MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE

WEST UKRAINIAN NATIONAL UNIVERSITY

Methodical instructions

for solution training tasks in discipline

«THEORY PROBABILITY AND MATHEMATICAL STATISTICS»

Ternopil – 2022

UDK 519.2

Reviewers:

S.V. Martyniuk – PhD, Candidate of physical and mathematical sciences, associate professor of the Department of Mathematics and Teaching Methods at Volodymyr Hnatyuk Ternopil National Pedagogical University.

O.S. Bashutska – PhD, Candidate of Economics, associate Professor of the Department of Economic Cybernetics and Informatics of the West Ukrainian National University;

approved at the meeting of the Department of Applied Mathematics, protocol № 1 of 26.08.2022.

Methodical instructions for solutions training tasks in discipline «Theory Probability and Mathematical Statistics» include examples of solution the tasks in the discipline «*Theory Probability and Mathematical Statistics*»

.

Plaskon S.A. Methodical instructions for solution training tasks in discipline «Theory Probability and Mathematical Statistics». - Ternopil: WUNU, 2022.- 15 p.

UDK 519.2

Responsible for release: O.M. Martyniuk, PhD, Candidate of physical and mathematical sciences, associate professor of the Department of Applied Mathematics,
Head of the Department of Applied Mathematics WUNU

Training tasks enable students to:

learn more deeply and consolidate the theoretical knowledge obtained at the lectures;
carry out qualitative and quantitative mathematical analysis of random events, random variables and systems of values;
•provide systematic mathematical processing of data;
• carry out statistical estimates (point and interval) of population parameters;
• use the correlation, regression and analysis of variance;
• use the results of research in the study of mathematical models of economic problems;
• to test statistical hypotheses.

Topic of the training – solution of practical task for statistical simple and analyze the result of calculation.

Student should take the task for solution according to your variant (two last numbers of your credit book) and solve it during time of training.

TASK.

For interval statistical distributions given in the conditions of tasks No. 1-50, it is necessary:

a) construct histograms of frequencies and relative frequencies;
b) to find the empirical distribution function and to construct its schedule;
c) compute summary characteristics of the sample;
d) find a confidence interval that cover with reliability γ = 0,95 mean breaking the whole party of statistical value;
e) find a confidence probability that the sample mean depart trom the general average in absolute value no more than 0.5.

Tasks number 1-10.
The data on the selective testing of strength yarns are given in the tables, in the first row of which are the partial intervals of the strength of the threads (kg), and in the second - the number of threads from the appropriate interval.

№ 1

$[x_i; x_{i+1})$	[1,4; 1,6)	[1,6; 1,8)	[1,8; 2)	[2; 2,2)	[2,2; 2,4)	[2,4; 2,6)	[2,6; 2,8]
n_i	6	10	18	25	20	13	8

№ 2

$[x_i; x_{i+1})$	[1,6; 1,8)	[1,8; 2)	[2; 2,2)	[2,2; 2,4)	[2,4; 2,6)	[2,6; 2,8)	[2,8; 3]
n_i	2	7	13	38	22	11	7

№ 3

$[x_i; x_{i+1})$	[2; 2,2)	[2,2; 2,4)	[2,4; 2,6)	[2,6; 2,8)	[2,8; 3)	[3; 3,2)	[3,2; 3,4]
n_i	7	12	21	35	13	8	4

№ 4

$[x_i; x_{i+1})$	[1,7; 2)	[2; 2,3)	[2,3; 2,6)	[2,6; 2,9)	[2,9; 3,2)	[3,2; 3,5)	[3,5; 3,8]
n_i	5	8	17	29	21	13	7

№ 5

$[x_i; x_{i+1})$	[1,5; 1,6)	[1,6; 1,7)	[1,7; 1,8)	[1,8; 1,9)	[1,9; 2)	[2; 2,1)	[2,1; 2,2]
n_i	6	9	18	29	20	12	6

№ 6

$[x_i; x_{i+1})$	[1,8; 2)	[2; 2,2)	[2,2; 2,4)	[2,4; 2,6)	[2,6; 2,8)	[2,8; 3)	[3; 3,2]
n_i	6	11	25	27	20	8	3

№ 7

$[x_i; x_{i+1})$	[2,1; 2,2)	[2,2; 2,3)	[2,3; 2,4)	[2,4; 2,5)	[2,5; 2,6)	[2,6; 2,7)	[2,7; 2,8]
n_i	6	13	17	27	19	10	8

№ 8

$[x_i; x_{i+1})$	[1,9; 2,1)	[2,1; 2,3)	[2,3; 2,5)	[2,5; 2,7)	[2,7; 2,9)	[2,9; 3,1)	[3,1; 3,3]
n_i	5	10	18	28	20	13	6

№ 9

$[x_i; x_{i+1})$	[1,3; 1,5)	[1,5; 1,7)	[1,7; 1,9)	[1,9; 2,1)	[2,1; 2,3)	[2,3; 2,5)	[2,5; 2,7]
n_i	5	9	18	30	20	12	6

№ 10

$[x_i; x_{i+1})$	[1,7; 1,9)	[1,9; 2,1)	[2,1; 2,3)	[2,3; 2,5)	[2,5; 2,7)	[2,7; 2,9)	[2,9; 3,1]
n_i	6	10	18	25	20	13	8

Tasks No. 11-20.

The results of the sample observation of the processing time of one part by the workers are given in the tables, in the first line of which there are partial time intervals (in min.), And in the second there is the number of workers, whose work time has fallen into the corresponding interval.

$[x_i; x_{i+1})$	[4; 4,4)	[4,4; 4,8)	[4,8; 5,2)	[5,2; 5,6)	[5,6; 6)	[6; 6,4)	[6,4; 6,8]
n_i	3	8	21	31	19	14	4

№ 12

$[x_i; x_{i+1})$	[5; 5,4)	[5,4; 5,8)	[5,8; 6,2)	[6,2; 6,6)	[6,6; 7)	[7; 7,4)	[7,4; 7,8]
n_i	2	6	10	35	20	10	7

№ 13

$[x_i; x_{i+1})$	[3,8; 4)	[4; 4,2)	[4,2; 4,4)	[4,4; 4,6)	[4,6; 4,8)	[4,8; 5)	[5; 5,2]
n_i	2	8	25	34	20	8	3

№ 14

$[x_i; x_{i+1})$	[6; 6,4)	[6,4; 6,8)	[6,8; 7,2)	[7,2; 7,6)	[7,6; 8)	[8; 8,4)	[8,4; 8,8]
n_i	2	7	20	35	19	12	5

№ 15

$[x_i; x_{i+1})$	[7; 7,2)	[7,2; 7,4)	[7,4; 7,6)	[7,6; 7,8)	[7,8; 8)	[8; 8,2)	[8,2; 8,4]
n_i	4	10	18	30	20	12	6

№ 16

$[x_i; x_{i+1})$	[4; 4,4)	[4,4; 4,8)	[4,8; 5,2)	[5,2; 5,6)	[5,6; 6)	[6; 6,4)	[6,4; 6,8]
n_i	6	12	17	33	20	10	2

№ 17

$[x_i; x_{i+1})$	[5,5; 5,9)	[5,9; 6,3)	[6,3; 6,7)	[6,7; 7,1)	[7,1; 7,5)	[7,5; 7,9)	[7,9; 8,3]
n_i	6	10	17	28	20	11	8

№ 18

$[x_i; x_{i+1})$	[6,5; 6,9)	[6,9; 7,3)	[7,3; 7,7)	[7,7; 8,1)	[8,1; 8,5)	[8,5; 8,9)	[8,9; 9,3]
n_i	3	10	20	32	16	12	7

№ 19

$[x_i; x_{i+1})$	[7,5; 7,9)	[7,9; 8,3)	[8,3; 8,7)	[8,7; 9,1)	[9,1; 9,5)	[9,5; 9,9)	[9,9; 10,3]
n_i	8	12	17	29	18	10	6

№ 20

$[x_i; x_{i+1})$	[5; 5,4)	[5,4; 5,8)	[5,8; 6,2)	[6,2; 6,6)	[6,6; 7)	[7; 7,4)	[7,4; 7,8]
n_i	4	10	18	33	17	12	6

Task №№ 21-30.

The study of the duration of work (in ths. Hours) of electric light bulbs is given in the tables.№ 21

$[x_i; x_{i+1})$	[2,1; 2,2)	[2,2; 2,3)	[2,3; 2,4)	[2,4; 2,5)	[2,5; 2,6)	[2,6; 2,7)	[2,7; 2,8]
n_i	2	8	22	40	12	10	6

№ 22

$[x_i; x_{i+1})$	[1,8; 2)	[2; 2,2)	[2,2; 2,4)	[2,4; 2,6)	[2,6; 2,8)	[2,8; 3)	[3; 3,2]
n_i	4	9	21	35	18	8	5

№ 23

$[x_i; x_{i+1})$	[1,6; 1,8)	[1,8; 2)	[2; 2,2)	[2,2; 2,4)	[2,4; 2,6)	[2,6; 2,8)	[2,8; 3]
n_i	2	8	15	35	20	12	8

№ 24

$[x_i; x_{i+1})$	[1,5; 1,7)	[1,7; 1,9)	[1,9; 2,1)	[2,1; 2,3)	[2,3; 2,5)	[2,5; 2,7)	[2,7; 2,9]
n_i	4	11	18	30	21	10	6

№ 25

$[x_i; x_{i+1})$	[2; 2,1)	[2,1; 2,2)	[2,2; 2,3)	[2,3; 2,4)	[2,4; 2,5)	[2,5; 2,6)	[2,6; 2,7]
n_i	5	10	17	32	20	12	4

№ 26

$[x_i; x_{i+1})$	[2,4; 2,6)	[2,6; 2,8)	[2,8; 3)	[3; 3,2)	[3,2; 3,4)	[3,4; 3,6)	[3,6; 3,8]
n_i	5	12	18	28	19	13	5

№ 27

$[x_i; x_{i+1})$	[2,3; 2,5)	[2,5; 2,7)	[2,7; 2,9)	[2,9; 3,1)	[3,1; 3,3)	[3,3; 3,5)	[3,5; 3,7]
n_i	3	9	15	33	21	13	9

№ 28

$[x_i; x_{i+1})$	[2; 2,2)	[2,2; 2,4)	[2,4; 2,6)	[2,6; 2,8)	[2,8; 3)	[3; 3,2)	[3,2; 3,4]
n_i	2	9	16	34	22	12	5

№ 29

$[x_i; x_{i+1})$	[1,4; 1,6)	[1,6; 1,8)	[1,8; 2)	[2; 2,2)	[2,2; 2,4)	[2,4; 2,6)	[2,6; 2,8]
n_i	4	12	18	33	19	11	3

№ 30

$[x_i; x_{i+1})$	[1,4; 1,6)	[1,6; 1,8)	[1,8; 2)	[2; 2,2)	[2,2; 2,4)	[2,4; 2,6)	[2,6; 2,8]
n_i	4	12	18	33	19	11	3

Задачі №№ 31-40.

Результати вибіркового вимірювання діаметрів валиків наведені у таблицях, перший рядок яких містить частинні інтервали діаметрів у мм, другий — число валиків, діаметри яких потрапили у відповідний інтервал.

№ 31

$[x_i; x_{i+1})$	[4,02; 4,04)	[4,04; 4,06)	[4,06; 4,08)	[4,08; 4,1)	[4,1; 4,12]
n_i	6	9	20	11	4

№ 32

$[x_i; x_{i+1})$	[2,06; 2,08)	[2,08; 2,1)	[2,1; 2,12)	[2,12; 2,14)	[2,14; 2,16]
n_i	5	8	22	11	4

№ 33

$[x_i; x_{i+1})$	[3,08; 3,1)	[3,1; 3,12)	[3,12; 3,14)	[3,14; 3,16)	[3,16; 3,18]
n_i	6	8	23	10	3

№ 34

$[x_i; x_{i+1})$	[3,12; 3,16)	[3,16; 3,2)	[3,2; 3,24)	[3,24; 3,28)	[3,28; 3,32]
n_i	4	7	25	9	5

№ 35

$[x_i; x_{i+1})$	[2,12; 2,16)	[2,16; 2,2)	[2,2; 2,24)	[2,24; 2,28)	[2,28; 2,32]
n_i	5	7	27	8	3

№ 36

$[x_i; x_{i+1})$	[3,28; 3,3)	[3,3; 3,32)	[3,32; 3,34)	[3,34; 3,36)	[3,36; 3,38]
n_i	2	9	28	10	1

№ 37

$[x_i; x_{i+1})$	[5,12; 5,16)	[5,16; 5,2)	[5,2; 5,24)	[5,24; 5,28)	[5,28; 5,32]
n_i	5	9	23	11	2

№ 38

$[x_i; x_{i+1})$	[4,92; 4,94)	[4,94; 4,96)	[4,96; 4,98)	[4,98; 5)	[5; 5,02]
n_i	5	9	21	11	4

№ 39

$[x_i; x_{i+1})$	[3,42; 3,46)	[3,46; 3,5)	[3,5; 3,54)	[3,54; 3,58)	[3,58; 3,62]
n_i	2	7	30	8	3

№ 40

$[x_i; x_{i+1})$	[4,56; 4,6)	[4,6; 4,64)	[4,64; 4,68)	[4,68; 4,72)	[4,72; 4,76]
n_i	5	11	24	9	1

Task №№ 41-50.

The results of tests on the strength of steel wires of equal length are given in the tables, the first row of which contains partial intervals of bursting forces (kg / mm2), the second - the number of wires whose bursting forces have fallen into an appropriate interval.

№ 41

$[x_i; x_{i+1})$	[30; 32)	[32; 34)	[34; 36)	[36; 38)	[38; 40)	[40; 42)	[42; 44]
n_i	6	11	18	29	19	12	5

№ 42

$[x_i; x_{i+1})$	[35; 38)	[38; 41)	[41; 44)	[44; 47)	[47; 50)	[50; 53)	[53; 56]
n_i	4	9	22	31	16	11	7

№ 43

$[x_i; x_{i+1})$	[27; 28,5)	[28,5; 30)	[30; 31,5)	[31,5; 33)	[33; 34,5)	[34,5; 36)	[36; 37,5]
n_i	3	12	24	35	18	6	2

№ 44

$[x_i; x_{i+1})$	[51; 55)	[55; 59)	[59; 63)	[63; 67)	[67; 71)	[71; 75)	[75; 79]
n_i	7	14	17	22	20	12	8

№ 45

$[x_i; x_{i+1})$	[42; 45,5)	[45,5; 49)	[49; 52,5)	[52,5; 56)	[56; 59,5)	[59,5; 63)	[63; 66,5]
n_i	2	14	22	34	18	15	5

№ 46

$[x_i; x_{i+1})$	[46; 49)	[49; 52)	[52; 55)	[55; 58)	[58; 61)	[61; 64)	[64; 67]
n_i	4	13	21	21	23	12	6

№ 47

$[x_i; x_{i+1})$	[26; 28,5)	[28,5; 31)	[31; 33,5)	[33,5; 36)	[36; 38,5)	[38,5; 41)	[41; 43,5]
n_i	3	8	17	44	18	6	4

№ 48

$[x_i; x_{i+1})$	[65; 69)	[69; 73)	[73; 77)	[77; 81)	[81; 85)	[85; 89)	[89; 93]
n_i	6	11	22	31	14	12	4

№ 49

$[x_i; x_{i+1})$	[40; 42)	[42; 44)	[44; 46)	[46; 48)	[48; 50)	[50; 52)	[52; 54]
n_i	7	13	19	24	21	10	6

№ 50

$[x_i; x_{i+1})$	[31; 33)	[33; 35)	[35; 37)	[37; 39)	[39; 41)	[41; 43)	[43; 45]
n_i	3	9	24	28	22	10	4

Examples solution of the task:

Task 1.In studying the question of production quotas weavers observed the next frequency yarn breakages similar loom stalls at various intervals of equal length *t*:
7, 1, 2, 1, 2, 5, 4, 4, 3, 2, 2, 6, 0, 1, 6,
5, 3, 2, 0, 1, 4, 3, 2, 1, 5, 3, 0, 4, 2, 3.
1) To make the statistical distribution of frequencies and relative frequencies among the cliffs yarn stalls;
2) build a landfill frequencies and relative frequencies;
3) find empirical distribution function and build its schedule;
4) Calculate the sample, average, variance, standard deviation, mode, median, range of variation, coefficient of variation.

1) The sample size n = 30. This version number is written as a series of variations:

0, 0, 0, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 6, 6, 7.

Variants: $x_1 = 0$, $x_2 = 1$, $x_3 = 2$, $x_4 = 3$, $x_5 = 4$, $x_6 = 5$, $x_7 = 6$, $x_8 = 7$.
Frequency: $n_1 = 3$, $n_2 = 5$, $n_3 = 7$, $n_4 = 5$, $n_5 = 4$, $n_6 = 3$, $n_7 = 2$, $n_8 = 1$.
As a result, we obtain the statistical distribution of frequencies:

x_i	0	1	2	3	4	5	6	7
n_i	3	5	7	5	4	3	2	1

Checking: $\sum n_i = 3 + 5 + 7 + 5 + 4 + 3 + 2 + 1 = 30 = n$.

According to the formula $w_i = n_i / n$ consistently calculate the relative frequencies: $w_1 = 3/30$, $w_2 = 5/30$, $w_3 = 7/30$, $w_4 = 5/30$, $w_5 = 4/30$, $w_6 = 3/30$, $w_7 = 2/30$, $w_8 = 1/30$. Control: $\sum w_i =$ 3/30 + 5/30 + 7/30 + + 5/30 + 4/30 + 3/30 + 2/30 + 1/30= 1. Thus, the statistical distribution of relative frequencies among the cliffs yarn has the stalls type:

x_i	0	1	2	3	4	5	6	7
w_i	3/30	5/30	7/30	5/30	4/30	3/30	2/30	1/30

2) Polygon of frequencies depicted in Fig.1.1 and polygon of relative frequencies — Fig.1.2.

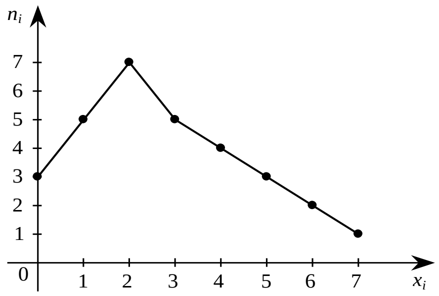


Fig.1.1.

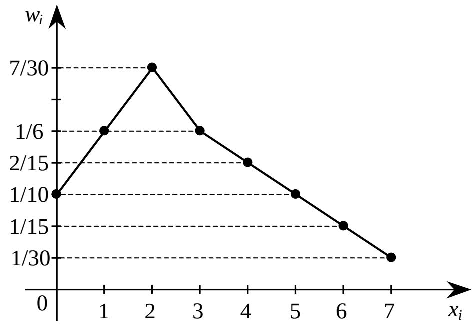


Fig.1.2.

3) The sample size $n = 30$. If $x \le 0$, then there is no options, lesser x, $n_x = 0$, and therefore $F^*(x) = n_x/30 = 0$.

Let $x \in (0; 1]$. Then variant $x = 0$ is less than x, so $n_x = 3$ and $F^*(x) = 3/30 = 0{,}1$.

If x satisfies the double inequality $1 < x \le 2$, while less than x are variants 0 and 1, the amount of which frequencies $n_x = 3 + 5 = 8$. So $F^*(x) = 8/30 = 4/15$ for $x \in (1; 2]$.

If x is performed by the double inequality $2 < x \le 3$, then variants 0, 1, 2 are lesser x, sum frequency which $n_x = 3 + 5 + 7 = 15$. Therefore, for $x \in (2; 3]$ $F^*(x) = 15/30 = 0{,}5$.

Similarly, we find the value of $F^*(x)$ for intervals (3; 4], (4; 5], (5; 6], (6; 7], (7; ∞]. As a result, we obtain the required empirical distribution function:

$$F^*(x) = \begin{cases} 0, & \text{якщо } x \le 0, \\ 1/10, & \text{якщо } 0 < x \le 1, \\ 4/15, & \text{якщо } 1 < x \le 2, \\ 1/2, & \text{якщо } 2 < x \le 3, \\ 2/3, & \text{якщо } 3 < x \le 4, \\ 4/5, & \text{якщо } 4 < x \le 5, \\ 9/10, & \text{якщо } 5 < x \le 6, \\ 29/30, & \text{якщо } 6 < x \le 7, \\ 1, & \text{якщо } x > 7. \end{cases}$$

Graph this function is shown in Fig. 1.3.

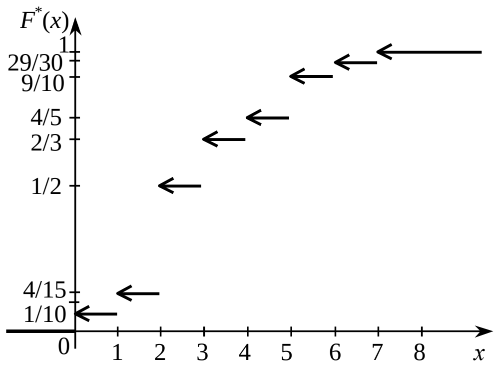


Рис. 1.3.

2) To calculate $\overline{x}_в$ and $D_в$ use the formulas (1.7) i (1.12):

$$\overline{x}_в = \frac{\sum_{i=1}^{k} x_i n_i}{n} = \frac{0\cdot3+1\cdot5+2\cdot7+3\cdot5+4\cdot4+5\cdot3+6\cdot2+7\cdot1}{30} = 84/30 = 2,8;$$

$$D_в = \overline{x^2} - (\overline{x}_в)^2 = \frac{\sum_{i=1}^{k} x_i^2 n_i}{n} - (\overline{x}_в)^2 =$$

$$= \frac{0^2\cdot3+1^2\cdot5+2^2\cdot7+3^2\cdot5+4^2\cdot4+5^2\cdot3+6^2\cdot2+7^2\cdot1}{30} - (2,8)^2 =$$

$= 338/30 - 7,84 = 3,4267;$

$$\sigma_в = \sqrt{D_в} = \sqrt{3,4267} = 1,8511.$$

Conclusion: The average number of yarn breakages stalls during the time interval is 2.8, and the average spread of numbers breaks stalls average of 2.8 is 1.8511.

Mode $Mo^* = 2,$ since variant 2 corresponds to the highest frequency of 7.

Median Me^* can be find, using variation row obtained in 1), or directly from the statistical distribution. Sum frequency first three variants of this distribution is 15 (half-volume of sample) and the next five - also 15. So the median is between variants 2 and 3: $Me^* = (2+3)/2 = 2,5.$ A mismatch $\overline{x}_в$, Mo^* and Me^* indicates the absence of strict symmetry of the distribution.

The coefficient of variation (according to the formula (1.14))

$$V = \frac{\sigma_в}{\overline{x}_в}\cdot100\% = \frac{1,8511}{2,8}\cdot100\% = 66,11\%.$$

Task 2. During the period between regular readjustment equipment made control measurements of thickness (in mm) 200 liners connecting rod bearings. The data are presented in Table. 1.2.

1) To make the interval statistical distribution of frequencies and relative frequencies of the sample. Based on the statistical distribution of the resulting interval:
2) build a histogram of frequencies and relative frequencies;
3) find empirical distribution function and build its schedule;
4) to calculate $\overline{x}_в, D_в, \sigma_в$, the median and mode. *Table. 1.2*

1,754	1,739	1,743	1,764	1,733	1,736	1,743	1,742
1,728	1,732	1,731	1,752	1,737	1,747	1,758	1,737
1,724	1,737	1,733	1,713	1,740	1,740	1,729	1,740
1,751	1,739	1,747	1,748	1,730	1,750	1,740	1,732
1,740	1,730	1,748	1,757	1,741	1,733	1,743	1,745
1,748	1,723	1,737	1,748	1,741	1,751	1,714	1,750
1,744	1,748	1,758	1,756	1,727	1,731	1,738	1,753
1,735	1,738	1,743	1,729	1,743	1,737	1,731	1,734
1,741	1,742	1,744	1,756	1,744	1,752	1,739	1,740
1,729	1,745	1,742	1,753	1,743	1,734	1,731	1,734
1,732	1,732	1,746	1,748	1,755	1,738	1,742	1,729

1,731	1,725	1,729	1,745	1,739	1,754	1,752	1,720
1,750	1,734	1,749	1,738	1,747	1,757	1,751	1,746
1,723	1,736	1,746	1,744	1,759	1,728	1,751	1,750
1,746	1,759	1,748	1,740	1,735	1,745	1,740	1,746
1,737	1,726	1,743	1,755	1,740	1,726	1,745	1,744
1,735	1,746	1,739	1,732	1,758	1,744	1,754	1,724
1,742	1,750	1,761	1,758	1,753	1,757	1,720	1,733
1,738	1,728	1,758	1,732	1,763	1,733	1,745	1,766
1,745	1,743	1,734	1,733	1,755	1,756	1,769	1,750
1,740	1,762	1,738	1,742	1,740	1,740	1,760	1,752
1,746	1,728	1,743	1,718	1,738	1,762	1,728	1,734
1,753	1,751	1,748	1,735	1,739	1,729	1,754	1,736
1,762	1,748	1,738	1,726	1,757	1,738	1,726	1,720
1,751	1,734	1,724	1,741	1,752	1,732	1,738	1,739

○ 1). All variants are in the range [1,713; 1.769] length of 0,056 mm. We divide it into 8 equal the length of intervals:
[1,713; 1,720) [1,720; 1.727) [1.727; 1,734) [1,734; 1.741)
[1.741; 1.748) [1.748; 1.755) [1.755; 1.762) [1.762; 1.769].
Each version of the table. 1.2 assign one of these partial intervals and find sum of option, enter the first, second, ..., eighth intervals. The result:

$n_1 = 3$, $n_2 = 13$, $n_3 = 32$, $n_4 = 49$, $n_5 = 42$, $n_6 = 35$, $n_7 = 19$, $n_8 = 7$, $n = \sum_{i=1}^{8} n_i = 200$.

The result is a statistical interval frequency distribution is shown in Table 1.3.

Table 1.3

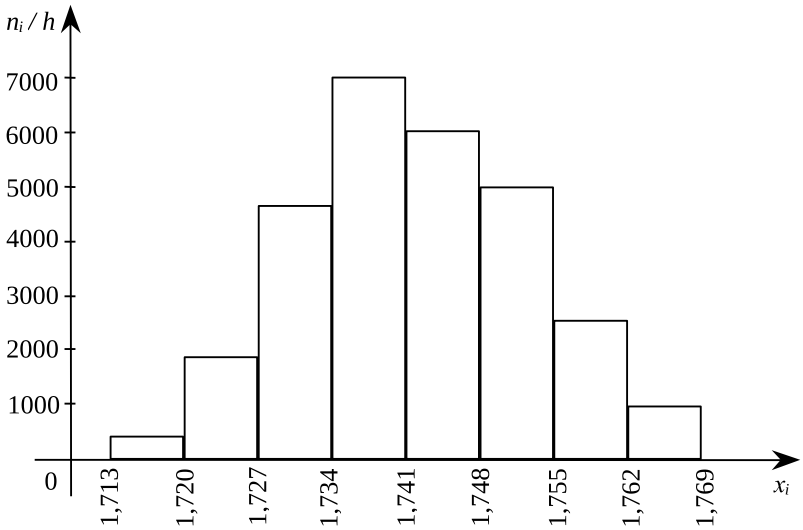
$[x_i; x_{i+1})$	n_i	$[x_i; x_{i+1})$	n_i
[1,713; 1,720)	3	[1,741; 1,748)	42
[1,720; 1,727)	13	[1,748; 1,755)	35
[1,727; 1,734)	32	[1,755; 1,762)	19
[1,734; 1,741)	49	[1,762; 1,769]	7

In view of the fact that the sample size n = 200, we write the inter-abundant statistical distribution of relative frequencies of the sample (Table. 1.4):

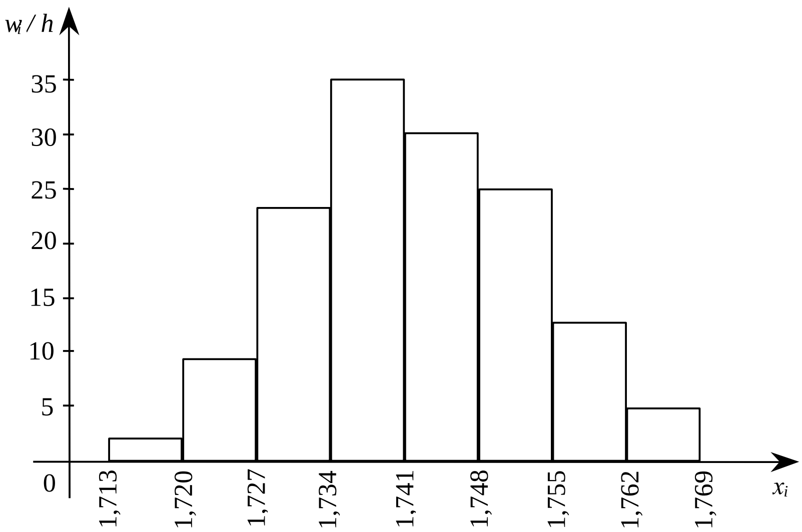
Table 1.4

$[x_i; x_{i+1})$	w_i	$[x_i; x_{i+1})$	w_i
[1,713; 1,720)	3/200	[1,741; 1,748)	42/200
[1,720; 1,727)	13/200	[1,748; 1,755)	35/200
[1,727; 1,734)	32/200	[1,755; 1,762)	19/200
[1,734; 1,741)	49/200	[1,762; 1,769]	7/200

2). The histogram of the frequency distribution h = 0,007 depicted in rys.1.4 and relative frequency histogram for rys.1.5.



Rys.1.4.



Rys.1.5.

3). The study quantitative trait X is a continuous random variable. Therefore, the probability distribution function $F(x)$, and empirical function $F^*(x)$ are continuous functions determined arhument x.
Let $x \le 1{,}713$. Then $n_x = 0$, as observed values quantitative features smaller than x, are absent. Consequently, $F^*(x) = 0$ for all x □ 1,713.

Find the value of the empirical distribution function for $x = 1{,}720$ - left end of the second interval. Thus, we assume that we can not do this for each internal point of the first interval. When $x = 1{,}720$ $n_x = 3$ i $F^*(1{,}720) = 3/200 = 0{,}015$.

For $x = 1{,}727$ $n_x = 3 + 13 = 16$, $F^*(1{,}727) = 16/200 = 0{,}08$.

For $x = 1{,}734$ $n_x = 16 + 32 = 48$, $F^*(1{,}734) = 48/200 = 0{,}24$.

Similarly, we find:

$F^*(1{,}741) = (48 + 49)/200 = 97/200 = 0{,}485$;

$F^*(1{,}748) = (97 + 42)/200 = 139/200 = 0{,}695$;

$F^*(1{,}755) = (139 + 35)/200 = 174/200 = 0{,}87$;

$F^*(1{,}762) = (174 + 19)/200 = 193/200 = 0{,}965$.

Finally, for $x > 1{,}769$ all 200 observed quantitative trait values smaller than x, i.e. $n_x = 200$.. So $F^*(x) = 1$ for $x > 1{,}769$.

Construct a graph of the empirical distribution function: first for intervals $(-\infty; 1{,}713]$ and

(1,769; ∞), then in these locations. In order to show the continuity of change $F^*(x)$, obtained neighboring points connect straight segments (Fig. 1.6).

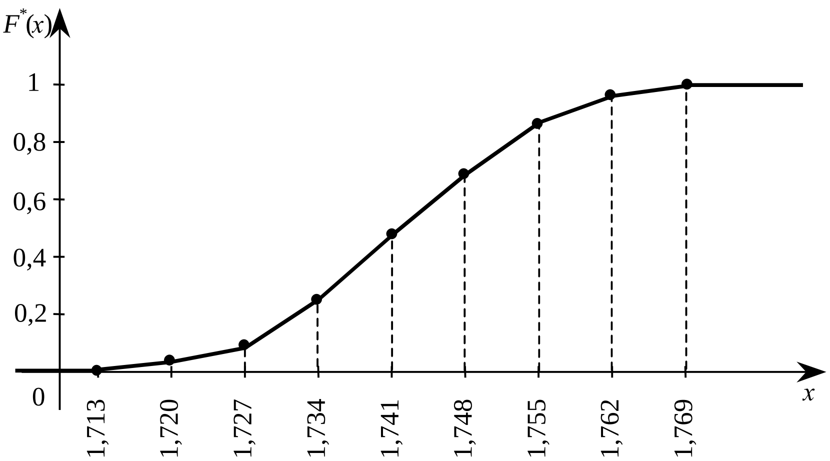


Fig. 1.6.

4). To find the numerical characteristics of the sample interval specified statistical distribution of claim 1, move on to the discrete distribution:

x_i	1,7165	1,7235	1,7305	1,7375	1,7445	1,7515	1,7585	1,7655
n_i	3	13	32	49	42	35	19	7

This "new" options are midpoints partial intervals.

$$\overline{x}_в = \frac{\sum_{i=1}^{k} x_i n_i}{n} =$$

$$= \frac{1,7165\cdot 3 + 1,7235\cdot 13 + 1,7305\cdot 32 + 1,7375\cdot 49 + 1,7445\cdot 42 + 1,7515\cdot 35 + 1,7585\cdot 19 + 1,7655\cdot 7}{200} = = 1,7417;$$

$$D_в = \overline{x^2} - (\overline{x}_в)^2 = \frac{\sum_{i=1}^{k} x_i^2 n_i}{n} - (\overline{x}_в)^2 =$$

$$= \frac{(1,7165)^2\cdot 3 + (1,7235)^2\cdot 13 + (1,7305)^2\cdot 32 + (1,7375)^2\cdot 49}{200} +$$

$$+ \frac{(1,7445)^2\cdot 42 + (1,7515)^2\cdot 35 + (1,7585)^2\cdot 19 + (1,7655)^2\cdot 7}{200} - (1,7417)^2 =$$

$$= 0,0001283;$$

$$\sigma_в = \sqrt{D_в} = \sqrt{0,0001283} = 0,0113269.$$

Since all partial distribution of the intervals have the same length, it is a modal interval [1,734; 1.741), which corresponds to the highest frequency 49. The value Mo^* contained within this interval is calculated by the formula (1.10).

$$Mo^* = x_m + \frac{n_m - n_{m-1}}{2n_m - n_{m-1} - n_{m+1}} h = 1,734 + \frac{49-32}{2\cdot 49 - 32 - 42}\cdot 0,007 =$$

$$= 1,734 + \frac{17}{24}\cdot 0,007 \approx 1,73896.$$

To calculate the median find the first partial median interval $[x_m; x_{m+1})$, that performs inequality:

$$F^*(x_m) < 0,5, \qquad F^*(x_{m+1}) > 0,5.$$

According to 3) $F^*(1,741)=0,485<0,5$, $F^*(1,748)=0,695>0,5$.

Therefore, $x_m = 1,741$, $x_{m+1} = 1,748$. According to the formula (1.9)

$$Mo^* = x_m + \frac{0,5 - F^*(x_m)}{F^*(x_{m+1}) - F^*(x_m)}(x_{m+1} - x_m) =$$
$$= 1,741 + \frac{0,5 - 0,485}{0,695 - 0,485}(1,748 - 1,741) =$$
$$= 1,741 + \frac{0,015}{0,21} \cdot 0,007 = 1,741 + 0,0005 = 1,7415.$$

The coefficient of variation calculated using the formula (1.14):

$$V = \frac{\sigma_{\hat{a}}}{\bar{x}_{\hat{a}}} \cdot 100\% = \frac{0,0113269}{1,7417} \cdot 100\% = 0,65\%.$$

Task 3. The test results on the strength of the steel wires of equal length are shown in Table. 2.1 interval distribution.

Таблиця 2.1

Breaking strength (kg / mm	[30; 32)	[32; 34)	[34; 36)	[36; 38)	[38; 40)	[40; 42)	[42; 44]
Number of wires	6	12	17	29	19	13	4

1) Find a confidence interval that cover with reliability γ = 0,95 mean breaking the whole party 4000 wires.
2) Find a confidence probability that the sample mean depart trom general the average in absolute value no more than 0.5 kg / mm2.

○

1) Sample is nosmal (n = 100 > 30), sample is unrepeating, because after checking facility (wire) can not be restore a bearing in the general population (wire terminated due to checking its strength). To find the limits of the confidence interval limit error Δ find the formula (2.32), previously know sample numeric characteristics $\bar{x}_в$ and $D_в$.

$x_i = \frac{x_k + x_{k+1}}{2}$	31	33	35	37	39	41	43
n_i	6	12	17	29	19	13	4

$$\bar{x}_в = \frac{\sum_{i=1}^{k} x_i n_i}{n} =$$
$$= \frac{31 \cdot 6 + 33 \cdot 12 + 35 \cdot 17 + 37 \cdot 29 + 39 \cdot 19 + 41 \cdot 13 + 43 \cdot 4}{100} = 36,96;$$
$$D_в = \overline{x^2} - (\bar{x}_в)^2 = \frac{\sum_{i=1}^{k} x_i^2 n_i}{n} - (\bar{x}_в)^2 =$$
$$= \frac{31^2 \cdot 6 + 33^2 \cdot 12 + 35^2 \cdot 17 + 37^2 \cdot 29 + 39^2 \cdot 19 + 41^2 \cdot 13 + 43^2 \cdot 4}{100} - (36,96)^2 =$$
= 9,0384

For the function tables of Laplace (tab. 3 applications) find the value of t using the equation $\Phi(t) = \gamma/2 = 0,95/2 = 0,475$, t = 1,96. Formula (2.32), where N = 4000,

$$\Delta = 1,96\sqrt{\frac{9,0384}{100}\left(1 - \frac{100}{4000}\right)} \approx 0,5818,$$

then left limit of the confidence interval

$$\bar{x}_{\text{в}} - \Delta = 36{,}96 - 0{,}5818 = 36{,}3782,$$
$$\bar{x}_{\text{в}} + \Delta = 36{,}96 + 0{,}5818 = 37{,}5418.$$

According to the table of Laplas (table 3) find the value t from equation $\Phi(t) = \gamma/2 = 0{,}95/2 = 0{,}475$, t = 1,96. N = 4000,

$$\Delta = 1{,}96\sqrt{\frac{9{,}0384}{100}\left(1 - \frac{100}{4000}\right)} \approx 0{,}5818,$$

Then left boundary of interal

$$\bar{x}_{\text{в}} - \Delta = 36{,}96 - 0{,}5818 = 36{,}3782,$$

right —

$$\bar{x}_{\text{в}} + \Delta = 36{,}96 + 0{,}5818 = 37{,}5418.$$

Finally, the desired confidence interval has the form (36.3782, 37.5418).

2) Quest is the probability $P(|\bar{x}_{\text{в}} - \bar{x}_{\text{г}}| \le 0{,}5)$, to be found by the formula

$$P(\,|\bar{x}_{\text{в}} - \bar{x}_{\text{г}}| < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\bar{\sigma}}\right),$$

as $\bar{x}_{\text{в}}$ is normally distributed random variable (according to Theorem 2.2), $M(\bar{x}_{\text{в}}) = \bar{x}_{\text{г}}$. Accordingly formula (2.25)

$$\bar{\sigma} = \bar{\sigma}'_{\bar{x}} = \sqrt{\frac{D_{\text{в}}}{n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{9{,}0384}{100}\left(1 - \frac{100}{4000}\right)} \approx 0{,}2968.$$

Then $\varepsilon/\bar{\sigma} = 0{,}5/0{,}2968 = 1{,}69$ and accordingly table. 3 application

$$P(|\bar{x}_{\text{в}} - \bar{x}_{\text{г}}| \le 0{,}5) = 2\Phi(0{,}5/0{,}2968) =$$
$$= 2\Phi(1{,}69) = 2 \cdot 0{,}45449 = 0{,}90898.$$