ОПОРНІ КОНСПЕКТИ
з «Теорії ймовірностей та математичної статистики»

Лекція 1.

Тема: ***Основні поняття Теорії ймовірностей.***

1. Предмет «Теорії ймовірностей».
2. Класичне означення ймовірності. Її властивості.
3. Основні формули комбінаторики.
4. Статистичне означення ймовірності. Відносна частота та її властивості
5. Геометрична ймовірність.

Мета. *Формувати поняття стохастичного експерименту, події, поняття відносної частоти та статистичної ймовірності;ознайомити з класичним означенням ймовірності, її властивостями та алгоритмом обчислення.*

1.

В житті та практичній діяльності людей часто доводиться мати справу з явищами, перебіг яких неможливо передбачити заздалегідь, вони залежать від багатьох умов. Про такі явища кажуть, що вони є випадковими.

Наприклад:

- попадання (непопадання) спортсмена в кошик за одного чи кількох киданнях м'яча;
- виграш (програш) в лотереї;
- кількість пасажирів в автобусі в даний момент часу тощо.

Взагалі, людська діяльність —це неперервний процес прийняття рішень в обставинах невизначеності чи випадковості. Яку встановити ціну, щоб продати товар і одержати прибуток? Яким повинен бути внесок при страхуванні, щоб страхова компанія не мала збитків? З такими та подібними їм запитаннями люди постійно стикаються в повсякденному житті. Тому варто вміти працювати з випадковими явищами і використовувати їх у житті, наукових дослідженнях тощо.

Наукою, що займається математичним аналізом випадкових явищ, зокрема, випадкових подій, є ***«Теорія ймовірностей»,*** її предмет — це вивчення ***закономірностей масових випадкових явищ***.

До основних понять відносять поняття стохастичного експерименту та події.

Забезпечення певного комплексу умов називають *випробуванням* або *дослідом*, а можливий результат випробування — *подією*. Наприклад, підкидання монети є випробуванням, а випадання «герба» або «номіналу» — подіями. Події позначатимемо великими латинськими літерами: A, B, C.

Подію називають *випадковою*, якщо вона може відбутися або не відбутися в даному випробуванні.

Достовірною називають подію, яка обов'язково відбудеться в даному випробуванні.

Неможливою називають подію, яка точно не відбудеться в даному випробуванні.

Зауважимо, що будь-яка подія пов'язана з певним випробуванням.

Дві події називають *сумісними*, якщо поява однієї з них не виключає появи іншої в одному й тому самому випробуванні.

Дві події називають *несумісними*, якщо вони не можуть відбутися одночасно в одному й тому самому випробуванні.

Попарно несумісні випадкові події A_1, A_2, ..., A_n утворюють *повну групу подій*, якщо внаслідок випробування одна з них обов'язково відбудеться. Наприклад, події «виграш», «програш» і «нічия» (для певного гравця) утворюють повну групу подій у випробуванні — грі в шахи двох суперників.

Елементарними подіями (*наслідками*) у певному випробуванні називають усі можливі результати цього випробування, які не можна розкласти на простіші. Множину всіх можливих елементарних подій називають *простором елементарних подій*, який позначають Ω. Наприклад, при підкиданні грального кубика простір елементарних подій утворюють події A_i = {випаде i очок}, $i = 1, 2, 3, 4, 5, 6$.

Елементарні події, при появі яких відбувається певна подія, називають *сприятливими* для цієї події. Наприклад, при підкиданні грального кубика для події C = {випаде непарне число очок} сприятливими є елементарні події A_1, A_3, A_5.

Кожну подію можна розглядати як деяку підмножину простору елементарних подій у даному випробуванні. Подію $\overline{A}$ називають *протилежною* до події A в даному випробуванні, якщо ці дві події несумісні й утворюють повну групу подій.

Приклад 1. У ящику містяться кульки білого та чорного кольору. Навмання з нього виймають одну кульку. Подія A = {вийнято кульку білого кольору}, подія B = {вийнято кульку чорного кольору}. Сумісні чи несумісні ці події?

Розв'язання. Ці події несумісні, тому що поява події A виключає можливість появи події B, і навпаки. У даному випробуванні події A і B є протилежними:

$$A = \overline{B},\ B = \overline{A}. \bullet$$

2.

Імовірність події A дорівнює відношенню кількості елементарних наслідків, сприятливих цій події, до кількості всіх рівноможливих елементарних наслідків у даному випробуванні.

Імовірність події A позначають $P(A)$, тому за означенням

$$P(A) = \frac{m}{n}, \tag{1.1}$$

де m — кількість елементарних наслідків, сприятливих події A; n — кількість усіх елементарних наслідків у даному випробуванні. З класичного означення ймовірності випливають такі властивості:

. *1.* $P(A) = 0$, якщо $A = \varnothing$ — неможлива подія.

2. $P(A) = 1$, якщо $A = \Omega$ — достовірна подія.

3. Якщо .A — випадкова подія, тоді $0 \le P(A) \le 1$.

Приклад 2. Знайти ймовірність того, що вибране випадковим чином двозначне число ділиться на: **а)** 3; **б)** 5.

Розв'язання. У даному разі випробування полягає в тому, що вибирається випадковим чином двозначне число. Наслідком такого випробування є одне з чисел від 10 до 99. Оскільки таких чисел 90, то $n = 90$.

а) Нехай подія $A = \{$вибране двозначне число ділиться на 3$\}$. Оскільки кожне третє з 90 двозначних чисел ділиться на 3, то сприятливими для події A є 30 наслідків, тобто $m = 30$. Тоді за формулою (1.1) імовірність події A

$$P(A) = \frac{30}{90} = \frac{1}{3}.$$

б) Нехай подія $B = \{$вибране двозначне число ділиться на 5$\}$. Загальна кількість наслідків випробування, як і в попередньому випадку, $n = 90$. Визначимо кількість чисел, які діляться на 5. Очевидно, що таких чисел буде $m = 18$ (кожне п'яте число ділиться на 5). Отже,

$$P(A) = \frac{18}{90} = \frac{1}{3}. \bullet$$

3.

При обчисленні ймовірностей подій досить часто потрібно підраховувати кількість елементарних подій (сприятливих деякій події або всіх можливих подій). Здебільшого це зумовлює великі труднощі, подолати які допомагає комбінаторика, що вивчає способи підрахунку кількості розміщень, перестановок, комбінацій.

Перш ніж представити деталі, нагадаємо, що вираз $n!$ читається «ен-факторіал» і означає добуток усіх натуральних чисел до n:

$$n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n,$$

причому вважають, що $0! = 1$.

Розміщеннями із l елементів по k називають множини із k елементів, вибраних із l елементів, які можуть розрізнятися між собою як складом елементів, так і їх порядком. Наприклад, розміщеннями із трьох елементів по два будуть такі множини: $\{1; 2\}$, $\{1; 3\}$, $\{2; 1\}$, $\{2; 3\}$, $\{3; 1\}$, $\{3; 2\}$. Кількість усіх розміщень із l елементів по k визначають за формулою

$$A_l^k = l \cdot (l-1) \cdot (l-2) \cdot \ldots \cdot (l-k+1) = \frac{l!}{(l-k)!}. \qquad (1.3)$$

Перестановками із l елементів називають множини із l елементів, що відрізняються лише їх порядком. Наприклад, перестановками із трьох елементів будуть такі множини: {1; 2; 3}, {1; 3; 2}, {2; 1; 3}, {2; 3; 1}, {3; 1; 2}, {3; 2; 1}. Кількість усіх перестановок із l елементів визначають так:

$$P_l = 1 \cdot 2 \cdot \ldots \cdot l = l!. \tag{1.4}$$

Комбінаціями із l елементів по k називають множини із k елементів, вибраних із l елементів, які розрізняються між собою тільки складом елементів. Наприклад, комбінаціями із трьох елементів по два будуть такі множини: {1; 2}, {1; 3}, {2; 3}. Кількість усіх комбінацій із l елементів по k визначають за формулою

$$C_l^k = \frac{(l-k+1)\cdot(l-k+2)\cdot\ldots\cdot l}{k!} = \frac{l!}{k!\cdot(l-k)!}. \tag{1.5}$$

Між переліченими поняттями існують такі співвідношення:

$$P_l = A_l^l = l!; \quad C_l^k = \frac{A_l^k}{P_k}; \quad C_l^k = C_l^{l-k};$$

$$C_l^0 = C_l^l = 1; \quad C_l^1 = l; \quad C_l^0 + C_l^1 + C_l^2 + \ldots + C_l^l = 2^l.$$

Приклад 3. Шістнадцять варіантів контрольної роботи написані на окремих картках і розподіляються випадковим чином серед 14 студентів, які сидять в одному ряду. Кожний студент отримує одну картку. Знайти ймовірність того, що: **а)** варіанти 1 і 2 не будуть використані; **б)** варіанти 1 і 2 видадуть студентам, які сидять поруч.

Розв’язання. Маємо випробування розподілу 16 білетів серед 14 студентів. У цьому разі події відрізняються одна від одної не лише номерами варіантів, що розподіляються серед студентів, а й порядком розподілу. Тому такі сполучення називають розміщеннями, а кількість таких розміщень визначається за формулою (1.3):

$$n = A_{14}^{16} = \frac{16!}{2!}.$$

а) Позначимо через A подію, яка полягає в тому, що варіанти 1 і 2 залишаться нерозподіленими. Тоді інші 14 білетів розподіляться серед 14 студентів. Такі сполучення називають перестановками, а їх кількість визначається за формулою (1.4):

$$m = P_{14} = 14!$$

Отже, застосувавши класичну формулу ймовірності (1), матимемо:

$$P(A) = \frac{m}{n} = \frac{14!\cdot 2!}{16!} = \frac{1}{15\cdot 8} \approx 0{,}008.$$

б) Нехай подія B полягає в тому, що варіанти 1 і 2 видані студентам, які сидять поруч. У ряду із 14 місць є 13 пар сусідніх місць, причому в кожній парі варіанти можуть розподілятися двома способами:

$$m_1 = 13 \cdot 2 = 26.$$

Інші 14 варіантів білетів розподіляються між 12 студентами

$$m_2 = A_{14}^{12} = \frac{14!}{2!}$$

способами. Тому події В сприяють

$$m = m_1 \cdot m_2 = 26 \cdot \frac{14!}{2!}$$

наслідків.

Отже, імовірність події B

$$P(B) = \frac{26 \cdot 14! \cdot 2!}{2! \cdot 16!} = \frac{13}{120} \approx 0{,}108. \bullet$$

Приклад 4. У податковій адміністрації зареєстровано 6 приватних і 4 державних підприємства. Знайти ймовірність того, що серед навмання вибраних трьох підприємств приватними будуть: **а)** три; **б)** два; **в)** не більше одного.

Розв'язання. Оскільки, не ставиться умова впорядкованості підмножини із вибраних трьох підприємств, то потрібно використати комбінації.

Тоді $n = C_{10}^3 = \frac{10!}{3! \cdot 7!} = \frac{7! \cdot 8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3 \cdot 7!} = 120$. Для пункту **а)** одержимо: $m = C_6^3 = 20$. $P(A) = \frac{20}{120} = \frac{1}{6}$.

Пункт **б)** відрізняється від попереднього тим, що вибрана тут група із трьох підприємств включає два приватні й одне державне. За правилом множення одержимо: $m = C_6^2 \cdot C_4^1 = 15 \cdot 4 = 60$. $P(B) = \frac{60}{120} = \frac{1}{2}$.

Для виконання пункту **в)** розкриємо зміст словосполучення «не більше одного». Воно означає «одне або жодного».

Тоді $m = C_6^1 \cdot C_4^2 + C_4^3 = 6 \cdot 6 + 4 = 40$. $P(C) = \frac{40}{120} = \frac{1}{3}$.

4.

Оскільки класичне означення ймовірності передбачає, що всі елементарні наслідки випробування рівноможливі, що важко обґрунтувати, то розглядають ще й *статистичне означення ймовірності*.

Відносною частотою події A називають відношення кількості випробувань, у яких подія A відбулася, до кількості всіх проведених випробувань. Відносну частоту події A позначають $W(A)$. Тоді

$$W(A) = \frac{M}{N},$$

де m — кількість випробувань, у яких відбулася подія A; n — кількість усіх проведених випробувань.

Порівняння означень ймовірності та відносної частоти дає можливість зробити висновок, що відносна частота є апостеріорною величиною, тобто величиною, яку обчислюють після випробування, на відміну від імовірності, яку обчислюють до випробування(апріорна величина).

Відносна частота володіє властивістю *стійкост*, яка полягає в тому, що в різних серіях випробувань відносна частота змінюється мало, коливаючись навколо деякого постійного числа.

Число, до якого прямує значення частоти події A при великій кількості випробувань, називають *ймовірністю події A*:

$$P(A)=\lim_{n\to\infty}W(A).$$

Крім того, для відносної частоти також можна записати подвійну нерівність

$$0\le W(A)\le 1.$$

5.

Один із недоліків класичного означення — його неможливо використати для випробувань з нескінченною кількістю наслідків випробувань.

Нехай множина всіх елементарних наслідків випробування нескінченна і утворює деяку множину Ω, усі елементарні наслідки рівноможливі, причому події *A* сприяють ті елементарні події, які утворюють множину $A\subseteq\Omega$. Тоді ймовірність події *A* дорівнює відношенню міри множини *A* до міри множини Ω, тобто

$$P(A)=\frac{m(A)}{m(\Omega)}. \tag{1.6}$$

Мірою множини на прямій, площині, у просторі є відповідно довжина, площа, об’єм геометричної фігури, яку утворює ця множина.

Рекомендована література:[1], С. 5–16; [2], С. 17–30; [5], С. 4–24

Лекція 2.

Тема: *Теореми множення і додаванняімовірностей та їх наслідки.*

1. ***Умовна імовірність. Теорема множення імовірностей.***
2. ***Теореми додавання імовірностей.***
3. ***Основна властивість подій, які утворюють повну групу.***
4. ***Алгоритм розв'язування задач з використанням теорем додавання та множення імовірностей.***
5. ***Імовірність появи хоча б однієї події.***
6. ***Імовірність відбуття тільки однієї події.***
7. ***Формула повної імовірності.***
8. ***Формули Байєса.***
9. ***Алгоритм розв'язування задач з використанням формул повної імовірності та Байєса.***

1.

Нехай випробування, в результаті якого може відбутися випадкова подія В, доповнюється умовою про відбуття випадкової події А. Тоді **імовірність події В, знайдена при умові, що подія *A* відбулася, називається умовною імовірністю події** *B* і позначається $P_A(B)$ або $P(B|A)$.

Дві події називаються **незалежними**, якщо імовірність однієї з них не змінюється від того, відбулася чи ні інша. Аналітичним критерієм незалежності подій *A* та *B* є рівність

$$P_A(B) = P_{\overline{A}}(B). \tag{2.1}$$

У випадку виконання (2.1) можна вважати, що

$$P_A(B) = P_{\overline{A}}(B) = P(B), \tag{2.2}$$

де остання імовірність називається **безумовною**.

Дві події називаються **залежними**, якщо імовірність однієї з них змінюється внаслідок відбуття іншої. Тобто, якщо події *A* та *B* залежні, то рівність (2.2) порушується:

$$P_A(B) \neq P_{\overline{A}}(B). \tag{2.3}$$

Теорема множення імовірностей

Імовірність добутку двох сумісних подій дорівнює добутку імовірностей однієї з них на умовну імовірність іншої події, обчислену в припущенні, що перша подія відбулася:

$$P(AB) = P(A)P_A(B)\big(= P(B)P_B(A)\big). \tag{2.4}$$

Зауваження. Продовження рівності в дужках вказує на «рівноправність» подій *A* та *B* з врахуванням комутативності дії множення ($AB = BA$).

Наслідок 1. Імовірність добутку *k* подій дорівнює добутку імовірності однієї з них на умовні імовірності всіх решти, причому імовірність кожної наступної події знаходиться в припущенні, що всі попередні події вже

відбулися:

$$P(A_1A_2A_3...A_k)=P(A_1)P_{A_1}(A_2)P_{A_1A_2}(A_3)...P_{A_1A_2\cdots A_{k-1}}(A_k). \qquad (2.5)$$

Наслідок 2. Якщо події А та В незалежні, тоді

$$\boldsymbol{P(AB)=P(A)\,P(B).} \qquad (2.6)$$

Для узагальнення наслідка 2 на випадок довільного числа подій розглянемо такі означення.

Декілька подій називаються попарно незалежними, якщо **кожні** дві з них незалежні. Наприклад, події A_1, A_2, A_3 попарно незалежні, якщо незалежні події A_1 і A_2, A_1 і A_3, A_2 і A_3.

Декілька подій називаються незалежними в сукупності (або просто **незалежними**), якщо вони попарно незалежні, а також є незалежними кожна з них і всі можливі добутки інших. Наприклад, якщо події A, B, C незалежні в сукупності, то незалежні події A і B, A і C, B і C, A і BC, B і AC, C і AB.

Виявляється, що **попарна незалежність декількох подій ще не гарантує їх незалежність в сукупності**.

Наслідок 3. Імовірність добутку *k* подій, незалежних в сукупності, дорівнює добутку імовірностей цих подій:

$$\boldsymbol{P(A_1A_2\ldots A_k)=P(A_1)\,P(A_2)\ldots P(A_k).} \qquad (2.7)$$

Приклад 1. В шухляді є 11 карток розрізної абетки: 4 картки з літерою «и», 5 — з літерою «р» і 2 — «м». Навмання витягуються три картки і розкладаються на стіл зліва направо. Знайти імовірність того, що в результаті отримається слово «мир».

○ Нехай A — отримання слова «мир», B_1 — перша картка має літеру «м», B_2 — друга картка має літеру «и», B_3 — третя — «р». Подія A відбудеться, якщо відбудеться **і** подія B_1, **і** подія B_2, **і** подія B_3, тобто $A=B_1B_2B_3$. Тоді за теоремою множення імовірностей для залежних подій (рівність (2.5) для $k=3$)

$$P(A)=P(B_1B_2B_3)=P(B_1)P_{B_1}(B_2)P_{B_1B_2}(B_3)=\frac{2}{11}\cdot\frac{4}{10}\cdot\frac{5}{9}=\frac{4}{99}.$$

Зауваження. Цю задачу можна було б розв'язати і за класичним означенням імовірності. Але наслідки випробування (витягування трьох карток) вже не є розміщеннями, оскільки серед 11 вихідних елементів є однакові (пригадайте означення розміщення!). Ці наслідки не можна також назвати і комбінаціями по цій самій причині (крім того, розташування літер суттєве). Тому для обчислення m та n треба використовувати більш складні формули комбінаторики.

Розглянемо ще одну задачу, до розв'язування якої стосується попереднє зауваження з деякими доповненнями.

Приклад 2. На чотирнадцяти окремих картках написано по одній літері: 5 карток з літерою «А», 4 — з літерою «І», 3 — з літерою «Р» і 2 – «Т». Навмання витягуються шість карток і розкладаються зліва направо. Яка імовірність того, що отримається слово «АРАРАТ»?

○ Позначимо випадкові події: B — отримання слова «АРАРАТ», A_1 — перша картка має літеру «А», A_2 — друга картка має літеру «Р», A_3 — третя картка має літеру «А», A_4 — четверта картка має літеру «Р», A_5 — п'ята картка має літеру «А», A_6 — шоста картка має літеру «Т»,

Подія B відбудеться, якщо відбудуться всі події $A_1, A_2, ..., A_6$, тобто $B = A_1A_2A_3A_4A_5A_6$. Використавши теорему множення імовірностей для залежних подій, отримаємо:

$$P(B) = P(A_1A_2A_3A_4A_5A_6) = P(A_1) \cdot P_{A_1}(A_2) \cdot P_{A_1A_2}(A_3) \cdot P_{A_1A_2A_3}(A_4) \cdot$$

$$P_{A_1A_2A_3A_4}(A_5) \cdot P_{A_1A_2A_3A_4A_5}(A_6) = \frac{5}{14} \cdot \frac{3}{13} \cdot \frac{4}{12} \cdot \frac{2}{11} \cdot \frac{3}{10} \cdot \frac{2}{9} = \frac{1}{3003} \approx 0,00033.$$

2.

Теорема 1. Імовірність суми двох несумісних подій дорівнює сумі імовірностей цих подій:

$$P(A + B) = P(A) + P(B). \qquad (2.8)$$

Наслідок. **Імовірність суми *k* попарно несумісних подій дорівнює сумі імовірностей цих подій:**

$$\boldsymbol{P(A_1 + A_2 + ... + A_k) = P(A_1) + P(A_2) + ... + P(A_k).} \qquad (2.9)$$

Теорема 2. Імовірність суми двох сумісних подій дорівнює сумі імовірностей цих подій без імовірності їх сумісної появи:

$$\boldsymbol{P(A + B) = P(A) + P(B) - P(AB).} \qquad (2.10)$$

Приклад 1. В кейсі є 5 акцій першого виду, 6 — другого і 3 — третього. Знайти імовірність того, що три навмання взяті акції виявляться одного і того ж виду.

○ Позначимо: A — три навмання взяті акції є одного виду, B_i — три взяті акції i-го виду (i = 1, 2, 3). Подія A відбувається тоді, коли відбуваються або подія B_1, або B_2, або B_3. Інших можливостей для появи A немає. Тому мають місце такі рівності: $A = B_1 + B_2 + B_3$, $P(A) = P(B_1 + B_2 + B_3)$. Події B_1, B_2, B_3 попарно несумісні і згідно з наслідком теореми 1 (рівність (2.9))

$$P(A) = P(B_1) + P(B_2) + P(B_3).$$

За класичним означенням імовірності

$$P(B_1) = \frac{C_5^3}{C_{14}^3} = \frac{5}{182}, \quad P(B_2) = \frac{C_6^3}{C_{14}^3} = \frac{5}{91}, \quad P(B_3) = \frac{C_3^3}{C_{14}^3} = \frac{1}{364}.$$

Остаточно

$$P(A) = \frac{5}{182} + \frac{5}{91} + \frac{1}{364} = \frac{31}{364}.$$

Покажемо, як можна знайти імовірності подій B_1, B_2 та B_3, використовуючи теорему множення імовірностей. З цією метою позначимо: $B_j^{(i)}$ — j-та відібрана акція (j = 1, 2, 3) є акцією i-го виду (i = 1, 2, 3). Тоді, зокрема, B_1 відбудеться, якщо всі акції (**і** перша, **і** друга, **і** третя) є акціями першого виду, тобто коли відбуваються всі події $B_1^{(1)}, B_2^{(1)}, B_3^{(1)}$.

Отже, $B_1 = B_1^{(1)} B_2^{(1)} B_3^{(1)}$, $P(B_1) = P\left(B_1^{(1)} B_2^{(1)} B_3^{(1)}\right)$ і згідно з рівністю (2.5) для випадку $k = 3$

$$P(B_1) = P\left(B_1^{(1)}\right) P_{B_1^{(1)}}\left(B_2^{(1)}\right) P_{B_1^{(1)} B_2^{(1)}}\left(B_3^{(1)}\right) = \frac{5}{14} \cdot \frac{4}{13} \cdot \frac{3}{12} = \frac{5}{182}.$$

Відмітимо, що вибір рівності (2.5) на противагу рівності (2.7) зумовлений залежністю $B_1^{(1)}, B_2^{(1)}, B_3^{(1)}$ (доведіть!).

Аналогічно

$$P(B_2) = P\left(B_1^{(2)}\right) P_{B_1^{(2)}}\left(B_2^{(2)}\right) P_{B_1^{(2)} B_2^{(2)}}\left(B_3^{(2)}\right) = \frac{6}{14} \cdot \frac{5}{13} \cdot \frac{4}{12} = \frac{5}{91},$$

$$P(B_3) = P\left(B_1^{(3)}\right) P_{B_1^{(3)}}\left(B_2^{(3)}\right) P_{B_1^{(3)} B_2^{(3)}}\left(B_3^{(3)}\right) = \frac{3}{14} \cdot \frac{2}{13} \cdot \frac{1}{12} = \frac{1}{364}.$$

Отже, другий метод знаходження імовірностей подій B_1, B_2 та B_3 дає ті ж самі результати.

Приклад 2. Підприємство планує здійснювати поставки двох видів виробів. Імовірність зриву поставок для першого виду виробів складає 0,05, а для другого 0,08. Згідно із проектом контракту при порушенні термінів поставок хоча б одного виду продукції до виробника застосовуються штрафні санкції, які приводять до нерентабельності виробництва обох видів виробів. Знайти імовірність нерентабельності виробництва цих виробів.

○ Позначимо: C — нерентабельність виробництва обох видів продукції, A та B — зрив поставки виробів першого та другого видів відповідно. Згідно з умовою задачі подія C відбувається, якщо відбувається **або** подія A, **або** подія B, **або** події A та B разом, тобто хоча б одна з них. Тому $C = A + B$, де A та B сумісні події. Згідно з теоремою додавання для сумісних подій

$$P(C) = P(A + B) = P(A) + P(B) - P(AB).$$

За умовою $P(A) = 0{,}05$, $P(B) = 0{,}08$ і можна вважати, що події A та B незалежні, тобто $P(AB) = P(A)\,P(B)$.

Остаточно $P(C) = 0{,}05 + 0{,}08 - 0{,}05 \cdot 0{,}08 = 0{,}126$.

Другий метод розв'язування. Подію C можна представити через більш прості події A та B ще й таким чином:

$$C = A\overline{B} + \overline{A}B + AB.$$

Доданки-події справа є попарно несумісними випадковими подіями. Справді, припустимо, що події $A\overline{B}$ та $\overline{A}B$ є сумісними, тобто можуть відбутися у випробуванні. Тоді отримується суперечність: подія A, зокрема, в одному випробуванні і відбувається, і не відбувається. Це протиріччя вказує на хибність припущення. Переконайтеся в попарній несумісності цих подій, використавши діаграми В'єнна. Використавши рівності (2.9) та (2.6) і незалежність подій A та $\overline{B}$, $\overline{A}$ та B, отримаємо:

$$P(C)=P(A\overline{B}+\overline{A}B+AB)=P(A\overline{B})+P(\overline{A}B)+P(AB)=$$
$$=P(A)P(\overline{B})+P(\overline{A})P(B)+P(A)P(B)=$$
$$=0,05\cdot 0,92+0,95\cdot 0,08+0,05\cdot 0,08=0,046+0,076+0,004=0,126,$$
де згідно із рівністю (2.11*)
$$P(\overline{A})=1-P(A)=1-0,05=0,95, \qquad P(\overline{B})=1-P(B)=1-0,08=0,92.$$
3.

Теорема. Сума імовірностей подій, які утворюють повну групу, дорівнює одиниці:
$$\boldsymbol{P(A_1)+P(A_2)+\ldots P(A_n)=1.} \tag{2.11}$$
Для $n=2$ події A_1, A_2 є протилежними, тому рівність (2.11) набуває такого виду:
$$P(A)+P(\overline{A})=1 \quad \text{або} \qquad p+q=1, \tag{2.11*}$$
де $p=P(A)$, $q=P(\overline{A})$.

В попередньому пункті теми ми розглянули два методи розв’язання задачі про підприємство. Проте на основі даного наслідку можна запропонувати ще один спосіб.

Приклад 1. Підприємство планує здійснювати поставки двох видів виробів. Імовірність зриву поставок для першого виду виробів складає 0,05, а для другого 0,08. Згідно із проектом контракту при порушенні термінів поставок хоча б одного виду продукції до виробника застосовуються штрафні санкції, які приводять до нерентабельності виробництва обох видів виробів. Знайти імовірність нерентабельності виробництва цих виробів.

Третій метод. Протилежною до події C є подія $\overline{C}$, яка полягає в тому, що внаслідок здійснення поставок продукції виробництво обох видів виробів є рентабельним. Подія $\overline{C}$ відбувається тоді, коли **і** першого виду вироби вчасно поставляються, **і** другого, тобто відбуваються події і $\overline{A}$, і $\overline{B}$. Тому $\overline{C}=\overline{A}\overline{B}$, $P(\overline{C})=P(\overline{A})P(\overline{B})=0,95\cdot 0,92=0,874$. Але згідно з (2.11*) $P(C)+P(\overline{C})=1$, звідки $P(C)=1-P(\overline{C})=1-0,874=0,126$.

Висновки. Перевага третього методу, зокрема, полягає в тому, що він дозволяє розв’язати задачу для довільного скінченного числа видів продукції. Що ж до отриманої відповіді, то знайдену імовірність слід інтерпретувати таким чином: в середньому в 126 випадках з кожної тисячі (12,6%) очікується нерентабельність виробництва обох видів продукції. Оскільки така імовірність достатньо велика, то керівництву підприємства потрібно подбати про зменшення імовірностей зривів поставок або (і) зменшення штрафних санкцій.

Приклад 3. У зв’язці шість різних ключів, з яких тільки одним можна відкрити замок. Навмання вибирається ключ і робиться спроба відкрити ним замок. Ключ, що не підійшов, більше не використовується. Знайти імовірність того, що для відкривання буде використано не більше трьох ключів.

○ Позначимо: A_k ($k = 1, 2, 3$) — замок буде відкрито k-тим за порядку відбору ключем, B — замок відкривається після використання не більше трьох ключів. Подія B відбудеться, якщо до замка підійде **або** перший ключ (відбувається A_1), **або** другий (при цьому перший ключ не підійшов — відбувається подія $\overline{A}_1 A_2$), **або** третій (перший і другий ключі не підійшли — відбувається подія $\overline{A}_1 \overline{A}_2 A_3$). Тобто вираз B через простіші події A_1, A_2, A_3 має такий вид:

$$B = A_1 + \overline{A}_1 A_2 + \overline{A}_1 \overline{A}_2 A_3.$$

Для знаходження $P(B)$ потрібно використати рівність (2.9), оскільки доданки є попарно несумісними подіями, а потім теорему множення імовірностей для залежних подій (обчисливши $P_{\overline{A}_1}(A_2)$ і $P_{A_1}(A_2)$, переконайтеся у тому, що події $\overline{A}_1$ і A_2 є залежними):

$$P(B) = P(A_1 + \overline{A}_1 A_2 + \overline{A}_1 \overline{A}_2 A_3) = P(A_1) + P(\overline{A}_1 A_2) + P(\overline{A}_1 \overline{A}_2 A_3) =$$

$$= P(A_1) + P(\overline{A}_1) P_{\overline{A}_1}(A_2) + P(\overline{A}_1) P_{\overline{A}_1}(\overline{A}_2) P_{\overline{A}_1 \overline{A}_2}(A_3) = \frac{1}{6} + \frac{5}{6}\cdot\frac{1}{5} + \frac{5}{6}\cdot\frac{4}{5}\cdot\frac{1}{4} = 0{,}5.$$

4.

Алгоритм розв’язування задач з використанням теорем додавання та множення імовірностей

1) Вводяться в розгляд подія, імовірність якої треба знайти, а також більш простіші події, імовірності яких відомі або можуть бути знайдені за класичним означенням.

2) «Шукана» випадкова подія (імовірність якої потрібно знайти) виражається через простіші події за допомогою алгебри подій, тобто операцій суми, добутку, заперечення (протилежної події). При цьому потрібно керуватися мнемонічними правилами: «+» $\leftrightarrow$ **або**, «×» $\leftrightarrow$ **і**.

3) В залежності від виду отриманого виразу використовуються теореми додавання імовірностей або(і) теорема множення імовірностей та її наслідки. При реалізації цього пункту необхідно з’ясовувати властивості подій (сумісність, несумісність, залежність, незалежність, протилежність або повноту пари чи групи подій).

Зауваження. Слід мати на увазі те, що в багатьох задачах реалізація пункту 2) неєдина. В таких випадках бажано вибрати найкомпактнішу, переконавшись у співпаданні остаточних результатів після виконання пункту 3). Якщо ж результати не співпадають, то необхідно перевірити правильність побудови в п. 2) або коректність виконання п. 3).

5.

Імовірність появи хоча б однієї події

Нехай в результаті випробування можуть відбутися події $A_1, A_2, ..., A_k$, імовірності появи кожної з яких відомі і які є **незалежними в сукупності**. Позначимо A — поява хоча б однієї із цих подій у випробуванні. Тоді згідно з означенням суми подій $A = A_1 + A_2 + ... + A_k$. Оскільки події $A_1, A_2, ..., A_k$ є сумісні, то теорема додавання імовірностей не «працює» для $k \geq 3$ (див. зауваження 3 п. 2.2). Разом з тим третій метод розв'язування задачі 2.2 вказує шлях до знаходження $P(A)$, підсумком якого є така рівність:

$$P(A) = 1 - q_1 q_2 ... q_k, \tag{2.12}$$

де $q_1 = P(\overline{A}_1), q_2 = P(\overline{A}_2), ..., q_k = P(\overline{A}_k)$.

В частинному випадку, коли $P(A_1) = P(A_2) = ... = P(A_k) = p$, отримується така формула

$$P(A) = 1 - q^k, \tag{2.12*}$$

де $q = 1 - p$.

Нарешті, якщо події $A_1, A_2, ..., A_k$ є **залежними** (не володіють властивістю незалежності в сукупності), тоді

$$P(A) = 1 - P(\overline{A}_1) P_{\overline{A}_1}(\overline{A}_2) ... P_{\overline{A}_1 \overline{A}_2 ... \overline{A}_{k-1}}(\overline{A}_k). \tag{2.13}$$

Приклад 1. Підприємство отримує від суміжників продукцію $П_1, П_2, П_3, П_4$ і виконає план, якщо вчасно отримає якісну продукцію в потрібній кількості. Ймовірність зриву поставок (як по кількості, так і по якості) по кожній із цих компонент технологічного процесу відповідно дорівнює 0,02; 0,01; 0,1; 0,05. Знайти імовірність невиконання заводом плану.

○ Позначимо: A — невиконання заводом плану, A_i — зрив поставки компоненти $П_i$ ($i = 1, 2, 3, 4$). Якщо відбувається хоча одна із подій A_1, A_2, A_3, A_4, тоді відбувається подія A. За умовою задачі $P(A_1) = 0{,}02$; $P(A_2) = 0{,}01$; $P(A_3) = 0{,}1$; $P(A_4) = 0{,}05$ і ці події можна вважати незалежними в сукупності. Використовуючи формулу (2.12), отримаємо:

$$q_1 = 1 - 0{,}02 = 0{,}98; \quad q_2 = 1 - 0{,}01 = 0{,}99;$$
$$q_3 = 1 - 0{,}1 = 0{,}9; \quad q_4 = 1 - 0{,}05 = 0{,}95,$$
$$\mathrm{P(A)} = 1 - 0{,}98 \cdot 0{,}99 \cdot 0{,}9 \cdot 0{,}95 = 1 - 0{,}829521 \approx 0{,}171.$$

Приклад 2. На залік виноситься 60 питань, кожне з яких надруковане на окремій картці. За правилами здачі цього предмету, студент навмання витягує 4 картки. Якщо він знатиме хоча б одне питання, тоді йому не потрібно повторно захищати розрахункові індивідуальні завдання (навіть якщо він і не отримує залік). Знайти імовірність того, що перший студент не буде повторно захищати свої індивідуальні завдання, якщо він підготував 40 питань.

○ Нехай A — студент не буде повторно захищати свої завдання, A_i — студент знає i-е витягнуте питання ($i = \overline{1,4}$). Подія A відбудеться при появі хоча б однієї із подій A_1, A_2, A_3, A_4. Але ці події не володіють властивістю незалежності у сукупності. Для доведення цього достатньо показати, наприклад, залежність подій A_1 та A_2:

$$P_{A_1}(A_2)=\frac{39}{59};\ P_{\overline{A}_1}(A_2)=\frac{40}{59}\quad\Rightarrow\quad P_{A_1}(A_2)\neq P_{\overline{A}_1}(A_2).$$

За формулою (2.13) для випадку $k = 4$

$$P(A)=1-P(\overline{A}_1)P_{\overline{A}_1}(\overline{A}_2)P_{\overline{A}_1\overline{A}_2}(\overline{A}_3)P_{\overline{A}_1\overline{A}_2\overline{A}_3}(\overline{A}_4)=$$

$$=1-\frac{20}{60}\cdot\frac{19}{59}\cdot\frac{18}{58}\cdot\frac{17}{57}=1-\frac{17}{1711}\approx 0,99.$$

Другий спосіб розв'язування. Протилежною до події A є $\overline{A}$— всіх чотирьох питань студент не знає. За класичним означенням

$$P(\overline{A})=\frac{C_{20}^4}{C_{60}^4}=\frac{20\cdot19\cdot18\cdot17}{60\cdot59\cdot58\cdot57}=\frac{17}{1711}.$$

Згідно із властивістю протилежності подій

$$P(A)=1-P(\overline{A})=1-\frac{17}{1711}\approx 0,99.$$

Приклад 3. Взимку при включенні запалення двигун почне працювати з імовірністю 0,7. Знайти імовірність того, що для запуску двигуна доведеться включати запалення менше чотирьох раз.

- Випробування – включення запалення.

Позначимо: B — двигун запрацює до четвертого включення запалення, A_i — двигун запускається при i-му включенні запалення, де i=1, 2, 3, оскільки 3<4. На перший погляд подія B відбудеться тоді, коли появиться хоча б одна із подій A_1, A_2, A_3. Це веде до використання формули (2.12). Проте, якщо відбудеться, наприклад, подія A_2 , то немає сенсу виключати запалення, а потім знову його включати (зокрема, не достовірно, що двигун запрацює при третьому включенні запалення, якщо він запрацював при другому включенні).

Отже, подія B відбудеться або при першій вдалій спробі (A_1), або при другій ($\overline{A}_1A_2$) або при третій ($\overline{A}_1\overline{A}_2A_3$):

$$B=A_1+\overline{A}_1A_2+\overline{A}_1\overline{A}_2A_3.$$

Використання теорем додавання та множення ймовірностей дає:

$$P(B)=P(A_1+\overline{A}_1A_2+\overline{A}_1\overline{A}_2A_3)=$$

$$=P(A_1)+P(\overline{A}_1)\cdot P(A_2)+P(\overline{A}_1)\cdot P(\overline{A}_2)\cdot P(A_3)=$$

$$=0,7+0,3\cdot0,7+0,3\cdot0,3\cdot0,7=0,7+0,21+0,063=0,973.$$

Висновки: 1. Використанню тієї чи іншої формули повинен передувати вдумливий аналіз умов задачі.

2. Імовірності вдалої спроби різко зменшуються із збільшенням номера спроби. Тому слід подбати про збільшення ефективності запалювальної системи і якість роботи акумулятора.

Приклад 3. Скільки карток спортлото «5 із 36» треба придбати, щоб з імовірністю, не менше від α, виграла хоча б одна із них?

○ Нехай придбано k (поки що невідоме число) карток. Введемо в розгляд події: A_j — виграла j-та картка $(j=\overline{1;k})$, A — виграшною виявиться хоча б одна із k карток (відбудеться хоча б одна із подій $A_1, A_2, ..., A_k$). Оскільки картки заповнюються довільним (випадковим чином) і загальне число різних їх заповнень дуже велике $\left(C_{36}^5 = 376992\right)$, то можна вважати події $A_1, A_2, ..., A_k$ незалежними у сукупності і виконаними рівності $P(A_1) = P(A_2) = ... = P(A_k) = p$. Навмання заповнена картка буде виграшною, якщо в результаті проведення тиражу буде вгадано 3 числа (B_3), або 4 (B_4), або 5 (B_5). Із врахуванням несумісності цих подій $p = P(B_3) + P(B_4) + P(B_5)$. За класичним означенням

$$P(B_3) = \frac{C_5^3 \cdot C_{31}^2}{C_{36}^5} = \frac{4650}{376992}, \quad P(B_4) = \frac{C_5^4 \cdot C_{31}^1}{C_{36}^5} = \frac{155}{376992},$$

$$P(B_5) = \frac{C_5^5}{C_{36}^5} = \frac{1}{376992}.$$

Тому $p = \dfrac{4650}{376992} + \dfrac{115}{376992} + \dfrac{1}{376992} = \dfrac{4766}{376992}, \quad q \approx 0{,}987.$

За формулою (2.12*) $P(A) = 1 - q^k$. Але за умовою k повинно бути таким, щоб виконувалась нерівність $P(A) \geq \alpha$, звідки $q^k \leq 1 - \alpha$. Прологарифмувавши ліву і праву частини і врахувавши те, що $\ln q < 0$, остаточно отримаємо:

$$k \geq \frac{\ln(1-\alpha)}{\ln q}, \text{ де } q \approx 0{,}987.$$

6.

Приклад 1. Імовірності появи кожної з трьох незалежних у сукупності подій A_1, A_2, A_3 відповідно рівні p_1, p_2, p_3. Знайти імовірність появи у випробуванні тільки однієї з цих подій.

○ Нехай B — відбуття тільки однієї із подій A_1, A_2, A_3. Поява події B передбачає **або** відбуття події A_1 (і не відбуття A_2 та A_3), **або** події A_2 (і не відбуття A_1 та A_3), **або** події A_3 (і не відбуття A_1 та A_2), тобто

$$B = A_1\overline{A}_2\overline{A}_3 + \overline{A}_1 A_2 \overline{A}_3 + \overline{A}_1\overline{A}_2 A_3.$$

Події-доданки є попарно несумісними (доведіть від протилежного), а співмножники кожного доданку — незалежні у сукупності. Використавши теореми суми, а потім добутку імовірностей, отримаємо:

$$P(B) = P(A_1\overline{A}_2\overline{A}_3 + \overline{A}_1 A_2 \overline{A}_3 + \overline{A}_1\overline{A}_2 A_3) = P(A_1\overline{A}_2\overline{A}_3) + P(\overline{A}_1 A_2 \overline{A}_3) +$$
$$+P(\overline{A}_1\overline{A}_2 A_3) = P(A_1)P(\overline{A}_2)P(\overline{A}_3) + P(\overline{A}_1)P(A_2)P(\overline{A}_3) +$$
$$+P(\overline{A}_1)P(\overline{A}_2)P(A_3) = p_1 q_2 q_3 + q_1 p_2 q_3 + q_1 q_2 p_3,$$

де $q_i = 1 - p_i$, $i = 1, 2, 3$.

Приклад 2. Виходячи з умови попереднього прикладу, знайти імовірність того, що у випробуванні відбудуться: а) всі; б) жодна; в) тільки дві; г) хоча б дві; д) не більше однієї; е) не більше двох.

○ а) $P\left(A_1A_2A_3\right)=p_1p_2p_3$; б) $P\left(\overline{A}_1\overline{A}_2\overline{A}_3\right)=q_1q_2q_3$;

в) $P\left(A_1A_2\overline{A}_3+A_1\overline{A}_2A_3+\overline{A}_1A_2A_3\right)=P\left(A_1A_2\overline{A}_3\right)+\left(A_1\overline{A}_2A_3\right)+\left(\overline{A}_1A_2A_3\right)=$

$=p_1p_2q_3+p_1q_2p_3+q_1p_2p_3$;

г) $P\left(A_1A_2\overline{A}_3+A_1\overline{A}_2A_3+\overline{A}_1A_2A_3+A_1A_2A_3\right)=$

$=P\left(A_1A_2\overline{A}_3\right)+\left(A_1\overline{A}_2A_3\right)+\left(\overline{A}_1A_2A_3\right)+\left(A_1A_2A_3\right)=$

$p_1p_2q_3+p_1q_2p_3+q_1p_2p_3+p_1p_2p_3$;

д) $P\left(\overline{A}_1\overline{A}_2\overline{A}_3+A_1\overline{A}_2\overline{A}_3+\overline{A}_1A_2\overline{A}_3+\overline{A}_1\overline{A}_2A_3\right)=$

$P\left(\overline{A}_1\overline{A}_2\overline{A}_3\right)+\left(A_1\overline{A}_2\overline{A}_3\right)+\left(\overline{A}_1A_2\overline{A}_3\right)+\left(\overline{A}_1\overline{A}_2A_3\right)=$

$q_1q_2q_3+p_1q_2q_3+q_1p_2q_3+q_1q_2p_3$;

е) протилежною подією до події, імовірність якої треба знайти, є $A_1A_2A_3$; далі використати властивість протилежної події (поміркуйте самостійно).

7.

Формула повної імовірності

Нехай подія A може відбутися тільки при умові появи однієї із подій $B_1, B_2, ..., B_n$, які утворюють повну групу. Нехай відомі імовірності $P(B_k)$, $P_{B_k}(A)$, $k=(\overline{1,n})$. Відповідь на питання: як знайти $P(A)$? — дає

Теорема. Імовірність події *A*, яка може відбутися тільки після появи однієї із подій $B_1, B_2, ..., B_n$, які утворюють повну групу, знаходиться за так званою формулою повної імовірності:

$$P(A)=P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)+\ldots+P(B_n)P_{B_n}(A). \tag{2.14}$$

Зауваження. Оскільки до випробування невідомо, після якої із подій $B_1, B_2, ..., B_n$ відбудеться подія ***A***, то ці події називаються **гіпотезами (припущеннями)**.

8.

Формули Байєса

Нехай виконуються умови теореми відносно подій $B_1, B_2, ..., B_n$ та A. Припустимо, що проведено випробування, в результаті якого відбулася подія A. Виникає питання: як «переоцінити» імовірності гіпотез $B_1, B_2, ..., B_n$ із врахуванням відбуття події A, тобто знайти умовні імовірності $P_A(B_k)$, $(k=\overline{1,n})$? Відповідь дають так звані формули Байєса:

$$P_A(B_k)=\frac{P(B_k)P_{B_k}(A)}{\sum\limits_{i=1}^{n}P(B_i)P_{B_i}(A)}, k=1,2,\ldots,n. \tag{2.15}$$

9.

Алгоритм розв’язування задач з використанням формул повної імовірності та Байєса

Рекомендується така послідовність розв’язування задач.

1) Формулюють гіпотези $B_1, B_2, ..., B_n$ і подію A. При цьому слід перевірити **повноту групи гіпотез**, а також те, що **подія *A* може відбутися тільки після появи однієї із гіпотез**.

2) Знаходяться імовірності гіпотез. Правильність розрахунків контролюється виконанням рівності $P(B_1) + P(B_2) + ... + P(B_n) = 1$. Обчислюються умовні імовірності $P_{B_1}(A), P_{B_2}(A), \ldots, P_{B_n}(A)$.

3) Вибирається формула повної імовірності або формули Байєса. Останні використовуються тоді, коли є інформація про відбуття випадкової події.

Приклад 1. Три верстати-автомати штампують однотипні деталі, що потрапляють на спільний конвейєр. Продуктивність другого автомату на 40% вища від продуктивності першого і вдвічі — від третього. Відсоток браку для кожного з автоматів дорівнює відповідно 3, 6, 2. а). Яка імовірність того, що навмання взята з конвейєра деталь виявиться бракованою? б). Навмання взята деталь виявилася бракованою. Що імовірніше: ця деталь виготовлена першим чи третім автоматом?

○ а) Першу частину задачі розв’язуємо за формулою повної імовірності, оскільки відсутня інформація про відбуття випадкової події.

При відборі довільним чином деталі з конвейєра невідомо, яким автоматом вона виготовлена. Цю **невизначеність не можна трактувати як відбуття випадкової події** (типова помилка студентів). Можна зробити припущення: будь-яка деталь (стандартна чи бракована) виготовлена або першим автоматом (подія B_1), або другим (B_2), або третім (B_3). Неважко переконатися в тому, що події B_1, B_2, B_3 утворюють повну групу (доведіть!).

Позначимо: A — навмання взята деталь бракована. Ясно, що подія A може відбутися після появи однієї із гіпотез, адже щоб говорити про бракованість деталі, потрібно її спочатку виготовити.

Знайдемо імовірності гіпотез. Для цього позначимо через a кількість деталей, що виготовить перший автомат за деякий проміжок часу. Тоді за умовою задачі за цей же проміжок часу другий автомат виготовить $1{,}4a$ деталей, а третій — $0{,}7a$. За класичним означенням імовірності

$$P(B_1) = \frac{a}{3{,}1a} = \frac{10}{31};\; P(B_2) = \frac{1{,}4a}{3{,}1a} = \frac{14}{31};\; P(B_3) = \frac{0{,}7a}{3{,}1a} = \frac{7}{31}.$$

Перевіримо обчислення. Оскільки гіпотези утворюють повну групу, то сума їх імовірностей повинна дорівнювати одиниці:

$$P(B_1) + P(B_2) + P(B_3) = \frac{10}{31} + \frac{14}{31} + \frac{7}{31} = 1.$$

Згідно з умовою задачі і за класичним означенням

$$P_{B_1}(A)=\frac{3}{100}=0,03;\ P_{B_2}(A)=\frac{6}{100}=0,06;\ P_{B_3}(A)=\frac{2}{100}=0,02.$$

Прочитайте ліві частини цих рівностей з урахуванням змісту подій A, B_1, B_2, B_3 (див. п. 2.2), а також переконайтеся у правильності цих рівностей.

Формула повної імовірності (2.14) у цьому випадку має такий вид:

$$P(A)=P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)+P(B_3)P_{B_3}(A).$$

Підставивши обчислені дані в праву частину, отримаємо

$$P(A)=\frac{10}{31}\cdot 0,03+\frac{14}{31}\cdot 0,06+\frac{7}{31}\cdot 0,02\approx 0,0401.$$

б) У другій частині задачі використаємо формули Байєса, оскільки є інформація про відбуття випадкової події (A — деталь виявилася бракованою). При цьому ми повинні знайти $P_A(B_1)$, $P_A(B_3)$ і порівняти їх. Але більшою з них буде та імовірність, чисельник в правій частині формули Байєса для якої буде більшим. Тобто, достатньо порівняти $P(B_1)P_{B_1}(A)$ і $P(B_3)P_{B_3}(A)$: $P(B_1)P_{B_1}(A)=\dfrac{0,3}{31}$, $P(B_3)P_{B_3}(A)=\dfrac{0,14}{31}$. Отже, $P_A(B_1)>P_A(B_3)$, і можна зробити висновок: більш імовірніше, що бракована деталь виготовлена першим автоматом.

Приклад 2. В кінці потокової лінії по виготовленню приладів встановлено два автомати-контролери, які визначають: має прилад вищу категорію якості, чи ні. Статистично встановлено, що 20% продукції задовільняють вимогам вищої категорії якості, а контролери роблять помилкові висновки стосовно якості приладу відповідно у 5% і 1% випадків. Випадково один і той самий прилад був перевірений обома автоматами: перший визначив, що прилад не задовільняє вимогам вищої категорії, а другий, що задовільняє. Якому із висновків вірити?

○ Стосовно двічі перевіреного приладу можна висунути два припущення: або він справді задовільняє вимогам вищої категорії якості (подія B_1) або не задовільняє (подія B_2). За умовою задачі $P(B_1)$=0,2, $P(B_2)$=0,8.

Розглянемо спочатку ситуацію відносно висновку першого контролера. Подія A – визнання ним відсутності у перевіреного приладу якостей вищої категорії. Ця випадкова подія відбулася, тому використаємо формулу Байєса.

За умовою у 5% випадків перший контролер помиляється відносно якості приладу, а тому у 95% робить правильні висновки. В зв’язку з цим $P_{B_1}(A)=0,05$, $P_{B_2}(A)=0,95$ і за формулою Байєса

$$P_A(B_2)=\frac{P(B_2)\cdot P_{B_2}(A)}{P(B_1)\cdot P_{B_1}(A)+P(B_2)\cdot P_{B_2}(A)}=\frac{0,8\cdot 0,95}{0,2\cdot 0,05+0,8\cdot 0,95}=0,987013.$$

Отже, імовірність того, що перший контролер зробив правильний висновок, дорівнює 0,987013.

Нехай A^* – визнання вищої категорії якості перевіреного приладу другим контролером. З використанням умови задачі отримаємо:

$P_{B_1}(A^*) = 0,99$, $P_{B_2}(A^*) = 0,01$,

$$P_{A^*}(B_1) = \frac{P(B_1) \cdot P_{B_1}(A^*)}{P(B_1) \cdot P_{B_1}(A^*) + P(B_2) \cdot P_{B_2}(A^*)} = \frac{0,2 \cdot 0,99}{0,2 \cdot 0,99 + 0,8 \cdot 0,01} = 0,961165$$

Тобто, імовірність того, що другий контролер зробив правильний висновок, складає 0,961165.

Порівняння знайдених імовірностей дозволяє зробити висновок: потрібно вірити висновку першого автомата-контролера (точніше, більш імовірніше, що перший контролер зробив правильний висновок).

Рекомендована література: [1], С. 29–39; [2], С. 31–53; [4], 4–23; [5], 30–54.

Лекція 3.

Тема: *Повторні незалежні випробування.*

1. ***Формула Бернуллі.***
10. ***Найімовірніше число появи події.***
11. ***Локальна формула Лапласа.***
12. ***Формула Пуассона.***
13. ***Інтегральна формула Лапласа.***
14. ***Ймовірність відхилення відносної частоти події від її постійної ймовірності.***
15. ***Алгоритм розв'язування задач для повторних незалежних випробувань.***

1.

На практиці часто зустрічаються випадки, коли проводиться не одне випробування, а декілька (можливо, дуже велике число). Такі випробування називаються **повторними**, а їх сукупність — **схемою повторних випробувань**, або **схемою Бернуллі**. В кожному із таких випробувань може відбутися одна і та ж випадкова подія A. Якщо $P(A)$ залишається незмінною для **кожного** випробування, то такі випробування будемо називати **незалежними** (**відносно події** A).

Формула Бернуллі

Теорема. Імовірність того, що в n повторних незалежних випробуваннях випадкова подія A відбудеться рівно m разів, знаходиться за формулою Бернуллі:

$$P_n(m) = C_n^m p^m q^{n-m}, \tag{3.1}$$

де C_n^m — число комбінацій, визначене формулою (1.4),

p — імовірність появи події A в **одному** випробуванні ($p = P(A)$),

q — імовірність **непояви** події A в одному випробуванні $(q = P(\overline{A}))$.

Приклад 1. Компанія володіє мережею ділерів на біржі. Імовірність того, що ділер буде грати вдало, становить 0,7. 1) Знайти імовірність того, що з п'яти ділерів будуть у збитках: а) два; б) хоча б два (вважається, що дії ділерів на біржі є незалежними).

○ Випробування — гра ділера. Оскільки ділерів є 5, то $n = 5$. 1) Подія A — збиткова гра ділера. За умовою $P(\overline{A}) = q = 0,7,$ тоді $p = 1 - q = 0,3$. Для випадку а) число появи події A $m = 2$, і потрібно знайти $P_5(2)$. Так як $n = 5$ — мале, то використаємо формулу Бернуллі:

$$P_5(2) = C_5^2(0,3)^2 \cdot (0,7)^3 = 10 \cdot 0,09 \cdot 0,343 = 0,3087.$$

б) $P_5(m \geq 2) = P_5(2 \leq m \leq 5)$ — шукана імовірність. Хоча вона візуально нагадує ліву частину інтегральної формули Лапласа, останню використовувати не можна, бо $npq = 5 \cdot 0,3 \cdot 0,7 = \ = 1,05 << 9$. Випадкова подія $(2 \le m \le 5)$ може відбутися тоді, коли або $(m = 2)$, або $(m = 3)$, або $(m = 4)$, або $(m = 5)$, тобто $(2 \le m \le 5) = (m = 2) + (m = 3) + (m = 4) + (m = 5)$. Випадкові події-доданки справа є попарно несумісними, тому згідно з теоремою додавання імовірностей

$P_5(2 \le m \le 5) = P(m = 2) + P(m = 3) + P(m = 4) + P(m = 5) =$

$= P_5(2) + P_5(3) + P_5(4) + P_5(5)$,

де в останній рівності було враховано, зокрема, те, що випадкова подія $(m = 2)$ — в п'яти випробуваннях подія A відбудеться рівно два рази, тобто $P(m = 2) = P_5(2)$. Використавши формулу Бернуллі, отримаємо:

$$P_5(3) = C_5^3(0,3)^3 \cdot (0,7)^2 = 0,1323;$$

$$P_5(4) = C_5^4(0,3)^4 \cdot 0,7 = 0,02835;$$

$$P_5(5) = C_5^5(0,3)^5 \cdot 0,7^0 = 0,00243.$$

Остаточно

$$P_5(2 \le m \le 5) = 0,3087 + 0,1323 + 0,028535 + 0,00243 = 0,47178.$$

Другий метод. Події $(m \ge 2)$ та $(m < 2)$ протилежні, тому

$P_5(m \ge 2) = 1 - P_5(m < 2) = 1 - [P_5(0) + P_5(1)] =$

$= 1 - [C_5^0(0,3)^0(0,7)^5 + C_5^1(0,3)^1(0,7)^4] =$

$= 1 - (0,16807 + 0,36015) = 0,47178.$

Висновок: другий метод значно швидше веде до мети, його ефективність ще більш відчутна при збільшенні числа доданків.

2.

Найімовірніше число появи події

В n повторних незалежних випробуваннях подія A може відбутися число разів

$$0, 1, ..., m_0-1, m_0, m_0+1, ..., n \qquad (3.2)$$

з відповідними імовірностями (які можна знайти за формулою Бернуллі):

$$P_n(0), P_n(1), ..., P_n(m_0-1), P_n(m_0), P_n(m_0+1), ..., P_n(n). \qquad (3.3)$$

Означення. Число m_0 в послідовності (3.2) називається найімовірнішим числом появи події в n незалежних випробуваннях (або модою), якщо йому відповідає найбільша імовірність в послідовності (3.3).

Найімовірніше число можна знайти з такої подвійної нерівності:

$$np - q \le m_0 \le np + p. \qquad (3.4)$$

Оскільки різниця між правою і лівою частинами цієї нерівності дорівнює 1, то можна зробити висновок, що максимальне число найімовірніших подій рівне двом.

Довести формулу (3.2) не складно використавши формулу Бернуллі $P_n(m) = C_n^m p^m q^{n-m}$, а також те, що $P_n(m_0-1) \le P_n(m_0) \le P_n(m_0+1)$, де m_0 — найімовірніше число появи події.

Врахувавши, що $P_n(m_0 - 1) = C_n^{m_0-1} p^{m_0-1} q^{n-(m_0-1)}$, $P_n(m_0) = C_n^{m_0} p^{m_0} q^{n-m_0}$,

$P_n(m_0+1)=C_n^{m_0+1}p^{m_0+1}q^{n-(m_0+1)}$, а також формулу числа комбінацій, доводимо формулу (3.2).

Приклад 1. Компанія володіє мережею ділерів на біржі. Імовірність того, що ділер буде грати вдало, становить 0,7. Знайти найімовірніше число ділерів, які будуть грати вдало, а також імовірність такої кількості.

○ Випробування — гра ділера. Оскільки ділерів є 5, то $n = 5$.

Подія A — вдала гра ділерів. За умовою задачі $n = 5$, $p = 0{,}7$, $q = 0{,}3$. Найімовірніше число m_0 ділерів, які будуть грати вдало, знайдемо за подвійною нерівністю (3.4)

$$np - q \le m_0 \le np + p.$$

Підставивши значення в ліву та праву частини, знайдемо $3{,}2 \le m_0 \le 4{,}2$, звідки з врахуванням того, що m_0 — ціле число, остаточно отримаємо: $m_0 = 4$. Нарешті, $P_5(m_0) = P_5(4) = C_5^4(0{,}7)^4\cdot 0{,}3 = 0{,}36015$.

3.

Локальна формула Лапласа

Користуватися формулою Бернуллі для великих значень n досить важко. Наприклад, при $p = 0{,}2$, $q = 0{,}8$ $P_{50}(30)=C_{50}^{30}(0{,}2)^{30}(0{,}8)^{20}$, де $C_{50}^{30}=4712921\cdot 10^7$.

Одну із реалізацій наближеного знаходження правої частини формули (3.1) дає

Локальна теорема Лапласа (Муавра—Лапласа).

Якщо імовірність p появи випадкової події A в кожному із n повторних випробувань залишається незмінною (причому $0 < p < 1$), а число випробувань достатньо велике, то імовірність того, що в n випробуваннях подія відбудеться m разів, знаходиться за наближеною формулою

$$P_n(m)\approx\frac{1}{\sqrt{npq}}\varphi(x), \tag{3.5}$$

де $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ — функція Гаусса, $x=\frac{m-np}{\sqrt{npq}}, q=1-p$.

Зауваження. Формула (3.5) називається **локальною формулою Лапласа**. Її точність зростає при збільшенні n.

Функція Гаусса протабульована і її значення наведені в табл. 1 додатку. Для правильного користування цією таблицею слід враховувати такі властивості функції Гаусса: 1) $\varphi(x)$ визначена для всіх $x \in R$; 2) $\varphi(x)$ — парна функція $(\varphi(-x)=\varphi(x))$; 3) для додатних x $\varphi(x)$ дуже швидко прямує до 0 при збільшенні x, зокрема $\varphi(3{,}99)=0{,}0001$. Згідно із цими властивостями в таблиці наведені значення функції Гаусса для x з проміжку [0; 5].

Практично можна вважати, що локальна формула Лапласа дає

добре наближення, якщо $npq > 9$. Якщо ж вимоги до точності значення вищі, то слід вимагати виконання нерівності $npq \ge 25$.

4.

Формула Пуассона

Нерівність $npq > 9$ навіть для великих n може не виконуватися (а отже, похибка при використанні локальної формули Лапласа буде дуже великою) у випадку рідкісних (малоімовірних) подій A, тобто таких подій, для яких p значно менше 0,1 ($p << 0,1$). В таких випадках слід користуватися іншим наближенням правої частини формули Бернуллі. Одне із них дається таким твердженням.

Теорема Пуассона. Якщо в кожному із n повторних випробувань імовірність p появи події A стала і мала ($p << 0,1$), а число випробувань n досить велике, то імовірність того, що подія A настане в цих випробуваннях рівно m разів, знаходиться за формулою

$$P_n(m) \approx \frac{\lambda^m e^{-\lambda}}{m!}, \qquad (3.6)$$

де $\lambda = np$.

Зауваження. Похибка в наближеній рівності (3.6), яка називається **формулою Пуассона**, тим менша, чим більше число випробувань n.

Значення функції $P(m) = \dfrac{\lambda^m e^{-\lambda}}{m!}$ двох змінних λ та m для деяких m та λ наведені в табл. додатків.

Відмітимо, що формула Пуассона використовується також до числа невідбуття події A, якщо $q << 0,1$, а nq невелике.

Приклад 1. При скануванні текстового матеріалу в середньому на кожну тисячу символів два помилкові. Знайти імовірність того, що після сканування тексту обсягом в 2 500 символів виявиться помилкових: а) шість символів; б) хоча б шість.

Випробування — сканування символа тексту, подія A — отримання помилкового символу. За умовою $n = 2500$, $p = P(A) = 0,002$.

а) Число появи події $m = 6$. Для знаходження $P_{2500}(6)$ скористаємося формулою Пуассона, оскільки n — велике, $p = 0,002 << 0,1$, $\lambda = np = 5 < 9$.

$$P_n(m) \approx \frac{\lambda^m e^{-\lambda}}{m!}, \quad P_{2500}(6) \approx \frac{5^6 \cdot e^{-5}}{6!} = 0,146223.$$

Останнє значення знайдене для функції $\dfrac{\lambda^m e^{-\lambda}}{m!}$ двох змінних λ та m в табл. додатків для значень $\lambda = 5$, $m = 6$.

б) Випадкова подія ($m \ge 6$), імовірність якої треба знайти, зображається через прості події таким чином:

$$(m \ge 6) = (m = 6) + (m = 7) + ... + (m = 2500).$$

Використовувати теорему додавання імовірностей в такому випадку практично неможливо в зв'язку з тим, що в правій частині є 2495 доданків (!). З другого боку, протилежною до події $(m \geq 6)$ є подія $(m < 6)$, для якої виконується рівність

$$(m < 6) = (m = 0) + (m = 1) + (m = 2) + (m = 3) + (m = 4) + (m = 5),$$

звідки після використання теореми додавання імовірностей отримуємо

$$P_{2500}(m < 6) = P_{2500}(0) + P_{2500}(1) + P_{2500}(2) + P_{2500}(3) + P_{2500}(4) + P_{2500}(5).$$

Кожний із доданків обчислюється за формулою Пуассона. В даному випадку ці імовірності знаходяться за табл. 2 додатків для $\lambda = 5$ та $m = 0, 1, ..., 5$. Тобто

$$P_{2500}(m < 6) = 0{,}00674 + 0{,}03369 + 0{,}08422 + 0{,}14037 + \\ + 0{,}17547 + 0{,}17547 = 0{,}61596.$$

З врахуванням протилежності подій

$$P_{2500}(m \geq 6) = 1 - P_{2500}(m < 6) = 1 - 0{,}61596 = 0{,}38404.$$

5.

Інтегральна формула Лапласа

Інтегральна теорема Лапласа (Муавра—Лапласа). Якщо імовірність *p* появи події *A* в кожному із *n* повторних випробувань є сталою ($0 < p < 1$), а число випробувань досить велике, то імовірність того, що подія *A* в цих випробуваннях відбудеться не менше m_1 разів і не більше m_2 разів, знаходиться за такою наближеною рівністю (інтегральною формулою Лапласа):

$$P_n(m_1 \leq m \leq m_2) \approx \Phi(x_2) - \Phi(x_1), \tag{3.7}$$

де $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-t^2/2} dt$ — функція Лапласа,

$$x_1 = \frac{m_1 - np}{\sqrt{npq}}, x_2 = \frac{m_2 - np}{\sqrt{npq}}.$$

Функція Лапласа протабульована і її значення наведені в табл. додатку. При користуванні цією таблицею потрібно враховувати такі властивості функції Лапласа:

1) $\Phi(x)$ визначена для всіх $x \in R$;

2) $\Phi(x)$ непарна функція ($\Phi(-x) = -\Phi(x)$);

3) $\Phi(x)$ монотонно зростає для всіх $x \in R$, при цьому $y = -0{,}5$ — лівостороння асимптота, а $y = 0{,}5$ — правостороння;

4) швидкість зростання $\Phi(x)$ на проміжку [0; 5] дуже висока (зокрема $\Phi(5) = 0{,}499997$), тому для всіх $x > 5$ з мізерною похибкою $\Phi(x) \approx 0{,}5$.

Точність інтегральної формули Лапласа тим більша, чим більше число випробувань *n*. Як і у випадку локальної, **інтегральна формула дає добрі наближення, якщо $npq > 9$**. Якщо ж вимоги до точності значно вищі, то потрібно вимагати виконання нерівності $npq \geq 25$.

Приклад 1. Два станки з програмним управлінням виготовляють однотипні деталі, які надходять на спільний конвейєр. Їх продуктивності відносяться як 2 : 3, причому перший виготовляє 35% деталей вищої якості, якими комплектуються вироби на експорт, другий — 10%. Знайти імовірність того, що з 400 навмання відібраних з конвейєра деталей вищої якості виявилось: а) хоча б 80; б) не більше 75.

Випробування — відбір деталі, за умовою $n = 400$. A — відібрана деталь має вищу якість. Знайдемо $P(A) = p$. Відібрана деталь (вищої якості чи ні) може бути виготовлена або першим станком (подія B_1), або другим (подія B_2). Ці гіпотези утворюють повну групу, а подія A може відбутися тільки після появи однієї із них. Тому $P(A)$ можна знайти за формулою повної імовірності

$$P(A) = P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A).$$

Якщо за деякий проміжок часу перший станок виготовить $2a$ деталей, то другий — $3a$ деталей. За класичними означенням імовірності

$$P(B_1) = \frac{2a}{5a} = 0,4,\ \ P(B_2) = \frac{3a}{5a} = 0,6,$$

$$P_{B_1}(A) = \frac{35}{100} = 0,35,\ \ P_{B_2}(A) = \frac{10}{100} = 0,1.$$

Підставивши знайдені імовірності у формулу повної імовірності, отримаємо

$P(A) = 0,4 \cdot 0,35 + 0,6 \cdot 0,1 = 0,2$. Отже, $p = 0,2$, $q = 1 - p = 0,8$.

а) Для знаходження імовірності $P_{400}(m \ge 80) = P_{400}(80 \le m \le 400)$ (див. зауваження до алгоритму) використовуємо інтегральну формулу Лапласа, оскільки $npq = 64 > 9$. У відповідності із (3.7)

$$x_1 = \frac{m_1 - np}{\sqrt{npq}} = \frac{80 - 400 \cdot 0,2}{\sqrt{400 \cdot 0,2 \cdot 0,8}} = 0;\ \ x_2 = \frac{m_2 - np}{\sqrt{npq}} = \frac{400 - 400 \cdot 0,2}{\sqrt{400 \cdot 0,2 \cdot 0,8}} = 40,$$

$$P_{400}(80 \le m \le 400) \approx \Phi(40) - \Phi(0) = 0,5 - 0 = 0,5.$$

При знаходженні $\Phi(40)$ враховувалося те, що $\Phi(x) = 0,5$ для $x > 5$, а $\Phi(0)$ знаходилося за табл. додатка.

б) Імовірність $P_{400}(m \le 75) = P_{400}(0 \le m \le 75)$ знову обчислюємо за інтегральною формулою Лапласа:

$$x_1 = \frac{0 - 400 \cdot 0,2}{8} = -10;\ \ x_2 = \frac{75 - 400 \cdot 0,2}{8} = -0,63,$$

$$P_{400}(0 \le m \le 75) \approx \Phi(-0,63) - \Phi(-10) = \Phi(10) - \Phi(0,63) =$$
$$= 0,5 - 0,2357 = 0,2643.$$

В останніх рівностях використана непарність $\Phi(x)$.

Зауваження. При знаходженні імовірностей в б) можна було досягти вищої точності, оскільки точні значення x_1 та x_2 рівні –0,625, а $\varphi(0,62) = 0,3292$, $\Phi(0,62) = 0,2324$. Для цього потрібно здійснити лінійну інтерполяцію:

$\varphi(0,625) = (\varphi(0,62) + \varphi(0,63))/2 = (0,3292 + 0,0,3271)/2 = 0,32815,$

$\Phi(0,625) = (\Phi(0,62) + \Phi(0,63)/2 = (0,2324 + 0,2357)/2 = 0,23405.$

Проте доцільність такого підходу визначається вимогами до точності відповіді, поставленими конкретною задачею.

Наскільки суттєво враховувати **всі** умови використання тієї чи іншої формули, ілюструє розв'язування наступної задачі.

6.

Імовірність відхилення відносної частоти події

від її постійної імовірності

Теорема. Якщо імовірність *p* появи випадкової події *A* в кожному з *n* повторних випробувань стала, а число випробувань досить велике, то імовірність того, що відхилення відносної частоти *m/n* події *A* від її імовірності *p* по абсолютній величині не перевищить заданого числа ε > 0, знаходиться за формулою

$$P\left(\left|m/n - p\right| \le \varepsilon\right) \approx 2\Phi\left(\varepsilon\sqrt{n/pq}\right) . \tag{3.8}$$

Приклад. Імовірність того, що кінескоп задовольняє вимогам вищого сорту, дорівнює 0,8. За наступні два місяці ВТК перевірив 900 кінескопів. Знайти з імовірністю 0,95 межі, в яких буде знаходитися число m кінескопів вищого гатунку серед перевірених.

Випробування — перевірка кінескопа. Подія A — кінескоп задовольняє якостям вищого сорту.

За умовою $n = 900$, $p = 0{,}8$, $q = 0{,}2$, $P\left(\left|m/n - 0{,}8\right| \le \varepsilon\right) = 0{,}95$. Потрібно знайти межі для числа m. Знайдемо спочатку ε, використавши згідно із формулою (3.8) рівність

$$2\Phi\left(\varepsilon\sqrt{900/(0{,}8\cdot 0{,}2)}\right) = 0{,}95 \quad \text{або} \quad \Phi(75\varepsilon) = 0{,}475.$$

За таблицею функції Лапласа знайдемо $\Phi(1{,}96) = 0{,}475$, тому $75\varepsilon = 1{,}96$, звідки $\varepsilon \approx 0{,}03$.

Таким чином, з імовірністю 0,95 відхилення відносної частоти числа кінескопів вищої якості від імовірності 0,8 задовольняє нерівності

$$\left|m/800 - 0{,}8\right| \le 0{,}03 \quad \text{або} \quad 0{,}77 \le m/800 \le 0{,}83,$$

і остаточно $616 \le m \le 664$. ●

7.

Алгоритм розв'язування задач для повторних незалежних випробувань

В цій темі, як і для попередніх, залишається актуальним питання вибору тієї чи іншої формули при розв'язуванні конкретних задач. Це зумовлено, по-перше, тим, що у всіх трьох формулах (Бернуллі, локальній формулі Лапласа та Пуассона) ліві частини однакові. З другого боку, при знаходженні імовірності $P_n(m_1 \le m \le m_2)$ зовсім не обов'язково (а деколи й помилково) використовувати інтегральну формулу Лапласа.

1. Обчислення $P_n(m)$.

а) Якщо n мале ($n \le 15$), то використовується формула Бернуллі для будь-яких значень p та q.

б) Якщо n велике, а p та q немалі, тобто при виконанні нерівності $npq > 9$, тоді використовується локальна формула Лапласа.

в) Якщо ж n велике, а p дуже мале (значно менше 0,1) і $\lambda = np \leq 9$, то застосовується формула Пуассона. При великому n, дуже малому q ($q << 0{,}1$) і при виконанні нерівності $\lambda' = nq \leq 9$ слід перейти до числа невиконання події A.

2. Знаходження $P_n(m_1 \leq m \leq m_2)$.

а) Якщо n мале ($n \leq 15$), тоді потрібно використати спочатку теорему додавання імовірностей, а потім формулу Бернуллі.

б) Для великих n і немалих p та q, тобто при виконанні нерівності $npq > 9$ використовується інтегральна формула Лапласа.

в) Для великих n і малих p використовується або теорема додавання імовірностей з наступним застосуванням формули Пуассона, або здійснюється перехід до протилежної події з наступним використанням теореми додавання імовірностей і формули Пуассона. При виборі однієї із альтернатив слід керуватися мінімізацією числа доданків в теоремі додавання імовірностей. Якщо n велике, а q мале і $\lambda' = nq \leq 9$, тоді потрібно перейти до числа невиконання події A, а потім виконати рекомендації початку цього підпункту.

Зауваження. Якщо в n повторних незалежних випробуваннях потрібно знайти імовірності $P(m \leq k)$ та $P(m \geq k)$, де k ціле число, що не перевищує n, тоді потрібно скористатися рівностями $P(m \leq k) = P_n(0 \leq m \leq k)$, $P(m \geq k) = P_n(k \leq m \leq n)$ і перейти до п. 2 алгоритму.

Рекомендована література: [1], С. 56–68; [2], С. 55–63; [4], С. 47–55; [5], С. 68–83.

Лекція 4.

Тема: *Дискретні випадкові величини та їх числові характеристики.*

1. ***Випадкові величини та їх види.***
2. ***Закон розподілу ймовірностей дискретної випадкової величини.***
3. ***Основні розподіли дискретних (цілочисельних) випадкових величин: рівномірний, біноміальний, Пуассонівський, геометричний, гіпергеометричний.***
4. ***Найпростіший потік подій.***
5. ***Дії над випадковими величинами.***
6. ***Числові характеристики дискретних випадкових величин та їх властивості (математичне сподівання, дисперсія, середньоквадратичне відхилення, початковий та центральний момент).***
7. ***Числові характеристики біноміального розподілу.***

1.

Випадковою називається величина, яка при випробуванні набирає єдиного значення із всіх можливих з деякою імовірністю, тобто наперед невідомо, яке конкретне можливе значення вона набере, оскільки це залежить від випадкових причин.

Дискретною (**перервною**) називається випадкова величина, можливі значення якої є ізольованими числами. Число можливих значень дискретної випадкової величини може бути скінченним або зчисленним. В останньому випадку можна встановити взаємно-однозначну відповідність між можливими значеннями і натуральними числами 1, 2, 3, ..., n, ...

Неперервною називається випадкова величина, можливі значення якої заповнюють суцільно деякий скінченний або нескінченний проміжок. Очевидно, що число можливих значень кожної неперервної величини нескінченне.

Зауваження. Означення неперервної випадкової величини має попередній характер. Уточнення буде зроблено в наступній темі.

Інформації про множину можливих значень недостатньо для повного описання випадкової величини (різні величини можуть мати однакові можливі значення). Потрібно ще знати, з якими імовірностями набираються можливі значення випадковою величиною.

2.

Законом розподілу імовірностей дискретної випадкової величини називається відповідність між можливими значеннями та імовірностями, з якими вони набираються випадковою величиною.

Таблична форма задання закону розподілу має такий вид

$$\begin{array}{c|cccc} X & x_1 & x_2 & \ldots & x_n \\ \hline P & p_1 & p_2 & \ldots & p_n \end{array}, \qquad (4.1)$$

де $p_i = P(X = x_i)$, $i = \overline{1,n}$. Оскільки в одному випробуванні випадкова величина набирає тільки одне із своїх можливих значень, то випадкові події $(X = x_1)$, $(X = x_2)$, ..., $(X = x_n)$ утворюють повну групу. Тому сума їх імовірностей дорівнює одиниці:

$$p_1 + p_2 + \ldots p_n = 1. \qquad (4.2)$$

Ця рівність називається **умовою нормування**.

Якщо множина можливих значень дискретної випадкової величини зчисленна: $x_1, x_2,..., x_n,...$, то ряд $\sum_{t=1}^{\infty} p_i$ збігається і його сума дорівнює одиниці.

Приклад 1. Складальник навмання бере дві деталі із контейнера, в якому знаходиться 20 деталей, серед яких 4 нестандартні. Знайти закон розподілу числа нестандартних деталей серед відібраних.

○ Нехай X — число нестандартних деталей серед двох відібраних. Можливі значення X — 0, 1, 2. Знайдемо відповідні імовірності, використовуючи класичне означення:

$$p_1 = P(X = 0) = m/n = C_{16}^2 / C_{20}^2 = 12/19,$$
$$p_2 = P(X = 1) = C_{16}^1 C_4^1 / C_{20}^2 = 32/95,$$
$$p_3 = P(X = 2) = C_4^2 / C_{20}^2 = 3/95.$$

Перевірка: $p_1 + p_2 + p_3 = 12/19 + 32/95 + 3/95 = 1.$

Шуканий закон розподілу має такий вид:

$$\begin{array}{c|ccc} X & 0 & 1 & 2 \\ \hline P & 12/19 & 32/95 & 3/95 \end{array}. \; \bullet$$

3.

Закон розподілу дискретної випадкової величини X можна задати також **аналітично**, тобто з допомогою формули $p_i = P(X = x_i) = g(x_i)$, $i = \overline{1,n}$. Всі нижче наведені закони задаються аналітично.

Рівномірний розподіл

Приклад 1. У зв’язці є п’ять ключів, з яких тільки один підходить до замка. Скласти закон розподілу числа ключів, що випробовуються при відкриванні замка, якщо ключ, що був у випробуванні, у наступних випробуваннях участі не бере. Знайти імовірність того, що число випробувань не перевищить двох.

○ Позначимо: X — число ключів, що випробовуються при відкриванні замка; A_i — випадкова подія, яка полягає в тому, що i-тий ключ відкриє замок $(i = \overline{1, 5})$. Можливі значення X: 1, 2, 3, 4, 5. Знайдемо відповідні імовірності закону розподілу, попередньо з’ясувавши структуру випадкових подій

$(X = 1)$, $(X = 2)$, ..., $(X = 5)$:

$$(X=1)=A_1,\quad (X=2)=\overline{A}_1A_2,\quad (X=3)=\overline{A}_1\overline{A}_2A_3,$$
$$(X=4)=\overline{A}_1\overline{A}_2\overline{A}_3A_4,\quad (X=5)=\overline{A}_1\overline{A}_2\overline{A}_3\overline{A}_4A_5.$$

Використавши теорему добутку для залежних подій, отримаємо:

$$p_1=1/5,\ \ p_2=P(\overline{A}_1)P_{\overline{A}_1}(A_2)=\frac{4}{5}\cdot\frac{1}{4}=\frac{1}{5},$$
$$p_3=P(\overline{A}_1)P_{\overline{A}_1}(\overline{A}_2)P_{\overline{A}_1\overline{A}_2}(A_3)=\frac{4}{5}\cdot\frac{3}{4}\cdot\frac{1}{3}=\frac{1}{5},$$
$$p_4=\frac{1}{5},\ \ p_5=\frac{1}{5}.$$

Шуканий розподіл має вид

X	1	2	3	4	5
P	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

Випадкова подія $(X \le 2)$ полягає в тому, що число випробувань ключів не перевищить двох. Ясно, що $(X \le 2) = (X = 1) + (X = 2)$. В силу теореми додавання імовірностей для несумісних подій $P(X \le 2) = P(X = 1) + P(X = 2) = 2/5$. ●

Характерною особливістю отриманого закону розподілу є рівність всіх імовірностей. Відповідна випадкова величина називається **рівномірно розподіленою**. В загальному випадку цілочисельна випадкова величина **розподілена за рівномірним законом** (**рівномірно розподілена**), якщо імовірності в законі розподілу мають такий вид: $p_k = P(X=k) = 1/n,\ k=\overline{1,n}.$

Біноміальний розподіл

Приклад 2. Імовірність появи події A в кожному із n повторних випробувань дорівнює p. Скласти закон розподілу випадкової величини X — числа появи події в n випробуваннях.

○ Можливі значення величини X такі: $x_1 = 0$, $x_2 = 1$, $x_3 = 2$, ..., $x_{n+1} = n$. Нехай m будь-яке із цих чисел. Тоді $(X = m)$ — випадкова подія, яка полягає в тому, що в n повторних незалежних випробуваннях ($P(A) = p$ залишається незмінною для кожного випробування) подія A відбудеться рівно m разів. Тоді її імовірність можна позначити $P_n(m)$ (в термінах § 3) і знайти за формулою Бернуллі: $P(X=m)=P_n(m)=C_n^m p^m q^{n-m}$, де $m = 0, 1, 2, ..., n$. Отже, аналітичний вираз закону розподілу імовірностей даної випадкової величини X має такий вид:

$$p_{m+1}=P(X=m)=C_n^m p^m q^{n-m}, \tag{4.3}$$

де $m = 0, 1, 2, ..., n$; $q = 1 - p$. ●

Закон розподілу цілочисельної випадкової величини, імовірності якого знаходяться за формулою Бернуллі (4.3), називається **біноміальним**.

Приклад 3. Імовірність того, що підприємець при перетині кордону декларує не весь товар, дорівнює 0, 4. За зміну митний контроль пройшло 4 підприємці. Скласти закон розподілу числа підприємців, які декларують весь товар.

○ Позначимо: X — число підприємців, які задекларують весь товар при перетині кордону, A — підприємець декларує весь товар. За умовою задачі $p = P(A) = 0{,}6$, $q = 0{,}4$, $n = 4$. Можливі значення X: 0, 1, 2, 3, 4. Відповідні імовірності знайдемо, використовуючи співвідношення (4.3):

$$p_1 = P(X=0) = C_4^0(0{,}6)^0(0{,}4)^4 = 0{,}0256;$$

$$p_2 = P(X=1) = C_4^1(0{,}6)^1(0{,}4)^3 = 0{,}1536;$$

$$p_3 = P(X=2) = C_4^2(0{,}6)^2(0{,}4)^2 = 0{,}3456;$$

$$p_4 = P(X=3) = C_4^3(0{,}6)^3(0{,}4)^1 = 0{,}3456;$$

$$p_5 = P(X=4) = C_4^4(0{,}6)^4(0{,}4)^0 = 0{,}1296.$$

Шуканий закон розподілу є біноміальним і має вид:

X	0	1	2	3	4
P	0,0256	0,1536	0,3456	0,3456	0,1296

. ●

Пуассонівський розподіл

Нехай X — випадкова величина, визначена в задачі 4.3, де n — велике, $p << 0{,}1$, $\lambda = np \le 9$. Тоді імовірності (4.3) в законі розподілу доцільно шукати за формулою Пуассона:

$$p_{m+1} = P(X=m) = \frac{\lambda^m e^{-\lambda}}{m!},\ \ m = 0,\ 1,\ 2,\ ...,\ n,$$

а відповідний розподіл називається пуассонівським.

Зауваження. В пуассонівському розподілі суттєво, що $n \to \infty$.

Геометричний розподіл

Приклад 4. Проводяться повторні випробування, в кожному з яких $P(A) = p > 0$. Випробування закінчуються, як тільки відбудеться подія A. Скласти закон розподілу випадкової величини X — числа випробувань, які потрібно провести до першої появи події A.

○ Можливі значення X — 1, 2, 3, ..., n, ... Подія A може відбутися або в 1-му, або в 2-му, ..., або в n-му, ..., випробуваннях, тобто $(X=1) = A$, $(X=2) = \overline{A}A$, $(X=3) = \overline{A}\overline{A}A$, ..., $(X=n) = \overline{A}\overline{A}...\overline{A}A$, ... Враховуючи незалежність випробувань і теорему множення імовірностей, отримаємо:

$$p_1 = p,\ \ p_2 = qp,\ \ p_3 = q^2 p,\ \ ...,\ \ p_n = q^{n-1}p,\ ... \ . \ \bullet \qquad (4.4)$$

Закон розподілу імовірностей, який визначається послідовністю імовірностей (4.4), називається **геометричним**. Назва закону зумовлена тим, що послідовність (4.4) утворює нескінченну **геометричну прогресію** з першим членом p і знаменником $q \in (0{,}1)$.

Приклад 5. Імовірність того, що виготовлений виріб вимагає додаткового регулювання, дорівнює p. Контролер перевіряє якість партії виробів, навмання вибираючи виріб. Якщо він вимагає додаткового регулювання, то наступні випробування припиняються, а вся партія відправляється на доробку. Якщо ж виріб стандартний, то контролер бере наступний виріб, тощо. Згідно із інструкцією контролер перевіряє не більше п'яти виробів.

1) Скласти закон розподілу числа перевірених контролером виробів. 2) Знайти імовірність доробки всієї партії виробів.

○ 1) Позначимо: X – число виробів, перевірених контролером, A_i – i-тий відібраний виріб вимагає додаткового регулювання ($i=\overline{1,5}$). За умовою $P(A_i)=p,\ i=\overline{1,5}$. Тоді $\overline{A_i}$ – i-тий виріб стандарний. $P(\overline{A_i})=1-p=q,\ i=\overline{1,5}$.

Можливі значення X: 1, 2, 3, 4, 5. Знайдемо імовірності, з якими X набирає ці значення. Випадкова подія (X=1) відбудеться тоді, коли перший відібраний прилад вимагатиме додатково регулювання, тобто (X=1)=A_1, звідки $p_1=P(X=1)=P(A_1)=p$. Відбуття події (X=2) означає, що перший виріб стандартний і другий вимагає регулювання: $(X=2)=\overline{A}_1\cdot A_2$. Використовуючи теорему множення імовірностей, отримаємо: $p_2=P(X=2)=P(\overline{A}_1\cdot A_2)=qp$. Аналогічно знаходимо, що $p_3=P(X=3)=q^2p$, $p_4=P(X=4)=q^3p$.

Нарешті, подія (X=5) відбудеться або тоді, коли чотири перші вироби стандартні, а п'ятий вимагає регулювання (партія відправляється на доробку), або коли всі п'ять виробів стандартні (партія пропукається), тобто

$$(X=5)=\overline{A}_1\overline{A}_2\overline{A}_3\overline{A}_4A_5+\overline{A}_1\overline{A}_2\overline{A}_3\overline{A}_4\overline{A}_5,$$

звідки $p_5=P(X=5)=q^4p+q^5=q^4(p+q)=q^4$, бо $p+q=1$.

Остаточно шуканий закон розподілу має такий вид:

X	1	2	3	4	5
P	p	qp	q^2p	q^3p	q^4

Для перевірки з'ясуємо, чи виконується умова нормування:

$$\sum_1^5 p_i=p+qp+q^2p+q^3p+q^4=p+qp+q^2p+q^3(p+q)=$$

$$=p+qp+q^2(p+q)=p+q(p+q)=p+q=1.$$

2) Партія пропускається контролером в єдиному випадку, коли всі п'ять перевірених виробів є стандартними, імовірність цієї випадкової події дорівнює q^5. Тому імовірність протилежної події (партія виробів відправляється на доробку) дорівнює $1-q^5$ або $1-(1-p)^5$. ●

Гіпергеометричний розподіл

Приклад 6. З партії N виробів, серед яких M стандартних ($M<N$), навмання вибирається n виробів ($n\le M$) без повернення кожного із них.

Знайти закон розподілу випадкової величини X — числа стандартних виробів серед відібраних.

○ Можливі значення X: 0, 1, 2, 3, ..., n. Нехай m одне із цих можливих значень. Тоді випадкова подія ($X = m$) полягає в тому, що серед n відібраних виробів m стандартні і $n - m$ нестандартних. Імовірність цієї події знайдемо за класичним означенням:

$$P(X = m) = C_M^m \cdot C_{N-M}^{n-m} \Big/ C_N^n, \;\; m = 0, 1, 2, ..., n. \quad \bullet \tag{4.5}$$

Закон розподілу імовірностей, який визначається співвідношенням (4.5), називається **гіпергеометричним**.

Приклад 7. У лотереї, яка проводиться з приводу презентації фірми, на 4000 білетів розігруються п’ять речей, вартість яких складає відповідно 100, 300, 400, 500 та 700 грн. Вартість одного білета – 5 грн. Скласти закон розподілу суми чистого виграшу для особи, яка має два білети.

○ Хай випадкова величина X – сума виграшу особи, яка має два лотерейні білети. Знайдемо можливі значення X. Якщо обидва білети невиграшні, тоді особа втрачає 2·5=10, тобто виграш складає –10. Нехай один білет невиграшний, а другий виграшний. Тоді із врахуванням вартості білетів можливі варіанти виграшу: 90, 290, 390, 490 та 690. Нарешті, обидва білети можуть бути виграшні, при цьому можливі варіанти: 100+300-10; 100+400-10; 100+500-10; 100+700-10; 300+400-10; 300+500-10; 300+700-10; 400+500-10; 400+700-10; 500+700-10.

Отже, можливі значення X в порядку розглянутих варіантів такі: а) –10; б)90, 290, 390, 490, 690; в) 390, 490, 590, 790, 690, 790, 990, 890, 1090, 1190.

Відмітимо, що серед них є співпадаючі, а також порушений порядок зростання. Поділ на групи зумовлений тим, що в межах кожної групи імовірності того, що X набере можливе значення, рівні.

Випробування полягає у відборі (придбанні) двох білетів. Всіх наслідків випробування є $C_{2000}^2 = 1999000$. Невиграшних білетів є 1995. Тому

$$P(X = -10) = C_{1995}^2 \big/ C_{2000}^2 = 1989015 \big/ C_{2000}^2 .$$

Кожне із можливих значень групи б) набирається величиною X тоді, коли один білет буде виграшним із конкретною (однією) вартістю виграшу, а другий – невиграшний. Тому

$$P(X = 90) = P(X = 290) = P(X = 390) = P(X = 490) = P(X = 690) =$$

$$= C_1^1 \cdot C_{1995}^1 \big/ C_{2000}^2 = 1995 \big/ C_{2000}^2 .$$

Кожне із можливих значень групи в) величина X набирає, коли обидва білети будуть виграшними, причому пара виграшів буде єдиною (за виключенням можливого значення 790, яке двічі повторюється). Тому

$$P(X = 390) = P(X = 490) = P(X = 590) = P(X = 690) = P(X = 890) =$$

$$= P(X = 990) = P(X = 1090) = P(X = 1190) = 1 \big/ C_{2000}^2 .$$

$$P(X = 790) = 2 \big/ C_{2000}^2 .$$

Тепер врахуємо те, що можливі значення 390, 490, 690 зустрічаються в групах б) та в). Використавши теорему додавання імовірностей, отримаємо:

$$P(X=390)=1995\big/C_{2000}^{2}+1\big/C_{2000}^{2}=1996\big/C_{2000}^{2},$$

$$P(X=490)=P(X=490)=1996\big/C_{2000}^{2}.$$

Таким чином, шуканий закон розподілу імовірностей випадкової величини X має такий вид:

X	10	0	90	90	90	90	90	90	90	90	090	190
P	$\frac{198901}{C_{2000}^{2}}$	$\frac{1995}{C_{2000}^{2}}$	$\frac{1995}{C_{2000}^{2}}$	$\frac{1996}{C_{2000}^{2}}$	$\frac{1996}{C_{2000}^{2}}$	$\frac{1}{C_{2000}^{2}}$	$\frac{1996}{C_{2000}^{2}}$	$\frac{2}{C_{2000}^{2}}$	$\frac{1}{C_{2000}^{2}}$	$\frac{1}{C_{2000}^{2}}$	$\frac{1}{C_{2000}^{2}}$	$\frac{1}{C_{2000}^{2}}$

Додавши всі імовірності і врахувавши, що $C_{2000}^{2}=1999000$, можна переконатися у виконанні умови нормування. ●

Самостійно розглянути випадок $M < n < N$.

4.

Найпростіший потік подій

Потоком подій називається послідовність подій, які відбуваються у випадкові моменти часу.

Приклади потоків: виклики на АТС, диспетчерські швидкої допомоги, аварійні служби газового та водоканалізаційного господарств, прихід клієнтів в банки, послідовність відмов обладнання тощо. Потоки вивчаються в теорії масового обслуговування — історично одній із галузей теорії імовірностей (в теперішній час це самостійна наукова дисципліна).

Найпростішим називається **потік подій,** який має властивості стаціонарності, відсутності післядії та ординарності.

Властивість стаціонарності полягає в тому, що імовірність появи $\boldsymbol{m}$ подій потоку за будь-який проміжок часу довжиною t залежить тільки від числа $\boldsymbol{m}$ та t і не залежить від початку відліку часу; при цьому різні проміжки часу не повинні перетинатися.

Властивість відсутності післядії визначає те, що імовірність появи $\boldsymbol{m}$ подій на довільному проміжку часу не залежить від появи чи не появи подій потоку в моменти часу, що передують початку цього проміжку.

Властивість ординарності полягає в тому, що поява двох і більшого числа подій за малий проміжок часу практично неможлива, тобто за малий проміжок часу може появитись не більше однієї події потоку.

Найпростішим називається **потік подій**, який має властивості **стаціонарності**, відсутності післядії та ординарності.

Інтенсивністю λ потоку називається середнє число подій, які вібуваються за одиницю часу.

Математичною моделлю найпростішого потоку подій є формула Пуассона

$$P_t(m)=(\lambda t)^m e^{-\lambda t}\big/m!, \tag{4.6}$$

з допомогою якої можна знайти імовірність появи m подій найпростішого потоку за проміжок часу довжиною t.

5.

Дії над випадковими величинами

Визначимо **добуток сталої величини C на дискретну випадкову величину X**, задану законом розподілу (4.1), як дискретну випадкову величину CX, закон розподілу якої має вид

$$\begin{array}{c|cccc} CX & Cx_1 & Cx_2 & \dots & Cx_n \\ \hline P & p_1 & p_2 & \dots & p_n \end{array}. \qquad (4.7)$$

Квадрат випадкової величини X, тобто X^2 — це нова дискретна випадкова величина, яка описується законом розподілу

$$\begin{array}{c|cccc} X^2 & x_1^2 & x_n^2 & \dots & x_n^2 \\ \hline P & p_1 & p_2 & \dots & p_n \end{array}. \qquad (4.8)$$

Нехай випадкова величина X задана розподілом $\begin{array}{c|cccc} X & x_1 & x_2 & \dots & x_n \\ \hline P & p_1 & p_2 & \dots & p_n \end{array}$, , а Y — законом розподілу $\begin{array}{c|cccc} Y & y_1 & y_2 & \dots & y_m \\ \hline P & g_1 & g_2 & \dots & g_m \end{array}$. Ці величини називаються **незалежними**, якщо випадкові події $(X = x_i)$, $(Y = y_j)$ незалежні при довільних $i = \overline{1,n}$ та $j = \overline{1,m}$. Поняття незалежності випадкових величин поширюється на довільне скінченне число випадкових величин. В протилежному випадку випадкові величини називаються **залежними**.

Дискретні випадкові величини $X_1, X_2, \dots, X_k$ називаються **незалежними у сукупності** (**взаємно незалежними**), якщо закон розподілу **кожної** із них не змінюється, якщо довільна випадкова величина або будь-які групи цих величин наберуть яке завгодно із своїх можливих значень.

Визначимо **суму випадкових величин** X та Y як випадкову величину $X+Y$, можливі значення якої рівні сумам кожного можливого значення X з кожним можливим значення Y, а імовірності можливих значень $X+Y$ для незалежних величин X та Y дорівнюють добуткам імовірностей доданків; для залежних величин — добуткам імовірностей одного доданку на умовну імовірність другого. Якщо деякі суми $x_i + y_j$ виявляються рівними між собою, тоді імовірність можливого значення суми дорівнює сумі відповідних імовірностей.

Приклад 1. Випадкові величини X та Y задані законами розподілу $\begin{array}{c|cc} X & 2 & 5 \\ \hline P & 0{,}8 & 0{,}2 \end{array}$, $\begin{array}{c|ccc} Y & 1 & 3 & 6 \\ \hline P & 0{,}3 & 0{,}5 & 0{,}2 \end{array}$. Знайти $X+Y$, якщо X та Y незалежні величини.

○ Якщо X та Y задані законами розподілу відповідно $\begin{array}{c|cc} X & x_1 & x_2 \\ \hline P & p_1 & p_2 \end{array}$,

Y	y_1	y_2	y_3
P	g_1	g_2	g_3

, то за означенням випадкова величина $X+Y$ описується законом розподілу

$X+Y$	x_1+y_1	x_1+y_2	x_1+y_3	x_2+y_1	x_2+y_2	x_2+y_3
P	p_1g_1	p_1g_2	p_1g_3	p_2g_1	p_2g_2	p_2g_3

.

Підставивши дані, отримуємо:

$X+Y$	$2+1$	$2+3$	$2+6$	$5+1$	$5+3$	$5+6$
P	$0,8\cdot 0,3$	$0,8\cdot 0,5$	$0,8\cdot 0,2$	$0,2\cdot 0,3$	$0,2\cdot 0,5$	$0,2\cdot 0,2$

.

Можливі значення 2 + 6 та 5 + 3 рівні, тому відповідні імовірності 0,16 та 0,1 додаємо, залишивши в остаточному законі розподілу $X+Y$ одне можливе значення 8:

$X+Y$	3	5	6	8	11
P	0,24	0,4	0,06	0,26	0,04

. ●

Визначимо **добуток незалежних випадкових величин** X та Y як випадкову величину XY, можливі значення якої рівні добуткам кожного можливого значення X на кожне можливе значення Y; імовірності можливих значень добутку XY дорівнюють добуткам імовірностей можливих значень співмножників. У випадку рівності добутків x_iy_j імовірність можливого значення XY рівна сумі відповідних імовірностей.

Наприклад, якщо X та Y визначені розподілами прикладу 1, то закон розподілу XY має такий вид:

XY	2	5	6	12	15	30
P	0,24	0,06	0,4	0,16	0,1	0,04

.

Приклад 2. Дослідження роботи блоків чотирьох типів в умовах перепаду напруги показало, що імовірність безвідмовної роботи на протязі часу T кожного із типів складає відповідно 0,8; 0,9; 0,7 та 0,75. Відбираються чотири блоки кожного типу. Скласти закон розподілу випадкової величини X – числа блоків, які безвідмовно працюватимуть у вказаних умовах.

○ Задачу можна розв’язати не одним способом.

Перший спосіб.

Нехай, події A_i – блок i-го типу працює безвідмовно, $i=\overline{1,4}$,

$\overline{A}_i$ – блок i-го типу вийшов з ладу.

За умовою $P(A_1)$=0,8, $P(A_2)$=0,9, $P(A_3)$=0,7, $P(A_4)$=0,75, $P(\overline{A}_1)=0,2$, $P(\overline{A}_2)=0,1$, $P(\overline{A}_3)=0,3$, $P(\overline{A}_4)=0,25$.

Можливі значення X – 0, 1, 2, 3, 4.

Знайдемо відповідні імовірності:

$P(X=0)=P(\overline{A}_1\overline{A}_2\overline{A}_3\overline{A}_4)=0,2\cdot 0,1\cdot 0,3\cdot 0,25=0,0015$;

$$P(X=1)=P(A_1\overline{A}_2\overline{A}_3\overline{A}_4+\overline{A}_1A_2\overline{A}_3\overline{A}_4+\overline{A}_1\overline{A}_2A_3\overline{A}_4+\overline{A}_1\overline{A}_2\overline{A}_3A_4)=$$
$$0,8\cdot 0,1\cdot 0,3\cdot 0,25+0,2\cdot 0,9\cdot 0,3\cdot 0,25+0,2\cdot 0,1\cdot 0,7\cdot 0,25+0,2\cdot 0,1\cdot 0,3\cdot 0,75=$$
$$=0,006+0,0135+0,0035+0,0045=0,0275;$$
$$P(X=2)=P(A_1A_2\overline{A}_3\overline{A}_4+A_1\overline{A}_2A_3\overline{A}_4+A_1\overline{A}_2\overline{A}_3A_4+\overline{A}_1A_2A_3\overline{A}_4+\overline{A}_1A_2\overline{A}_3A_4+$$
$$\overline{A}_1\overline{A}_2A_3A_4)=0,1685;$$
$$P(X=3)=P(A_1A_2A_3\overline{A}_4+A_1A_2\overline{A}_3A_4+A_1\overline{A}_2A_3A_4+\overline{A}_1A_2A_3A_4)=0,4245;$$
$$P(X=4)=P(A_1A_2A_3A_4)=0,378$$

Остаточно отримаємо шуканий закон розподілу:

X	0	1	2	3	4
P	0,0015	0,0275	0,1685	0,4245	0,378

(*)

Другий спосіб. Позначимо випадкові величини: X_k – число блоків k-го типу, які безвідмовно працюватимуть в умовах перепаду напруги ($k=\overline{1,4}$).

Оскільки тип представлений тільки одним блоком, то X_k може набирати два можливих значення: 0 (блок вийшов з ладу) і 1 (блок безвідмовно працюватиме). Із врахуванням умови задачі отримаємо закони розподілу.

X_1	0	1
P	0,2	0,8

X_2	0	1
P	0,1	0,9

X_3	0	1
P	0,3	0,7

X_4	0	1
P	0,25	0,75

Число (**загальне**) блоків, які безвідмовно працюватимуть у вказаних умовах, складається із числа блоків кожного типу, що витримають випробування:

$\boldsymbol{X=X_1+X_2+X_3+X_4}$.

Позначимо $Y=X_1+X_2$; $Z=Y+X_3$; $X=Z+X_4$.

За аналогією із розв'язуванням прикладу 1 послідовно отримаємо:

$Y=X_1+X_2$	0	1	2
P	0,02	0,26	0,72

$Z=Y+X_3$	0	1	2	3
P	0,006	0,092	0,398	0,504

і, нарешті, закон розподілу для $X=Z+X_4$, який співпадає із (*).

Третій спосіб. Закон розподілу для X можна отримати ще й чисто механічно, перемноживши двочлени (біноми):

$$\Upsilon_4(z)=(0,8z+0,2)(0,9z+0,1)(0,7z+0,3)(0,75z+0,25).$$

В результаті отримаємо

$$\Upsilon_4(z)=0,378z^4+0,4245z^3+0,1685z^2+0,0275z+0,0015.$$

При цьому кожний із п'яти коефіцієнтів при z^k (k=0,1,2,3,4) дорівнює імовірності $P(X=k)$ (порівняйте із (*)). ●

Зауваження. Функція

$$Y_n(z)=\prod_{i=1}^{n}(p_iz+q_i)=(p_1z+q_1)(p_2z+q_2)\cdot\ldots\cdot(p_nz+q_n),$$

коефіцієнти розкладу якої по степеням $z^k (k=0,1,2,\ldots,n)$ дорівнюють імовірностям того, що при випробуванні випадкова величина X набере можливе значення k, називається **твірною функцією величини *X***. Твірна функція використовується також при знаходженні імовірностей $P_n(m)$ для n повторних випробувань таких, що в першому випробуванні випадкова подія A відбувається з імовірністю p_1, в другому – p_2, …, в n-му випробуванні – p_n, а імовірності непояви події A відповідно рівні $q_1,q_2,\ldots,q_n$. В такому випадку $Y_n(z)$ називається **твірною функцією імовірностей $P_n(m)$**.

6.

Числові характеристики дискретних випадкових величин та їх властивості

Закон розподілу імовірностей дає повну інформацію про дискретну випадкову величину. Проте в багатьох практично важливих випадках економісту-досліднику буває достатньо знати одне або кілька чисел, пов'язаних із випадковою величиною, які дають менш повне, але більш наочне описання випадкової величини. Такі числа, які сумарно описують випадкову величину, називаються її **числовими характеристиками**.

Математичне сподівання

Математичним сподіванням дискретної випадкової величини X називається сума добутків всіх її можливих значень на відповідні імовірності:

$$M(X) = x_1p_1 + x_2p_2 + \ldots + x_np_n. \tag{4.9}$$

Якщо величина X може набирати зчисленну множину значень (наприклад, як у випадку геометричного або пуассонівського розподілу), то при умові, що цей ряд абсолютно збіжний,

$$M(X) = \sum_{i=1}^{\infty} x_i p_i. \tag{4.10}$$

Імовірносний зміст математичного сподівання: **$M(X)$ приблизно дорівнює** (тим точніше, чим більше число випробувань) **середньому арифметичному спостережених значень випадкової величини**.

Властивості математичного сподівання.

1. **Математичне сподівання сталої величини дорівнює цій сталій**: $M(C) = C$.

2. **Сталий множник можна виносити за знак математичного сподівання**: $M(CX) = CM(X)$.

3. **Математичне сподівання суми двох випадкових величин дорівнює сумі математичних сподівань доданків**: $M(X + Y) = M(X) + M(Y)$.

Наслідок. Математичне сподівання суми кількох випадкових величин дорівнює сумі математичних сподівань доданків.

4. **Математичне сподівання добутку двох незалежних випадкових величин дорівнює добутку їх математичних сподівань**: $M(XY) = M(X) \cdot M(Y)$.

Наслідок. Математичне сподівання добутку кількох незалежних у сукупності випадкових величин дорівнює добутку їх математичних сподівань.

Приклад 1. Існує три методи експрес-контролю партії виробів. Кожний із них безпомилково виявляє якісні вироби, а при ідентифікації некондиційних допускає помилки, тобто вони визнаються якісними. Число таких виробів для кожного із методів є відповідно випадковими величинами X, Y та Z, закони розподілу яких мають такий вид:

X	0	1	3	4
P	0,5	0,4	0,06	0,04

,

Y	0	1	2	3
P	0,6	0,2	0,1	0,1

,

Z	0	1	2	3	4
P	0,8	0,1	0,04	0,03	0,03

.

Який із методів експрес-контролю кращий?

○ Нехай a — кількість стандартних виробів у партії. Тоді згідно з умовою задачі кожен із методів контролю відповідно дасть $a + X$, $a + Y$, $a + Z$ стандартних деталей після перевірки всієї партії, де кожний другий доданок у сумах — це помилка при ідентифікації бракованих виробів. Мінімальне значення її (у середньому) визначить кращий метод. Характеристикою середнього значення є математичне сподівання. Тому знайдемо

$M(X) = 0 \cdot 0,5 + 1 \cdot 0,4 + 3 \cdot 0,06 + 4 \cdot 0,04 = 0,74;$

$M(Y) = 0 \cdot 0,6 + 1 \cdot 0,2 + 2 \cdot 0,1 + 3 \cdot 0,1 = 0,7;$

$M(Z) = 0 \cdot 0,8 + 1 \cdot 0,1 + 2 \cdot 0,04 + 3 \cdot 0,03 + 4 \cdot 0,03 = 0,39.$

Отже, кращим методом експрес-контролю є третій. ●

Дисперсія

Для оцінки розкиду можливих значень випадкової величини навколо середнього значення (математичного сподівання) використовують, зокрема, числову характеристику, яку називають **дисперсією**.

Дисперсією (розкидом) випадкової величини X називається математичне сподівання квадрату відхилення X від $M(X)$:

$$D(X) = M[X - M(X)]^2. \qquad (4.11)$$

Згідно із цим означенням дисперсія характеризує середню величину розкиду можливих значень випадкової величини навколо її матсподівання (середньої) в квадратних одиницях.

Якщо X розподілена за законом (4.1), то з врахуванням закону (4.8) випадкова величина $[X - M(X)]^2$ має такий закон розподілу:

$[X-M(X)]^2$	$[x_1-M(X)]^2$	$[x_2-M(X)]^2$	...	$[x_n-M(X)]^2$
P	p_1	p_2	...	p_n

.

Використавши означення дисперсії та математичного сподівання, отримаємо формулу для обчислення $D(X)$:

$$D(X) = M[X-M(X)]^2 = [x_1 - M(X)]^2 p_1 + [x_2 - M(X)]^2 p_2 + \\ + \ldots + [x_n - M(X)]^2 p_n = \sum_{i=1}^{n} [x_i - M(X)]^2 p_i. \qquad (4.11^*)$$

Зменшення об'єму обчислень досягається за рахунок використання **розрахункової формули** для обчислення дисперсії:

$$D(X) = M(X^2) - [M(X)]^2. \tag{4.12}$$

Застосувавши означення математичного сподівання до закону розподілу $\begin{array}{c|cccc} X^2 & x_1^2 & x_n^2 & \ldots & x_n^2 \\ \hline P & p_1 & p_2 & \ldots & p_n \end{array}$., отримаємо **числову реалізацію розрахункової формули**:

$$D(X) = \sum_{i=1}^{n} x_i^2 p_i - [M(X)]^2. \tag{4.12*}$$

Властивості дисперсії.

1. **Дисперсія сталої дорівнює нулю**: $D(C) = 0$.

2. **Сталий множник можна виносити за знак дисперсії, піднісши його до квадрату**: $D(CX) = C^2 D(X)$.

3. **Дисперсія суми двох незалежних випадкових величин дорівнює сумі дисперсій цих величин**: $D(X + Y) = D(X) + D(Y)$.

Наслідок. Дисперсія суми кількох незалежних у сукупності випадкових величин дорівнює сумі дисперсій цих величин.

4. **Дисперсія різниці двох незалежних випадкових величин дорівнює сумі їх дисперсій**: $D(X - Y) = D(X) + D(Y)$.

Зауваження. Навіть якщо X та Y незалежні, то в загальному випадку виконується нерівність $D(XY) \neq D(X)\, D(Y)$.

Середнє квадратичне відхилення

Незручність використання дисперсії в деяких випадках зумовлена тим, що вона має розмірність, яка дорівнює квадрату розмірності випадкової величини. Наприклад, якщо можливі значення X вимірюються в кг, то $D(X)$ в (кг)^2.

Середнім квадратичним відхиленням випадкової величини X називається квадратний корінь із дисперсії:

$$\sigma(X) = \sqrt{D(X)}. \tag{4.13}$$

$\sigma(X)$ характеризує середню величину розкиду можливих значень X навколо $M(X)$ (середньої) в лінійних одиницях.

Розглянемо типові задачі.

Приклад 2. Знайти середнє квадратичне відхилення випадкової величини $Z = 2X - 5Y - 30$, якщо X та Y — незалежні випадкові величини, $D(X) = 0{,}25$, $D(Y) = 0{,}2$.

○ Рівність випадкових величин зумовлює рівність їх дисперсій: $D(Z) = D(2X - 5Y - 30)$. Використання властивостей дисперсії та умов задачі дає такий ланцюжок рівностей:

$$D(Z) = D(2X + (-5Y) + (-30)) = D(2X) + D(-5Y) + D(-30) =$$
$$= 2^2 D(X) + (-5)^2 D(Y) + 0 = 4 \cdot 0{,}25 + 25 \cdot 0{,}2 = 6.$$

Тоді згідно із (4.13) $\sigma(Z)=\sqrt{D(Z)}=\sqrt{6}\approx 2{,}45$. •

Приклад 3. Знайти закон розподілу дискретної випадкової величини на основі числових характеристик цієї величини та певної інформації про закон її розподілу:

а) $\begin{array}{c|ccc} X & 3 & 5 & x_3 \\ \hline P & 0{,}1 & 0{,}6 & p_3 \end{array}$, $M(X)=5{,}7$;

б) $\begin{array}{c|ccc} Y & -2 & 1 & 2 \\ \hline P & p_1 & p_2 & p_3 \end{array}$, $M(Y)=0{,}9$ $M(Y^2)=3{,}1$;

в) $\begin{array}{c|cc} Z & z_1 & z_2 \\ \hline P & 0{,}8 & p_2 \end{array}$, $M(Z)=1{,}2$ $\sigma(Z)=0{,}4$.

○ а) Сума імовірностей в законі розподілу дорівнює одиниці, тому $0{,}1+0{,}6+p_3=1$, звідки $p_3=0{,}3$. Згідно з означенням (4.9) математичного сподівання і умовою задачі $M(X)=3\cdot 0{,}1+5\cdot 0{,}6+$ $+x_3\cdot 0{,}3=5{,}7$ або $0{,}3x_3=4$, звідки $x_3=\dfrac{40}{3}$. Знайдений закон розподілу випадкової величини має такий вид: $\begin{array}{c|ccc} X & 3 & 5 & 40/3 \\ \hline P & 0{,}1 & 0{,}6 & 0{,}3 \end{array}$.

б) За аналогією із випадком а) і використавши рівність $M(Y^2)=y_1^2p_1+y_2^2p_2+y_3^2p_3$, отримаємо таку систему рівнянь відносно невідомих p_1, p_2, p_3:

$$\begin{cases} p_1+p_2+p_3=1, \\ -2p_1+p_2+2p_3=0{,}9, \\ 4p_1+p_2+4p_3=3{,}1. \end{cases}$$

Перше рівняння розв'яжемо відносно p_1, а друге помножимо на 2 і додамо до третього. В результаті система набуде такого виду:

$$\begin{cases} p_1=1-p_2-p_3, \\ -2p_1+p_2+2p_3=0{,}9, \\ 3p_2+8p_3=4{,}9. \end{cases} \Leftrightarrow \begin{cases} p_1=1-p_2-p_3, \\ 3p_2+4p_3=2{,}9, \\ 3p_2+8p_3=4{,}9. \end{cases} \Leftrightarrow \begin{cases} p_1=1-p_2-p_3, \\ 3p_2+4p_3=2{,}9, \\ 4p_3=2. \end{cases}$$

Остаточно $p_1=0{,}2$, $p_2=0{,}3$, $p_3=0{,}5$.

в) Оскільки $p_2=1-0{,}8=0{,}2$, то для знаходження двох невідомих z_1 і z_2 складемо систему рівнянь:

$$\begin{cases} 0{,}8z_1+0{,}2z_2=1{,}2, \\ \sqrt{0{,}8z_1^2+0{,}2z_2^2-(1{,}2)^2}=0{,}4. \end{cases} \Leftrightarrow \begin{cases} z_2=6-4z_1, \\ 0{,}8z_1^2+0{,}2z_2^2=1{,}6. \end{cases} \Leftrightarrow \begin{cases} z_2=6-4z_1, \\ z_1^2-2{,}4z_1+1{,}4=0. \end{cases}$$

Звідки $z_1^{(1)}=1$, $z_1^{(2)}=1{,}4$, $z_2^{(1)}=2$, $z_2^{(2)}=0{,}4$.

Отже, умовам задачі задовольняють два закони розподілу імовірностей:

$$\begin{array}{c|cc} Z^{(1)} & 1 & 2 \\ \hline P & 0{,}8 & 0{,}2 \end{array}; \quad \begin{array}{c|cc} Z^{(2)} & 1{,}4 & 0{,}4 \\ \hline P & 0{,}8 & 0{,}2 \end{array}. \quad \bullet$$

Приклад 4. Маршрут руху інкасаторського автомобіля від банку до його філії перетинає п'ять перехресть, регульованих світлофорами, які з імовірностями відповідно 0,2, 0,5, 0,8, 0,4, 0,5 дозволяють рух без зупинки. Знайти середнє число зупинок спецавтомобіля на перехрестях по цьому маршруту, якщо світлофори працюють незалежно один від інших (відсутня «зелена лінія»).

○ Позначимо: X — число зупинок автомобіля на перехрестях по всьому маршруту, X_i — число зупинок при проходженні i-ого перехрестя, де $i = 1, 2, ..., 5$. Тоді загальне число зупинок буде складатися із числа зупинок на кожному світлофорі: $X = X_1 + X_2 + X_3 + X_4 + X_5$. За властивістю математичного сподівання $M(X) = \sum_{i=1}^{5} M(X_i)$, при цьому $M(X)$ характеризує середнє число зупинок по всьому маршруту. Можливі значення X_1: 0, 1. Випадкова подія $(X_1 = 1)$ полягає в тому, що на першому світлофорі автомобіль зупинився, а $(X_1 = 0)$ — не зупинився. За умовою задачі $P(X_1 = 0) = 0{,}2$, тоді $P(X_1 = 1) = 1 - 0{,}2 = 0{,}8$. Отже, закон розподілу випадкової величини X_1 має такий вид: $\begin{array}{c|cc} X_1 & 0 & 1 \\ \hline P & 0{,}2 & 0{,}8 \end{array}$, звідки $M(X_1) = 0 \cdot 0{,}2 + 1 \cdot 0{,}8 = 0{,}8$. Аналогічно знаходимо: $M(X_2) = 0{,}5$, $M(X_3) = 0{,}2$, $M(X_4) = 0{,}6$, $M(X_5) = 0{,}5$. Остаточно $M(X) = 0{,}8 + 0{,}5 + 0{,}2 + 0{,}6 + 0{,}5 = 2{,}6$. ●

Приклад 5. Автоматична лінія після сервісної наладки випускає стандартні вироби з імовірністю 0,98. Чергова наладка здійснюється після появи першого нестандартного виробу. Знайти середнє число виробів, виготовлених між двома черговими сервісними наладками лінії.

○ Нехай випадкова величина X — число виробів, виготовлених лінією між її двома черговими сервісними наладками. Всі ці вироби стандартні за виключенням останнього нестандартного, поява якого і зумовлює здійснення чергової наладки при умові, що перший виріб не виявився бракованим. Позначимо: A — поява нестандартного виробу. За умовою задачі $p = P(A) = 1 - 0{,}98 = 0{,}02$ залишається незмінною для кожного із випробувань (виготовлення виробу автоматичною лінією). Тоді можна стверджувати (див. задачу 4.5), що X розподілена за геометричним законом і її закон розподілу має такий вид:

$$\begin{array}{c|cccccc} X & 1 & 2 & 3 & \dots & n & \dots \\ \hline P & p & qp & q^2 p & \dots & q^{n-1} p & \dots \end{array},$$

де $p = 0{,}02$, $q = 0{,}98$.

Шукане середнє число виробів, виготовлених між двома черговими наладками лінії, дорівнює $M(X)$. Використовуючи формулу (4.10), а також те, що $1 + q + q^2 + ... + q^n + ... = 1/(1 - q)$, отримаємо

$$M(X) = p + 2qp + 3q^2 p + \ldots + nq^{n-1} p + \ldots = p\sum_{n=1}^{\infty} nq^{n-1} =$$

$$= p\left(\frac{d}{dx}\sum_{n=1}^{\infty} x^n\right)\Bigg|_{x=q} = p\left(\frac{d}{dx}\left(\frac{1}{1-x}\right)\right)\Bigg|_{x=q} = p\cdot\frac{1}{(1-q)^2} = \frac{1}{p} = \frac{1}{0{,}02} = 50.$$

Відмітимо, що правомірність почленного диференціювання степеневого ряду забезпечується його рівномірною збіжністю, оскільки він є сумою нескінченної геометричної прогресії із знаменником $q = 0{,}98 < 1$. ●

Початкові та центральні моменти

Узагальненням розглянутих вище основних числових характеристик випадкової величини, які описують середину розподілу (математичне сподівання) і розкид (дисперсія та середнє квадратичне відхилення), є поняття моментів випадкової величини. В теорії імовірностей розглядаються моменти двох видів: початкові та центральні.

Початковим моментом порядку *k* дискретної випадкової величини X називається математичне сподівання величини X^k:

$$\nu_k = M(X^k) = \sum x_i^k p_i. \tag{4.14}$$

Зокрема $\nu_1 = M(X)$, $\nu_2 = M(X^2)$, тому розрахункова формула для обчислення дисперсії набирає такого виду:

$$D(X) = \nu_2 - \nu_1^2.$$

Крім моментів самої величини X доцільно (особливо для неперервної величини) розглядати моменти відхилення $X - M(X)$.

Центральним моментом порядку *k* називається математичне сподівання величини $(X - M(X))^k$:

$$\mu_k = M[(X - M(X))^k] = \sum [x_i - M(X)]^k p_i. \tag{4.15}$$

Зокрема, $\mu_1 = M[X - M(X)] = 0$, $\mu_2 = M[(X - M(X))^2] = D(X)$.

Зв'язок між початковими та центральними моментами отримуються із таких співвідношень:

$$\mu_2 = \nu_2 - \nu_1^2,\ \ \mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3,\ \ \mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4.$$

Моменти більш високих порядків використовуються рідко.

Зауваження. Назва моменту центральним зумовлена розглядом степенів відхилень від $M(X)$ — «центру» можливих значень, а початковим — відхилень від нуля ($M(X) = M(X - 0)$, або «початку».

Приклад 6. Дискретна випадкова величина X задана законом розподілу

X	1	3
P	0,6	0,4

. Знайти початкові та центральні моменти першого, другого і третього порядків.

○ За формулами (4.14) для $k = 1, 2, 3$ отримуємо початкові моменти 1-го, 2-го та 3-го порядків:

$\nu_1 = M(X) = 1\cdot 0{,}6 + 3\cdot 0{,}4 = 1{,}8,$

$\nu_2 = 1^2 \cdot 0,6 + 3^2 \cdot 0,4 = 4,3,$

$\nu_3 = 1^3 \cdot 0,6 + 3^3 \cdot 0,4 = 11,4.$

Центральні моменти цих порядків знаходимо за допомогою формул (4.15):

$\mu_1 = M[X - M(X)] = 0,$

$\mu_2 = (1-1,8)^2 \cdot 0,6 + (3-1,8)^2 \cdot 0,4 = 0,96,$

$\mu_3 = (1-1,8)^3 \cdot 0,6 + (3-1,8)^3 \cdot 0,4 = 0,9984.$ ●

7.

Числові характеристики основних законів розподілу

Приклад 1. Знайти числові характеристики випадкової величини, розподіленої за **біноміальним законом**.

○ Дана випадкова величина X — число появи події A в n повторних незалежних випробуваннях, в кожному з яких $P(A) = p$ (див. питання 4.3). Нехай X_k — число появ події A в одному k-му випробуванні $(k = 1,2...,n)$. Тоді $X = X_1 + X_2 + ... + X_n$, а закон розподілу X_k має такий вид:

$$\begin{array}{c|cc} X_k & 0 & 1 \\ \hline P & q & p \end{array}, \quad k = 1,2...,n.$$

Знайдемо числові характеристики величини X_k:

$$M(X) = \sum_{i=1}^{2} x_i p_i = 0 \cdot q + 1 \cdot p = p,$$

$$D(X) = \sum_{i=1}^{n} [x_i - M(X)]^2 p_i = (0-p)^2 q + (1-p)^2 p = p^2 q + q^2 p = pq(p+q) = pq,$$

оскільки $p+q=1$.

Використавши відповідні властивості математичного сподівання і дисперсії (при цьому враховується незалежність випадкових величин X_1, X_2, ..., X_n), отримаємо:

$$M(X) = M(X_1 + X_2 + ... + X_n) = \sum_{k=1}^{n} M(X_k) = \sum_{k=1}^{n} p = np,$$

$$D(X) = D(X_1 + X_2 + ... + X_n) = \sum_{k=1}^{n} D(X_k) = \sum_{k=1}^{n} pq = npq.$$

Отже,

$$M(X) = np, \quad D(X) = npq. \; \bullet \qquad (4.16)$$

Використання отриманих формул ілюструє

Приклад 2. Фірма розглядає питання про об’єм запасів деталей одного типу на пункті гарантійного обслуговування. Статистично встановлено, що імовірність виходу з ладу протягом гарантійного терміну деталі цього типу дорівнює 0,1. Скільки в середньому потрібно мати на пункті деталей такого типу і яка величина розкиду можливих кількостей, якщо в регіоні обслуговування пункту зареєстровано 180 виробів, комплектуючою яких є ця деталь?

○ Нехай A — вихід з ладу деталі протягом гарантійного терміну, випробування — робота деталі протягом цього часу (в складі виробу). Число випробувань $n = 180$. Вони є незалежними, бо $P(A) = 0{,}1$ для кожного із випробувань. Нехай X — сумарне число деталей, які вийдуть із ладу. Тоді середнє число деталей, які потрібно замінити, визначається математичним сподіванням величини X. Згідно з (4.16) $M(X) = 0{,}1 \cdot 180 = 18$. Величину розкиду доцільно характеризувати середнім квадратичним відхиленням, оскільки одиниці вимірювання його — лінійні. Тому за другою формулою (4.16) $\sigma(X) = \sqrt{D(X)} = \sqrt{npq} = \sqrt{16{,}2} \approx 4{,}02$. Самостійно перевірте, що найімовірніше число деталей, які вийдуть із ладу за гарантійний термін, також дорівнює 18. ●

Приклад 3. Знайти математичне сподівання і дисперсію випадкової величини, розподіленої за законом Пуассона.

○ Згідно із означенням величина X розподілена за законом Пуассона, якщо вона набирає можливі значення 0, 1, 2, ..., m, ... з імовірностями

$$p_{m+1} = P(X = m) = \frac{\lambda^m e^{-\lambda}}{m!}, \quad m = 0,\ 1,\ 2,\ \ldots\ .$$

Використавши рівності $m! = (m-1)! \cdot m$; 0!=1 та розклад

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \ldots + \frac{\lambda^m}{m!} + \ldots,$$

знайдемо математичне сподівання:

$$M(X) = \sum_{i=1}^{\infty} x_i p_i = \sum_{m=0}^{\infty} m \cdot \frac{\lambda^m e^{-\lambda}}{m!} = \sum_{m=1}^{\infty} \frac{\lambda^m}{(m-1)!} \cdot e^{-\lambda} =$$

$$= \lambda \cdot e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} = \lambda \cdot e^{-\lambda} (1 + \lambda + \frac{\lambda^2}{2!} + \ldots) = \lambda \cdot e^{-\lambda} e^{\lambda} = \lambda.$$

Для обчислення дисперсії використаємо розрахункову формулу (4.12): $D(X) = M(X^2) - [M(X)]^2$. Тоді

$$M(X^2) = \sum_{i=1}^{\infty} x_i^2 p_i = \sum_{m=0}^{\infty} m^2 \cdot \frac{\lambda^m e^{-\lambda}}{m!} = e^{-\lambda} \sum_{m=1}^{\infty} m \frac{\lambda^m}{(m-1)!} =$$

$$= e^{-\lambda} \sum_{m=1}^{\infty} \frac{[(m-1)+1]}{(m-1)!} \cdot \lambda^m = e^{-\lambda} \sum_{m=2}^{\infty} \frac{\lambda^2 \lambda^{m-2}}{(m-2)!} + e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda \cdot \lambda^{m-1}}{(m-1)!} =$$

$$= (e^{-\lambda}\lambda^2 + e^{-\lambda}\lambda)(1 + \lambda + \frac{\lambda^2}{2!} + \ldots) = (\lambda^2 + \lambda) \cdot e^{-\lambda} e^{\lambda} = \lambda^2 + \lambda;$$

$$D(X) = (\lambda^2 + \lambda) - \lambda^2 = \lambda\ .\ ●$$

Приклад 4. Знайти дисперсію рівномірно розподіленої випадкової величини.

○ Рівномірний закон розподілу має такий вид:

X	1	2	3	...	5
P	$1/n$	$1/n$	$1/n$	...	$1/n$

.

При знаходженні числових характеристик використаємо формули:

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2},\ \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

За формулами (4.9) і (4.12*)

$$M(X) = \sum_{k=1}^{n} x_k p_k = \sum_{k=1}^{n} k \cdot \frac{1}{n} = \frac{1}{n}\sum_{k=1}^{n} k = \frac{n+1}{2},$$

$$D(X) = \sum_{k=1}^{n} x_k^2 p_k - [M(X)]^2 = \sum_{k=1}^{n} k^2 \cdot \frac{1}{n} - (\frac{n+1}{2})^2 = \frac{1}{n}\cdot\frac{n(n+1)(2n+1)}{6} -$$

$$-\frac{(n+1)^2}{4} = (n+1)(\frac{2n+1}{6} - \frac{n+1}{4}) = \frac{n^2-1}{12}. \bullet$$

Приклад 5. Обчислити дисперсію випадкової величини, розподіленої за геометричним законом .

○ Згідно із означенням закон розподілу випадкової величини, геометрично роподіленої, має вид

X	1	2	3	...	n	...
P	p	qp	q^2p	...	$q^{n-1}p$	...

.

В задачі 4.18 отримана формула $M(X) = 1/p$.

Тоді за розрахунковою формулою (4.12)

$$D(X) = \sum_{k=1}^{\infty} x_k^2 p_k - [M(X)]^2 = \sum_{k=1}^{\infty} k^2 q^{k-1} p - \frac{1}{p^2} = p\sum_{k=1}^{\infty} k^2 q^{k-1} - \frac{1}{p^2}.$$

Обчислимо суму отриманого ряду:

$$\sum_{k=1}^{\infty} k^2 q^{k-1} = 1 + 2^2 q + 3^2 q^2 + \ldots + n^2 q^{n-1} + \ldots = (x + 2x^2 + 3x^3 + \ldots + nx^n + \ldots)'\Big|_{x=q} =$$

$$= \left[x(1 + 2x + 3x^2 + \ldots + nx^{n-1} + \ldots)\right]'\Big|_{x=q} = \left[x(x + x^2 + x^3 + \ldots + x^n + \ldots)'\right]'\Big|_{x=q} =$$

$$= \left[x\left(\frac{x}{1-x}\right)'\right]'\Big|_{x=q} = \left[\frac{x}{(1-x)^2}\right]'\Big|_{x=q} =$$

$$= \frac{(1-x)^2 + x\cdot 2(1-x)}{(1-x)^4}\Big|_{x=q} = \frac{1+x}{(1-x)^3}\Big|_{x=q} = \frac{1+q}{(1-q)^3} = \frac{1+q}{p^3}.$$

Тоді остаточно отримаємо

$$D(X) = p\cdot\frac{1+q}{p^3} - \frac{1}{p^2} = \frac{1}{p^2}(1+q-1) = \frac{q}{p^2}. \bullet$$

Рекомендована література: [1], С. 76–126; [2], С. 64–100; [4], С. 68–90; [5], С. 90–115.

Лекція 5.

Тема: ***Неперервні випадкові величини та їх числові характеристики.***

1. ***Функція розподілу ймовірностей і її властивості.***
2. ***Густина розподілу ймовірностей та її властивості.***
3. ***Числові характеристики неперервних випадкових величин.***

1.

Функцією розподілу імовірностей називається функція $F(x)$ детермінованого (невипадкового) аргументу x, яка чисельно дорівнює імовірності того, що в результаті випробування випадкова величина X набере значення, менше від x, тобто

$$F(x) = P(X < x). \tag{5.1}$$

Іноді функцію розподілу називають ще інтегральною.

Уточнимо означення неперевної випадкової величини: випадкову величину називають **неперервною**, якщо її функція розподілу імовірностей є неперервною на області її визначення, а похідна від функції розподілу неперервна у всіх точках, за виключенням, можливо, скінченного числа точок на довільному скінченному інтервалі.

Властивості функції розподілу імовірностей

Властивість 1. Область визначення функції розподілу — $R^1 = (-\infty;\infty)$, **а область значень — відрізок [0; 1].**

Властивість 2. $\boldsymbol{F(x)}$ **— неспадна функція, тобто для будь-якої пари чисел** $\boldsymbol{x_1, x_2}$ **з нерівності** $\boldsymbol{x_2 > x_1}$ **випливає нерівність** $\boldsymbol{F(x_2) \geq F(x_1)}$**.**

Наслідок 1. Імовірність того, що випадкова величина при випробуванні набере можливого значення з проміжку [*a*; *b*), дорівнює приросту функції розподілу на цьому проміжку:

$$P(a \leq X < b) = F(b) - F(a). \tag{5.2}$$

Наслідок 2. Імовірність того, що неперервна випадкова величина *X* набере при випробуванні одне конкретне можливе значення, дорівнює нулю.

Зауваження. Рівність $P(X = x_1) = 0$, де x_1 — конкретне можливе значення величини X, не означає, що подія $X = x_1$ є неможливою (на відміну від класичного означення імовірності). Нагадайте геометричне означення імовірності.

Наслідок 2 з використанням формули (5.2) та теореми додавання імовірностей дозволяє отримати такий ланцюжок рівностей:

$$\begin{aligned} P(a < X < b) &= \big(P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b)\big) = \\ &= F(b) - F(a). \end{aligned} \tag{5.3}$$

Властивість 3. Якщо всі можливі значення випадкові величини

належать відрізку [a; b], то $F(x) = 0$ для всіх $x \le a$ і $F(x) = 1$ для всіх $x > b$.

Наслідок. Якщо можливими значеннями неперервної випадкової величини є всі дійсні числа, тоді мають місце такі граничні співвідношення:

$$\lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to \infty} F(x) = 1.$$

2.

Густиною розподілу імовірностей неперервної випадкової величини X називається функція $f(x)$, яка дорівнює першій похідній від функції розподілу $F(x)$:

$$f(x) = F'(x). \tag{5.4}$$

Із означення (5.4) слідує, що $F(x)$ є первісною для густини розподілу. Нижче буде наведено формулу для знаходження $F(x)$, якщо відома функція $f(x)$. Таким чином, задання однієї із функцій $f(x)$ та $F(x)$ дозволяє знайти іншу. Тому в літературі ці функції називають **законами розподілу неперервної випадкової величини**.

Властивості густини розподілу

Властивість 1. **Область визначення функції $f(x)$ – R^1, а область значень проміжок $[0, \infty)$.**

Властивість 2. Невласний інтеграл від густини розподілу в межах від $-\infty$ до ∞ дорівнює одиниці:

$$\int_{-\infty}^{\infty} f(x)dx = 1. \tag{5.5}$$

Рівність (5.5) називають **умовою нормування** неперервної випадкової величини.

Зауваження. Якщо $[\alpha, \beta]$ - мінімальний проміжок, в якому містяться **всі** можливі значення неперервної випадкової величини, тоді умова нормування набирає такого виду:

$$\int_{\alpha}^{\beta} f(x)dx = 1. \tag{5.5*}$$

Мають місце такі формули:

$$P\left(a \overset{(\le)}{<} X \overset{(\le)}{<} b \right) = \int_a^b f(x)dx, \tag{5.6}$$

$$F(x) = P(X < x) = P(-\infty < X < x) = \int_{-\infty}^{x} f(x)dx. \tag{5.7}$$

Відмітимо, що рівність (5.5) є аналогом рівності одиниці суми імовірностей із закону розподілу дискретної випадкової величини.

З’ясуємо **імовірносний зміст** густини розподілу імовірностей. Означення $f(x)$ (5.4) та похідної функції, а також рівність (5.3) дають такі рівності

$$f(x)=\lim_{\Delta x\to 0}\frac{F(x+\Delta x)-F(x)}{\Delta x}=\lim_{\Delta x\to 0}\frac{P(x<X<x+\Delta x)}{\Delta x}, \qquad (5.8)$$

які дозволяють записати наближену рівність

$$P(x<X<x+\Delta x)\approx f(x)\Delta x. \qquad (5.9)$$

з точністю до нескінченно малих вищого порядку відносно Δx.

Права частина (5.8) пояснює назву функції $f(x)$ за аналогією із змістом густини маси, розподіленої на відрізку, а наближена рівність (5.9) — «вклад» величини $f(x)$ у величину імовірності події $(x<X<x+\Delta x)$ для кожного відрізка довжиною Δx.

Приклад 1. Випадкова величина X задана густиною розподілу

$$f(x)=\begin{cases}0 & \text{при } x\le -\pi/2,\\ a\cos x & \text{при } -\pi/2<x\le \pi/2,\\ 0 & \text{при } x>\pi/2.\end{cases}$$

Знайти параметр a і функцію розподілу імовірностей, а також імовірність того, що при випробуванні величина X набере значення з інтервалу $(-\pi/4,\ \pi/3)$.

○ Параметр a знайдемо з рівності (5.6):

$$\int_{-\infty}^{\infty} f(x)dx=1.$$

Розіб'ємо невласний інтеграл на три інтеграли у відповідності із інтервалами задання густини розподілу:

$$\int_{-\infty}^{\infty} f(x)dx=\int_{-\infty}^{-\pi/2} f(x)dx+\int_{-\pi/2}^{\pi/2} f(x)dx+\int_{\pi/2}^{\infty} f(x)dx=$$

$$=\int_{-\infty}^{-\pi/2} 0dx+\int_{-\pi/2}^{\pi/2} a\cos x dx+\int_{\pi/2}^{\infty} 0dx=a\int_{-\pi/2}^{\pi/2}\cos x dx=$$

$$=a\sin x\Big|_{-\pi/2}^{\pi/2}=a(\sin(\pi/2)-\sin(-\pi/2))=2a.$$

Отже, параметр a задовільняє рівнянню $2a=1$, звідки $a=0{,}5$. При знаходженні $F(x)$ врахуємо це значення параметра і використаємо рівність (5.7). Відмітимо, що вид густини розподілу в даній задачі вказує на те, що функція розподілу на трьох проміжках також буде задаватися різними аналітичними виразами. В зв'язку із цим нехай $x\le -\pi/2$. Тоді

$$F(x)=\int_{-\infty}^{x} f(x)dx=\int_{-\infty}^{x} 0dx=0.$$

Для $x\in(-\pi/2;\ \pi/2]$

$$F(x)=\int_{-\infty}^{x} f(x)dx=\int_{-\infty}^{-\pi/2} f(x)dx+\int_{-\pi/2}^{x} f(x)dx=\int_{-\infty}^{-\pi/2} 0dx+\int_{-\pi/2}^{x} 0{,}5\cos x dx=$$

$$=0{,}5\sin x\Big|_{-\pi/2}^{x}=0{,}5(\sin x-\sin(-\pi/2))=0{,}5(1+\sin x).$$

Нарешті, якщо $x>\pi/2$, то

$$F(x)=\int_{-\infty}^{x} f(x)dx=\int_{-\infty}^{-\pi/2} 0dx+\int_{-\pi/2}^{\pi/2} 0,5\cos x dx+\int_{\pi/2}^{x} 0dx=0,5\sin x\Big|_{-\pi/2}^{\pi/2}=1.$$

Таким чином, функція розподілу для даної величини X має такий вигляд:

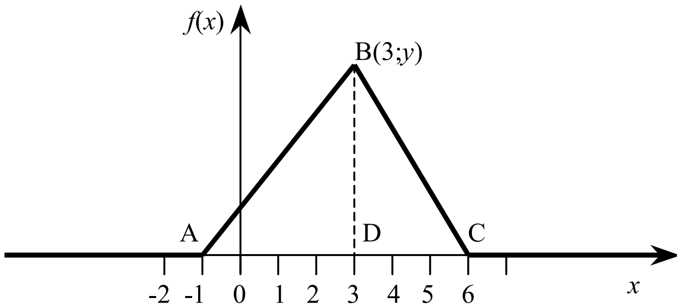
$$F(x)=\begin{cases} 0 & \text{при } x\le -\pi/2,\\ 0,5(1+\sin x) & \text{при } -\pi/2<x\le \pi/2,\\ 1 & \text{при } x>\pi/2. \end{cases}$$

Використаємо формулу (5.6) для знаходження шуканої імовірності:

$$P(-\pi/4<X<\pi/3)=\int_{-\pi/4}^{\pi/3} f(x)dx=\int_{-\pi/4}^{\pi/3} 0,5\cos x dx=0,5\sin x\Big|_{-\pi/4}^{\pi/3}=$$

$$=0,5(\sin(\pi/3)-\sin(-\pi/4))=0,5\left(\sqrt{3}/2+\sqrt{2}/2\right)=\left(\sqrt{3}+\sqrt{2}\right)/4. \ \bullet$$

Приклад 2. Графік густини розподілу неперервної випадкової величини X зображено на мал.5.3. Знайти аналітичні вирази густини розподілу та функції розподілу імовірностей цієї величини. Побудувати графік $F(x)$ і обчислити $P(1<X<4)$, використавши інформацію про $F(x)$ та $f(x)$.



Мал.5.3.

○ Знайдемо ординату точки B, використовуючи умову нормування (5.5*):

$$\int_{-1}^{6} f(x)dx=S_{\Delta ABC}=\frac{1}{2}AC\cdot BD=\frac{1}{2}7y=1,$$

звідки $y=\dfrac{2}{7}$.

На проміжку $[-1,3]$ $f(x)$ дорівнює ординаті відрізка прямої, що проходить через точки $A(-1;0)$ та $B(3;2/7)$. Знайдемо рівняння прямої AB:

$$\frac{y-y_A}{y_B-y_A}=\frac{x-x_A}{x_B-x_A}\Rightarrow\frac{y-0}{\frac{2}{7}-0}=\frac{x-(-1)}{3-(-1)}\Rightarrow y=\frac{x+1}{14}.$$

Отже, $f(x)=\dfrac{x+1}{14}$, якщо $x\in[-1;3]$.

На проміжку $[3;6]$ $f(x)$ дорівнює ординаті відрізка прямої, що проходить через точки $B(3;\frac{2}{7})$ і $C(6;0)$.

Рівняння прямої BC:

$$\frac{y-y_B}{y_B-y_B}=\frac{x-x_B}{x_B-x_B}\Rightarrow\frac{y-\frac{2}{7}}{0-\frac{2}{7}}=\frac{x-3}{6-3}\Rightarrow y=\frac{2(6-x)}{21}.$$

Отже, $f(x)=\frac{2}{21}(6-x)$, якщо $x\in[3;6]$.

В підсумку отримаємо такий аналітичний вираз для густини роподілу:

$$f(x)=\begin{cases}0\ , & \text{при } x\le -1,\\ \frac{x+1}{14}, & \text{при } -1<x\le 3,\\ \frac{2(6-x)}{21}, & \text{при } 3<x\le 6,\\ 0\ , & \text{при } x>6.\end{cases}\tag{5.16}$$

Для знаходження функції розподілу $F(x)$ використаємо формулу (5.7). Нехай $x\le -1$, тоді $f(x)$=0 і $F(x)=\int\limits_{-\infty}^{x}f(x)dx=\int\limits_{-\infty}^{x}0dx=0.$

Якщо $x\in(-1;3]$, то

$$F(x)=\int\limits_{-\infty}^{x}f(x)dx=\int\limits_{-\infty}^{-1}f(x)dx+\int\limits_{-1}^{x}f(x)dx=F(-1)+\int\limits_{-1}^{x}\frac{x+1}{14}dx=$$

$$=0+\frac{1}{14}\int\limits_{-1}^{x}(x+1)d(x+1)=\frac{1}{28}(x+1)^2\Big|_{-1}^{x}=\frac{1}{28}(x+1)^2.$$

Для $x\in(3;6]$

$$F(x)=\int\limits_{-\infty}^{x}f(x)dx=\int\limits_{-\infty}^{3}f(x)dx+\int\limits_{3}^{x}f(x)dx=F(3)+\int\limits_{3}^{x}\frac{2(6-x)}{21}dx=$$

$$=\frac{1}{28}(3+1)^2-\frac{2}{21}\int\limits_{3}^{x}(6-x)d(6-x)=\frac{4}{7}-\frac{1}{21}(6-x)^2\Big|_{3}^{x}=$$

$$=\frac{4}{7}-\frac{1}{21}(6-x)^2+\frac{3}{7}=1-\frac{(6-x)^2}{21}.$$

Нарешті, якщо $x>6$, то

$$F(x)=\int\limits_{-\infty}^{x}f(x)dx=\int\limits_{-\infty}^{6}0dx+\int\limits_{6}^{x}f(x)dx=F(6)+\int\limits_{6}^{x}0dx=1-\frac{(6-6)^2}{21}+0=1.$$

Таким чином, функція розподілу для даної випадкової величини X має такий вигляд:

$$F(x)=\begin{cases}0\ , & \text{якщо } x\le -1,\\ \dfrac{1}{28}(x+1)^2, & \text{якщо } -1<x\le 3,\\ 1-\dfrac{(6-x)^2}{21}, & \text{якщо } 3<x\le 6,\\ 1\ , & \text{якщо } x>6.\end{cases}$$

Графік функції $F(x)$ зображено на мал.5.4.

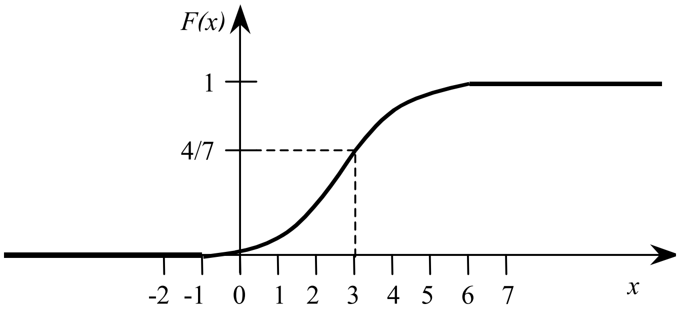
Згідно з формулою (5.3)

$P(1<X<4)=F(4)-F(1)$.

Значення аргументів 4 і 1 належать відповідно третьому та другому інтервалам, тому

$$F(4)=1-\frac{(6-4)^2}{21}=1-\frac{4}{21}=\frac{17}{21},$$

$$F(1)=\frac{1}{28}(1+1)^2=\frac{1}{7},$$



Мал.5.4.

$$P(1<X<4)=\frac{17}{21}-\frac{1}{7}=\frac{14}{21}=\frac{2}{3}.$$

Цю же саму імовірність знайдемо з допомогою формули (5.6):

$$P(1<X<4)=\int_1^4 f(x)dx=\int_1^3 f(x)dx+\int_3^4 f(x)dx=$$

$$=\int_1^3\frac{(x+1)}{14}dx+\int_3^4\frac{2(6-x)}{21}dx=\frac{1}{14}\int_1^3(x+1)d(x+1)-\frac{2}{21}\int_3^4(6-x)d(6-x)=$$

$$=\frac{1}{14}\frac{(x+1)^2}{2}\Bigg|_1^3-\frac{2}{21}\frac{(6-x)^2}{2}\Bigg|_3^4=\frac{1}{28}[(3+1)^2-(1+1)^2]-\frac{1}{21}[(6-4)^2-(6-3)^2]=$$

$$=\frac{3}{7}+\frac{5}{21}=\frac{2}{3}.\ \bullet$$

3.

Математичним сподіванням неперервної випадкової величини X, всі можливі значення якої належать скінченному відрізку $[a;\ b]$, називається визначений інтеграл

$$M(X)=\int_a^b xf(x)dx. \tag{5.10}$$

Якщо можливі значення X належать R^1, тоді

$$M(X)=\int_{-\infty}^{\infty} xf(x)dx, \tag{5.10*}$$

де за припущенням невласний інтеграл збігається абсолютно.

Як і для випадку дискретної величини, **дисперсією неперервної випадкової величини** називається математичне сподівання квадрату її відхилення від математичного сподівання. Із врахуванням означення $M(X)$:

$$D(X)=\int_a^b \left[x-M(X)\right]^2 f(x)dx, \tag{5.11}$$

якщо всі можливі значення X належать скінченному відрізку $[a; b]$, і

$$D(X)=\int_{-\infty}^{\infty} \left[x-M(X)\right]^2 f(x)dx, \tag{5.11*}$$

якщо можливі значення X заповнюють R^1.

Розрахункові формули для обчислення дисперсії:

$$D(X)=\int_a^b x^2 f(x)dx-\left[M(X)\right]^2, \tag{5.12}$$

$$D(X)=\int_{-\infty}^{\infty} x^2 f(x)dx-\left[M(X)\right]^2. \tag{5.12*}$$

Середнє квадратичне відхилення неперервної випадкової величини визначається, як і для випадку дискретної, рівністю

$$\sigma(X)=\sqrt{D(X)}. \tag{5.13}$$

Модою $Mo(X)$ дискретної випадкової величини X називається те її можливе значення, якому відповідає найбільша імовірність її появи. Для неперервної випадкової величини X модою $Mo(X)$ називається можливе значення X, якому відповідає локальний максимум густини розподілу. Випадкова величина може мати кілька мод. У цьому випадку вона називається **многомодальною**. Зустрічаються також розподіли, що не мають максимуму. Такі розподіли називаються **антимодальними**.

Медіаною $Me(X)$ неперервної випадкової величини Х називається те її можливе значення, для якого виконується рівність $P(X < Me(X)) = P(X > Me(X))$.

Початковим моментом *k*-ого порядку неперервної випадкової величини X, заданої густиною розподілу $f(x)$, називається математичне сподівання величини X^k, тобто

$$\nu_k=\int_{-\infty}^{\infty} x^k f(x)dx, \quad k=1, 2, 3, \ldots . \tag{5.14}$$

Центральним моментом *k*-ого порядку неперервної випадкової величини X називається математичне сподівання величини $[X - M(X)]^k$:

$$\mu_k = \int_{-\infty}^{\infty} [x - M(X)]^k f(x)dx, \ \ k = 1, 2, 3, \dots . \tag{5.15}$$

Зокрема, якщо всі можливі значення X належать проміжку (a, b), то

$$\nu_k = \int_a^b x^k f(x)dx, \qquad \mu_k = \int_a^b [x - M(X)]^k f(x)dx.$$

Очевидно, що $\nu_1 = M(X), \ \mu_1 = 0, \ \mu_2 = D(X).$

Формули зв'язку центральних та початкових моментів, які наведені в «Темі 4», залишаються в силі і для випадку неперервних випадкових величин.

Центральний момент третього порядку характеризує **асиметрію закону розподілу випадкової величини**, тобто порушення симетрії кривої розподілу. Якщо $\mu_3 = 0$, то говорять, що випадкова величина симетрично поділена відносно $M(X)$ (крива розподілу симетрична відносно прямої $x = M(X)$. Оскільки μ_3 має вимірність випадкової величини в кубі, то впроваджують безрозмірну величину — коефіцієнт асиметрії: $A_s = \mu_3 / \sigma^3$.

Центральний момент четвертого порядку використовується для характеристики гостровершинності чи плосковершинності закону розподілу (кривої розподілу). Ці властивості описуються з допомогою **ексцесу**: $E_s = \mu_4 / \sigma^4 - 3$,
де число 3 віднімається в зв'язку з тим, що для найпоширенішого нормального закону розподілу (див. наступну тему) $\mu_4 / \sigma^4 = 3$, а отже, $E_s = 0$. Таким чином, ексцес характеризує гостровершинність чи плосковершинність в порівнянні із **нормальною кривою** (кривою розподілу нормального закону).

Приклад 1. Випадкова величина X задана густиною розподілу

$$f(x) = \begin{cases} 2\cos 2x, & \text{якщо } x \in (0; \pi/4), \\ 0, & \text{якщо } x \notin (0; \pi/4). \end{cases}$$

Знайти: а) інтервал, в якому знаходяться всі можливі значення X; б) $\sigma(X)$; в) моду $Mo(X)$; г) медіану $Me(X)$.

○ **а)** Згідно із наближеною рівністю (5.9) неперервна випадкова величина X набирає можливих значень при випробуванні в кожному інтервалі, на якому густина розподілу відмінна від нуля. Для нашого випадку $f(x) \neq 0 \ \forall \ x \in (0; \pi/4)$. Отже, в інтервалі $(0; \pi/4)$ знаходяться всі можливі значення X.

б) Для обчислення $\sigma(X)$ потрібно знайти спочатку $M(X)$ та $D(X)$. При знаходженні цих характеристик скористаємося інтегруванням частинами:

$$M(X) = \int_0^{\pi/4} x 2\cos 2x dx = 2\int_0^{\pi/4} x\cos 2x dx = 2\left(x \cdot \frac{1}{2}\sin 2x \Big|_0^{\pi/4} - \frac{1}{2}\int_0^{\pi/4} \sin 2x dx \right) =$$

$$u = x, \quad du = dx, \quad dv = \cos 2x dx, \quad v = \frac{1}{2}\sin 2x$$

$$= \frac{\pi}{4}\sin\frac{\pi}{2} + \frac{1}{2}\cos 2x\Big|_0^{\pi/4} = (\pi - 2)/4.$$

Дисперсію знайдемо за розрахунковою формулою (5.12):

$$M(X^2) = \int_0^{\pi/4} x^2 2\sqrt{2}\cos 2x dx = 2\int_0^{\pi/4} x^2 \cos 2x dx = 2\left[\left(\frac{x^2}{2}\sin 2x\right)\Big|_0^{\pi/4} - \int_0^{\pi/4} \sin 2x dx\right] =$$

$$u{=}x^2 \quad du{=}2xdx \quad dv{=}\cos 2xdx \quad v{=}\frac{1}{2}\sin 2x \qquad u{=}x \quad du{=}dx \quad dv{=}\sin 2xdx \quad v{=}-\frac{1}{2}\cos 2x$$

$$= \left(\frac{\pi^2}{16} + x\cos 2x\Big|_0^{\pi/4} - \frac{1}{2}\int_0^{\pi/4}\cos 2x dx\right) = \left(\pi^2/16 - \frac{1}{2}\sin 2x\Big|_0^{\pi/4}\right) = \left(\pi^2/16 - 1/2\right).$$

$$D(X) = M(X^2) - \left[M(X)\right]^2 = \left(\pi^2/16 - 1/2\right) - (\pi-2)^2/16 = (\pi-3)/4.$$

$$\sigma(X) = \sqrt{D(X)} = \sqrt{\pi - 3}/2 .$$

в) Оскільки функція $f(x) = 2\cos 2x$ у відкритому інтервалі $(0,\pi/4)$ не має максимуму, то X моди не має.

г) Знайдемо медіану Me, виходячи із її означення: $P(X < Me) = P(X > Me)$. Події $(X < Me)$ та $(X \geq Me)$ протилежні, тому $P(X < Me) + P(X \geq Me) = 1$. Але у відповідності із наслідком 2 вл. 2 функції розподілу $P(X = Me) = 0$, тому $P(X < Me) = 1/2$. Проте

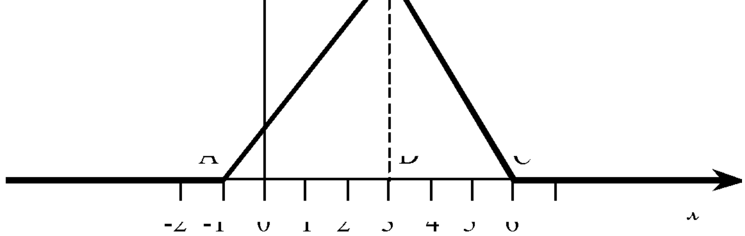
$$P\left(X < Me\right) = P\left(-\infty < X < Me\right) = \int_{-\infty}^{Me} f(x)dx = \int_{-\infty}^{0} 0dx + \int_0^{Me} 2\cos 2x dx = \sin 2x\Big|_0^{Me} = \sin 2Me.$$

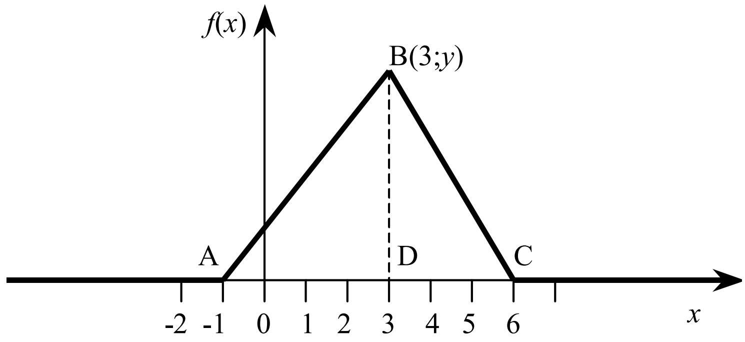
В результаті отримуємо рівняння відносно Me:

$$\sin 2\,Me = 1/2, \quad \text{звідки } 2\,Me = \arcsin(1/4).$$

Остаточно $Me = \left(\arcsin 1/4\right)/2$. •

Приклад 2. Графік густини розподілу неперервної випадкової величини X зображено на мал.5.1. Знайти: **а)** середнє квадратичне відхилення; **б)** моду $Mo(X)$; **в)** медіану $Me(X)$.





Мал. 5.1

○ Згідно з умовою густина розподілу випадкової величини X визначається виразом (5.16), суттєвою особливістю якого є те, що $f(x)$ визначається різними аналітичними виразами, відмінними від нуля, на різних інтервалах.

а) Для обчислення $\sigma(X)$ знайдемо спочатку $M(X)$ та $D(X)$ за формулами (5.10) та (5.12):

$$M(X)=\int_{-1}^{6} xf(x)dx=\int_{-1}^{3} xf(x)dx+\int_{3}^{6} xf(x)dx=$$

$$=\int_{-1}^{3} x\frac{x+1}{14}dx+\int_{3}^{6} x\frac{2(6-x)}{21}dx=\frac{1}{14}\int_{-1}^{3}(x^2+x)dx+\frac{2}{21}\int_{3}^{6}(6x-x^2)dx=$$

$$=\frac{1}{14}\left(\frac{x^3}{3}+\frac{x^2}{2}\right)\Bigg|_{-1}^{3}+\frac{2}{21}\left(6\cdot\frac{x^2}{2}-\frac{x^3}{3}\right)\Bigg|_{3}^{6}=\frac{8}{3};$$

$$M(X^2)=\int_{-1}^{6} x^2 f(x)dx=\int_{-1}^{3} x^2\frac{x+1}{14}dx+\int_{3}^{6} x^2\frac{2(6-x)}{21}dx=$$

$$=\frac{1}{14}\int_{-1}^{3}(x^3+x^2)dx+\frac{2}{21}\int_{3}^{6}(6x^2-x^3)dx=$$

$$=\frac{1}{14}\left(\frac{x^4}{4}+\frac{x^3}{3}\right)\Bigg|_{-1}^{3}+\frac{2}{21}\left(6\cdot\frac{x^3}{3}-\frac{x^4}{4}\right)\Bigg|_{3}^{6}=\frac{529}{42};$$

$$D(X)=M(X^2)-[M(X)]^2=\frac{529}{42}-\left(\frac{8}{3}\right)^2=\frac{691}{126};$$

$$\sigma(X)=\sqrt{D(X)}=\sqrt{\frac{691}{126}}\approx 2{,}342.$$

б) Із мал.5.1, на якому зображений графік густини розподілу, видно, що максимальне значення $f(x)$ досягає в т. x=3. Отже, $Mo(X)$=3.

в) Всі можливі значення даної випадкової величини знаходяться або на проміжку (–1; 3], або на (3; 6], але на кожному із них густина розподілу задається різними аналітичними виразами. З другого боку, медіана задовольняє рівнянню $P(X<Me)$=0,5.

Тому якщо $P(X\le 3)\ge 0{,}5$, то $Me(X)\in(1;3]$, якщо ж $P(X\le 3)<0{,}5$, то $Me(X)\in(3;6]$.

Використавши формулу (5.3), отримаємо

$$P(X \le 3) = P(-\infty < X \le 3) = F(3) - F(-\infty) = \frac{1}{28}(x+1)^2\Big|_{x=3} - 0 = \frac{16}{28} = \frac{4}{7} > 0{,}5.$$

Отже, $Me(X) \in (1;3]$. Тоді для знаходження Me отримаємо рівняння:

$$P(X < Me) = F(Me) - F(-\infty) = \frac{1}{28}(Me+1)^2 - 0 = 0{,}5$$,або $(Me+1)^2=14$,

звідки $Me = \sqrt{14} - 1 \approx 2{,}742$. •

Рекомендована література: [1], С. 76–126; [2], С. 64–100; [4], С. 68–90; [5], С. 90–115.

Лекція 6.

Тема: ***Основні закони неперервних випадкових величин.***

1. ***Нормальний закон: імовірнісний зміст параметрів розподілу; нормальна крива та вплив параметрів розподілу на її форму; ймовірність попадання у заданий інтервал; знаходження ймовірності заданого відхилення; правило трьох сигм.***
2. ***Закон рівномірного розподілу.***
3. ***Показників закон. Гамма-розподіл та розподіл Ерланга.***
4. ***Розподіл хі-квадрат.***

1.

Серед всіх випадкових величин, з якими доводиться зустрічатися на практиці та наукових дослідженнях, особливе місце займають величини, що мають нормальний закон розподілу імовірностей. Це зумовлено як найбільшою частотою їх зустрічей, так і великим практичним значеним інших випадкових величин, пов'язаних із нормальним розподілом.

Випадкова величина називається **нормально розподіленою (розподіленою за нормальним законом),** якщо її густина розподілу має такий вид

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-a)^2/2\sigma^2} \tag{6.1}$$

1.1. Імовірносний зміст параметрів розподілу

Густина розподілу (6.1) повністю визначається двома параметрами: a та σ. Імовірносний зміст цих параметрів визначається такими рівностями:

$$M(X)=a,\ \sigma=\sqrt{D(X)} \tag{6.2}$$

Для доведення цих співвідношень слід використати формули:

$$M(X)=\int_{-\infty}^{+\infty}xf(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}xe^{-(x-a)^2/2\sigma^2}dx$$

$$D(X)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left(x-a\right)^2e^{-(x-a)^2/2\sigma^2}dx$$

Для знаходження кожного із цих невласних інтегралів потрібно виконати заміну змінної

$$x=\sigma z+a,\ dx=\sigma dz \tag{6.3}$$

при цьому в першому інтегралі після простих перетворень скористатися

Інтегралом Пуассона $\int_{-\infty}^{+\infty}e^{-z^2/2}dz=\sqrt{2\pi}$, а в другому – проінтегрувати частинами. Всі викладки рекомендуються здійснити самостійно.

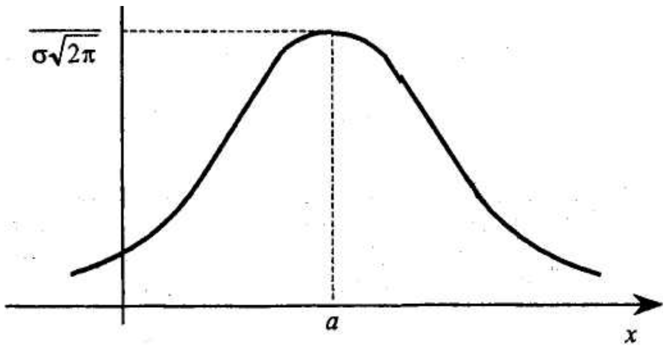
1.2. Нормальна крива та вплив параметрів розподілу на її форму

Графік густини нормального розподілу називається нормальною кривою (крива Гаусса).

Наведемо підсумки дослідження функції (6.1) методами диференціального числення.

1) $D(f) = R^1$, $E(f) = (0,\infty)$.
2) $\lim\limits_{|x|\to\infty} f(x) = 0$, тобто вісь $0x$ є лівосторонньою та правосторонньою горизонтальною асимптотою графіка функції.
3) Нормальна крива симетрична відносно прямої $x = a$, оскільки $x - a$ міститься у $f(x)$ в квадраті.
4) В точці $x = a$ функція $f(x)$ набуває максимальне значення $\dfrac{1}{\sigma\sqrt{2\pi}}$

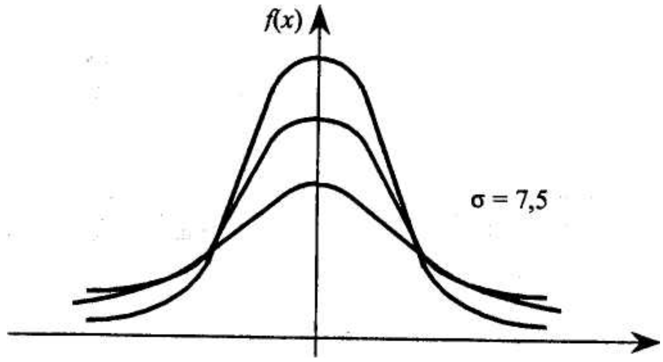
В результаті отримується такий ескіз нормальної кривої (мал. 6.1).



Мал. 6.1.

З'ясуємо, як впливають на форму і розташування нормальної кривої зміни значень параметрів a та σ.

З елементарної математики відомо, що графіки функцій $f(x)$ і $f(x–b)$ мають однакову форму, причому графік $f(x–b)$ отримується із графіка $f(x)$ внаслідок паралельного пересування вправо (в додатньому напрямку осі $0x$) на b одиниць при $b > 0$ або вліво при $b < 0$. Отже, зростання параметра a приводить до пересування нормальної кривої (без зміни форми) вправо, а якщо a зменшується — то вліво.



Мал. 6.2.

При зміні параметра a відбувається «деформування» нормальної кривої. Це зумовлено, з одного боку, тим, що зміна σ веде до зміни висоти кривої в точці a (зменшення а приводить до збільшення максимального значення $f(x)$, а збільшення σ — до зменшення висоти). З другого боку, згідно із властивістю 2 густини розподілу, площа криволінійної трапеції, обмеженої зверху нормальною кривою, дорівнює одиниці. Тому збільшення висоти кривої внаслідок зменшення σ приводить до більшої крутизни віток кривої, а збільшення σ зумовлює більшу пологість віток. На мал. 6.2 зображені нормальні криві для різних значень σ та $a = 0$.

1.3. Імовірність попадання у заданий інтервал

Нехай випадкова величина X розподілена за нормальним законом. Тоді за формулою (5.7) імовірність того, що Xпри випробуванні набере значення з інтервалу (а, Р), визначається рівністю

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(x-a)^2 / 2\sigma^2} dx$$

Мета наступних перетворень полягає у зведенні інтеграла справа до значень функції Лапласа. Для досягнення її здійснимо заміну змінних (6.3) і скористаємося рівністю

$$\int_c^d g(z)dz = \int_0^d g(z)dz - \int_0^c g(z)dz,$$

де інтегрована функція $g(z)>0$. В результаті остаточно отримаємо

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right) \tag{6.4}$$

1.4. Знаходження імовірності заданого відхилення

В практично важливих задачах виникає необхідність в знаходженні імовірності того, що відхилення нормально розподіленої випадкової величини від a по абсолютній величині менше від заданого додатного числа ε, тобто $P(|X–a|<\varepsilon)$.

Випадкова подія $(|X–a|<\varepsilon)$ рівносильна події $a–\varepsilon<X<a+\varepsilon$, імовірність якої можна знайти за формулою 6.4, де $\alpha = a–\varepsilon$, $\beta = a+\varepsilon$. Врахувавши непарність функції Лапласа $\Phi(x)$, остаточно отримаємо рівність

$$P(|X - a| < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right) \tag{6.5}$$

Ця рівність, зокрема, показує, що імовірність відхилення для одного і того ж ε буде тим більша, чим меншим буде σ.

1.5. Правило трьох сигм

З'ясуємо одну характеристичну особливість нормально розподілених величин. Розглянемо випадкові події $(|X–a|<\sigma)$, $(|X–a|<2\sigma)$ і т. д. і знайдемо їх імовірності формулою (6.5) і табл. додатків. В результаті отримаємо:

$P(|X–a|<\sigma)=2\Phi(1)=0{,}6826$,

$P(|X–a|<2\sigma)=2\Phi(2)=0{,}9594$,

$P(|X-a|<3\sigma)=2\Phi(3)=0{,}9973$,
$P(|X-a|<4\sigma)=2\Phi(4)=0{,}999936$,
$P(|X-a|<5\sigma)=2\Phi(5)=0{,}999994$.

Як видно з обчислень, інтервалом **практично всіх** можливих значень нормально розподіленої величини є інтервал ($a-3\sigma$; $a+3\sigma$) оскільки імовірності попадання в наступні інтервали збільшуються незначно. Це означає, що імовірність того, що абсолютна величина відхилень нормально розподіленої величини перевищить потроєне середнє квадратичне відхилення, дуже мала (0,0027). Таким чином, знаючи $M(X)$ та $\sigma(X)$, орієнтовно можна знайти інтервал її практично всіх можливих значень (при умові, що X— нормально розподілена).

Це правило називається **правилом трьох сигм.**

Отже, якщо досліджувана випадкова величина характеризується вказаною вище особливістю, тоді можна висувати гіпотезу про нормальність її розподілу.

Приклад 1. Коробки із шоколадом упаковуються автоматично, при цьому середня маса однієї коробки становить 1,04 кг. Відомо, що тільки 2,5% коробок мають масу, меншу від 1 кг. Припускаючи, що маса коробок розподілена нормально, знайти середнє квадратичне відхилення.

○ Позначимо: X — маса навмання взятої коробки із шоколадом. За умовою X — нормально розподілена випадкова величина, $M(X) = 1{,}04$ і $P(X < 1) = 0{,}025$. Протилежною до події ($X < 1$) є подія ($1 \le X < \infty$), тому $P(1 \le X < \infty) = 1 - P(X < 1) = 1 - 0{,}025 = 0{,}975$. Для знаходження σ використаємо формулу (6.3), де $a = 1{,}04$, $\alpha = 1$, $\beta = \infty$, тобто

$$P(1 \le X < \infty) = 0{,}975 = \Phi\left(\frac{\infty - 1{,}04}{\sigma}\right) - \Phi\left(\frac{1 - 1{,}04}{\sigma}\right),$$

або

$$0{,}975 = 0{,}5 + \Phi\left(\frac{0{,}04}{\sigma}\right),$$

оскільки $\Phi(x)$ — непарна функція і $\Phi(x) = 0{,}5$ для $x > 5$.

Остаточно рівняння набуває такого вигляду

$$\Phi\left(\frac{0{,}04}{\sigma}\right) = 0{,}475.$$

За табл. додатків $\Phi(1{,}96) = 0{,}475$, тому

$$0{,}04/\sigma = 1{,}96, \text{ звідки } \sigma = 0{,}04/1{,}96 \approx 0{,}0204. \quad \bullet$$

Приклад 2. На автоматичному токарному станку виготовляють болти, номінальна довжина яких 40 мм. В процесі роботи станка спостерігаються випадкові відхилення від вказаного розміру, які розподілені за нормальним законом з математичним сподіванням 0 і середнім квадратичним відхиленням 1 мм. При контролі бракуються всі болти, розміри яких відрізняються від номінального більше, ніж на допуск в 2 мм. Знайти імовірність того, що навмання взятий болт виявиться бракованим.

○ Нехай випадкова величина X — відхилення розміру навмання взятого болта від номінального. За умовою задачі X розподілена за нормальним законом і $M(X) = 0$, $\sigma(X) = 1$. Потрібно знайти $P\left(\,|X| > 2\right)$. Але випадкова подія $\left(\,|X| > 2\right)$ протилежна до події $\left(\,|X| \le 2\right)$, тому $P\left(\,|X| > 2\right) = 1 - P\left(\,|X| \le 2\right)$.

Для знаходження імовірності в правій частині рівності можна використати формулу (6.4) з такими значеннями параметрів: $a = 0$, $\varepsilon = 2$, $\sigma = 1$.

Тоді $P\left(\,|X| \le 2\right) = 2\Phi(2) = 2 \cdot 0{,}4772 = 0{,}9544,$

$$P\left(\,|X| > 2\right) = 1 - 0{,}9544 = 0{,}0456. \quad \bullet$$

Приклад 3. Випадкова величина X розподілена за нормальним законом із математичним сподіванням $a = 0$ і середнім квадратичним відхиленням σ. При якому значенні σ імовірність того, що при випробуванні величина X набуде можливого значення з інтервалу (1, 2), буде максимальною?

○ За формулою (6.3)

$$P(1 < X < 2) = \Phi\left(\frac{2}{\sigma}\right) - \Phi\left(\frac{1}{\sigma}\right) = \frac{1}{\sqrt{2\pi}}\left(\int\limits_0^{2/\sigma} e^{-x^2/2}dx - \int\limits_0^{1/\sigma} e^{-x^2/2}dx\right).$$

Ця імовірність є функцією однієї змінної σ і для знаходження її максимального значення знайдемо критичні точки, продиференціювавши обидві частини і прирівнявши до нуля похідну:

$$\frac{d}{d\sigma}P(1 < X < 2) = \frac{1}{\sqrt{2\pi}}\left(e^{-(2/\sigma)^2/2}\,(2/\sigma)' - e^{-(1/\sigma)^2/2}\,(1/\sigma)'\right) =$$

$$= \frac{1}{\sqrt{2\pi}}\left(e^{-2/\sigma^2}(-2/\sigma^2) - e^{-1/(2\sigma^2)}(-1/\sigma^2)\right) = 0.$$

Звідси $2e^{-2/\sigma^2} = e^{-1/2\sigma^2}$. Прологарифмувавши ліву і праву частину за основою e, після простих перетворень отримаємо:

$$\sigma = \sqrt{3/(2\ln 2)} \approx \sqrt{3/(2 \cdot 0{,}693)} \approx 1{,}471.$$

Можна перевірити, що $\dfrac{d}{d\sigma}P(1 < X < 2)$ при проходженні через знайдену точку змінює знак з «+» на «–», а тому $\sigma \approx 1{,}471$ надає максимального значення імовірності випадкової події $(1 < X < 2)$. $\bullet$

Приклад 4. Випадкова величина X розподілена за нормальним законом із математичним сподіванням 7,6. Імовірність того, що при випробуванні X набере значення з інтервалу (7,65; 7,68), дорівнює 0,25. Знайти $P(7{,}52<X<7{,}55)$.

○ Нормальна крива є симетричною відносно прямої x=7,6, а інтервали однакової довжини (7,52; 7,55) і (7,65; 7,68) – симетричні відносно точки x=7,6. Тому рівними є площі криволінійних трапецій, основами яких є вказані відрізки і які обмежені зверху нормальною кривою. За формулою (5.6) $P(7{,}52 < X < 7{,}55) = \int\limits_{7{,}52}^{7{,}55} f(x)dx$, де визначений інтеграл – площа криволінійної трапеції із основою (7,52; 7,55), яка обмежена зверху частиною нормальної кривої. Із рівності площ розглянутих криволінійних трапецій отримуємо шукану імовірність: $P(7{,}52<X<7{,}55)=0{,}25$. •

2.

2. Закон рівномірного розподілу

Випадкова величина називається **розподіленою за рівномірним** законом **(рівномірно розподіленою),** якщо її густина розподілу має такий вид

$$f(x) = \begin{cases} C = \text{const}, & \text{якщо } x \in [a;b] \\ 0, & \text{якщо } x \notin [a;b] \end{cases}$$

Знайдемо значення сталої C, використавши другу властивість густини розподілу. Геометричний зміст цієї властивості означає рівність одиниці площі, обмеженої кривою розподілу, тобто $C\,(b-a) = 1$, звідки $C = 1/(b-a)$. Отже, густина рівномірно розподіленої випадкової величини має такий вид:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & \text{якщо } x \in [a;b] \\ 0, & \text{якщо } x \notin [a;b] \end{cases} \tag{6.6}$$

Відрізок $[a,\ b]$ називається **відрізком концентрації** рівномірного **розподілу.**

З'ясуємо походження назви **рівномірний** розподіл. З цією метою розглянемо інтервали (a_1, b_1), (a_2, b_2), такі, що їхні кінці належать відрізку $[a, b]$ і $b_1 - a_1 = b_2 - a2$. Тоді за формулою (5.7)

$$P\left(a_1 < X < b_1\right) = \int\limits_{a_1}^{b_1} f(x)dx = \frac{1}{b-a}\int\limits_{a_1}^{b_1} dx = \frac{b_1 - a_1}{b-a},$$

$$P\left(a_2 < X < b_2\right) = \frac{b_2 - a_2}{b-a}$$

Отже, імовірності попадання при випробуванні значень величини X в будь-які інтервали однакової довжини із відрізка $[a, b]$ **рівні.**

Знайдемо функцію розподілу імовірностей рівномірно розподіленої випадкової величини X

$$F(x) = \begin{cases} 0, & x < a, \\ \dfrac{x-a}{b-a}, & a \le x \le b, \\ 1, & x > b, \end{cases}$$

також числові характеристики $M(X) = \dfrac{a+b}{2}$, $D(X) = \dfrac{(b-a)^2}{12}$, $\sigma(X) = \dfrac{b-a}{2\sqrt{3}}$.

В якості самостійної роботи пропонується виконати розрахунки, а також побудувати графіки функцій $f(x)$ та $F(x)$. Рівномірний розподіл мають випадкові величини, які характеризують помилки вимірювань при допомозі приладів із крупними поділками, коли значення заокруглюються до найближчого цілого.

Приклад 1. Для рівномірно розподіленої випадкової величини X знайти: 1) функцію розподілу імовірностей та побудувати її графік; 2) математичне сподівання; 3) дисперсію.

○ Густину (6.5) рівномірного розподілу запишемо в такому вигляді

$$f(x)=\begin{cases}0, \quad x<a,\\ 1/(b-a), \quad a\le x\le b,\\ 0, \quad x>b.\end{cases}$$

1) Використаємо формулу (5.7)

$$F(x)=\int_{-\infty}^{x} f(x)dx.$$

Якщо $x<a$, тоді $F(x)$=0. Нехай x□$[a, b]$, тоді

$$F(x)=\int_{-\infty}^{x} f(x)dx=\int_{-\infty}^{a} f(x)dx+\int_{a}^{x} f(x)dx=\int_{-\infty}^{a} 0dx+\int_{a}^{x}\frac{1}{b-a}dx=\frac{1}{b-a}x\Big|_{a}^{x}=\frac{x-a}{b-a}.$$

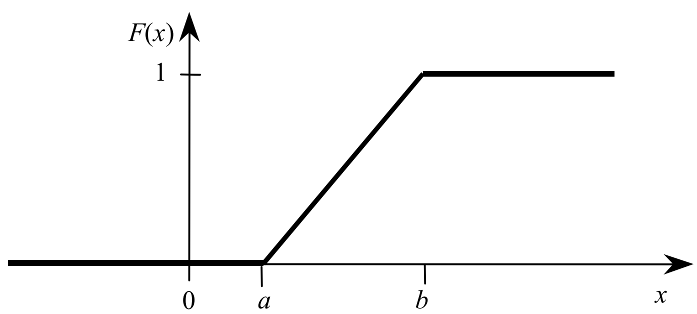
Якщо ж $x>b$, то

$$F(x)=\int_{-\infty}^{x} f(x)dx=\int_{-\infty}^{a} f(x)dx+\int_{a}^{b} f(x)dx+\int_{b}^{x} f(x)dx=\int_{-\infty}^{a} 0dx+\int_{a}^{b}\frac{dx}{b-a}+\int_{b}^{x}0dx=\frac{x}{b-a}\Big|_{a}^{b}=1.$$

Отже, шукана функція розподілу має такий вид

$$F(x)=\begin{cases}0, \quad x<a,\\ \dfrac{x-a}{b-a}, \quad a\le x\le b,\\ 1, \quad x>b.\end{cases}$$

Її графік зображено на мал.6.2



Мал.6.2.

2) Для знаходження математичного сподівання використаємо формули (5.10*) та (6.5)

$$M(X)=\int_{-\infty}^{\infty} xf(x)dx=\int_{-\infty}^{a} xf(x)dx+\int_{a}^{b} xf(x)dx+\int_{b}^{\infty} xf(x)dx=$$
$$=\int_{-\infty}^{a} x0dx+\int_{a}^{b} x\frac{1}{(b-a)}dx+\int_{b}^{\infty} x0dx=0+\frac{1}{(b-a)}\int_{a}^{b} xdx+0=$$
$$=\frac{1}{(b-a)}\cdot\frac{x^2}{2}\bigg|_a^b=\frac{b^2}{2(b-a)}-\frac{a^2}{2(b-a)}=\frac{(b-a)(b+a)}{2(b-a)}=\frac{b+a}{2}.$$

3) Дисперсію знайдемо, використавши формули (5.12*) та (6.5)

$$D(X)=\int_{-\infty}^{\infty} x^2 f(x)dx-\left[M(X)\right]^2=\int_{-\infty}^{a} x^2 f(x)dx+\int_{a}^{b} x^2 f(x)dx+$$
$$+\int_{b}^{\infty} x^2 f(x)dx-\left(\frac{b+a}{2}\right)^2=\int_{-\infty}^{a} x^2 0dx+\int_{a}^{b} x^2\frac{1}{(b-a)}dx+\int_{b}^{\infty} x^2 0dx-$$
$$-\frac{(b+a)^2}{4}=0+\frac{1}{(b-a)}\int_{a}^{b} x^2dx+0-\frac{(b+a)^2}{4}=\frac{x^3}{3(b-a)}\bigg|_a^b-\frac{(b+a)^2}{4}=$$
$$=\frac{b^3-a^3}{3(b-a)}-\frac{(b+a)^2}{4}=\frac{b^2+ab+a^2}{3}-\frac{b^2+2ab+a}{4}=\frac{b^2-2ab+a^2}{12}=\frac{(b-a)^2}{12}.\ \bullet$$

Приклад 2. Перехрестя обладнане автоматичним світлофором, на якому зелене та червоне світло горить протягом 1 хв. та 0,75 хв. відповідно. Автомобіль під'їжджає до перехрестя у випадковий момент часу, не пов'язаний із роботою світлофора. Знайти імовірність того, що він проїде перехрестя, не зупиняючись.

○ Позначимо випадкову величину X — момент приїзду автомобіля до перехрестя. Її можливі значення знаходяться в проміжку (0; 1,75), довжина якого визначається періодом зміни кольорів на світлофорі. Зрозуміло, що можливість приїзду автомобілем в будь-який момент часу даного інтервалу з (0; 1,75) рівна можливості приїзду в момент часу іншого інтервалу такої ж довжини. Тому X — рівномірно розподілена з густиною розподілу (6.5), де $a=0$, $b=1{,}75$. Автомобіль проїде, не зупиняючись, перехрестя, якщо відбудеться випадкова подія $(0<X<1)$, тобто під час «дії» зеленого світла світлофора. Тоді за формулою (5.6)

$$P(0<X<1)=\int_0^1 f(x)dx=\frac{1}{1{,}75}\int_0^1 dx=\frac{4}{7}.\quad\bullet$$

3.

Показниковий закон

Випадкова величина X розподіляється за **показниковим (експоненціальним) законом (показниково розподілена),** якщо її густина розподілу імовірностей має такий вид

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & \text{якщо } x\ge 0\\ 0, & \text{якщо } x<0\end{cases}\qquad(6.7)$$

де стала $\lambda > 0$ називається параметром експоненціального розподілу. Прикладом неперервної випадкової величини, розподіленої за показниковим розподілом, може служити час між появами двох послідовних подій найпростішого потоку, де λ — інтенсивність потоку.

Знову за схемою знаходження функції розподілу ймовірностей випадкової величини можна знайти функцію розподілу імовірностей показниково розподіленої величини *X:*

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, \text{ якщо } x \geq 0, \\ 0, \qquad \text{якщо } x < 0. \end{cases}$$

Приклад 1. Знайти числові характеристики показниково розподіленої випадкової величини *X*.

○ За формулами (5.10*) та (6.6):

$$M(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{0} x \cdot 0dx + \int_{0}^{\infty} x\lambda e^{-\lambda x}dx = \lambda\int_{0}^{\infty} xe^{-\lambda x}dx =$$

$$= \begin{vmatrix} u = x, & dv = e^{-\lambda x}dx \\ du = dx, & v = -\dfrac{1}{\lambda}e^{-\lambda x} \end{vmatrix} = \lambda\left\{ -\frac{x}{\lambda}e^{-\lambda x}\Big|_0^{\infty} + \frac{1}{\lambda}\int_0^{\infty} e^{-\lambda x}dx \right\} = -\frac{1}{\lambda}e^{-\lambda x}\Big|_0^{\infty} = \frac{1}{\lambda}.$$

Для знаходження $D(X)$ скористаємось формулами (5.12*) та (6.6):

$$D(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - \left[M(X)\right]^2 = \lambda\int_0^{\infty} x^2 e^{-\lambda x}dx - \frac{1}{\lambda^2};$$

$$\int_0^{\infty} x^2 e^{-\lambda x}dx = \begin{vmatrix} u = x^2, & du = 2xdx \\ dv = e^{-\lambda x}dx, & v = -\dfrac{1}{\lambda}e^{-\lambda x} \end{vmatrix} =$$

$$= -\frac{x^2}{\lambda}e^{-\lambda x}\Big|_0^{\infty} + 2\lambda^{-1}\int_0^{\infty} xe^{-\lambda x}dx = \begin{vmatrix} u = x, & du = dx \\ dv = e^{-\lambda x}dx, & v = -\dfrac{1}{\lambda}e^{-\lambda x} \end{vmatrix} =$$

$$= \frac{2}{\lambda}\left\{ -\frac{x}{\lambda}e^{-\lambda x}\Big|_0^{\infty} + \frac{1}{\lambda}\int_0^{\infty} e^{-\lambda x}dx \right\} = \frac{-2}{\lambda^3}e^{-\lambda x}\Big|_0^{\infty} = \frac{2}{\lambda^3};$$

$$D(X) = \lambda\frac{2}{\lambda^3} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}; \ \sigma(X) = \frac{1}{\lambda}. \qquad \bullet$$

Приклад 2. Статистично встановлено, що середній час очікування наступного покупця, який підійде до каси, дорівнює 0,2 хвилини. Час чекання касиром чергового покупця можна вважати випадковою величиною, розподіленою за показниковим законом. Для заміни стрічки касового апарату потрібно 2 хвилини. Знайти імовірність того, що за час заміни стрічки не утвориться черга покупців.

○ Нехай *X* – час чекання касиром чергового покупця. За умовою *X* розподілена за показниковим законом і $M(X)$=0,2. Але $M(X)=\frac{1}{\lambda}$ (див. приклад 1). Тому $1/\lambda=0,2$ звідки λ=5.

Нехай випадкова подія A – за час заміни стрічки касового апарату не утвориться черга. Тоді протилежна до неї подія

$\overline{A}=[M(X)\le X\le 2]$.

Але оскільки $M(X)$=0,2, то за формулою (5.6)

$$P(\overline{A})=P[M(X)\le X\le 2]=P(0,2\le X\le 2)=\int\limits_{0,2}^{2} f(x)dx=$$

$$=\int\limits_{0,2}^{2} 5e^{-5x}dx=-\int\limits_{0,2}^{2} e^{-5x}d(-5x)=-e^{-5x}\Big|_{0,2}^{2}=-e^{-10}+e^{-1}=e^{-1}-e^{-10}.$$

Отже, $P(A)=1-e^{-1}+e^{-10}$. ●

Показниковий закон розподілу може використовуватися також в якості однієї із можливих математичних моделей в теорії надійності. Параметр λ в теорії надійності називається інтенсивністю відмови елемента. Характерною для експоненціального закону є притаманна лише йому властивість — відсутність післядії, тобто минуле для цього закону не впливає на передбачення у майбутньому. Має місце і зворотне твердження.

Гамма-розподіл та розподіл Ерланга

Моделювання складних явищ та процесів (зокрема, економічних) передбачає використання адекватних математичних моделей. Показниковий розподіл є частинним випадком так званого гамма-розподілу, який успішно використовується в техніко-економічному моделюванні і який, зокрема, достатньо добре описує час безвідмовної роботи різних технічних пристроїв та обладнання.

Невід'ємна неперервна випадкова величина X має **гамма-розподіл імовірностей**, якщо густина розподілу має такий вид:

$$f(x)=\begin{cases}\dfrac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, & \text{якщо } x\ge 0\\ 0, & \text{якщо } x<0\end{cases} \qquad (6.8)$$

де $\alpha>0$, $\lambda>0$ — параметри розподілу $\Gamma(\alpha)=\int\limits_{0}^{\infty} z^{\alpha-1}e^{-z}dz$ — гамма-функція (інтеграл Ейлера).

Наведемо властивості гамма-функції, корисні при користуванні гамма-розподілом:

1) $\Gamma(1)=1$; 2) $\Gamma(\alpha+1)=\alpha\,\Gamma(\alpha)$; 3) якщо $\alpha=n$ — ціле невід'ємне число, то $\Gamma(n+1)=n(n-1)(n-2)\cdot\ldots\cdot 2\cdot 1=n!$; 4) $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$.

Із врахуванням властивості 1) видно, що при $\alpha = 1$ з (6.8) отримуємо (6.7), тобто при $\alpha = 1$ отримуємо показниковий розподіл.

Використання властивостей гамма-функції дозволяє знайти числові характеристики гамма-розподіленої випадкової величини *X:*

$$M(X) = \frac{\alpha}{2},\ D(X) = \frac{\alpha}{\lambda^2},\ \sigma(X) = \frac{\sqrt{\alpha}}{\lambda}. \qquad (6.9)$$

Якщо параметр α набуває цілих значень ($\alpha = k = 1,2,3,...$), То гамма-розподіл перетворюється на розподіл Ерланга *к-го* порядку:

$$f(x) = \begin{cases} \dfrac{\lambda(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x}, & \text{якщо } x \geq 0 \\ 0, & \text{якщо } x < 0 \end{cases}$$

Розподіл Ерланга знаходить важливе застосування в теорії масового обслуговування.

Числові характеристики цього розподілу отримують із співвідношень (6.9), де $\alpha = k$.

4.

Частинним випадком гамма-розподілу при $\alpha = k/2$, де $k = 1,2,3,...$, а $\lambda = 1/2$, є так званий розподіл χ^2 (*хі*-квадрат), значення якого в математичній статистиці важко переоцінити; параметр k називається в цьому випадку числом ступенів вільності розподілу χ^2. Сформулюємо інший підхід до визначення розподілу χ^2.

Нехай задано k незалежних випадкових величин $X_1, X_2, ..., X_k$ кожна з яких має нормований нормальний закон розподілу імовірностей, тобто $M(Xi) = 0$, $\sigma(X_i) = 1, i = \overline{1;k}$. Тоді випадкова величина $X = \sum_{i=1}^{k} X_i^2$ матиме розподіл χ^2 із k ступенями вільності, тобто густина розподілу імовірностей X матиме такий вид:

$$f(x) = \begin{cases} 0, & \text{якщо } x < 0 \\ \dfrac{1}{2^{k/2}\Gamma\left(k/2\right)} x^{(k/2)-1} e^{-x/2}, & \text{якщо } x \geq 0 \end{cases} \qquad (6.10)$$

Числові характеристики випадкової величини, розподіленої за законом χ^2, визначаються рівностями:

$$M(X) = k,\ D(X) = 2k,\ \sigma(X) = \sqrt{2k}.$$

Із збільшення числа ступенів вільності розподіл χ^2 «повільно» наближається до нормального.

Рекомендована література: [1], С. 116–126; [2], С. 127–155; [4], С. 80–90; [5], С. 140–151.

Лекція 7.

Тема: **Системи випадкових величин.**

1. ***Закон розподілу ймовірностей двовимірної дискретної випадкової величини.***
2. ***Функція розподілу двохвимірної випадкової величини та її властивості.***
3. ***Густина розподілу ймовірностей двохвимірної випадкової величини та її властивості.***
4. ***Умовні закони розподілу. Залежні і не залежні випадкові величини. Умовне математичне сподівання.***
5. ***Рівняння регресії.***
6. ***Числові характеристики системи двох випадкових величин.***
7. ***Кореляційний момент. Коефіцієнт кореляції.***
8. ***Система довільного скінченого числа випадкових величин.***
9. ***Кореляційна матриця.***
10. ***Нормальний закон розподілу двохвимірної випадкової величини.***

1.

Системою випадкових величин називається сукупність випадкових величин, яка розглядається як єдине ціле.

Нехай X та Y є випадковими величинами, які або в одному і тому ж випробуванні набирають своїх можливих значень, або розглядаються одночасно. Будемо позначати (X, Y) систему цих величин (двовимірну випадкову величину), де кожну із величин X та Y називатимемо складовими або компонентами. Геометрично (X, Y) можна інтерпретувати як випадкову точку $A(X, Y)$ на площині $x0y$ або випадковий вектор $\vec{0A}$.

Розглянемо дискретну двовимірну випадкову величину (X, Y), тобто величину, складові якої є дискретними величинами. Множина можливих значень Ω_{XY} такої випадкової величини містить скінченне або зчисленне число точок (x_i, y_j), де $i, j = \overline{1,\infty}$, тобто

$$\Omega_{XY} = \{(x_1, y_1); (x_1, y_2); ...; (x_2, y_1); (x_2, y_2); ...; (x_3, y_1); (x_3, y_2); ...\}.$$

Позначимо:

$$p_{ij} = P\big((X = x_i)\cdot(Y = y_j)\big) = P(X = x_i; Y = y_j), \qquad (7.1)$$

тобто p_{ij} — це імовірність того, що при випробуванні складова X набирає можливого значення x_i і одночасно з цим складова Y набирає значення y_j, де $(x_i, y_j) \in \Omega_{XY}$, крапка — множення подій.

Законом розподілу імовірностей або **розподілом дискретної двовимірної випадкової величини** (X, Y) називається відповідність між парами чисел $(x_i, y_j) \in \Omega_{XY}$ та імовірностями p_{ij}.

Закон розподілу дискретної двовимірної величини (X, Y) задається таблицею, вид якої для випадку скінченного числа точок множини Ω_{XY} визначається табл. 7.1.

Таблиця 7.1

$X \backslash Y$	y_1	y_2	...	y_j	...	y_n
x_1	p_{11}	p_{12}	...	p_{1j}	...	p_{1n}
x_2	p_{21}	p_{22}	...	p_{2j}	...	p_{2n}
...	...	...	...	...	...	...
x_i	p_{i1}	p_{i2}	...	p_{ij}	...	p_{in}
...	...	...	...	...	...	...
x_m	p_{m1}	p_{m2}	...	p_{mj}	...	p_{mn}

Перший стовпець таблиці містить всі можливі значення складової X, а перший рядок — всі можливі значення складової Y. В клітинці, яка розміщена на перетині рядка «x_i» та стовпця «y_j» вказана імовірність p_{ij} того, що вектор (X, Y) набере значення (x_i, y_j), де $i = \overline{1,m}, j = \overline{1,n}.$

Всі можливі (випадкові) події $(X = x_i; Y = y_j)$ при $i = 1, 2, ... m$, $j = 1, 2, ... n$ утворюють повну групу, а тому $\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij} = 1.$ Цю рівність називають **умовою нормування** для системи (X, Y).

Значення закону розподілу системи двох дискретних випадкових величин дозволяє знайти закони розподілу її складових.

Справді, із врахування повноти групи подій $(Y = y_1)$, $(Y = y_2)$, ..., $(Y = y_n)$ для $i = \overline{1,m}$ має місце рівність

$$(X = x_i) = (X = x_i)(Y = y_1) + (X = x_i)(Y = y_2) + ... + (X = x_i)(Y = y_n),$$

доданки якої є попарно несумісними подіями. Використання теореми додавання імовірностй і позначень (7.1) дає рівність

$$P(X = x_i) = p_{i1} + p_{i2} + ... + p_{in}, \quad (i = \overline{1,m}). \tag{7.2}$$

Отже, для знаходження $P(X = x_i)$ потрібно додати імовірності рядка «x_i» табл. 7.1. Аналогічно просумувавши імовірності стовпця «y_j», отримаємо:

$$P(Y = y_j) = \sum_{i=1}^{m} p_{ij} \quad (j = \overline{1,n}). \tag{7.3}$$

Задача 7.1. Якість продукції характеризується двома випадковими параметрами: X та Y. Знайти закони розподілу складових випадкової величини $Z = (X, Y)$, закон розподілу якої наведений в табл. 7.2.

Таблиця 7.2

$X \backslash Y$	0	0,1	0,2	0,3
2	0,2	0,05	0,1	0,05
3	0	0,1	0,15	0,15
4	0	0,15	0	0,05

– Використаємо рівності (7.2):

$$P(X = 2) = 0{,}2 + 0{,}05 + 0{,}1 + 0{,}05 = 0{,}4;$$
$$P(X = 3) = 0 + 0{,}1 + 0{,}15 + 0{,}15 = 0{,}4;$$
$$P(X = 4) = 0 + 0{,}15 + 0 + 0{,}05 = 0{,}2.$$

Закон розподілу складової X має вид

X	2	3	4
P	0,4	0,4	0,2

.

За формулами (7.3)

$$P(Y = 0) = 0{,}2 + 0 + 0 = 0{,}2;$$
$$P(Y = 0{,}1) = 0{,}05 + 0{,}1 + 0{,}15 = 0{,}3;$$
$$P(Y = 0{,}2) = 0{,}1 + 0{,}15 + 0 = 0{,}25;$$
$$P(Y = 0{,}3) = 0{,}05 + 0{,}15 + 0{,}05 = 0{,}25,$$

звідки отримується закон розподілу складової Y:

Y	0	0,1	0,2	0,3
P	0,2	0,3	0,25	0,25

.

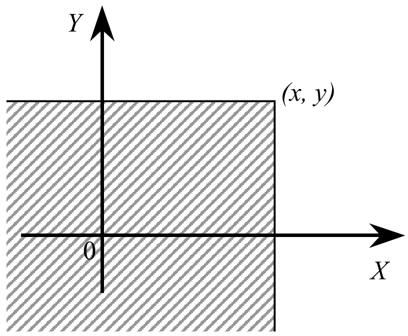
2.

Розглянемо двовимірну випадкову величину (X, Y), складові якої можуть бути дискретними або неперервними величинами і множина можливих значень якої є Ω_{XY}.

Функцією розподілу (**імовірностей**) двовимірної випадкової величини (X, Y) називається функція $F(x, y)$ детермінованих аргументів x та y, яка дорівнює імовірності того, що при випробуванні X набере значення, менше від x, і при цьому Y набере значення, менше від y:

$$F(x, y) = P\{(X < x)\cdot(Y < y)\} = P(X < x, Y < y). \tag{7.4}$$

З геометричної точки зору функція розподілу є імовірність попадання випадкової точки (X, Y) в нескінченний квадрант з вершиною в точці (x, y), який знаходиться лівіше і нижче цієї точки (мал. 7.1).



Мал. 7.1.

Розглянемо **властивості функції розподілу** двовимірної випадкової величини.

1°. $0 \le F(x, y) \le 1$.

2°. $F(x, y)$ є неспадною функцією по кожному аргументу, тобто

$$F(x_2, y) \ge F(x_1, y), \text{ якщо } x_2 > x_1;$$
$$F(x, y_2) \ge F(x, y_1), \text{ якщо } y_2 > y_1.$$

3°. Для $F(x, y)$ виконуються граничні співвідношення

$$\lim_{x\to-\infty} F(x,y) = \lim_{y\to-\infty} F(x,y) = \lim_{\substack{x\to-\infty\\ y\to-\infty}} F(x,y) = 0,\ \lim_{\substack{x\to+\infty\\ y\to+\infty}} F(x,y) = 1.$$

4°. Якщо один із аргументів $F(x, y)$ прямує до $+\infty$, то вона перетворюється у функцію розподілу випадкової величини, яка відповідає іншому аргументу:

$$F(x,+\infty) = F_1(x), F(+\infty, y) = F_2(y),$$

де $F_1(x)$ і $F_2(y)$ — відповідно функції розподілу випадкових величин X та Y.

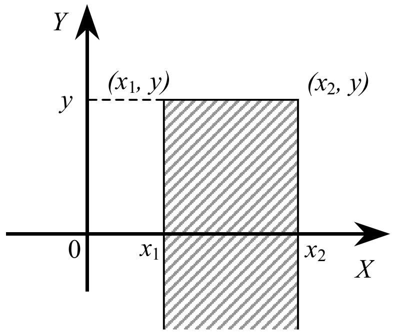
5°. $F(x, y)$ неперервна зліва по кожному аргументу.

Властивості 1°-5° є **характеристичними**. Це означає, що кожна функція $F(x, y)$, яка володіє цими властивостями, є функцією розподілу двовимірної випадкової величини (X, Y).

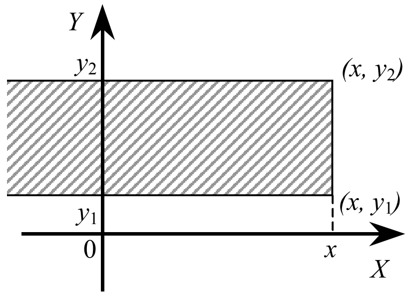
Використання розглянутих властивостей і теореми додавання імовірностей дозволяє отримати такі твердження.

Наслідок 1. Імовірність попадання при випробуванні випадкової точки (X, Y) в нескінченну напівсмугу дорівнює приросту функції розподілу по одному із аргументів (мал. 7.2, 7.3):

$$P\{(x_1 < X < x_2)\cdot(Y < y)\} = F(x_2, y) - F(x_1, y),$$
$$P\{(X < x)\cdot(y_1 < Y < y_2)\} = F(x, y_2) - F(x, y_1).$$



Мал. 7.2.

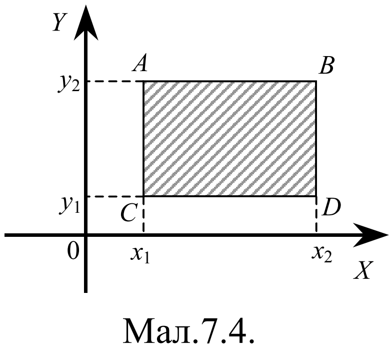


Мал. 7.3.

Наслідок 2. Імовірність попадання при випробуванні випадкової точки (X, Y) в прямокутник, сторони якого паралельні координатним осям, обчислюється за формулою

$$P\{(x_1 < X < x_2)\cdot(y_1 < Y < y_2)\} =$$
$$= F(x_2, y_2) - F(x_1, y_2) - \left(F(x_2, y_1) - F(x_1, y_1)\right).$$

Вказана імовірність знаходиться наступним чином: від імовірності попадання випадкової точки в півсмугу AB віднімається імовірність попадання випадкової точки в півсмугу CD (мал. 7.4).



Мал.7.4.

3.

Функція розподілу описує системи любих випадкових величин (як дискретних, так і неперервних). Неперервну двовимірну величину, тобто величину, складові якої є неперервними, можна задати також і густиною розподілу імовірностей. Знання густини розподілу дозволяє знайти імовірність попадання випадкової точки в області довільної конфігурації (необов'язково прямокутні, як для випадку функції розподілу), а опис розподілу системи зробити більш наочним.

Тут і надалі будемо припускати, що функція розподілу $F(x, y)$ скрізь неперервна і має скрізь, за виключенням, можливо, скінченного числа кривих, неперервну мішану частинну похідну другого порядку $\frac{\partial^2 F(x,y)}{\partial x \partial y}$.

Для відношення

$$\frac{P\{(x \le X < x + \Delta x) \cdot (y \le Y < y + \Delta y)\}}{\Delta x \Delta y},$$

імовірностний зміст якого очевидний, використаємо до чисельника формулу (7.5) та теорему Лагранжа і здійснимо граничний перехід в процесі, коли $\Delta x \to 0$, $\Delta y \to 0$. В результаті отримаємо $\frac{\partial^2 F(x,y)}{\partial x \partial y}$.

Густиною сумісного розподілу імовірностей $f(x, y)$ двовимірної неперервної випадкової величини (X, Y) називається друга мішана частинна похідна від функції розподілу:

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}. \tag{7.6}$$

Геометрично функцію $f(x, y)$ можна зобразити певною поверхнею, яку називають **поверхнею розподілу**.

Вкажемо підхід до знаходження імовірності попадання неперервної двовимірної випадкової величини (*X, Y*) в довільну область *D*. Із означення густини розподілу отримується рівність

$$P\{(x \le X < x + \Delta x) \cdot (y \le Y < y + \Delta y)\} = \left(f(x,y) + \varepsilon\right)\Delta x \Delta y, \tag{7.7}$$

де $\varepsilon \to 0$ при $\Delta x \to 0$, $\Delta y \to 0$. Наблизимо область D сукупністю n елементарних прямокутників, сторони яких паралельні осям координат в прямокутній декартовій системі координат $x0y$. Використовуючи для кожного із них рівність (7.7) і здійснюючи граничний перехід в процесі, коли максимальна довжина сторони кожного прямокутника прямує до нуля, отримаємо шукану формулу

$$P\left((X,Y) \in D\right) = \iint\limits_D f(x,y)\,dxdy. \tag{7.8}$$

Ця формула, зокрема, дозволяє знайти **функцію розподілу двовимірної випадкової величини за відомою густиною розподілу**. Згідно із (7.4) D є нескінченним квадрантом з вершиною (x, y), розташованим лівіше і нижче цієї вершини (мал. 7.1), а тому підсумком формул (7.4) та (7.8) є така рівність:

$$F(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(\xi,\zeta)\,d\xi d\zeta. \tag{7.9}$$

Сформулюємо властивості густини розподілу імовірностей двовимірної неперервної випадкової величини.

1°. $f(x, y) \geq 0.$ 2°. $\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x, y)dxdy = 1.$

Задача 7.2. Двовимірна випадкова величина (X, Y) розподілена із сталою густиною розподілу всередині квадрата D, вершини якого мають координати (0; 0), (0; 2), (2; 0), (2; 2).

Знайти густину імовірностей $f(x, y)$ і функцію розподілу.

– Згідно з умовою задачі густина розподілу імовірностей має такий вид

$$f(x, y) = \begin{cases} C, & \text{якщо } (x, y) \in D, \\ 0, & \text{якщо } (x, y) \notin D, \end{cases}$$

де C — стала. Знайдемо C, використавши властивість 2° густини розподілу

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x, y)dxdy = \iint\limits_D Cdxdy = C\int\limits_0^2 dx\int\limits_0^2 dy = 2C\int\limits_0^2 dx = 4C = 1,$$

звідки C = 1/4. Отже,

$$f(x, y) = \begin{cases} 0{,}25, & \text{якщо } (x, y) \in D, \\ 0, & \text{якщо } (x, y) \notin D. \end{cases}$$

Функцію розподілу знайдемо за формулою (7.9).

Нехай $x \leq 0$ або $y \leq 0$, тоді $f(x, y) = 0$ і

$$F(x, y) = P\{(X < x) \cdot (Y < y)\} = \int\limits_{-\infty}^{x} d\xi \int\limits_{-\infty}^{y} f(\xi, \zeta)d\zeta = 0.$$

Якщо $0 < x \leq 2$ і $0 < y \leq 2$, тоді

$$F(x, y) = P\{(X < x) \cdot (Y < y)\} = \int\limits_{-\infty}^{x} d\xi \int\limits_{-\infty}^{y} f(\xi, \zeta)d\zeta =$$

$$= \int\limits_0^x d\zeta \int\limits_0^y 0{,}25 d\xi = 0{,}25\int\limits_0^x y d\xi = 0{,}25xy.$$

Нехай $x > 2$ і $y \in (0, 2]$, тоді

$$F(x, y) = P\{(X < x) \cdot (Y < y)\} = \int\limits_{-\infty}^{x} d\xi \int\limits_{-\infty}^{y} f(\xi, \zeta)d\zeta =$$

$$= \int\limits_0^y d\zeta \int\limits_0^2 0{,}25 d\xi = 0{,}5\int\limits_0^y d\zeta = 0{,}5y.$$

В цьому випадку враховувалося те, що $f(x, y) = 0$ за межами квадрата D.

Нехай $x \in (0, 2]$ і $y > 2$, тоді

$$F(x, y) = P\{(X < x) \cdot (Y < y)\} = \int\limits_{-\infty}^{x} d\xi \int\limits_{-\infty}^{y} f(\xi, \zeta)d\zeta = 0{,}25\int\limits_0^x d\xi \int\limits_0^2 d\zeta = 0{,}5x.$$

Нарешті, якщо $x > 2$ і $y > 2$, то

$$F(x, y) = \int\limits_{-\infty}^{x} d\xi \int\limits_{-\infty}^{y} f(\xi, \zeta)d\zeta = 0{,}25\int\limits_0^2 d\xi \int\limits_0^2 d\zeta = 1.$$

Таким чином, шукана функція розподілу даної двовимірної випадкової величини має такий вид:

$$F(x, y) = \begin{cases} 0, & \text{якщо } x \leq 0 \text{ і } y \leq 0, \\ 0{,}25xy, & \text{якщо } x \in (0{,}2] \text{ і } y \in (0{,}2], \\ 0{,}5y, & \text{якщо } x > 2 \text{ і } y \in (0{,}2], \\ 0{,}5x, & \text{якщо } x \in (0{,}2] \text{ і } y > 2, \\ 1, & \text{якщо } x > 2 \text{ і } y > 2. \end{cases} \qquad 2$$

Знайдемо **густини розподілів імовірностей складових** X та Y двовимірної неперервної випадкової величини (X, Y), якщо відома її густина розподілу.

Знаючи закон розподілу двовимірної випадкової величини, можна знайти закони розподілу складових (п. 7.2). Якщо двовимірна випадкова величина (X, Y) є неперервною і її функція розподілу задається рівністю (7.9), тоді згідно із властивістю 4° функції розподілу складові X та Y описуються такими функціями розподілу

$$F_1(x) = \int_{-\infty}^{x} d\xi \int_{-\infty}^{\infty} f(\xi,\zeta)d\zeta, \tag{7.10}$$

$$F_2(y) = \int_{-\infty}^{y} d\zeta \int_{-\infty}^{\infty} f(\xi,\zeta)d\xi. \tag{7.11}$$

Диференціюючи ці вирази по відповідним змінним і використовуючи теорему про диференціювання визначеного інтеграла по змінній верхній межі, отримуємо вирази для густин розподілів імовірностей складових X та Y:

$$f_1(x) = \frac{d}{dx}F_1(x) = \frac{d}{dx}\left(\int_{-\infty}^{x} d\zeta \int_{-\infty}^{\infty} f(\xi,\zeta)d\xi \right) = \int_{-\infty}^{\infty} f(x,y)dy. \tag{7.12}$$

$$f_2(y) = \frac{d}{dy}F_2(y) = \frac{d}{dy}\left(\int_{-\infty}^{y} d\zeta \int_{-\infty}^{\infty} f(\xi,\zeta)d\xi \right) = \int_{-\infty}^{\infty} f(\xi,y)d\xi. \tag{7.13}$$

4.

Якщо відома функція розподілу двовимірної випадкової величини, тоді за формулами (7.10)-(7.13) можна знайти розподіли складових. Проте обернену задачу в загальному випадку розв’язати неможливо. Потрібно додатково знати ще залежність між величинами-складовими. Ця залежність буде охарактеризована з допомогою **умовних законів розподілу**.

Розглянемо спочатку систему двох дискретних величин (X, Y), закон розподілу якої наведений у табл. 7.1.

Відомо, що у випадку залежності двох випадкових подій A та B умовна імовірність $P_A(B)$ відрізняється від її безумовної імовірності $P(B)$. В цьому випадку із теореми 1 (§ 2, п.2.1)

$$P_A(B) = P(AB)/P(A), \tag{7.14}$$

де умовну імовірність $P_A(B)$ часто позначають ще таким чином: $P(B|A)$. Аналогічне положення має місце і для випадкових величин.

Припустимо, що в результаті випробування величина X набрала значення x_1. При цьому величина Y може набрати одне із своїх можливих значень $y_1, y_2, ..., y_n$. Позначимо умовну імовірність того, що Y набере, наприклад, значення y_1 при умові, що $X = x_1$ через $P(y_1|x_1)$, тобто в позначеннях п. 2:

$$P_{(X=x_1)}(Y = y_1) = p(y_1 \mid x_1).$$

Ця імовірність, взагалі кажучи, не дорівнює безумовній імовірності $P(Y = y_1)$.

В загальному випадку умовні імовірності складової Y позначимо таким чином:

$$p(y_j \mid x_i) \quad (j = 1,2,...,n; j = 1,2,...,m).$$

Умовним розподілом складової Y при $X = x_i$ $(i = \overline{1,m})$ називається сукупність умовних імовірностей

$$p(y_1 \mid x_i), p(y_2 \mid x_i), ..., p(y_n \mid x_i),$$

обчислених у припущенні, що випадкова подія ($X = x_i$) відбулася.

Умовним розподілом складової X при $Y = y_j$ $(j = \overline{1,n})$ називається сукупність умовних імовірностей

$$p(x_1 \mid y_j), p(x_2 \mid y_j), ..., p(x_m \mid y_j),$$

обчислених у припущенні, що випадкова подія ($Y = y_j$) відбулася.

Використовуючи формулу (7.14) для відомого закону розподілу двовимірної дискретної випадкової величини, можна обчислити умовні закони розподілу складових. Наприклад, умовний закон розподілу Y в припущенні, що подія ($X = x_i$) відбулася, знаходиться за формулою

$$p(y_j \mid x_1) = \frac{p_{1i}}{P(X = x_1)}, \quad (j = 1,2,...,n),$$

де, нагадаємо,

$$p_{1j} = P(X = x_1, Y = y_j) = P\{(X = x_1) \cdot (Y = y_j)\},$$

$$P(X = x_1) = \sum_{j=1}^{n} p_{1j}.$$

В загальному випадку умовні закони розподілу складової Y визначаються рівностями

$$p(y_j \mid x_i) = p_{ij} \Big/ \sum_{j=1}^{n} p_{ij}, \quad j = \overline{1,n}; i = \overline{1,m}. \tag{7.15}$$

а складової X:

$$p(x_i \mid y_j) = p_{ij} \Big/ \sum_{i=1}^{m} p_{ij}, \quad i = \overline{1,m}; j = \overline{1,n}. \tag{7.16}$$

Зауваження. Сума імовірностей умовного розподілу як для складової X, так і для складової Y, дорівнює одиниці. Справді, оскільки згідно із (7.3) для фіксованого y_j $\sum_{i=1}^{m} p_{ij} = P(Y = y_j)$, то

$$\sum_{i=1}^{m} p(x_i \mid y_j) = \sum_{i=1}^{m} p_{ij} / P(Y = y_j) = P(Y = y_j) / P(Y = y_j) = 1.$$

Аналогічно $\sum_{j=1}^{n} p(y_j \mid x_i) = 1.$

Ці властивості умовних розподілів використовуються для контролю обчислень.

Нехай тепер (X, Y) — неперервна двовимірна випадкова величина.

Визначимо вираз для **умовної функції розподілу складової** X:

$$F(x \mid y) = P\{(X < x) \mid (Y < y)\}.$$

З цією метою введемо в розгляд дві події $B = (X < x)$ і $A = (y \le Y < y + \Delta y)$, для яких у відповідності із (7.14) при умові $P(A) \ne 0$ отримаємо з наступним використанням формули (7.8):

$$\begin{aligned} P\{(X < x) \mid (y \le Y < y + \Delta y)\} &= \frac{P\{(X < x) \cdot (y \le Y < y + \Delta y)\}}{P(y \le Y < y + \Delta y)} = \\ &= \int_{-\infty}^{x} \int_{y}^{y+\Delta y} f(\xi, \zeta) d\xi d\zeta \Big/ \int_{-\infty}^{\infty} \int_{y}^{y+\Delta y} f(\xi, \zeta) d\xi d\zeta. \end{aligned} \tag{7.17}$$

Припустимо, що густина $f(x, y)$ неперервна в розглядуваній області.

Розділимо в рівності (7.12) чисельник і знаменник правої частини на Δy і перейдемо до границі в обох частинах рівності в процесі, коли $\Delta y \to 0$. Тоді згідно із теоремою про середнє остаточно отримаємо

$$F(x \mid y) = P\{(X < x) \mid (Y = y)\} = \int_{-\infty}^{x} f(\xi, y)d\xi / f_2(y),$$

де $f_2(y)$ визначається рівністю (7.13).

Аналогічно отримаємо вираз для **умовної функції розподілу складової** Y:

$$F(y \mid x) = P\{(Y < y) \mid (X = x)\} = \int_{-\infty}^{y} f(x, \xi)d\xi / f_1(x),$$

де $f_1(x)$ визначається рівністю (7.12).

Диференціюючи умовні функції розподілу $F(x|y)$ та $F(y|x)$ по x та y відповідно, отримуємо формули **обчислення умовних густин розподілу імовірностей (умовних законів розподілу) складових** X та Y неперервної двовимірної випадкової величини (X, Y):

$$\varphi(x \mid y) = \frac{\partial F(x \mid y)}{\partial x} = \frac{f(x, y)}{f_2(y)}; \tag{7.18}$$

$$\psi(y \mid x) = \frac{\partial F(y \mid x)}{\partial y} = \frac{f(x, y)}{f_1(x)}. \tag{7.18}$$

Таким чином, **густина розподілу імовірностей неперервної двовимірної випадкової величини** (X, Y) **дорівнює густині розподілу однієї із складових, помноженій на умовну густину розподілу іншої складової, обчислену при умові, що перша складова набрала задане значення**:

$$f(x, y) = f_1(x)\psi(y \mid x) \quad \text{або} \quad f(x, y) = f_2(y)\varphi(x \mid y).$$

Це твердження називається **теоремою множення густин імовірностей**.

Вкажемо **наступні критерії незалежності випадкових величин**.

Теорема 1. Для того, щоб випадкові величини X та Y, які є складовими двовимірної випадкової величини (X, Y), були незалежними, необхідно і достатньо, щоб функція розподілу величини (X, Y) дорівнювала добутку функцій розподілу складових:

$$F(x, y) = F_1(x)\, F_2(y).$$

Теорема 2. Для того, щоб неперервні випадкові величини X та Y, що є складовими двовимірної випадкової величини (X, Y), були незалежними, необхідно і достатньо, щоб густина розподілу системи (X, Y) дорівнювала добутку густин розподілу складових X та Y:

$$f(x, y) = f_1(x)\, f_2(y).$$

Теорема 3. Якщо двовимірна випадкова величина (X, Y) дискретна, то для незалежності її складових необхідно і достатньо виконання рівностей

$$p_{ij} = P(X = x_i)\, P(Y = y_j),$$

для довільних x_i та y_j.

5.

Важливою характеристикою умовного розподілу імовірностей є умовне математичне сподівання.

Умовним математичним сподіванням дискретної випадкової величини Y при $X = x$ (де x — певне можливе значення випадкової величини X) називається сума добутків можливих значень Y на відповідні умовні імовірності:

$$M(Y \mid X = x) = \sum_{j=1}^{n} y_j p(y_j \mid x). \tag{7.20}$$

Аналогічно визначається математичне сподівання дискретної випадкової величини X при $Y = y$:

$$M(X \mid Y = y) = \sum_{i=1}^{m} x_i p(x_i \mid y). \tag{7.21}$$

Для неперервних випадкових величин X та Y

$$M(Y \mid X = x) = \int_{-\infty}^{\infty} y\psi(y \mid x)dy, \tag{7.20*}$$

$$M(X \mid Y = y) = \int_{-\infty}^{\infty} x\varphi(x \mid y)dx, \tag{7.21*}$$

де $\psi(y|x)$ та $\varphi(x|y)$ відповідні умовні густини розподілів системи (X, Y).

Із наведених означень випливає, що при зміні значення x, взагалі кажучи, змінюється і $M(Y|X = x)$. Це означає, що $M(Y|X = x)$ можна розглядати як функцію аргумента x, тобто

$$M(Y \mid X = x) = g(x). \tag{7.22}$$

Функція $g(x)$ називається **функцією регресії *Y* на *X***, рівняння (7.22) називається **рівнянням регресії *Y* на *X***, а графік функції $g(x)$ — **лінією регресії *Y* на *X***.

Аналогічно умовне математичне сподівання $M(X|Y = y)$ можна розглядати як функцію аргументу:

$$M(X \mid Y = y) = q(y). \tag{7.23}$$

Функція $q(y)$ називається **функцією регресії *X* на *Y***, рівняння (7.23) називається **рівнянням регресії *X* на *Y***, а графік функції $q(y)$ — **лінією регресії *X* на *Y***.

6.

Нехай (X, Y) — двовимірна дискретна випадкова величина, закон розподілу якої заданий табл. 7.1. Тоді **математичне сподівання для кожної складової** визначається рівностями:

$$M(X) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i p_{ij}, \qquad M(Y) = \sum_{i=1}^{m}\sum_{j=1}^{n} y_j p_{ij}.$$

Математичне сподівання для складових системи двох неперервних випадкових величин визначається за формулами:

$$M(X) = \int_{-\infty}^{\infty} xf_1(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy,$$

$$M(Y) = \int_{-\infty}^{\infty} yf_2(y)dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy,$$

де $f(x, y)$ — густина розподілу імовірностей системи (X, Y), $f_1(x)$ та $f_2(y)$ — густини розподілів складових X та Y відповідно.

Дисперсії складових системи (X, Y) визначаються формулами:

$$D(X)=\begin{cases}\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}\left[(x_i-M(X)\right]^2 p_{ij} & \text{- для дискретної в.в.}(X,Y);\\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\left[(x_i-M(X)\right]^2 f(x,y)dxdy & \text{- для неперервної в.в.}(X,Y);\end{cases}$$

$$D(X)=\begin{cases}\sum\limits_{i=1}^{m}x_i^2P(X=x_i)-\left[M(X)\right]^2 & \text{- для дискретної в.в.}(X,Y),\ \text{якщо відомі імовірності } P(X=x_i);\\ \int\limits_{-\infty}^{\infty}x^2 f_1(x)dx-\left[M(X)\right]^2 & \text{- для неперервної в.в.}(X,Y),\ \text{якщо відома густина } f_1(\text{x});\end{cases}$$

$$D(Y)=\begin{cases}\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}\left[(y_i-M(Y)\right]^2 p_{ij} & \text{- для дискретної в.в.}(X,Y);\\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\left[(y-M(Y)\right]^2 f(x,y)dxdy & \text{- для неперервної в.в.}(X,Y);\end{cases}$$

$$D(Y)=\begin{cases}\sum\limits_{i=1}^{m}y_j^2P(Y=y_j)-\left[M(X)\right]^2 & \text{- для дискретної в.в.}(X,Y),\ \text{якщо відомі імовірності } P(Y=y_j);\\ \int\limits_{-\infty}^{\infty}y^2 f_2(y)dy-\left[M(Y)\right]^2 & \text{- для неперервної в.в.}(X,Y),\ \text{якщо відома густина } f_2(\text{y}).\end{cases}$$

Сукупність математичних сподівань $M(X)$ і $M(Y)$ являє собою характеристику положення системи (X, Y). Геомерично — це координати «середньої» точки на площині, навколо якої здійснюється розсіювання значень системи (X, Y) як випадкової точки. Числові характеристики $D(X)$ та $D(Y)$ характеризують розсіювання випадкової точки в напрямку осей $0x$ та $0y$ відповідно.

Серед числових характеристик двовимірної випадкової величини особливу роль відіграють кореляційний момент та коефіцієнт кореляції, які описують зв'язок між складовими цієї величини.

Кореляційним моментом (**моментом зв'язку**) системи двох випадкових величин (X, Y) називається математичне сподівання добутку відхилень цих величин від своїх математичних сподівань:

$$K_{XY}=M\left[(X-M(X))(Y-M(Y))\right].$$

Використання властивостей математичного сподівання дозволяє записати кореляційний момент у такому вигляді:

$$K_{XY}=M(XY)-M(X)M(Y).$$

Для обчислення кореляційного момента дискретних випадкових величин користуються формулами:

$$K_{XY}=\sum_{i=1}^{m}\sum_{j=1}^{n}\left(x_i-M(X)\right)\left(y_j-M(Y)\right)p_{ij},$$
$$K_{XY}=\sum_{i=1}^{m}\sum_{j=1}^{n}x_i y_j p_{ij}-M(X)M(Y), \tag{7.24}$$

а для неперервних випадкових величин — формулами:

$$K_{XY}=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\left(x-M(X)\right)\left(y-M(Y)\right)f(x,y)dxdy,$$
$$K_{XY}=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}xyf(x,y)dxdy-M(X)M(Y). \tag{7.25}$$

Розмірність кореляційного моменту дорівнює добутку розмірностей величин X та Y. Тому для одних і тих же величин величина K_{XY} має різні значення в залежності від того, в яких одиницях вимірюються ці величини. Для усунення цього недоліку вводиться нова числова характеристика — коефіцієнт кореляції.

Коефіцієнтом кореляції r_{XY} випадкових величин X та Y називається відношення кореляційного момента до добутку середніх квадратичних відхилень цих величин:

$$r_{xy} = K_{XY} / (\sigma_x \sigma_y).$$

Властивості коефіцієнта кореляції.

1°. r_{xy} є безрозмірною величиною.

2°. $|r_{xy}| \le 1.$

3°. Якщо $r_{xy} = \pm 1,$ то між складовими X та Y випадкової величини (X, Y) існує лінійна функціональна залежність: $Y = \alpha_1 X + \alpha_0$, де α_1, α_0 — дійсні числа.

4°. Якщо $r_{xy} \neq 0,$ то складові X та Y двовимірної випадкової величини (X, Y) залежні.

Коефіцієнт кореляції r_{XY} характеризує силу лінійного зв'язку між випадковими величинами X та Y: чим ближче $|r_{xy}|$ до одиниці, тим тісніший лінійний зв'язок; чим ближче $|r_{xy}|$ до нуля, тим слабший лінійний зв'язок.

Випадкові величини X та Y нахиваються **корельованими**, якщо $r_{xy} \neq 0,$ і **некорельованими**, якщо $r_{xy} = 0.$

Згідно із властивістю 4° коефіцієнта кореляції **із корельованості двох випадкових величин випливає їх залежність**. Проте **із залежності випадкових величини ще не слідує їх корельованість**, тобто вони можуть бути або корельованими, або некорельованими.

З другого боку, **із незалежності двох випадкових величин випливає їх некорельованість**, але **із некорельованості випадкових величини ще не можна зробити висновок про їх незалежність**.

Ілюструє останнє твердження

Задача 7.8. Двовимірна випадкова величина (X, Y) розподілена рівномірно в крузі $x^2 + y^2 \le r^2.$ Довести, що X та Y — залежні випадкові величини, але некорельовані.

– За умовою задачі густина розподілу величини (X, Y) має такий вид:

$$f(x, y) = \begin{cases} C = \text{const} & \text{при } x^2 + y^2 \le r^2, \\ 0 & \text{при } x^2 + y^2 > r^2. \end{cases}$$

Використавши властивість 2° густини розподілу імовірності, одержимо $C = 1/(\pi r^2).$ За формулами (7.12), (7.13) знайдемо густину розподілу кожної складової:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_{-r}^{r} \left(1/(\pi r^2)\right) dy = \left(1/(\pi r^2)\right) y\Big|_{-r}^{r} = 2/(\pi r);$$

$$f_2(y)=\int_{-\infty}^{\infty} f(x,y)dx = 2/(\pi r).$$

Оскільки рівність $f(x, y) = f_1(x) f_2(y)$ не виконується, то робимо висновок, що X та Y є залежними випадковими величинами.

Кореляційний момент системи (X, Y) обчислимо за другою із формул (7.25):

$$K_{XY}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(xy)dxdy - M(X)M(Y)=$$

$$=\frac{1}{\pi r^2}\int_{-r}^{r} xdx \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} ydy - \frac{1}{\left(\pi r^2\right)^2}\int_{-r}^{r} xdx \int_{-r}^{r} ydy = 0.$$

Отже, $r_{XY}=0$ і величини X та Y є некорельованими. 2

7.

Розглянуті вище двовимірні випадкові величини дозволяють природнім чином узагальнити методи їх задання та вивчення властивостей на випадок довільного скінченного числа складових. В зв'язку із цим ми обмежимося лише мінімальною інформацією, звернувши особливу увагу тільки на деякі із основних особливостей, що виникають внаслідок збільшення вимірності випадкової величини в порівнянні із двовимірним випадком.

Система випадкових величин $X_1, X_2, ..., X_n$ інтерпретується як випадкова точка або випадковий вектор в n-вимірному просторі. Описання цієї системи може здійснюватися функцією розподілу, густиною розподілу імовірностей або розподілом імовірностей $P_{X_1X_2...X_n}$ n-вимірної випадкової величини.

Функцією розподілу системи *n* випадкових величин $(X_1, X_2, ..., X_n)$ називається функція детермінованих аргументів $x_1, x_2, ..., x_n$, яка дорівнює імовірності сумісного виконання нерівностей $X_i < x_i$ $(i=\overline{1,n})$:

$$F(x_1,x_2,...,x_n)=P\{(X_1<x_1)\cdot(X_2<x_2)\cdot...\cdot(X_n<x_n)\},$$

де крапками позначено добуток (перетин) подій.

Ця функція є неспадною по кожному із аргументів x_i, при решті фіксованих. Вона прямує до нуля, якщо хоча б один аргумент прямує до $-\infty$.

Якщо із системи випадкових величин $X_1, X_2, ..., X_n$ виділити деяку підсистему, наприклад $X_1, X_2, ..., X_k$, де $k < n$, то функція розподілу для цієї підсистеми матиме вигляд $F(x_1, x_2, ..., x_k, \infty, ..., \infty)$. Зокрема, функцію розподілу для кожної із величин, що входять до складу системи, одержимо, якщо вся решта аргументів прямуватиме до $+\infty$. Наприклад, $F(x_1) = F(x_1, \infty, ..., \infty)$. Якщо всі аргументи функції розподілу прямують до ∞, то вона прямує до 1.

Густина розподілу імовірностей *n*-вимірної неперервної випадкової величини $(X_1, X_2, ..., X_n)$ визначається рівністю

$$f(x_1,x_2,...,x_n)=\frac{\partial^n F(x_1,x_2,...,x_n)}{\partial x_1 \partial x_2 ... \partial x_n}.$$

Ця функція невід'ємна і для неї виконується умова нормування

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} f(x_1,x_2,...,x_n)dx_1dx_2...dx_n = 1.$$

Густина розподілу для підсистеми $(X_1, X_2, ..., X_k)$, де $k < n$, визначається рівністю

$$f(x_1, x_2, ..., x_k) = \underbrace{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}}_{n-k} f(x_1, x_2, ..., x_n)dx_{k+1}dx_{k+2}...dx_n.$$

Умовним законом розподілу підсистеми випадкових величин $(X_1, X_2, ..., X_k)$ називається такий закон, який отримується за умови, що решта з них $(X_{k+1}, X_{k+2}, ..., X_n)$ набули своїх можливих значень $x_{k+1}, x_{k+2}, ..., x_n$.

Умовна густина розподілу імовірностей визначається за формулою

$$f(x_1, x_2, ..., x_k \mid x_{k+1}, ..., x_n) = \frac{f(x_1, x_2, ..., x_n)}{f(x_{k+1}, x_{k+2}, ..., x_n)}.$$

Необхідною і достатньою умовою незалежності складових n-вимірної випадкової величини $(X_1, X_2, ..., X_n)$ є виконання рівності

$$f(x_1, x_2, ..., x_n) = f(x_1)f(x_2)...f(x_n).$$

Розглянемо числові характеристики n-вимірних випадкових величин.

Математичне сподівання неперервної величини X_i, що входить до системи, обчислюється за формулою

$$M(X_i) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} x_i f(x_1, x_2, ..., x_n)dx_1dx_2...dx_n, \quad (i = \overline{1, n}),$$

а дисперсія цієї ж величини (розрахункова формула) —

$$D(X_i) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} x_i^2 f(x_1, x_2, ..., x_n)dx_1dx_2...dx_n - [M(X_i)]^2.$$

Кореляційні моменти для кожної пари (X_i, Y_j) складових n-вимірної випадкової величини $(X_1, X_2, ..., X_n)$ обчислюється за формулою

$$K_{X_iY_j} = K_{ij} = M\{(X_i - M(X_i))(X_j - M(X_j))\}.$$

Зокрема, якщо $i = j$, то $K_{ij} = D(X_i) = \sigma_i^2$.

Всі кореляційні моменти випадкової величини $(X_1, X_2, ..., X_n)$ розміщують у вигляді квадратної таблиці, яка називається **кореляційною матрицею системи *n* випадкових величин**:

$$\begin{pmatrix} K_{11} & K_{12} & K_{13} & ... & K_{1n} \\ K_{21} & K_{22} & K_{23} & ... & K_{2n} \\ ... & ... & ... & ... & ... \\ K_{n1} & K_{n2} & K_{n3} & ... & K_{nn} \end{pmatrix}.$$

З врахуванням рівності $K_{ij} = K_{ji}$, $K_{ii} = \sigma_i^2$, кореляційна матриця набирає такого виду:

$$\begin{pmatrix} \sigma_1^2 & K_{12} & K_{13} & ... & K_{1n} \\ K_{12} & \sigma_2^2 & K_{23} & ... & K_{2n} \\ ... & ... & ... & ... & ... \\ K_{1n} & K_{2n} & K_{3n} & ... & \sigma_n^2 \end{pmatrix}.$$

У випадку некорельованості кожної пари системи $(X_1, X_2, ..., X_n)$ кореляційна матриця стає діагональною:

$$\begin{pmatrix} \sigma_1^2 & 0 & 0 & ... & 0 \\ 0 & \sigma_2^2 & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & \sigma_n^2 \end{pmatrix}.$$

Знаючи кореляційні моменти, можна визначити **коефіцієнти кореляції для кожної пари випадкових величин *n*-вимірної випадкової величини**. Такі коефіцієнти кореляції називаються парними і визначаються за формулою

$$r_{ij} = \frac{K_{ij}}{\sigma_i \sigma_j}.$$

Із парних коефіцієнтів кореляції утворюється квадратна матриця, яка називається **нормованою кореляційною матрицею**:

$$\begin{pmatrix} 1 & r_{12} & r_{13} & \dots & r_{1n} \\ r_{12} & 1 & r_{23} & \dots & r_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ r_{1n} & r_{2n} & r_{3n} & \dots & 1 \end{pmatrix}.$$

8.

На практиці часто доводиться мати справу із двовимірними випадковими величинами, розподіленими за нормальним законом.

Розподіл імовірностей системи (*X, Y*) називається нормальним (двовимірна випадкова величина розподілена за нормальним законом), якщо густина розподілу ймовірностей

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1 - r_{xy}^2}} e^{-\frac{1}{2(1 - r_{xy}^2)}\left[\frac{(x-a)^2}{\sigma_x^2} + \frac{(y-b)^2}{\sigma_y^2} - 2r_{xy}\frac{(x-a)(y-b)}{\sigma_x\sigma_y}\right]},$$

де п'ять параметрів розподілу мають такий імовірносний зміст: $a = M(X),\ b = M(Y),\ \sigma_x = \sqrt{D(X)},\ \sigma_y = \sqrt{D(Y)},\ r_{xy}$ — коефіцієнт кореляції величин X та Y.

Можна довести, що якщо система (X, Y) розподілена за нормальним законом із параметрами $a, b, \sigma_x, \sigma_y, r_{xy}$, то її складові також розподілені за нормальним законом з параметрами, рівними відповідно a, σ_x та b, σ_y.

Вище (див. п. 7.6) було показано, що із некорельованості двох випадкових величин, які є складовими двовимірної величини, ще не слідує (в загальному випадку) незалежність цих величин. В зв'язку із цим доцільно з'ясувати співвідношення між вказаними властивостями складових двовимірної випадкової величини, розподіленої за нормальним законом.

Нехай X та Y некорельовані, тобто $r_{xy} = 0$. Тоді

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left[\frac{(x-a)^2}{\sigma_x^2} + \frac{(y-b)^2}{\sigma_y^2}\right]} =$$

$$= \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma_x^2}} \cdot \frac{1}{\sigma_y\sqrt{2\pi}} e^{-\frac{(x-b)^2}{2\sigma_y^2}} = f_1(x) \cdot f_2(y),$$

де $f_1(x)$ і $f_2(y)$ густини розподілу імовірностей випадкових величин X та Y відповідно.

Отже, при $r_{xy} = 0$ отримуємо рівність $f(x, y) = f_1(x)\, f_2(y)$, яка означає, що складові X та Y є незалежними випадковими величинами. Якщо ж X та Y є незалежними випадковими величинами, то вони є і некорельованими (п. 7.6).

Лекція 8.

Тема: ***Функція випадкових величин.***

1. ***Функція одного випадкового аргументу та її математичне сподівання.***
2. ***Функції двох випадкових величин.***
3. ***Розподіл С'юдента, розподіл Фішера-Снедокора.***

1.

Якщо кожному можливому значенню випадкової величини (одновимірної) X відповідає одне можливе значення випадкової величини Y, то Y називається функцією випадкового аргумента X:

$$Y = \varphi(X). \tag{8.1}$$

Якщо **аргумент** X — **дискретна випадкова величина**, закон розподілу імовірностей якої має такий вид:

$$\begin{array}{c|cccc} X & x_1 & x_2 & \dots & x_n \\ \hline P & p_1 & p_2 & \dots & p_n \end{array}.$$

то Y також дискретна випадкова величина.

Оскільки випадкові події $(X = x_i)$ та $(Y = \varphi(x_i))$, де $i = \overline{1,n}$, рівносильні, то їх імовірності рівні. Тому випадкова величина Y розподілена за таким законом:

$$\begin{array}{c|cccc} Y & \varphi(x_1) & \varphi(x_2) & \dots & \varphi(x_n) \\ \hline P & p_1 & p_2 & \dots & p_n \end{array}. \tag{8.2}$$

Якщо серед можливих значень Y є однакові, то в остаточному законі розподілу слід залишити одне із них, поставивши йому у відповідність суму імовірностей повторюваних значень Y.

Використавши означення математичного сподівання до закону розподілу (8.2), отримаємо:

$$M(Y) = \sum_{i=1}^{n} \varphi(x_i) p_i. \tag{8.3}$$

Нехай **аргумент** X **функції** (8.1) **є неперервною випадковою величиною**. Намітимо підхід до знаходження законів розподілу (функції розподілу та густини розподілу імовірностей) випадкової величини Y у загальному випадку функції φ, що дозволить оцінити труднощі, які виникають на цьому шляху, а також виділити простіші випадки, що часто зустрічаються на практиці.

Функція розподілу випадкової величини *Y*, яка пов'язана із величиною X функціональною залежністю (8.1) має вигляд

$$G(y) = P(Y < y) = \sum_i P(X \in \alpha_i(y)) = \sum_i \int_{c_i(y)}^{d_i(y)} f(x)dx, \tag{8.4}$$

де $\alpha_i(y) = (c_i(y); d_i(y)) = \{x : \varphi(x) < y\}$, $f(x)$ — функція густини розподілу імовірностей випадкової величини X.

Відмітимо, що межі інтервалів $\alpha_i(y)$ можуть бути виражені як явні функції y для кожної конкретної функції $y = \varphi(x)$.

Густину розподілу імовірностей $g(y)$ величини Y можна знайти, продиференціювавши ліву і праву частини рівностей (8.4).

Розглянутий загальний випадок функції $y = \varphi(x)$ дозволяє очікувати, що випадок монотонної функції φ суттєво спростить знаходження законів розподілу випадкової величини Y.

Якщо $y = \varphi(x)$ — диференційована строго зростаюча або строго спадаюча функція, обернена до якої $x = \psi(y)$, то густина розподілу імовірностей випадкової величини Y має такий вид:

$$g(y) = f[\psi(y)]\,|\psi'(y)|. \tag{8.5}$$

Відмітимо, що для знаходження $M(X)$ не обов'язково знати густину розподілу величини Y. По аналогії із рівністю (8.3) можна отримати, що

$$M(Y) = \int_{-\infty}^{\infty} \varphi(x) f(x)dx.$$

Якщо ж $g(y)$ відома, тоді має місце така рівність:

$$M(Y) = \int_{-\infty}^{\infty} y g(y)dy.$$

Дисперсія функції розподілу випадкової величини *Y*, яка пов'язана із величиною X функціональною залежністю (8.1) визначається формулою:

$$D(Y) = \begin{cases} \sum_{i=1}^{n} [\varphi(x_i) - M(Y)]^2 p_i & - \text{для дискретної в.в.}(X); \\ \int_{-\infty}^{\infty} [\varphi(x) - M(Y)]^2 f(x)dx & - \text{для неперервної в.в.}(X). \end{cases}$$

Якщо ж відома $g(y)$, тоді має місце рівність:

$$D(Y) = \int_{-\infty}^{\infty} [y - M(Y)]^2 g(y)dy.$$

Випадкова величина Y називається **логарифмічно нормально** (**логнормально**) **розподіленою**, якщо її густина розподілу імовірностей має вид:

$$g(y) = \begin{cases} 0, & \text{якщо } y \le 0, \\ \dfrac{1}{y\sigma_x\sqrt{2\pi}} e^{-\frac{(\ln y - a_x)^2}{2\sigma_x^2}}, & \text{якщо } y > 0. \end{cases} \tag{8.6}$$

Випадкова величина Y має **розподіл** χ, якщо її густина розподілу імовірностей має вид:

$$g(y) = \begin{cases} 0, & \text{якщо } y \le 0, \\ \dfrac{1}{2^{k/2-1}\Gamma(k/2)} y^{k-1} e^{-y^2/2}, & \text{якщо } y > 0. \end{cases} \tag{8.7}$$

2.

Розглянемо систему двох неперервних випадкових величин (X, Y) із густиною розподілу імовірностей $f(x, y)$. Нехай випадкова величина Z пов'язана із випадковими величинами X та Y функціональною залежністю $Z = \varphi(X, Y)$.

Функція розподілу випадкової величини Z визначається за формулою

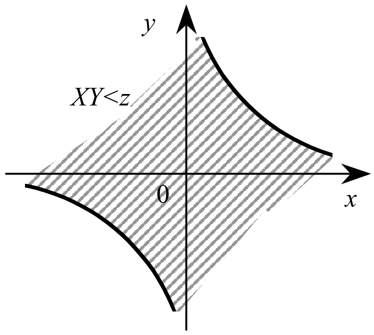
$$G(z) = P(\varphi(X,Y) < z) = P((X,Y) \in D) = \iint_D f(x,y)dxdy, \qquad (8.8)$$

де $D = \{(x,y) : \varphi(x,y) < z\}$.

Величина z входить в праву частину рівності (8.8) неявно через межі інтегрування. Явний вид функції $z = \varphi(x, y)$ дозволяє в принципі знайти ці межі як функцію z.

Густина розподілу імовірностей випадкової величини Z є похідною від функції розподілу: $g(z) = G'(z)$.

Якщо, наприклад, $Z = XY$, то для визначення функції розподілу і густини розподілу потрібно побудувати криву $xy = z$ на площині $x0y$ і вибрати ту її частину, для якої $XY < z$. В розглянутому випадку — це заштрихована частина на мал. 8.1.



Мал. 8.1.

Тоді

$$G(z) = \int_{-\infty}^{0} dx \int_{z/x}^{\infty} f(x,y)dy + \int_{0}^{\infty} dx \int_{\infty}^{z/x} f(x,y)dy;$$

$$g(z) = G'(z) = -\int_{-\infty}^{0} \frac{1}{x} f\left(x, \frac{z}{x}\right)dx + \int_{0}^{\infty} \frac{1}{x} f\left(x, \frac{z}{x}\right)dx.$$

Математичне сподівання та дисперсія функції двох випадкових величин $Z = \varphi(X, Y)$ визначаються формулами:

$$M(Z) = \begin{cases} \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} \varphi(x_i, y_j) p_{ij} & \text{– для дискретних в.в. } X \text{ та } Y; \\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \varphi(x,y) f(x,y)dxdy & \text{– для неперервних в.в. } X \text{ та } Y; \end{cases}$$

$$D(Z) = \begin{cases} \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} [\varphi(x_i, y_j) - M(Z)]^2 p_{ij} & \text{– для дискретних в.в. } X \text{ та } Y; \\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} [\varphi(x,y) - M(Z)]^2 f(x,y)dxdy & \text{– для неперервних в.в. } X \text{ та } Y, \end{cases}$$

де p_{ij} — імовірність того, що вектор (X, Y) набере значення (x_i, y_j) при $i = \overline{1,n}, \quad j = \overline{1,m}$.

3.

Нехай випадкова величина X має розподіл $\chi/\sqrt{k}$, випадкова величина Y розподілена за нормованим нормальним законом із параметрами $M(Y) = 0$, $\sigma(Y) = 1$. Якщо X та Y незалежні випадкові величини, тоді **випадкова величина $Z = Y/X$ називається розподіленою за законом Ст'юдента (t-розподіл)**.

Густина розподілу імовірностей Ст'юдента має вид:

$$g(z)=\frac{1}{\sqrt{\pi k}}\cdot\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma(\mathrm{k}/2)}\left(1+z\,2\!\!\diagup_{k}\right)^{-(k+1)/2},\quad -\infty<z<\infty.$$

Розподіл Ст'юдента має k ступенів вільності і при зростанні k він швидко наближається до нормального.

Якщо Y та X — незалежні випадкові величини, розподілені за законом χ^2 із ступенями вільності k_1 та k_2 відповідно, тоді величина

$$F=\frac{Y/k_1}{X/k_2}$$

має розподіл, який називається **розподілом *F* Фішера-Снедокора із ступенями вільності k_1 та k_2.**

Густина цього розподілу

$$g(z)=\begin{cases}0, & \text{при } z\le 0,\\ C_0\,\dfrac{z^{(k_1-2)/2}}{(k_2+k_1 z)^{(k_2+k_1)/2}} & \text{при } z>0,\end{cases}$$

де

$$C_0=\frac{\Gamma\left(\frac{k_1+k_2}{2}\right)k_1^{k_1/2}k_2^{k_2/2}}{\Gamma(k_1/2)\Gamma(k_2/2)}.$$

Лекція 9.

Тема: ***Закон великих чисел.***

1. ***Лема та нерівність Чебишева.***
2. ***Теорема Чебишева (стійкість середніх).***
3. ***Теорема Бернуллі (стійкість відносних частот).***
4. ***Центральна гранична теорема Ляпунова.***

1.

Лема Чебишева. Якщо всі можливі значення випадкової величини Y невід'ємні, тоді імовірність того, що вона при випробуванні набере значення, більше від додатного числа b, не більша від дробу, чисельник якого — математичне сподівання від Y, а знаменник — число b:

$$P(Y > b) \le M(Y)/b. \qquad (9.1)$$

Нерівність Чебишева. Імовірність того, що відхилення випадкової величини X від її математичного сподівання за абсолютною величиною не більше від додатного числа ε, не менша, ніж $1 - D(X)/\varepsilon^2$:

$$P\left(\left|X - M(X)\right| \le \varepsilon \right) \ge 1 - D(X)/\varepsilon^2. \qquad (9.2)$$

Зауваження. Якщо випадкова величина X розподілена за біноміальним законом (X — число появи події в n повторних незалежних випробуваннях), тоді нерівність Чебишева набере такого виду:

$$P\left(\left|m - np\right| \le \varepsilon \right) \ge 1 - \frac{npq}{\varepsilon^2} \qquad (9.3)$$

або

$$P\left(\left|\frac{m}{n} - p\right| \le \varepsilon \right) \ge 1 - \frac{pq}{n\varepsilon^2}, \qquad (9.4)$$

де m — число появи події A в n повторних незалежних випробуваннях, $p = P(A)$, m/n — відносна частота появи події A.

Доведення нерівності Чебишева не складно провести за допомогою леми Чебишева, розглянувши нерівність $\left(X - M(X)\right)^2 \le \varepsilon^2$.

Приклад 1. Випадкова величина задана законом розподілу

X	0,3	0,6
P	0,2	0,8

.

Використовуючи нерівність Чебишева, оцінити імовірність того, що X відхилиться від свого математичного сподівання на величину, яка за абсолютною величиною не перевищить 0,2. Перевірити точність отриманої оцінки.

○ Знайдемо $M(X)$ та $D(X)$

$M(X) = 0,3 \cdot 0,2 + 0,6 \cdot 0,8 = 0,54,$

$D(X) = M(X^2) - [M(X)]^2 = 0{,}3^2 \cdot 0{,}2 + 0{,}6^2 \cdot 0{,}8 - (0{,}54)^2 = 0{,}0144.$Використавши нерівність (7.2) з ε = 0,3, отримаємо

$P\left(|X - 0{,}54| \le 0{,}2 \right) \ \ge 1 - 0{,}0144/0{,}04 = 0{,}64.$

Випадкова подія $|X - 0{,}54| \le 0{,}2$ рівносильна події $0{,}34 \le X \le 0{,}74$, яка відбувається із врахуванням закону розподілу, тоді і тільки тоді, коли X=0,6. Тому $P\left(|X - 0{,}54| \le 0{,}2 \right) = P(0{,}34 \le X \le 0{,}74) = P(X = 0{,}6) = 0{,}8.$

Отже, нижня межа оцінки імовірності дорівнює 0,64, а справжнє значення імовірності дорівнює 0,8. ●

Приклад 2. Імовірність появи події в кожному випробуванні дорівнює 0,25. Використовуючи нерівність Чебишева, оцінити імовірність того, що число X появи події міститься в межах від 150 до 250 включно, якщо буде проведено 800 випробувань. Перевірити точність отриманої оцінки.

○ Використаємо нерівність (7.3), де $X = m$, $np = 800 \cdot 0{,}25 = 200$ і попередньо треба знайти ε. Випадкову подію ($150 \le m \le 250$) можна записати в такій рівносильній формі: $\left(|m - 200| \le 50 \right)$. А тому в (7.3) ε = 50 і

$$P(150 \le m \le 250) = P\left(|m - 200| \le 50 \right) \ \ge 1 - \frac{800 \cdot 0{,}25 \cdot 0{,}75}{50^2} = 0{,}94.$$

Тобто, остаточно $P(150 \le m \le 250) \ge 0{,}94$.

Перевіримо точність отриманого результату, використавши інтегральну формулу Лапласа (3.7) ($npq = 800 \cdot 0{,}25 \cdot 0{,}75 = 150 >> 9$):

$P_n(m_1 \le m \le m_2) \approx \Phi(x_2) - \Phi(x_1)$,

де $x_2 = \dfrac{m_2 - np}{\sqrt{npq}} = \dfrac{250 - 800 \cdot 0{,}25}{\sqrt{150}} = \dfrac{50}{12{,}247} = 4{,}08$,

$x_1 = \dfrac{m_1 - np}{\sqrt{npq}} = \dfrac{150 - 200}{\sqrt{150}} = \dfrac{-50}{12{,}247} = -4{,}08$,

$P_n(150 \le m \le 250) \approx \Phi(4{,}08) - \Phi(-4{,}08) = 2\Phi(4{,}08) = 2 \cdot 0{,}4999992 = 0{,}9999984$.

Отже, нижня межа оцінки імовірності дорівнює 0,94, а справжнє значення імовірності дорівнює 0,9999984. ●

2.

Теорема Чебишева. Якщо $X_1, X_2, ..., X_n$ попарно незалежні випадкові величини, дисперсії яких рівномірно обмежені $\left(D(X_i) \le C, i = \overline{1,n} \right)$, тоді для довільного ε > 0 і достатньо великого n імовірність події

$$\left| \frac{X_1 + X_2 + ... + X_n}{n} - \frac{M(X_1) + M(X_2) + ... + M(X_n)}{n} \right| \le \varepsilon \qquad (9.5)$$

буде як завгодно близькою до одиниці.

При доведені теореми встановлюється правильність нерівності

$$P\left(\left|\frac{X_1+X_2+...+X_n}{n}-\frac{M(X_1)+M(X_2)+...+M(X_n)}{n}\right|\le\varepsilon\right)\ge 1-C\big/(n\varepsilon^2). \quad (9.6)$$

Приклад 1. Дисперсія кожної із 2500 незалежних випадкових величин не перевищує 5. Оцінити імовірність того, що абсолютна величина відхилення середнього арифметичного цих випадкових величин від середнього арифметичного їх математичних сподівань не перевищить 0,4.

○ Для задачі виконуються обидві умови теореми Чебишева, а тому можна використати нерівність (7.6), де n = 2500, ε = 0,4, C = 5. Тоді

$$P\left(\left|\left(\sum_{i=1}^{2500}X_i-\sum_{i=1}^{2500}M(X_i)\right)\Big/2500\right|\le 0,4\right)\ge 1-\frac{5}{2500(0,4)^2}=\frac{79}{80}.$$

Отже, шукана імовірність оцінюється знизу числом 79/80. ●

Приклад 2. Випадкові величини X_1, X_2, …, X_n, … є незалежними, причому для k=1, 2, …, n, … закон розподілу величини X_k має такий вид:

X_k	$-\sqrt{k}$	0	$\sqrt{k}$
P	$1/(4k)$	$1-2/k$	$1/(4k)$

.

Чи можна використати теорему Чебишева до цієї послідовності випадкових величин?

○ Можливість використання теореми Чебишева означає виконання обох умов теореми відносно послідовності випадкових величин. Згідно з умовою задачі перша з них виконується. Перевіримо виконання другої умови – рівномірну обмеженість дисперсій випадкових величин. Для цього знайдемо $D(X_k)$, де k=1, 2, …, n, … .

Оскільки $M(X_k)=\dfrac{-\sqrt{k}}{4k}+0\cdot(1-\dfrac{2}{k})+\dfrac{\sqrt{k}}{4k}=0$, то за розрахунковою формулою

$$D(X_k)=M(X_k^2)-\left[M(X_k)\right]^2=\left(-\sqrt{k}\right)^2\cdot\frac{1}{4k}+0\cdot\left(1-\frac{2}{k}\right)+\left(-\sqrt{k}\right)^2\cdot\frac{1}{4k}=0,5.$$

Отже, дисперсії випадкових величин X_1, X_2, …, X_n, … обмежені в сукупності сталою C=0,5, а тому до даної послідовності випадкових величин можна використати теорему Чебишева. ●

3.

Теорема Бернуллі. Якщо в кожному із n повторних випробувань імовірність p появи події A стала, тоді як завгодно близька до одиниці імовірність того, що відхилення відносної частоти події A від імовірності p за абсолютною величиною буде як завгодно малим, якщо число випробувань достатньо велике.

З нерівності (9.6) можна отримати нерівність

$$P\left(\left|m\big/n-p\right|\le\varepsilon\right)\ge 1-\frac{0,25}{n\varepsilon^2}. \quad (9.7)$$

Приклад 3. Відомо, що 80% виробів механічного цеху є першосортними. Оцінити імовірність того, що відносна частоту виробів першого сорту серед 20 000 виготовлених відрізнятиметься від імовірності виготовлення виробу першого сорту не більше, ніж на 0,02 в той чи інший бік. Перевірити точність отриманої оцінки.

○ Подія A — виготовлений виріб першосортний. Тоді $P(A) = 0{,}8$, $n = 20\ 000$, $\varepsilon = 0{,}02$. Використаємо нерівність (7.7):

$$P\left(\left| m/n - 0{,}8 \right| \le 0{,}02 \right) \ge 1 - \frac{0{,}25}{20000 \cdot 0{,}0004} = 31/32.$$

Таким чином, шукана імовірність оцінюється знизу числом 31/32.

Для перевірки точності отриманої оцінки використаємо формулу (3.8):

$$P\left(\left| m/n - p \right| \le \varepsilon \right) \approx 2\Phi\left(\varepsilon\sqrt{n/(pq)} \right),$$

$\varepsilon=0{,}02$, $n=20000$, $p=0{,}8$, $q=0{,}2$,

$$P\left(\left| m/n - 0{,}8 \right| \le 0{,}02 \right) \approx 2\Phi\left(0{,}02\sqrt{\frac{20000}{0{,}8 \cdot 0{,}2}} \right) = 2\Phi(7{,}07) = 2 \cdot 0{,}5 = 1.$$

Оскільки npq=3200 і 7,07>5, то похибки у наближеній рівності (3.8) і $\Phi(7{,}07) \approx 0{,}5$ є надзвичайно малими, тобто подія $\left(\left| m/n - 0{,}8 \right| \le 0{,}02 \right)$ є майже достовірною. ●

4.

В теоремі Чебишева не використовувалася інформація про закони розподілу випадкових величин. Разом з тим для задач теорії і практики важливим є таке питання: за яким законом розподіляється сума достатньо великого числа випадкових величин?

Випадкову величину

$$Z_n = \left[\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} M(X_i) \right] \Big/ \sqrt{\sum_{i=1}^{n} D(X_i)}. \tag{9.8}$$

будемо називати **нормованою сумою** або **центрованою випадковою величиною**.

Теорема Ляпунова. Якщо $X_1, X_2, ..., X_n$ — незалежні випадкові величини, які мають математичні сподівання m_i, дисперсії σ_i^2 і скінченні абсолютні центральні моменти третього порядку $|\mu_3|$, що задовільняють умовам

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} |\mu_3|_i}{\left(\sum_{i=1}^{n} \sigma_i^2 \right)^{3/2}} = 0, \tag{9.9}$$

тоді при необмеженому збільшенні n закон розподілу нормованої суми (9.8) збігається за імовірністю до нормального закону з параметрами $a = 0$, $\sigma = 1$.

Зміст умови (9.9) полягає в тому, що дисперсія кожної випадкової величини $X_i, i=\overline{1,n},$ складає лише малу частину в загальній дисперсії суми $\sum_{i=1}^{n} X_i$.

Наслідок теореми Ляпунова. Якщо виконуються умови теореми Ляпунова, тоді випадкова величина $\sum_{i=1}^{n} X_i$ для великих n з достатнім ступенем точності розподілена за нормальним законом з параметрами

$$a=\sum_{i=1}^{n} m_i, \quad \sigma^2=\sum_{i=1}^{n}\sigma_i^2.$$

В практичних задачах центральну граничну теорему Ляпунова часто використовують для обчислення імовірності того, що сума кількох випадкових величин набере значення, яке належить вказаному інтервалу.

Підставою такого використання є так званий

Частковий наслідок теореми Ляпунова. Якщо X_1, X_2, ..., X_n – незалежні випадкові величини, у яких існують рівні математичні сподівання $M(X_i)=m$, дисперсії $D(X_i)=\sigma^2$ і абсолютні центральні моменти третього порядку $M[|X_i-m|^3]=\mu_3$ $i=\overline{1,n}$, то закон розподілу суми $Y_n=X_1+X_2+...+X_n$ при $n\to\infty$ необмежено наближається до нормального закону з параметрами $a=\sum_{i=1}^{n} m$, $\sigma^2=\sum_{i=1}^{n}\sigma_i^2$.

Приклад 1. Розмір виплати кожному клієнту банку випадковий. Статистичні дані цього банку показали, що середня виплата одному клієнту складає 1000 грн., а середнє квадратичне відхилення цієї ж величини – 360 грн. Вважаючи, що виплати окремим клієнтам є незалежними, знайти, скільки повинно бути готівки в банку, щоб з імовірністю, не меншою 0,95, грошей вистачило б на обслуговування 80 клієнтів. Знайти цю саму величину у додатковому припущенні, що абсолютний центральний момент третього порядку для виплати кожного клієнту є сталим.

○ Нехай випадкова величина X_i — виплата i-му клієнту, де $i=\overline{1,\ 80}$; x — невідома кількість готівки в банку для обслуговування 80 клієнтів. За умовою $M(X_i)=1000$ і $\sigma(X_i)=360$ для $i=\overline{1,\ 80}$,

$$P(X_1+X_2+...+X_{80}\le x)\ge 0{,}95 \qquad (9.10)$$

Оскільки інформація про закон розподілу випадкової величини $X_1+X_2+\ldots+X_{80}$ відсутня, то для знаходження невідомого числа x використаємо нерівність Чебишева для величини $\bar{X}=(X_1+X_2+...+X_{80})/80$:

$$P\left(\ \left|\bar{X}-M(\bar{X})\right|\le\varepsilon\right)\ \ge 1-D(\bar{X})/\varepsilon^2. \qquad (9.11)$$

Враховуючи властивості математичного сподівання і дисперсії, незалежність випадкових величин $X_1, X_2, \ldots, X_{80}$, а також умову задачі, отримаємо:

$$M(\bar{X}) = M\left(\sum_{i=1}^{80} X_i / 80\right) = \left[\sum_{i=1}^{80} M(X_i)\right] \Big/ 80 = \left(\sum_{i=1}^{80} 1000\right) \Big/ 80 = 1000,$$

$$D(\bar{X}) = D\left(\sum_{i=1}^{80} X_i / 80\right) = \frac{1}{80^2}\sum_{i=1}^{80} D(X_i) = \frac{1}{80^2}\sum_{i=1}^{80} \sigma^2(X_i) = \frac{1}{80^2}\sum_{i=1}^{80} 360^2 = 360^2 \big/ 80 = 1620.$$

Випадкова подія $\left(\sum_{i=1}^{80} X_i \le x\right)$ рівносильна події

$$\left(\frac{X_1 + X_2 + \ldots + X_{80}}{80} - 1000 \le \frac{x}{80} - 1000\right) \text{ або } \left[\left|\bar{X} - M(\bar{X})\right| \le \frac{x}{80} - 1000\right],$$

звідки

$$P(X_1 + X_2 + \ldots + X_{80} \le x) = P\left[\left|\bar{X} - M(\bar{X})\right| \le \frac{x}{80} - 1000\right].$$

Тому якщо ліва частина нерівності (9.11), де $\varepsilon = \frac{x}{80} - 1000$, буде не меншою від 0,95, тоді виконається нерівність (9.10). Отже, для знаходження невідомого числа x отримаємо нерівність

$$1 - \frac{1620}{\left(\frac{x}{80} - 1000\right)^2} \ge 0{,}95\ ,$$

яка рівносильна нерівностям

$$\frac{1620}{\left(\frac{x}{80} - 1000\right)^2} \le 0{,}05 \Leftrightarrow \left(\frac{x}{80} - 1000\right)^2 \ge 32400 \Leftrightarrow \frac{x}{80} - 1000 \ge 180 \Leftrightarrow x \ge 94400.$$

Врахуємо тепер додаткову умову про сталість для кожного клієнта абсолютного центрального моменту третього порядку величини виплати. Тоді згідно із спрощеним наслідком теореми Ляпунова випадкова величина $Y_{80} = \sum_{i=1}^{80} X_i$ наближено розподілена за нормальним законом із параметрами

$$a = \sum_{i=1}^{80} M(X_i) = 80000,\ \sigma = \sqrt{\sum_{i=1}^{80} \sigma^2(X_i)} = \sqrt{80 \cdot 360^2} = 3219{,}94.$$

З одного боку

$$P(Y_{80} \le x) = P(-\infty < Y_{80} \le x) = \Phi\left(\frac{x - a}{\sigma}\right) - \Phi\left(\frac{-\infty - a}{\sigma}\right) =$$

$$= \Phi\left(\frac{x - 80000}{3219{,}94}\right) + 0{,}5,$$

а за умовою $P(Y_{80} \le x) \ge 0{,}95$. Отже, x можна знайти із нерівності

$$\Phi\left(\frac{x-80000}{3219{,}94}\right) \ge 0{,}45.$$

Використавши табл. додатків і монотонне зростання функції Лапласа, отримаємо

$$\frac{x-80000}{3219{,}94} \ge 1{,}645, \qquad \text{звідки } x \ge 85296{,}801. \bullet$$

Приклад 2. Кожна із 200 незалежних випадкових величин $X_1, X_2, \ldots, X_{200}$ рівномірно розподілена на проміжку [0; 0,8]. Знайти:

1) наближений закон розподілу для випадкової величини $Y = \sum_{i=1}^{200} X_i$;

2) P(Y<0,7).

○ Для рівномірно розподіленої на проміжку [a;b] величини X

$$M(X) = \frac{b+a}{2},\ D(X) = \frac{(b-a)^2}{12}.$$

Тому $M(X_i)$=0,4, $D(X_i)$=4/75, $i = \overline{1,200}$.

Очевидно, що і абсолютні центральні моменти третього порядку для кожної із величин $X_1, X_2, \ldots, X_{200}$ також рівні. За спрощеним наслідком теореми Ляпунова випадкову величину Y можна вважати розподіленою за нормальним законом із параметрами $a = \sum_{i=1}^{200} M(X_i) = 200 \cdot 0{,}4 = 80$,

$$\sigma = \sqrt{\sum_{i=1}^{200} D(X_i)} = \sqrt{32/3} = 4\sqrt{2}.$$

У відповідності із рівністю (6.1) густина розподілу величини Y має такий вид $f(x) = \dfrac{\sqrt{3}}{8\sqrt{\pi}} e^{-\frac{3(x-80)^2}{64}}$.

2) За формулою (6.3)

$$P(Y > 0{,}7) = P(0{,}7 < Y < \infty) = \Phi\left(\frac{\infty - a}{\sigma}\right) - \Phi\left(\frac{0{,}7-a}{\sigma}\right) = 0{,}5 - \Phi\left(\frac{0{,}7-80}{3{,}27}\right) = 0{,}5 + \Phi(24{,}25) \approx 1.$$

Рекомендована література: [1], С. 162–169, [2], С. 101–111; [3], С. 99–102, 129–139; [4], С. 134–137.

Лекція 10.

Тема: ***Вступ в математичну статистику. Вибірковий метод.***

1. ***Задачі математичної статистики.***
2. ***Генеральна та вибіркова сукупності.***
3. ***Способи утворення вибіркової сукупності.***
4. ***Статистичний розподіл вибірки.***
5. ***Емпірична функція розподілу та її властивості.***
6. ***Графічне зображення статистичних розподілів (полігон та гістограма).***
7. ***Числові характеристики вибірки.***

1.

Перша задача математичної статистики — вказати метод відбору і групування статистичних даних, а також знаходження числа необхідних випробувань (статистичних даних).

Друга задача математичної статистики полягає в розробці методів аналізу статистичних даних в залежності від цілей дослідження. Одна з них — оцінка невідомих:

— імовірності випадкової події;

— функції розподілу ймовірностей (густини розподілу);

— параметрів розподілу, вид якого відомий;

— залежності випадкової величини від однієї або кількох випадкових величин.

Друга мета — це перевірка статистичних гіпотез про вид невідомого розподілу або про величину параметрів розподілу, вид якого відомий.

2.

Генеральною називається вся сукупність однотипних об’єктів, яка підлягає вивченню. Множину цих об’єктів позначатимемо через Ω.

Об’єкти множини Ω можуть характеризуватися однією або кількома ознаками. Ці ознаки можуть бути кількісними та якісними.

Іноді проводять **суцільне** обстеження, тобто обстежують **кожний** елемент множини Ω відносно ознаки, яка досліджується. Проте на практиці суцільне обстеження застосовують порівняно рідко. Це зумовлено тим, що при перевірці об’єкт частково або повністю знищується, або дослідження об’єктів вимагає великих матеріальних витрат. Деколи провести суцільне обстеження фізично неможливо. В таких випадках природно відібрати із генеральної сукупності обмежене число об’єктів, дослідити їх на кількісну чи якісну ознаку, а потім робити висновки про всю генеральну сукупність.

Вибірковою сукупністю або просто **вибіркою** називається сукупність **випадково** відібраних об’єктів із генеральної сукупності. Вибірка утворює підмножину V множини Ω ($V \subset \Omega$).

Обсягом сукупності (генеральної або вибіркової) називається число об'єктів цієї сукупності. Надалі обсяг генеральної сукупності позначатимемо літерою N, а вибіркової — n.

Для правомірності висновків про досліджувану ознаку об'єктів генеральної сукупності на підставі опрацювання вибірки необхідно, щоб об'єкти вибірки правильно представляли генеральну сукупність, тобто вибірка повинна володіти властивістю репрезентативності (представницькості). Випадковість відбору об'єктів у вибіркову сукупність і використання закону великих чисел дозволяють вирішити питання про репрезентативність вибірки.

Точність результатів вибіркового спостереження, в кінцевому підсумку, буде залежати від способу відбору об'єктів, ступеня коливання досліджуваної ознаки в генеральній сукупності та від обсягу вибірки.

3.

На практиці використовуються різні способи утворення вибірки, які принципово розподіляються на два види:

1) відбір, що не вимагає розчленування генеральної сукупності на частини (**простий** (**власне випадковий**) **відбір**);

2) відбір, при якому генеральна сукупність розбивається на частини (**типовий** відбір, **механічний** відбір, **серійний** відбір, **комбінований** відбір).

Простим випадковим (**власне випадковим**) називається такий відбір, при якому об'єкти відбираються по одному випадковим чином із усієї генеральної сукупності. Проста випадкова вибірка може бути **повторною** або **безповторною**. **Повторною** називається вибірка, при утворенні якої відібраний об'єкт (перед відбором наступного) повертається в генеральну сукупність. **Безповторною** називається вибірка, в процесі утворення якої відібраний об'єкт в генеральну сукупність не повертається.

Нехай досліджується кількісна ознака X об'єктів генеральної сукупності Ω. Для скорочення припустимо, що вона є одновимірною випадковою величиною. Після опрацювання n об'єктів вибіркової сукупності отримуються n чисел $x_1, x_2, \ldots, x_n$, які називаються **варіантами** і утворюють **ряд варіант** або **простий статистичний ряд**.

Первинна обробка ряду варіант полягає у групуванні рівних варіант цього ряду. Ряд варіант розташуємо в порядку зростання і у відповідності з цим перенумеруємо їх. В результаті одержимо послідовність чисел $x_1 \le x_2 \le x_3 \le \ldots \le x_n$, яка називається варіаційним рядом. Якщо серед цієї послідовності є однакові варіанти, то ми їх знову перенумеруємо, залишаючи один і той самий номер однаковим варіантам. Нехай у варіаційному ряді варіанта x_1 повторюється n_1 разів, x_2 — n_2 разів, ..., x_k — n_k разів. Числа n_i називаються **частотами** (**абсолютними частотами**), а їх відношення до обсягу вибірки $n_i/n = w_i$ — **відносними частотами**. Із цих означень випливають рівності:

$$\sum_{i=1}^{k} n_i = n, \quad \sum_{i=1}^{k} w_i = 1. \tag{10.1}$$

4.

Статистичним розподілом вибірки називається відповідність між варіантами та частотами або відносними частотами:

$$\begin{array}{c|cccc} x_i & x_1 & x_2 & \dots & x_k \\ \hline n_i & n_1 & n_2 & \dots & n_k \end{array}, \tag{10.2}$$

$$\begin{array}{c|cccc} x_i & x_1 & x_2 & \dots & x_k \\ \hline w_i = {n_i}/{n} & w_1 & w_2 & \dots & w_k \end{array}. \tag{10.3}$$

В більшості випадків статистичний розподіл вибірки у вигляді (10.2) або (10.3) використовується тоді, коли ряд варіант є реалізацією **дискретної** випадкової величини X. Якщо ж X є неперервною випадковою величиною, тоді статистичний розподіл вибірки задається у вигляді відповідності між інтервалами і частотами або відносними частотами тих варіант, які потрапляють у ці інтервали, тобто у вигляді таблиць:

$$\begin{array}{c|cccc} [x_i, x_{i+1}) & [x_1, x_2) & [x_2, x_3) & \dots & [x_k, x_{k+1}] \\ \hline n_i & n_1 & n_2 & \dots & n_k \end{array}, \tag{10.4}$$

$$\begin{array}{c|cccc} [x_i, x_{i+1}) & [x_1, x_2) & [x_2, x_3) & \dots & [x_k, x_{k+1}] \\ \hline w_i = {n_i}/{n} & w_1 & w_2 & \dots & w_k \end{array}. \tag{1.5}$$

Ці таблиці називаються **інтервальним статистичним розподілом вибірки**. При побудові інтервального статистичного розподілу на основі ряду варіант розглядається k інтервалів однакової довжини. Кількість інтервалів можна визначати наближено за формулою Стерджеса: $k = 1+3{,}322 \lg n$. При цьому $k \in [5;6]$ при $20 \le n \le 30$; $k \in [7;8]$ при $60 \le n \le 70$; $k \in [8;9]$ при $70 \le n \le 200$; $k \in [9;15]$ при $n > 200$.

Статистичні розподіли вибірки (10.3) та (10.5) називаються **емпіричними** (**дослідними**) **розподілами випадкової величини** X (кількісної ознаки об’єктів генеральної сукупності).

Одна із задач математичної статистики — оцінка (наближене знаходження) невідомої функції розподілу $F(x)$ імовірностей кількісної ознаки X об’єктів генеральної сукупності. За означенням $F(x) = P(X < x)$. В розпорядженні дослідника є статистичні дані, згруповані в статистичному розподілі частот або відносно частот. Тому, врахувавши властивість стійкості відносної частоти, доцільно імовірність події $(X < x)$ наближати відносною частотою цієї ж події.

5.

Емпіричною функцією розподілу (**функцією розподілу вибірки**) називається функція $F^*(x)$ детермінованого аргумента x, яка дорівнює відносній частоті появи події $(X < x)$ для даної вибірки значень випадкової величини X, тобто

$$F^*(x) = W(X < x) = n_x/n, \tag{10.6}$$

де n_x — сума частот тих варіант, які менші від x, n — обсяг вибірки.

На противагу $F^*(x)$ функцію $F(x)$ називають теоретичною функцією розподілу.

Із означення емпіричної функції випливають такі її властивості, аналогічні властивостям теоретичної функції розподілу:

1) $D(F^*) = R$, $E(F^*) = [0, 1]$;
2) $F^*(x)$ — неспадна функція;
3) якщо x_1 — найменша варіанта, тоді $F^*(x) = 0$ для $x \le x_1$; якщо x_k — найбільша варіанта, тоді $F^*(x) = 1$ для $x > x_k$.

6.

В процесі аналізу статистичних даних суттєву роль відіграє геометрична ілюстрація цих даних. Для наочності будують різні графіки статистичних розподілів, зокрема полігон і гістограму.

Нехай статистичний розподіл вибірки визначається таблицями (10.2) або (10.3).

Полігоном частот (**частотним многокутником**) називається ламана, прямолінійні відрізки якої з'єднують сусідні точки (x_1, n_1), (x_2, n_2), …, (x_k, n_k). Для побудови полігона на осі абсцис відкладають варіанти, а на осі ординат — відповідні їм частоти.

Полігоном відносних частот називається ламана, прямолінійні відрізки якої з'єднують сусідні точки (x_1, w_1), (x_2, w_2), …, (x_k, w_k). При побудові полігона відносних частот на осі абсцис відкладають варіанти x_i, а на осі ординат — відповідні їм відносні частоти w_i.

У випадку неперервної ознаки доцільно будувати гістограму. Нехай розподіли (10.4) та (10.5) такі, що довжина кожного із частинних інтервалів дорівнює одному і тому ж числу h.

Гістограмою частот називається сходинкова фігура, що складається із прямокутників, основами яких є частинні інтервали довжиною h, а висоти дорівнюють відношенню n_i/h (**густина частоти**). Для побудови гістограми частот на осі абсцис відкладаються частинні інтервали, а над ними проводяться прямолінійні відрізки, паралельні осі абсцис на віддалі n_i/h. Площа i-го частинного прямокутника дорівнює $hn_i/h = n_i$, тобто сумі частот тих варіант, що потрапляють в i-ий інтервал. Тому площа гістограми частот дорівнює сумі всіх частот вибірки n.

Порівняння двох гістограм дозволяє зробити висновок про те, що виразність гістограми суттєво залежить від обрання довжини h частинних інтервалів.

Гістограмою відносних частот називається сходинкова фігура, що складається із прямокутників, основами яких є частинні інтервали довжиною h, а висоти дорівнюють відношенню w_i/h (**густини відносної частоти**). Для побудови гістограми відносних частот на осі абсцис відкладаються частинні інтервали, а над ними проводяться відрізки, паралельні осі абсцис на віддалі w_i/h. Площа i-ого частинного прямокутника дорівнює $hw_i/h = w_i$, тобто відносній частоті тих варіант, що потрапили в i-ий інтервал. Отже, площа гістограми відносних частот дорівнює одиниці.

7.

Побудова статистичних розподілів вибірки (10.2) або (10.4) та їх графічне зображення — це тільки перший крок на шляху розв'язування задач математичної статистики. Наступний крок передбачає знаходження числових характеристик, які у компактній формі виражають найбільш суттєві особливості статистичного розподілу вибірки і слугують оцінками (наближеними значеннями) невідомих параметрів розподілу кількісної ознаки генеральної сукупності.

Середня вибіркова (**середня арифметична варіант**) статистичного розподілу (10.2) визначається формулою

$$\bar{x}_в = \left(\sum_{i=1}^{k} x_i n_i\right) \Big/ n. \tag{10.7}$$

Якщо всі *n* варіант різні, тоді (1.7) набирає такого виду:

$$\bar{x}_в = \left(\sum_{i=1}^{n} x_i\right) \Big/ n. \tag{10.7*}$$

Якщо вибірка задається інтервальним статистичним розподілом частот (10.4), тоді при знаходженні $\bar{x}_в$ потрібно перейти до дискретного розподілу (10.2), "нові" варіанти якого є серединами інтервалів, а потім використати формулу (10.7).

Крім вказаної середньої, у статистиці застосовують ще й **структурні середні**, які не залежать від значень варіант, що розташовані на краях розподілу, а пов'язані із рядом частот. До структурних середніх належать медіана та мода.

Медіаною Me^* дискретного статистичного розподілу вибірки (10.2) називається таке число, яке ділить варіаційний ряд, що "породжує" цей розподіл, на дві рівні за кількістю варіант частини. Якщо число варіант непарне, тобто $n = 2m + 1$, тоді $Me^* = x_{m+1}$. Якщо ж обсяг вибірки є парним числом, тобто $n = 2m$, тоді медіана дорівнює середньому арифметичному "середньої" (медіанної) пари варіант:

$$Me^* = (x_m + x_{m+1})/2.$$

Медіаною для інтервального статистичного розподілу називається таке число Me^*, для якого виконується рівність:

$$F^*(Me^*) = 0,5, \tag{10.8}$$

де $F^*(x)$ — емпірична функція цього розподілу.

Формула для обчислення медіани має такий вид:

$$Me^* = x_m + \frac{0,5 - F^*(x_m)}{F^*(x_{m+1}) - F^*(x_m)}(x_{m+1} - x_m), \tag{10.9}$$

де $[x_m, x_{m+1})$ – так званий медіанний частинний інтервал ($1 \le m \le k$) для якого виконуються нерівності $F^*(x_m) < 0,5$, $F^*(x_{m+1}) > 0,5$.

Модою Mo^* дискретного статистичного розподілу (10.2) називається варіанта, якій відповідає найбільша частота.

Мода для інтервального статистичного розподілу обчислюється таким чином. Спочатку визначається модальний інтервал $[x_m, x_{m+1})$, тобто такий інтервал, для якого $n_m / h_m = \max\limits_{1 \le i \le k}\{n_i / h_i\}$, де h_i — довжина частинного інтервалу

$[x_m, x_{m+1})$, n_i — число варіант з цього інтервалу. Значення Mo^* міститься всередині модального інтервалу і обчислюється за інтерполяційною формулою

$$Mo^* = x_m + \frac{n_m - n_{m-1}}{2n_m - n_{m-1} - n_{m+1}} h_m. \qquad (10.10)$$

Розглянемо деякі числові характеристики розсіювання варіант навколо середньої вибіркової.

Дисперсією вибірковою статистичного розподілу (10.2) називається середнє арифметичне квадратів відхилень варіант від середньої вибіркової:

$$D_в = \frac{1}{n}\sum_{i=1}^{k}(x_i - \overline{x}_в)^2 n_i. \qquad (10.11)$$

$D_в$ характеризує середню величину розкиду варіант навколо $\overline{x}_в$ в квадратних одиницях.

На практиці зручніше користуватися так званою **розрахунковою формулою для обчислення дисперсії**:

$$D_в = \overline{x^2} - (\overline{x}_в)^2 = \frac{1}{n}\sum_{i=1}^{k} x_i^2 n_i - (\overline{x}_в)^2. \qquad (10.12)$$

Недоліком $D_в$ є її розмірність. Для виправлення цього недоліку використовується інша числова характеристика: **середнє квадратичне відхилення вибіркове**

$$\sigma_в = \sqrt{D_в}. \qquad (10.13)$$

Коливність окремих значень варіант характеризують показники варіації. Найпростішим із них є показник **розмаху варіації** R, який дорівнює різниці між найбільшою та найменшою варіантами розподілу: $R = x_{max} - x_{min}$. Розмах варіації використовується при статистичному вивченні якості продукції.

Якщо $\overline{x}_в$ відмінна від нуля, тоді для порівняння двох статистичних розподілів з точки зору їх розмірності відносно середньої вибіркової вводиться показник **коефіцієнт варіації**, який дорівнює відношенню середнього квадратичного відхилення до середньої вибіркової і виражений у відсотках:

$$V = \frac{\sigma_в}{\overline{x}_в} \cdot 100\%. \qquad (10.14)$$

Вибірковою часткою називається відношення числа m об’єктів вибірки з ознакою α до обсягу вибірки: $w = m/n$. Ознакою α може бути стандартність виробу, сортність продукції, стать людини тощо. За змістом w є відносною частотою випадкової події, яка полягає в тому, що навмання відібраний об’єкт із генеральної сукупності має ознаку α.

Рекомендована література: [2] С. 185–196, [3] С. 5–48; [4], С. 141–145; [5], С. 171–195.

Лекція 11.

Тема: ***Статистичне оцінювання.***

1. ***Точкові статистичні оцінки параметрів розподілу та їхні властивості.***
2. ***Оцінка середньої генеральної для простої вибірки (повторної та безповторної).***
3. ***Оцінка генеральної частки для простої вибірки.***
4. ***Середні квадратичні помилки простої вибірки.***
5. ***Виправлена дисперсія вибіркова.***
6. ***Інтервальні статистичні оцінки. Довірчі інтервали для оцінок $\bar{x}_г$ та p для немалих і є малих вибірок.***
7. ***Знаходження мінімального обсягу вибірки.***
8. ***Довірчі інтервали для $D_Г, \sigma_Г$ у випадку малої вибірки.***

1.

Нехай досліджується неперервна кількісна ознака X об'єктів генеральної сукупності з метою знаходження невідомого закону розподілу. В розпорядженні дослідника є статистичні дані вибірки. Припустимо, що з тих чи інших міркувань висунута гіпотеза про нормальний закон розподілу ознаки X (дослідження питання про правильність цієї гіпотези або хибність буде проведене в наступному параграфі). Оскільки нормальний розподіл повністю визначається двома параметрами a та σ, то виникає необхідність оцінити їх, тобто знайти наближені значення, використовуючи **тільки** спостережені варіанти $x_1, x_2, ..., x_n$ вибірки обсягом n. Параметр $a = M(X)$, математичне сподівання характеризує середнє арифметичне спостережених можливих значень випадкової величини X. З другого боку, $\bar{x}_в$ — це середнє арифметичне варіант. Тому (поки що інтуїтивно) доцільно наближати a середнім вибірковим:

$$a \approx \bar{x}_в = \left(\sum_{i=1}^{k} x_i n_i\right) \Big/ n. \tag{11.1}$$

Аналогічно

$$\sigma \approx \sigma_в = \sqrt{\sum_{i=1}^{k} (x_i - \bar{x}_в)^2 n_i \Big/ n}. \tag{11.2}$$

Ще раз відмітимо, що праві частини рівностей (11.1), (11.2) є **випадковими величинами**, оскільки об'єкти у вибірку потрапляють **випадковим чином**.

Нехай тепер кількісна ознака X об'єктів генеральної сукупності є дискретною випадковою величиною. Припустимо, що статистичні розподіли генеральної та вибіркової сукупності описуються таблицями:

$$\begin{array}{c|cccc} x_i & x_1 & x_2 & \dots & x_m \\ \hline N_i & N_1 & N_2 & \dots & N_m \end{array}, \quad N = \sum_{i=1}^{m} N_i, \tag{11.3}$$

$$\begin{array}{c|cccc} x_i & x_1 & x_2 & \dots & x_m \\ \hline n_i & n_1 & n_2 & \dots & n_m \end{array}, \qquad n = \sum_{i=1}^{m} n_i, \tag{11.4}$$

де N та n — обсяги генеральної та вибіркової сукупностей відповідно. Розподіл (11.3) є гіпотетичним — він завжди буде для нас невідомим, бо в протилежному випадку відпала б необхідність у дослідженні вибіркової сукупності. Нарешті, зауважимо, що деякі із частот розподілу (11.4) можуть дорівнювати нулю, що відповідає ситуації, коли значення кількісної ознаки об'єктів генеральної сукупності не зустрілися серед варіант вибірки (порівняйте розподіл (11.4) з (1.2)).

За аналогією із числовими характеристиками вибірки наступні формули визначають **числові характеристики генеральної сукупності**:

$$\bar{x}_г = \left(\sum_{i=1}^{m} x_i N_i\right) \Big/ N, \tag{11.5}$$

$$D_г = \left(\sum_{i=1}^{m} (x_i - \bar{x}_г)^2 N_i\right) \Big/ N, \tag{11.6}$$

$$\sigma_г = \sqrt{D_г}. \tag{11.7}$$

Ці **числа** невідомі, і оцінки (наближення) їх дають такі рівності:

$$\bar{x}_г \approx \bar{x}_в, \qquad D_г \approx D_в, \qquad \sigma_г \approx \sigma_в. \tag{11.8}$$

Наведені приклади дозволяють зробити деякі висновки. Ліві частини **наближених** рівностей (11.1), (11.2) та (11.8) є невідомими параметрами "відомого" закону розподілу або числовими характеристиками генеральної сукупності; вони є невідомими числами для конкретної генеральної сукупності. Праві частини цих рівностей є функціями випадкових величин, які для **фіксованої** вибірки статистичних даних набирають числові значення, що можна зобразити точками. Це дозволяє назвати їх **точковими статистичними оцінками** відповідних параметрів або числових характеристик генеральної сукупності.

Позначимо узагальнено символом Θ ліві частини наближених рівностей (11.1), (11.2), (11.8) (а також багатьох інших, що можна отримати для інших законів розподілу), а символом Θ^* — праві частини цих рівностей.

Точковою статистичною оцінкою, **вибірковою функцією** або **статистикою** числового параметра Θ називається функція вибіркових значень (варіант) $\Theta^* = \Theta^*(x_1, x_2, ..., x_n)$, яка в певному статистичному сенсі є близькою до справжнього значення цього параметра.

Незміщеною називається точкова статистична оцінка Θ^*, математичне сподівання якої дорівнює оцінюваному параметру Θ при довільному обсязі вибірки, тобто

$$M(\Theta^*) = \Theta. \tag{11.9}$$

Зміщеною називається оцінка, для якої не виконується рівність (11.9).

Нехай для оцінювання параметра Θ можуть бути використані незміщені точкові оцінки $\Theta_1^*, \Theta_2^*, ..., \Theta_k^*$. Оцінка Θ_m^*, $1 \le m \le k$, називається **ефективною**, якщо при заданому обсязі n вибірки для неї виконується рівність

$$D(\Theta_m^*) = \min_{1 \le i \le k} D(\Theta_i^*).$$

Оцінка Θ^* називається **спроможною** оцінкою параметра Θ, якщо при $n \to \infty$ вона збігається по імовірності до Θ, тобто для як завгодно малого $\varepsilon > 0$ має місце граничний перехід

$$\lim_{n\to\infty} P\left(\left|\Theta - \Theta^*\right| < \varepsilon \right) = 1.$$

2.

Теорема 2.1. Для повторної вибірки обсягом *n* середня вибіркова $\bar{x}_в$ є незміщеною і спроможною оцінкою невідомої середньої генеральної $\bar{x}_г$. Якщо *n* досить велике, тоді $\bar{x}_в$ з достатнім ступенем точності розподілена за нормальним законом з параметрами:

$$a = \bar{x}_г, \qquad \sigma = \sqrt{D_г / n}. \tag{11.10}$$

Теорема 2.2. Для безповторної вибірки обсягом *n* середня вибіркова $\bar{x}_в$ є незміщеною оцінкою невідомої середньої генеральної $\bar{x}_г$. Для досить великих *n* $\bar{x}_в$ з достатнім ступенем точності розподілена за нормальним законом з параметрами:

$$a = \bar{x}_г, \qquad \sigma = \sqrt{\frac{D_г}{n} \cdot \frac{N-n}{N-1}}, \tag{11.11}$$

де *N* — обсяг генеральної сукупності.

Зауваження. Обсяг генеральної сукупності *N*, як правило, дуже великий. Тому заміна в знаменнику другої рівності (11.11) $N-1$ на *N* невідчутна для σ. В зв'язку із цим надалі будемо користуватися рівністю

$$\sigma = \sqrt{\frac{D_г}{n} \cdot \left(1 - \frac{n}{N}\right)}. \tag{11.12}$$

3.

Нехай досліджується якісна ознака об'єктів генеральної сукупності, число яких є скінченним. **Генеральною часткою** будемо називати відношення числа *M* об'єктів генеральної сукупності, що володіють ознакою α, до обсягу *N* генеральної сукупності:

$$p = M/N.$$

Теорема 2.3. Для повторної вибірки обсягом *n* середня вибіркова частка *w* є незміщеною і спроможною точковою оцінкою невідомої генеральної частки *p*. Якщо *n* є досить великим, тоді *w* з достатнім ступенем точності розподілена за нормальним законом з параметрами:

$$a = p, \qquad \sigma = \sqrt{pq/n}, \tag{11.13}$$

де $q = 1 - p.$

Теорема 2.4. Для безповторної вибірки обсягом *n* вибіркова частка *w* є незміщеною точковою оцінкою невідомої генеральної частки *p*. Якщо *n* є досить великим, тоді *w* з достатнім ступенем точності розподілена за нормальним законом з параметрами:

$$a = p, \qquad \sigma = \sqrt{\frac{pq}{n} \cdot \frac{N-n}{N-1}}, \tag{11.14}$$

Зауваження. Оскільки обсяг генеральної сукупності *N* в більшості випадків дуже великий, то заміна в знаменнику другої рівності (11.14) $N-1$ на *N* невідчутна для σ. З огляду на це надалі будемо користуватися рівністю

$$\sigma = \sqrt{\frac{pq}{n}\left(1 - \frac{n}{N}\right)}. \qquad (11.15)$$

4.

Випадкові величини $\bar{x}_{в}$ та w є незміщеними точковими статистичними оцінками невідомих $\bar{x}_{г}$ та p відповідно. Їх реалізація або можливі значення, знайдені на основі даних простої вибірки (повторної або безповторної), не співпадають із оцінюваними параметрами. І кожне таке неспівпадання або відхилення природно називати **помилкою репрезентативності оцінки**, зумовленою тим, що досліджується не вся генеральна сукупність, а лише її частина (вибіркова сукупність). Для статистики дуже важливою є інформація про середню величину таких помилок.

Середньою квадратичною помилкою (**СКП**) при оцінюванні невідомих середньої генеральної $\bar{x}_{г}$ та генеральної частки p називається середнє квадратичне відхилення середньої вибіркової $\bar{x}_{в}$ та вибіркової частки w відповідно.

Із врахуванням результатів теорем 2.1–2.4 можна вказати наступні формули для визначення середніх квадратичних помилок:

СКП середньої вибіркової повторної вибірки (див. (11.10))

$$\bar{\sigma}_{\bar{x}} = \sqrt{D_{г}/n}; \qquad (11.16)$$

СКП середньої вибіркової безповторної вибірки (див. (11.12))

$$\bar{\sigma}'_{\bar{x}} = \sqrt{\frac{D_{г}}{n}\left(1 - \frac{n}{N}\right)}; \qquad (11.17)$$

СКП вибіркової частки повторної вибірки (див. (11.13))

$$\bar{\sigma}_{w} = \sqrt{pq/n}; \qquad (11.18)$$

СКП вибіркової частки безповторної вибірки (див. (11.15))

$$\bar{\sigma}'_{w} = \sqrt{\frac{pq}{n}\left(1 - \frac{n}{N}\right)}. \qquad (11.19)$$

На практиці користуватися формулами (11.16)–(11.19) неможливо, оскільки для цього необхідно знати або дисперсію генеральну, або генеральну частку (проаналізуйте, наскільки обтяжливою є ця умова, згадавши початкову умову задачі оцінювання). Цю принципову трудність можна усунути за рахунок заміни у формулах СКП дисперсії генеральної $D_{г}$ на дисперсію вибіркову $D_{в}$, а генеральної частки p — на вибіркову частку w. Проте при такій заміні потрібно попередньо впевнитися, чи не з'явиться ще додаткова систематична помилка за рахунок зміщеності оцінок $D_{в}$ в обох задачах оцінювання. Якби вони виявилися зміщеними, то при заміні слід було б ввести "виправлені" оцінки невідомих дисперсій генеральних. На шляху реалізації вказаного вище підходу корисними є такі твердження.

Теорема 2.5. Математичне сподівання дисперсії вибіркової в задачі про оцінювання невідомої середньої генеральної для повторної вибірки визначається рівністю

$$M(D_{в}) = \frac{n-1}{n} D_{г}, \qquad \mathbf{(11.20)}$$

а для безповторної —

$$M(D_в) = \frac{n-1}{n} \cdot \frac{N}{N-1} D_г, \qquad (11.21)$$

де n та N відповідно обсяги вибіркової та генеральної сукупностей.

Теорема 2.6. Математичне сподівання дисперсії вибіркової в задачі про оцінювання невідомої генеральної частки для повторної вибірки визначається рівністю

$$M(D_в) = \frac{n-1}{n} pq, \qquad (11.22)$$

а для безповторної —

$$M(D_в) = \frac{n-1}{n} \cdot \frac{N}{N-1} pq. \qquad (11.23)$$

5.

Зміст теорем 2.5 та 2.6 полягає в тому, що дисперсія вибіркова є зміщеною оцінкою дисперсії генеральної в задачах оцінювання невідомих $\bar{x}_г$ та p у випадку простої вибірки (повторної та безповторної). При цьому формули (11.20)–(11.23) вказують на заниження значень дисперсії генеральної за рахунок наявності в кожній із них множника $(n-1)/n$.

Проте, цю зміщеність легко “виправити”: достатньо дисперсію вибіркову помножити на дріб $n/(n-1)$.

Виправленою дисперсією називається числова характеристика S^2, яка визначається рівністю

$$S^2 = \frac{n}{n-1} D_в.$$

Для **немалих** вибірок (обсягом $n \ge 30$), замінивши у формулах (11.16)–(11.19) дисперсію генеральну $D_г$ вибірковою $D_в$, а генеральну частку p — вибірковою w, отримаємо використовувані на практиці формули для знаходження СКП:

СКП середньої вибіркової повторної вибірки

$$\bar{\sigma}_{\bar{x}} = \sqrt{D_в/n} = \sigma_в / \sqrt{n}; \qquad (11.24)$$

СКП середньої вибіркової безповторної вибірки

$$\bar{\sigma}'_{\bar{x}} = \sqrt{\frac{D_в}{n}\left(1 - \frac{n}{N}\right)}; \qquad (11.25)$$

СКП вибіркової частки повторної вибірки

$$\bar{\sigma}_w = \sqrt{\frac{w(1-w)}{n}}; \qquad (11.26)$$

СКП вибіркової частки безповторної вибірки

$$\bar{\sigma}'_w = \sqrt{\frac{w(1-w)}{n}\left(1 - \frac{n}{N}\right)}. \qquad (11.27)$$

6.

Точкова статистична оцінка Θ^* не співпадає (за виключанням рідкісних випадків) із справжнім значенням невідомого параметра Θ. Тому завжди виникає похибка при заміні невідомого параметра його оцінкою, тобто

$\left|\Theta-\Theta^{*}\right|>0.$ Величина похибки при цьому невідома, хоча потрібно знати, до яких помилок може призвести вказана вище заміна.

Інтервальною називається статистична оцінка, яка визначається **двома числами** — кінцями інтервалу. Перевагою інтервальних оцінок є те, що вони дозволяють встановити точність і надійність оцінок.

Число $\delta>0,$ яке фігурує у нерівності

$$\left|\Theta-\Theta^{*}\right|<\delta, \tag{11.28}$$

природно назвати **точністю** оцінки Θ^{*}. Проте говорити про виконання нерівності (11.28) можна тільки в імовірносному сенсі, бо Θ^{*} — випадкова величина. Тобто, для достатньо малого δ ця нерівність є випадковою подією.

Надійністю (довірчою імовірністю) оцінки Θ по Θ^{*} називається імовірність γ виконання нерівності (11.28):

$$P\left(\left|\Theta-\Theta^{*}\right|<\delta\right)=\gamma. \tag{11.29}$$

Довірчим називається інтервал $(\Theta^{*}-\delta;\ \Theta^{*}+\delta),$ який із заданою надійністю γ покриває невідомий параметр Θ.

Довірчі інтервали для оцінок $\bar{x}_{\text{г}}$ та p для немалих вибірок

Найбільше відхилення середньої вибіркової (або вибіркової частки) від середньої генеральної (або генеральної частки), яке можливе для заданої довірчої імовірності γ, називається **граничною помилкою Δ**.

Гранична помилка знаходиться за формулою

$$\Delta=t\bar{\sigma}, \tag{11.30}$$

де t — корінь рівняння $2\Phi(t)=\gamma.$

Підставивши в рівність (11.30) вирази СКП (11.24)–(11.27), отримаємо **придатні для практики формули граничної помилки**:

середньої вибіркової повторної вибірки

$$\Delta=t\sqrt{D_{\text{в}}/n}; \tag{11.31}$$

середньої вибіркової безповторної вибірки

$$\Delta=t\sqrt{\frac{D_{\text{в}}}{n}\left(1-\frac{n}{N}\right)}; \tag{11.32}$$

частки вибіркової повторної вибірки

$$\Delta=t\sqrt{w(1-w)/n}; \tag{11.33}$$

частки вибіркової безповторної вибірки

$$\Delta=t\sqrt{\frac{w(1-w)}{n}\left(1-\frac{n}{N}\right)}. \tag{11.34}$$

В результаті $(\bar{x}_{\text{в}}-\Delta;\ \bar{x}_{\text{в}}+\Delta)$ є довірчим інтервалом, який з надійністю γ покриває невідому середню генеральну $\bar{x}_{\text{г}}$. Аналогічно $(w-\Delta;\ w+\Delta)$ — довірчий інтервал, який з тією ж надійністю покриває невідому генеральну частку p.

Зауваження. В деяких випадках може бути відомою числова характеристика $\sigma_{\text{г}}$ (наприклад, як інформація технологічного процесу). Тоді $D_{\text{в}}$ у формулах (11.31), (11.32) слід замінити на $\sigma_{\text{г}}^{2}$.

7.

Перед утворенням вибіркової сукупності необхідно з’ясувати, яким повинен бути її обсяг.

Наступні формули визначають мінімальні обсяги вибірки при оцінюванні невідомих:

а) середньої генеральної

$$n = \frac{t^2 D_г}{\Delta^2} \tag{11.35}$$

для повторної вибірки,

$$n' = \frac{nN}{n + N} \tag{11.36}$$

для безповторної вибірки (n визначається (11.35));

б) генеральної частки

$$n = \frac{t^2 pq}{\Delta^2} \tag{11.37}$$

для повторної вибірки,

$$n' = \frac{nN}{n + N} \tag{11.38}$$

для безповторної вибірки (n визначається (11.37));
де N — обсяг генеральної сукупності.

8.

Нехай про досліджувану кількісну ознаку X відомо тільки те, що вона розподілена за нормальним законом. Ставиться задача: побудувати довірчий інтервал для оцінки невідомого параметра $a = M(X)$ (або $\bar{x}_г$ у випадку скінченності обсягу генеральної сукупності) за даними повторної вибірки малого обсягу n, заданою довірчою імовірністю γ і якщо невідомий параметр $\sigma = \sigma_г$.

У [8] доведено, що довірчий інтервал для невідомого параметра a ($\bar{x}_г$) при невідомому $\sigma\,(\sigma_г)$ з надійністю γ має такий вид:

$$\bar{x}_в - t(\gamma, n-1)\frac{S}{\sqrt{n}} < a < \bar{x}_в + t(\gamma, n-1)\frac{S}{\sqrt{n}}, \tag{11.39}$$

де

$$S^2 = \sum_{i=1}^{n} (x_i - \bar{x}_в)^2 \Big/ (n-1), \tag{11.40}$$

параметр $t = t(\gamma, n\text{-}1)$ – корінь рівняння

$$P\left(\left| \frac{\bar{x}_в - a}{S/\sqrt{n}} \right| < t \right) = P\left(|T| < t \right) = 2\int_0^t g_k(t)dt = \gamma, \tag{11.41}$$

який можна знайти за табл.4 додатків в залежності від заданої довірчої імовірності γ і числа ступенів вільності $k = n - 1$; $g_k(t)$ – густина розподілу Ст’юдента.

Довірчі інтервали для $D_г$ та $\sigma_г$ у випадку малої вибірки.

При знаходженні мінімального обсягу вибірки необхідною є інформація про дисперсію генеральну (середнє квадратичне відхилення генеральне). Часто в розпорядженні дослідника є тільки вибірка малого обсягу.

Можна довести (див. [8]), що довірчий інтервал для оцінки невідомої дисперсії генеральної $D_г=\sigma^2$ з надійністю γ має такий вид:

$$\frac{(n-1)S^2}{\chi_2^2} < \sigma^2 < \frac{(n-1)S^2}{\chi_1^2}, \tag{11.42}$$

де S визначається рівністю (11.40), значення χ_1^2 і χ_2^2 знаходяться за табл.6 додатків з використанням рівнянь

$$P\left(\chi^2(k) > \chi_1^2(p;k)\right) = p\,, \ \ p = \frac{1+\gamma}{2}; \tag{11.43}$$

$$P\left(\chi^2(k) > \chi_2^2(p;k)\right) = p\,, \ \ p = \frac{1-\gamma}{2}; \tag{11.44}$$

$k = n - 1$ – число ступенів вільності закону розподілу χ^2.

Із подвійної нерівності (11.42) отримаємо рівносильну подвійну нерівність

$$\frac{S\sqrt{n-1}}{\chi_2} < \sigma < \frac{S\sqrt{n-1}}{\chi_1}, \tag{11.45}$$

яка визначає довірчий інтервал для оцінки невідомої $\sigma(\sigma_г)$.

Зауваження. В табл. 6 додатків наведені значення $\chi^2(p;k)$, що задовільняють рівняння $P\left(\chi^2(k) > \chi^2(p;k)\right) = p$ тільки для $k = n - 1 \le 30$, а також для $k = 40, 50, 100$. Це зумовлено тим, що при зростанні k закон розподілу χ^2 наближається до нормального.

Рекомендована література: [2] С. 197–252, [3] С. 48–95; [4], С. 141–145; [5], С. 202–218.

Лекція 12.

Тема: ***Перевірка статистичних гіпотез.***

1. ***Статистичні гіпотези та їхні види.***
2. ***Статистичний критерій перевірки основної гіпотези. Потужність критерію.***
3. ***Параметричні статистичні гіпотези.***
4. ***Критерій узгодженості Пірсона та Колмогорова (на прикладі перевірки гіпотези про нормальний закон розподілу).***
5. ***Критерій однорідності двох виборок (критерій Смирнова).***

1.

Статистичною називається гіпотеза про вид невідомого розподілу випадкової величини (кількісної ознаки об'єктів генеральної сукупності) або про параметри відомого розподілу.

Поряд із висунутою гіпотезою розглядають і гіпотезу, яка суперечить їй. Тому надалі будемо припускати, що у нас є дві гіпотези: H_0 та H_1, які не перетинаються. Гіпотезу H_0 будемо називати **основною** або **нульовою**, а гіпотезу H_1 — **конкуруючою** або **альтернативною**.

Гіпотези розрізняються за числом припущень. **Простою** називається гіпотеза, яка містить тільки одне припущення. **Складною** називається гіпотеза, яка складається із скінченного або нескінченного числа простих гіпотез.

Висунута статистична гіпотеза може бути правильною або хибною. Для перевірки її правильності використовують статистичні дані і статистичні методи, тому перевірку називають **статистичною**.

Для перевірки основної гіпотези потрібно мати критерій правильності цієї гіпотези. Як вже зазначалося, в розпорядженні дослідника є тільки вибірка $X_1, X_2, ..., X_n$, де X_i — значення кількісної ознаки i-ого об'єкта вибірки $(i=\overline{1,n})$, або значення вибіркової частки ω у випадку вивчення якісної ознаки об'єктів генеральної сукупності.

2.

Статистичним критерієм (або просто **критерієм**, чи **статистикою**) називається випадкова великина K, яка використовується для перевірки основної гіпотези і закон розподілу якої (точний або наближений) відомий. Для кожного конкретного випадку величина K спеціально підбирається і може позначатися різними літерами: U або Z, якщо вона нормально розподілена, F або v^2 — по закону Фішера-Снедокора, T — по закону Ст'юдента, χ^2 — по закону "хі-квадрат", K — по закону Колмогорова і т. д.

Можливі значення випадкової величини (критерію) K розбиваються на дві непорожні множини Q та $\overline{Q}$ ($Q \cap \overline{Q} = \varnothing$) такі, що Q складається із значень критерію, при яких H_0 приймається, а $\overline{Q}$ — із тих значень критерію, при яких H_0 відхиляється (а отже, приймається H_1). Множину $\overline{Q}$ називають

критичною областю, а множину Q — **областю прийняття гіпотези**, або **областю допустимих значень**.

Для конкретної вибірки обчислюється значення критерію як функції варіант. Отримане значення позначається K_{cn} і називається **спостереженим значеням критерію**.

Сформулюємо **основний принцип статистичної перевірки статистичної гіпотези**: якщо спостережене значення критерію належить критичній області $\left(K_{cn} \in \overline{Q}\right)$, тоді гіпотезу H_0 відхиляють; якщо спостережене значення критерію належить області допустимих значень $(K_{cn} \in Q)$, тоді гіпотезу H_0 приймають.

Зауваження. Припустимо, що для конкретних гіпотез H_0 і H_1 множини Q і $\overline{Q}$ **визначені**. В ряді посібників по математичній статистиці під статистичним критерієм розуміється правило, яке реалізує основний принцип статистичної перевірки статистичної гіпотези.

Оскільки висновок про правильність гіпотези робиться за результатами скінченної вибірки, а вона може бути "невдалою", то завжди існує ризик прийняти хибне рішення. При цьому можуть бути допущені помилки двох родів.

Якщо буде відхилена гіпотеза H_0 (і прийнята H_1), в той час як насправді правильною є H_0, тоді це є **помилка першого роду**; її імовірність позначають α:

$$\alpha = P\left(K_{cn} \in \overline{Q}/H_0\right) = P(H_1/H_0)\ ,$$

де $P(H_1/H_0)$ — імовірність того, що буде прийнята гіпотеза H_1, якщо насправді для генеральної сукупності правильною є гіпотеза H_0. Число α називають **рівнем значущості**.

Якщо буде прийнята гіпотеза H_0, в той час як насправді правильною є H_1, тоді буде допущена **помилка другого роду**, її імовірність позначають β:

$$\beta = P(K_{cn} \in Q/H_1) = P(H_0/H_1).$$

В цьому параграфі розглянемо один із методів перевірки статистичних гіпотез, який передбачає виконання таких кроків:

1) формулювання основної і конкуруючої гіпотез;

2) вибір відповідного рівня значущості α;

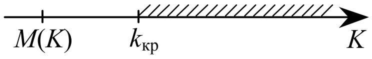
3) вибір статистичного критерія K перевірки гіпотези;

4) знаходження критичної області $\overline{Q}$ і області Q прийняття гіпотези;

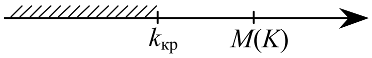
5) формулювання правила перевірки гіпотези: гіпотеза H_0 приймається при заданому рівні значущості α, якщо $K_{cn} \in Q$; гіпотеза H_0 відкидається, якщо $K_{cn} \in \overline{Q}$.

Розглянемо деякі особливості структури критичної області $\overline{Q}$. Якщо $\overline{Q}$ розташована зліва **і** справа від математичного сподівання випадкової величини K, то критична область називається **двосторонньою**, а критерій K — **двостороннім критерієм значущості**. Якщо ж $\overline{Q}$ розташована зліва **або** справа від математичного сподівання випадкової величини K, то критична область називається **односторонньою**, а критерій — **одностороннім**.

Реалізація основного принципу перевірки конкретної основної гіпотези H_0 передбачає знаходження множин Q та $\overline{Q}$. Оскільки ці множини числової осі не перетинаються, то існують точки, які їх розділяють. Такі точки називають **критичними точками** і позначаються $k_{\text{кр}}$. **Правосторонньою** називається критична область, яка визначається нерівністю $K > k_{\text{кр}}$, де $k_{\text{кр}} > M(K)$:



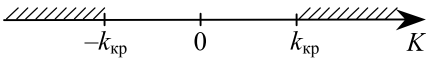
Лівосторонньою називається критична область, яка визначається нерівністю $K < k_{\text{кр}}$, де $k_{\text{кр}} < M(K)$:



Двосторонньою називається критична область, яка визначається сукупністю нерівностей

$$\left[\begin{array}{l} K < k_{\text{кр}}^{(1)}, \\ K > k_{\text{кр}}^{(2)}, \end{array}\right.$$

де $k_{\text{кр}}^{(1)} < M(K) < k_{\text{кр}}^{(2)}$. Зокрема, якщо критичні точки $k_{\text{кр}}^{(1)}$ та $k_{\text{кр}}^{(2)}$ відрізняються тільки знаком, тоді двостороння критична область визначається нерівністю $|K| > k_{\text{кр}}$:



Інтуїтивно зрозуміло, що множини Q та $\overline{Q}$ залежать від вибраного рівня значущості α, а тому слід очікувати залежності критичної точки $k_{\text{кр}}$ від α. Переконаємося в цьому, розглянувши підхід до використання основного принципу статистичної перевірки статистичної гіпотези у випадку правосторонньої критичної області. Аналіз решти випадків проводиться аналогічно. Для вибраного рівня значущості α $\overline{Q} = \{K > k_{\text{кр}}\}$. Критична точка $k_{\text{кр}}$ при умові правильності основної гіпотези H_0 задовольняє рівності

$$P\left(K > k_{\text{кр}} \right) = \alpha, \tag{12.1}$$

яка трактується таким чином: при правильності H_0 малоімовірно (з врахуванням малості α), що спостережене значення критерію K виявиться більшим від $k_{\text{кр}}$.

Для кожного практично важливого критерію складено таблиці, за допомогою яких знаходиться $k_{\text{кр}}$, що задовольняє рівність (12.1). Наступний крок — обчислення спостереженого значення $K_{сп}$ критерію за даною вибіркою. А далі в силу вступає **правило**: якщо $K_{сп} > k_{\text{кр}}$, то гіпотезу H_0 відхиляють (вважається, що вона є хибною), якщо ж $K_{сп} < k_{\text{кр}}$, тоді H_0 приймають (вважається, що H_0 — правильна).

Якщо основна гіпотеза H_0 містить твердження про параметри розподілу, вид якого відомий, тоді статистичний критерій перевірки цієї гіпотези називається **параметричним**. Якщо ж у гіпотезі H_0 мова ведеться про невідомий розподіл, тоді відповідний критерій називається **критерієм узгодженості**.

3.

Розгляд статистичних гіпотез розпочнемо із параметричних гіпотез. Це зумовлено відносною простотою досліджень таких гіпотез в порівнянні із критеріями узгодженості.

Нижче будуть наведені тільки підсумки аналізу, який в повному обсязі можна знайти, зокрема, у посібнику [8].

Порівняння середньої вибіркової із гіпотетичною

середньою генеральною нормальної сукупності

Нехай кількісна ознака X генеральної сукупності розподілена нормально з параметрами a та σ. При цьому середня генеральна a хоча і невідома, але є підстави припускати, що вона дорівнює гіпотетичному значенню a_0. Для перевірки гіпотези $a = a_0$ потрібно знайти середню вибіркову $\bar{x}_в$ і з'ясувати, істотно чи ні $\bar{x}_в$ відхиляється від a_0.

Розв'язування окресленої задачі суттєво залежить від того, чи відомий параметр σ.

А. Дисперсія генеральної сукупності відома

Нехай із генеральної сукупності, кількісна ознака якої нормально розподілена, організована вибірка обсягом n і знайдена середня вибіркова $\bar{x}_в$, при цьому $D_г = \sigma^2$ — відома. Потрібно при заданому рівні значущості α перевірити основну гіпотезу про рівність середньої генеральної a гіпотетичному значенню a_0, тобто $H_0 : a = a_0$.

Оскільки середня вибіркова $\bar{X}$ (вибірка поки не фіксована) є незміщеною оцінкою середньої генеральної, тобто $M(\bar{X}) = a$, то гіпотезу H_0 можна записати ще й таким чином: $M(\bar{X}) = a_0$.

Отже, потрібно перевірити, що математичне сподівання середньої вибіркової дорівнює гіпотетичній середній генеральній.

В якості критерія перевірки гіпотези H_0 використаємо випадкову величину:

$$U = (\bar{X} - a_0)/\sigma(\bar{X}) = (\bar{X} - a_0)\sqrt{n}/\sigma.$$

Ця величина розподілена нормально, причому якщо гіпотеза H_0 правильна, то $M(U) = 0, \ \sigma(U) = 1$.

Позначимо через $U_{спост}$ спостережене значення критерія для фіксованої вибірки:

$$U_{спост} = (\bar{x}_в - a_0)\sqrt{n}/\sigma. \qquad (12.2)$$

Критичну область побудуємо в залежності від виду конкуруючої гіпотези.

Правило 1. Гіпотеза $H_0 : a = a_0$ при конкуруючій гіпотезі $H_1 : a \neq a_0$ для рівня значущості α приймається, якщо $|U_{спост}| < u_{кр}$, де $u_{кр}$ — корінь рівняння

$$\Phi(u_{кр}) = (1 - \alpha)/2\ . \qquad (12.3)$$

Якщо $|U_{спост}| > u_{кр}$, то гіпотеза H_0 відкидається.

Правило 2. При конкуруючій гіпотезі $H_1 : a > a_0$ критичну точку правосторонньої критичної області для рівня значущості α знаходять за рівністю (табл. 3 додатків):

$$\Phi(u_{\text{кр}}) = (1-2\alpha)/2. \tag{12.4}$$

Якщо $U_{\text{спост}} < u_{\text{кр}}$, то основна гіпотеза H_0 приймається. Якщо ж $U_{\text{спост}} > u_{\text{кр}}$ — H_0 відкидається.

Правило 3. При конкуруючій гіпотезі $H_1 : a < a_0$ знаходять корінь $u_{\text{кр}}$ рівняння (12.4) (табл. 3 додатків).

Якщо $U_{\text{спост}} > -u_{\text{кр}}$, то нема підстав відкидати гіпотезу H_0. Якщо ж $U_{\text{спост}} < -u_{\text{кр}}$, то H_0 відкидається.

Б. Дисперсія генеральної сукупності невідома

В цьому випадку в якості критерія перевірки основної гіпотези береться випадкова величина

$$T = (\overline{X} - a_0)\sqrt{n}/S, \tag{12.5}$$

де $\overline{X}$ та S — середня та “виправлене” середнє квадратичне відхилення для нефіксованої вибірки обсягу n.

Величина T розподілена за законом Ст’юдента із $k = n-1$ ступенями вільності.

Симетричність розподілу Ст’юдента дозволяє здійснювати побудову критичних областей (в залежності від виду конкуруючої гіпотези) аналогічно тому, як це робилося у випадку відомого параметра σ.

Правило 1*. Для того, щоб при заданому рівні значущості α перевірити основну гіпотезу $H_0 : a = a_0$ про рівність невідомої середньої генеральної a (нормально розподіленої генеральної сукупності з невідомою дисперсією) гіпотетичному значенню a_0 при конкуруючій гіпотезі $H_1 : a \neq a_0$, потрібно обчислити спостережене значення критерія

$$T_{\text{спост}} = (\overline{x}_в - a_0)\sqrt{n}/S$$

і за таблицею критичних точок розподілу Ст’юдента (табл. 5 додатків) по заданому рівню значущості α, розміщеному у верхньому рядку таблиці і числу ступенів вільності $k = n-1$ знайти критичну точку $t_{\text{двост.кр}}(\alpha, k)$.

Якщо $|T_{\text{спост}}| < t_{\text{двост.кр}}$, то нема підстав відкидати основну гіпотезу H_0.

Якщо ж $|T_{\text{спост}}| > t_{\text{двост.кр}}$, то гіпотеза H_0 відкидається на користь альтернативної гіпотези H_1.

Зауваження. Критична точка $t_{\text{двост.кр}} = t_\alpha$ є коренем рівняння

$$\int_0^{t_\alpha} g_k(t)dt = (1-\alpha)/2,$$

де $g_k(t)$ — густина розподілу Ст’юдента, $k = n-1$ — число ступенів вільності. Це рівняння є аналогом рівняння (12.3).

Правило 2*. При конкуруючій гіпотезі $H_1 : a > a_0$ по рівню значущості α, розміщеному в нижньому рядку табл. 5 додатків, і числу ступенів вільності $k = n-1$ знаходимо критичну точку $t_{\text{правост.кр}}(\alpha, k)$ правосторонньої критичної області.

Якщо $T_{\text{спост}} < t_{\text{правост.кр}}$, то нема підстав відкидати основну гіпотезу H_0.

Зауваження. Вибір критичної точки в табл. 5 додатків по нижньому рядку значень рівня значущості зумовлений аналогом рівності (12.4) для випадку правосторонньої критичної області, тобто $t_{\text{правост.кр}} = t_{2\alpha}$ є коренем рівняння

$$\int_0^{t_{2\alpha}} g_k(t) = (1-2\alpha)/2.$$

Правило 3*. При конкуруючій гіпотезі $H_1 : a < a_0$ спочатку знаходять "допоміжну" критичну точку $t_{\text{правост.кр}}(\alpha; k)$, а потім покладають межу лівосторонньої критичної області:

$$t_{\text{правост.кр}} = t_{\text{лівост.кр}}.$$

Якщо $T_{\text{спост}} > -t_{\text{правост.кр}}$, то гіпотеза H_0 приймається.

Якщо $T_{\text{спост}} < -t_{\text{правост.кр}}$, то гіпотезу H_0 відкидають.

Порівняння виправленої дисперсії вибіркової з гіпотетичною дисперсією

генеральною нормальної сукупності

Нехай кількісна ознака генеральної сукупності розподілена за нормальним законом, причому дисперсія генеральна хоча і невідома, проте є підстави припускати, що вона дорівнює гіпотетичному значенню σ_0^2.

На підставі виправленої дисперсії вибіркової S^2, знайденої для вибірки обсягом *n*, при заданому рівні значущості α потрібно перевірити основну гіпотезу H_0, яка полягає в тому, що дисперсія генеральна дорівнює гіпотетичному значенню σ_0^2. Враховуючи незміщеність S^2 як оцінки дисперсії генеральної, основну гіпотезу можна записати таким чином:

$$H_0 : M(S^2) = \sigma_0^2. \tag{12.6}$$

Іншими словами, потрібно з'ясувати, істотно чи неістотно відрізняються виправлена вибіркова і гіпотетична генеральна дисперсії.

В якості критерія перевірки гіпотези (12.6) візьмемо випадкову величину

$$\chi^2 = (n-1)S^2/\sigma_0^2, \tag{12.7}$$

де права частина є випадковою величиною, розподіленою за законом χ^2 з $k = n-1$ ступенями вільності.

Критична область будується в залежності від виду конкуруючої гіпотези.

Правило 1. Для того, щоб при заданому рівні значущості α перевірити основну гіпотезу $H_0 : \sigma^2 = \sigma_0^2$ при конкуруючій гіпотезі $H_1 : \sigma^2 > \sigma_0^2$, потрібно обчислити спостережене значення критерія $\chi^2_{\text{спост}}$ за формулою (12.7) і за табл. 6 додатків по заданому рівню значущості α і числу ступенів вільності $k = n-1$ знайти критичну точку $\chi^2_{\text{кр}}(\alpha;k)$.

Якщо $\chi^2_{\text{спост}} < \chi^2_{\text{кр}}$, то нема підстав відкидати основну гіпотезу. Якщо $\chi^2_{\text{спост}} > \chi^2_{\text{кр}}$, то гіпотеза H_0 відкидається на користь гіпотези H_1.

Правило 2. Для того, щоб при заданому рівні значущості α перевірити гіпотезу $H_0 : \sigma^2 = \sigma_0^2$ при конкуруючій гіпотезі $H_1 : \sigma^2 \neq \sigma_0^2$, потрібно обчислити спостережене значення критерія $\chi^2_{\text{спост}}$ за формулою (12.7) і за табл. 6 додатків знайти ліву критичну точку $\chi^2_{\text{кр}}(1-\alpha/2;k)$ і праву критичну точку $\chi^2_{\text{кр}}(\alpha/2;k)$.

Якщо $\chi^2_{\text{лівост.кр}} < \chi^2_{\text{спост}} < \chi^2_{\text{правост.кр}}$, то нема підстав відкидати гіпотезу H_0.

Якщо $\chi^2_{\text{спост}} < \chi^2_{\text{лівост.кр}}$ або $\chi^2_{\text{спост}} > \chi^2_{\text{правост.кр}}$, то основна гіпотеза H_0 відкидається.

Правило 3*.* При конкуруючій гіпотезі $H_1 : \sigma^2 < \sigma_0^2$ знаходиться критична точка $\chi^2_{\text{кр}}(1-\alpha/2;k)$ лівосторонньої критичної області.

Якщо $\chi^2_{\text{спост}} > \chi^2_{\text{кр}}(1-\alpha/2;k)$, то нема підстав відкидати основну гіпотезу.

Якщо $\chi^2_{\text{спост}} < \chi^2_{\text{кр}}(1-\alpha/2;k)$, то гіпотеза H_0 відкидається.

4.

Одна із найважливіших задач математичної статистики — знаходження невідомого закону розподілу випадкової величини X — кількісної ознаки об'єктів генеральної сукупності. В лекції 10 були висвітлені питання опрацювання вибірки з метою отримання інформації про вид емпіричного розподілу та його характеристик: середньої ознаки, величини розкиду, симетричності розподілу. Потім на основі цих даних підбирається той розподіл, який найкраще апроксимує дослідний розподіл випадкової величини.

Після вибору виду розподілу необхідно знайти (хоча б наближено) параметри того закону розподілу, який характеризує досліджувану випадкову величину.

Оскільки в більшості практично важливих випадків параметри теоретичного закону розподілу є або математичним сподіванням, або дисперсією випадкової великими, або виражаються через них [7, 5.3.4 Ч1], то отримані вище висновки дають можливість знайти ці параметри за дослідними даними. Наприклад, для нормального розподілу параметри a та σ визначається рівностями $a = \bar{x}_в$, $\sigma = \sigma_в$ (або $\sigma = S$). Таким чином, можна говорити про "відомість" теоретичного розподілу досліджуваної ознаки X. При цьому лапки вказують на гіпотетичність такого знання.

Критерій узгодженості Пірсона $\left(\chi^2\right)$ грунтується на виборі певної міри розбіжності між теоретичним і емпіричним (дослідним) розподілами. При цьому задачу перевірки узгодженості можна сформулювати таким чином: на основі вибірки спостережених значень деякої випадкової величини X потрібно визначити, що емпіричний розподіл належить певному розподілу (нормальному, показниковому, біноміальному і т. д.) із *визначеними* параметрами — гіпотеза H_0 проти альтернативної гіпотези H_1 — розподіл не належить вибраному розподілу.

Нехай у відповідності із гіпотезою H_0 **відома функція розподілу імовірностей *F*(*x*) досліджуваної ознаки *X***. Позначимо через $S_1, S_2, ..., S_m$ множини, на які розбивається вся область можливих значень випадкової величини *X*; ці множини — або інтервали для **неперервної** випадкової величини, або групи окремих значень **дискретної** випадкової величини, які не мають спільних точок. Тоді можна обчислити імовірності того, що при випробуванні випадкова величина *X* набере значення із множини S_i, тобто

$$p_i = P(X \in S_i), \quad i = \overline{1,m}. \tag{12.8}$$

При цьому всі $p_i > 0, \ i = \overline{1,m},$ і $\sum\limits_{i=1}^{m} p_i = 1.$

Нехай $n_1, n_2, ..., n_m$ — відповідні емпіричні групові частоти, тобто суми частот тих варіант (значень випадкової величини *X* із вибірки), що потрапляють відповідно у множини $S_1, S_2, ..., S_m$.

Якщо гіпотеза H_0 правильна, то статистичний розподіл вибірки можна розглядати як емпіричний аналог для генерального розподілу, який визначається функцією $F(x)$. Це означає, що n_i є частотою (абсолютною) появи випадкової події $(X \in S_i), \ i = \overline{1,m},$ в послідовності із *n* спостережень. Отже, в першому розподілі кожній множині S_i ставляться у відповідність відносна частота $n_i/n,$ а в другому — імовірність p_i. Тоді за методом найменших квадратів в якості міри розходження між статистичним розподілом відносних частот вибірки і теоретичним розподілом імовірностей отримується міра розбіжності виду

$$\chi^2 = \sum_{i=1}^{m} \frac{(n_i - np_i)^2}{np_i} \tag{12.9}$$

така, що при збільшенні обсягу вибірки розподіл величини χ^2 наближається до граничного розподілу χ^2 із $k = m - r - 1$ ступенями вільності, де *m* — число інтервалів або груп, на які розбита вся множина спостережених даних, *r* — число параметрів гіпотетичного розподілу імовірностей, що оцінюється за даними вибірки.

Зауваження. Числа

$$n_i^0 = np_i, \quad i = \overline{1,m}, \tag{12.10}$$

називаються **теоретичними частотами** відбуття випадкових подій $(X \in S_i)$. Враховуючи рівності (12.10), із (12.9) отримаємо таку міру розбіжності між теоретичними і емпіричними частотами:

$$\chi^2 = \sum_{i=1}^{m} \frac{(n_i - n_i^0)^2}{n_i^0}. \tag{12.9*}$$

Випадкова величина, що визначається рівностями (12.9) або (12.9*), називається **критерієм Пірсона**.

Для перевірки гіпотези H_0 задамо рівень значущості α і за табл. 6 додатків знайдемо критичну точку $\chi^2_{\text{кр}}(\alpha; k),$ де $p = \alpha, \ k = m - r - 1.$ Таким чином, правостороння критична область визначається нерівністю $\chi^2 > \chi^2_{\text{кр}}(\alpha;k)$, а область прийняття нульової гіпотези — нерівністю $\chi^2 < \chi^2_{\text{кр}}(\alpha;k)$. За результатами вибірки обчислюємо $\chi^2_{\text{сп}}$. Якщо $\chi^2_{\text{сп}} < \chi^2_{\text{кр}}$, то гіпотезу H_0 приймаємо, а у випадку $\chi^2_{\text{сп}} > \chi^2_{\text{кр}}$ H_0 відкидаємо.

5.

Використання критерія узгодженості Пірсона проілюструємо для випадку перевірки гіпотези H_0: кількісна ознака X об'єктів генеральної сукупності розподілена за нормальним законом.

Нехай статистичний розподіл має такий вид:

$$\begin{array}{c|cccc} [\,x_i;x_{i+1}) & [\,x_1;x_2) & [\,x_2;x_3) & \dots & [x_m;x_{m+1}] \\ \hline n_i & n_1 & n_2 & \dots & n_m \end{array}, \tag{12.11}$$

де $x_{i+1} - x_i = h, \quad i = \overline{1,m}, \quad \sum_{i=1}^{k} n_i = n,$ число частинних інтервалів визначається наближеною рівністю $m \approx \log_2 n$. Якщо ж статистичний розподіл є дискретним, тоді потрібно перейти до інтервального (див. розв'язування задачі 1.2). Знайдемо для розподілу (12.11) $\bar{x}_в$ та $D_в$. Тоді вважаємо, що густина розподілу досліджуваної ознаки X має такий вид:

$$f(x) = \frac{1}{\sigma_в \sqrt{2\pi}} e^{-\frac{(x-\bar{x}_в)^2}{2\sigma_в^2}}.$$

Враховуючи те, що можливі значення нормально розподіленої випадкової величини заповнюють всю дійсну вісь, в якості множин $S_1, S_2, \dots S_m$ візьмемо інтервали

$$(-\infty; x_2), (x_2; x_3), \dots (x_m; \infty).$$

Для знаходження імовірностей $p_i = P(x_i < X < x_{i+1}), \quad i = \overline{1,m},$ де $x_1 = -\infty, \quad x_{m+1} = \infty,$ використаємо формулу

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right),$$

в якій a та σ — параметри нормального розподілу.
Отже, за табл. 3 додатків обчислимо

$$p_i = \Phi\left(\frac{x_{i+1} - \bar{x}_в}{\sigma_в}\right) - \Phi\left(\frac{x_i - \bar{x}_в}{\sigma_в}\right), \quad i = \overline{1,m},$$

поклавши $x_1 = -\infty, \quad x_{m+1} = \infty.$

В результаті за формулою (12.9) можна знайти для даної вибірки (розподілу (12.11)) $\chi^2_{\text{сп}}$. Для заданого рівня значущості α і ступеня вільності $k = m - 3$ (число параметрів нормального розподілу $r = 2$) за табл. 6 додатків знаходимо $\chi^2_{\text{кр}}(\alpha;k)$. Тоді гіпотеза H_0 приймається, якщо $\chi^2_{\text{сп}} < \chi^2_{\text{кр}}(\alpha;k)$, і відхиляється у випадку $\chi^2_{\text{сп}} > \chi^2_{\text{кр}}(\alpha;k)$.

Зауваження. Вказане вище число ступенів вільності $k = m - 3$ відноситься тільки до того випадку, коли обидва параметри нормального закону розподілу знаходяться за даними вибірки, тобто коли замість точних значень a і σ використовуються їх емпіричні значення $\bar{x}_в$ та $\sigma_в$. Якщо значення a точно відоме (наприклад, у випадку знаходження відхилень від еталону), то число ступенів вільності дорівнює $k = m - 2.$ Якщо ж відомі обидва параметри, то число ступенів вільності $k = m - 1.$ На практиці така ситуація зустрічається рідко, а тому для отримання числа ступенів вільності не менше п'яти потрібно, щоб число частинних інтервалів було не меншим восьми.

Критерій Пірсона можна використовувати як для випадку неперервної ознаки *X*, так і для дискретної. Проте недоліком його є те, що випадкова величина $\chi^2,$ як правило, залежить від групування варіаційного ряду вибірки. Тому іноді доцільним є використання інших критеріїв узгодженості.

Рекомендована література: [2] С. 281–346, [3] С. 95–149; [4], С. 141–145; [5], С. 223–249.

Лекція 13.

Тема: *Елементи кореляційного і регресійного аналізу.*

1. ***Поняття стохастичності та стохастичної залежності, кореляції та регресії.***
2. ***Основні задачі кореляційного та регресійного аналізу.***
3. ***Лінійні емпіричні рівняння парної кореляції.***
4. ***Вибірковий коефіцієнт лінійної кореляції та його властивості.***
5. ***Нелінійна парна кореляція.***

1.

Дві випадкові величини можуть бути або незалежними, або пов'язані функціональною залежністю, або залежністю, яка називається стохастичною. Функціональна залежність, розглянута в [7, § 8], досить рідко зустрічається в економічних дослідженнях. Дуже часто доводиться вивчати такі випадкові величини, для яких зміна можливих значень (при проведенні випробувань) приводить до зміни закону (умовного) розподілу іншої (див. [7, § 7]). Такий зв'язок між випадковими величинами називається **стохастичним**; він виникає тоді, коли на обидві величини впливають випадкові фактори, серед яких є спільні.

Найбільш важливим випадком стохастичного зв'язку є так званий **кореляційний** зв'язок, який встановлює залежність між значеннями випадкової величини X і умовним математичним сподіванням $M(Y \mid X = x)$ *) випадкової величини Y:

$$M(Y \mid X = x) = g(x), \tag{13.1}$$

або між значенням випадкової величини Y і умовним математичним сподіванням $M(X \mid Y = y)$ *) випадкової величини X:

$$M(X \mid Y = y) = q(y). \tag{13.2}$$

Функції $g(x)$ і $q(y)$ називаються **функціями регресії** Y на X та X на Y відповідно, а їх графіки — лініями регресії Y на X та X на Y, рівняння (13.1) та (13.2) називаються **рівняннями регресії *Y* на *X* та *X* на *Y*** відповідно.

Використовуючи інформацію §§ 7, 8 [7] можна отримати важливу для наступного викладу основну властивість регресії величини Y на величину X: якщо $g(x)$ є функцією регресії величини Y на X, то математичне сподівання квадрата відхилення величини Y від функції $g(x)$ менше, ніж від будь-якої функції $h(X) \neq g(X)$, тобто

$$M\left[\,Y - g(X)\right]^2 < M\left[\,Y - h(X)\right]^2. \tag{13.3}$$

Аналогічною властивістю володіє і регресія величини X на величину Y.

Прикладами кореляційного зв'язку є стохастична взаємозалежність між: 1) окремими параметрами тіла людини або тварини; 2) обсягом виробництва підприємства та коефіцієнтом використання основних засобів.

На практиці сумісний закон розподілу випадкових величин X та Y (двовимірної випадкової величини (X; Y) [7, §7]) невідомий. В розпорядженні дослідника є двовимірна вибірка

$$(x^{(1)}; y^{(1)}), (x^{(2)}; y^{(2)}), ..., (x^{(n)}; y^{(n)}), \qquad (13.4)$$

згрупувавши (для великих n) дані якої, можна отримати двовимірний статистичний розподіл, наведений в табл. 4.1.

Таблиця 13.1

X	Y						
	y_1	y_2	...	y_j	...	y_m	n_{x_i}
x_1	n_{11}	n_{12}	...	n_{1j}	...	n_{1m}	n_{x_1}
x_2	n_{21}	n_{22}	...	n_{2j}	...	n_{2m}	n_{x_2}
...	...	...	...	...	...	...	...
x_i	n_{i1}	n_{i2}	...	n_{ij}	...	n_{im}	n_{x_i}
...	...	...	...	...	...	...	...
x_k	n_{k1}	n_{k2}	...	n_{kj}	...	n_{km}	n_{x_k}
n_{y_j}	n_{y_1}	n_{y_2}	...	n_{y_j}	...	n_{y_m}	n

Тут n_{ij} — частота спільної появи ознак x_i, y_j (пари $(x_i; y_j)$); $\sum\limits_{i=1}^{k}\sum\limits_{j=1}^{m} n_{ij} = n$ — обсяг вибірки;

$$n_{x_i} = \sum_{j=1}^{m} n_{ij}, \quad (i = \overline{1;k}); \qquad n_{y_j} = \sum_{i=1}^{k} n_{ij}, \quad (j = \overline{1;k}).$$

Кожному значенню X відповідає ряд значень Y, тобто зміна значень X приводить до зміни умовного статистичного розподілу $Y \mid x.$ Аналогічно зміна значень Y веде до зміни умовного статистичного розподілу $X \mid y.$

Наприклад, при $X = x_1$ та $X = x_2$ відповідні умовні статистичні розподіли величини Y мають такі види:

$$\begin{array}{c|cccc} Y \mid x_1 & y_1 & y_2 & ... & y_m \\ \hline n_i & n_{11} & n_{12} & ... & n_{1m} \end{array}; \qquad \begin{array}{c|cccc} Y \mid x_2 & y_1 & y_2 & ... & y_m \\ \hline n_i & n_{21} & n_{22} & ... & n_{2m} \end{array}.$$

Статистичною називається така залежність між величинами X та Y, для якої зміна спостережених значень однієї із величин зумовлює зміну умовного статистичного розподілу іншої.

Умовною середньою $\bar{y}_x$ називається середнє арифметичне спостережених значень Y, які відповідають значенню $X = x$. Наприклад, згідно із табл. 4.1

$$\bar{y}_{x_2} = \frac{y_1 n_{21} + y_2 n_{22} + ... + y_j n_{2j} + ... + y_m n_{2m}}{n_{x_2}}.$$

Умовною середньою $\bar{x}_y$ називається середнє арифметичне спостережених значень X, які відповідають значенню $Y = y$. Наприклад, у відповідності із табл. 13.1

$$\bar{x}_{y_1} = \frac{x_1 n_{11} + x_2 n_{21} + ... + x_k n_{k1}}{n_{y_1}}.$$

Можна довести, що умовні середні є незміщеними і спроможними точковими статистичними оцінками відповідних умовних математичних сподівань:

$$\bar{y}_x \approx M(Y \mid X = x); \qquad \bar{x}_y \approx M(X \mid Y = y). \tag{13.5}$$

Вкажемо підхід до знаходження статистичних наближень (емпіричних функцій) $\hat{g}(x) = \hat{g}(x; a_0, a_1, ..., a_m)$ та $\hat{q}(y) = \hat{q}(y; b_0, b_1, ..., b_m)$ невідомих функцій регресій $g(x)$ та $q(y)$ відповідно, а отже, з врахуванням співвідношень (13.5) і емпіричних рівнянь *Y* на *X*

$$\bar{y}_x = \hat{g}(x, a_0, a_1, ..., a_m) \tag{13.6}$$

та *X* на *Y*

$$\bar{x}_y = \hat{q}(y, b_0, b_1, ..., b_m) \tag{13.7}$$

де $a_0, a_1, ..., a_m$ та $b_0, b_1, ..., b_m$ — невідомі параметри.

На першому кроці в прямокутній системі координат *xOy* будується послідовність пар чисел (13.4). Отримана сукупність точок називається **кореляційним полем** або **діаграмою розсіювання**. Конфігурація кореляційного поля дозволяє висунути гіпотезу про вид функції $\hat{g}(x)$. Наприклад, у випадку кореляційного поля, зображеного на рис. 4.1, $\hat{g}(x)$ доцільно шукати у вигляді лінійної функції $a_1 x + a_0$, а у випадку рис. 4.2 у вигляді параболічної функції другого порядку $a_2 x^2 + a_1 x + a_0$.

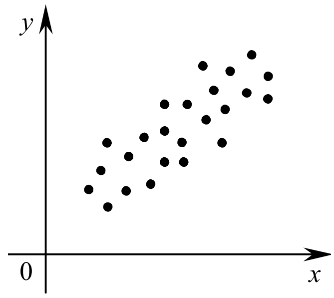


Рис. 13.1.

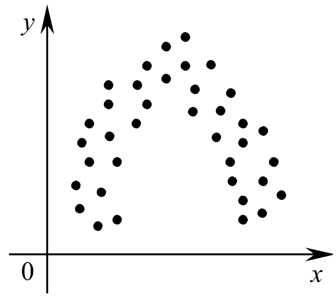


Рис. 13.2.

На другому кроці потрібно знайти невідомі параметри. Ця задача розв'язується з допомогою так званого **методу найменших квадратів** (МНК), суть якого полягає в наступному. Оскільки невідома функція регресії *Y* на *X* $g(X)$ згідно із (13.3) мінімізує величину $M[\,Y - h(X)]^2$, а оцінкою $M(Y \mid x)$ є статистична середня $\bar{y}_x$, що відповідає певним спостереженим значенням *X* = *x*, то емпірична лінія регресії (13.6) повинна задовольняти рівності

$$F(a_0, a_1, ..., a_m) \equiv \sum_{i=1}^{n} \left[\, y^{(i)} - \hat{g}(x^{(i)}; a_0, a_1, ..., a_m) \right]^2 \Rightarrow \min. \tag{13.8}$$

Права частина цієї рівності є сума квадратів віддалей вздовж осі 0*y* від точок $(x^{(i)}; y^{(i)})$ послідовності (13.4) до відповідних точок лінії регресії із тією ж абсцисою.

Для випадку згрупованих даних із табл. 4.1 рівність (13.8) набирає такого виду

$$F(a_0, a_1, ..., a_m) \equiv \sum_{i=1}^{k} \sum_{j=1}^{m} \left[\, y_j - \hat{g}(x_i; a_0, a_1, ..., a_m) \right]^2 n_{ij} \Rightarrow \min. \tag{13.8*}$$

Таким чином, емпірична функція $\hat{g}(x; a_0, a_1, ..., a_m)$ повинна усереднити (згладити) спостережені дані $(x^{(i)}; y^{(i)})$, $i = \overline{1,n}$, з послідовності (13.4) або (x_i, y_j), $i = \overline{1,k}$, $j = \overline{1,m}$, із табл. 4.1. При цьому невідомі параметри $a_0, a_1, ..., a_m$ повинні мінімізувати функцію $F(a_0, a_1, ..., a_m)$; необхідною умовою цього є виконання системи рівнянь

$$\left\{ \frac{\partial F(a_0, a_1, ..., a_m)}{\partial a_i} = 0, \quad i = \overline{0, m}, \right. \tag{13.9}$$

яка називається **системою нормальних рівнянь**.

Метод найменших квадратів дозволяє **при встановленому виді емпіричної функції регресії** $\hat{g}(x; a_0, a_1, ..., a_m)$ так знайти невідомі параметри $a_0, a_1, ..., a_m$, що вона буде найкращою оцінкою функції регресії $g(x)$ в тому розумінні, що сума квадратів відхилень спостережених значень випадкової величини Y від відповідних ординат емпіричної функції буде найменшою.

Аналогічно знаходиться емпірична функція $\hat{q}(y; b_0, b_1, ..., b_m)$ регресії X на Y.

Зауваження. Враховуючи випадковий характер організації вибірки, можна зробити висновок, що для нефіксованої вибірки знайдені параметри є випадковими величинами.

Знаходження емпіричних рівнянь регресії — це тільки перший крок дослідження статистичних зв'язків між випадковими величинами. Наступним є встановлення сили або тісноти цих зв'язків. Відомо [7, п. 7.6], що мірою лінійного зв'язку (стохастичного) між двома випадковими величинами X та Y є коефіцієнт кореляції

$$r = r_{XY} = \frac{K_{XY}}{\sigma_X \sigma_Y} = \frac{M(XY) - M(X)M(Y)}{\sigma_X \sigma_Y}, \tag{13.10}$$

де K_{XY} — кореляційний момент або коефіцієнт коваріації (сумісної варіації), який ще часто позначають $\operatorname{cov}(X, Y)$, $\sigma_X = \sqrt{D(X)}$, $\sigma_Y = \sqrt{D(Y)}$. Зокрема, якщо величини X та Y незалежні, то коефіцієнт кореляції $r = 0$; якщо $r = \pm 1$, то X та Y пов'язані **лінійною** функціональною залежністю. Звідси випливає, що коефіцієнт кореляції r вимірює силу (тісноту) **лінійного** зв'язку між X та Y.

Використовуючи метод моментів, тобто замінивши числові характеристики їх статистичними оцінками, можна отримати точкову статистичну оцінку коефіцієнта кореляції — вибірковий (емпіричний) коефіцієнт кореляції.

$$r_в = r_{xy} = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x \sigma_y}, \tag{13.11}$$

де для незгрупованих даних (13.4)

$$\overline{xy} = \sum_{i=1}^{n} x^{(i)} y^{(i)} \big/ n, \tag{13.12}$$

а для згрупованих даних із табл. 4.1

$$\overline{xy} = \sum_{i=1}^{k} \sum_{j=1}^{m} n_{ij}\, x_i y_j \big/ n = \sum n_{xy}\, xy \big/ n, \tag{13.13}$$

σ_x та σ_y — середні квадратичні відхилення вибіркові для ознак X та Y відповідно.

Зауваження. Оскільки кореляційний момент рівносильно (13.10) визначається ще й таким чином: $K_{XY} = M\{\, [\, X - M(X)]\, [Y - M(Y)]\, \}$, то формулу (13.11) можна подати і в такому вигляді

$$r_в = \frac{\sum_{i=1}^{n}(x^{(i)} - \bar{x})(y^{(i)} - \bar{y})}{n\sigma_x\sigma_y}. \tag{13.11*}$$

Оскільки $r_в$ для нефіксованої вибірки є випадковою величиною, то знайдене значення $r_в$ для конкретної вибірки може суттєво відрізнятись від r. В зв'язку із цим необхідно побудувати довірчі інтервали для оцінки r, а також перевірити статистичну гіпотезу про некорельованість X та Y $(r = o)$.

У випадку нелінійності хоча б однієї із функцій регресії $g(x)$ та $q(y)$ коефіцієнт кореляції r вже не дає інформації про силу зв'язку між X та Y. Натомість ступінь концентрації розподілу поблизу лінії регресії показує **кореляційне відношення** Y на X:

$$\eta_{Y/X} = 1 - \frac{\sigma_{Y/X}^2}{\sigma_Y^2}, \tag{13.14}$$

де

$$\sigma_{Y/X}^2 = D(Y \mid X = x) = M\left\{ \left[Y - g(x) \right]^2 \mid X = x \right\}. \tag{13.15}$$

Із означення випливає, що кореляційне відношення змінюється в межах від 0 до 1 включно; воно дорівнює нулю тоді і тільки тоді, коли $\sigma_{Y/X}^2 = 0$, тобто весь розподіл зосереджений на лінії регресії (має місце функціональна залежність). Це відношення дорівнює нулю тоді і тільки тоді, коли лінія регресій Y на X являє собою горизонтальну пряму, що проходить через центр розподілу, тобто якщо Y та X некорельовані.

Аналогічно вводиться означення кореляційного відношення X на Y $\eta_{X/Y}$.

Можна довести, що у всіх випадках виконуються нерівності

$$r^2 \le \eta_{X/Y}^2, \quad r^2 \le \eta_{Y/X}^2. \tag{13.16}$$

За статистичними даними двовимірної вибірки можна знайти вибіркові кореляційні відношення $\hat{\eta}_{Y/X}$ та $\hat{\eta}_{X/Y}$, які є точковими статистичними оцінками $\eta_{Y/X}$ та $\eta_{X/Y}$ відповідно.

Сукупність методів оцінки кореляційних характеристик і перевірка статистичних гіпотез про них за даними вибірки називається **кореляційний аналізом**. В кореляційному аналізі використовуються такі основні методи:

1) побудова кореляційного поля і складання кореляційної таблиці;

2) знаходження вибіркового (емпіричного) коефіцієнта кореляції або кореляційного відношення;

3) перевірка статистичних гіпотез про значущість (істотність) зв'язку.

В практично важливих задачах одна із випадкових величин є "результуючою", а інша — від якої вона залежить (факторна величина). Наприклад, залежність урожайності сільськогосподарських культур Y від маси внесених добрив X. Це зразок односторонньої залежності між випадковими величинами, яка досліджується в регресійному аналізі.

Більш того, факторна величин може бути **детермінованою** (невипадковою), тобто значення якої не тільки відоме, але й ним можна керувати. При дослідженні лінійної залежності результуючої величини Y від факторної величини X загальний вираз емпіричного рівняння регресії має такий вид

$$\bar{y} = a_0 + a_1 x. \qquad (13.17)$$

В загальному випадку емпіричне рівняння регресії визначається таким чином

$$\bar{y} = \hat{g}(x; a_0, a_1, ..., a_m), \qquad (13.18)$$

де $a_0, a_1, ..., a_m$ — невідомі параметри.

2.

Тепер, коли визначені об'єкти дослідження цього параграфа (рівняння регресії (13.6), (13.7), (13.17), (13.18)), можемо сформулювати основні задачі регресійного аналізу.

Регресійний аналіз — це дослідження односторонніх статистичних залежностей між випадковими величинами. При цьому деякі із факторних величин можуть бути невипадковими величинами.

Задачі регресійного аналізу:

1) визначення форми залежностей;

2) знаходження функції регресії;

3) побудова точкових та інтервальних оцінок параметрів функції регресії;

4) знаходження точкових та інтервальних оцінок умовних математичних сподівань, необхідних для визначення меж, в яких із заданою надійністю будуть міститися середні значення досліджуваної величини, якщо інші пов'язані з нею величини набирають певних значень;

5) перевірка узгодженості знайденої емпіричної функції регресії спостереженим даним.

Основна мета регресійного аналізу — **теоретично обгрунтований і статистично надійний точковий та інтервальний прогноз** значень залежної величини або умовного математичного сподівання цієї величини.

Зауваження. Якщо рівняння регресії описує об'єкт дослідження із економічної сфери і воно обґрунтоване в теоретично-економічному відношенні, то його називають **економетричним рівнянням**.

3.

Найбільш простим випадком є той, коли **обидві функції регресії** $g(x)$ і $q(y)$ в рівняннях регресії (13.1) та (13.2) є лінійними, тобто обидві лінії регресії є прямими лініями; вони називаються **прямими регресії**. В цьому випадку будемо говорити про лінійну кореляцію між випадковими величинами X та Y.

Можна довести, що коли сумісний розподіл імовірностей величин X і Y є **нормальним розподілом** на площині (див. [7, п. 7.8]), тоді **кореляційний зв'язок завжди є лінійним**. В зв'язку із цим слід очікувати лінійного кореляційного зв'язку між статистично залежними випадковими величинами

X та Y, якщо кожну із них можна розглядати як суму великого числа незалежних або майже незалежних випадкових доданків.

Нехай конфігурація кореляційного поля, отриманого внаслідок зображення у вигляді точки в прямокутній декартовій системі координат $x0y$ кожної із пар чисел вибіркової послідовності (13.4), дозволяє висунути припущення про лінійну кореляційну залежність між X та Y. Тобто рівняння регресії (13.1) та (13.2) в цьому випадку набирають відповідно такого виду

$$M(Y \mid X = x) = \alpha_1 x + \alpha_0, \tag{13.19}$$

$$M(X \mid Y = y) = \beta_1 y + \beta_0, \tag{13.20}$$

де α_i, β_i $(i = 1,0)$ — сталі.

Тоді емпіричні рівняння регресії Y на X та X на Y будемо шукати у вигляді

$$\overline{y}_x = a_1 x + a_0, \tag{13.21}$$

$$\overline{x}_y = b_1 y + b_0, \tag{13.22}$$

де a_0, a_1 та b_0, b_1 — невідомі параметри, які є **точковими статистичними оцінками** відповідних **чисел** рівнянь (13.19) і (13.20).

Використавши метод найменших квадратів як для незгрупованих даних(13.4), так і для згрупованих даних з табл.4.1, можна отримати **систему нормальних рівнянь**

$$\begin{cases} \overline{x^2} a_1 + \overline{x} a_0 = \overline{xy}, \\ \overline{x} a_1 + a_0 = \overline{y}, \end{cases} \tag{13.23}$$

для знаходження параметрів рівняння регресії (Y на X) (13.21).

Система (13.23) має єдиний розв'язок

$$a_1 = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\overline{x^2} - (\overline{x})^2}, \quad a_0 = \overline{y} - \overline{x} a_1. \tag{13.24}$$

Коефіцієнт a_1 рівняння (13.21) емпіричної прямої регресії Y на X називається **вибірковим (емпіричним або статистичним) коефіцієнтом регресії *Y* на *X*** і позначається $a_{Y/X}$. Оскільки $\overline{x^2} - (\overline{x})^2 = \sigma_x^2$, а

$$\overline{xy} - \overline{x} \cdot \overline{y} = K^*_{XY} (= \overline{\mathrm{cov}}(X, Y)) \tag{13.25}$$

— **вибірковий (емпіричний або статистичний) кореляційний момент** або **вибіркова коваріація**, то першу формулу (13.24) можна записати в такому виді

$$a_{Y/X} = a_1 = \frac{K^*_{XY}}{\sigma_X^2} \left(\equiv \frac{\overline{\mathrm{cov}}(X, Y)}{\sigma_X^2} \right), \tag{13.26}$$

а шукане рівняння (13.21) із врахуванням другої рівності (13.25) набуде вигляду

$$\overline{y}_x - \overline{y} = \frac{K^*_{XY}}{\sigma_X^2} (x - \overline{x}). \tag{13.27}$$

Це рівняння показує, що емпірична пряма регресії Y на X проходить через точку з координатами $(\overline{x}, \overline{y})$, яка називається **середньою точкою кореляційного поля**.

Аналогічно можна отримати емпіричне рівняння прямої X на Y, якщо мінімізувати сумарні квадрати відхилень точок $(\bar{x}_i, y_i)$, $i=\overline{1,m}$, від шуканої прямої, тобто прямої

$$\bar{x}_y - \bar{x} = b_{X/Y}(y - \bar{y}), \tag{13.28}$$

де вибірковий (емпіричний або статистичний) коефіцієнт регресії X на Y визначається за формулою

$$b_{X/Y} = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\sigma_y^2} = \frac{K_{XY}^*}{\sigma_y^2}\left(\equiv \frac{\overline{\mathrm{cov}}(X,Y)}{\sigma_y^2}\right). \tag{13.29}$$

У формулах вибіркових коефіцієнтів регресії (13.26) і (13.29) чисельники співпадають, а знаменники завжди додатні, оскільки є дисперсіями вибірковими випадкових величин X та Y відповідно. Тому $a_{Y/X}$ і $b_{X/Y}$ мають однакові знаки. Відмітимо також, що вибірковий коефіцієнт регресії $a_{Y/X}(b_{X/Y})$ — це міра, яка на основі вибіркових даних в середньому вказує на вплив зміни змінної X (або Y) на змінну Y (або X).

Рівністю (13.11) **формально** був визначений вибірковий (емпіричний) коефіцієнт кореляції $r_в$ як точкова статистична оцінка коефіцієнта кореляції $r = r_{XY}$ — міри лінійного кореляційного зв'язку між випадковими величинами X та Y. Проте вид емпіричних коефіцієнтів регресії Y на X та X на Y вказує на природний зв'язок $r_в$ із $a_{Y/X}$ і $b_{X/Y}$ (за аналогією із зв'язком в теорії кореляції як розділу теорії імовірностей). А тому виникає необхідність детально вивчити властивості $r_в$ як **оцінки** сили лінійного кореляційного зв'язку між величинами X та Y. При цьому слід очікувати властивостей, аналогічних із властивостями r_{XY}.

4.

Вибірковим (емпіричним або статистичним) коефіцієнтом кореляції $r_в = r_{XY}$ випадкових величини X та Y, між якими припускається лінійний кореляційний зв'язок, називається відношення емпіричного кореляційного моменту (коефіцієнта коваріації) $K^*(X,Y)(=\overline{\mathrm{cov}}(X,Y))$ до добутку середніх квадратичних відхилень вибіркових σ_x та σ_y:

$$r_в = r_{xy} = \frac{K_{XY}^*}{\sigma_x\sigma_y}\left(= \frac{\overline{\mathrm{cov}}(X,Y)}{\sigma_x\sigma_y}\right) = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\sigma_x\sigma_y}, \tag{13.30}$$

де $\overline{xy}$ для незгрупованих даних (13.4) вибірки обчислюється за формулою (13.12), а для згрупованих даних із табл. 4.1 — за (13.13).

Зауваження. Для випадку, коли x_i та y_i $(i=\overline{1,k})$ є великими числами, а обсяг вибірки $n \ge 50$, то зручніше користуватися такою формулою

$$r_в = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]\left[\sum_{i=1}^{n}(y_i - \bar{y})^2\right]}}. \tag{13.30*}$$

Використовуючи формули (13.26) і (13.29), отримаємо вираз через емпіричні коефіцієнти регресій для вибіркового коефіцієнта кореляції:

$$r_в = \pm\sqrt{\frac{(K_{XY}^*)^2}{\sigma_x^2\sigma_y^2}} = \pm\sqrt{\frac{K_{XY}^*}{\sigma_x^2}\cdot\frac{K_{XY}^*}{\sigma_y^2}} = \pm\sqrt{a_{Y/X}b_{X/Y}},$$

де знак перед коренем визначається знаком емпіричних коефіцієнтів регресій.

З другого боку, і емпіричні коефіцієнти регресій можна виразити через вибірковий коефіцієнт кореляції:

$$a_{Y/X} = \frac{K_{XY}^*}{\sigma_x^2} = \frac{K_{XY}^*}{\sigma_x\sigma_y}\cdot\frac{\sigma_y}{\sigma_x} = r_в\frac{\sigma_y}{\sigma_x}, \qquad b_{X/Y} = \frac{K_{XY}^*}{\sigma_y^2} = \frac{K_{XY}^*}{\sigma_x\sigma_y}\cdot\frac{\sigma_x}{\sigma_y} = r_в\frac{\sigma_x}{\sigma_y}.$$

Тоді емпіричне рівняння прямих регресій із врахуванням цих формул можна записати в такому вигляді:

$$\overline{y}_x - \overline{y} = r_в\frac{\sigma_y}{\sigma_x}(x - \overline{x}), \tag{13.31}$$

$$\overline{x}_y - \overline{x} = r_в\frac{\sigma_x}{\sigma_y}(y - \overline{y}). \tag{13.32}$$

Наведемо еквівалентні (13.30) вирази вибіркового коефіцієнта кореляції.

Нехай $\{(x_i, y_i), i = \overline{1,n}\}$ — незгруповані дані двовимірної вибірки обсягом n. Величину

$$D_{yx} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}\left[y_i - \overline{y} - r_в\frac{\sigma_y}{\sigma_x}(x_i - \overline{x})\right]^2 \tag{13.33}$$

природно називати дисперсією спостережених значень y_i навколо емпіричної прямої регресії (13.31) Y на X, врахувавши при цьому, що геометрично різниця $y_i - \widehat{y}_i$ означає відхилення по ординаті точки (x_i, y_i) від точки $(x_i, \widehat{y}_i)$ прямої (13.31).

Можна довести, що

$$D_{yx} = \sigma_y^2(1 - r_в^2), \qquad r_в = \pm\sqrt{1 - \frac{D_{yx}}{\sigma_y^2}}, \tag{13.30**}$$

де знак перед коренем вибирається у відповідності із знаком коефіцієнта регресії.

Аналогічно ввівши дисперсію спостережених значень x_i навколо емпіричної прямої регресії (13.32) X на Y за формулою

$$D_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \widehat{x}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}\left[x_i - \overline{x} - r_в\frac{\sigma_x}{\sigma_y}(y_i - \overline{y})\right]^2,$$

можна отримати рівність

$$r_в = \pm\sqrt{1 - \frac{D_{xy}}{\sigma_x^2}}. \tag{13.30***}$$

Властивості вибіркового коефіцієнта кореляції

Властивість 1. Величина $r_в$ є безрозмірною, тобто вона не залежить від вибору одиниць виміру випадкових величин X та Y.

Властивість 2. Вибірковий коефіцієнт кореляції $r_в$ не перевищує за абсолютною величиною одиницю, тобто $|r_в| \le 1$.

Властивість 3. Вибірковий коефіцієнт кореляції $r_в = \pm 1$ тоді і тільки тоді, коли між випадковими величинами X та Y існує лінійний функціональний зв'язок.

Властивість 4. Якщо між випадковими величинами X і Y відсутній хоча б один із кореляційних зв'язків, то вибірковий коефіцієнт кореляції $r_в$ дорівнює нулю.

Властивість 5. Рівність $r_в = \pm 1$ є необхідною і достатньою умовою співпадань регресій Y на X і X на Y.

Із розглянутих властивостей $r_в$ можна зробити висновок про те, що вибірковий коефіцієнт кореляції є мірою тісноти (сили) лінійного кореляційного зв'язку між випадковими величинами X і Y. Справді, якщо $|r_в| = 1,$ то між X і Y існує лінійний функціональний зв'язок, а якщо $r_в = 0,$ то лінійний кореляційний зв'язок відсутній. А із формул (13.30**) і (13.30***) випливає, що у випадку збільшення $|r_в|$ до одиниці сила кореляційного зв'язку зростає, оскільки сума квадратів відхилень спостережених значень від прямих регресій прямує до нуля. Якщо ж $|r_в| \to 0,$ то сила кореляційного зв'язку зменшується, бо сума квадратів відхилень зростає.

6.

Розглянемо тепер випадок, коли хоча б одна із двох послідовностей точок

$$(x_1, \overline{y}_{x_1}), (x_2, \overline{y}_{x_2}), \ldots, (x_k, \overline{y}_{x_k}), \tag{13.34}$$

$$(y_1, \overline{x}_{y_1}), (y_2, \overline{x}_{y_2}), \ldots, (y_m, \overline{x}_{y_m}) \tag{13.35}$$

дає підстави зробити висновок про існування нелінійного кореляційного зв'язку між випадковими величинами X та Y. На основі цих статистичних даних потрібно оцінити параметри нелінійної регресії та силу кореляційної залежності.

Нехай, наприклад, конфігурація точок (13.34) дозволяє зробити припущення про наявність параболічної кореляції другого порядку. В цьому випадку вибіркове рівняння регресії Y на X слід шукати в такому вигляді

$$\overline{y}_x = a_2 x^2 + a_1 x + a_0, \tag{13.36}$$

де $a_i\,(i = 0,1,2)$ — невідомі параметри.

Користуючись методом найменших квадратів, можна отримати систему нормальних рівнянь:

$$\begin{cases} \overline{x^4} a_2 + \overline{x^3} a_1 + \overline{x^2} a_0 = \overline{\overline{y}_x x^2}, \\ \overline{x^3} a_2 + \overline{x^2} a_1 + \overline{x} a_0 = \overline{\overline{y}_x x}, \\ \overline{x^2} a_2 + \overline{x} a_1 + a_0 = \overline{\overline{y}_x}, \end{cases} \tag{13.37}$$

де

$$\overline{x^l} = \left(\sum_{i=1}^{k} x_i^l n_{x_i} \right) \Big/ n, \quad l = \overline{1,4}, \qquad \overline{\overline{y}_x x^l} = \left(\sum_{i=1}^{k} \overline{y}_{x_i} x_i^l n_{x_i} \right) \Big/ n, \quad l = \overline{0,2}. \tag{13.38}$$

Оскільки $\overline{\overline{y}_x} = \overline{y},$ то розв'язавши третє рівняння системи (13.37) відносно a_0 і підставивши в (13.36), після простих перетворень отримаємо

$$\overline{y}_x = \overline{y} + a_1(x - \overline{x}) + a_2(x^2 - \overline{x^2}). \qquad (13.36^*)$$

Знайдені із системи рівнянь (13.37) параметри a_1 та a_2 підставимо в (13.36*); в підсумку отримуємо шукане рівняння регресії *Y* на *X*.

У випадку параболічної регресії (другого порядку) *X* на *Y*

$$\overline{x}_y = b_2 y^2 + b_1 y + b_0$$

невідомі параметри b_2, b_1, b_0 знаходяться як розв'язок такої системи нормальних рівнянь

$$\begin{cases} \overline{y^4} b_2 + \overline{y^3} b_1 + \overline{y^2} b_0 = \overline{\overline{x}_y y^2}, \\ \overline{y^3} b_2 + \overline{y^2} b_1 + \overline{y} b_0 = \overline{\overline{x}_y y}, \\ \overline{y^2} b_2 + \overline{y} b_1 + b_0 = \overline{x}. \end{cases}$$

Рекомендована література: [2] С. 253–280, [3] С. 153–220; [4], С. 141–145; [5], С. 250–300.

Основна література

1. Єрьоменко В. О., Шинкарик М. І. Теорія ймовірностей. Навчальний посібник для студентів економічних спеціальностей. — Тернопіль: Економічна думка, 2000. — 176 с.
2. Гмурман В. Е. Теория вероятностей и математическая статистика. — М.: Высшая школа, 1977. — 480 с.
3. Єрьоменко В. О., Шинкарик М. І. Математична статистика. Навчальний посібник для студентів економічних спеціальностей. — Тернопіль: Економічна думка, 2002. — 248 с.
4. Бугір М. К. Практикум з теорії ймовірності та математичної статистики. Навчальний посібник. — Тернопіль: ЦМДС, 1998. — 164 с.
5. Єрьоменко В.О., Шинкарик М.І., Бабій Р.М., Процик А.І. Практикум з теорії ймовірностей та математичної статистики. – Тернопіль: Економічна думка, 2005. – 317с.
6. Жлуктенко В. І., Наконечний С. І. Теорія ймовірностей і математична статистика: Навч.методичний посібник у 2-х ч. – ч. І Теорія ймовірностей . – К.: КНЕУ, 2000. -304с.
7. Жлуктенко В. І., Наконечний С. І. Теорія ймовірностей і математична статистика: Навч.методичний посібник у 2-х ч. – ч. ІІ Математична статистика . – К.: КНЕУ, 2003. -316с.

Додаткова література

1. Кремер М.Ш. Теория вероятностей и математическая статистика. – М.: ЮНИТИДАНА , 2002. - 543с.
2. Кибзун А.И., Коротков Е.Р., Наумов А.И., Сиротин А.Н. Теоря вероятностей и математическая статистика. Базовый курс с примерами и задачами / Учебн. пособие. М.: ФИЗМАТЛИТ, 2002. - 224с.
3. Малыхин В. И. Математика в экономике. Учебное пособие. — М. ИНФРА-М, 2002. — 352 с.
4. Павлова Л., Дітчук Р. Елементи комбінаторики і стохастики. — Тернопіль, Підручники і посібники, 2005. — 160 с.