів:

$$\overline{d} = \frac{\sum d}{n} = \frac{30}{5} = 6 \text{ ц/га.}$$

Знайдемо середнє квадратичне відхилення окремих різниць d від середньої різниці $\overline{d}$:

$$S_d = \sqrt{\frac{\sum (d - \overline{d})^2}{n-1}} = \sqrt{\frac{197{,}20}{4}} = 7{,}021 .$$

Визначимо середню помилку середньої різниці:

$$\mu_{\overline{d}} = \frac{S_d}{\sqrt{n}} = \frac{7{,}021}{\sqrt{5}} = 3{,}14 \text{ ц/га.}$$

За таблицею Стьюдента значення t при кількості ступенів вільності 4 і з прийнятою нами значимістю 0,05 буде становити $t_{0,05} =$ 2,7764.

Звідси гранична помилка середньої різниці складає:

$$\Delta = t \cdot \mu = 2{,}7764 \cdot 3{,}14 = 8{,}75 \text{ ц/га.}$$

Таким чином, фактична середня різниця складає 6 ц/га, а різниця між середніми тільки в результаті випадкового коливання складає 8,75 ц/га, це означає, що різниця між врожайністю сортів ярого ячменю не є суттєвою.

До такого ж висновку прийдемо, якщо фактичне нормоване відхилення зіставимо з табличним:

$$t_{\Phi} = \frac{\overline{d}}{\mu_{\overline{d}}} = \frac{6}{3{,}14} = 1{,}9108 .$$

Табличне нормоване відхилення при значимості 0,05 $t_T = 2{,}7764$, що дає можливість зробити висновок про те, що середня різниця між врожайністю обох сортів не суттєва.

На основі теорії малої вибірки оцінюється точність вибіркової середньої, тобто визначається ймовірність того, що різниця між вибірковою і генеральною середньою не перевищить задану абсолютну величину.

Розглянемо розв'язання цієї задачі на конкретному прикладі. Припустимо, що на ділянці лісу вибірково обстежено 10 дерев з метою визначення діаметра крон. Вибіркова середня склала 4,8 м при середньому квадратичному відхиленні рівному 0,9 м. Потрібно визначити ймовірність того, що розходження між генеральною і вибірковою середньою не перевищить 0,6 м, тобто $P(\left|\tilde{x} - \overline{x}\right| \le 0{,}6 = ?$

Розв'язок. Знаходимо середню помилку вибірки:

$$\mu_{\text{м.в.}} = \frac{S_{\text{м.в.}}}{\sqrt{n-1}} = \frac{0{,}9}{\sqrt{10-1}} = \frac{0{,}9}{3{,}0} = 0{,}3 .$$

Виходячи із заданого значення граничної помилки $\Delta \le 0{,}6$ знаходимо t:

$$t = \frac{\Delta}{\mu_{\text{м.в.}}} = \frac{0{,}6}{0{,}3} = 2 .$$

За таблицею розподілу ймовірностей Стьюдента знаходимо S(t) для t = 2 і n = 10. S(t) = 0,962

S(t) = P(t) – ймовірність того, що стандартизована різниця між вибірковою і генеральною середньою має величину t.

Звідси: $P(\left|\tilde{x} - \overline{x}\right| \le t \cdot \mu_{\text{м.в.}}) = 2S(t) - 1 = 2 \cdot 0{,}962 - 1 = 1{,}924 - 1 = 0{,}924$ ,

тобто ймовірність того, що розходження між вибірковою і генеральною середньою не перевищить 0,6 м, дорівнює 0,924.

Коротко відповідь можна записати наступним чином:

$$P(4{,}2 \le \overline{x} \le 5{,}4) = 0{,}924 .$$

Отже, можна зробити висновок, що ймовірність нормованих відхилень генеральної середньої від вибіркової середньої на абсолютну величину, що перевищує 0,6 м досить висока (0,924). Тому різниця між генеральною і вибірковою середніми легко могла перевищити 0,6 м.

Користуючись даними попереднього прикладу можна з ймовірністю Р= 0,954 знайти межі, в яких знаходиться генеральна середня (діаметр крони дерев).

За умовою прикладу $\tilde{x} = 4{,}8$м; $P = 0{,}954$; $\mu_{м.в.} = 0{,}3$;
$n = 9$ (10-1).

Знаходимо $t_ф$ за формулою:

$$2S(t_ф) - 1 = 0{,}954;$$

$$2S(t_ф) = 1{,}954;$$

$$S(t_ф) = 0{,}977.$$

За таблицею ймовірностей при $n = 9$ і $S = 0{,}977$ знаходимо $t_ф$, яка дорівнює 2,3.

Межі генеральної середньої $(\bar{x})$ знаходимо за формулою:

$$\bar{x} = \tilde{x} \pm t_ф \cdot \mu_{м.в.} = 4{,}8 \pm 2{,}3 \cdot 0{,}3 = 4{,}8 \pm 0{,}69 .$$

Таким чином, з ймовірністю 0,954 можна стверджувати, що діаметр крон дерев в генеральній сукупності буде знаходитись в межах від 4,11 м до 5,49 м.

## 13.5. Способи поширення даних вибіркового спостереження на генеральну сукупність

Кінцевою практичною метою всякого вибіркового спостереження є поширення його характеристик на генеральну сукупність.

Існують два способи розповсюдження даних вибіркового спостереження: 1) спосіб прямого перерахування; 2) спосіб коефіцієнтів.

*Спосіб прямого перерахування* застосовують в тому випадку, коли на основі вибірки розраховують об’ємні показники генеральної сукупності, використовуючи для цього вибіркові середню або частку. В першому випадку середній розмір ознаки, визначений в результаті вибіркового спостереження, множиться на кількість одиниць генеральної сукупності.

Встановивши, наприклад, в результаті вибіркового спостереження продуктивності тваринництва, яке перебуває в особистій власності, що середній річний надій молока на одну корову в області становить 2500 кг, і знаючи, що всього в області в особистій власності є 20000 корів, можемо отримати величину валового надою молока: 2500 · 20000 = 50 000 000 кг або 50 тис. т. Якщо при цьому відомо, що середня помилка вибірки з певною ймовірністю дорівнює ±20 кг, а, отже, генеральна середня з тією ж ймовірністю коливається в межах від 2480 до 2520 кг, то загальний валовий надій молока з врахуванням помилки вибірки буде коливатися від 49,6 тис. т до 50, 4 тис. т.

Способом прямого перерахування можна поширювати дані вибіркового спостереження також на основі співвідношення чисельності вибіркової і генеральної сукупностей. Так, наприклад, при 2 %-му вибірковому обстеженні якості продукції в вибірку попало 100 одиниць, з яких 20 одиниць виявились неякісними. Помилка вибірки склала ± 2 одиниці. Використовуючи співвідношення чисельності вибіркової і генеральної сукупностей (2% = 1/50), можна встановити, що число бракованих одиниць продукції для всієї партії, яка складається з 5000 одиниць, буде рівним 1000 ± 100 (20 · 50 ± 2 · 50), тобто коливатиметься в межах від 900 до 1100 одиниць.

*Спосіб поправочних коефіцієнтів* застосовується в тих випадках, коли вибіркове спостереження проводиться з метою перевірки і уточнення результатів суцільного спостереження. В даному випадку, співставляючи дані вибіркового спостереження із суцільним, вираховують поправочний коефіцієнт, який використовують для внесення поправок в матеріали суцільного спостереження. Так, по закінченні переписів худоби проводять 10 %-не контрольне вибіркове обстеження худоби, яка знаходиться в особистій власності населення, і якщо в результаті контрольного обходу виявляється недооблік, тоді дані перепису коригуються на процент недообліку.

Припустимо, що в результаті перепису худоби, яка знаходиться в особистій власності населення району є 15000 свиней, в тому числі в населених пунктах, де проводиться 10 %-на вибірка, – 1200 свиней. В результаті контрольних обходів в цих же населених пунктах було обчислено 1215 свиней. Тобто, при переписі було недообліковано 15 голів свиней, що складає 1,25 % (15 від 1200). За допомогою цього поправочного коефіцієнта уточнюють матеріали перепису суцільного спостереження. В даному випадку по району всього недообліковано 187 голів свиней $\left(\frac{1{,}25 \cdot 15000}{100}\right)$.

Отже, загальна кількість свиней, які знаходяться в особистій власності населення району з поправкою на недооблік, складає 1387 голів (1200 + 187).

Вибіркове спостереження в даний час знаходить все більше застосування, а тому багато питань теорії і практики вибіркового спостереження вимагають подальшого його детального вивчення і вдосконалення.

### *Контрольні запитання*

1. Які форми несуцільного спостереження застосовуються в статистиці?
2. Що ви розумієте під вибірковим спостереженням?
3. Необхідність і умови застосування вибіркового спостереження в економічній практиці.
4. Теоретичні основи застосування вибіркового методу спостереження.
5. Що ви розумієте під генеральною сукупністю?
6. Що ви розумієте під вибірковою сукупністю?
7. Узагальнюючі характеристики генеральної і вибіркової сукупності.
8. Які ви знаєте способи відбору одиниць із генеральної сукупності?
9. Поняття про індивідуальний і груповий відбір.
10. Що ви розумієте під повторним і безповторним відбором?
11. Які ви знаєте види вибірки?
12. Поняття про власне випадкову вибірку.
13. Поняття про механічну вибірку.
14. Поняття про типову вибірку.
15. Поняття про серійну вибірку.
16. Поняття про багатоступінчасту і багатофазну вибірки.
17. Поняття про взаємопроникаючу, квантильну вибірки.
18. Поняття про направлений відбір.
19. Особливості малої вибірки.
20. Чому виникають помилки при вибірковому спостереженні?
21. Як визначається помилка вибірки для середньої і частки?
22. Як поширюються дані вибіркового спостереження на генеральну сукупність?
23. Визначення необхідної чисельності вибірки для середньої і частки при повторному і безповторному відборі.
24. Мета комбінування суцільного вибіркового спостереження.
25. Наведіть приклади використання вибіркового спостереження статистикою України.

# Розділ 14. Перевірка статистичних гіпотез

## 14.1. Загальні поняття про статистичні гіпотези

**Гіпотеза** (гр. hipothesis – основа, припущення) – наукове припущення, що висувається для пояснення сутності, тенденцій розвитку, очікуваних наслідків яких-небудь економічних явищ і процесів і вимагає перевірки на досвіді та теоретичного обґрунтування, щоб стати достовірною науковою теорією з метою їх передбачення і підтвердження.

Необхідність перевірки статистичних гіпотез виникає у різних сферах економічної і соціальної діяльності людей. Оцінювання певної ознаки генеральної сукупності здійснюється на основі цієї ж ознаки в вибірковій сукупності із врахуванням помилки репрезентативності. А по відношенню властивостей генеральної сукупності висувається деяка гіпотеза про величину середньої, дисперсії, характер розподілу, форму і тісноту зв’язку між досліджуваними змінними. Перевірку гіпотези проводять на основі виявлення узгодження фактичних і теоретичних даних. Якщо розбіжності між порівнюваними даними знаходяться в межах випадкових помилок, гіпотезу приймають.

**Статистичною гіпотезою** називається припущення про властивість певних параметрів генеральної сукупності, які перевіряються за даними вибіркового спостереження.

**Параметричною** називають статистичну гіпотезу властивості випадкової величини якої оцінюють кількісно, а **непараметричною** – якісно.

**Нульовою (основною, робочою)** називають гіпотезу, яку потрібно перевірити. Її позначають ($H_o$).

**Альтернативною (конкуруючою)** називають гіпотезу, протилежну нульовій і позначають через ($H_a$).

Нульову гіпотезу перевіряють виходячи з передбачення про відсутність розбіжності між невідомим параметром генеральної сукупності $\sigma$ і заданою величиною А. Її зміст записують після двокрапки - $H_o$: $\gamma$=A.

Нульовій гіпотезі протиставляють альтернативну, при формуванні якої враховують відхилення ($\gamma$-A). Якщо відхилення додатні – $H_a$: $\gamma$>A, від'ємні – $H_a$: $\gamma$<A, а для обох варіантів – $H_a$: $\gamma \neq$A.

Якщо вибіркові дані узгоджуються з нульовою гіпотезою, вона приймається, а якщо ці дані суперечать нульовій гіпотезі, вона відхиляється.

Перевіряють нульову гіпотезу відносно альтернативної за даними вибіркового спостереження.

Суть перевірки статистичної гіпотези зводиться до вибору з можливих двох взаємовиключаючих (альтернативних) рішень одного вірного. Наприклад, при вивченні впливу стажу роботи робітників на їхню продуктивність праці можлива альтернатива: а) збільшення стажу роботи робітників сприяє росту продуктивності праці; б) збільшення стажу роботи робітників не сприяє їх продуктивності праці.

Прикладами статистичних гіпотез можуть бути припущення про те, що урожайність озимої пшениці одного сорту в господарствах району підпорядкована закону нормального розподілу; середня продуктивність корів однієї породи перевищує середню продуктивність корів іншої породи; середня надійність виробі, що виробляються на однотипних верстатах в цехах підприємствах однакова і т.д.

Так як нульовій гіпотезі завжди протистоїть альтернативна гіпотеза, яка заперечує її, однак, при формальному підході, будь-яку з цих гіпотез можна розглядати як нульову. Тому вибір однієї з конкуруючих гіпотез як нульової потребує статистичного обґрунтування.

Принципи обґрунтування та прийняття рішень в умовах випадкової варіації досліджуваних чинників розроблені вченими статистиками і математиками: К. Пірсоном, Е. Нейманом, А.Н. Колмогоровим, Б.С. Ястремським, В.І. Романовським, Р.Фішером та ін.

Аргументація економічного обґрунтування при формуванні нульової гіпотези повинна враховувати також можливі помилки при перевірці статистичних гіпотез.

## 14.2. Помилки при перевірці статистичних гіпотез. Статистичні критерії і критична область

Статистичні гіпотези перевірять за даними вибіркового спостереження, які у зв'язку з обмеженістю обсягу вибірки зумовлюють можливість прийняття неправильних висновків. Тому існує потреба у статистичній перевірці правильності прийнятого рішення. В результаті перевірки статистичних гіпотез можна виявити два випадки, коли було прийняте неправильне рішення, тобто здійснено два роди помилок.

**Помилка першого роду** – відхилена нульова гіпотеза, хоча насправді вона є вірною.

**Помилка другого роду** – прийнята нульова гіпотеза, коли насправді правильної є альтернативна гіпотеза.

Ймовірність допустити помилку першого роду $\alpha$ – характеризує ризик 1, який ще називають **рівнем значущості**, а ймовірність допустити помилку другого роду $\beta$ – характеризує ризик 2. Так як $\alpha$ більше нуля, тому завжди існує ризик допустити помилку $\beta$.

**Статистичним критерієм** називається правило, за яким нульова гіпотеза або приймається або відхиляється. Для цього використовують випадкову величину, розподіл якої відомий. Залежно від закону розподілу використовують наступні критерії:

а) z-критерій нормального розподілу; б) t-критерій нормального розподілу і розподілу Стюдента: в) F-критерій за законом розподілу Фішера-Снеденора; г) $\chi^2$ – за законом ($\chi_i$ - квадрат) розподілу Пірсона та ін.

Для перевірки статистичних гіпотез використовують два види статистичних критеріїв: а) параметричні; б) непараметричні.

**Параметричні критерії** (z, t, F, $\chi^2$ та ін.) застосовуються у випадках, коли розподіл випадкової величини у досліджуваній сукупності підпорядкований повному відомому закону.

**Непараметричні критерії** (критерії знаків, Вілконсона, Уайта та ін.) застосовуються тоді, коли досліджуваний розподіл значно відрізняється від нормального.

Параметричні критерії ефективніші від непараметричних, ще вони використовуються лише для досліджування сукупності з нормальним або близьким до нього розподілом, а непараметричні – для будь-якого розподілу при взаємо незалежних даних спостереження.

Вибірковий простір як множину можливих вибіркових значень статистичний критерій поділяє на дві пересічні області:

а) критичну область; б) область допустимих значень (область прийняття нульової гіпотези)

Точки, які розсікають вибірковий простір на область допустимих значень і критичну область називають **критичними точками**. Критичні області, в свою чергу, поділяються на односторонню ( право-лівосторонні) і двосторонню.

**Одностороння** (право чи лівостороння) критична область визначається наступними нерівностями: для правосторонньої критичної області $k>k_{кр}$ ($k_{кр}$ – додатне число), для лівосторонньої критичної області $k<k_{кр}$ ($k_{кр}$ – від'ємне число).

**Двостороння** критична область визначається нерівностями:

$$k< k_1,\ k> k_2,$$

де $k_2> k_1$, або $|k| > k_{кр}$, де $k_{кр}> 0$.

де k – відповідний критерій;

$k_{кр}$ – критична точка;

$k_2$, $k_1$ –права і ліва сторона двосторонньої критичної області.

Вибір однієї з двох критичних областей залежить від конкретних умов і поставленого завдання перед дослідженням. Якщо гіпотеза альтернативна

$H_a$: $\overline{x_1} \neq \overline{x_2}$ – використовують двосторонню критичну область, а при гіпотезах

$H_a$: $\overline{x_1} > \overline{x_2}$ – односторонню (право-чи лівосторонню) критичну область.

Критичну область будують таким чином, щоб вона чітко відрізняла нульову гіпотезу від альтернативної.

Критерій перевірки гіпотез повинен зводити до мінімуму ризик допущення помилок. Тут також важливим є визначення ймовірності недопущення помилки другого роду. Таку ймовірність називають потужністю критерію.

**Потужністю критерію** називають ймовірність відхилення нульової гіпотези, коли правильною є альтернативна гіпотеза $H_a$ (1-α), тобто не допускається помилка другого ряду. З цією метою потрібно вибрати найпотужніший критерій.

Потужність критерію можна підвищити за рахунок:

а) збільшення рівня значущості ( від α= 0,10; 0,05; 0,01; 0,001; 0,0001 та ін.);

б) збільшення чисельності вибірки (n).

За результатами перевірки гіпотези роблять відповідні висновки: якщо фактичне значення критерію потрапляє в критичну область, нульова гіпотеза відхиляється, а якщо це значення потрапляє в область допустимих значень, нульова гіпотеза приймається.

Для кожного критерію складені спеціальні статистичні таблиці, за якими знаходять його табличне значення, яке порівнюють із фактичним значенням досліджуваного явища. Якщо фактичне значення критерію за даними вибіркового спостереження менше або дорівнює табличному нульова гіпотеза приймається, а якщо фактичне значення більше табличного – нульову гіпотезу відхиляють.

**Наведемо приклади:** а) якщо при перевірці впливу певного чинника на результативну ознаку за допомогою t критерію Стьюдента виявиться, що $t_ф > t_Т$, нульову гіпотезу відхиляють, а вплив чинника на результативну ознаку вважають вірогідним (істотним); б) якщо перевіряють істотність різниці між середніми двох і більше малих вибірок $t_ф > t_Т$, роблять висновок про те, що відмінності між середніми досить значні, а тому вони не можуть бути результатом випадкових причин, такі відмінності визначаються істотними, вірогідними; в) якщо виявиться, що $t_ф < t_Т$, роблять зворотні висновки: досліджуваний чинник на результативну ознаку не впливає, тобто його вплив неістотний, невірогідний, а різниця між середніми це також неістотна, невірогідна, в такому разі нульова гіпотеза приймається.

### 14.3. Загальна схема перевірки статистичних гіпотез

Перевірка гіпотези у випадку застосування параметричних критеріїв складається з наступних етапів [ 20,166]:

- оцінки вихідної інформації та описання статистичної моделі вибіркової сукупності;
- формулювання нульової та альтернативної гіпотез;
- встановлення рівня значущості, за допомогою якого контролюється помилка першого роду;
- вибір найпотужнішого критерію для перевірки нульової гіпотези, який контролюватиме ймовірність появи помилки другого роду;
- обчислення фактичного значення критерію;
- встановлення табличного значення критерію;
- зіставлення фактичного і табличного значення критерію з метою формулювання висновків за результатами перевірки нульової гіпотези.

Перевірка статистичних гіпотез дозволяє розв’язати задачі двох основних типів: а) про істотність відмінностей між параметрами статистичних сукупностей; б) оцінити вірогідність відмінностей між середніми, дисперсіями, коефіцієнтами кореляції, регресії та ін.

Задачі другого типу використовуються для перевірки гіпотез про відповідність вибіркового розподілу теоретичному, близькість двох фактичних розподілів, однорідність складу декількох сукупностей тощо.

Задачі першого типу пов’язані з перевіркою статистичних гіпотез, пов’язаних із застосуванням параметричних критеріїв і припущення нормального розподілу в генеральній сукупності.

Схема перевірки гіпотези задач першого типу залежить від її характеру, особливостей наявної інформації, обсягу вибіркової сукупності тощо.

Вибір конкретної схеми перевірки статистичної гіпотези ґрунтується на наступних основних засадах.

1. В залежності від обсягу вибіркової сукупності статистичну гіпотезу перевіряють через такі критерії: а) для великих за обсягом

вибірок ($n > 30$) використовують t – критерій нормального розподілу; б) для малих вибірок ($n <30$) – t-критерій розподілу Стьюдента.

2. В залежності від рівності вибірок за чисельністю, які, в свою чергу, можуть бути рівними і нерівними. Ці властивості враховують при перевірці гіпотез про істотність відмінностей між середніми, наприклад, при обчисленні середньої помилки двох вибіркових середніх.

3. В залежності від характеру формування вибіркових сукупностей:

а) для незалежних вибірок статистичній оцінці підлягає різниця середніх; для загальних вибірок – середня різниця.

4. В залежності від рівності вибіркових дисперсій при перевірці гіпотез щодо середніх можливі два випадки: а) коли дисперсії рівні ($\sigma_1^2 = \sigma_2^2$);

б) коли дисперсії нерівні ($\sigma_1^2 \neq \sigma_2^2$). Перевірку гіпотези про рівність двох дисперсій у генеральній сукупності проводять через F-критерій розподілу Фішера, який ґрунтується на співвідношенні двох вибіркових скоригованих дисперсій ($F_ф = S_1^2 : S_2^2, де S_1^2 > S_2^2$), що замінюють значення невідомих дисперсій у генеральних сукупностях. Потім знаходять за спеціальними таблицями при відповідному числі ступенів волі і заданому рівні значущості теоретичних критерій Фішера $F_т$, і якщо $F_ф > F_т$, - нульову гіпотезу відхиляють, а якщо $F_ф < F_т$, то нульову гіпотезу приймають.

При перевірці гіпотез відносно законів розподілу генеральних сукупностей може виникнути потреба у розв'язанні трьох видів задач другого типу: а) про узгодженість фактичного і теоретичного розподілів; б) про незалежність розподілу двох ознак; в) про однорідність двох і більше фактичних розподілів.

Такі гіпотези перевіряють через критерії згоди К. Пірсона ($\chi^2$), О.М. Колмогорова ($\lambda$), Б.С. Ястремського (L), В.І. Романовського (R), Р. Фішера (z) та ін. Ці критерії дозволяють встановити узгодженість досліджуваних розподілів з теоретичними, а також істотність розбіжності між ними.

### 14.4. Перевірка статистичних гіпотез відносно середніх величин

Середня величина є однією з найважливіших узагальнюючих характеристик відносно яких найчастіше висуваються статистичні гіпотези.

При формулюванні нульової гіпотези про рівність середніх у генеральній сукупності виходять з припущення, що обидві вибірки взяті з нормально розподіленої генеральної сукупності з математичним сподіванням, рівним $\overline{x}$ і з дисперсією $\sigma_0^2$. За цим припущенням $\widetilde{x_1} = \widetilde{x_2} = \overline{x}$. Однак, фактичні вибіркові середні $\widetilde{x_1}$ і $\widetilde{x_2}$ ніколи не можуть бути рівними через випадковий метод відбору досліджуваних одиниць з генеральної сукупності. В цьому випадку перевіряють істотність розбіжності між $\widetilde{x_3}$ і $\widetilde{x_4}$ і з'ясовують, в яких межах знаходиться їх різниця ( у межах можливої випадкової варіації , або за цими межами). Задача перевірити гіпотези заключається у перевірці істотності різниці:

$$\left|\tilde{x}_1 - \overline{x}\right| - \left|\tilde{x}_2 - \overline{x}\right| = \left|\tilde{x}_1 - \tilde{x}_2\right| = \Delta_\phi .$$

Помилки вибіркових середніх $(\mu)$ визначають за формулами:

$$\mu_1 = \frac{S_1}{\sqrt{n_1}}; \qquad \mu_2 = \frac{S_2}{\sqrt{n_2}},$$

де скореговані середні квадратичні відхилення:

$$S_1 = \sqrt{\frac{\Sigma\left(x_{1i} - \tilde{x}_1\right)^2}{n_1 - 1}}; \qquad S_1 = \sqrt{\frac{\Sigma\left(x_{2i} - \tilde{x}_2\right)^2}{n_1 - 1}}, \text{або}$$

$$S_1 = \sqrt{\frac{\Sigma x_{1i}^2 - n_1 \tilde{x}_2^{1}}{n_1 - 1}}; \qquad S_2 = \sqrt{\frac{\Sigma x_{2i}^2 - n_2 \tilde{x}_2^{2}}{n_2 - 1}}.$$

Середню помилку різниці двох вибіркових середніх визначають за формулою:

$$\overline{\mu}_{1-2} = \sqrt{\mu_1^2 + \mu_2^2}.$$

Фактичне значення t – критерію визначають за формулою:

$$t_\phi = \frac{\left|\tilde{X}_1 - \tilde{X}\right|}{\overline{\mu}_{1-2}} = \frac{\Delta_\phi}{\overline{\mu}_{1-2}}.$$

Таблиця значення t-критерію при певному рівні значущості і числі ступенів волі варіації для великих вибірок ($n \geq 30$) визначають за таблицею для t-критерію нормального розподілу, а для малих вибірок ($n < 30$) – за спеціальною таблицею для t-критерію Стьюдента.

Якщо вибіркове значення t-критерію потрапляє в область допустимих значень ($t_ф < t_Т$), нульова гіпотеза про рівність середніх приймається, а якщо вибіркове значення t-критерію потрапляє в критичну область ($t_ф < t_Т$), нульова гіпотеза відхиляється.

Розглядаємо приклад перевірки статистичної гіпотези для оцінки достовірності різниці середніх при малій виборці ( вибірки незалежні).

Маємо данні про живу вагу телят дослідної і контрольної групи у 3-х місячному віці (табл. 14.1):

Таблиця 14.1

Розрахункова таблиця

| Спостереження | Жива вага телят, кг | | Розрахункові дані | |
|---|---|---|---|---|
| | дослідна група $x_1$ | контрольна група $x_2$ | $x_1^2$ | $x_2^2$ |
| 1 | 95 | 65 | 9025 | 4225 |
| 2 | 98 | 75 | 9604 | 5625 |
| 3 | 105 | 73 | 11025 | 5329 |
| 4 | 112 | 67 | 12544 | 4489 |
| 5 | 100 | 70 | 10000 | 4900 |
| 6 | 99 | 64 | 9801 | 4096 |
| 7 | 111 | 74 | 12321 | 5476 |
| 8 | 115 | 72 | 13225 | 5184 |
| 9 | 125 | 69 | 15625 | 4761 |
| 10 | 90 | 71 | 8100 | 5041 |
| Разом: | 1050 | 700 | 111270 | 49126 |

Порівняння живої ваги телят по двох групах показує, що більш висока вага телят у дослідній групі в якій у раціон годівлі входили крім незбираного молока ще й концентрати. Але, так як чисельність вибірки невелика, не виключається можливість, що розбіжність між живою вагою телят отримані в результаті дії випадкових причин.

Потрібно статистично оцінити різницю між двома середніми живої ваги телят у дослідній і контрольній групах і в якості міри ця різниця викликана згодовуванням в дослідній групі крім незбираного молока ще й концентратів.

Після перевірки гіпотези зможемо зробити висновок про те, що різниця між середніми лежить в межах випадкових коливань, або ж ця різниця досить суттєва, що не узгоджується з нульовою гіпотезою про випадковий характер відмінностей між середніми.

Довівши друге припущення зможемо відхилити перше і зробити висновок про те, що раціон кормів суттєво впливає на живу вагу телят.

Умовою задачі передбачено, що обидві групи телят відібрані із нормально розподіленої генеральної сукупності, формування груп випадкове, а тому оцінювати будемо різницю між середніми.

Визначимо середню вагу телят по групах:

$$\tilde{x}_1 = \frac{\sum x_1}{n_1} = \frac{1050}{10} = 150\text{кг}; \qquad \tilde{x}_2 = \frac{\sum x_2}{n_2} = \frac{700}{10} = 70\text{кг}.$$

Знайдемо фактичну різницю між середнім двох вибірок:

$$\Delta_{ф} = \tilde{x}_1 - \tilde{x}_2 = 105 - 70 = 35 \text{ кг}.$$

Формулюємо нульову і альтернативних гіпотези:

$$H_o : \overline{x_1} = \overline{x_2}; \qquad H_a : \overline{x_1} \neq \overline{x_2}.$$

Приймемо рівень значущості α = 0,05, гарантуючи прийняття або відхилення нульової гіпотези із ймовірністю помилки в 5 випадках із 100.

Виберемо найпотужніший t-критерій Стьюдента для перевірки нашої гіпотези.

Так як за альтернативною гіпотезою $\tilde{x}_1$ може бути більшим або меншим за $\tilde{x}_2$, то критична область повинна бути встановлена з двох сторін (двостороння критична область), яка при α = 0,05 буде знаходитись в межах – всі значення вище, ніж верхня 2,5 % і нижче 2,5 % точки розподілу t-критерію Стьюдента.

Отже, нульова гіпотеза буде відхилена, якщо фактичне значення t-критерію буде більшим за табличне ($t_ф > t_Т$), або прийнята, якщо ($t_ф < t_Т$).

Вирахуємо середнє квадратичне відхилення ваги для повної групи телят:

$$S_1 = \sqrt{\frac{\Sigma x_1^2 - n\tilde{x}_2^{1}}{n_1 - 1}} = \sqrt{\frac{111270 - 10 \cdot 105^2}{10 - 1}} =$$

$$= \sqrt{\frac{111270 - 110250}{9}} = \sqrt{113,33} = 10,6 \text{ кг}.$$

$$S_2 = \sqrt{\frac{\Sigma x_2^2 - n\ \tilde{x}_2^2}{n_2 - 1}} = \sqrt{\frac{49126 - 10 \cdot 70^2}{10 - 1}} =$$

$$= \sqrt{\frac{49126 - 49000}{9}} = \sqrt{14,00} = 3,7 \text{ кг.}$$

Обчислимо середні помилки вибіркових середніх по кожній групі:

$$\mu_1 = \frac{S_1}{\sqrt{n_1}} = \frac{10,6}{\sqrt{10}} = \frac{10,64}{3,162} = 3,37; \quad \mu_2 = \frac{S_2}{\sqrt{n_2}} = \frac{3,742}{3,162} = 1,18.$$

Середню помилку різниці двох вибіркових середніх визначемо за формулою:

$$\bar{\mu}_{1-2} = \sqrt{\mu_1^2 + \mu_2^2} = \sqrt{3,37^2 + 1,18^2} = \sqrt{12,7493} = 3,57.$$

Число ступенів вільності двох вибірок буде становити:

$$k = (n_1 - 1) + (n_2 - 1) = 9 + 9 = 18.$$

При рівні значущості α = 0,05 і ступенях вільності k=18

$$t_T = 2,1009 \approx 2,1.$$

Розрахуємо фактичне значення t-критерій Стьюдента:

$$t_\phi = \frac{\Delta_\phi}{\bar{\mu}_{1-2}} = \frac{35}{3,57} = 9,8.$$

Співставимо фактичне і табличне значення t-критерій Стюдента

$$t_ф > t_{T;} \qquad 9,8 > 2,1.$$

Оскільки $t_ф > t_T$ (вибіркове значення критерію попадає в критичну область), нульова гіпотеза про рівність середніх у генеральних сукупностях відхиляється.

В нашому прикладі фактичне нормоване відхилення значно перевищує табличне, а тому можна зробити висновок, що різниця ваги двох середніх є не випадковою, а цілком достовірна.

До такого самого висновку можна прийти зіставивши фактичну різницю між обома середніми з граничною помилкою.

Гранична помилка для двох вибіркових середніх буде дорівнювати:

$$\Delta_{0,05} = t \cdot \bar{\mu}_{1-2} = 2,1 \cdot 3,57 = 7,497 \approx 7,50 кг.$$

Порівнявши фактичну різницю між обома середніми $\Delta_\phi = 3,5кг.$, з граничною помилкою $\Delta_{0,05} = 7,5кг.$, бачимо, що перша значно перевищує другу. Це свідчить про те, що різниця у середній вазі телят в дослідній і контрольній групах зумовлена дією досліджуваного чинника.

Для оцінки відмінності двох залежних вибіркових середніх застосовується середня різниця.

Розглянемо приклад. При порівнянні врожайності двох сортів озимої пшениці отримані наступні дані (табл. 14.2):

Таблиця 14.2

Розрахункова таблиця

| №: повторностей (n) | Урожайність, ц/га | | Різниця врожайності d=А-Б | Квадрат різниці врожайності $d^2$ |
|---|---|---|---|---|
| | Сорт „А” | Сорт „Б” | | |
| 1 | 38,5 | 33,2 | 5,3 | 28,09 |
| 2 | 39,9 | 34,8 | 5,1 | 26,01 |
| 3 | 46,0 | 41,0 | 5,0 | 25,00 |
| 4 | 44,5 | 39,6 | 4,9 | 24,01 |
| 5 | 41,1 | 31,4 | 9,7 | 94,09 |
| Суми | 210,0 | 180,0 | 30,0 | 197,20 |
| Середні | 42,0 | 36,0 | 6,0 | × |

Потрібно дати статистичну оцінку результатам досліду, тобто встановити істотність різниці врожайності озимої пшениці між середніми двох сортів, при рівні значущості α=0,05, використавши найпотужніший критерій перевірки нульової гіпотези t-критерій Стьюдента. Врожайність обох сортів можна порівняти попарно в межах кожної повторності.

Сформулюємо нульову і альтернативну гіпотези:

$$H_0 : \bar{d} = 0; \qquad \mathrm{H_a} : \bar{\mathrm{d}} \neq 0;$$

Вирахуємо середню різницю врожайності обох сортів:

$$\bar{d} = \frac{\sum d}{n} = \frac{30}{5} = 6 \text{ ц/га},$$

$$\text{або } \tilde{d} = \left|\tilde{x}_1 - \tilde{x}_2\right| = 42 - 36 = 6 \text{ ц/га}.$$

Знайдемо скореговане середнє квадратичне відхилення окремих різниць d від середньої різниці $\tilde{\mathrm{d}}$ :

$$S_{\tilde{d}} = \sqrt{\frac{\sum d^2 - n\tilde{d}^2}{n-1}} = \sqrt{\frac{197,2 - 5 \cdot 6^2}{5-1}} = \sqrt{\frac{197,2 - 180,0}{4}} = \sqrt{4,3} = 2.074.$$

Визначимо середню помилку середньої різниці:

$$\mu_{\tilde{d}} = \frac{S_{\tilde{d}}}{\sqrt{n}} = \frac{2,074}{\sqrt{5}} = \frac{2,074}{2,236} = 0,927.$$

Обчислимо фактичне значення t-критерію Стьюдента:

$$t_{\phi} = \frac{\left|\tilde{d}\right|}{\mu_{\tilde{d}}} = \frac{6}{0,927} = 6,4690.$$

Встановимо число ступенів вільності, виходячи з числа пар взаємопов'язаних різниць:

$$k = n-1 = 5-1 = 4.$$

За таблицею Стьюдента значення t-критерію при кількості ступенів вільності k=4 і з прийнятою значущістю $\alpha = 0,05$ становитиме:

$$t_{T(0,05)} = 2,7764.$$

Перевіряємо фактичне і табличне значення t-критеріїв:

$$t_{ф} > t_{T(0,05)}; \quad 6,4690 > 2,7764.$$

Оскільки, фактичне значення критерію значно вище за табличне, а величина середньої різниці між урожайністю двох сортів озимої пшениці істотна, то нульова гіпотеза відхиляється.

Аналітичні висновки отримаємо, порівнявши можливу граничну помилку з фактичною середньою різницею.

Гранична помилка середньої різниці складає:

$$\Delta_{0,05} = t \cdot \mu_{\tilde{d}} = 2,7764 \cdot 0,927 = 2,6 \text{ ц/га}.$$

Таким чином, фактична середня різниця складає 6 ц/га, а різниця між середніми в результаті випадкового коливання складає 2,6 ц/га, тобто $\Delta_{\phi} > \Delta_{T(0,05)}$ (6,0 >2,6), це означає, що різниця між урожайністю окремих сортів озимої пшениці вірогідна.

Серед завдань, які пов'язані з оцінкою гіпотез про середні величини, виділяють також дві групи, коли: а) дисперсія генеральної сукупності відома; б) дисперсія генеральної сукупності невідома і її потрібно замінити на дисперсію вибіркових даних.

Гіпотези щодо середніх величин перевіряють відповідно до логічних принципів у викладеній вище послідовності.

Розглянемо розв'язок цих завдань на конкретних прикладах.

Спочатку звернемось до прикладу коли дисперсія відома.

При вибірковому обстеженні 25 абонементів встановлено, що середньодобова тривалість розмова одного абонемента по телефону складає 18 хвилин $(\tilde{x})$. Припустимо, що данні про тривалість розмови розподіляється нормально з дисперсією $S^2 = 25$.

Перевіримо на рівні значущості α=0,05 нульову гіпотезу $H_0 : \tilde{x}_{H_o} = 20$ хв., проти альтернативної гіпотези $H_a : \tilde{x}_{H_a} \neq 20$ хв.

Так, як дисперсія відома, визначимо фактичне значення t-критерію:

$$t_{ф} = \frac{|\tilde{x} - \tilde{x}H_0|}{S} \cdot \sqrt{n} = \frac{|18-20|}{\sqrt{25}} \cdot \sqrt{25} = \frac{|2|}{5} \cdot 5 = 2.$$

Оскільки альтернативна гіпотеза $Ha : \tilde{x}_{H_a} \neq 20$, вибираємо двосторонню критичну область з межами f (t) =1-α = 1-0,05=0,95.

За стандартною таблицею значень функції нормованого відхилення Лапласа знаходимо, що рівню ймовірності р = 0,95 відповідає табличне значення $t_T = 1{,}96$.

Якщо $t_ф > t_T$ (2>1,96), нульова гіпотеза відхиляється і робиться висновок про те, що середньодобова тривалість розмови одного абонемента суттєво відрізняється від показника 20 хв.

Тепер розглянемо приклад, коли дисперсія невідома.

При вибірковому обстеженні 10 абонементів встановлено, що середньодобова тривалість їхньої розмови по телефону відповідно становила, хв: 15; 19; 17; 22; 20; 16; 25; 19; 18; 23. Припустивши, що дані середньодобової тривалості одного абонемента розподілені нормально, перевіримо на рівні значущості α = 0,05 нульову гіпотезу $H_0 : \tilde{x}_{H_o} = 18$ хв. при альтернативній гіпотезі $Ha : \tilde{x}_{H_a} = 19$ хв.

Обчислимо вибіркові середню арифметичну $H_0 : \tilde{X}_{H_o} = 18\ \tilde{x}$ і дисперсію $S_в^2$ ( табл. 14.3):

Таблиця 14.3

Розрахункова таблиця

| Номер спостереження | $x_i$ | $x_i - \tilde{x}$ | $(x_i - \tilde{x})^2$ |
|---|---|---|---|
| 1 | 15 | - 4,4 | 19,36 |
| 2 | 19 | - 0,4 | 0,16 |
| 3 | 17 | - 2,4 | 5,76 |
| 4 | 22 | 2,6 | 6,76 |
| 5 | 20 | 0,6 | 0,36 |
| 6 | 16 | - 3,4 | 11,56 |
| 7 | 25 | 5,6 | 31,36 |
| 8 | 19 | - 0,4 | 0,16 |
| 9 | 18 | - 1,4 | 1,96 |
| 10 | 23 | 3,6 | 12,96 |
| Суми | 194 | х | 90,40 |
| Середні | 19,4 | х | 9,04 |

$$\tilde{x} = \frac{\sum x_i}{n} = \frac{194}{10} = 19,4 \text{ хв.}; \qquad S_в^2 = \frac{\sum(x_i - \tilde{x})^2}{n} = \frac{90,4}{10} = 9,04;$$

$$S_в = \sqrt{S_в^2} = \sqrt{9,04} = 3,0066 \approx 3 \text{ хв.}$$

Оскільки дисперсія генеральної сукупності невідома, число ступенів вільності дорівнює k= n-1, скористаємось дисперсією вибіркової сукупності $S_в^2$ =9 ( а $S_в$ =3) і визначимо фактичне значення нормованого відхилення:

$$t_ф = \frac{\tilde{x} - \tilde{x}_{но}}{S_b} \cdot \sqrt{n-1} = \frac{19,4 - 18}{3} \cdot 3 = 1,4 .$$

У зв'язку з тим, що $H_a$: $\tilde{x}_{H_a} = 19 > \tilde{x}_{H_o}$ рівно як і для $H_a$: $\tilde{x}_{H_a} < \tilde{x}_H$ вибираємо односторонню критичну область, межі якої S(t)=1-2 $\alpha$ =1-2·0,05=0,9.

За стандартною таблицею t-розподілу Стьюдента, при числі ступенів вільності k=10-1=9 і рівні значущості $\alpha$ = 0,1 (2 . 0,05) знаходимо, що $t_T$= 1,83.

Оскільки $t_ф < t_T$ ( 1,4 < 1,83), то нульова гіпотеза не відхиляється тобто вона не протирічить вибірковим даним. Отже, середньодобова тривалість розмови по телефону одного абонементу суттєво не відрізняється від 18 хв.

Якщо гіпотеза про величину центру розподілу перевіряється за результатами малої вибірки, то відношення різниці середніх до стандартної помилки вибірки підпадає під розподіл Стьюдента з ступенями вільності k= n-1 ( як у попередньому прикладі), тобто

$$t_{\text{ф}} = \frac{|\tilde{x}_1 - \tilde{x}_2|}{\mu_{\tilde{x}}} = \frac{|\tilde{x}_1 - \tilde{x}_2|}{\frac{S}{\sqrt{n-1}}} = \frac{|\tilde{x}_1 - \tilde{x}_2|}{S} \cdot \sqrt{n-1}.$$

Розглянемо приклад. За результатами вибіркової перевірки 17 виробів встановлено, що собівартість одного виробу за традиційною технологією становила ( $\tilde{x}_{\text{Т}} = 25\text{грн.}$ ), а за прогресивною ( $\tilde{x}_{\text{п}} = 20\text{грн.}$ ), дисперсія – $S^2_{\text{в}} = 9, \quad S_{\text{в}} = 3$ грн.

Визначимо фактичну величину t-критерію:

$$t_{\text{ф}} = \frac{|\tilde{x}_{\text{П}} - \tilde{x}_{\text{Т}}|}{\frac{S}{\sqrt{n-1}}} = \frac{|20-25|}{3} \cdot \sqrt{17-1} = \frac{|5|}{3} \cdot 4 = 6,67.$$

За таблицею Стьюдента значення t-критерію для числа ступенів вільності k = n-1=17-1=16, і рівні значущості 0,05 становить $t_{\text{Т}}$ = 2,12. Так як фактичне значення t-критерію значно перевищує табличне ($t_{\text{ф}}$=6,67 > $t_{\text{Т}(\ 0,05)}$=2,12), то нульова гіпотеза відхиляється, тобто прогресивна технологія веде до суттєвого зниження собівартості виробів.

Перевірку гіпотези про суттєвість розбіжностей двох вибіркових часток покажемо на такому прикладі. Нехай, за результатами вибіркового обстеження сімей двох областей України були отримані дані про їх забезпеченість основними товарами культурно-побутового призначення в залежності від середньодушових доходів сімей (табл. 14.4):

Таблиця 14.4

| Область | Кількість обстежених сімей, тис. $n_i$ | Забезпеченість сімей товарами культурно-побутового призначення, % $w_i$ |
|---|---|---|
| А | 12 | 75 |
| Б | 10 | 80 |

Різниця двох вибіркових часток складає 5 % (80-75). Потрібно переконатись, чи можна вважати несуттєвими розбіжності в частці забезпеченості сімей товарами культурно-побутового призначення в залежності від їх середньодушового доходу. Нульова гіпотеза заключається в тому, що немає суттєвих розбіжностей у забезпеченості сімей цими товарами в залежності від середньодушових доходів.

Для оцінки генеральної частки використаємо середню зважену із часток, отриманих за результатами вибіркового обстеження сімей в кожній області.

$$P = \frac{\sum w_i n_i}{\sum n_i} = \frac{0,75 \cdot 12 + 0,80 \cdot 10}{12 + 10} = \frac{17}{22} = 0,7727,$$

тобто, оцінка частки забезпеченості сімей товарами культурно-побутового призначення в залежності від середньодушового доходу в генеральній сукупності становить 77,3 %.

Середню помилку різниці двох часток при справедливості нульової гіпотези обчислимо за формулою:

$$\bar{\mu}_{W_1 - W_2} = \sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}} = \sqrt{p(1-p) \cdot (\frac{1}{n_1} + \frac{1}{n_2})} =$$

$$= \sqrt{0,7727 \cdot 0,0073 \cdot (\frac{1}{18000} + \frac{1}{10000})} =$$

$$= \sqrt{0,1756347 \cdot 0,0000833} = \sqrt{0,0000146} = 0,00382.$$

Таким чином, середня помилка різниці двох вибіркових часток складає 0,382 %.

Оскільки обидві вибірки досить великого обсягу, потрібно скористатись таблицею нормованої функції Лапласа для знаходження значення коефіцієнта довір'я t при ймовірності 0,95 або 0,99. Цим значенням ймовірності відповідають табличні значення: $t_{(0,95)}$= 1,96; $t_{(0,99)}$= 2,58.

Визначимо фактичне значення t-критерію:

$$t_{\phi} = \frac{|w_1 - w_2|}{\bar{\mu}_{W_1 - W_2}} = \frac{|0,75 - 0,80|}{0,00382} = \frac{|0,05|}{0,00382} = 13,089.$$

Оскільки фактичне значення t-критерію значно більше теоретичних $t_{\phi} > t_{T(0,99)} > t_{T(0,95)}$ (13,089 > 2,58 > 1,96) нульова гіпотеза відхиляється і робиться висновок про суттєвість розбіжностей частки забезпеченості сімей товарами культурно-побутового призначення в залежності від їх середньодушового грошового доходу.

Можна також визначити максимально можливу величину розбіжностей двох вибіркових часток із заданою ймовірністю за формулою:

$$\Delta_{W_1-W_2} = t \cdot \overline{\mu}_{W_1-W_2}.$$

При ймовірності р=0,95: $\Delta_{W_1-W_2}$ =1,96 · 0,00382=0,00749.

При ймовірності р=0,99: $\Delta_{W_1-W_2}$ =2,58 · 0,00382=0,00985.

Оскільки фактична розбіжність двох вибіркових часток 0,05 більша визначених граничних помилок, нульову гіпотезу потрібно відхилити.

Таким чином, обидва варіанти перевірки гіпотези про суттєвість відмінностей частки забезпеченості сімей товарами культурно-побутового призначення в залежності від середньодушових доходів за результатами вибіркових обстежень в двох областях показали однаковий результат.

## 14.5 Перевірка статистичних гіпотез відносно законів розподілу

Поряд з перевіркою статистичних гіпотез відносно середніх в окремих випадках потрібно обов'язково перевірити гіпотези щодо характеру розподілу. Розподіли в генеральній сукупності підпорядковується певному статистичному закону. Перевірка статистичної гіпотези заключається в тому, щоб на основі порівняння фактичних частот з теоретичними зробити висновок про відповідність фактичного розподілу теоретичному.

Варіацію рядів розподілу можна описати певною функцією теоретичної кривої. Серед найбільш розповсюджених є **крива нормального розподілу.** Вона використовується як стандарт для порівняння інших розподілів.

Нормальний розподіл близький до інших одновершинних розподілів, а тому його часто використовують як перше наближення при статистичному моделюванні. Деякі розподіли приводять до нормального виду через заміну значень „x" їх логарифмами „lgx". Логарифмічною нормальною кривою можна описати ряд асиметричних розподілів.

Частоти теоретичної кривої нормального розподілу визначають за формулою:

$$f' = \sum f \frac{i}{S}\ f(t), \quad \text{де} \quad f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} .$$

Інтегральна функція розподілу має вигляд:

$$F(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_{-x}^{+x} e^{-\frac{t^2}{2}} dt .$$

Функція F(x) табульована, її значення знаходять по спеціальній таблиці.

Аналітично нормальний розподіл описується таким рівнянням:

$$\hat{y}_t = \frac{1}{S\sqrt{2\pi}} e^{-\frac{t^2}{2}},$$

де $\hat{y}_t$ – ордината кривої нормального розподілу;

t – нормоване відхилення, яке дорівнює $\frac{x - \tilde{x}}{S}$;

S – середнє квадратичне відхилення;

$\pi$ – величина відношення довжини кола до діаметру, $\pi$=3,1415;

е – основа натуральних логарифмів, е =2, 7182.

Для перевірки відповідності фактичного розподілу нормальному, частоти фактичного розподілу порівнюють з теоретичними, які характерні для нормального розподілу. З цією метою за фактичними даними вираховують теоретичні частоти кривої нормального розподілу. Тобто, фактичну криву розподілу вирівнюють кривою нормального розподілу.

Після розрахунку теоретичних частот виникає потреба перевірки висунутої нульової гіпотези про відповідність чи невідповідність теоретичного закону розподілу, прийнятого в якості математичної моделі для емпіричного розрахунку. При цьому виходять з того, що якщо відхилення між фактичними і теоретичними частотами можна вважати випадковими, тоді нульова гіпотеза про те, що прийнятий теоретичний розподіл відповідає даному емпіричному, не відхиляється.

Математична статистика дає ряд показників, за якими судять наскільки фактичний розподіл узгоджується з нормальним. Такі показники називаються критеріями узгодження, які виступають у вигляді деякої величини, котра оцінює досліджуване явище з певною ймовірністю.

Статистика використовує критерії узгодження К. Пірсона ($\chi^2$), А.Н. Колмогорова ($\lambda$), Б.С. Ястремського (L), В.І. Романовського (R), Р. Фішера (z), Вілконсона та ін.

Одними з основних і найбільш розповсюджених в критерії К. Пірсона $\chi^2$ і А.Н. Колмогорова $\lambda$.

**Критерій Пірсона** $\chi^2$ обчислюють за формулою:

$$\chi^2 = \sum \frac{(f - f')^2}{f'},$$

де: $f$ і $f'$ - відповідно фактичні і теоретичні частоти.

За спеціальними таблицями визначають ймовірність досліджуваного значення $\chi^2$ в залежності від числа ступенів вільності. Число ступенів вільності визначають за формулою: k=m-r, де m – число груп; r – число обмежених зв'язків. Якщо фактичне $\chi^2_ф$ менше табличного $\chi^2_ф$ ($\chi^2_ф < \chi^2_т$), тоді при прийнятому рівні значущості розходження між фактичними і теоретичними частотами вважаються випадковими, нульова гіпотеза про закон розподілу приймається.

Розглянемо на прикладах застосування критерію $\chi^2$ для доведення гіпотези про правильність вибору типу розподілу господарств за врожайністю гречки. ( табл. 14.5).

Таблиця 14.5.

Розрахункова таблиця

| Урожайність гречки, ц/га (x) | Кількість господарств (f) | x | $\lvert x-\bar{x}\rvert$ | $t=\frac{x-\bar{x}}{S}$ | $f(t)$ | $f'$ | $f-f'$ | $(f-f')^2$ | $\frac{(f-f')^2}{f}$ |
|---|---|---|---|---|---|---|---|---|---|
| 13-15 | 4 | 14 | 5,76 | 2,23 | 0,0332 | 2 | 2 | 4 | 2,00 |
| 15-17 | 7 | 16 | 3,76 | 1,46 | 0,1374 | 11 | -4 | 16 | 1,45 |
| 17-19 | 27 | 18 | 1,76 | 0,68 | 0,3166 | 25 | 2 | 4 | 0,16 |
| 19-21 | 35 | 20 | 0,24 | 0,09 | 0,3973 | 31 | 4 | 16 | 0,52 |
| 21-23 | 17 | 22 | 2,24 | 0,87 | 0,2732 | 21 | -4 | 16 | 0,76 |
| 23-25 | 6 | 24 | 4,24 | 1,64 | 0,1040 | 8 | -2 | 4 | 0,50 |
| 25-27 | 4 | 26 | 6,24 | 2,42 | 0,0213 | 2 | 2 | 4 | 2,00 |
| Разом: | 100 | × | × | × | × | 100 | × | × | 7,39 |

$$\tilde{x}=\frac{\sum xf}{\sum f}=\frac{1976}{100}=19,76 \text{ ц/га}; \qquad S^2=\frac{\sum(x-\tilde{x})^2 f}{\sum f}=\frac{666,24}{100}=6,6624;$$

$$S=\sqrt{6,6624}=2,58 \text{ ц/га}.$$

$$\sum f\frac{i}{S}=100\frac{2}{2,58}=77,5\ ; \qquad f_1'=\sum f\frac{i}{S}\cdot f(t)=77,5\cdot 0,0332=2;$$

$$f_1'=77,5\cdot 0,1374=11; \text{ і т.д.}$$

$$\chi^2_ф=\sum\frac{(f-f')^2}{f'}=7,39; \quad k=m\text{-}r=7\text{-}3=4; \quad \chi^2_{Т(0,95)}=14,86.$$

Оскільки фактичний критерій $\chi^2_ф$ значно менший табличних при двох ймовірностях ($\chi^2_ф$=7,39 < $\chi^2_{Т(0,99)}$=13,28 < $\chi^2_{Т(0,95)}$=14,86), нульова гіпотеза приймається, а з ймовірністю 0,99 чи 0,95 можна вважати доведеним, що тип розподілу прийнятий правильно, тобто, що розподіл господарств за врожайністю гречки є нормальним.

До такого самого висновку можна дійти за допомогою непараметричних критеріїв згоди, таких як критерій Колмогорова (λ), критерій Уайта, критерій Уїлксона та ін.

**Критерій згоди А.Н. Колмогорова (λ)** оцінює близькість фактичного розподілу до теоретичного шляхом знаходження величини (D), тобто максимальної різниці нагромаджених (кумулятивних) часток (частот) фактичного і теоретичного розподілів.

Критерій Колмогорова визначають за формулою:

$$\lambda = \mathrm{D}\sqrt{n}\ ,$$

де: D – абсолютна максимальна різниця кумулятивних часток D = max $\left|S_d - S_{d^1}\right|$, або частот D= max $\left|S_f - S_{f'}\right|$ емпіричного і теоретичного розподілу;

n – число спостережень (чисельність одиниць сукупності).

Якщо розподіл задано в частотах, тоді формула матиме вигляд:

$$\lambda = \frac{D}{\sqrt{n}}\ .$$

Критерій згоди Колмогорова ($\lambda$) простий в розрахунках, не передбачає використання стандартних таблиць для того оцінювання, не потребує визначення показника кількості ступенів вільності. Математичною статистикою доведено, що при кількості вибіркової сукупності більшої за 25 одиниць (n > 25) граничні значення критерію ($\lambda_Т$), які відповідають трьом порогам довірчої ймовірності дорівнюють: при p=0,95, $\lambda_Т$=1,36; при p=0,99, $\lambda_Т$= 1,63; при p= 0,999, $\lambda_Т$= 1,95.

Якщо $\lambda_ф \geq \lambda_{Т,}$ то з відповідною ймовірністю можна стверджувати, що розбіжність між фактичними і теоретичними розподілами істотні ( суттєві), і, навпаки, якщо $\lambda_ф < \lambda_Т$ – розбіжності між фактичними і теоретичними розподілами вважають не суттєвими, а тип розподілу вибраний правильно.

Методику розрахунку цього показника розглянемо на прикладі даних табл. 14.5.

Таблиця 14.6

Розрахунок критерію (λ)

| Номер групи | Нагромаджені частоти | | Відхилення $\lvert S_f - S_{f'} \rvert$ |
|---|---|---|---|
| | Емпіричні $S_f$ | Теоретичні $S_{f'}$ | |
| 1 | 4 | 2 | 2 |
| 2 | 11 | 13 | 2 |
| 3 | 38 | 38 | 0 |
| 4 | 73 | 69 | 4 |
| 5 | 90 | 90 | 0 |
| 6 | 96 | 98 | 2 |
| 7 | 100 | 100 | 0 |

$$\lambda_{\text{ф}} = \frac{D}{\sqrt{n}} = \frac{4}{\sqrt{100}} = \frac{4}{10} = 0,4.$$

Отримана величина фактичного критерію $\lambda_{\phi(p=0,95)=1,36}$, тому розбіжності між фактичним і нормальним розподілом є невірогідними, тобто розходження між частотами знаходяться в межах випадкових коливань, а тому можна зробити висновок про те, що розподіл господарств за врожайністю гречки являється нормальним.

Для встановлення вірогідності двох фактичних розподілів, отриманих при вибірці з однієї генеральної сукупності, але з різною кількістю одиниць ($f_1 \neq f_2$), критерій Колмогорова ($\lambda$) обчислюють за формулою:

$$\lambda_{\text{ф}} = \left| \frac{\sum f_1^i}{f_1} - \frac{\sum f_2^i}{f_2} \right| \max \sqrt{\frac{f_1 \cdot f_2}{f_1 + f_2}},$$

де $\frac{\sum f_1^i}{f_1}$ - суми нагромаджених частот по кожному інтервалу першого ряду розподілу, поділені на обсяг першої вибірки;

$\frac{\sum f_2^i}{f_2}$ - суми нагромаджених частот по кожному інтервалу другого ряду розподілу, поділені на обсяг другої вибірки;

$\left| \frac{\sum f_1^i}{f_1} - \frac{\sum f_2^i}{f_2} \right|$ max – максимальне абсолютне значення різниці

частот від ділення нагромаджених частот на обсяг вибірок;

$f_1, f_2$ - обсяги одиниць вибіркової сукупності першого і другого рядів розподілу.

Звернемось до прикладу. Доповнимо попередню задачу ще одним емпіричним рядом розподілу і обчислимо критерій Колмогорова ($\lambda$) для оцінки розбіжностей між емпіричними рядами розподілу з неоднаковими обсягами вибірки (табл. 14.7).

Таблиця 14.7

Розрахункова таблиця.

| Номер груп | Частоти | | Нагромаджені частоти | | Розрахункові данні | | |
|---|---|---|---|---|---|---|---|
| | $f_1$ | $f_2$ | $f_1^i$ | $f_2^i$ | $\frac{f_1^i}{f_1}$ | $\frac{f_2^i}{f_2}$ | $d = \left\lvert \frac{f_1^i}{f_1} - \frac{f_2^i}{f_2} \right\rvert$ |
| 1 | 4 | 8 | 4 | 8 | 0,04 | 0,04 | 0,00 |
| 2 | 7 | 14 | 11 | 22 | 0,11 | 0,11 | 0,00 |
| 3 | 27 | 50 | 38 | 72 | 0,38 | 0,36 | **0,02** |
| 4 | 35 | 72 | 73 | 144 | 0,73 | 0,72 | 0,01 |
| 5 | 17 | 38 | 90 | 182 | 0,90 | 0,91 | 0,01 |
| 6 | 6 | 13 | 96 | 193 | 0,96 | 0,96 | 0,00 |
| 7 | 4 | 7 | 100 | 200 | 1,00 | 1,00 | 0,00 |
| Разом: | 100 | 200 | × | × | × | × | × |

У таблиці 14.7 наведені два емпіричні ряди розподілу господарств за врожайністю гречки, відібраних з однієї генеральної сукупності у різних обсягах.

В останній графі цієї таблиці максимальна різниця нагромаджених частот становить d max =0,02. Тоді, за вище наведеною формулою фактичний критерій $\lambda$ дорівнюватиме:

$$\lambda_ф = d\max \sqrt{\frac{f_1 \cdot f_2}{f_1 + f_2}} = 0,02 \cdot \sqrt{\frac{100 \cdot 200}{100 + 200}} = 0,02 \cdot \sqrt{\frac{20000}{300}} = 0,02 \cdot 8,16 = 0,163.$$

Фактична величина критерію $\lambda_ф$ значно менше граничних критичних значень $\lambda$ для трьох порогів ймовірності ($\lambda_ф$ =0,163 <

$\lambda_{0,05}$=1,36< $\lambda_{0,01}$=1,63< $\lambda_{(0,001)} = 1{,}95)$. Отже, розбіжності між порівнюваними емпіричними рядами розподілу несуттєві, що свідчить про те, що обидві вибірки репрезентують досліджувану генеральну сукупність.

Якщо під знаком радикала маємо громіздкі числа, критерій ($\lambda$) можна обчислити за дещо видозміненою формулою:

$$\lambda_{ф}^{2} = d^{2}\max\frac{f_1\cdot f_2}{f_1+f_2} = 0,02^{2}\cdot\frac{20000}{300} = 0,0004\cdot 66,6667 = 0,0267,$$

$$\text{звідси } \lambda_{ф} = \sqrt{\lambda_{ф}^{2}} = \sqrt{0,0267} = 0,163.$$

**Критерій згоди В.І.Романовского (R)** також використовують для оцінки наближення фактичного розподілу до теоретичного. Його обчислюють за формулою:

$$R = \frac{\chi^{2}-k}{\sqrt{2\cdot k}}.$$

Скористаємось розрахунками наведеними в табл. 14.5 обчислимо критерій R:

$$R = \frac{7,39-4}{\sqrt{2\cdot 4}} = \frac{3,3900}{2,8284} = 1,1985 \approx 1,2.$$

Якщо при дослідженні наближення фактичного розподілу до теоретичного величина критерію згоди Романовського менше трьох (R=1,2< 3), це дозволяє стверджувати про можливість прийняття цього розподілу за законом його розподілу. Тобто, що розподіл господарств за врожайністю гречки є нормальним.

**Критерій згоди Б.С. Ястремського (L)** використовують для прямої відповіді на питання про міру розбіжності між фактичним і теоретичним розподілами. Критерій визначають за формулою:

$$L = \frac{\chi^{2}-n}{\sqrt{2n+4Q}},$$

де: $Q=\dfrac{(f_i - f_i^{'})^{2}}{f_i(1-p_i)}$ - при кількості груп менше 20 ($n < 20$), дорівнює 0,6.

За даними попередніх прикладів наведемо розрахунок критерію згоди Ястремського L:

$$L = \frac{7,39-7}{\sqrt{2\cdot 7+4\cdot 0,6}} = \frac{0,39}{4,02} = 0,097.$$

Оскільки, критерій згоди Ястремського значно менший трьох (L=0,097<3), то з ймовірністю 0,997 можна стверджувати, що розподіл господарств за врожайністю гречки є нормальним.

Таким чином, для перевірки гіпотези про відповідність чи невідповідність теоретичного закону розподілу емпіричному можна використати любий з наведених критеріїв, які забезпечать дослідження законів розподілу з різною точністю, надійністю і трудомісткістю.

### 14.6. Перевірка статистичних гіпотез про істотність розбіжностей між дисперсіями

Перевірка статистичних гіпотез про рівність дисперсій має велике практичне значення, особливо при вирішенні технічних завдань, наприклад, при аналізі стабільності виробничого процесу до і після впровадження нової техніки і технології, точності приладів, машин і устаткування; однорідності двох сукупностей відносно таких ознак, як стаж роботи, рівень продуктивності праці; в біологічних дослідженнях тощо. На оцінці різниць дисперсій побудований один з найбільш ефективних статистичних прийомів кількісного дослідження видів причинно-наслідкових залежностей – метод дисперсійного аналізу, який детально викладений у розділі 8.

Потреба у перевірці гіпотези про рівність дисперсій виникає при порівнянні середніх величин, коли передбачається, що генеральні дисперсії рівні. Однак, вибіркові дисперсії, як правило, нерівні, тому потрібно під час перевірки статистичної гіпотези про рівність середніх перевірити гіпотезу про істотність різниці дисперсій.

Гіпотеза про рівність двох дисперсій перевіряється за допомогою F-критерія Фішера, а для оцінки істотності відмінностей ряду дисперсій при однаковій чисельності вибірок користуються критерієм Кохрана (q), при неоднакових вибіркових сукупностях – критерієм Бартлета (M).

**F-критерій Фішера** являє собою відношення двох вибіркових дисперсій $S_1^2$ і $S_2^2$ при відповідних ступенях вільності $k_1$ і $k_2$.

$$F_ф = \frac{S_1^2}{S_2^2}, \text{ де } S_1^2 > S_2^2.$$

Обчислений фактичний F-критерій за даними вибіркових сукупностей порівнюють з табличним ($F_т$) з урахуванням для кожної дисперсії ступенів вільності варіації ($k_1$ і $k_2$) при заданому рівні значущості (α). Стандартні таблиці F-критерію являють собою розподіл відношень більшої дисперсії до меншої, тому F-критерій завжди буде більший одиниці (F>1), а за критичну область приймають коли $F_ф \geq F_т$.

Якщо при перевірці статистичних гіпотез про рівність дисперсій ($H_o$: $S_1^2 = S_2^2$) виявиться, що $F_ф > F_т$ нульова гіпотеза відхиляється, а при $F_ф < F_т$ – приймається.

Якщо за альтернативною гіпотезою $H_a$: $S_1^2 \neq S_2^2$ вибирають двосторонній критерій, а якщо $H_a$: $S_1^2 > S_2^2$ – вибирають односторонній критерій.

Послідовність перевірки статистичної гіпотези про істотність різниці між двома дисперсіями розглянемо на прикладі. Для перевірки гіпотези про рівність дисперсій (оцінки істотності різниці дисперсій) показників тривалості горіння електричних лампочок проведено механічний 5-ти і 10-ти відсотковий відбір однотипних підприємств з генеральної сукупності отримані вибірки з різною чисельністю одиниць спостереження: $n_1$=25; $n_2$=15; $S_1^2 = 12$; $S_2^2 = 10$.

Сформулюємо нульову і альтернативну гіпотези:

$$H_o: S_1^2 = S_2^2; \; H_a: S_1^2 > S_2^2.$$

Приймемо рівень значущості α = 0,05.

Обчислимо фактичне значення F-критерію:

$$F_ф = \frac{S_1^2}{S_2^2} = \frac{12}{10} = 1,2 .$$

Табличне значення F-критерію при $k_1=n_1-1=25-1=24$, $k_2=n_2-1=15-1=14$ і α=0,05 (p=0,95) дорівнює $F_{т(0,05)}=2,35$.

Оскільки фактичне значення F-критерію значно менше табличного ($F_ф=1,2<F_{т(0,05)}=2,35$) нульова гіпотеза про рівність дисперсій приймається, а це означає, що дисперсія ознаки (тривалість горіння лампочок) в меншій сукупності підприємств майже не відрізняється, тобто різниця між порівнювальними дисперсіями є неістотною (випадковою).

Для великих вибірок перевірка статистичної гіпотези про рівність двох дисперсій здійснюється через t-критерій нормального розподілу, який визначається за формулою:

$$t = \frac{S_1 - S_2}{\bar{\mu}_{1-2}},$$

де: $\bar{\mu}_{1-2}$ - середня з помилок вибіркових середніх квадратичних відхилень.

Тут: $\mu_1 = \dfrac{S_1}{\sqrt{2(n_1-1)}}$; $\mu_2 = \dfrac{S_2}{\sqrt{2(n_2-1)}}$, звідси: $\bar{\mu}_{1-2} = \sqrt{\mu_2^2 + \mu_1^2}$ .

При необхідності отримання оцінки істотності декількох (більше двох) дисперсій з однакової чисельності вибіркових сукупностей використовують критерій Кохраща (q).

**Критерій Кохрана (q)** визначають діленням максимальної з порівнювальних дисперсій до суми всіх дисперсій.

$$q_ф = \frac{S_{max}^2}{S_1^2 + S_2^2 + S_3^2 + .... + S_n^2}.$$

Фактичну величину критерію Кохрана ($q_ф$) порівнюють з табличним значенням ($q_T$) з врахуванням кількості ступенів вільності (k=n-1) і заданому рівні значущості ( $\alpha$ ).

Якщо фактична величина критерію Кохрана менша табличного ($q_ф < q_T$), нульова гіпотеза приймається, а якщо ($q_ф > q_T$), нульова гіпотеза відхиляється, дисперсії визнають неоднорідними, оскільки їх відмінність істотна.

Розглянемо приклад. З метою дослідження годинної продуктивної праці робітників заводу було відібрано власне випадковим відбором три групи робітників по 9 осіб з цехів заводу. Для перевірки гіпотези про істотність різниць дисперсій отримані такі параметри: $n_1 = n_2 = n_3 = 9$ осіб; $S_1^2 = 6$; $S_2^2 = 6$; $S_3^2 = 3$; k= n-1=9-1=8; $\alpha$ =0,05.

Визначимо фактичний критерій Кохрана ($q_ф$):

$$q_ф = \frac{6}{6+5+3} = \frac{6}{14} = 0,4286.$$

Табличне значення критерію при кількості ступенів вільності k=8, ймовірності p=0,95 дорівнює $q_{T(0,05)}$=0,3043.

Оскільки $q_ф$=0,4286>$q_{T(0,05)}$=0,3043 нульова гіпотеза відхиляється. Тобто дисперсії вважають неоднорідними, так як їх різниця досить суттєва.

Одним з найбільш потужніших критеріїв перевірки гіпотези про однорідність дисперсій вважається критерій Бартлета (М). Він використовується для оцінки істотності відмінності декількох дисперсій, вирахуваних з неоднакових вибірок. Теоретичною основою застосування цього критерію є припущення, що розподіл ознак у досліджуваних сукупностях є нормальним.

Суть розрахунку критерію Бартлета (М) заключається у порівнянні середніх дисперсій обчислених за формулами середньої арифметичної і середньої геометричної [26,161]:

$$\overline{\sigma_{а}^{2}} = \frac{\sum \sigma_{i}^{2} n_{i}}{\sum n_{i}}; \qquad \overline{\sigma_{г}^{2}} = \sum n_{i} \sqrt{\Pi\left[(\sigma_{i}^{2})^{ni}\right]}.$$

Якщо порівнювальні дисперсії рівні, то обчислені середня арифметична і середня геометрична з дисперсій збігатимуться.

**Критерій Бартлета (М)** для перевірки дисперсій визначається за формулою

$$M = \ln \frac{\overline{\sigma_{а}^{2}}}{\overline{\sigma_{г}^{2}}} \cdot \sum n_{i}.$$

Перетворивши натуральні логарифми у десяткові формула Бартлета матиме вигляд:

$$M=2,3026\ (\lg \overline{\sigma_{а}^{2}} \sum n_{i} - \sum n_{i}\ \lg \overline{\sigma_{г}^{2}}).$$

Порівнявши відношення М:С, де $C= 1+ \dfrac{\sum \frac{1}{n_{i}} - \frac{1}{\sum n_{1}}}{3(m-1)}$,

то тому розподіл відповідатиме розподілу $\chi^2$ (хі-квадрат) з кількістю ступені вільності k=p-1 де p – кількість дисперсій, які порівнюються. Критичні значення критерію (М) знаходять за стандартними таблицями розподілу $\chi^2$ при обраній довірчій ймовірності і кількості степенів вільності.

Розглянемо послідовність перевірки статистичної гіпотези за критеріями Бартлета на конкретному прикладі. В результаті статистичного аналізу вибіркових сукупностей населення трьох областей Західного регіону України щодо середньодушового вкладу в Ощадний банк отримані наступні данні: $n_1$=100; $n_2$=120; $n_3$=130; $\sigma_{1}^{2} = 22$; $\sigma_{2}^{2} = 21$; $\sigma_{3}^{2}$ =20.

З метою перевірки нульової гіпотези істотності відмінностей дисперсій отриманих з неоднакових вибіркових сукупностей, обчислимо необхідні параметри:

$$\overline{\sigma_{а}^{2}} = \frac{\sum \sigma_{i}^{2} n_{i}}{\sum n_{i}} = \frac{22 \cdot 100 + 21 \cdot 120 + 200 \cdot 130}{100 + 120 + 130} = \frac{7320}{350} = 20,91;$$

$$\lg \overline{\sigma_{а}^{2}} = \lg 20,91 = 1,3203;$$

$$\sum n_{i} \lg \overline{\sigma_{\alpha}^{2}} = 100 \cdot \lg 22 + 120 \cdot \lg 21 + 130 \lg 20 =$$
$$= 100 \cdot 1.3424 + 120 \cdot 1.3222 + 130 \cdot 1.3010 = 462.034;$$

$M= 2,3026 \cdot (\lg \overline{\sigma_{а}^{2}} \sum n_{i} - \sum n_{i}\ \lg \overline{\sigma_{г}^{2}})$=2,3026·(1,3203·350-462,034)=2,3026·(462,105-462,034)=2,3026·0,071=0,16348.

$$C=1+\frac{\sum\frac{1}{n_i}-\frac{1}{\sum n_i}}{3(m-1)}1+\frac{\frac{1}{100}+\frac{1}{120}+\frac{1}{130}-\frac{1}{350}}{3\cdot 2}=1+\frac{0,0231685}{6}=1,00386.$$

Розрахункова величина критерію Бартлета дорівнюватиме:

$$\chi^2_ф=\frac{M}{C}=\frac{0,16348}{1,00386}=0,163.$$

Табличне значення $\chi^2_T$ при ймовірності р=0,95 і числі ступенів вільності k=p-1=3-1=2 становитиме $\chi^2_{T(p=0,95)}=6,0.$

Оскільки $\chi^2_ф=0,163<\chi^2_{T(з=0,95)}=6,0$ можемо зробити висновок про неістотність відмінностей у дисперсіях, тобто ці відмінності є випадковими, а отже, результати статистичних спостережень підтверджують гіпотезу, яка перевірялась.

## 14.7 Перевірка гіпотези про належність спостережень, що виділяються, до досліджуваної генеральної сукупності

Під час проведення статистичного спостереження у складі зібраних даних можуть зустрічатися поодинокі зареєстровані значення ознак, які помітно відрізняються від загального рівня. Такі відмінності можуть виникати внаслідок: а) помилок у проведенні спостереження; б) випадкового збігу різного роду окремих несуттєвих обставин; в) порушення однорідності досліджуваної сукупності.

Механічно вилучати варіанти значень ознак, які сильно відрізняються від середньої арифметичної не можна. Для вилучення спостережень які виділяються із подальшої обробки потрібне застосування обґрунтованих статистичних критеріїв.

Припустимо, що розподіл результатів у звичайних умовах статистичних спостережень відповідає закону нормального розподілу з параметричними $(\overline{x})$ і $(\sigma)$. В результаті проведення однієї із серій таких спостережень отримані „n" значень $x_1 \le x_2 \le x_3 \le .......\le x_n$, серед яких є максимальне значення „$x_n$" чи мінімальне „$x_1$", що різко відрізняються за своєю величиною від решти „n-1" спостережень. Потрібно виявнити, чи відносяться ці значення до досліджуваної генеральної сукупності, чи їх поява є наслідком якихось випадкових неординарних обставин?

Нульовою гіпотезою в даному випадку є передбачення того, що „$x_n$» і „$x_1$" належать до тієї ж сукупності, як і всі інші „n-1" спостереження або порушення загальних умов формування рівня ознак в сукупності.

Перевірка цієї гіпотези заключається у порівнянні за величиною „$x_n$» чи „$x_1$" з певною критичною межею „x". Якщо спостереження, що виділяється є найбільше, тоді „$x_n$" порівнюють з верхньою допустимою межею, вибраною таким чином, щоб ймовірність перевершити її дорівнювала рівню значущості. В даному випадку матимемо критичну область такого вигляду:

$$p\,(x_n > \bar{x} + t \cdot \sigma) = \alpha\,.$$

Нульова гіпотеза відхиляється, якщо „$x_n$" перевершує за своєю величиною вказану межу.

Якщо ж спостереження, що виділяється є найменше „$x_1$", його порівнюють з нижньою допустимою межею, яка буде рівною:

$$\bar{x} - t \cdot \sigma, \text{ тобто } p(\,x_1 < \bar{x} - t \cdot \sigma) = \alpha\,.$$

Якщо одночасно досліджують максимальне і мінімальне значення ознак, то критична область матиме вигляд:

$$p(|x - \bar{x}| > t \cdot \sigma) = \alpha\,.$$

Розглянемо приклад. Нехай маємо наступні дані, наведені у таблиці 14.8.

Таблиця 14.8.

| Кількість спостережень | Мінімальні значення | | Максимальні значення | | Різниця сумірних значень | | Середнє значення | Середнє квадратичне відхилення у вибірці |
|---|---|---|---|---|---|---|---|---|
| | $x_1$ | $x_2$ | $x_{n-1}$ | $x_n$ | $x_2$-$x_1$ | $x_n$- $x_{n-1}$ | $\bar{x}$ | $\sigma$ |
| 100 | 10 | 16 | 140 | 188 | 6 | 48 | 67,3 | 45,3 |

Мінімальне значення „$x_1$" мало відрізняється від наступного за ним значення „$x_2$" у рангованому ряду, тому використаємо критичну область $p\,(x_n > \bar{x} + t \cdot \sigma)$ і перевіримо належність спостереження, що виділяється $x_n$=188 до результатів досліджуваної сукупності.

При рівні значущості $\alpha$=0,01 значення другої функції нормованого відхилення Лапласа для критичної області яка розглядається дорівнює:

$$\frac{1}{2}f(t) = \int_0^t e^{-\frac{t^2}{2}}dt = \frac{1}{2} - 0,01 = 0,49.$$

За таблицею другої функції нормованого відхилення знаходимо, що цьому значенню відповідає t= 2,33. Таким чином, верхня допустима межа значень ознаки, яка не може бути перевищена з ймовірністю 0,99 становитиме 172,8 (67,3+2,33 · 45,3). Значення яке виділяється $x_n$=188, виходить за розраховану межу, тому з ймовірністю 0,99 можна вважати, що $x_n$=188 не належить до досліджуваної сукупності і повинно бути виключене з подальших розрахунків.

На практиці часто виникають випадки, коли параметри генеральної сукупності "$\overline{x}$" і "σ" невідомі, а тому для перевірки гіпотези про спостереження які виділяються використовують відповідні параметри вибіркового спостереження. Однак, особливо при малих вибірках, ці оцінки є не цілком надійні. Тому для вилучення спостережень, що виділяються за результатами малих вибірок використовують критерій Ф. Груббса (К).

**Критерій Ф. Груббса (К)** базується на відношенні двох сум квадратів відхиленнь [8, 188]:

а) для випробовування найбільшого спостереження, яке виділяється у вибірковій сукупності обсягом «n» з нормальним розподілом, обчислюють відношення:

$$\frac{S_n^2}{S^2} = \frac{\sum_{i=1}^{n-1}(xi - \overline{x}_n)^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \leq К,$$

де $x_1 \leq x_2 \leq x_3 \leq .......\leq x_n$ ; $\overline{x}_n = \frac{\sum_{i=2}^{n-1} x_i}{n-1}$; $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$;

б) для випробовування найменшого спостереження, яке виділяється у вибірковій сукупності „n” з нормальним розподілом, обчислюють відношення:

$$\frac{S_n^2}{S^2} = \frac{\sum_{i=2}^{n}(x_i - \overline{x}_1)^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \leq К,$$

де $\overline{x}_1 = \frac{\sum_{i=2}^{n} x_i}{n-1}$.

Обчислену величину відношення ($К_ф$) порівнюють з табличною величиною ($К_Т$) при певному числі спостережень і заданому рівні значущості.

Табличний критерій Груббса ($К_ф$) характеризує ту граничну величину розбіжностей у сумах квадратів відхилень, котра з ймовірністю (1- $\alpha$ ) пояснюються випадковими причинами.

Якщо фактичний критерій буде рівним або меншим за табличний ($К_ф \leq К_Т$), тоді найбільше чи найменше спостереження не вилучається, а якщо фактичний критерій буде більший за табличний ($К_ф > К_Т$), то ймовірність розбіжностей у сумах квадратів відхилень внаслідок випадкових причин дорівнює рівню значущості „ $\alpha$ ” і в силу малої ймовірності вважається подією практично неможливою. В таких випадках, спостереження які виділяються вилучають і в подальших розрахунках використовують спостереження (n-1), що залишились.

Розглянемо застосування критерію Ф. Груббса на прикладі за такими даними: відхилення ваги деталей від нормальної наступні (г): 0,06; 0,09; 0,10; 0,11; 0,13; 0,14; 0,15; 0,16; 0,24. Потрібно встановити, чи не містять результати помилки спостереження, передбачаючи, що розподіл ваги деталей в генеральній сукупності відповідає закону нормального розподілу.

Обчислимо необхідні параметри.

Середні відхилення ваги всіх деталей:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{1,18}{9} = 0,1311г.$$

Вилучимо максимальне відхилення ваги деталі 0,24

$$\bar{x}_n = \frac{\sum x_i - x_n}{n-1} = \frac{1,18-0,24}{9-1} = \frac{0,94}{8} = 0,1175 \text{ г.}$$

Суму квадратів відхилення від „ $\bar{x}$ ”:

$$\sum(x_i - \bar{x}_n)^2 = [(0,06\text{-}0,1311)^2 + (0,09\text{-}0,1311)^2 + (0,10\text{-}0,1311)^2 + \\ +(0,11\text{-}0,1311)^2 - (0,13\text{-}0,1311)^2 + (0,14\text{-}0,1311)^2 + (0,15\text{-}0,1311)^2 + \\ +(0,16\text{-}0,1311)^2 + (0,24\text{-}0,1311)^2 = 0,021289.$$

Суму квадратів відхилення (n-1) значень від „ $\bar{x}_n$ ”:

$$\sum(x_i - \bar{x}_n)^2 = [(0,06\text{-}0,1175)^2 + (0,09\text{-}0,1175)^2 + (0,10\text{-}0,1175)^2 + \\ +(0,11\text{-}0,1175)^2 + (0,13\text{-}0,1175)^2 + (0,14\text{-}0,1175)^2 + (0,15\text{-}0,1175)^2 + \\ +(0,16\text{-}0,1175)^2 = 0,007952.$$

Відхилення двох сум квадратів відхилення:

$$К_ф = \frac{\sum(x_i - \bar{x}_n)^2}{\sum(x_i - \bar{x})^2} = \frac{0,007952}{0,021289} = 0,3735.$$

Таблиця 14.9

Витяг із таблиці Ф. Груббса.

| Кількість спостережень (n) | При рівні значущості ($\alpha$) | | Кількість спостережень (n) | Кількість спостережень (n) | |
|---|---|---|---|---|---|
| | 0,01 | 0,05 | | 0,01 | 0,05 |
| 3 | 0,0001 | 0,0027 | ... | ... | ... |
| 4 | 0,0100 | 0,0494 | 15 | 0,4401 | 0,5559 |
| 5 | 0,0442 | 0,1270 | ... | ... | ... |
| 6 | 0,0928 | 0,2032 | 20 | 0,5393 | 0,6379 |
| 7 | 0,1447 | 0,2696 | ... | ... | ... |
| 8 | 0,1948 | 0,3261 | 25 | 0,6071 | 0,6923 |
| 9 | 0,2411 | 0,3742 | ... | ... | ... |
| 10 | 0,2931 | 0,4154 | ... | ... | ... |

При кількості спостережень n=9 і рівні значущості 0,01 табличний критерій $К_Т$=0,2411 менший фактичного ($К_ф$=0,3735 > $К_{Т(0,01)}$=0,2411), тому відхилення від номінальної ваги 0,24 г потрібно віднести до помилок спостереження, із досліджуваної сукупності вилучити і подальші розрахунки проводити з кількості спостережень, що залишились (n-1=9-1=8).

**Критерій Дж. Ірвіна ($\lambda$)** також використовується як один з варіантів перевірки статистичної гіпотези про належність спостережень, які виділяються, до досліджуваних даних генеральної сукупності.

Цей критерій визначається як відношення різниці між „$\overline{x}_n$” і „$\overline{x}_{n-1}$” до середнього квадратичного відхилення в генеральній сукупності.

$$\lambda = \frac{x_n - x_{n-1}}{\sigma}, \text{ де } x_1 \le x_2 \le x_3 \le .......\le x_{n-1} \le x_n;$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}.$$

Таблиця 14.10

Значення Р (λ), де λ критерій Дж. Ірвіна (витяг)

| Обсяг вибірки (n) | Значення ймовірності при λ, яке дорівнює | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1,0 | 1,1 | 1,2 | 1,3 | 1,4 | 1,5 | 1,6 | 1,7 | 1,8 | 1,9 | 2,0 |
| 10 | 0,152 | 0,121 | 0,096 | 0,075 | 0,059 | 0,045 | 0,038 | 0,026 | 0,020 | 0,015 | 0,011 |
| 20 | 0,107 | 0,082 | 0,062 | 0,047 | 0,035 | 0,026 | 0,019 | 0,014 | 0,010 | 0,007 | 0,005 |
| 30 | 0,089 | 0,068 | 0,050 | 0,037 | 0,027 | 0,020 | 0,014 | 0,010 | 0,007 | 0,005 | 0,004 |
| ... | | | | | | | | | | | |
| 60 | 0,065 | 0,048 | 0,034 | 0,025 | 0,017 | 0,012 | 0,009 | 0,006 | 0,004 | 0,003 | 0,002 |
| 70 | 0,061 | 0,044 | 0,032 | 0,022 | 0,016 | 0,011 | 0,008 | 0,005 | 0,004 | 0,002 | 0,002 |
| 80 | 0,058 | 0,041 | 0,030 | 0,021 | 0,015 | 0,010 | 0,007 | 0,005 | 0,003 | 0,002 | 0,001 |

У таблиці Дж. Ірвіна наводиться ймовірність того, що різниця між „$\overline{x}_n$” і „$\overline{x}_{n-1}$” відрізняється від середнього квадратичного відхилення у власне випадковій вибірці обсягом „n” одиниць більше, ніж у λ разів.

Використовуючи дані прикладу, наведеного у табл. 14.8, визначимо критерій Ірвіна λ:

$$\lambda = \frac{x_n - x_{n-1}}{\sigma} = \frac{48,0}{45,3} = 1,06.$$

За даними табл.14.10, при n≥ 80 і λ=1,06 Р(λ)=0,05 ($\frac{0,058+0,041}{2}$).

У зв'язку з малою ймовірністю появи таких розбіжностей між суміжними значеннями ($x_n$-$x_{n-1}$) у рангованому ряду отриманих даних можна стверджувати, що значення $x_n$=188, не належить до досліджуваної сукупності, і в подальшій обробці інформації використовувати обсяг сукупності, що залишилась, тобто 99 значень (100-1).

### *Контрольні запитання*

1. Що собою являє статистична гіпотеза?
2. Що називається статистичною гіпотезою?
3. Що таке нульова гіпотеза?
4. Для чого використовується альтернативна гіпотеза?
5. Які помилки виникають при перевірці статистичних гіпотез?
6. Що собою являють помилки першого і другого роду?
7. Що називається статистичним критерієм?
8. Які види критеріїв використовує статистика при перевірці гіпотез?

9. Коли застосовують параметричні критерії?
10. Коли застосовують непараметричні критерії?
11. Що таке критична область?
12. Що розуміється під областю допустимих значень?
13. Що називають потужністю критерію?
14. Охарактеризуйте загальну схему перевірки статистичних гіпотез.
15. Охарактеризуйте загальну схему перевірки статистичних гіпотез відносно середніх величин.
16. За якою формулою і для чого вираховують t – критерій Стьюдента?
17. Охарактеризуйте загальну схему перевірки статистичних гіпотез відносно законів розподілу.
18. Що собою являє крива нормального розподілу і за якою формулою вона визначається?
19. За якою формулою визначаються теоретичні частоти?
20. За якою формулою і для чого обчисляють критерій Пірсона ($\chi^2$)?
21. За якою формулою і для чого вираховують критерій згоди А.Н.Колмогорова ($\lambda$)?
22. За якою формулою і з якою метою визначають критерій згоди В.І.Романовського (R)?
23. За якою формулою визначається і для чого використовується критерій згоди Б.С.Ястремського?
24. Охарактеризуйте загальну схему перевірки статистичних гіпотез про істотність розбіжностей між дисперсіями.
25. Що собою являє і за якою формулою обчислюється F-критерій Фішера?
26. За якою формулою і для чого вираховують критерій Кохрана (q)?
27. Що собою являє критерій Бартлета (М)?
28. Охарактеризуйте загальну схему перевірки гіпотез про наявність спостережень, що виділяються, до досліджуваної генеральної сукупності?
29. За якою формулою і з якою метою обчисляють критерій Ф.Груббса (К)?
30. За якою формулою і для чого вираховують критерій Дж. Ірвіна ($\lambda$)?
31. Наведіть приклади застосування критеріїв перевірки статистичних гіпотез.

# ДОДАТКИ

**Додаток 1.**

## Таблиця вирахувань значень за рядом Фур'є

Для визначення сезонності як періодичної функції Фур'є за $\pi$ береться число місяців року (число кварталів за три роки), тоді ряд динаміки по відношенню до значень визначається у вигляді наступних значень $y$:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | $\pi/6$ | $\pi/3$ | $\pi/2$ | $2\pi/3$ | $5\pi/6$ | $\pi$ | $7\pi/6$ | $4\pi/3$ | $3\pi/2$ | $5\pi/3$ | $11\pi/6$ |
| $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | $y_8$ | $y_9$ | $y_{10}$ | $y_{11}$ | $y_{12}$ |

Значення ***cos k t*** і ***sin k t*** для різних значень *t*

| ***t*** | ***cos t*** | ***cos 2t*** | ***sin t*** | ***sin 2t*** |
|---|---|---|---|---|
| **0** | 1 | 1 | 0 | 0 |
| $\pi/6$ | 0,866 | 0,5 | 0,5 | 0,866 |
| $\pi/3$ | 0,5 | -0,5 | 0,866 | -0,866 |
| $\pi/2$ | 1 | -1 | 1 | 0 |
| $2\pi/3$ | -0,5 | -0,5 | 0,866 | -0,866 |
| $5\pi/6$ | -0,866 | 0,5 | 0,5 | -0,866 |
| $\pi$ | -1 | 1 | 0 | 0 |
| $7\pi/6$ | -0,866 | 0,5 | -0,5 | 0,866 |
| $4\pi/3$ | -0,5 | -0,5 | -0,866 | 0,866 |
| $3\pi/2$ | 0 | -1 | -1 | 0 |
| $5\pi/3$ | 0,5 | -0,5 | -0,866 | -0,866 |
| $11\pi/6$ | 0,866 | 0,5 | -0,5 | -0,866 |

**Додаток 2.**

## Таблиця 5 %-го і 1 %-го рівнів імовірностей коефіцієнтів кореляції

| Розмір вибірки | Додатні значення | | Від'ємні значення | |
|---|---|---|---|---|
| | 5 %-ний рівень | 1 %-ний рівень | 5 %-ний рівень | 1 %-ний рівень |
| 5 | 0,253 | 0,297 | -0,753 | -0,798 |
| 6 | 0,354 | 0,447 | -0,708 | -0,863 |
| 7 | 0,370 | 0,510 | -0,674 | -0,799 |
| 8 | 0,371 | 0,531 | -0,625 | -0,764 |
| 9 | 0,366 | 0,533 | -0,593 | -0,737 |
| 10 | 0,360 | 0,525 | -0,564 | -0,705 |
| 11 | 0,353 | 0,515 | -0,539 | -0,679 |
| 12 | 0,348 | 0,505 | -0,516 | -0,655 |
| 13 | 0,341 | 0,495 | -0,497 | -0,634 |
| 14 | 0,335 | 0,485 | -0,479 | -0,615 |
| 15 | 0,328 | 0,475 | -0,462 | -0,697 |
| 20 | 0,299 | 0,342 | -0,399 | -0,524 |
| 25 | 0,276 | 0,398 | -0,356 | -0,473 |
| 30 | 0,257 | 0,370 | -0,324 | -0,433 |
| 35 | 0,242 | 0,347 | -0,300 | -0,401 |
| 40 | 0,229 | 0,329 | -0,279 | -0,376 |
| 45 | 0,218 | 0,313 | -0,262 | -0,256 |
| 50 | 0,208 | 0,301 | -0,248 | -0,339 |

**<u>Додаток 3</u>.**

**Значення функції**

$$\Phi(X)=\frac{2}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt=\frac{1}{\sqrt{2\pi}}\int_{-x}^{+x} e^{-\frac{t^2}{2}}dt$$

| Цілі і десяті долі | Соті долі | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 0,0 | 0,0000 | 0,0080 | 0,0160 | 0,0239 | 0,0319 | 0,0399 | 0,0478 | 0,0558 | 0,0638 | 0,0717 |
| 0,1 | 0,0797 | 0,0876 | 0,0955 | 0,1034 | 0,1113 | 0,1192 | 0,1271 | 0,1350 | 0,1428 | 0,1507 |
| 0,2 | 0,1585 | 0,1663 | 0,1741 | 0,1819 | 0,1897 | 0,1974 | 0,2051 | 0,2128 | 0,2205 | 0,2282 |
| 0,3 | 0,2358 | 0,2434 | 0,2510 | 0,2586 | 0,2661 | 0,2737 | 0,2812 | 0,2886 | 0,2960 | 0,3035 |
| 0,4 | 0,3108 | 0,3182 | 0,3255 | 0,3328 | 0,3401 | 0,3473 | 0,3545 | 0,3616 | 0,3688 | 0,3759 |
| 0,5 | 0,3829 | 0,3899 | 0,3969 | 0,4039 | 0,4108 | 0,4177 | 0,4245 | 0,4313 | 0,4381 | 0,4448 |
| 0,6 | 0,4515 | 0,4581 | 0,4647 | 0,4713 | 0,4778 | 0,4843 | 0,4907 | 0,4971 | 0,5035 | 0,5098 |
| 0,7 | 0,5161 | 0,5223 | 0,5285 | 0,5346 | 0,5407 | 0,5467 | 0,5527 | 0,5587 | 0,5646 | 0,6705 |
| 0,8 | 0,5763 | 0,5821 | 0,5878 | 0,5935 | 0,5991 | 0,6047 | 0,6102 | 0,6157 | 0,6211 | 0,6265 |
| 0,9 | 0,6319 | 0,6372 | 0,6424 | 0,6476 | 0,6528 | 0,6579 | 0,6629 | 0,6679 | 0,6729 | 0,7778 |
| 1,0 | 0,6827 | 0,6875 | 0,6923 | 0,6970 | 0,7017 | 0,7063 | 0,7109 | 0,7154 | 0,7199 | 0,7243 |
| 1,1 | 0,7287 | 0,7330 | 0,7373 | 0,7415 | 0,7457 | 0,7499 | 0,7540 | 0,7580 | 0,7620 | 0,7660 |
| 1,2 | 0,7699 | 0,7737 | 0,7775 | 0,7813 | 0,7850 | 0,7887 | 0,7923 | 0,7959 | 0,7994 | 0,8029 |
| 1,3 | 0,8064 | 0,8098 | 0,8132 | 0,8165 | 0,8198 | 0,8230 | 0,8262 | 0,8293 | 0,8324 | 0,8355 |
| 1,4 | 0,8385 | 0,8415 | 0,8444 | 0,8473 | 0,8501 | 0,8529 | 0,8557 | 0,8584 | 0,8611 | 0,8638 |
| 1,5 | 0,8654 | 0,8690 | 0,8715 | 0,8740 | 0,8764 | 0,8789 | 0,8812 | 0,8836 | 0,8859 | 0,8882 |

**Продовження додатку 3.**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1,6 | 0,8904 | 0,8926 | 0,8948 | 0,8969 | 0,8990 | 0,9011 | 0,9031 | 0,9051 | 0,9070 | 0,9090 |
| 1,7 | 0,9109 | 0,9127 | 0,9146 | 0,9164 | 0,9181 | 0,9199 | 0,9216 | 0,9233 | 0,9249 | 0,9265 |
| 1,8 | 0,9281 | 0,9297 | 0,9312 | 0,9327 | 0,9342 | 0,9357 | 0,9371 | 0,9385 | 0,9399 | 0,9412 |
| 1,9 | 0,9426 | 0,9439 | 0,9451 | 0,9464 | 0,9476 | 0,9488 | 0,9500 | 0,9512 | 0,9523 | 0,9534 |
| 2,0 | 0,9545 | 0,9556 | 0,9566 | 0,9576 | 0,9586 | 0,9596 | 0,9606 | 0,9616 | 0,9625 | 0,9634 |
| 2,1 | 0,9643 | 0,9651 | 0,9660 | 0,9668 | 0,9676 | 0,9684 | 0,9692 | 0,9700 | 0,9707 | 0,9715 |
| 2,2 | 0,9722 | 0,9729 | 0,9736 | 0,9743 | 0,9749 | 0,9756 | 0,9762 | 0,9768 | 0,9774 | 0,9780 |
| 2,3 | 0,9786 | 0,9791 | 0,9797 | 0,9802 | 0,9807 | 0,9812 | 0,9817 | 0,9822 | 0,9827 | 0,9832 |
| 2,4 | 0,9836 | 0,9841 | 0,9845 | 0,9849 | 0,9853 | 0,9857 | 0,9861 | 0,9865 | 0,9869 | 0,9872 |
| 2,5 | 0,9876 | 0,9879 | 0,9883 | 0,9886 | 0,9889 | 0,9892 | 0,9895 | 0,9898 | 0,9901 | 0,9904 |
| 2,6 | 0,9907 | 0,9910 | 0,9912 | 0,9915 | 0,9917 | 0,9920 | 0,9922 | 0,9924 | 0,9926 | 0,9928 |
| 2,7 | 0,9931 | 0,9933 | 0,9935 | 0,9937 | 0,9939 | 0,9940 | 0,9942 | 0,9944 | 0,9946 | 0,9947 |
| 2,8 | 0,9949 | 0,9951 | 0,9952 | 0,9953 | 0,9955 | 0,9965 | 0,9958 | 0,9959 | 0,9960 | 0,9961 |
| 2,9 | 0,9963 | 0,9964 | 0,9965 | 0,9966 | 0,9967 | 0,9968 | 0,9969 | 0,9970 | 0,9971 | 0,9972 |
| 3,0 | 0,9973 | 0,9974 | 0,9975 | 0,9976 | 0,9976 | 0,9977 | 0,9978 | 0,9979 | 0,9979 | 0,9980 |
| 3,1 | 0,9981 | 0,9981 | 0,9982 | 0,9983 | 0,9983 | 0,9984 | 0,9984 | 0,9985 | 0,9985 | 0,9986 |
| 3,2 | 0,9986 | 0,9987 | 0,9987 | 0,9988 | 0,9989 | 0,9989 | 0,9989 | 0,9990 | 0,9990 | 0,9990 |
| 3,3 | 0,9990 | 0,9991 | 0,9991 | 0,9991 | 0,9992 | 0,9992 | 0,9992 | 0,9992 | 0,9993 | 0,9993 |
| 3,4 | 0,9993 | 0,9994 | 0,9994 | 0.9994 | 0,9994 | 0,9994 | 0,9995 | 0,9995 | 0,9995 | 0,9995 |
| 3,5 | 0,9995 | 0,9996 | 0,9996 | 0,9996 | 0,9996 | 0,9996 | 0,9996 | 0,9996 | 0,9997 | 0,9997 |
| 3,6 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9998 | 0,9998 | 0,9998 |
| 3,7 | 0,9998 | 0,9998 | 0,9998 | 0,9998 | 0,9998 | 0,9998 | 0,9998 | 0,9998 | 0,9998 | 0,9998 |
| 3,8 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 |
| 3,9 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 |
| 4,0 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 |

**Додаток 4.**

**Значення функції** $$f(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$$

| Цілі і десяті долі | Соті долі | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 0,0 | 0,3989 | 0,3989 | 0,3989 | 0,3988 | 0,3986 | 0,3984 | 0,3982 | 0,3980 | 0,3977 | 0,3973 |
| 0,1 | 3970 | 3965 | 3956 | 3956 | 3951 | 3945 | 3939 | 3932 | 3925 | 3918 |
| 0,2 | 3910 | 3902 | 3894 | 3885 | 3876 | 3867 | 3857 | 3847 | 3836 | 3825 |
| 0,3 | 3814 | 3802 | 3790 | 3778 | 3765 | 3752 | 3739 | 3726 | 3712 | 3697 |
| 0,4 | 3683 | 3668 | 3653 | 3637 | 3621 | 3605 | 3589 | 3572 | 3555 | 3538 |
| 0,5 | 3521 | 3503 | 3485 | 3467 | 3448 | 3429 | 3410 | 3391 | 3372 | 3352 |
| 0,6 | 3332 | 3312 | 3292 | 3271 | 3251 | 3230 | 3209 | 3187 | 3166 | 3144 |
| 0,7 | 3123 | 3101 | 3079 | 3056 | 3034 | 3011 | 2989 | 2966 | 2943 | 2920 |
| 0,8 | 2897 | 2874 | 2850 | 2827 | 2803 | 2780 | 2756 | 2732 | 2709 | 2685 |
| 0,9 | 2661 | 2637 | 2613 | 2589 | 2565 | 2541 | 2516 | 2492 | 2468 | 2444 |
| 1,0 | 2420 | 2396 | 2371 | 2347 | 2323 | 2299 | 2275 | 2251 | 2227 | 2203 |
| 1,1 | 2179 | 2155 | 2131 | 2107 | 2083 | 2059 | 2036 | 2012 | 1989 | 1965 |
| 1,2 | 1942 | 1919 | 1895 | 1872 | 1849 | 1826 | 1804 | 1781 | 1758 | 1736 |
| 1,3 | 1714 | 1691 | 1669 | 1647 | 1626 | 1604 | 1582 | 1561 | 1539 | 1518 |
| 1,4 | 1497 | 1476 | 1456 | 1435 | 1415 | 1394 | 1374 | 1354 | 1334 | 1315 |
| 1,5 | 1295 | 1276 | 1257 | 1238 | 1219 | 1200 | 1182 | 1163 | 1145 | 1127 |
| 1,6 | 1109 | 1092 | 1074 | 1057 | 1040 | 1023 | 1006 | 0989 | 0973 | 0957 |
| 1,7 | 0940 | 0925 | 0909 | 0893 | 0878 | 0863 | 0848 | 0833 | 0818 | 0804 |
| 1,8 | 0790 | 0775 | 0761 | 0748 | 0734 | 0721 | 0707 | 0694 | 0681 | 0669 |
| 1,9 | 0656 | 0644 | 0632 | 0620 | 0608 | 0596 | 0584 | 0573 | 0562 | 0551 |
| 2,0 | 0,0540 | 0,0529 | 0,0519 | 0,0508 | 0,0498 | 0,0488 | 0,0478 | 0,0468 | 0,0459 | 0,0449 |
| 2,1 | 0440 | 0431 | 0422 | 0413 | 0404 | 0396 | 0387 | 0379 | 0371 | 0363 |
| 2,2 | 0355 | 0347 | 0339 | 0332 | 0325 | 0317 | 0310 | 0303 | 0297 | 0290 |

**Продовження додатку 4.**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|
| 2,3 | 0283 | 0277 | 0270 | 0264 | 0258 | 0252 | 0246 | 0241 | 0235 | 0229 |
| 2,4 | 0224 | 0219 | 0213 | 0208 | 0203 | 0198 | 0194 | 0189 | 0184 | 0180 |
| 2,5 | 0175 | 0171 | 0167 | 0163 | 0158 | 0154 | 0151 | 0147 | 0143 | 0139 |
| 2,6 | 0136 | 0132 | 0129 | 0126 | 0122 | 0119 | 0116 | 0113 | 0110 | 0107 |
| 2,7 | 0104 | 0101 | 0099 | 0096 | 0093 | 0091 | 0088 | 0086 | 0084 | 0081 |
| 2,8 | 0079 | 0077 | 0075 | 0073 | 0071 | 0069 | 0067 | 0065 | 0063 | 0061 |
| 2,9 | 0060 | 0058 | 0056 | 0055 | 0053 | 0051 | 0050 | 0048 | 0047 | 0046 |
| 3,0 | 0,0044 | 0,0043 | 0,0042 | 0,0041 | 0,0039 | 0,0038 | 0,0037 | 0,0036 | 0,0035 | 0,0034 |
| 3,1 | 0033 | 0032 | 0031 | 0030 | 0029 | 0028 | 0027 | 0026 | 0025 | 0025 |
| 3,2 | 0024 | 0023 | 0022 | 0022 | 0021 | 0020 | 0020 | 0019 | 0018 | 0018 |
| 3,3 | 0017 | 0017 | 0016 | 0016 | 0015 | 0015 | 0014 | 0014 | 0013 | 0013 |
| 3,4 | 0012 | 0012 | 0012 | 0011 | 0011 | 0010 | 0010 | 0010 | 0009 | 0009 |
| 3,5 | 0009 | 0008 | 0008 | 0008 | 0008 | 0007 | 0007 | 0007 | 0007 | 0006 |
| 3,6 | 0006 | 0006 | 0006 | 0005 | 0005 | 0005 | 0005 | 0005 | 0005 | 0004 |
| 3,7 | 0004 | 0004 | 0004 | 0004 | 0004 | 0004 | 0003 | 0003 | 0003 | 0003 |
| 3,8 | 0003 | 0003 | 0003 | 0003 | 0003 | 0002 | 0002 | 0002 | 0002 | 0002 |
| 3,9 | 0002 | 0002 | 0002 | 0002 | 0002 | 0002 | 0002 | 0002 | 0001 | 0001 |
| 4,0 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
| 4,1 | 0,0001338 | | | | | | | | | |
| 4,5 | 0,0000160 | | | | | | | | | |
| 5,0 | 0,0000015 | | | | | | | | | |

**Додаток 5.**

## Інтегральна функція нормального розподілу

$$F(t)=\frac{2}{\sqrt{2\pi}}\int_0^t e^{-\frac{t^2}{2}}dt$$

| *t* | *F(t)* | *t* | *F(t)* | *t* | *F(t)* | *t* | *F(t)* |
|---|---|---|---|---|---|---|---|
| 0,0 | 0,0000 | 1,0 | 0,6827 | 2,0 | 0,9545 | 3,0 | 0,9973 |
| 0,1 | 0,0797 | 1,1 | 0,7287 | 2,1 | 0,9643 | 3,1 | 0,9981 |
| 0,2 | 0,1585 | 1,2 | 0,7699 | 2,2 | 0,9722 | 3,2 | 0,9986 |
| 0,3 | 0,2358 | 1,3 | 0,8064 | 2,3 | 0,9786 | 3,3 | 0,9990 |
| 0,4 | 0,3108 | 1,4 | 0,8385 | 2,4 | 0,9836 | 3,4 | 0,9993 |
| 0,5 | 0,3829 | 1,5 | 0,8664 | 2,5 | 0,9676 | 3,5 | 0,9995 |
| 0,6 | 0,4515 | 1,6 | 0,8904 | 2,6 | 0,9907 | 3,6 | 0,9997 |
| 0,7 | 0,5161 | 1,7 | 0,9109 | 2,7 | 0,9931 | 3,7 | 0,9998 |
| 0,8 | 0,5763 | 1,8 | 0,9281 | 2,8 | 0,9949 | 3,8 | 0,9999 |
| 0,9 | 0,6319 | 1,9 | 0,9426 | 2,9 | 0,9963 | 3,9 | 0,9999 |

## Додаток 6.

Значення п'яти- і однопроцентних верхніх меж відхилення величини $F$ в залежності від ступенів вільності $k_1$ і $k_2$ (п'ятипроцентні межі $F_5$, $k_1$ $k_2$, набрані звичайним шрифтом, однопроцентні $F_1$ $k_1$ $k_2$ – жирним)

| $k_2$ – степені вільності для меншої дисперсії; $k_2$ | $k_1$ – степені вільності для більшої дисперсії: 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 14 | 16 | 20 | 24 | 30 | 40 | 50 | 75 | 100 | 200 | 500 | ∞ | $k_2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 161 | 200 | 216 | 225 | 230 | 234 | 237 | 239 | 241 | 242 | 243 | 244 | 245 | 246 | 248 | 249 | 250 | 251 | 252 | 253 | 253 | 254 | 254 | 254 | 1 |
| | **4052** | **4999** | **5403** | **5625** | **5764** | **5889** | **5928** | **5981** | **6022** | **6056** | **6082** | **6106** | **6142** | **6169** | **6208** | **6234** | **6258** | **6286** | **6302** | **6323** | **6334** | **6352** | **6361** | **6366** | |
| 2 | 18,51 | 19,00 | 19,16 | 19,25 | 19,30 | 19,33 | 19,36 | 19,37 | 19,38 | 19,39 | 19,40 | 19,41 | 19,42 | 19,43 | 19,44 | 19,45 | 19,46 | 19,47 | 19,47 | 19,48 | 19,49 | 19,49 | 19,50 | 19,50 | 2 |
| | **98,49** | **99,01** | **99,17** | **99,25** | **99,30** | **99,33** | **99,34** | **99,36** | **99,38** | **99,40** | **99,41** | **99,42** | **99,43** | **99,44** | **99,45** | **99,46** | **99,47** | **99,48** | **99,48** | **99,49** | **99,49** | **99,49** | **99,50** | **99,50** | |
| 3 | 10,13 | 9,55 | 9,28 | 9,12 | 9,01 | 8,94 | 8,88 | 8,84 | 8,81 | 8,78 | 8,76 | 8,74 | 8,71 | 8,69 | 8,66 | 8,64 | 8,62 | 8,60 | 8,58 | 8,57 | 8,56 | 8,54 | 8,54 | 8,53 | 3 |
| | **34,12** | **30,81** | **29,46** | **28,71** | **28,24** | **27,91** | **27,67** | **27,49** | **27,34** | **27,23** | **27,13** | **27,05** | **26,92** | **26,83** | **26,69** | **26,60** | **26,50** | **26,41** | **26,35** | **26,27** | **26,23** | **26,18** | **26,14** | **26,12** | |
| 4 | 7,71 | 6,94 | 6,59 | 6,39 | 6,26 | 6,16 | 6,09 | 6,04 | 6,00 | 5,96 | 5,93 | 5,91 | 5,87 | 5,84 | 5,80 | 5,77 | 5,74 | 5,71 | 5,70 | 5,68 | 5,66 | 5,65 | 5,64 | 5,63 | 4 |
| | **21,20** | **18,00** | **16,69** | **15,98** | **15,52** | **15,21** | **14,98** | **14,80** | **14,66** | **14,54** | **14,45** | **14,37** | **14,24** | **14,15** | **14,02** | **13,93** | **13,83** | **13,74** | **13,69** | **13,61** | **13,57** | **13,52** | **13,48** | **13,46** | |
| 5 | 6,61 | 5,79 | 5,41 | 5,19 | 5,05 | 4,95 | 4,88 | 4,82 | 4,78 | 4,74 | 4,70 | 4,68 | 4,64 | 4,60 | 4,56 | 4,53 | 4,50 | 4,46 | 4,44 | 4,42 | 4,40 | 4,38 | 4,37 | 4,36 | 5 |
| | **16,26** | **13,27** | **12,06** | **11,39** | **10,97** | **10,67** | **10,45** | **10,27** | **10,15** | **10,05** | **9,96** | **9,89** | **9,77** | **9,68** | **9,55** | **9,47** | **9,38** | **9,29** | **9,24** | **9,17** | **9,13** | **9,07** | **9,04** | **9,02** | |
| 6 | 5,99 | 5,14 | 4,76 | 4,53 | 4,39 | 4,28 | 4,21 | 4,15 | 4,10 | 4,06 | 4,03 | 4,00 | 3,96 | 3,92 | 3,87 | 3,84 | 3,81 | 3,77 | 3,75 | 3,72 | 3,71 | 3,69 | 3,68 | 3,67 | 6 |
| | **13,74** | **10,92** | **9,78** | **9,15** | **8,75** | **8,47** | **8,26** | **8,10** | **7,98** | **7,87** | **7,79** | **7,72** | **7,60** | **7,52** | **7,39** | **7,31** | **7,23** | **7,14** | **7,09** | **7,02** | **6,99** | **6,94** | **6,90** | **6,88** | |
| 7 | 5,59 | 4,74 | 4,35 | 4,12 | 3,97 | 3,87 | 3,79 | 3,73 | 3,68 | 3,63 | 3,60 | 3,57 | 3,52 | 3,49 | 3,44 | 3,41 | 3,38 | 3,34 | 3,32 | 3,29 | 3,28 | 3,25 | 3,24 | 3,23 | 7 |
| | **12,25** | **9,55** | **8,45** | **7,85** | **7,46** | **7,19** | **7,00** | **6,84** | **6,71** | **6,62** | **6,54** | **6,47** | **6,35** | **6,27** | **6,15** | **6,07** | **5,98** | **5,90** | **5,85** | **5,78** | **5,75** | **5,70** | **5,67** | **5,65** | |
| 8 | 5,32 | 4,46 | 4,07 | 3,84 | 3,69 | 3,58 | 3,50 | 3,44 | 3,39 | 3,34 | 3,31 | 3,28 | 3,23 | 3,20 | 3,15 | 3,12 | 3,08 | 3,05 | 3,03 | 3,00 | 2,98 | 2,96 | 2,94 | 2,93 | 8 |
| | **11,26** | **8,65** | **7,59** | **7,01** | **6,63** | **6,37** | **6,19** | **6,03** | **5,91** | **5,82** | **5,74** | **5,67** | **5,56** | **5,48** | **5,36** | **5,28** | **5,20** | **5,11** | **5,06** | **5,00** | **4,96** | **4,91** | **4,88** | **4,86** | |
| 9 | 5,12 | 4,26 | 3,86 | 3,63 | 3,48 | 3,37 | 3,29 | 3,23 | 3,18 | 3,13 | 3,10 | 3,07 | 3,02 | 2,98 | 2,93 | 2,90 | 2,86 | 2,82 | 2,80 | 2,77 | 2,76 | 2,73 | 2,72 | 2,71 | 9 |
| | **10,56** | **8,02** | **6,99** | **6,42** | **6,06** | **5,80** | **5,62** | **5,47** | **5,35** | **5,26** | **5,18** | **5,11** | **5,00** | **4,92** | **4,80** | **4,73** | **4,64** | **4,56** | **4,51** | **4,45** | **4,41** | **4,36** | **4,33** | **4,31** | |
| 10 | 4,96 | 4,10 | 3,71 | 3,48 | 3,33 | 3,22 | 3,14 | 3,07 | 3,02 | 2,97 | 2,94 | 2,91 | 2,86 | 2,82 | 2,77 | 2,74 | 2,70 | 2,67 | 2,64 | 2,61 | 2,59 | 2,56 | 2,55 | 2,54 | 10 |
| | **10,04** | **7,56** | **6,55** | **5,99** | **5,64** | **5,39** | **5,21** | **5,06** | **4,95** | **4,85** | **4,78** | **4,71** | **4,60** | **4,52** | **4,41** | **4,33** | **4,25** | **4,17** | **4,12** | **4,05** | **4,01** | **3,96** | **3,93** | **3,91** | |
| 11 | 4,84 | 3,98 | 3,59 | 3,36 | 3,20 | 3,09 | 3,01 | 2,95 | 2,90 | 2,86 | 2,82 | 2,79 | 2,74 | 2,70 | 2,65 | 2,61 | 2,57 | 2,53 | 2,50 | 2,47 | 2,45 | 2,42 | 2,41 | 2,40 | 11 |
| | **9,85** | **7,20** | **6,22** | **5,67** | **5,32** | **5,07** | **4,88** | **4,74** | **4,63** | **4,54** | **4,46** | **4,40** | **4,29** | **4,21** | **4,10** | **4,02** | **3,94** | **3,86** | **3,80** | **3,74** | **3,70** | **3,66** | **3,62** | **3,60** | |
| 12 | 4,75 | 3,88 | 3,49 | 3,26 | 3,11 | 3,00 | 2,92 | 2,85 | 2,80 | 2,76 | 2,72 | 2,69 | 2,64 | 2,60 | 2,54 | 2,50 | 2,46 | 2,42 | 2,40 | 2,36 | 2,35 | 2,32 | 2,31 | 2,30 | 12 |
| | **9,33** | **6,93** | **5,95** | **5,41** | **5,06** | **4,82** | **4,65** | **4,50** | **4,39** | **4,30** | **4,22** | **4,16** | **4,05** | **3,98** | **3,86** | **3,78** | **3,70** | **3,61** | **3,56** | **3,49** | **3,46** | **3,41** | **3,38** | **3,36** | |
| 13 | 4,67 | 3,80 | 3,41 | 3,18 | 3,02 | 2,92 | 2,84 | 2,77 | 2,72 | 2,67 | 2,63 | 2,60 | 2,55 | 2,51 | 2,46 | 2,42 | 2,38 | 2,34 | 2,32 | 2,28 | 2,26 | 2,24 | 2,22 | 2,21 | 13 |
| | **9,07** | **6,70** | **5,74** | **5,20** | **4,86** | **4,62** | **4,44** | **4,30** | **4,19** | **4,10** | **4,02** | **3,96** | **3,85** | **3,78** | **3,67** | **3,59** | **3,51** | **3,42** | **3,37** | **3,30** | **3,27** | **3,21** | **3,18** | **3,16** | |
| 14 | 4,60 | 3,74 | 3,34 | 3,11 | 2,96 | 2,85 | 2,77 | 2,70 | 2,65 | 2,60 | 2,56 | 2,53 | 2,48 | 2,44 | 2,39 | 2,35 | 2,31 | 2,27 | 2,24 | 2,21 | 2,19 | 2,16 | 2,14 | 2,13 | 14 |
| | **8,86** | **6,51** | **5,56** | **5,03** | **4,69** | **4,46** | **4,28** | **4,14** | **4,03** | **3,94** | **3,86** | **3,80** | **3,70** | **3,62** | **3,51** | **3,43** | **3,34** | **3,26** | **3,21** | **3,14** | **3,11** | **3,06** | **3,02** | **3,00** | |
| 15 | 4,54 | 3,68 | 3,29 | 3,06 | 2,90 | 2,79 | 2,70 | 2,64 | 2,59 | 2,55 | 2,51 | 2,48 | 2,43 | 2,39 | 2,33 | 2,29 | 2,25 | 2,21 | 2,18 | 2,15 | 2,12 | 2,10 | 2,08 | 2,07 | 15 |
| | **8,68** | **6,36** | **5,42** | **4,89** | **4,56** | **4,32** | **4,14** | **4,00** | **3,89** | **3,80** | **3,73** | **3,67** | **3,56** | **3,48** | **3,36** | **3,29** | **3,20** | **3,12** | **3,07** | **3,00** | **2,97** | **2,92** | **2,89** | **2,87** | |
| 16 | 4,49 | 3,63 | 3,24 | 3,01 | 2,85 | 2,74 | 2,66 | 2,59 | 2,54 | 2,49 | **2,45** | 2,42 | 2,37 | 2,33 | 2,28 | 2,24 | 2,20 | 2,16 | 2,13 | 2,09 | 2,07 | 2,04 | 2,02 | 2,01 | 16 |
| | **8,53** | **6,23** | **5,29** | **4,77** | **4,44** | **4,20** | **4,03** | **3,89** | **3,78** | **3,69** | **3,61** | **3,55** | **3,45** | **3,37** | **3,25** | **3,18** | **3,10** | **3,01** | **2,96** | **2,89** | **2,86** | **2,80** | **2,77** | **2,75** | |
| 17 | 4,45 | 3,59 | 3,20 | 2,96 | 2,81 | 2,70 | 2,62 | 2,55 | 2,50 | 2,45 | 2,41 | 2,38 | 2,33 | 2,29 | 2,23 | 2,19 | 2,15 | 2,11 | 2,08 | 2,04 | 2,02 | 1,99 | 1,97 | 1,96 | 17 |
| | **8,40** | **6,11** | **5,18** | **4,67** | **4,34** | **4,10** | **3,93** | **3,79** | **3,68** | **3,59** | **3,52** | **3,45** | **3,35** | **3,27** | **3,16** | **3,08** | **3,00** | **2,92** | **2,86** | **2,79** | **2,76** | **2,70** | **2,67** | **2,65** | |

**Додаток 7.**

## Критичні значення F-критерія

| $k_1$ / $k_2$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 10 | 20 |
|---|---|---|---|---|---|---|---|---|---|
| а) рівень значимості $\alpha = 0,05$ | | | | | | | | | |
| 1 | 161,4 | 199,5 | 215,7 | 224,6 | 230,2 | 234,0 | 238,9 | 242,0 | 248,0 |
| 2 | 18,51 | 19,00 | 19,16 | 19,25 | 19,30 | 19,33 | 19,37 | 19,39 | 19,44 |
| 3 | 10,13 | 9,55 | 9,28 | 9,12 | 9,01 | 8,94 | 8,84 | 8,78 | 8,66 |
| 4 | 7,71 | 6,94 | 6,59 | 6,39 | 6,26 | 6,16 | 6,04 | 5,96 | 5,80 |
| 5 | 6,61 | 5,79 | 5,41 | 5,19 | 5,05 | 4,95 | 4,82 | 4,74 | 4,56 |
| 6 | 5,99 | 5,14 | 4,76 | 4,53 | 4,39 | 4,28 | 4,15 | 4,06 | 3,87 |
| 7 | 5,59 | 4,74 | 4,35 | 4,12 | 3,97 | 3,87 | 3,73 | 3,63 | 3,44 |
| 8 | 5,32 | 4,46 | 4,07 | 3,84 | 3,69 | 3,58 | 3,44 | 3,34 | 3,15 |
| 9 | 5,12 | 4,26 | 3,86 | 3,63 | 3,48 | 3,37 | 3,23 | 3,13 | 2,93 |
| 10 | 4,96 | 4,10 | 3,71 | 3,48 | 3,33 | 3,22 | 3,07 | 2,97 | 2,77 |
| 11 | 4,82 | 3,98 | 3,59 | 3,63 | 3,20 | 3,09 | 2,95 | 2,86 | 2,65 |
| 12 | 4,75 | 3,88 | 3,49 | 3,26 | 3,11 | 3.00 | 2,85 | 2,76 | 2,54 |
| 14 | 4,60 | 3,74 | 3,34 | 3,11 | 2,96 | 2,85 | 2,70 | 2,60 | 2,39 |
| 16 | 4,49 | 3,63 | 3,24 | 3,01 | 2,85 | 2,74 | 2,59 | 2,49 | 2,28 |
| 18 | 4,41 | 3,55 | 3,16 | 2,93 | 2,77 | 2,66 | 2,51 | 2,41 | 2,19 |
| 20 | 4,35 | 3,49 | 3,10 | 2,87 | 2,71 | 2,60 | 2,45 | 2,35 | 2,12 |
| 30 | 4,17 | 3,32 | 2,92 | 2,69 | 2,53 | 2,42 | 2,27 | 2,16 | 1,93 |
| 40 | 4,08 | 3,23 | 2,84 | 2,61 | 2,45 | 2,34 | 2,18 | 2,12 | 1,84 |
| 60 | 4,00 | 3,15 | 2,76 | 2,52 | 2,37 | 2,25 | 2,10 | 2,04 | 1,75 |
| 120 | 3,92 | 3,07 | 2,68 | 2,45 | 2,29 | 2,17 | 2,02 | 1,90 | 1,65 |
| | 3,84 | 2,99 | 2,60 | 2,37 | 2,21 | 2,09 | 1,94 | 1,83 | 1,57 |

**Продовження додатку 7.**

| б) рівень значимості $\alpha = 0,01$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 4052 | 4999 | 5403 | 5625 | 5764 | 5859 | 5981 | 6056 | 6208 |
| 2 | 98,49 | 99,00 | 99,17 | 99,25 | 99,30 | 99,33 | 99,36 | 99,40 | 99,45 |
| 3 | 34,12 | 30,81 | 29,46 | 28,71 | 28,24 | 27,91 | 27,49 | 27,23 | 26,69 |
| 4 | 21,20 | 18,00 | 16,69 | 15,98 | 15,52 | 15,21 | 14,89 | 14,54 | 14,02 |
| 5 | 16,26 | 13,27 | 12,06 | 11,39 | 10,97 | 10,67 | 10,27 | 10,05 | 10,55 |
| 6 | 13,74 | 10,92 | 9,78 | 9,15 | 8,75 | 8,47 | 8,10 | 7,87 | 7,39 |
| 7 | 12,25 | 9,55 | 8,45 | 7,85 | 7,46 | 7,19 | 6,84 | 6,62 | 6,15 |
| 8 | 11,26 | 8,65 | 7,59 | 7,01 | 6,63 | 6,37 | 6,03 | 5,82 | 5,36 |
| 9 | 10,56 | 8,02 | 6,99 | 6,42 | 6,02 | 5,80 | 5,47 | 5,26 | 4,80 |
| 10 | 10,04 | 7,56 | 6,55 | 5,99 | 5,64 | 5,39 | 5,06 | 4,85 | 4,41 |
| 11 | 9,65 | 7,20 | 6,22 | 5,64 | 5,32 | 5,07 | 4,74 | 4,54 | 4,10 |
| 12 | 9,33 | 6,93 | 5,95 | 5,41 | 5,06 | 4,82 | 4,50 | 4,30 | 3,86 |
| 14 | 8,86 | 6,51 | 5,56 | 5,03 | 4,69 | 4,46 | 4,14 | 3,94 | 3,51 |
| 16 | 8,58 | 6,23 | 5,29 | 4,77 | 4,44 | 4,20 | 3,89 | 3,69 | 3,25 |
| 18 | 8,28 | 6,01 | 5,09 | 4,58 | 4,25 | 4,01 | 3,71 | 3,51 | 3,07 |
| 20 | 8,10 | 5,85 | 4,94 | 4,43 | 4,10 | 3,87 | 3,56 | 3,37 | 2,94 |
| 30 | 7,56 | 5,39 | 4,51 | 4,02 | 3,70 | 3,47 | 3,17 | 2,98 | 2,55 |
| 40 | 7,31 | 5,18 | 4,31 | 3,83 | 3,51 | 3,29 | 2,99 | 2,80 | 2,37 |
| 60 | 7,08 | 4,98 | 4,13 | 3,65 | 3,34 | 3,12 | 2,82 | 2,63 | 2,20 |
| 120 | 6,85 | 4,79 | 3,95 | 3,48 | 3,17 | 2,96 | 2,66 | 2,47 | 2,03 |
| | 6,64 | 4,60 | 3,78 | 3,32 | 3,02 | 2,80 | 2,51 | 2,32 | 1,87 |

**Додаток 8.**

## Таблиця критерію *F* для 5%-го рівня значимості (ймовірність «нульової гіпотези» 0,05)

| $K_2$-ступені свободи для внутрішньо | $K_1$ - ступені свободи для міжгрупової дисперсії | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 2 | 18,51 | 19,00 | 19,16 | 19,25 | 19,30 | 19,33 | 19,36 | 13,97 | 19,38 |
| 4 | 7,71 | 6,94 | 6,59 | 6,39 | 6,26 | 6,16 | 6,09 | 6,04 | 6,00 |
| 6 | 5,99 | 5,14 | 4,76 | 4,53 | 4,39 | 4,28 | 4,21 | 4,15 | 4,10 |
| 8 | 5,32 | 4,46 | 4,07 | 3,84 | 3,69 | 3,58 | 3,50 | 3,44 | 3,39 |
| 9 | 5,12 | 4,26 | 3,86 | 3,63 | 3,48 | 3,37 | 3,29 | 3,23 | 3,18 |
| 10 | 4,96 | 4,10 | 3,71 | 3,48 | 3,33 | 3,22 | 3,14 | 3,07 | 3,02 |
| 12 | 4,75 | 3,88 | 3,49 | 3,26 | 3,11 | 3,00 | 2,92 | 2,85 | 2,80 |
| 14 | 4,60 | 3,74 | 3,34 | 3,11 | 2,96 | 2,85 | 2,77 | 2,70 | 2,65 |
| 16 | 4,49 | 3,63 | 3,24 | 3,01 | 2,85 | 2,74 | 2,66 | 2,59 | 2,54 |
| 18 | 4,41 | 3,55 | 3,16 | 2,93 | 2,77 | 2,66 | 2,58 | 2,51 | 2,46 |
| 20 | 4,35 | 3,49 | 3,10 | 2,87 | 2,71 | 2,60 | 2,52 | 2,45 | 2,40 |
| 24 | 4,26 | 3,40 | 3,01 | 2,78 | 2,62 | 2,51 | 2,42 | 2,36 | 2,30 |
| 25 | 4,24 | 3,38 | 2,99 | 2,76 | 2,60 | 2,49 | 2,41 | 2,34 | 2,28 |
| 27 | 4,21 | 3,35 | 2,96 | 2,73 | 2,57 | 2,46 | 2,37 | 2,30 | 2,25 |
| 30 | 4,17 | 3,32 | 2,92 | 2,69 | 2,53 | 2,42 | 2,34 | 2,27 | 2,21 |
| 40 | 4,08 | 3,23 | 2,84 | 2,61 | 2,45 | 2,34 | 2,25 | 2,18 | 2,12 |
| 50 | 4,03 | 3,18 | 2,79 | 2,56 | 2,40 | 2,29 | 2,20 | 2,13 | 2,07 |
| 60 | 4,00 | 3,15 | 2,76 | 2,52 | 2,37 | 2,25 | 2,17 | 2,10 | 2,04 |
| 80 | 3,96 | 3,11 | 2,72 | 2,48 | 2,33 | 2,21 | 2,12 | 2,05 | 1,99 |
| 100 | 3,94 | 3,09 | 2,70 | 2,46 | 2,30 | 2,19 | 2,10 | 2,03 | 1,97 |
| 200 | 3,89 | 3,04 | 2,65 | 2,41 | 2,26 | 2,14 | 2,05 | 1,98 | 1,92 |

**Додаток 9**

## Розподіл $\chi^2$

| Число ступенів волі | Ймовірність більшого значення | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0,995 | 0,990 | 0,975 | 0,950 | 0,900 | 0,750 | 0,500 | 0,250 | 0,100 | 0,050 | 0,025 | 0,010 | 0,005 |
| 1 | | | | | 0,02 | 0,10 | 0,45 | 1,32 | 2,71 | 3,84 | 5,02 | 6,63 | 7,88 |
| 2 | 0,01 | 0,02 | 0,05 | 0,10 | 0,21 | 0,58 | 1,39 | 2,77 | 4,61 | 5,99 | 7,38 | 9,21 | 10,60 |
| 3 | 0,07 | 0,11 | 0,22 | 0,35 | 0,58 | 1,21 | 2,37 | 4,11 | 6,25 | 7,81 | 9,35 | 11,34 | 12,84 |
| 4 | 0,21 | 0,30 | 0,48 | 0,71 | 1,06 | 1,92 | 3,36 | 5,39 | 7,78 | 9,49 | 11,14 | 13,28 | 14,86 |
| 5 | 0,41 | 0,55 | 0,83 | 1,15 | 1,61 | 2,67 | 4,35 | 6,63 | 9,24 | 11,07 | 12,83 | 15,09 | 16,75 |
| 6 | 0,68 | 0,87 | 1,24 | 1,64 | 2,20 | 3,45 | 5,36 | 7,84 | 10,64 | 12,59 | 14,45 | 16,81 | 18,55 |
| 7 | 0,99 | 1,24 | 1,69 | 2,17 | 2,83 | 4,25 | 6,35 | 9,04 | 12,02 | 14,07 | 16,01 | 18,48 | 20,28 |
| 8 | 1,34 | 1,65 | 2,18 | 2,73 | 3,49 | 5,07 | 7,34 | 10,22 | 13,36 | 15,51 | 17,53 | 20,09 | 21,96 |
| 9 | 1,73 | 2,09 | 2,70 | 3,33 | 4,17 | 5,90 | 8,34 | 11,39 | 14,68 | 16,92 | 19,02 | 21,67 | 23,59 |
| 10 | 2,16 | 2,56 | 3,25 | 3,94 | 4,87 | 6,74 | 9,34 | 12,55 | 15,99 | 18,31 | 20,48 | 23,21 | 25,19 |
| 11 | 2,60 | 3,05 | 3,82 | 4,57 | 5,58 | 7,58 | 10,34 | 13,70 | 17,28 | 19,68 | 21,92 | 24,72 | 26,76 |
| 12 | 3,07 | 3,57 | 4,40 | 5,23 | 6,30 | 8,44 | 11,34 | 14,85 | 18,55 | 21,03 | 23,34 | 26,22 | 25,30 |
| 13 | 3,57 | 4,11 | 5,01 | 5,89 | 7,04 | 9,30 | 12,34 | 15,98 | 19,81 | 22,36 | 24,74 | 27,69 | 29,82 |
| 14 | 4,07 | 4,66 | 5,63 | 6,57 | 7,79 | 10,17 | 13,34 | 17,12 | 21,06 | 23,68 | 26,12 | 29,14 | 31,32 |
| 15 | 4,60 | 5,23 | 6,27 | 7,26 | 8,55 | 11,04 | 14,34 | 18,25 | 22,31 | 25,00 | 27,49 | 30,58 | 32,80 |
| 16 | 5,14 | 5,81 | 6,91 | 7,96 | 9,31 | 11,91 | 15,34 | 19,37 | 23,54 | 26,30 | 28,85 | 32,00 | 34,27 |
| 17 | 5,70 | 6,41 | 7,56 | 8,67 | 10,09 | 12,79 | 16,34 | 20,49 | 24,77 | 27,59 | 30,19 | 33,41 | 35,72 |
| 18 | 6,26 | 7,01 | 8,23 | 9,39 | 10,86 | 13,68 | 17,34 | 21,60 | 25,99 | 28,87 | 31,53 | 34,81 | 37,16 |
| 19 | 6,84 | 7,63 | 8,91 | 10,12 | 11,65 | 14,56 | 18,34 | 22,72 | 27,20 | 30,14 | 32,85 | 36,19 | 38,58 |
| 20 | 7,43 | 8,26 | 9,59 | 10,85 | 12,44 | 15,45 | 19,34 | 22,83 | 28,41 | 31,41 | 34,17 | 37,57 | 40,00 |

**Додаток 10.**

## Значення $\chi^2$ - критерію Пірсона при рівні значимості 0,10, 0,05, 0,01

| *d.f.* | 0,10 | 0,05 | 0,01 | *d.f.* | 0,10 | 0,05 | 0,01 |
|---|---|---|---|---|---|---|---|
| 1 | 2,71 | 3,84 | 6,63 | 21 | 29,62 | 32,67 | 38,93 |
| 2 | 4,61 | 5,99 | 9,21 | 22 | 30,81 | 33,92 | 40,29 |
| 3 | 6,25 | 7,81 | 11,34 | 23 | 31,01 | 35,17 | 41,64 |
| 4 | 7,78 | 9,49 | 13,28 | 24 | 33,20 | 36,42 | 42,98 |
| 5 | 9,24 | 11,07 | 15,09 | 25 | 34,38 | 37,65 | 44,31 |
| 6 | 10,64 | 12,59 | 16,81 | 26 | 35,56 | 38,89 | 45,64 |
| 7 | 12,02 | 14,07 | 18,48 | 27 | 36,74 | 40,11 | 46,96 |
| 8 | 13,36 | 15,51 | 20,09 | 28 | 37,92 | 41,34 | 48,28 |
| 9 | 14,68 | 16,92 | 21,67 | 29 | 39,09 | 42,56 | 49,59 |
| 10 | 50,99 | 18,31 | 23,21 | 30 | 40,26 | 43,77 | 50,89 |
| 11 | 17,28 | 19,68 | 24,72 | 40 | 51,80 | 55,76 | 63,69 |
| 12 | 18,55 | 21,03 | 26,22 | 50 | 63,17 | 67,50 | 76,15 |
| 13 | 19,81 | 22,36 | 27,69 | 60 | 74,40 | 79,08 | 88,38 |
| 14 | 21,06 | 23,68 | 29,14 | 70 | 85,53 | 90,53 | 100,42 |
| 15 | 22,31 | 25,00 | 30,58 | 80 | 96,58 | 101,88 | 112,33 |
| 16 | 23,54 | 26,30 | 32,00 | 90 | 107,56 | 113,14 | 124,12 |
| 17 | 24,77 | 27,59 | 33,41 | 100 | 118,50 | 124,34 | 135,81 |
| 18 | 25,99 | 28,87 | 34,81 | | | | |
| 19 | 27,20 | 30,14 | 36,19 | | | | |
| 20 | 28,41 | 31,14 | 37,57 | | | | |

**Додаток 11.**

## Квантилі $\chi^2$ – розподілу

| $k$ \ $p$ | 0,025 | 0,050 | 0,10 | 0,90 | 0,95 | 0,975 |
|---|---|---|---|---|---|---|
| 1 | 0,01 | 0,04 | 0,02 | 2,71 | 3,84 | 5,02 |
| 2 | 0,05 | 0,10 | 0,21 | 4,61 | 5,99 | 7,38 |
| 3 | 0,22 | 0,35 | 0,58 | 6,25 | 7,82 | 9,35 |
| 4 | 0,48 | 0,71 | 1,06 | 7,78 | 9,49 | 11,14 |
| 5 | 0,83 | 1,15 | 1,61 | 9,24 | 11,07 | 12,03 |
| 6 | 1,24 | 1,64 | 2,20 | 10,65 | 12,59 | 14,45 |
| 7 | 1,69 | 2,17 | 2,83 | 12,02 | 14,07 | 16,01 |
| 8 | 2,18 | 2,73 | 3,49 | 13,36 | 15,51 | 17,54 |
| 9 | 2,70 | 3,33 | 4,17 | 14,68 | 16,92 | 19,02 |
| 10 | 3,25 | 3,94 | 4,87 | 15,99 | 18,31 | 20,48 |
| 11 | 3,82 | 4,58 | 5,58 | 17,28 | 19,68 | 21,92 |
| 12 | 4,40 | 5,23 | 6,30 | 18,55 | 21,03 | 23,34 |
| 13 | 5,01 | 5,89 | 7,04 | 19,81 | 22,36 | 24,74 |
| 14 | 5,63 | 6,57 | 7,79 | 21,06 | 23,69 | 26,12 |
| 15 | 6,26 | 7,26 | 8,55 | 22,31 | 25,00 | 27,49 |
| 16 | 6,91 | 7,96 | 9,31 | 23,54 | 26,30 | 28,85 |
| 17 | 7,56 | 8,67 | 10,09 | 24,77 | 27,59 | 30,19 |
| 18 | 8,23 | 9,39 | 10,87 | 25,09 | 28,87 | 31,53 |
| 19 | 8,91 | 10,12 | 11,65 | 27,20 | 30,14 | 32,85 |
| 20 | 9,59 | 10,85 | 12,44 | 28,41 | 31,41 | 34,17 |
| 22 | 10,98 | 12,34 | 14,04 | 30,81 | 33,92 | 36,78 |
| 24 | 12,40 | 13,85 | 15,66 | 33,20 | 36,42 | 39,36 |
| 26 | 13,84 | 15,38 | 17,29 | 35,56 | 38,89 | 41,92 |
| 28 | 15,31 | 16,93 | 18,94 | 37,92 | 41,34 | 44,46 |
| 30 | 16,79 | 18,49 | 20,60 | 40,26 | 43,77 | 46,90 |
| 35 | 20,57 | 22,47 | 24,80 | 46,06 | 49,00 | 53,20 |
| 40 | 24,43 | 26,51 | 29,05 | 51,81 | 55,76 | 59,34 |
| 45 | 28,37 | 30,61 | 33,35 | 57,51 | 61,66 | 65,41 |
| 50 | 32,36 | 34,76 | 37,69 | 63,17 | 67,51 | 71,42 |

**<u>Додаток 12.</u>**

**Значення верхнього $\alpha$% межі $\chi^2\alpha$ в залежності від ймовірності $P$ ($\chi^2 > \chi^2\alpha$) та числа ступенів свободи $\chi^2$ – розподілу**

| Число ступенів свободи $k$ | Ймовірність $P$ ($\chi^2 > \chi^2\alpha$) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 0,99 | 0,98 | 0,95 | 0,90 | 0,80 | 0,70 | 0,50 | 0,30 |
| 1 | 1,64 | 2,7 | 3,8 | 5,4 | 6,6 | 7,9 | 9,5 | 10,83 |
| 2 | 3,22 | 4,6 | 6,0 | 7,8 | 9,2 | 11,6 | 12,3 | 13,8 |
| 3 | 4,64 | 6,3 | 7,8 | 9,8 | 11,3 | 12,8 | 14,8 | 16,3 |
| 4 | 6,0 | 7,8 | 9,5 | 11,7 | 13.3 | 13,9 | 16,9 | 18,5 |
| 5 | 7,3 | 9,2 | 11,1 | 13,4 | 15,1 | 16,3 | 18,9 | 20,5 |
| 6 | 8,6 | 10,6 | 12,6 | 15,0 | 16,8 | 18,6 | 20,7 | 22,5 |
| 7 | 9,8 | 12,0 | 14,1 | 16,6 | 18,5 | 20,3 | 22,6 | 24,3 |
| 8 | 11,0 | 13,4 | 15,5 | 18,2 | 20,1 | 21,9 | 24,3 | 26,1 |
| 9 | 12,2 | 14,7 | 16,9 | 19,7 | 21,7 | 23,6 | 26,1 | 27,9 |
| 10 | 13,4 | 16,0 | 18,3 | 21,2 | 23,3 | 25,2 | 27,7 | 29,6 |
| 11 | 14,6 | 17,3 | 19,7 | 22,6 | 24,7 | 26,8 | 29,4 | 31,3 |
| 12 | 15,8 | 18,5 | 21,0 | 24,1 | 26,2 | 28,3 | 31,0 | 32,9 |
| 13 | 17,0 | 19,8 | 22,4 | 25,5 | 27,7 | 29,8 | 32,5 | 34,5 |
| 14 | 18,2 | 21,1 | 23,7 | 26,9 | 29,1 | 31,0 | 34,0 | 36,1 |
| 15 | 19,3 | 22,3 | 25,0 | 28,3 | 30,6 | 32,5 | 35,5 | 37,7 |
| 16 | 20,5 | 23,5 | 26,3 | 29,6 | 32,0 | 34,0 | 37,0 | 39,2 |
| 17 | 21,6 | 24,8 | 27,6 | 31,0 | 33,4 | 35,5 | 38,5 | 40,8 |
| 18 | 22,8 | 26,0 | 28,9 | 32,3 | 34,8 | 37,0 | 40,0 | 42,3 |
| 19 | 23,9 | 27,2 | 30,1 | 33,7 | 36,2 | 38,5 | 41,5 | 43,8 |
| 20 | 25,0 | 28,4 | 31,4 | 35,0 | 37,6 | 40,0 | 43,0 | 45,3 |
| 21 | 26,2 | 29,6 | 32,7 | 36,3 | 38,9 | 41,5 | 44,5 | 46,8 |
| 22 | 27,3 | 30,8 | 33,9 | 37,7 | 40,3 | 42,5 | 46,0 | 48,3 |
| 23 | 28,4 | 32,0 | 35,2 | 39,0 | 41,6 | 44,0 | 47,5 | 49,7 |
| 24 | 29,6 | 33,2 | 36,4 | 40,3 | 43,0 | 45,5 | 48,5 | 51,2 |
| 25 | 30,7 | 34,4 | 37,7 | 41,6 | 44,3 | 47,0 | 50,0 | 52,6 |
| 26 | 31,8 | 35,6 | 38,9 | 42,9 | 45,6 | 48,0 | 51,5 | 54,1 |
| 27 | 32,9 | 36,7 | 40,1 | 44,1 | 47,0 | 49,5 | 53,0 | 55,5 |
| 28 | 34.0 | 37,9 | 41,3 | 45,4 | 48,3 | 51,0 | 54,5 | 56,9 |
| 29 | 35,1 | 39,1 | 42,6 | 46,7 | 49,9 | 52.5 | 56,0 | 58,3 |
| 30 | 36,3 | 40,3 | 43,8 | 48,0 | 50,9 | 54,0 | 57,5 | 59,7 |

**Додаток 13.**

## Критичні значення кореляційного відношення $\eta^2$ і коефіцієнта детермінації $R^2$

| а) рівень значимості $\alpha = 0{,}05$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $k_2$ \ $k_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 10 | 20 |
| 3 | 0,771 | 865 | 903 | 924 | 938 | 947 | 959 | 967 | 983 |
| 4 | 658 | 776 | 832 | 865 | 887 | 902 | 924 | 937 | 967 |
| 5 | 569 | 699 | 764 | 806 | 835 | 854 | 885 | 904 | 948 |
| 6 | 500 | 632 | 704 | 751 | 785 | 811 | 847 | 871 | 928 |
| 7 | 444 | 575 | 651 | 702 | 739 | 768 | 810 | 839 | 908 |
| 8 | 399 | 527 | 604 | 657 | 697 | 729 | 775 | 807 | 887 |
| 9 | 362 | 488 | 563 | 618 | 659 | 692 | 742 | 777 | 867 |
| 10 | 332 | 451 | 527 | 582 | 624 | 659 | 711 | 749 | 847 |
| 11 | 306 | 420 | 495 | 550 | 593 | 628 | 682 | 722 | 828 |
| 12 | 283 | 394 | 466 | 521 | 564 | 600 | 655 | 696 | 809 |
| 14 | 247 | 348 | 417 | 471 | 514 | 550 | 607 | 650 | 773 |
| 16 | 219 | 312 | 378 | 429 | 477 | 507 | 564 | 609 | 740 |
| 18 | 197 | 283 | 345 | 394 | 435 | 470 | 527 | 573 | 709 |
| 20 | 179 | 259 | 318 | 364 | 404 | 432 | 495 | 540 | 680 |
| 22 | 164 | 238 | 294 | 339 | 377 | 410 | 466 | 511 | 653 |
| 24 | 151 | 221 | 273 | 316 | 353 | 385 | 440 | 484 | 628 |
| 26 | 140 | 206 | 256 | 297 | 332 | 363 | 417 | 461 | 605 |
| 28 | 130 | 193 | 240 | 279 | 314 | 344 | 396 | 439 | 583 |
| 30 | 122 | 182 | 227 | 264 | 297 | 326 | 373 | 419 | 563 |
| 32 | 115 | 171 | 214 | 250 | 282 | 310 | 360 | 401 | 544 |
| 34 | 108 | 162 | 203 | 238 | 268 | 296 | 344 | 384 | 526 |
| 36 | 102 | 153 | 102 | 226 | 256 | 282 | 329 | 368 | 509 |
| 38 | 097 | 146 | 184 | 218 | 245 | 271 | 316 | 355 | 493 |
| 40 | 093 | 139 | 176 | 207 | 234 | 259 | 304 | 342 | 479 |
| 50 | 075 | 113 | 143 | 170 | 194 | 216 | 254 | 288 | 416 |
| 60 | 063 | 095 | 121 | 144 | 165 | 184 | 218 | 249 | 368 |
| 80 | 047 | 072 | 093 | 110 | 127 | 142 | 170 | 196 | 298 |
| 100 | 038 | 058 | 075 | 090 | 103 | 116 | 140 | 161 | 251 |
| 120 | 032 | 049 | 063 | 075 | 087 | 098 | 119 | 137 | 217 |
| 200 | 019 | 030 | 038 | 046 | 053 | 060 | 073 | 086 | 139 |
| 400 | 010 | 015 | 019 | 023 | 027 | 031 | 038 | 044 | 074 |

**Продовження додатку 13.**

б) рівень значимості $\alpha = 0{,}01$

| $k_2$ \ $k_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 10 | 20 |
|---|---|---|---|---|---|---|---|---|---|
| 3 | 0,919 | 954 | 967 | 975 | 979 | 982 | 987 | 989 | 994 |
| 4 | 841 | 900 | 926 | 941 | 951 | 958 | 967 | 973 | 986 |
| 5 | 765 | 842 | 879 | 901 | 916 | 928 | 943 | 953 | 974 |
| 6 | 696 | 785 | 830 | 859 | 879 | 894 | 915 | 929 | 961 |
| 7 | 636 | 732 | 784 | 818 | 842 | 860 | 887 | 904 | 946 |
| 8 | 585 | 684 | 740 | 778 | 806 | 827 | 858 | 879 | 931 |
| 9 | 540 | 641 | 700 | 741 | 771 | 795 | 829 | 854 | 914 |
| 10 | 501 | 602 | 663 | 706 | 738 | 764 | 802 | 829 | 898 |
| 11 | 467 | 567 | 629 | 673 | 707 | 734 | 775 | 805 | 882 |
| 12 | 437 | 536 | 598 | 643 | 678 | 707 | 750 | 782 | 865 |
| 14 | 388 | 482 | 544 | 590 | 626 | 656 | 703 | 738 | 834 |
| 16 | 342 | 438 | 498 | 544 | 581 | 612 | 660 | 698 | 803 |
| 18 | 315 | 401 | 459 | 504 | 541 | 572 | 622 | 669 | 774 |
| 20 | 288 | 369 | 426 | 470 | 506 | 537 | 588 | 627 | 746 |
| 22 | 265 | 342 | 396 | 440 | 475 | 506 | 557 | 597 | 720 |
| 24 | 246 | 319 | 371 | 413 | 448 | 478 | 529 | 569 | 695 |
| 26 | 229 | 298 | 349 | 389 | 423 | 453 | 503 | 543 | 672 |
| 28 | 214 | 280 | 329 | 368 | 401 | 430 | 480 | 520 | 650 |
| 30 | 201 | 264 | 311 | 349 | 381 | 410 | 498 | 498 | 630 |
| 32 | 190 | 250 | 295 | 332 | 364 | 391 | 438 | 479 | 611 |
| 34 | 179 | 237 | 280 | 316 | 347 | 374 | 420 | 459 | 592 |
| 36 | 170 | 226 | 267 | 302 | 332 | 358 | 403 | 443 | 574 |
| 38 | 162 | 215 | 255 | 289 | 318 | 344 | 389 | 426 | 558 |
| 40 | 155 | 206 | 244 | 277 | 305 | 310 | 374 | 412 | 542 |
| 50 | 125 | 168 | 201 | 229 | 254 | 276 | 315 | 351 | 475 |
| 60 | 106 | 142 | 171 | 196 | 218 | 238 | 273 | 305 | 423 |
| 80 | 080 | 109 | 132 | 151 | 169 | 186 | 215 | 242 | 345 |
| 100 | 065 | 088 | 107 | 123 | 138 | 152 | 177 | 201 | 292 |
| 120 | 054 | 074 | 090 | 104 | 117 | 129 | 151 | 171 | 253 |
| 200 | 033 | 045 | 055 | 064 | 072 | 080 | 094 | 106 | 165 |
| 400 | 016 | 023 | 028 | 033 | 037 | 041 | 049 | 056 | 086 |

**Додаток 14.**

**Критичні значення коефіцієнтів кореляції для рівнів значимості 0,05, 0,01**

| *d.f.* | $\alpha = 0,05$ | $\alpha = 0,01$ | *d.f.* | $\alpha = 0,05$ | $\alpha = 0,01$ |
|---|---|---|---|---|---|
| 1 | 0,996917 | 0,9998766 | 17 | 0,4555 | 0,5751 |
| 2 | 0,995000 | 0,990000 | 18 | 0,4438 | 0,5614 |
| 3 | 0,8783 | 0,95873 | 19 | 0,4329 | 0,5487 |
| 4 | 0,8114 | 0,91720 | 20 | 0,4227 | 0,5368 |
| 5 | 0,7545 | 0,8745 | 25 | 0,3809 | 0,4869 |
| 6 | 0,07067 | 0,8343 | 30 | 0,3494 | 0,4487 |
| 7 | 0,6664 | 0,7977 | 35 | 0,3246 | 0,4182 |
| 8 | 0,6319 | 0,7646 | 40 | 0,3044 | 0,3932 |
| 9 | 0,6021 | 0,7348 | 45 | 0,2875 | 0,3721 |
| 10 | 0,5760 | 0,7079 | 50 | 0,2732 | 0,3541 |
| 11 | 0,5529 | 0,6835 | 60 | 0,2500 | 0,3248 |
| 12 | 0,5324 | 0,6614 | 70 | 0,2919 | 0,3017 |
| 13 | 0,5139 | 0,6411 | 80 | 0,2172 | 0,2830 |
| 14 | 0,4973 | 0,6226 | 90 | 0,2050 | 0,2673 |
| 15 | 0,4821 | 0,6055 | 100 | 0,1946 | 0,2540 |
| 16 | 0,4683 | 0,5897 | | | |

**Додаток 15.**

***Z*-перетворення. Значення величини *z* для значень *r***

| *r* | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0,0 | 0,0000 | 0,0100 | 0,0200 | 0,0300 | 0,0400 | 0,0501 | 0,0601 | 0,0701 | 0,0802 | 0,0902 |
| 0,1 | 0,1003 | 0,1105 | 0,1206 | 0,1308 | 0,1409 | 0,1511 | 0,1614 | 0,1717 | 0,1820 | 0,1923 |
| 0,2 | 0,2027 | 0,2132 | 0,2237 | 0,2342 | 0,2448 | 0,2554 | 0,2661 | 0,2769 | 0,2877 | 0,2986 |
| 0,3 | 0,3095 | 0,3206 | 0,3317 | 0,3428 | 0,3541 | 0,3654 | 0,3769 | 0,3884 | 0,4001 | 0,4118 |
| 0,4 | 0,4236 | 0,4356 | 0,4477 | 0,4599 | 0,4722 | 0,4847 | 0,4973 | 0,5101 | 0,5230 | 0,5361 |
| 0,5 | 0,5493 | 0,5627 | 0,5763 | 0,5901 | 0,6042 | 0,6184 | 0,6328 | 0,6475 | 0,6625 | 0,6777 |
| 0,6 | 0,6931 | 0,7089 | 0,7250 | 0,7414 | 0,7582 | 0,7753 | 0,7928 | 0,8107 | 0,8291 | 0,8480 |
| 0,7 | 0,8673 | 0,8872 | 0,9076 | 0,9287 | 0,9505 | 0,9730 | 0,9962 | 1,0203 | 1,0454 | 1,0714 |
| 0,8 | 1,0986 | 1,1270 | 1,1568 | 1,1881 | 1,2212 | 1,2562 | 1,2933 | 1,3331 | 1,3758 | 1,4219 |
| 0,9 | 1,4722 | 1,5275 | 1,5890 | 1,6584 | 1,7380 | 1,8318 | 1,9459 | 2,0923 | 2,2976 | 2,6467 |

**Додаток 16.**

**Значення $\alpha$ - процентних меж $t_{\alpha k}$ залежно від $k$ ступенів свободи і заданого рівня значення $\alpha$ для розподілу Стьюдента**

| $k$ \ $\alpha$ | 10,0 | 5,0 | 2,5 | 2,0 | 1,0 | 0,5 | 0,3 | 0,2 | 0,1 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 6,314 | 12,706 | 25,452 | 31,821 | 63,657 | 127,3 | 212,2 | 318,3 | 636,6 |
| 2 | 2,920 | 4,303 | 6,205 | 6,965 | 9,925 | 14,089 | 18,216 | 22,327 | 31,600 |
| 3 | 2,353 | 3,182 | 4,177 | 4,541 | 5,841 | 7,453 | 8,891 | 10,214 | 12,922 |
| 4 | 2,132 | 2,776 | 3,495 | 3,747 | 4,604 | 5,597 | 6,435 | 7,173 | 8,610 |
| 5 | 2,015 | 2,571 | 3,163 | 3,365 | 4,032 | 4,773 | 5,376 | 5,893 | 6,869 |
| 6 | 1,943 | 2,447 | 2,969 | 3,143 | 3,707 | 4,317 | 4,800 | 5,208 | 5,959 |
| 7 | 1,895 | 2,447 | 2,841 | 2,998 | 3,499 | 4,029 | 4,442 | 4,785 | 5,408 |
| 8 | 1,860 | 2,306 | 2,752 | 2,696 | 3,355 | 3,833 | 4,199 | 4,501 | 5,041 |
| 9 | 1,833 | 2,262 | 2,685 | 2,821 | 3,250 | 3,690 | 4,024 | 4,297 | 4,781 |
| 10 | 1,812 | 2,228 | 2,634 | 2,764 | 3,169 | 3,581 | 3,892 | 4,144 | 4,587 |
| 12 | 1,782 | 2,179 | 2,560 | 2,681 | 3,055 | 3,428 | 3,706 | 3,930 | 4,318 |
| 14 | 1,761 | 2,145 | 2,510 | 2,624 | 2,977 | 3,326 | 3,583 | 3,787 | 4,140 |
| 16 | 1,746 | 2,120 | 2,473 | 2,583 | 2,921 | 3,252 | 3,494 | 3,686 | 4,015 |
| 18 | 1,734 | 2,101 | 2,445 | 2,552 | 2,878 | 3,193 | 3,428 | 3,610 | 3,922 |
| 20 | 1,725 | 2,086 | 2,423 | 2,528 | 2,845 | 3,153 | 3,376 | 3,552 | 3,849 |
| 22 | 1,717 | 2,074 | 2,405 | 2,508 | 2,819 | 3,119 | 3,335 | 3,505 | 3,792 |
| 24 | 1,71 | 2,064 | 2,391 | 2,492 | 2,797 | 3,092 | 3,302 | 3,467 | 3,745 |
| 26 | 1,706 | 2,056 | 2,379 | 2,479 | 2,779 | 3,067 | 3,274 | 3,435 | 3,704 |
| 28 | 1,701 | 2,048 | 2,369 | 2,467 | 2,763 | 3,047 | 3,250 | 3,408 | 3,674 |
| 30 | 1,697 | 2,042 | 2,360 | 2,457 | 2,750 | 3,030 | 3,230 | 3,386 | 3,646 |
| ∞ | 1,645 | 1,960 | 2,241 | 2,326 | 2,576 | 2,807 | 2,968 | 3,090 | 3,291 |

**Витяг з таблиць розподілу Стьюдента для малих вибірок для деяких значень “$k$” і “$t$”**

| $t$ \ $k$ | 4 | 5 | 9 | 10 | 15 | 20 | 25 | Нормальний розподіл |
|---|---|---|---|---|---|---|---|---|
| 1,0 | 0,813 | 0,818 | 0,828 | 0,830 | 0,833 | 0,835 | 0,838 | 0,841 |
| 2,0 | 0,942 | 0,949 | 0,962 | 0,963 | 0,968 | 0,970 | 0,973 | 0,977 |
| 3,0 | 0,980 | 0,985 | 0,992 | 0,993 | 0,995 | 0,996 | 0,997 | 0,999 |

**Додаток 17.**

**Значення коефіцієнта кореляції рангів Спірмена для двохсторонніх меж рівня значимості $\alpha$**

| $n$ \ $\alpha$ | 0,20 | 0,10 | 0,05 | 0,02 | 0,01 | 0,002 |
|---|---|---|---|---|---|---|
| 4 | 0,8000 | 0,8000 | | | | |
| 5 | 0,7000 | 0,8000 | 0,9000 | 0,9000 | | |
| 6 | 0,6000 | 0,7714 | 0,8286 | 0,8857 | 0,9429 | |
| 7 | 0,5357 | 0,6786 | 0,7450 | 0,8571 | 0,8929 | 0,9643 |
| 8 | 0,5000 | 0,6190 | 0,7143 | 0,8095 | 0,8571 | 0,9286 |
| 9 | 0,4667 | 0,5833 | 0,6833 | 0,7667 | 0,8167 | 0,9000 |
| 10 | 0,4424 | 0,5515 | 0,6364 | 0,7333 | 0,7818 | 0,8667 |
| 11 | 0,4182 | 0,5273 | 0,6091 | 0,7000 | 0,7455 | 0,8364 |
| 12 | 0,3986 | 0,4965 | 0,5804 | 0,6713 | 0,7273 | 0,8182 |
| 13 | 0,3791 | 0,4780 | 0,5549 | 0,6429 | 0,6978 | 0,7912 |
| 14 | 0,3626 | 0,4593 | 0,5341 | 0,6220 | 0,6747 | 0,7670 |
| 15 | 0,3500 | 0,4429 | 0,5179 | 0,6000 | 0,6536 | 0,7464 |
| 16 | 0,3382 | 0,4265 | 0,5000 | 0,5824 | 0,6324 | 0,7265 |
| 17 | 0,3260 | 0,4118 | 0,4853 | 0,5637 | 0,6152 | 0,7083 |
| 18 | 0,3148 | 0,3994 | 0,4716 | 0,5480 | 0,5975 | 0,6904 |
| 19 | 0,3070 | 0,3895 | 0,4579 | 0,5333 | 0,5825 | 0,6737 |
| 20 | 0,2977 | 0,3789 | 0,4451 | 0,5203 | 0,5684 | 0,6586 |
| 21 | 0,2909 | 0,3688 | 0,4351 | 0,5078 | 0,5545 | 0,6455 |
| 22 | 0,2829 | 0,3597 | 0,4241 | 0,4963 | 0,5426 | 0,6318 |
| 23 | 0,2767 | 0,3518 | 0,4150 | 0,4852 | 0,5306 | 0,6186 |
| 24 | 0,2704 | 0,3435 | 0,4061 | 0,4748 | 0,5200 | 0,6070 |
| 25 | 0,2646 | 0,3362 | 0,3977 | 0,4654 | 0,5100 | 0,5962 |
| 26 | 0,2588 | 0,3299 | 0,3894 | 0,4564 | 0,5002 | 0,5856 |
| 27 | 0,2540 | 0,3236 | 0,3822 | 0,4481 | 0,4915 | 0,5757 |
| 28 | 0,2490 | 0,3175 | 0,3749 | 0,4401 | 0,4828 | 0,5660 |
| 29 | 0,2443 | 0,3113 | 0,3685 | 0,4320 | 0,4744 | 0,5567 |
| 30 | 0,2400 | 0,3059 | 0,3620 | 0,4251 | 0,4665 | 0,5479 |

**Додаток 18.**

**Значення z для бісеріального коефіцієнта кореляції при різних *p*. (Для $p \leq 0{,}50$ знаходять $z$ для $q = 1 - p$)**

| P | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0,5 | 0,3889 | 3988 | 3984 | 3978 | 3969 | 3958 | 3944 | 3928 | 3909 | 3888 |
| 0,6 | 3863 | 3837 | 3808 | 3776 | 3741 | 3704 | 3664 | 3622 | 3576 | 3528 |
| 0,7 | 3477 | 3429 | 3366 | 3307 | 3244 | 3178 | 3109 | 3037 | 2961 | 2882 |
| 0,8 | 2800 | 2714 | 2624 | 2531 | 2433 | 2332 | 2226 | 2116 | 2000 | 1880 |
| 0,9 | 1755 | 1624 | 1487 | 1343 | 1191 | 1031 | 0862 | 0680 | 0484 | 0267 |

**Додаток 19.**

**Таблиця значень функції Лапласа** $\Phi(x) = -\frac{1}{\sqrt{2\pi}} \int\limits_0^X e^{-\frac{z^2}{2}} dz$

| *x* | *Φ(x)* | *x* | *Φ(x)* | *x* | *Φ(x)* | *x* | *Φ(x)* |
|---|---|---|---|---|---|---|---|
| 0,00 | 0,0000 | 0,39 | 0,1617 | 0,78 | 0,2823 | 1,17 | 0,3790 |
| 0,01 | 0,0040 | 0,40 | 0,1564 | 0,79 | 0,2852 | 1,18 | 0,3810 |
| 0,02 | 0,0080 | 0,41 | 0,1691 | 0,80 | 0,2881 | 1,19 | 0,3830 |
| 0,03 | 0,0120 | 0,42 | 0,1628 | 0,81 | 0,2910 | 1,20 | 0,3849 |
| 0,04 | 0,0160 | 0,43 | 0,1664 | 0,82 | 0,2939 | 1,21 | 0,3869 |
| 0,05 | 0,0199 | 0,44 | 0,1700 | 0,83 | 0,2967 | 1,22 | 0,3883 |
| 0,06 | 0,0239 | 0,45 | 0,1736 | 0,84 | 0,2995 | 1,23 | 0,3907 |
| 0,07 | 0,0279 | 0,46 | 0,1772 | 0,85 | 0,3023 | 1,24 | 0,3925 |
| 0,08 | 0,0319 | 0,47 | 0,1808 | 0,86 | 0,3051 | 1,25 | 0,3944 |
| 0,09 | 0,0359 | 0,48 | 0,1844 | 0,87 | 0,3078 | 1,26 | 0,3962 |
| 0,10 | 0,0398 | 0,49 | 0,1879 | 0,88 | 0,3106 | 1,27 | 0,3980 |
| 0,11 | 0,0438 | 0,50 | 0,1915 | 0,89 | 0,3133 | 1,28 | 0,3997 |
| 0,12 | 0,0478 | 0,51 | 0,1950 | 0,90 | 0,3159 | 1,29 | 0,4015 |
| 0,13 | 0,0517 | 0,52 | 0,1985 | 0,91 | 0,3186 | 1,30 | 0,4032 |
| 0,14 | 0,8557 | 0,53 | 0,2019 | 0,92 | 0,3212 | 1,31 | 0,4049 |
| 0,15 | 0,0596 | 0,54 | 0,2054 | 0,93 | 0,3238 | 1,32 | 0,4066 |
| 0,16 | 0,0636 | 0,55 | 0,2088 | 0,94 | 0,3264 | 1,33 | 0,4082 |
| 0,17 | 0,0675 | 0,56 | 0,2123 | 0,95 | 0,3289 | 1,34 | 0,4099 |
| 0,18 | 0,0714 | 0,57 | 0,2157 | 0,96 | 0,3315 | 1,35 | 0,4115 |
| 0,19 | 0,0753 | 0,58 | 0,2190 | 0,97 | 0,3340 | 1,36 | 0,4131 |
| 0,20 | 0,0793 | 0,59 | 0,2224 | 0,98 | 0,3365 | 1,37 | 0,4147 |
| 0,21 | 0,0832 | 0,60 | 0,2257 | 0,99 | 0,3389 | 1,38 | 0,4162 |
| 0,22 | 0,0871 | 0,61 | 0,2291 | 1,00 | 0,3413 | 1,39 | 0,4177 |
| 0,23 | 0,0910 | 0,62 | 0,2324 | 1,01 | 0,3438 | 1,40 | 0,4192 |
| 0,24 | 0,0948 | 0,63 | 0,2357 | 1,02 | 0,3461 | 1,41 | 0,4207 |
| 0,25 | 0,0987 | 0,64 | 0,2389 | 1,03 | 0,3485 | 1,42 | 0,4222 |
| 0,26 | 0,1026 | 0,65 | 0,2422 | 1,04 | 0,3508 | 1,43 | 0,4236 |
| 0,27 | 0,1064 | 0,66 | 0,2454 | 1,05 | 0,3531 | 1,44 | 0,4251 |
| 0,28 | 0,1103 | 0,67 | 0,2486 | 1,06 | 0,3554 | 1,45 | 0,4265 |
| 0,29 | 0,1141 | 0,68 | 0,2517 | 1,07 | 0,3577 | 1,46 | 0,4279 |
| 0,30 | 0,1179 | 0,69 | 0,2549 | 1,08 | 0,3599 | 1,47 | 0,4292 |
| 0,31 | 0,1217 | 0,70 | 0,2580 | 1,09 | 0,3621 | 1,48 | 0,4306 |
| 0,32 | 0,1255 | 0,71 | 0,2611 | 1,10 | 0,3643 | 1,49 | 0,4319 |
| 0,33 | 0,1293 | 0,72 | 0,2642 | 1,11 | 0,3665 | 1,50 | 0,4332 |
| 0,34 | 0,1331 | 0,73 | 0,2673 | 1,12 | 0,3686 | 1,51 | 0,4345 |
| 0,35 | 0,1368 | 0,74 | 0,2703 | 1,13 | 0,3708 | 1,52 | 0,4357 |
| 0,36 | 0,1406 | 0,75 | 0,2734 | 1,14 | 0,3729 | 1,53 | 0,4370 |
| 0,37 | 0,1443 | 0,76 | 0,2764 | 1,15 | 0,3749 | 1,54 | 0,4382 |
| 0,38 | 0,1480 | 0,77 | 0,2794 | 1,16 | 0,3770 | 1,55 | 0,4394 |

**Продовження додатку 19.**

| $x$ | $Ф(x)$ | $x$ | $Ф(x)$ | $x$ | $Ф(x)$ | $x$ | $Ф(x)$ |
|---|---|---|---|---|---|---|---|
| 1,56 | 0,4406 | 1,82 | 0,4656 | 2,16 | 0,4846 | 2,68 | 0,4963 |
| 1,57 | 0,4418 | 1,83 | 0,4664 | 2,18 | 0,4854 | 2,70 | 0,4965 |
| 1,58 | 0,4429 | 1,84 | 0,4671 | 2,20 | 0,4861 | 2,72 | 0,4967 |
| 1,59 | 0,4441 | 1,85 | 0,4678 | 2,22 | 0,4868 | 2,74 | 0,4969 |
| 1,60 | 0,4452 | 1,86 | 0,4686 | 2,24 | 0,4875 | 2,76 | 0,4971 |
| 1,61 | 0,4463 | 1,87 | 0,4693 | 2,26 | 0,4881 | 2,78 | 0,4973 |
| 1,62 | 0,4474 | 1,88 | 0,4699 | 2,28 | 0,4887 | 2,80 | 0,4974 |
| 1,63 | 0,4484 | 1,89 | 0,4706 | 2,30 | 0,4893 | 2,82 | 0,4976 |
| 1,64 | 0,4495 | 1,90 | 0,4713 | 2,32 | 0,4898 | 2,84 | 0,4977 |
| 1,65 | 0,4505 | 1,91 | 0,4719 | 2,34 | 0,4904 | 2,86 | 0,4979 |
| 1,66 | 0,4515 | 1,92 | 0,4726 | 2,36 | 0,4909 | 2,90 | 0,4981 |
| 1,67 | 0,4525 | 1,93 | 0,4732 | 2,38 | 0,4913 | 2,92 | 0,4982 |
| 1,68 | 0,4535 | 1,94 | 0,4738 | 2,40 | 0,4918 | 2,94 | 0,4984 |
| 1,69 | 0,4545 | 1,95 | 0,4744 | 2,42 | 0,4922 | 2,96 | 0,49846 |
| 1,70 | 0,4554 | 1,96 | 0,4750 | 2,44 | 0,4927 | 2,98 | 0,49856 |
| 1,71 | 0,4564 | 1,97 | 0,4756 | 2,46 | 0,4931 | 3,00 | 0,49865 |
| 1,72 | 0,4573 | 1,98 | 0,4761 | 2,48 | 0,4934 | 3,20 | 0,49931 |
| 1,73 | 0,4582 | 1,99 | 0,4767 | 2,50 | 0,4938 | 3,40 | 0,49966 |
| 1,74 | 0,4591 | 2,00 | 0,4772 | 2,52 | 0,4941 | 3,60 | 0,49984 |
| 1,75 | 0,4599 | 2,02 | 0,4783 | 2,54 | 0,4945 | 3,80 | 0,499928 |
| 1,76 | 0,4608 | 2,04 | 0,4793 | 2,56 | 0,4948 | 4,00 | 0,499968 |
| 1,77 | 0,4616 | 2,06 | 0,4803 | 2,58 | 0,4951 | 5,00 | 0,499997 |
| 1,78 | 0,4625 | 2,08 | 0,4812 | 2,60 | 0,4953 | | |
| 1,79 | 0,4633 | 2,10 | 0,4821 | 2,62 | 0,4956 | | |
| 1,80 | 0,4641 | 2,12 | 0,4830 | 2,64 | 0,4959 | | |
| 1,81 | 0,4649 | 2,14 | 0,4838 | 2,66 | 0,4961 | $x > 5$ | 0,5 |

**Додаток 20.**

## Таблиця розрахунку середніх темпів динаміки

| Середній темп | Коефіцієнти динаміки | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ( $\overline{T}$ ) | $\sqrt{\ }$ | $\sqrt[3]{\ }$ | $\sqrt[4]{\ }$ | $\sqrt[5]{\ }$ | $\sqrt[6]{\ }$ | $\sqrt[7]{\ }$ | $\sqrt[8]{\ }$ | $\sqrt[9]{\ }$ | $\sqrt[10]{\ }$ |
| 0,960 | 0,922 | 0,885 | 0,849 | 0,815 | 0,783 | 0,751 | 0,721 | 0,692 | 0,665 |
| 0,961 | 0,923 | 0,887 | 0,853 | 0,820 | 0,788 | 0,757 | 0,727 | 0,699 | 0,672 |
| 0,962 | 0,925 | 0,890 | 0,856 | 0,824 | 0,793 | 0,762 | 0,733 | 0,706 | 0,679 |
| 0,963 | 0,927 | 0,893 | 0,860 | 0,828 | 0,797 | 0,768 | 0,740 | 0,712 | 0,686 |
| 0,964 | 0,929 | 0,896 | 0,864 | 0,832 | 0,802 | 0,774 | 0,746 | 0,719 | 0,693 |
| 0,965 | 0,931 | 0,899 | 0,867 | 0,837 | 0,807 | 0,779 | 0,752 | 0,726 | 0,700 |
| 0,9655 | 0,932 | 0,900 | 0,869 | 0,839 | 0,810 | 0,782 | 0,755 | 0,729 | 0,704 |
| 0,966 | 0,933 | 0,901 | 0,871 | 0,841 | 0,813 | 0,785 | 0,758 | 0,732 | 0,708 |
| 0,967 | 0,935 | 0,904 | 0,874 | 0,845 | 0,818 | 0,791 | 0,765 | 0,739 | 0,715 |
| 0,968 | 0,937 | 0,907 | 0,878 | 0,850 | 0,823 | 0,796 | 0,771 | 0,746 | 0,722 |
| 0,969 | 0,939 | 0,910 | 0,882 | 0,851 | 0,828 | 0,802 | 0,777 | 0,753 | 0,730 |
| 0,970 | 0,941 | 0,913 | 0,885 | 0,859 | 0,833 | 0,808 | 0,784 | 0,760 | 0,737 |
| 0,971 | 0,943 | 0,915 | 0,889 | 0,863 | 0,838 | 0,814 | 0,790 | 0,767 | 0,745 |
| 0,972 | 0,945 | 0,918 | 0,893 | 0,868 | 0,843 | 0,820 | 0,797 | 0,774 | 0,753 |
| 0,973 | 0,947 | 0,921 | 0,896 | 0,872 | 0,848 | 0,826 | 0,803 | 0,782 | 0,761 |
| 0,974 | 0,949 | 0,924 | 0,900 | 0,877 | 0,854 | 0,832 | 0,810 | 0,789 | 0,768 |
| 0,975 | 0,951 | 0,927 | 0,904 | 0,881 | 0,859 | 0,838 | 0,817 | 0,796 | 0,776 |
| 0,9755 | 0,952 | 0,928 | 0,905 | 0,883 | 0,862 | 0,841 | 0,820 | 0,800 | 0,780 |
| 0,976 | 0,953 | 0,930 | 0,907 | 0,886 | 0,864 | 0,844 | 0,823 | 0,804 | 0,784 |
| 0,977 | 0,954 | 0,933 | 0,911 | 0,890 | 0,870 | 0,850 | 0,830 | 0,811 | 0,792 |
| 0,978 | 0,956 | 0,935 | 0,915 | 0,895 | 0,875 | 0,856 | 0,837 | 0,819 | 0,801 |
| 0,979 | 0,958 | 0,938 | 0,919 | 0,899 | 0,880 | 0,862 | 0,844 | 0,826 | 0,809 |
| 0,980 | 0,960 | 0,941 | 0,922 | 0,904 | 0,886 | 0,868 | 0,851 | 0,834 | 0,817 |
| 0,981 | 0,962 | 0,944 | 0,926 | 0,908 | 0,891 | 0,874 | 0,858 | 0,841 | 0,825 |
| 0,982 | 0,964 | 0,947 | 0,930 | 0,913 | 0,897 | 0,881 | 0,865 | 0,849 | 0,834 |
| 0,983 | 0,966 | 0,950 | 0,934 | 0,918 | 0,902 | 0,887 | 0,872 | 0,857 | 0,842 |
| 0,984 | 0,968 | 0,953 | 0,937 | 0,922 | 0,908 | 0,893 | 0,879 | 0,865 | 0,851 |
| 0,985 | 0,970 | 0,956 | 0,941 | 0,927 | 0,913 | 0,900 | 0,886 | 0,873 | 0,860 |
| 0,9855 | 0,971 | 0,957 | 0,943 | 0,930 | 0,916 | 0,903 | 0,890 | 0,877 | 0,864 |
| 0,986 | 0,972 | 0,959 | 0,945 | 0,932 | 0,919 | 0,906 | 0,893 | 0,881 | 0,868 |
| 0,987 | 0,974 | 0,961 | 0,949 | 0,937 | 0,924 | 0,912 | 0,901 | 0,889 | 0.877 |
| 0,988 | 0,976 | 0,964 | 0,953 | 0,941 | 0,930 | 0,919 | 0,908 | 0,897 | 0,886 |
| 0,989 | 0,978 | 0,967 | 0,957 | 0,946 | 0,936 | 0,925 | 0,915 | 0,905 | 0,895 |
| 0,990 | 0,980 | 0,970 | 0,961 | 0,951 | 0,941 | 0,932 | 0,923 | 0,913 | 0,904 |
| 0,991 | 0,982 | 0,973 | 0,964 | 0,956 | 0,947 | 0,939 | 0,930 | 0,922 | 0,914 |

**Продовження додатку 20.**

| $(\overline{T})$ | $\sqrt{\ }$ | $\sqrt[3]{\ }$ | $\sqrt[4]{\ }$ | $\sqrt[5]{\ }$ | $\sqrt[6]{\ }$ | $\sqrt[7]{\ }$ | $\sqrt[8]{\ }$ | $\sqrt[9]{\ }$ | $\sqrt[10]{\ }$ |
|---|---|---|---|---|---|---|---|---|---|
| 0,992 | 0,984 | 0,976 | 0,968 | 0,961 | 0,953 | 0,945 | 0,938 | 0,930 | 0,923 |
| 0,993 | 0,986 | 0,979 | 0,972 | 0,965 | 0,959 | 0,952 | 0,945 | 0,939 | 0,932 |
| 0,994 | 0,988 | 0,982 | 0,976 | 0,970 | 0,964 | 0,959 | 0,953 | 0,947 | 0,942 |
| 0,995 | 0,990 | 0,985 | 0,980 | 0,975 | 0,970 | 0,965 | 0,961 | 0,956 | 0,951 |
| 0,9955 | 0,991 | 0,987 | 0,982 | 0,978 | 0,973 | 0,969 | 0,965 | 0,960 | 0,956 |
| 0,996 | 0,992 | 0,988 | 0,981 | 0,980 | 0,976 | 0,972 | 0,968 | 0,965 | 0,961 |
| 0,997 | 0,994 | 0,991 | 0,988 | 0,985 | 0,982 | 0,979 | 0,976 | 0,973 | 0,970 |
| 0,998 | 0,996 | 0,994 | 0,992 | 0,990 | 0,988 | 0,986 | 0,984 | 0,982 | 0,980 |
| 0,999 | 0,998 | 0,997 | 0,996 | 0,995 | 0,994 | 0,993 | 0,992 | 0,991 | 0,990 |
| 1,001 | 1,002 | 1,003 | 1,004 | 1,005 | 1,006 | 1,007 | 1,008 | 1,009 | 1,010 |
| 1,002 | 1,004 | 1,006 | 1,008 | 1,010 | 1,012 | 1,014 | 1,016 | 1,018 | 1,020 |
| 1,0025 | 1,005 | 1,0075 | 1,010 | 1,0126 | 1,015 | 1,018 | 1,020 | 1,023 | 1,025 |
| 1,003 | 1,006 | 1,009 | 1,012 | 1,015 | 1,018 | 1,021 | 1,024 | 1,027 | 1,030 |
| 1,004 | 1,008 | 1,012 | 1,016 | 1,020 | 1,024 | 1,028 | 1,032 | 1,037 | 1,041 |
| 1,005 | 1,010 | 1,015 | 1,020 | 1,025 | 1,030 | 1,035 | 1,041 | 1,046 | 1,051 |
| 1,0055 | 1,011 | 1,017 | 1,022 | 1,028 | 1,033 | 1,039 | 1,045 | 1,051 | 1,056 |
| 1,006 | 1,012 | 1,018 | 1,024 | 1,030 | 1,036 | 1,043 | 1,049 | 1,055 | 1,062 |
| 1,007 | 1,014 | 1,021 | 1,028 | 1,035 | 1,043 | 1,050 | 1,057 | 1,065 | 1,072 |
| 1,0075 | 1,015 | 1,023 | 1,030 | 1,0380 | 1,046 | 1,054 | 1,062 | 1,070 | 1,078 |
| 1,008 | 1,016 | 1,024 | 1,032 | 1,041 | 1,049 | 1,057 | 1,066 | 1,074 | 1,083 |
| 1,009 | 1,018 | 1,027 | 1,036 | 1,046 | 1,055 | 1,065 | 1,074 | 1,084 | 1,094 |
| 1,010 | 1,0200 | 1,0300 | 1,0400 | 1,0500 | 1,0605 | 1,0710 | 1,0820 | 1,0930 | 1,1040 |
| 1,011 | 1,0220 | 1,0330 | 1,0440 | 1,0550 | 1,0670 | 1,0790 | 1,0910 | 1,1030 | 1,1150 |
| 1,012 | 1,0240 | 1,0360 | 1,0480 | 1,0606 | 1,0740 | 1,0870 | 1,1000 | 1,1130 | 1,1260 |
| 1,0125 | 1,0252 | 1,0380 | 1,0510 | 1,0641 | 1,0774 | 1,0909 | 1,1045 | 1,1183 | 1,1323 |
| 1,013 | 1,0262 | 1,0395 | 1,0530 | 1,0667 | 1,0806 | 1,0946 | 1,1088 | 1,1232 | 1,1378 |
| 1,014 | 1,0282 | 1,0426 | 1,0572 | 1,0720 | 1,0870 | 1,1022 | 1,1176 | 1,1332 | 1,1491 |
| 1,015 | 1,0302 | 1,0456 | 1,0613 | 1,0772 | 1,0934 | 1,1098 | 1,1264 | 1,1433 | 1,1604 |
| 1,016 | 1,0323 | 1,0488 | 1,0656 | 1,0826 | 1,0999 | 1,1175 | 1,1354 | 1,1536 | 1,1721 |
| 1,017 | 1,0343 | 1,0519 | 1,0698 | 1,0880 | 1,1065 | 1,1253 | 1,1444 | 1,1638 | 1,1836 |
| 1,0175 | 1,0353 | 1,0534 | 1,0718 | 1,0906 | 1,1097 | 1,1291 | 1,1489 | 1,1690 | 1,1890 |
| 1,018 | 1,0363 | 1,0549 | 1,0739 | 1,0932 | 1,1129 | 1,1329 | 1,1533 | 1,1741 | 1,1952 |
| 1,019 | 1,0384 | 1,0581 | 1,0782 | 1,0987 | 1,1196 | 1,1409 | 1,1626 | 1,1847 | 1,2070 |
| 1,020 | 1,0404 | 1,0612 | 1,0824 | 1,1040 | 1,1261 | 1,1486 | 1,1716 | 1,1950 | 1,2190 |
| 1,021 | 1,0424 | 1,0643 | 1,0866 | 1,1094 | 1,1327 | 1,1565 | 1,1808 | 1,2051 | 1,2309 |
| 1,022 | 1,0445 | 1,0675 | 1,0910 | 1,1150 | 1,1395 | 1,1646 | 1,1902 | 1,2164 | 1,2432 |
| 1,023 | 1,0465 | 1,0706 | 1,0952 | 1,1204 | 1,1462 | 1,1726 | 1,1996 | 1,2272 | 1,2554 |
| 1,024 | 1,0486 | 1,0738 | 1,0996 | 1,1260 | 1,1530 | 1,1807 | 1,2090 | 1,2380 | 1,2677 |

**Продовження додатку 20.**

| $(\overline{T})$ | $\sqrt{\ }$ | $\sqrt[3]{\ }$ | $\sqrt[4]{\ }$ | $\sqrt[5]{\ }$ | $\sqrt[6]{\ }$ | $\sqrt[7]{\ }$ | $\sqrt[8]{\ }$ | $\sqrt[9]{\ }$ | $\sqrt[10]{\ }$ |
|---|---|---|---|---|---|---|---|---|---|
| 1,025 | 1,0506 | 1,0769 | 1,1038 | 1,1314 | 1,1597 | 1,1887 | 1,2184 | 1,2489 | 1,2801 |
| 1,026 | 1,0527 | 1,0801 | 1,1082 | 1,1370 | 1,1666 | 1,1969 | 1,2280 | 1,2599 | 1,2927 |
| 1,027 | 1,0547 | 1,0832 | 1,1124 | 1,1424 | 1,1732 | 1,2049 | 1,2374 | 1,2708 | 1,3051 |
| 1,0275 | 1,0558 | 1,0848 | 1,1146 | 1,1452 | 1,1767 | 1,2091 | 1,2423 | 1,2765 | 1,3116 |
| 1,028 | 1,0568 | 1,0864 | 1,1168 | 1,1481 | 1,1802 | 1,2132 | 1,2472 | 1,2821 | 1,3180 |
| 1,029 | 1,0588 | 1,0895 | 1,1211 | 1,1536 | 1,1870 | 1,2214 | 1,2568 | 1,2932 | 1,3307 |
| 1,030 | 1,0609 | 1,0927 | 1,1255 | 1,1593 | 1,1941 | 1,2290 | 1,2668 | 1,3048 | 1,3349 |
| 1,031 | 1,0630 | 1,0960 | 1,1300 | 1,1650 | 1,2011 | 1,2383 | 1,2767 | 1,3163 | 1,3571 |
| 1,032 | 1,0650 | 1,0991 | 1,1345 | 1,1706 | 1,2080 | 1,2467 | 1,2866 | 1,3278 | 1,3703 |
| 1,0325 | 1,0661 | 1,1007 | 1,1365 | 1,1734 | 1,2115 | 1,2509 | 1,2915 | 1,3335 | 1,3768 |
| 1,033 | 1,0671 | 1,1023 | 1,1387 | 1,1763 | 1,2151 | 1,2552 | 1,2966 | 1,3394 | 1,3836 |
| 1,034 | 1,0692 | 1,1055 | 1,1431 | 1,1820 | 1,2222 | 1,2637 | 1,3067 | 1,3511 | 1,3970 |
| 1,035 | 1,0712 | 1,1087 | 1,1475 | 1,1877 | 1,2293 | 1,2723 | 1,3168 | 1,3629 | 1,4106 |
| 1,036 | 1,0733 | 1,1119 | 1,1519 | 1,1934 | 1,2364 | 1,2809 | 1,3270 | 1,3748 | 1,4243 |
| 1,037 | 1,0754 | 1,1152 | 1,1565 | 1,1993 | 1,2437 | 1,2897 | 1,3394 | 1,3869 | 1,4382 |
| 1,0375 | 1,0764 | 1,1168 | 1,1587 | 1,2021 | 1,2472 | 1,2940 | 1,3425 | 1,3928 | 1,4450 |
| 1,038 | 1,0774 | 1,1183 | 1,1608 | 1,2049 | 1,2507 | 1,2982 | 1,3475 | 1,3987 | 1,4518 |
| 1,039 | 1,0795 | 1,1216 | 1,1653 | 1,2107 | 1,2579 | 1,3070 | 1,3580 | 1,4110 | 1,4660 |
| 1,040 | 1,0820 | 1,1253 | 1,1703 | 1,2171 | 1,2658 | 1,3164 | 1,3691 | 1,4239 | 1,4809 |
| 1,041 | 1,0837 | 1,1281 | 1,1743 | 1,2224 | 1,2725 | 1,3247 | 1,3790 | 1,4355 | 1,4944 |
| 1,042 | 1,0858 | 1,1314 | 1,1789 | 1,2284 | 1,2800 | 1,3338 | 1,3898 | 1,4482 | 1,5090 |
| 1,0425 | 1,0868 | 1,1330 | 1,1811 | 1,2313 | 1,2836 | 1,3381 | 1,3950 | 1,4543 | 1,5161 |
| 1,043 | 1,0878 | 1,1346 | 1,1834 | 1,2343 | 1,2874 | 1,3428 | 1,4005 | 1,4607 | 1,5235 |
| 1,044 | 1,0899 | 1,1379 | 1,1880 | 1,2403 | 1,2949 | 1,3519 | 1,4114 | 1,4735 | 1,5383 |
| 1,045 | 1,0920 | 1,1411 | 1,1924 | 1,2461 | 1,3022 | 1,3608 | 1,4220 | 1,4860 | 1,5529 |
| 1,046 | 1,0941 | 1,1444 | 1,1970 | 1,2521 | 1,3097 | 1,3699 | 1,4329 | 1,4988 | 1,5677 |
| 1,047 | 1,0962 | 1,1477 | 1,2016 | 1,2581 | 1,3172 | 1,3791 | 1,4439 | 1,5118 | 1,5829 |
| 1,0475 | 1,0973 | 1,1494 | 1,2040 | 1,2612 | 1,3211 | 1,3839 | 1,4496 | 1,5185 | 1,5906 |
| 1,048 | 1,0983 | 1,1510 | 1,2062 | 1,2641 | 1,3248 | 1,3884 | 1,4550 | 1,5248 | 1,5980 |
| 1,049 | 1,1004 | 1,1543 | 1,2109 | 1,2702 | 1,3324 | 1,3977 | 1,4662 | 1,5380 | 1,6134 |
| 1,050 | 1,1025 | 1,1576 | 1,2155 | 1,2763 | 1,3401 | 1,4071 | 1,4775 | 1,5514 | 1,6290 |
| 1,051 | 1,1046 | 1,1609 | 1,2201 | 1,2823 | 1,3477 | 1,4164 | 1,4886 | 1,5645 | 1,6443 |
| 1,052 | 1,1067 | 1,1642 | 1,2247 | 1,2884 | 1,3554 | 1,4259 | 1,5000 | 1,5780 | 1,6600 |
| 1,0525 | 1,1078 | 1,1660 | 1,2272 | 1,2916 | 1,3594 | 1,4308 | 1,5059 | 1,5850 | 1,6682 |
| 1,053 | 1,1088 | 1,1676 | 1,2295 | 1,2947 | 1,3633 | 1,4356 | 1,5117 | 1,5918 | 1,6762 |
| 1,054 | 1,1109 | 1,1709 | 1,2341 | 1,3007 | 1,3709 | 1,4449 | 1,5229 | 1,6051 | 1,6918 |
| 1,055 | 1,1130 | 1,1742 | 1,2388 | 1,3069 | 1,3788 | 1,4546 | 1,5346 | 1,6190 | 1,7080 |
| 1,056 | 1,1151 | 1,1775 | 1,2434 | 1,3130 | 1,3865 | 1,4641 | 1,5461 | 1,6327 | 1,7241 |

**Продовження додатку 20.**

| $(\overline{T})$ | $\sqrt{\ }$ | $\sqrt[3]{\ }$ | $\sqrt[4]{\ }$ | $\sqrt[5]{\ }$ | $\sqrt[6]{\ }$ | $\sqrt[7]{\ }$ | $\sqrt[8]{\ }$ | $\sqrt[9]{\ }$ | $\sqrt[10]{\ }$ |
|---|---|---|---|---|---|---|---|---|---|
| 1,057 | 1,1172 | 1,1809 | 1,2482 | 1,3193 | 1,3945 | 1,4740 | 1,5580 | 1,6468 | 1,7407 |
| 1,0575 | 1,1183 | 1,1826 | 1,2506 | 1,3225 | 1,3985 | 1,4789 | 1,5639 | 1,6538 | 1,7489 |
| 1,058 | 1,1194 | 1,1843 | 1,2530 | 1,3257 | 1,4026 | 1,4840 | 1,5701 | 1,6612 | 1,7575 |
| 1,059 | 1,1215 | 1,1877 | 1,2578 | 1,3320 | 1,4106 | 1,4938 | 1,5819 | 1,6752 | 1,7740 |
| 1,060 | 1,1236 | 1,1910 | 1,2625 | 1,3383 | 1,4186 | 1,5037 | 1,5939 | 1,6895 | 1,7909 |
| 1,061 | 1,1257 | 1,1944 | 1,2673 | 1,3446 | 1,4266 | 1,5136 | 1,6059 | 1,7039 | 1,8078 |
| 1,062 | 1,1278 | 1,1977 | 1,2720 | 1,3509 | 1,4347 | 1,5237 | 1,6182 | 1,7195 | 1,8250 |
| 1,0625 | 1,1289 | 1,1995 | 1,2745 | 1,3542 | 1,4388 | 1,5287 | 1,6242 | 1,7257 | 1,8336 |
| 1,063 | 1,1300 | 1,2012 | 1,2769 | 1,3573 | 1,4428 | 1,5337 | 1,6303 | 1,7330 | 1,8422 |
| 1,064 | 1,1321 | 1,2046 | 1,2817 | 1,3637 | 1,4510 | 1,5439 | 1,6427 | 1,7478 | 1,8597 |
| 1,065 | 1,1342 | 1,2079 | 1,2864 | 1,3700 | 1,4590 | 1,5538 | 1,6548 | 1,7624 | 1,8770 |
| 1,066 | 1,1364 | 1,2114 | 1,2914 | 1,3766 | 1,4675 | 1,5644 | 1,6677 | 1,7778 | 1,8951 |
| 1,067 | 1,1385 | 1,2148 | 1,2962 | 1,3830 | 1,4757 | 1,5716 | 1,6801 | 1,7927 | 1,9128 |
| 1,0675 | 1,1396 | 1,2165 | 1,2986 | 1,3863 | 1,4799 | 1,5798 | 1,6864 | 1,8002 | 1,9217 |
| 1,068 | 1,1406 | 1,2182 | 1,3010 | 1,3895 | 1,4840 | 1,5849 | 1,6927 | 1,8078 | 1,9307 |
| 1,069 | 1,1428 | 1,2210 | 1,3060 | 1,3961 | 1,4924 | 1,5954 | 1,7055 | 1,8232 | 1,9490 |
| 1,070 | 1,1449 | 1,2250 | 1,3108 | 1,4026 | 1,5008 | 1,6059 | 1,7183 | 1,8386 | 1,9673 |
| 1,071 | 1,1470 | 1,2284 | 1,3156 | 1,4090 | 1,5090 | 1,6161 | 1,7308 | 1,8537 | 1,9853 |
| 1,072 | 1,1492 | 1,2319 | 1,3206 | 1,4157 | 1,5176 | 1,6269 | 1,7440 | 1,8696 | 2,0042 |
| 1,0725 | 1,1503 | 1,2337 | 1,3231 | 1,4190 | 1,5219 | 1,6322 | 1,7505 | 1,8774 | 2,0135 |
| 1,073 | 1,1513 | 1,2353 | 1,3255 | 1,4223 | 1,5261 | 1,6375 | 1,7570 | 1,8853 | 2,0229 |
| 1,074 | 1,1535 | 1,2389 | 1,3306 | 1,4291 | 1,5349 | 1,6485 | 1,7705 | 1,9015 | 2,0422 |
| 1,075 | 1,1556 | 1,2423 | 1,3355 | 1,4357 | 1,5434 | 1,6592 | 1,7836 | 1,9174 | 2,0612 |
| 1,076 | 1,1578 | 1,2458 | 1,3405 | 1,4424 | 1,5520 | 1,6670 | 1,7937 | 1,9300 | 2,0767 |
| 1,077 | 1,1599 | 1,2492 | 1,3454 | 1,4490 | 1,5606 | 1,6808 | 1,8102 | 1,9496 | 2,0997 |
| 1,0775 | 1,1610 | 1,2510 | 1,3480 | 1,4525 | 1,5651 | 1,6864 | 1,8171 | 1,9579 | 2,1096 |
| 1,078 | 1,1621 | 1,2527 | 1,3504 | 1,4557 | 1,5692 | 1,6916 | 1,8235 | 1,9657 | 2,1190 |
| 1,079 | 1,1642 | 1,2562 | 1,3554 | 1,4625 | 1,5780 | 1,7027 | 1,8372 | 1,9823 | 2,1389 |
| 1,080 | 1,1664 | 1,2597 | 1,3605 | 1,4693 | 1,5868 | 1,7137 | 1,8508 | 1,9989 | 2,1588 |
| 1,081 | 1,1686 | 1,2633 | 1,3656 | 1,4762 | 1,5958 | 1,7251 | 1,8648 | 2,0158 | 2,1791 |
| 1,082 | 1,1707 | 1,2667 | 1,3706 | 1,4830 | 1,6046 | 1,7362 | 1,8786 | 2,0326 | 2,1993 |
| 1,0825 | 1,1718 | 1,2685 | 1,3732 | 1,4865 | 1,6091 | 1,7419 | 1,8856 | 2,0412 | 2,2096 |
| 1,083 | 1,1729 | 1,2703 | 1,3757 | 1,4899 | 1,6136 | 1,7475 | 1,8925 | 2,0496 | 2,2197 |
| 1,084 | 1,1751 | 1,2738 | 1,3808 | 1,4968 | 1,6225 | 1,7588 | 1,9065 | 2,0666 | 2,2402 |
| 1,085 | 1,1772 | 1,2773 | 1,3859 | 1,5037 | 1,6315 | 1,7702 | 1,9207 | 2,0810 | 2,2611 |
| 1,086 | 1,1794 | 1,2808 | 1,3909 | 1,5105 | 1,6404 | 1,7815 | 1,9347 | 2,1011 | 2,2818 |
| 1,087 | 1,1816 | 1,2844 | 1,3961 | 1,5176 | 1,6496 | 1,7931 | 1,9491 | 2,1187 | 2,3030 |
| 1,0875 | 1,1827 | 1,2862 | 1,3987 | 1,5211 | 1,6542 | 1,7989 | 1,9563 | 2,1275 | 2,3137 |

**Продовження додатку 20.**

| $(\overline{T})$ | $\sqrt{\ }$ | $\sqrt[3]{\ }$ | $\sqrt[4]{\ }$ | $\sqrt[5]{\ }$ | $\sqrt[6]{\ }$ | $\sqrt[7]{\ }$ | $\sqrt[8]{\ }$ | $\sqrt[9]{\ }$ | $\sqrt[10]{\ }$ |
|---|---|---|---|---|---|---|---|---|---|
| 1,088 | 1,1837 | 1,2879 | 1,4012 | 1,5245 | 1,6587 | 1,8047 | 1,9635 | 2,1363 | 2,3243 |
| 1,089 | 1,1859 | 1,2914 | 1,4063 | 1,5315 | 1,6678 | 1,8162 | 1,9778 | 2,1538 | 2,3455 |
| 1,090 | 1,1882 | 1,2950 | 1,4116 | 1,5386 | 1,6771 | 1,8280 | 1,9925 | 2,1718 | 2,3673 |
| 1,091 | 1,1903 | 1,2986 | 1,4168 | 1,5457 | 1,6864 | 1,8399 | 2,0073 | 2,1900 | 2,3893 |
| 1,092 | 1,1925 | 1,3022 | 1,4220 | 1,5528 | 1,6957 | 1,8517 | 2,0221 | 2,2081 | 2,4112 |
| 1,0925 | 1,1936 | 1,3040 | 1,4246 | 1,5564 | 1,7004 | 1,8577 | 2,0295 | 2,2172 | 2,4223 |
| 1,093 | 1,1946 | 1,3057 | 1,4271 | 1,5598 | 1,7049 | 1,8635 | 2,0368 | 2,2262 | 2,4332 |
| 1,094 | 1,1968 | 1,3093 | 1,4324 | 1,5670 | 1,7143 | 1,8754 | 2,0517 | 2,2446 | 2,4556 |
| 1,095 | 1,1990 | 1,3129 | 1,4376 | 1,5742 | 1,7237 | 1,8875 | 2,0668 | 2,2631 | 2,4781 |
| 1,096 | 1,2012 | 1,3165 | 1,4429 | 1,5814 | 1,7332 | 1,8996 | 2,0820 | 2,2819 | 2,5010 |
| 1,097 | 1,2034 | 1,3201 | 1,4481 | 1,5886 | 1,7427 | 1,9117 | 2,0971 | 2,3005 | 2,5236 |
| 1,0975 | 1,2045 | 1,3219 | 1,4508 | 1,5923 | 1,7475 | 1,9179 | 2,1049 | 2,3101 | 2,5353 |
| 1,098 | 1,2056 | 1,3237 | 1,4534 | 1,5958 | 1,7522 | 1,9239 | 2,1124 | 2,3194 | 2,5467 |
| 1,099 | 1,2078 | 1,3274 | 1,4588 | 1,6032 | 1,7619 | 1,9365 | 2,1282 | 2,3389 | 2,5705 |
| 1,100 | 1,2100 | 1,3310 | 1,4641 | 1,6105 | 1,7716 | 1,9488 | 2,1437 | 2,3581 | 2,5939 |
| 1,101 | 1,2122 | 1,3346 | 1,4694 | 1,6178 | 1,7812 | 1,9611 | 2,1592 | 2,3773 | 2,6174 |
| 1,102 | 1,2144 | 1,3388 | 1,4748 | 1,6252 | 1,7910 | 1,9737 | 2,1750 | 2,3969 | 2,6414 |
| 1,1025 | 1,2155 | 1,3401 | 1,4775 | 1,6289 | 1,7959 | 1,9800 | 2,1830 | 2,4068 | 2,6535 |
| 1,103 | 1,2166 | 1,3419 | 1,4801 | 1,6326 | 1,8008 | 1,9863 | 2,1909 | 2,4166 | 2,6655 |
| 1,104 | 1,2188 | 1,3456 | 1,4855 | 1,6400 | 1,8106 | 1,9989 | 2,2068 | 2,4363 | 2,6897 |
| 1,105 | 1,2210 | 1,3492 | 1.4909 | 1,6474 | 1,8204 | 2,0115 | 2,2227 | 2,4561 | 2,7140 |
| 1,106 | 1,2232 | 1,3529 | 1,4963 | 1,6549 | 1,8303 | 2,0243 | 2,2389 | 2,4762 | 2,7387 |
| 1,107 | 1,2254 | 1,3565 | 1,5016 | 1,6623 | 1,8402 | 2,0371 | 2,2551 | 2,4964 | 2,7635 |
| 1,1075 | 1,2266 | 1,3585 | 1,5045 | 1,6662 | 1,8453 | 2,0437 | 2,2634 | 2,5067 | 2,7762 |
| 1,108 | 1,2277 | 1,3603 | 1,5072 | 1,6700 | 1,8504 | 2,0502 | 2,2716 | 2,5169 | 2,7887 |
| 1,109 | 1,2299 | 1,3640 | 1,5127 | 1,6776 | 1,8605 | 2,0633 | 2,2882 | 2,5376 | 2,8142 |
| 1,110 | 1,2321 | 1,3676 | 1,5180 | 1,6850 | 1,8704 | 2,0761 | 2,3045 | 2,5580 | 2,8394 |
| 1,111 | 1,2343 | 1,3713 | 1,5235 | 1,6926 | 1,8805 | 2,0892 | 2,3211 | 2,5787 | 2,8649 |
| 1,112 | 1,2365 | 1,3750 | 1,5290 | 1,7002 | 1,8906 | 2,1023 | 2,3378 | 2,5996 | 2,8908 |
| 1,1125 | 1,2377 | 1,3769 | 1,5318 | 1,7041 | 1,8958 | 2,1091 | 2,3464 | 2,6104 | 2,9041 |
| 1,113 | 1,2388 | 1,3788 | 1,5346 | 1,7080 | 1,9010 | 2,1158 | 2,3549 | 2,6210 | 2,9172 |
| 1,114 | 1,2410 | 1,3825 | 1,5401 | 1,7157 | 1,9113 | 2,1292 | 2,3719 | 2,6423 | 2,9435 |
| 1,115 | 1,2432 | 1,3862 | 1,5456 | 1,7233 | 1,9215 | 2,1425 | 2,3889 | 2,6636 | 2,9699 |
| 1,116 | 1,2455 | 1,3900 | 1,5512 | 1,7311 | 1,9319 | 2,1560 | 2,4061 | 2,6852 | 2,9967 |
| 1,117 | 1,2477 | 1,3937 | 1,5568 | 1,7389 | 1,9424 | 2,1697 | 2,4236 | 2,7072 | 3,0239 |
| 1,118 | 1,2499 | 1,3974 | 1,5623 | 1,7467 | 1,9528 | 2,1832 | 2,4408 | 2,7288 | 3,0508 |
| 1,19 | 1,2522 | 1,4012 | 1,5679 | 1,7545 | 1,9633 | 2,1969 | 2,4583 | 2,7508 | 3,0781 |
| 1,120 | 1,2544 | 1,4049 | 1,5735 | 1,7623 | 1,9738 | 2,2107 | 2,4760 | 2,7731 | 3,1059 |

**Продовження додатку 20.**

| $(\overline{T})$ | $\sqrt{\ }$ | $\sqrt[3]{\ }$ | $\sqrt[4]{\ }$ | $\sqrt[5]{\ }$ | $\sqrt[6]{\ }$ | $\sqrt[7]{\ }$ | $\sqrt[8]{\ }$ | $\sqrt[9]{\ }$ | $\sqrt[10]{\ }$ |
|---|---|---|---|---|---|---|---|---|---|
| 1,121 | 1,2566 | 1,4086 | 1,5790 | 1,7701 | 1,9843 | 2,2244 | 2,4936 | 2,7953 | 3,1335 |
| 1,122 | 1,2589 | 1,4125 | 1,5848 | 1,7781 | 1,9950 | 2,2384 | 2,5115 | 2,8179 | 3,1617 |
| 1,1225 | 1,2600 | 1,4144 | 1,5877 | 1,7822 | 2,0005 | 2,2456 | 2,5207 | 2,8295 | 3,1761 |
| 1,123 | 1,2611 | 1,4162 | 1,5904 | 1,7860 | 2,0057 | 2,2524 | 2,5294 | 2,8405 | 3,1899 |
| 1,124 | 1,2684 | 1,4201 | 1,5962 | 1,7941 | 2,0166 | 2,2667 | 2,5478 | 2,8637 | 3,2188 |
| 1,125 | 1,2656 | 1,4238 | 1,6018 | 1,8020 | 2,0273 | 2,2807 | 2,5658 | 2,8865 | 3,2473 |
| 1,126 | 1,2679 | 1,4277 | 1,6076 | 1,8102 | 2,0388 | 2,2951 | 2,5843 | 2,9099 | 3,2765 |
| 1,127 | 1,2701 | 1,4314 | 1,6132 | 1,8181 | 2,0490 | 2,3092 | 2,6025 | 2,9330 | 3,3055 |
| 1,1275 | 1,2713 | 1,4334 | 1,6162 | 1,8223 | 2,0546 | 2,3166 | 2,6120 | 2,9450 | 3,3205 |
| 1,128 | 1,2724 | 1,4353 | 1,6190 | 1,8262 | 2,0600 | 2,3237 | 2,6211 | 2,9566 | 3,3350 |
| 1,129 | 1,2746 | 1,4390 | 1,6246 | 1,8342 | 2,0708 | 2,3379 | 2,6395 | 2,9800 | 3,3644 |
| 1,130 | 1,2769 | 1,4429 | 1,6305 | 1,8425 | 2,0820 | 2,3527 | 2,6586 | 3,0042 | 3,3947 |
| 1,131 | 1,2792 | 1,4468 | 1,6363 | 1,8507 | 2,0931 | 2,3673 | 2,6774 | 3,0281 | 3,4248 |
| 1,132 | 1,2814 | 1,4505 | 1,6420 | 1,8587 | 2,1040 | 2,3817 | 2,6961 | 3,0520 | 3,4549 |
| 1,1325 | 1,2826 | 1,4525 | 1,6450 | 1,8630 | 2,1098 | 2,3893 | 2,7059 | 3,0644 | 3,4704 |
| 1,133 | 1,2837 | 1,4544 | 1,6478 | 1,8670 | 2,1153 | 2,3966 | 2,7153 | 3,0764 | 3,4856 |
| 1,134 | 1,2860 | 1,4583 | 1,6537 | 1,8753 | 2,1266 | 2,4116 | 2,7348 | 3,1013 | 3,5169 |
| 1,135 | 1,2882 | 1,4621 | 1,6595 | 1,8835 | 2,1378 | 2,4264 | 2,7540 | 3,1258 | 3,5478 |
| 1,136 | 1,2905 | 1,4660 | 1,6654 | 1,8919 | 2,1492 | 2,4415 | 2,7735 | 3,1507 | 3,5792 |
| 1,137 | 1,2928 | 1,4699 | 1,6713 | 1,9003 | 2,1606 | 2,4566 | 2,7932 | 3,1759 | 3,6110 |
| 1,1375 | 1,2939 | 1,4718 | 1,6742 | 1,9044 | 2,1663 | 2,4642 | 2,8030 | 3,1884 | 3,6268 |
| 1,138 | 1,2950 | 1,4737 | 1,6771 | 1,9085 | 2,1719 | 2,4716 | 2,8127 | 3,2009 | 3,6426 |
| 1,139 | 1,2973 | 1,4776 | 1,6830 | 1,9169 | 2,1833 | 2,4868 | 2,8325 | 3,2262 | 3,6746 |
| 1,140 | 1,2996 | 1,4815 | 1,6889 | 1,9253 | 2,1948 | 2,5021 | 2,8524 | 3,2517 | 3,7069 |
| 1,141 | 1,3019 | 1,4855 | 1,6950 | 1,9340 | 2,2067 | 2,5178 | 2,8728 | 3,2779 | 3,7401 |
| 1,142 | 1,3042 | 1,4894 | 1,7009 | 1,9124 | 2,2182 | 2,5332 | 2,8929 | 3,3037 | 3,7728 |
| 1,1425 | 1,3053 | 1,4913 | 1,7038 | 1,9466 | 2,2240 | 2,5409 | 2,9030 | 3,3167 | 3,7893 |
| 1,143 | 1,3064 | 1,4932 | 1,7067 | 1,9508 | 2,2293 | 2,5487 | 2,9132 | 3,3293 | 3,8060 |
| 1,144 | 1,3087 | 1,4972 | 1,7128 | 1,9594 | 2,2416 | 2,5644 | 2,9337 | 3,3562 | 3,8395 |
| 1,145 | 1,3110 | 1,5011 | 1,7188 | 1,9680 | 2,2534 | 2,5801 | 2,9542 | 3,3826 | 3,8731 |
| 1,146 | 1,3133 | 1,5050 | 1,7247 | 1,9765 | 2,2651 | 2,5958 | 2,9748 | 3,4091 | 3,9068 |
| 1,147 | 1,3156 | 1,5090 | 1,7303 | 1,9852 | 2,2770 | 2,6117 | 2,9956 | 3,4360 | 3,9411 |
| 1,148 | 1,3179 | 1,5129 | 1,7368 | 1,9938 | 2,2889 | 2,6277 | 3,0166 | 3,4631 | 3,9756 |
| 1,149 | 1,3202 | 1,5169 | 1,7429 | 2,0026 | 2,3010 | 2,6438 | 3,0377 | 3,4903 | 4,0104 |
| 1,150 | 1,3225 | 1,5209 | 1,7490 | 2,0114 | 2,3131 | 2,6601 | 3,0591 | 3,5180 | 4,0457 |
| 1,151 | 1,3248 | 1,5248 | 1,7550 | 2,0200 | 2,3250 | 2,6761 | 3,0802 | 3,5453 | 4,0806 |
| 1,1525 | 1,3283 | 1,5309 | 1,7644 | 2,0335 | 2,3436 | 2,7010 | 3,1129 | 3,5876 | 4,1347 |
| 1,153 | 1,3294 | 1,5328 | 1,7673 | 2,0377 | 2,3395 | 2,7090 | 3,1235 | 3,6014 | 4,1524 |

**Додаток 21.**

## Таблиця випадкових чисел

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| 5880 | 1257 | 6163 | 4439 | 7276 | 6353 | 6912 | 0731 | 9033 | 5294 |
| 9083 | 4260 | 5277 | 4998 | 4298 | 5204 | 3965 | 4028 | 8936 | 5148 |
| 1762 | 8713 | 1189 | 1090 | 8989 | 7273 | 3213 | 1935 | 9321 | 4820 |
| 2023 | 2509 | 1740 | 0424 | 8924 | 0005 | 1969 | 1636 | 7237 | 1227 |
| 7965 | 3855 | 4765 | 0703 | 1678 | 0841 | 7543 | 0308 | 9732 | 1289 |
| 7690 | 0480 | 8098 | 9629 | 4819 | 7219 | 7214 | 5128 | 3853 | 1921 |
| 9292 | 0426 | 9573 | 4903 | 5916 | 6517 | 8368 | 3270 | 6641 | 0038 |
| 0867 | 1651 | 7016 | 4220 | 2533 | 6345 | 8227 | 1904 | 5138 | 2537 |
| 0505 | 2127 | 8255 | 5276 | 2233 | 3956 | 4118 | 8199 | 6380 | 6340 |
| 6290 | 9795 | 1112 | 5765 | 2575 | 6837 | 3336 | 9222 | 7403 | 8345 |
| 6323 | 2615 | 3410 | 3365 | 1117 | 2417 | 3176 | 2434 | 5240 | 5455 |
| 8672 | 8536 | 2966 | 5773 | 5412 | 8114 | 0903 | 4697 | 6919 | 4569 |
| 1422 | 5507 | 7596 | 0670 | 3013 | 1351 | 3886 | 3268 | 9469 | 2534 |
| 2553 | 1473 | 5113 | 5735 | 1469 | 9545 | 9331 | 5303 | 9914 | 6394 |
| 0438 | 4376 | 3328 | 8649 | 8327 | 0110 | 4549 | 7955 | 5275 | 2890 |
| 2851 | 2157 | 0047 | 7085 | 1129 | 0460 | 6821 | 8323 | 2572 | 8962 |
| 7962 | 2753 | 3077 | 8718 | 7418 | 8004 | 1425 | 3706 | 8822 | 1494 |
| 3837 | 4098 | 0220 | 1217 | 4732 | 0150 | 1637 | 1097 | 1040 | 7372 |
| 8542 | 4126 | 9274 | 2251 | 0607 | 4301 | 8730 | 7690 | 6235 | 3477 |
| 0139 | 0765 | 8039 | 9484 | 3577 | 7859 | 1976 | 0623 | 1418 | 6685 |
| 6687 | 1943 | 4307 | 0579 | 8171 | 8224 | 8641 | 7034 | 3595 | 3875 |
| 6242 | 5582 | 5872 | 3197 | 4919 | 2792 | 5991 | 4058 | 9769 | 1918 |
| 6859 | 9606 | 0522 | 4993 | 0345 | 8958 | 1289 | 5825 | 6941 | 7685 |
| 6590 | 1932 | 6043 | 3623 | 1978 | 4112 | 1795 | 8465 | 2110 | 8045 |
| 3482 | 0478 | 0221 | 6738 | 7323 | 5643 | 4767 | 0106 | 2272 | 9862 |
| 5489 | 5583 | 3156 | 0835 | 1988 | 3912 | 0938 | 7460 | 0869 | 4420 |
| 3522 | 0935 | 7877 | 5665 | 7020 | 9555 | 7379 | 7124 | 7878 | 5544 |
| 7555 | 7579 | 2550 | 2487 | 9477 | 0864 | 2349 | 1012 | 8250 | 2633 |
| 5759 | 3554 | 5080 | 9074 | 7001 | 6249 | 3234 | 6368 | 9102 | 2672 |
| 6303 | 6895 | 3371 | 3196 | 7231 | 2918 | 7380 | 0438 | 7547 | 2644 |
| 7351 | 5634 | 5323 | 2623 | 7803 | 8374 | 2191 | 0464 | 0696 | 9529 |
| 7068 | 7803 | 8832 | 5119 | 6350 | 0120 | 5026 | 3684 | 5667 | 0304 |
| 3613 | 1428 | 1796 | 8447 | 0503 | 5654 | 3254 | 7336 | 9536 | 1944 |
| 5143 | 4534 | 2105 | 0368 | 7890 | 2473 | 4240 | 8562 | 9435 | 1422 |
| 9815 | 5144 | 7649 | 8638 | 6137 | 8070 | 5345 | 4865 | 2456 | 5708 |
| 5780 | 1277 | 6316 | 1013 | 2867 | 9938 | 3930 | 3203 | 5696 | 1769 |
| 1187 | 0951 | 5991 | 5245 | 5700 | 5564 | 7352 | 0891 | 6249 | 6568 |
| 4184 | 2179 | 4554 | 9083 | 2254 | 2435 | 2965 | 5154 | 1209 | 7069 |
| 2916 | 2972 | 9885 | 0275 | 0144 | 8043 | 8122 | 3213 | 7666 | 0230 |
| 5524 | 1341 | 9860 | 6565 | 6981 | 9842 | 0171 | 2284 | 2707 | 3008 |

**Продовження додатку 21.**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| 0146 | 5291 | 2354 | 5694 | 0377 | 5336 | 5440 | 9585 | 3415 | 2358 |
| 4920 | 2826 | 5238 | 5402 | 7937 | 1993 | 4332 | 2327 | 6875 | 5230 |
| 7978 | 1947 | 6380 | 3425 | 7267 | 7285 | 1130 | 7722 | 0164 | 8573 |
| 7453 | 0653 | 3645 | 7497 | 3969 | 8682 | 4191 | 2976 | 9361 | 9334 |
| 1473 | 6938 | 4899 | 5348 | 1641 | 3652 | 0852 | 5296 | 4538 | 4456 |
| 8162 | 8797 | 8000 | 4707 | 1880 | 9660 | 8446 | 1883 | 9768 | 0881 |
| 5645 | 4219 | 0807 | 3301 | 4279 | 4168 | 4305 | 9937 | 3120 | 5547 |
| 2042 | 1192 | 1175 | 8851 | 6432 | 4635 | 5757 | 6656 | 1660 | 5369 |
| 5470 | 7702 | 6958 | 9080 | 5925 | 8519 | 0127 | 9233 | 2452 | 7341 |
| 4504 | 1730 | 6005 | 1704 | 0345 | 3275 | 4738 | 4862 | 2556 | 8333 |

**Додаток 22.**

## Критичні значення критерію Кохрана

| $\nu_1\nu_2$ | Рівень значущості 0,05 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 36 | 144 | |
| 2 | 9985* | 9750 | 9392 | 9057 | 8772 | 8534 | 8332 | 8159 | 6602 | 5813 | 5000 |
| 3 | 9669 | 8709 | 7977 | 7457 | 7007 | 6771 | 6530 | 6333 | 4748 | 4031 | 3333 |
| 4 | 9065 | 7679 | 6841 | 6287 | 5895 | 5598 | 5365 | 5175 | 3720 | 3093 | 2500 |
| 5 | 8412 | 6838 | 5981 | 5441 | 5065 | 4783 | 4564 | 4564 | 4067 | 2513 | 2000 |
| 6 | 7803 | 6161 | 2521 | 4803 | 4447 | 4164 | 3980 | 3817 | 2612 | 2119 | 1667 |
| 7 | 7271 | 5612 | 4800 | 4307 | 3974 | 3726 | 3535 | 3384 | 2278 | 1833 | 1429 |
| 8 | 6798 | 5157 | 4377 | 3910 | 3595 | 3362 | 3185 | 3043 | 2022 | 1616 | 1250 |
| 9 | 6385 | 4775 | 4027 | 3584 | 3286 | 3067 | 2901 | 2768 | 1820 | 1446 | 1111 |
| 10 | 6020 | 4450 | 3733 | 3311 | 3029 | 2823 | 2666 | 2541 | 1655 | 1308 | 1000 |
| 12 | 5410 | 3924 | 3264 | 2880 | 2624 | 2439 | 2299 | 2187 | 1403 | 1100 | 0833 |
| 15 | 4709 | 3346 | 2758 | 2419 | 2195 | 2034 | 1911 | 1815 | 1144 | 0889 | 0667 |
| 20 | 3894 | 2705 | 2205 | 1921 | 1735 | 1602 | 1501 | 1422 | 0879 | 0675 | 0500 |
| 24 | 3434 | 2354 | 1 907 | 1656 | 1496 | 1374 | 1287 | 1216 | 0743 | 0567 | 0417 |
| 30 | 2929 | 1980 | 1593 | 1377 | 1237 | 1137 | 1061 | 1002 | 0604 | 0457 | 0333 |
| 40 | 2370 | 1576 | 1259 | 1082 | 0968 | 0887 | 0827 | 0780 | 0462 | 0347 | 0250 |
| 60 | 1737 | 1131 | 0895 | 0765 | 0682 | 0625 | 0583 | 0316 | 0234 | 0167 | 0167 |
| 120 | 0998 | 0632 | 0495 | 0419 | 0571 | 0337 | 0292 | 0292 | 0165 | 0120 | 0083 |
| ∞ | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 |

*Тут і далі значення дано після коми

**Додаток 23.**

**Розділ критерії Дурбіна-Уотсона для додатної автокореляції (Р-0,95)**

| *n* | ν-1 | | ν-2 | | ν-3 | | ν-4 | | ν-5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $d_1$ | $d_2$ | $d_1$ | $d_2$ | $d_1$ | $d_2$ | $d_1$ | $d_2$ | $d_1$ | $d_2$ |
| 15 | 1,08 | 1,36 | 0,95 | 1,54 | 0,82 | 1,75 | 0,69 | 1,97 | 0,56 | 2,21 |
| 16 | 1,10 | 1,37 | 0,98 | 1,54 | 0,86 | 1,73 | 0,74 | 1,93 | 0,62 | 2,15 |
| 17 | 1,13 | 1,38 | 1,02 | 1,54 | 0,90 | 1,71 | 0,78 | 1,90 | 0,67 | 2,10 |
| 18 | 1,16 | 1,39 | 1,05 | 1,53 | 0,96 | 1,69 | 0,82 | 1,87 | 0,71 | 2,06 |
| 19 | 1,18 | 1,40 | 1,08 | ,53 | 0,97 | 0,86 | 1,85 | 1,85 | 0,75 | 2,02 |
| 20 | 1,20 | 1,41 | 1,10 | ,54 | 1,00 | 1,68 | 0,90 | 1,83 | 0,79 | 1,99 |
| 21 | 1,22 | 1,42 | 1,13 | ,54 | 1,03 | 1,67 | 0,93 | 1,81 | 0,83 | 1,96 |
| 22 | 1,24 | 1,43 | 1,15 | ,54 | 1,05 | 1,66 | 0,96 | 1,80 | 0,86 | 1,94 |
| 23 | 1,26 | 1,44 | 1,17 | 1,54 | 1,08 | 1,60 | 0,99 | 1,79 | 0,90 | 1,92 |
| 24 | 1,27 | 1,45 | 1,19 | 1,55 | 1,10 | 1,66 | 1,01 | 1,78 | 0,93 | 1,90 |
| 25 | 1,29 | 1,45 | 1,21 | 1,55 | 1,12 | 1,66 | 1,04 | 1,77 | 0,95 | 1,99 |
| 26 | 1,30 | 1,46 | 1,22 | 1,55 | 1,14 | 1,65 | 1,06 | 1,76 | 0,98 | 1,88 |
| 27 | 1,32 | 1,47 | 1,24 | 1,56 | 1,16 | 1,65 | 1,08 | 1,76 | 1,01 | 1,86 |
| 28 | 1,33 | 1,48 | 1,26 | 1,56 | 1,18 | 1,65 | 1,10 | 1,75 | 1,03 | 1,85 |
| 29 | 1,34 | 1,48 | 1,27 | 1,56 | 1,20 | 1,65 | 1,12 | 1,74 | 1,05 | 1,84 |
| 30 | 1,35 | 1,49 | 1,28 | 1,57 | 1,21 | 1,65 | 1,14 | 1,74 | 1,047 | 1,73 |
| 40 | 1,44 | 1,54 | 1,39 | 1,60 | 1,34 | 1,66 | 1,29 | 1,72 | 1,23 | 1,79 |
| 50 | 1,50 | 1,59 | 1,46 | 1,63 | 1,42 | 1,67 | 1,38 | 1,72 | 1,34 | 1,77 |
| 60 | 1,55 | 1,62 | 1,51 | 1,65 | 1,48 | 1,69 | 1,44 | 1,73 | 1,41 | 1,77 |
| 70 | 1,58 | 1,64 | 1,55 | 1,67 | 1,52 | 1,70 | 1,49 | 1,74 | 1,46 | 1,77 |
| 80 | 1,61 | 1,66 | 1,59 | 1,69 | 1,56 | 1,72 | 1,53 | 1,74 | 1,51 | 1,77 |
| 90 | 1,63 | 1,68 | 1,61 | 1,70 | 1,59 | 1,73 | 1,57 | 1,75 | 1,54 | 1,78 |
| 100 | 1,65 | 1,69 | 1,63 | 1,72 | 1,61 | 1,74 | 1,59 | 1,76 | 1,57 | 1,78 |

**Додаток 24.**

## Значення деяких функцій, що часто зустрічаються в статичних розрахунках

| x | $x^2$ | $x^3$ | $x^4$ | $\sqrt{x}$ | $\sqrt{10x}$ | $\frac{1}{x}$ | lg x |
|---|---|---|---|---|---|---|---|
| 1,0 | 1,00 | 1 ,000 | 1,000 | 1,000 | 3,162 | 1 ,0000 | 0,0000 |
| 1,1 | 1,21 | 1,331 | 1,464 | 1,049 | 3,317 | 0,9091 | 0,0414 |
| 1,2 | 1,44 | 1,728 | 2,074 | 1,095 | 3,464 | 0,8333 | 0,0792 |
| 1,3 | 1,69 | 2,197 | 2,856 | 1,140 | 3,606 | 0,7692 | 0,1139 |
| 1,4 | 1,96 | 2,744 | 3,842 | 1,183 | 3,742 | 0,7143 | 0,1461 |
| 1,5 | 2,25 | 3,375 | 5,063 | 1,225 | 3,873 | 0,6667 | 0,1761 |
| 1,6 | 2,56 | 4,096 | 6,551 | 1 ,265 | 4,000 | 0,6250 | 0,2041 |
| 1,7 | 2,89 | 4,913 | 8,352 | 1,304 | 4,123 | 0,5882 | 0,2304 |
| 1,8 | 3,24 | 5,832 | 10,50 | 1,342 | 4,243 | 0,5556 | 0,2553 |
| 1,9 | 3,61 | 6,859 | 13,03 | 1,378 | 4,359 | 0,5263 | 0,2788 |
| 2,0 | 4,00 | 8,000 | 16,00 | 1,414 | 4,472 | 0,5000 | 0,3010 |
| 2,1 | 4,41 | 9,261 | 19,45 | 1,449 | 4,583 | 0,4762 | 0,3222 |
| 2,2 | 4,84 | 10,65 | 23,43 | 1,483 | 4,690 | 0,4546 | 0,3424 |
| 2,3 | 5,29 | 12,17 | 27,98 | 1,517 | 4,796 | 0,4348 | 0,3617 |
| 2,4 | 5,76 | 13,82 | 33,18 | 1,549 | 4,899 | 0,4167 | 0,3802 |
| 2,5 | 6,25 | 15,63 | 39,06 | 1,581 | 5,000 | 0,4000 | 0,3979 |
| 2,6 | 6,76 | 17,58 | 45,70 | 1,612 | 5,099 | 0,3846 | 0,4 1 50 |
| 2,7 | 7,29 | 19,68 | 53,14 | 1 ,643 | 5,196 | 0,3704 | 0,4314 |
| 2,8 | 7,84 | 21,95 | 61,47 | 1,673 | 5,292 | 0,3571 | 0,4472 |
| 2,9 | 8,41 | 24,39 | 70,73 | 1,703 | 5,385 | 0,3448 | 0,4624 |
| 3,0 | 9,00 | 27,00 | 81.00 | 1,732 | 5,477 | 0,3333 | 0,4771 |
| 3,1 | 9,61 | 29,79 | 92,35 | 1,761 | 5,568 | 0,3226 | 0,4914 |
| 3,2 | 10,24 | 32,77 | 104,9 | 1,789 | 5,657 | 0,3125 | 0,5051 |
| 3,3 | 10,89 | 35,94 | 118,6 | 1,817 | 5,745 | 0,3030 | 0,5185 |
| 3,4 | 11,56 | 39,30 | 133,6 | 1,844 | 5,831 | 0,2941 | 0,5315 |
| 3,5 | 12,25 | 42,88 | 150,1 | 1,871 | 5,916 | 0,2857 | 0,5441 |

**Продовження додатка 24.**

| x | $x^2$ | $x^3$ | $x^4$ | $\sqrt{x}$ | $\sqrt{10x}$ | $\frac{1}{x}$ | lg x |
|---|---|---|---|---|---|---|---|
| 3,6 | 12,96 | 46,66 | 168,0 | 1,897 | 6,000 | 0,2773 | 0,5563 |
| 3,7 | 13,69 | 50,65 | 187,4 | 1,924 | 6,083 | 0,2703 | 0,5682 |
| 3,8 | 14,44 | 54,87 | 208,5 | 1,949 | 6,164 | 0,2632 | 0,5798 |
| 3,9 | 15,21 | 59,32 | 231,3 | 1,975 | 6,245 | 0,2564 | 0,5911 |
| 4,0 | 16,00 | 64,00 | 256,0 | 2,000 | 6,325 | 0,2500 | 0,6021 |
| 4,1 | 16,81 | 68,92 | 282,6 | 2,025 | 6,403 | 0,2439 | 0,6128 |
| 4,2 | 17,64 | 74,09 | 311,2 | 2,049 | 6,481 | 0,2381 | 0,6232 |
| 4,3 | 18,49 | 79,51 | 341,9 | 2,074 | 6,557 | 0,2326 | 0,6335 |
| 4.4 | 19,36 | 85,18 | 374,8 | 2,098 | 6,633 | 0,2273 | 0,6435 |
| 4,5 | 20,25 | 91,13 | 410,1 | 2,121 | 6,708 | 0,2222 | 0,6532 |
| 4,6 | 21,16 | 97,34 | 447,7 | 2,145 | 6,782 | 0,2174 | 0,6628 |
| 4,7 | 22,09 | 103,8 | 488,0 | 2,168 | 6,856 | 0,2128 | 0,6721 |
| 4,3 | 23,04 | 110,6 | 530,8 | 2,191 | 6,928 | 0,2083 | 0,6812 |
| 4,9 | 24,01 | 117,6 | 576,5 | 2,214 | 7,000 | 0,2041 | 0,6902 |
| 5.0 | 25,00 | 125,0 | 625,0 | 2,236 | 7,071 | 0,2000 | 0,6990 |
| 5.1 | 26,01 | 132,7 | 676,5 | 2,258 | 7,141 | 0,1961 | 0,7076 |
| 5,2 | 27,04 | 140.6 | 731,2 | 2,280 | 7,211 | 0,1923 | 0,7160 |
| 5.3 | 28,09 | 148,9 | 789,0 | 2.302 | 7,280 | 0,1887 | 0,7243 |
| 5,4 | 29,16 | 157,6 | 850,3 | 2,324 | 7,348 | 0,1852 | 0,7324 |
| 5,5 | 30,25 | 166,4 | 915,1 | 2,345 | 7,416 | 0,1818 | 0,7404 |
| 5,6 | 31,36 | 175,6 | 983,4 | 2,366 | 7,483 | 0,1786 | 0,7482 |
| 5,7 | 32,49 | 185,2 | 1056 | 2,387 | 1,550 | 0,1754 | 0,7559 |
| 5,8 | 33,64 | 195,1 | 1132 | 2,408 | 7,616 | 0,1724 | 0,7634 |
| 5,9 | 34,81 | 205,4 | 1212 | 2,429 | 7,681 | 0,1695 | 0,7709 |
| 6,1 | 37,21 | 227.0 | 1385 | 2,470 | 7,810 | 0,1639 | 0,7853 |
| 6.2 | 38,44 | 238,3 | 1478 | 2,490 | 7,874 | 0,1613 | 0,7924 |
| 6,3 | 39,69 | 250,0 | 1575 | 2,510 | 7,937 | 0,1587 | 0.7993 |
| 6,4 | 40,96 | 262,1 | 1678 | 2,530 | 8,000 | 0,1563 | 0,8062 |
| 6,5 | 42,25 | 274,6 | 1785 | 2,550 | 8,062 | 0,1539 | 0,8129 |
| 6,6 | 43,56 | 287,5 | 1897 | 2,569 | 8,124 | 0,1515 | 0,8195 |
| 6.7 | 44,89 | 300,8 | 2015 | 2,588 | 8,185 | 0,1493 | 0,8261 |
| 6,8 | 46,24 | 314.4 | 2138 | 2,608 | 8,246 | 0,1471 | 0,8325 |

**Продовження додатка 24.**

| x | $x^2$ | $x^3$ | $x^4$ | $\sqrt{x}$ | $\sqrt{10x}$ | $\frac{1}{x}$ | lg x |
|---|---|---|---|---|---|---|---|
| 6,9 | 47,61 | 328,5 | 2267 | 2,627 | 8,307 | 0,1449 | 0,8188 |
| 7,0 | 49,00 | 343,0 | 2401 | 2,646 | 8,367 | 0,1429 | 0,8451 |
| 7,1 | 50,41 | 357,9 | 2541 | 2,665 | 8,426 | 0,1409 | 0,8513 |
| 7,2 | 51,84 | 373,2 | 2687 | 2,683 | 8,485 | 0,1389 | 0,8573 |
| 7,3 | 53,29 | 389,0 | 2840 | 2,702 | 8,544 | 0,1370 | 0,8633 |
| 7,4 | 54,76 | 405,2 | 2999 | 2,720 | 8,602 | 0,1351 | 0,8692 |
| 7,5 | 56,25 | 421,9 | 3164 | 2,739 | 8,660 | 0,1333 | 0,8751 |
| 7,6 | 57,76 | 439,0 | 3336 | 2,757 | 8,718 | 0,1316 | 0,8808 |
| 7,7 | 59,29 | 456,5 | 3515 | 2,775 | 8,775 | 0,1299 | 0,8865 |
| 7,8 | 60,84 | 474,6 | 3702 | 2,793 | 8,832 | 0,1282 | 0,8921 |
| 7,9 | 62,41 | 493,0 | 3895 | 2,811 | 8,888 | 0,1266 | 0,8976 |
| 8,0 | 64,00 | 512,0 | 4096 | 2,828 | 8,944 | 0,1250 | 0,9031 |
| 8,1 | 65,61 | 531,4 | 4305 | 2,846 | 9,000 | 0,1235 | 0,9085 |
| 8,2 | 67,24 | 551,4 | 4521 | 2,864 | 9,055 | 0,1220 | 0,9138 |
| 8,3 | 68,89 | 571,8 | 4746 | 2,881 | 9,110 | 0,1205 | 0,9191 |
| 8,4 | 70,56 | 592,7 | 4979 | 2,898 | 9,165 | 0,1191 | 0,9243 |
| 8,5 | 72,25 | 614,1 | 5220 | 2,915 | 9,220 | 0,1177 | 0,9294 |
| 8,6 | 73,96 | 636,1 | 5470 | 2,933 | 9,274 | 0,1163 | 0,9345 |
| 8,7 | 75,69 | 658,5 | 5729 | 2,950 | 9,327 | 0,1149 | 0,9395 |
| 8,8 | 77,44 | 681,5 | 5997 | 2,966 | 9,381 | 0,1136 | 0,9445 |
| 8,9 | 79,21 | 705,0 | 6274 | 2,983 | 9,434 | 0,1124 | 0,9494 |
| 9,0 | 81,00 | 729,0 | 6561 | 3,000 | 9,487 | 0.1111 | 0,9542 |
| 9,1 | 82,81 | 753,6 | 6857 | 3,017 | 9,539 | 0,1099 | 0,9590 |
| 9,2 | 84,64 | 778,7 | 7164 | 3,033 | 9,592 | 0,1087 | 0,9638 |
| 9,3 | 86,49 | 804,4 | 7481 | 3,050 | 9,644 | 0,1075 | 0,9685 |
| 9,4 | 88,36 | 830,6 | 7807 | 3,066 | 9,695 | 0,1064 | 0,9731 |
| 9,5 | 90,25 | 857,4 | 8145 | 3,082 | 9,747 | 0,1053 | 0,9777 |
| 9,6 | 92,16 | 884,7 | 8493 | 3,098 | 9,798 | 0,1042 | 0,9823 |
| 9,7 | 94,09 | 912,7 | 8853 | 3,114 | 9,849 | 0,1031 | 0,9868 |
| 9,8 | 96,04 | 941,2 | 9224 | 3,131 | 9,899 | 0,1020 | 0,9912 |
| 9,9 | 98,01 | 970,3 | 9606 | 3,146 | 9,950 | 0,1010 | 0,9956 |

# СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ


1. Адамов В.Е. Факторный индексный анализ. – М.: Статистика, 1987. – 199 с.
2. Аллен Р. Экономические индексы. – М.: Статистика, 1980. – 256 с.
3. Бек В.Л. Теорія статистики: Курс лекцій. Навчальний посібник. – Київ: ЦІЛ, 2003. – 288 с.
4. Вайну Я.Я. – Ф. Корреляция рядов динамики. – М.: Статистика, 1977. – 119 с.
5. Головач А.В., Черноскулова З.А. Экономико-статистический анализ потребления и спроса. – К.: «Техника», 1978. – 184 с.
6. Герчук Я.П. Графические методы в статистике. – М.: Статистика, 1972. – 78 с.
7. Єрина А.М., Пальян З.О. Теорія статистики: Практикум. – К.: Товариство «Знання», КОО, 1997. – 432 с.: іл.
8. Ефимова М.Р., Петрова Е.В., Румянцев В.Н. Общая теория статистики: Учебник. – М.: ИНФРА. – М. 1998. – 416 с.
9. Збірник задач з теорії статистики: Навчальний посібник / Є.І.Ткач, І.М.Шост, З.О.Насінник та ін. / За ред. Є.І.Ткача. – Тернопіль: Друк "Лідер", 2003. – 72 с.
10. Кендэл М. Временные ряды. – М.: Финансы и статистика, 1981. – 199 с.
11. Кендэл М. Ранговые корреляции. – М.: Статистика, 1975. – 241 с.
12. Кильдишев Г.С., Аболонцев Ю.И. Многомерные группировки. – М.: Статистика, 1978. – 160 с.
13. Кильдишев Г.С., Френкель А.А. Анализ временных рядов и прогнозирование. – М.: Статистика, 1973. – 103 с.
14. Ковтун Н.В., Столяров Г.С. Загальна теорія статистики: Курс лекцій. – К.: Четверта хвиля, 1996. – 144 с.: іл.
15. Копрен Г. Методы выборочного исследования. – М.: Статистика, 1976. – 440 с.
16. Кулинич О.І. Теорія статистики: Підручник. 2-е доп. и дооп. видання. – К. – д.: Державне Центрально-Українське видавництво, 1996. – 228 с.
17. Липкин М.И. Кривые распределения в экономических исследованиях. – М.: Статистика, 1972. – 114 с.
18. Лугінін О.Є., Білоусова С.В. Статистика: Підручник. – К.: Центр навчальної літератури, 2005. – 580 с.
19. Лугінін О.Є., Фомішин С.В. Статистика навчальної економіки та світового господарства: Навчальний посібник. – К.: Центр навчальної літератури, 2006. – 502 с.

20. Мармоза А.Т. Теорія статистики: Навчальний посібник. – К.: Ельга, Ніка-Центр, 2003. – 392 с.
21. Михель В.М. Динамические ряды: Учеб. пособие. – М.: МИНХ, 1969. – 95 с.
22. Михок Г., Урсяну В. Выборочный метод и статистическое оценивание. – М.: Статистика, 1982. – 245 с.
23. Монстеллер Ф., Тьюки Д.К. Анализ данных и регрессия. – М.: Финансы и статистика, 1982. – 239 с.
24. Общая теория статистики: Статистическая методология в изучении комплексной деятельности: Учебник / А.И.Харламов, О.Э.Башина, В.Т.Бабурин и др.: под ред. А.А.Спирина, О.Э.Башиной. – М.: Финансы и статистика, 1996. – 296 с.: ил.
25. Общая теория статистики: Учебник / Г.С.Кильдишев, В.Е.Овсиенко, П.М.Рабинович, Т.В.Рябушкин. – М.: Статистика, 1980. – 423 с.
26. Опря А.Т. Статистика (з програмованою формою контролю знань). Математична статистика. Теорія статистики. Навчальний посібник. – К.: Центр навчальної літератури, 2005. – 472 с.
27. Пасхавер И.С. Средние величины в статистике. – М.: Статистика, 1979. – 279 с.
28. Ряузов Н.Н. Общая теория статистики: Учебник. – М.: Финансы и статистика, 1984. – 343 с.
29. Сборник задач по общей теории статистики: Учеб. пособие / В.Е.Голованова, Ю.Г.Королев и др. – 2-е изд., перераб. и доп. – М.: Финансы и статистика, 1986. – 191 с.
30. Сигел, Ендрю. Практическая бизнес-статистика: Пер. с англ. – М.: Издательский дом «Вильямс», 2004. – 1056 с.: ил. – Парал. Тит. англ.
31. Статистика підприємництва: Підручник / П.Г.Вашків, П.І.Пастер, В.П.Сторожук, Є.І.Ткач: за ред. П.Г.Вашкова, В.П.Сторожука. – К.: “Слобожанщина”, 1999. – 600 с.
32. Статистика: Навч.-метод. посібник для самост. вивч. дисц. / А.М.Єріна, Р.М.Моторин, А.В.Головач та ін.: За заг. ред. А.М.Єріної, Р.М.Моторина. – К.: КНЕІ, 2001. – 448 с.
33. Статистика: Підручник / А.В.Головач, А.М.Єріна, О.В.Козирев та ін. / За ред. А.В.Головача. – К.: Вища школа, 1993. – 623 с.
34. Статистика: Підручник / С.С.Герасименко, А.В.Головач, А.М.Єріна та ін. / За наук. ред. д-ра екон. наук С.С.Герасименка. – 2-е вид., перероб. и доп. – К.: КНЕОІ, 2000. – 467 с.

35. Статистика: теоретичні засади і прикладні аспекти. Навчальний посібник. / Р.В.Фещур, А.Ф.Барвінський, В.П.Кічор та ін.; За наук. ред. Р.В.Фещура. – 2-е вид. оновлене і доповнене. – Львів: «Інтелект-Захід», 2003. – 576 с.
36. Теорія статистики: Навчальний посібник / Вашків П.Г., Пастер П.І., Сторожук В.П., Ткач Є.І. – К.: Либідь, 2001. – 320 с.
37. Ткач Є.І. Загальна теорія статистики: Підручник. – Тернопіль: Лідер, 2004. – 388 с.
38. Уманець Т.В. Загальна теорія статистики: Навч. посіб. – К.: Знання, 2006. – 239 с.
39. Уманець Т.В., Пігарєв Ю.Б. Статистика: Навч. посіб. – К.: Вікар, 2003. – 623 с.
40. Урланис Б.Ц. Общая теория статистики: Учебник. – М.: Статистика, 1973. – 439 с.
41. Четыркин Е.М. Статистические методы прогнозирования. – М.: Статистика, 1977. – 200 с.
42. Штангрет А.М., Копилюк О.І. Статистика: Навчальний посібник. – К.: Центр навчальної літератури, 2005. – 232 с.
43. Юл. Дж., Кендэл М. Дж. Теория статистики. – М.: Госполитиздат, 1960. – 779 с.

# ЗМІСТ

НАВЧАЛЬНЕ ВИДАННЯ

Євген Іванович ТКАЧ
Володимир Петрович СТОРОЖУК

# ЗАГАЛЬНА ТЕОРІЯ СТАТИСТИКИ

*3-тє видання*

ПІДРУЧНИК

Керівник видавничих проектів – Б. А. Сладкевич
Друкується в авторській редакції
Дизайн обкладинки – Б. В. Борисов

Підписано до друку 19.01.2009. Формат 60х84 1/16.
Друк офсетний. Гарнітура PetersburgC.
Умовн. друк. арк. 24,8.
Наклад 1000 прим.

Видавництво "Центр учбової літератури"
вул. Електриків, 23
м. Київ, 04176
тел./факс 425-01-34, тел. 451-65-95, 425-04-47, 425-20-63
8-800-501-68-00 (безкоштовно в межах України)
e-mail: office@uabook.com
сайт: WWW.CUL.COM.UA

Свідоцтво ДК №2458 від 30.03.2006