PhD Misiura Ie.Iu. (доцент Місюра Є.Ю.)

Theme: Problems of mathematical statistics. Primary processing of statistical data. Statistical estimates of distribution parameters. Statistical evaluation methods in international trade

PART 2. Checking the statistical hypothesis

BASIC POINTS

- **to construct a continuous statistical series**
- **to calculate the numerical characteristics of a continuous statistical series**
- **to make the assumption about the distribution law (the hypothesis)**
- **to check the hypothesis relative to the distribution law**

The sample is given by:

99	145	127	128	114	118
103	155	135	129	125	106
112	95	143	131	108	111
121	101	96	99	113	113
132	115	102	107	98	139

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.1 Define the maximum and minimum values of this sample:

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.1 Define the maximum and minimum values of this sample:

$$x_{\max} = 155 \qquad x_{\min} = 95$$

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.1 Define the maximum and minimum values of this sample:

$$x_{\max} = 155 \qquad x_{\min} = 95$$

1.2 Find the range:

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.1 Define the maximum and minimum values of this sample:

$$x_{\max} = 155 \qquad x_{\min} = 95$$

1.2 Find the range:

$$R = x_{\max} - x_{\min} = 155 - 95 = 60$$

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.1 Define the maximum and minimum values of this sample:

$$x_{\max} = 155 \qquad x_{\min} = 95$$

1.2 Find the range:

$$R = x_{\max} - x_{\min} = 155 - 95 = 60$$

1.3 Find the sample size:

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.1 Define the maximum and minimum values of this sample:

$$x_{max} = 155 \qquad x_{min} = 95$$

1.2 Find the range:

$$R = x_{max} - x_{min} = 155 - 95 = 60$$

1.3 Find the sample size:

$$n = 30$$

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.4 Calculation of a number of intervals (this number is a whole number or an integer):

$$k = 1 + 3.322 \cdot \lg n = 1 + 3.322 \cdot 1{,}477 = 5{,}907 \approx 6$$

The obtained number $k = 5{,}907$ we approximate until an integer $k \approx 6$

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.4 Calculation of a number of intervals (this number is a whole number or an integer):

$$k = 1 + 3.322 \cdot \lg n = 1 + 3.322 \cdot 1{,}477 = 5{,}907 \approx 6$$

The obtained number $k = 5{,}907$ we approximate until a whole number $k \approx 6$

1.5 Calculate a length of each interval:

$$h = \frac{R}{k} = \frac{60}{6} = 10$$

Let’s construct the continuous statistic series.

1.6 Find $x_{beginning}$ and x_{end} by the following formulas:

$$x_{beginning} = x_{\min} - \frac{h}{2}$$

$$x_{end} = x_{\max} + \frac{h}{2}$$

Let's construct the continuous statistic series.

1.6 Find $x_{beginning}$ and x_{end} by the following formulas:

$$x_{beginning} = x_{\min} - \frac{h}{2}$$

$$x_{end} = x_{\max} + \frac{h}{2}$$

$$x_{beginning} = 95 - \frac{10}{2} = 95 - 5 = 90$$

$$x_{end} = 155 + \frac{10}{2} = 155 + 5 = 160$$

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.7 Let's form the intervals $[x_i, x_{i+1})$, where $x_{i+1} = x_i + h$ (form the 1-st and 2-nd columns).

Here $x_1 = x_{beginnig}$ $x_2 = x_1 + h$

$x_3 = x_2 + h$ $x_{i+1} = x_i + h$

.

$x_{last} = x_{end}$

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.7 Let's form the intervals $[x_i, x_{i+1})$, where $x_{i+1} = x_i + h$ (form the 1-st and 2-nd columns).

Here $x_1 = x_{beginnig} = 90$

$$x_2 = x_1 + h = 90 + 10 = 100$$

$$x_3 = x_2 + h = 100 + 10 = 110$$

$$x_4 = x_3 + h = 110 + 10 = 120$$

$$x_5 = x_4 + h = 120 + 10 = 130$$

$$x_6 = x_5 + h = 130 + 10 = 140$$

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

Here

$$x_1 = x_{beginnig} = 90$$

$$x_2 = x_1 + h = 90 + 10 = 100$$

$$x_3 = x_2 + h = 100 + 10 = 110$$

$$x_4 = x_3 + h = 110 + 10 = 120$$

$$x_5 = x_4 + h = 120 + 10 = 130$$

$$x_6 = x_5 + h = 130 + 10 = 140$$

$$x_7 = x_6 + h = 140 + 10 = 150$$

$$x_8 = x_7 + h = 150 + 10 = 160$$

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

№	$[x_i, x_{i+1})$				
1	$[90,100)$				
2	$[100,110)$				
3	$[110,120)$				
4	$[120,130)$				
5	$[130,140)$				
6	$[140,150)$				
7	$[150,160)$				
	sum				

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.8 Find the frequency m_i of data belonging to a given interval $[x_i, x_{i+1})$

.

.

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.8 Find the frequency m_i of data belonging to a given interval $[x_i, x_{i+1})$

99	145	127	128	114	118
103	155	135	129	125	106
112	95	143	131	108	111
121	101	96	99	113	113
132	115	102	107	98	139

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

№	$[x_i, x_{i+1})$	m_i			
1	$[90,100)$	5			
2	$[100,110)$	6			
3	$[110,120)$	7			
4	$[120,130)$	5			
5	$[130,140)$	4			
6	$[140,150)$	2			
7	$[150,160)$	1			
	sum				

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

№	$[x_i, x_{i+1})$	m_i			
1	$[90,100)$	5			
2	$[100,110)$	6			
3	$[110,120)$	7			
4	$[120,130)$	5			
5	$[130,140)$	4			
6	$[140,150)$	2			
7	$[150,160)$	1			
	sum	**30**			

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.9 Calculate $\overline{x}_i$ for each interval $[x_i, x_{i+1})$ (it is **the midpoint** of the corresponding interval) by the formula:

$$\overline{x}_i = \frac{x_i + x_{i+1}}{2}$$

For example,

$$[90,100)$$

.

$$\overline{x}_1 = \frac{x_1 + x_2}{2} = \frac{90 + 100}{2} = 95$$

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

№	$[x_i, x_{i+1})$	m_i	$\overline{x}_i$		
1	$[90,100)$	5	95		
2	$[100,110)$	6	105		
3	$[110,120)$	7	115		
4	$[120,130)$	5	125		
5	$[130,140)$	4	135		
6	$[140,150)$	2	145		
7	$[150,160)$	1	155		
	sum	**30**			

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

1.10 Find the relative frequencies $w_i = \frac{m_i}{n}$ for each interval $[x_i, x_{i+1})$

Then we check the condition: $\sum_{i=1}^{k} w_i = 1$

For example, .

$$w_1 = \frac{m_1}{n} = \frac{5}{30}$$

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

№	$[x_i, x_{i+1})$	m_i	$\bar{x}_i$	$w_i = \frac{m_i}{n}$	
1	$[90,100)$	5	95	5/30	
2	$[100,110)$	6	105	6/30	
3	$[110,120)$	7	115	7/30	
4	$[120,130)$	5	125	5/30	
5	$[130,140)$	4	135	4/30	
6	$[140,150)$	2	145	2/30	
7	$[150,160)$	1	155	1/30	
	sum	**30**			

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

№	$[x_i, x_{i+1})$	m_i	$\bar{x}_i$	$w_i = \frac{m_i}{n}$	
1	$[90,100)$	5	95	5/30	
2	$[100,110)$	6	105	6/30	
3	$[110,120)$	7	115	7/30	
4	$[120,130)$	5	125	5/30	
5	$[130,140)$	4	135	4/30	
6	$[140,150)$	2	145	2/30	
7	$[150,160)$	1	155	1/30	
	sum	**30**		**1**	

2. CALCULATION OF THE NUMERICAL CHARACTERISTICS OF A CONTINUOUS STATISTICAL SERIES

2.1. Mean

$$\overline{x} = \sum_{i=1}^{k} \overline{x}_i w_i$$

For this example we have:

$$\overline{x} = \sum_{i=1}^{7} \overline{x_i} \cdot w_i = \overline{x_1} \cdot w_1 + \overline{x_2} \cdot w_2 + \overline{x_3} \cdot w_3 + \overline{x_4} \cdot w_4 + \overline{x_5} \cdot w_5 + \overline{x_6} \cdot w_6 + \overline{x_7} \cdot w_7$$

$$\overline{x} = 95 \cdot \frac{5}{30} + 105 \cdot \frac{6}{30} + 115 \cdot \frac{7}{30} + 125 \cdot \frac{5}{30} + 135 \cdot \frac{4}{30} + 145 \cdot \frac{2}{30} + 155 \cdot \frac{1}{30} = \frac{3520}{30} \approx 117{,}33$$

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

№	$[x_i, x_{i+1})$	$\overline{x}_i$	$w_i = \frac{m_i}{n}$	$\overline{x_i w_i}$	
1	$[90,100)$	95	5/30	15,83	
2	$[100,110)$	105	6/30	21	
3	$[110,120)$	115	7/30	26,83	
4	$[120,130)$	125	5/30	20,83	
5	$[130,140)$	135	4/30	18	
6	$[140,150)$	145	. 2/30	9,67	
7	$[150,160)$	155	1/30	5,17	
	sum		**1**	**117,33**	

2. CALCULATION OF THE NUMERICAL CHARACTERISTICS OF A CONTINUOUS STATISTICAL SERIES

2.2. Variance

$$S_x^2 = \sum_{i=1}^{k} \overline{x_i^2} w_i - \left(\overline{x}\right)^2 = \overline{x^2} - \left(\overline{x}\right)^2$$

For this example we have:

$$\overline{x^2} = \sum_{i=1}^{7} \overline{x_i^2} \cdot w_i = \overline{x_1^2} \cdot w_1 + \overline{x_2^2} \cdot w_2 + \overline{x_3^2} \cdot w_3 + \overline{x_4^2} \cdot w_4 + \overline{x_5^2} \cdot w_5 + \overline{x_6^2} \cdot w_6 + \overline{x_7^2} \cdot w_7$$

$$\overline{x^2} = 95^2 \cdot \frac{5}{30} + 105^2 \cdot \frac{6}{30} + 115^2 \cdot \frac{7}{30} + 125^2 \cdot \frac{5}{30} + 135^2 \cdot \frac{4}{30} + 145^2 \cdot \frac{2}{30} + 155^2 \cdot \frac{1}{30} = \frac{420950}{30} \approx 14031{,}67$$

$$S_x^2 = \overline{x^2} - \left(\overline{x}\right)^2 = 14031{,}67 - 117{,}33^2 = 265{,}338$$

1. CONSTRUCTION OF A CONTINUOUS STATISTICAL SERIES

№	$[x_i, x_{i+1})$	$\overline{x}_i$	$w_i = \frac{m_i}{n}$	$\overline{x_i w_i}$	$\overline{x_i^2 w_i}$
1	$[90,100)$	95	5/30	15,83	1504,167
2	$[100,110)$	105	6/30	21	2205
3	$[110,120)$	115	7/30	26,83	3085,833
4	$[120,130)$	125	5/30	20,83	2604,167
5	$[130,140)$	135	4/30	18	2430
6	$[140,150)$	145	2/30	9,67	1401,667
7	$[150,160)$	155	1/30	5,17	800,8333
	sum		**1**	**117,33**	**14031,67**

2. CALCULATION OF THE NUMERICAL CHARACTERISTICS OF A CONTINUOUS STATISTICAL SERIES

2.2. Variance

$$S_x^2 = \sum_{i=1}^{k} \overline{x_i^2} w_i - \left(\overline{x}\right)^2 = \overline{x^2} - \left(\overline{x}\right)^2$$

For this example we have:

$$S_x^2 = \overline{x^2} - \left(\overline{x}\right)^2 = 14031{,}67 - 117{,}33^2 = 265{,}338$$

2. CALCULATION OF THE NUMERICAL CHARACTERISTICS OF A CONTINUOUS STATISTICAL SERIES

2.3. Root-mean-square deviation

$$S_x = \sqrt{S_x^2}$$

For this example we have:

$$S_x = \sqrt{S_x^2} = \sqrt{265{,}338} \approx 16{,}29$$

3. THE ASSUMPTION ABOUT THE DISTRIBUTION LAW (THE HYPOTHESIS)

Let’s consider the ratio $\frac{w_i}{h}$ for each interval $[x_i, x_{i+1})$ and and plot the histogram.

For example,

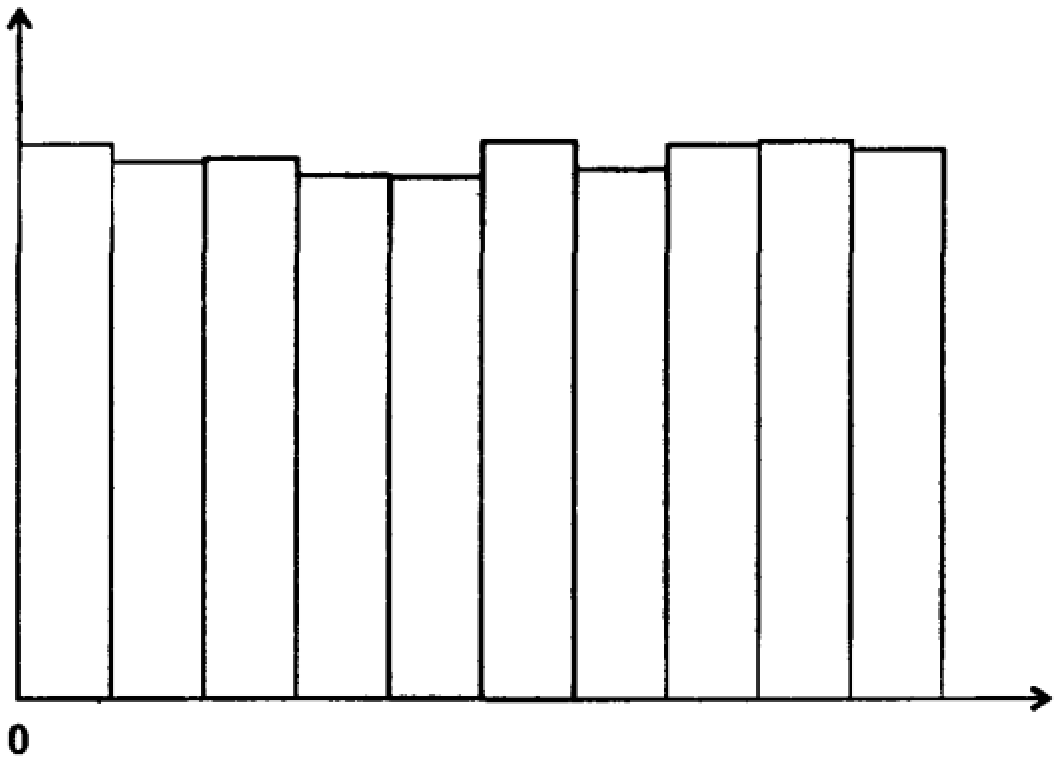
$$\frac{w_1}{h} = \frac{5}{300} \approx 0,017$$

A graphical representation of a continuous statistical series is called *a histogram*.

Here $\frac{w_i}{h}$ is the height of the corresponding rectangle on the interval $[x_i, x_{i+1})$

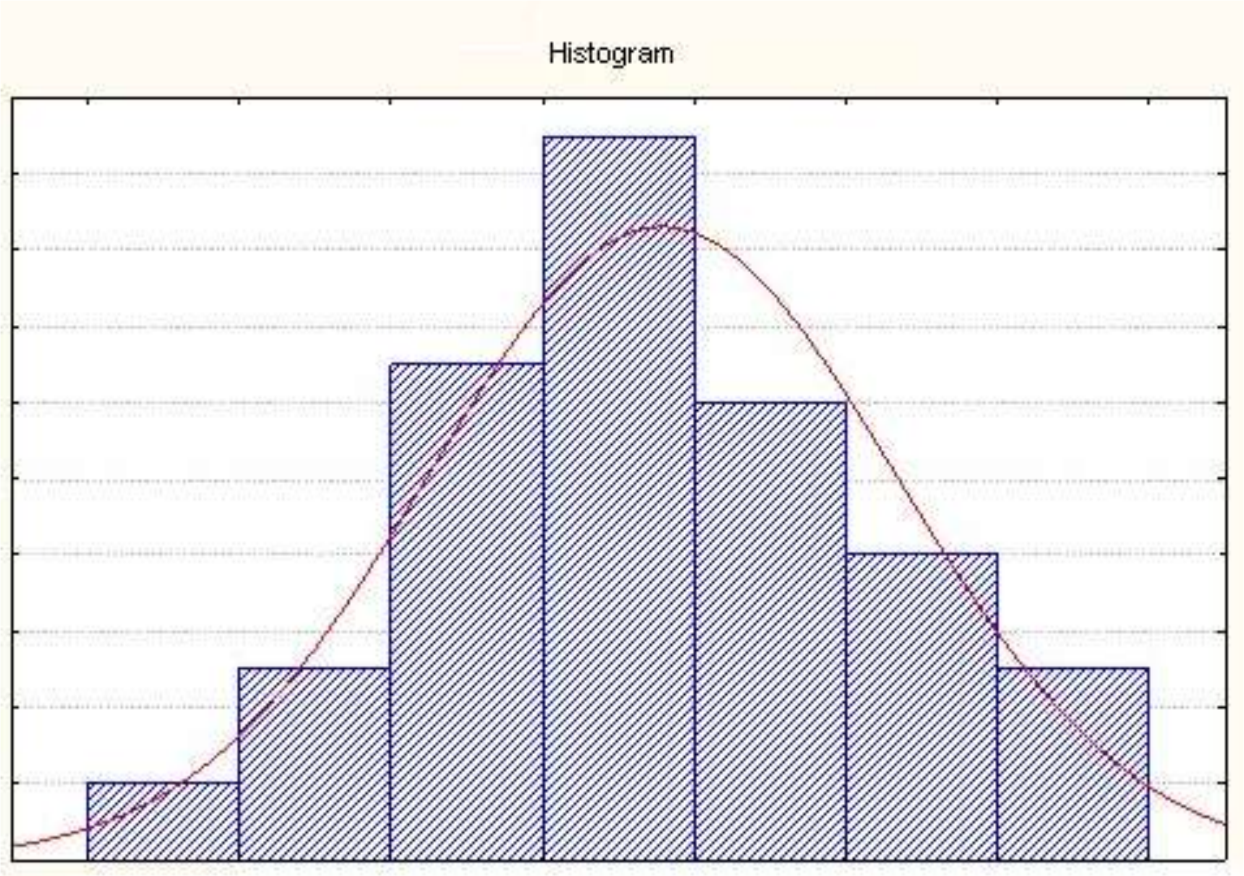
3. THE ASSUMPTION ABOUT THE DISTRIBUTION LAW (THE HYPOTHESIS)

- Let’s consider the basic distribution laws:
- a) the uniform distribution



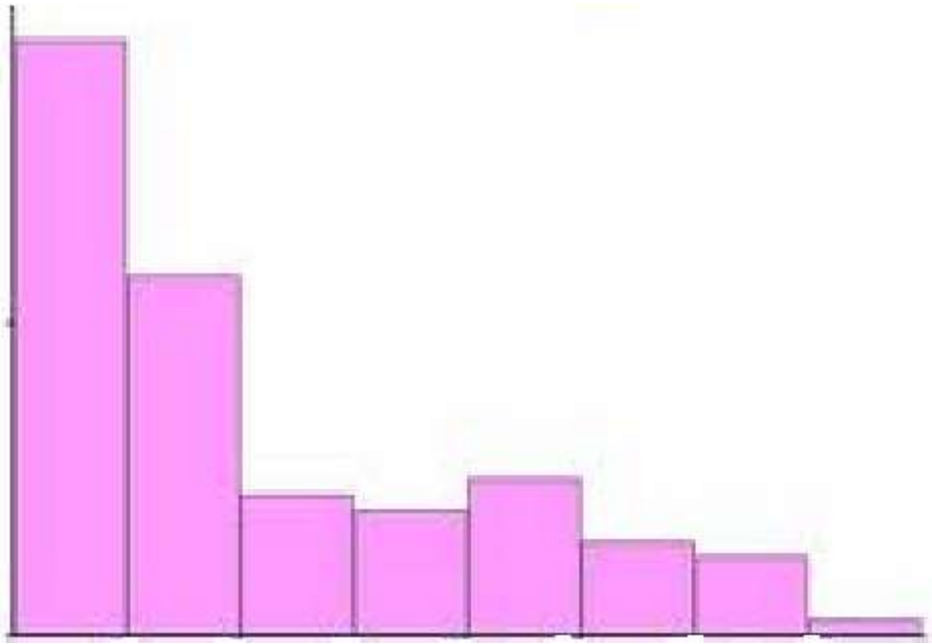
3. THE ASSUMPTION ABOUT THE DISTRIBUTION LAW (THE HYPOTHESIS)

- Let’s consider the basic distribution laws:
- b) the normal distribution



3. THE ASSUMPTION ABOUT THE DISTRIBUTION LAW (THE HYPOTHESIS)

- Let’s consider the basic distribution laws:
- c) the exponential distribution



3. THE ASSUMPTION ABOUT THE DISTRIBUTION LAW (THE HYPOTHESIS)

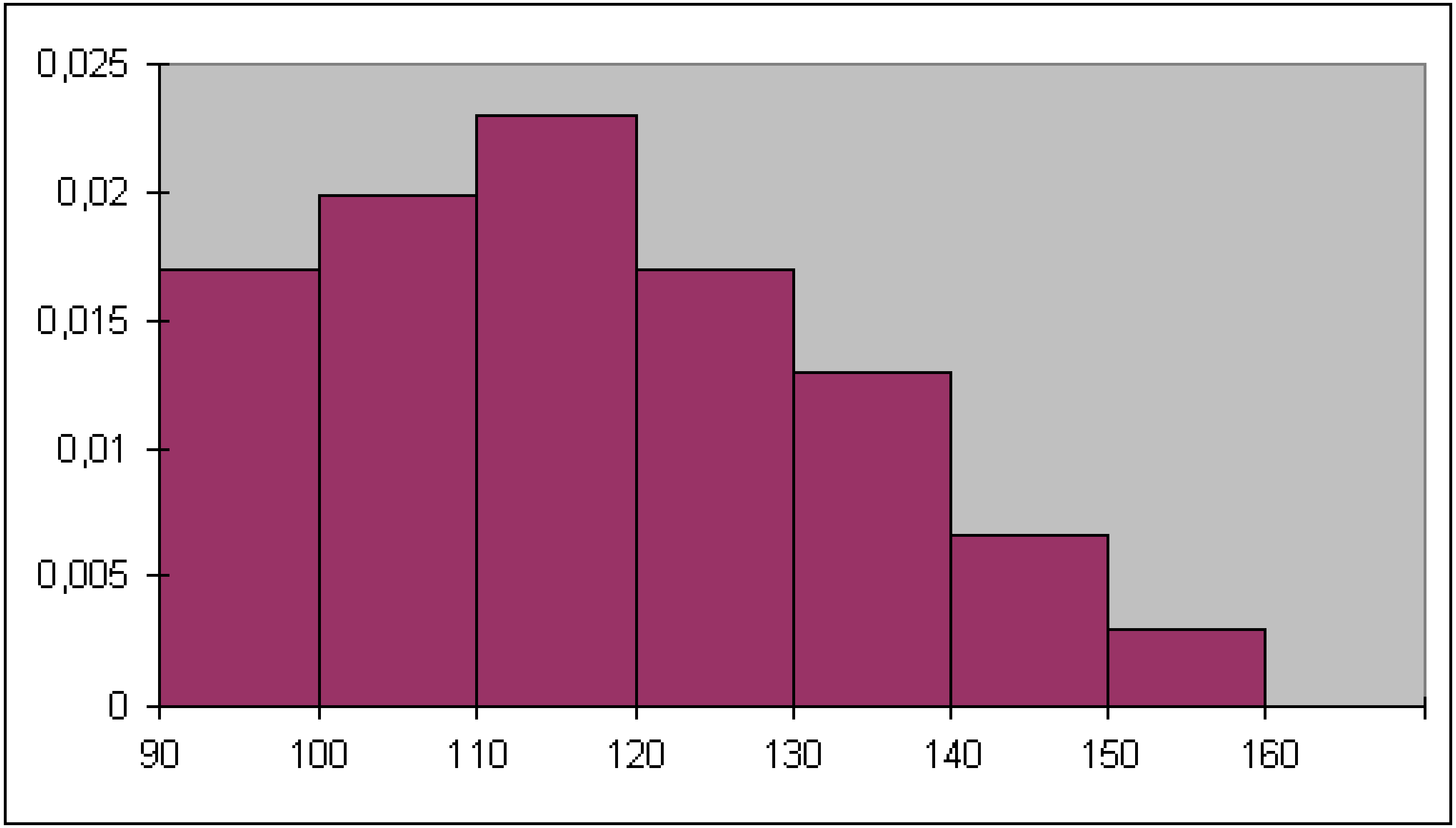
According to a form of a histogram we should make **an assumption** about a distribution law.

After this it's necessary to check this assumption with the help of **chi-square test for goodness of fit.**

3. THE ASSUMPTION ABOUT THE DISTRIBUTION LAW

№	$[x_i, x_{i+1})$	m_i	$\overline{x}_i$	$w_i = \frac{m_i}{n}$	$\frac{w_i}{h}$
1	$[90,100)$	5	95	5/30	5/300=0,017
2	$[100,110)$	6	105	6/30	6/300=0,02
3	$[110,120)$	7	115	7/30	7/300=0,023
4	$[120,130)$	5	125	5/30	5/300=0,017
5	$[130,140)$	4	135	4/30	4/300=0,013
6	$[140,150)$	2	145	2/30	2/300=0,0066
7	$[150,160)$	1	155	1/30	1/300=0,0033
	sum	**30**		**1**	

3. THE ASSUMPTION ABOUT THE DISTRIBUTION LAW



3. THE ASSUMPTION ABOUT THE DISTRIBUTION LAW

According to the histogram of this example we assume that the given sample has a normal distribution.

Let's check this assumption with the help of **chi-square test for goodness of fit.**

.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW (TABLE 1)

№	ξ_i	$t_i = \frac{\xi_i - \bar{x}}{s_{x\,s}}$	$\Phi(t_i)$	$\hat{F}_i$	p_i	$\hat{m}_i = n \cdot p_i$	m_i	$m_i - \hat{m}_i$	$\chi_i^2 = \frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1									
2									
				.					
k									
sum									χ_s^2

Let’s construct the table for χ^2 test.

1. The 2-nd column contains the upper limit of each interval $[x_i, x_{i+1})$. This value is denoted by ξ_i. For example, $\xi_1 = 100$

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

№	ξ_i	$t_i = \frac{\xi_i - \bar{x}}{s_{xs}}$	$\Phi(t_i)$	$\hat{F}_i$	p_i	$\hat{m}_i = n \cdot p_i$	m_i	$m_i - \hat{m}_i$	$\chi_i^2 = \frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	100								
2	110								
3	120								
4	130								
5	140								
6	150								
7	160								
sum									χ_s^2

2. $S_s = S_{xs} = 16{,}29$

3. Find the parameter $t_i = \dfrac{\xi_i - \overline{x}}{S_s}$ for each interval $[x_i, x_{i+1})$.

For example,

$$t_1 = \frac{100 - 117{,}33}{16{,}29} \approx -1{,}06$$

and so on.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

№	ξ_i	$t_i = \dfrac{\xi_i - \bar{x}}{s_{x\,s}}$	$\Phi(t_i)$	$\hat{F}_i$	p_i	$\hat{m}_i = n \cdot p_i$	m_i	$m_i - \hat{m}_i$	$\chi_i^2 = \dfrac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	100	-1,06							
2	110	-0,45							
3	120	0,16							
4	130	0,78							
5	140	1,39		.					
6	150	2,01							
7	160	2,62							
sum									χ_s^2

4. Find $\Phi(t_i)$ using the table for $\Phi(x)$.

For example, let's find $\Phi(t_1)=\Phi(-1{,}06)$.

Using the property $\Phi(-x)=-\Phi(x)$,

we get

$$\Phi(-1{,}06)=-\Phi(1{,}06)$$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0754
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2258	0,2291	0,2324	0,2356	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2996	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621

4. Find

$$\Phi(-0,45) = -\Phi(0,45)$$

Місюра Є.Ю.)

x	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0754
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2258	0,2291	0,2324	0,2356	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2996	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621

4. Find

$$\Phi(0{,}16)$$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0754
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2258	0,2291	0,2324	0,2356	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2996	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621

4. Find

$$\Phi(0{,}78)$$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0754
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2258	0,2291	0,2324	0,2356	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2996	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621

4. Find

$$\Phi(1{,}39)$$

x	0	1	2	3	4	5	6	7	8	9
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3906	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,1049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4274	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4430	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4648	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4700	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4762	0,4757
2,0	0,4772	0,4778	0,4783	0,4788	0,4796	0,4798	0,4803	0,4808	0,48Ï2	0,4817

4. Find

$$\Phi\left(2{,}01\right)$$

x	0	1	2	3	4	5	6	7	8	9
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3906	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,1049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4274	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4430	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4648	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4700	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4762	0,4757
2,0	0,4772	0,4778	0,4783	0,4788	0,4796	0,4798	0,4803	0,4808	0,48Ï2	0,4817

4. Find

$$\Phi(2{,}62)$$

x	0	1	2	3	4	5	6	7	8	9
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4874	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4903	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4924	0,4927	0,4929	0,4930	0,1932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4958	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4973
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4980	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4986	0,4986	0,4987	0,4987	0,4988	0,4988	0,4988	0,4989	0,4989	0,4990

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

№	ξ_i	$t_i = \frac{\xi_i - \bar{x}}{s_{xs}}$	$\Phi(t_i)$	$\hat{F}_i$	p_i	$\hat{m}_i = n \cdot p_i$	m_i	$m_i - \hat{m}_i$	$\chi_i^2 = \frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	100	-1,06	-0,3554						
2	110	-0,45	-0,1736						
3	120	0,16	0,0636						
4	130	0,78	0,2823						
5	140	1,39	0,4177	.					
6	150	2,01	0,4778						
7	160	2,62	0,4956						
sum									χ_s^2

5. Find the values $\hat{F}_i = \frac{1}{2} + \Phi(t_i)$, where $\hat{F}_i$ is the value of the distribution function for the experimental data.

For example,

$$\hat{F}_1 = \frac{1}{2} + \Phi(t_1) = \frac{1}{2} + (-0{,}3554) = 0{,}1446$$

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

№	ξ_i	$t_i = \frac{\xi_i - \bar{x}}{s_{xs}}$	$\Phi(t_i)$	$\hat{F}_i$ 0	p_i	$\hat{m}_i = n \cdot p_i$	m_i	$m_i - \hat{m}_i$	$\chi_i^2 = \frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	100	-1,06	-0,3554	0,1446					
2	110	-0,45	-0,1736	0,3264					
3	120	0,16	0,0636	0,5636					
4	130	0,78	0,2823	0,7823					
5	140	1,39	0,4177	0,9177 .					
6	150	2,01	0,4778	0,9778					
7	160	2,62	0,4956	0,9956					
sum									χ_s^2

6. Find the "theoretical" probability that *X* lies in the i-th interval, i.e. $p_i = \hat{F}_i - \hat{F}_{i-1}$ $\hat{F}_0 = 0$

For example,

.

$$p_1 = \hat{F}_1 - \hat{F}_0 = 0{,}1446 - 0 = 0{,}1446$$

6. Find the “theoretical” probability that *X* lies in the i-th interval, i.e. $p_i = \hat{F}_i - \hat{F}_{i-1}$ $\hat{F}_0 = 0$

For example,

$$p_1 = \hat{F}_1 - \hat{F}_0 = 0{,}1446 - 0 = 0{,}1446$$

$$p_2 = \hat{F}_2 - \hat{F}_1 = 0{,}3264 - 0{,}1446 = 0{,}1818$$

It’s necessary to remember that $\sum_{i=1}^{k} p_i \approx 1$

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

№	ξ_i	$t_i = \dfrac{\xi_i - \bar{x}}{s_{xs}}$	$\Phi(t_i)$	$\hat{F}_i$ 0	p_i	$\hat{m}_i = n \cdot p_i$	m_i	$m_i - \hat{m}_i$	$\chi_i^2 = \dfrac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	100	-1,06	-0,3554	0,1446	0,1446				
2	110	-0,45	-0,1736	0,3264	0,1818				
3	120	0,16	0,0636	0,5636	0,2372				
4	130	0,78	0,2823	0,7823	0,2187				
5	140	1,39	0,4177	0,9177	0,1354				
6	150	2,01	0,4778	0,9778	0,0601				
7	160	2,62	0,4956	0,9956	0,0178				
sum					**0,9956**				χ_s^2

7. Calculate $\hat{m}_i = n \cdot p_i$ (the “theoretical” frequency).

For example,

$$\hat{m}_1 = n \cdot p_1 = 30 \cdot 0{,}1446 = 4{,}338$$

It’s necessary to remember that $\sum_{i=1}^{k} \hat{m}_i \approx n$

In this case we have $\sum_{i=1}^{7} \hat{m}_i \approx 30$

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

№	ξ_i	$t_i = \frac{\xi_i - \bar{x}}{s_{xs}}$	$\Phi(t_i)$	$\hat{F}_i$ 0	p_i	$\hat{m}_i = n \cdot p_i$	m_i	$m_i - \hat{m}_i$	$\chi_i^2 = \frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	100	-1,06	-0,3554	0,1446	0,1446	4,338			
2	110	-0,45	-0,1736	0,3264	0,1818	5,454			
3	120	0,16	0,0636	0,5636	0,2372	7,116			
4	130	0,78	0,2823	0,7823	0,2187	6,561			
5	140	1,39	0,4177	0,9177	0,1354	4,062			
6	150	2,01	0,4778	0,9778	0,0601	1,803			
7	160	2,62	0,4956	0,9956	0,0178	0,5355			
sum					**0,9956**	**29,8695**			χ_s^2

8. Let’s take the value m_i from the previous table.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

№	ξ_i	$t_i = \frac{\xi_i - \bar{x}}{s_{xs}}$	$\Phi(t_i)$	$\hat{F}_i$ 0	p_i	$\hat{m}_i = n \cdot p_i$	m_i	$m_i - \hat{m}_i$	$\chi_i^2 = \frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	100	-1,06	-0,3554	0,1446	0,1446	4,338	5		
2	110	-0,45	-0,1736	0,3264	0,1818	5,454	6		
3	120	0,16	0,0636	0,5636	0,2372	7,116	7		
4	130	0,78	0,2823	0,7823	0,2187	6,561	5		
5	140	1,39	0,4177	0,9177	0,1354	4,062	4		
6	150	2,01	0,4778	0,9778	0,0601	1,803	2		
7	160	2,62	0,4956	0,9956	0,0178	0,5355	1		
sum					**0,9956**	**29,8695**	**30**		χ_s^2

9. Calculate $m_i - \hat{m}_i$.

For example,

$$m_1 - \hat{m}_1 = 5 - 4{,}338 = 0{,}662$$

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

№	ξ_i	$t_i = \frac{\xi_i - \bar{x}}{s_{xs}}$	$\Phi(t_i)$	$\hat{F}_i$ 0	p_i	$\hat{m}_i = n \cdot p_i$	m_i	$m_i - \hat{m}_i$	$\chi_i^2 = \frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	100	-1,06	-0,3554	0,1446	0,1446	**4,338**	**5**	**0,662**	
2	110	-0,45	-0,1736	0,3264	0,1818	5,454	6		
3	120	0,16	0,0636	0,5636	0,2372	7,116	7		
4	130	0,78	0,2823	0,7823	0,2187	6,561	5		
5	140	1,39	0,4177	0,9177	0,1354	4,062	4		
6	150	2,01	0,4778	0,9778	0,0601	1,803	2		
7	160	2,62	0,4956	0,9956	0,0178	0,5355	1		
sum					0,9956	29,8695	**30**		χ_s^2

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

№	ξ_i	$t_i = \frac{\xi_i - \bar{x}}{s_{xs}}$	$\Phi(t_i)$	$\hat{F}_i$ 0	p_i	$\hat{m}_i = n \cdot p_i$	m_i	$m_i - \hat{m}_i$	$\chi_i^2 = \frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	100	-1,06	-0,3554	0,1446	0,1446	4,338	5	0,662	
2	110	-0,45	-0,1736	0,3264	0,1818	5,454	6	0,546	
3	120	0,16	0,0636	0,5636	0,2372	7,116	7	-0,116	
4	130	0,78	0,2823	0,7823	0,2187	6,561	5	-1,561	
5	140	1,39	0,4177	0,9177	0,1354	4,062	4	-0,062	
6	150	2,01	0,4778	0,9778	0,0601	1,803	2	0,197	
7	160	2,62	0,4956	0,9956	0,0178	0,5355	1	0,4645	
sum					**0,9956**	**29,8695**	30		χ_s^2

10. Calculate $\chi_i^2 = \frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$ and

a sum $\sum_{i=1}^{k} \chi_i^2 = \chi_s^2$.

For example,

$$\chi_1^2 = \frac{(m_1 - \hat{m}_1)^2}{\hat{m}_1} = \frac{(5 - 4{,}338)^2}{4{,}338} = \frac{0{,}662^2}{4{,}338} = 0{,}1010$$

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

№	ξ_i	$t_i = \frac{\xi_i - \overline{x}}{s_{xs}}$	$\Phi(t_i)$	$\hat{F}_i$ 0	p_i	$\hat{m}_i = n \cdot p_i$	m_i	$m_i - \hat{m}_i$	$\chi_i^2 = \frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	100	-1,06	-0,3554	0,1446	0,1446	4,338	5	0,662	0,1010
2	110	-0,45	-0,1736	0,3264	0,1818	5,454	6	0,546	0,0547
3	120	0,16	0,0636	0,5636	0,2372	7,116	7	-0,116	0,0019
4	130	0,78	0,2823	0,7823	0,2187	6,561	5	-1,561	0,3714
5	140	1,39	0,4177	0,9177	0,1354	4,062	4	-0,062	0,00095
6	150	2,01	0,4778	0,9778	0,0601	1,803	2	0,197	0,0215
7	160	2,62	0,4956	0,9956	0,0178	0,5355	1	0,4645	0,4029
sum					**0,9956**	**29,8695**	30		**=0,95435** χ_s^2

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

α is the significance level (рівень значимості),

$\gamma = 1 - \alpha$ is the confidence probability (довірча ймовірність).

Usually γ equals 95%, 99%, 99,9%,

then α equals 5%, 1%.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

α is the significance level (рівень значимості).

α equals 5%, 1%.

Let's consider two significance levels:

$$\alpha = 0{,}01 \qquad\qquad \alpha = 0{,}05$$

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

11. With the help of the table of χ^2 test we define $\chi^2_{tabl}(0{,}01; k^*)$ and $\chi^2_{tabl}(0{,}05; k^*)$,

where k^* is called a number of degrees of freedom and $k^* = k - r - 1$:

k is a number of the intervals;

r is a number of parameters of the theoretical law ($r = 2$ for a normal law (a and σ^2)).

DEGREES of FREEDOM

The **degrees of freedom** in a statistical calculation represent how many values involved in a calculation have the freedom to vary. The degrees of freedom can be calculated to help ensure the statistical validity of chi-square test. This test is commonly used to compare observed data with data that would be expected to be obtained according to a specific hypothesis.

DEGREES of FREEDOM

Degrees of freedom are the number of independent values that a statistical analysis can estimate. You can also think of it as the number of values that are free to vary as you estimate parameters.

DEGREES of FREEDOM

Degrees of freedom are equal to your sample size minus the number of parameters you need to calculate during an analysis. It is usually a positive whole number.

DEGREES of FREEDOM

Degrees of freedom is a combination of how much data you have and how many parameters you need to estimate. It indicates how much independent information goes into a parame-ter estimate.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

Thus,

$$df = k^* = k - r - 1$$

uniform (a, b) dist. law $r = 2$

exp. dist. law (λ) $r = 1$

normal law (a, σ^2) $r = 2$

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

Thus,

$$df = k^{*} = k - r - 1$$

$$df = k^{*} = 7 - 2 - 1 = 4$$

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

Thus,

$$df = k^* = k - r - 1$$

$$df = k^* = 7 - 2 - 1 = 4$$

$$\alpha = 0{,}01$$

$$\chi^2_{tabl}(0{,}01;4) =$$

a number of degrees of freedom k^*	significance level α					
	0,01	0,05	0,1	0,90	0,95	0,99
1	6,6	3,8	2,71	0,02	0,004	0,0002
2	9,2	6,0	4,61	0,21	0,1	0,02
3	11,3	7,8	6,25	0,58	0,35	0,12
4	13,3	9,5	7,78	1,06	0,71	0,30
5	15,1	11,1	9,24	1,61	1,15	0,55
6	16,8	12,6	10,6	2,20	1,64	0,87
7	18,5	14,1	12,0	2,83	2,17	1,24
8	20,1	15,5	13,4	3,49	2,73	1,65
9	21,7	16,9	14,7	4,17	3,33	2,09
10	23,2	18,3	16,0	4,87	3,94	2,56
11	24,7	19,7	17,3	5,58	4,57	3,05
12	26,2	21,0	18,5	6,30	5,23	3,57
13	27,7	22,4	19,8	7,04	5,89	4,11
14	29,1	23,7	21,1	7,79	6,57	4,66
15	30,6	25,0	22,3	8,5	7,26	5,23
16	32,0	26,3	23,5	9,31	7,98	5,81
17	33,4	27,6	24,8	10,1	8,67	6,41
18	34,8	28,9	26,0	10,9	9,39	7,01
19	36,2	30,1	27,2	11,7	10,1	7,63
20	37,6	31,4	28,4	12,4	10,9	8,26
21	38,9	32,7	29,6	13,2	11,6	8,90
22	40,3	33,9	30,6	14,0	12,63	9,54
23	41,6	35,2	32,0	14,8	13,1	10,2
24	43,0	36,4	33,2	15,7	13,8	10,9
25	44,3	37,7	34,4	16,5	14,6	11,5
26	45,6	38,9	35,6	17,3	15,4	12,2
27	47,0	40,1	36,7	18,1	16,2	12,9
28	48,3	41,3	37,9	18,9	16,9	13,6
29	49,6	42,6	39,1	19,8	17,7	14,3
30	50,9	43,8	40,3	20,6	18,5	15,0

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

Thus,

$$df = k^* = k - r - 1$$

$$k^* = 7 - 2 - 1 = 4$$

$$\alpha = 0{,}01 \qquad \alpha = 0{,}05$$

$$\chi^2_{tabl}(0{,}01;4) = 13{,}3 \qquad \chi^2_{tabl}(0{,}05;4) =$$

a number of degrees of freedom k^*	significance level α					
	0,01	0,05	0,1	0,90	0,95	0,99
1	6,6	3,8	2,71	0,02	0,004	0,0002
2	9,2	6,0	4,61	0,21	0,1	0,02
3	11,3	7,8	6,25	0,58	0,35	0,12
4	13,3	9,5	7,78	1,06	0,71	0,30
5	15,1	11,1	9,24	1,61	1,15	0,55
6	16,8	12,6	10,6	2,20	1,64	0,87
7	18,5	14,1	12,0	2,83	2,17	1,24
8	20,1	15,5	13,4	3,49	2,73	1,65
9	21,7	16,9	14,7	4,17	3,33	2,09
10	23,2	18,3	16,0	4,87	3,94	2,56
11	24,7	19,7	17,3	5,58	4,57	3,05
12	26,2	21,0	18,5	6,30	5,23	3,57
13	27,7	22,4	19,8	7,04	5,89	4,11
14	29,1	23,7	21,1	7,79	6,57	4,66
15	30,6	25,0	22,3	8,5	7,26	5,23
16	32,0	26,3	23,5	9,31	7,98	5,81
17	33,4	27,6	24,8	10,1	8,67	6,41
18	34,8	28,9	26,0	10,9	9,39	7,01
19	36,2	30,1	27,2	11,7	10,1	7,63
20	37,6	31,4	28,4	12,4	10,9	8,26
21	38,9	32,7	29,6	13,2	11,6	8,90
22	40,3	33,9	30,6	14,0	12,63	9,54
23	41,6	35,2	32,0	14,8	13,1	10,2
24	43,0	36,4	33,2	15,7	13,8	10,9
25	44,3	37,7	34,4	16,5	14,6	11,5
26	45,6	38,9	35,6	17,3	15,4	12,2
27	47,0	40,1	36,7	18,1	16,2	12,9
28	48,3	41,3	37,9	18,9	16,9	13,6
29	49,6	42,6	39,1	19,8	17,7	14,3
30	50,9	43,8	40,3	20,6	18,5	15,0

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

Thus,

$$k^* = k - r - 1$$

$$k^* = 7 - 2 - 1 = 4$$

$$\alpha = 0{,}01$$

$$\alpha = 0{,}05$$

$$\chi^2_{tabl}(0{,}01;4) = 13{,}3$$

$$\chi^2_{tabl}(0{,}05;4) = 9{,}5$$

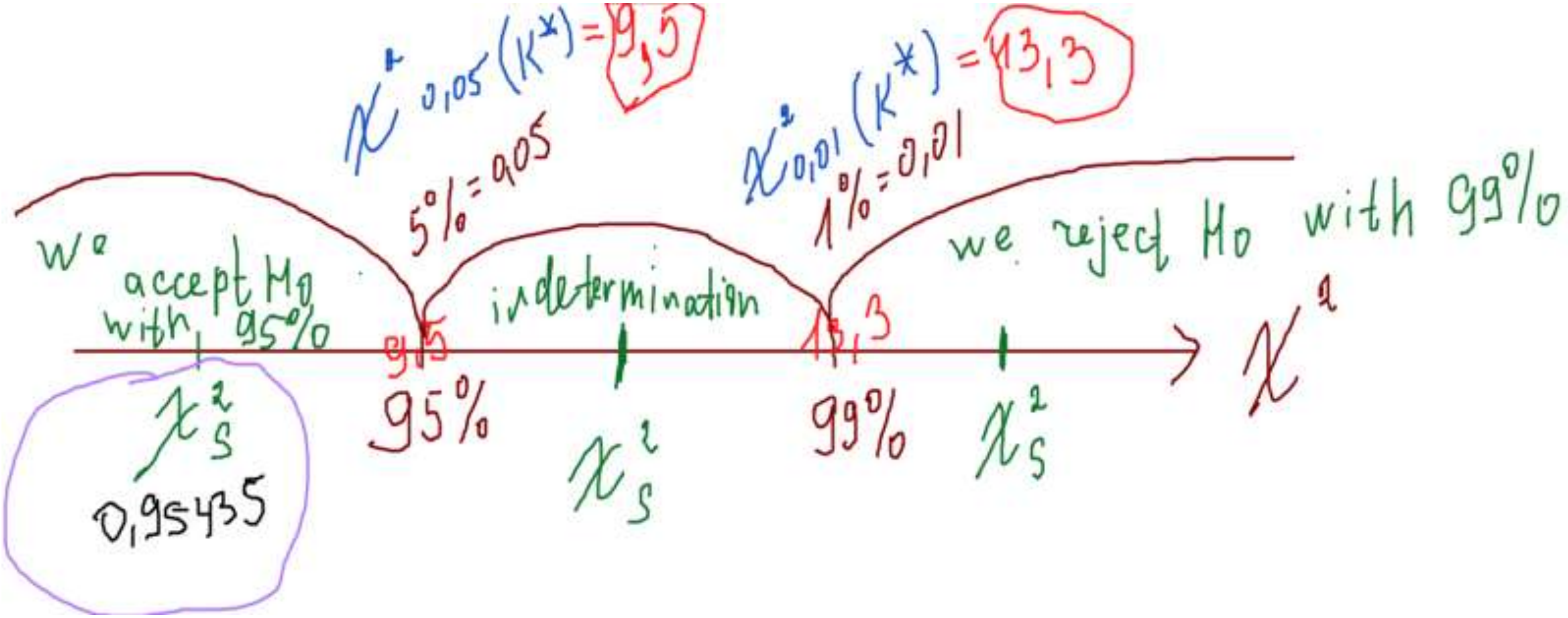
4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW

Thus,

$$k^* = 7 - 2 - 1 = 4$$

$$\chi^2_{tabl}(0{,}01;4) = 13{,}3$$

$$\chi^2_{tabl}(0{,}05;4) = 9{,}5$$



Let's consider χ^2 test for goodness (згода) of fit (вирівнювання)

The rule:

1) if

$$\chi_s^2 > \chi_{tabl}^2\left(0{,}01; k^*\right) > \chi_{tabl}^2\left(0{,}05; k^*\right)$$

then we reject (відкидати) a hypothesis H_0 *, i. e.*

a normal law does not correspond the experimental data;

Let's consider χ^2 test for goodness (згода) of fit (вирівнювання)

The rule:

2) if

$$\chi^2_{tabl}\left(0{,}05; k^*\right) < \chi^2_s < \chi^2_{tabl}\left(0{,}01; k^*\right)$$

then we have an indeterminateness relative to the distribution law;

It means that we can accept (приймати) or reject (відкидати) a hypothesis H_0

Let's consider χ^2 test for goodness (згода) of fit (вирівнювання)

The rule:

3) if

$$\chi_s^2 < \chi_{tabl}^2\left(0{,}05;k^*\right) < \chi_{tabl}^2\left(0{,}01;k^*\right)$$

then we accept (приймати) the assumption that the sample came from a normal distribution law (a hypothesis H_0).

Let's consider χ^2 test for goodness (згода) of fit (вирівнювання)

We compare χ_s^2 with the tabular values $\chi_{tabl}^2(0{,}01;\,4)$ and $\chi_{tabl}^2(0{,}05;\,4)$.

We have

$$\chi_s^2 = 0{,}95435$$

$$\chi_{tabl}^2(0{,}01;4) = 13{,}3$$

$$\chi_{tabl}^2(0{,}05;4) = 9{,}5$$

Let's consider χ^2 test for goodness (згода) of fit (вирівнювання)

We have $\chi_s^2 = 0,95435$

$$\chi_{tabl}^2(0,01;4) = 13,3$$

$$\chi_{tabl}^2(0,05;4) = 9,5$$

$$0,95435 < 9,5 < 13,3$$

Let's consider χ^2 test for goodness (згода) of fit (вирівнювання)

We have $$\chi_s^2 = 0,95435$$

$$\chi_{tabl}^2(0,01;4) = 13,3$$

$$\chi_{tabl}^2(0,05;4) = 9,5$$

$$0,95435 < 9,5 < 13,3$$

$$\chi_s^2 < \chi_{tabl}^2(0,05;4) < \chi_{tabl}^2(0,01;4)$$

Let's consider χ^2 test for goodness (згода) of fit (вирівнювання)

We have

$$0,95435 < 9,5 < 13,3$$

$$\chi_s^2 < \chi_{tabl}^2(0,05;\,4) < \chi_{tabl}^2(0,01;\,4)$$

The third inequality is fulfilled. It means we accept the assumption that this sample came from the normal distribution law (the hypothesis H_0).

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW (TABLE 2)

$[x_i; x_{i+1})$	x_i'							

.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW (TABLE 2)

$[x_i; x_{i+1})$	x_i'							

$$x_i' = \frac{x_i + x_{i+1}}{2}$$ is a midpoint of the interval

.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW (TABLE 2)

$[x_i; x_{i+1})$	x'_i	m_i						

m_i is a frequency of the interval

.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW (TABLE 2)

$[x_i; x_{i+1})$	x'_i	m_i	t_i					

$$t_i = \frac{x'_i - \overline{x}_s}{S_x}$$

is a standardized value of the interval

.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW (TABLE 2)

$[x_i;x_{i+1})$	x_i'	m_i	t_i	$\varphi(t_i)$				

$\varphi(t_i)$ is a value of Laplace differential function (see table 1)

.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW (TABLE 2)

$[x_i; x_{i+1})$	x_i'	m_i	t_i	$\varphi(t_i)$	f_i			

$$f_i = \frac{\varphi(t_i)}{S_x}$$

is a value of the differential function of a normal law

.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW (TABLE 2)

$[x_i; x_{i+1})$	x_i'	m_i	t_i	$\varphi(t_i)$	f_i	$\tilde{p}_i = h_i f_i$		

$\tilde{p}_i = h_i f_i$ is the "theoretical" probability of a normal law

.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW (TABLE 2)

$[x_i; x_{i+1})$	x'_i	m_i	t_i	$\varphi(t_i)$	f_i	$\tilde{p}_i = h_i f_i$		

$\tilde{p}_i = h_i f_i$ is the "theoretical" probability of a normal law

.

$$h_i = \Delta x_i = x_{i+1} - x_i$$

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW (TABLE 2)

$[x_i; x_{i+1})$	x_i'	m_i	t_i	$\varphi(t_i)$	f_i	$\tilde{p}_i = h_i f_i$	$\tilde{m}_i = n\tilde{p}_i$	

$\tilde{m}_i$ is the "theoretical" frequency of a normal law

.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW (TABLE 2)

$[x_i; x_{i+1})$	x_i'	m_i	t_i	$\varphi(t_i)$	f_i	$\tilde{p}_i = h_i f_i$	$\tilde{m}_i = n\tilde{p}_i$	χ_i^2

$$\chi_i^2 = \frac{(m_i - \tilde{m}_i)^2}{\tilde{m}_i}$$

is a value of chi-square

.

4. CHECKING THE HYPOTHESIS RELATIVE TO THE DISTRIBUTION LAW (TABLE 2)

$[x_i; x_{i+1})$	x'_i	m_i	t_i	$\varphi(t_i)$	f_i	$\tilde{p}_i = h_i f_i$	$\tilde{m}_i = n\tilde{p}_i$	χ_i^2

$$\chi_i^2 = \frac{(m_i - \tilde{m}_i)^2}{\tilde{m}_i}$$ is a value of chi-square

.

$$\chi^2 = \sum_{i=1}^{k} \frac{(m_i - \tilde{m}_i)^2}{\tilde{m}_i}$$

EXAMPLE

Check the hypothesis about a normal distribution law.

Intervals of ***X***	10 – 14	14 – 18	18 – 22	22 – 26	26 – 30
Frequencies	3	8	15	10	4

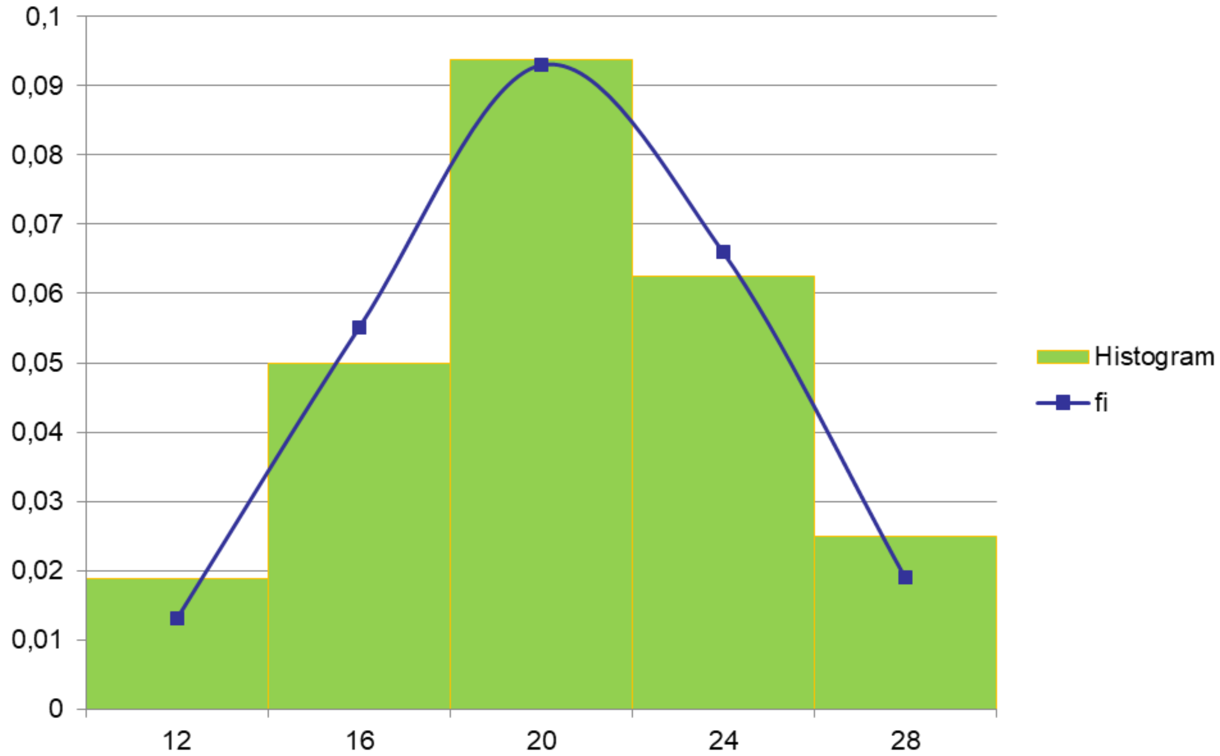
EXAMPLE

Check the hypothesis about a normal distribution law.

Intervals of ***X***	10 – 14	14 – 18	18 – 22	22 – 26	26 – 30
Frequencies	3	8	15	10	4

Let's assume the hypothesis H_0 : the random variable ***X*** is distributed by a normal law

x'_i	12	16	20	24	28	**sum**
w_i	0,075	0,2	0,375	0,25	0,1	1
$x'_i w_i$	0,9	3,2	7,5	6	2,8	20,4
$x'^2_i w_i$	10,8	51,2	150	144	78,4	434,4
$t_i = \frac{x_i - \bar{x}_S}{S_x}$	-1,97	-1,03	-0,09	0,84	1,78	–
$\varphi(t_i)$	0,0573	0,2347	0,3973	0,2803	0,0818	–
$f_i = \frac{\varphi(t_i)}{S_x}$	0,013	0,055	0,093	0,066	0,019	–
$\tilde{m}_i = h_i n f_i$	2,15	8,79	14,88	10,50	3,06	39,4
χ^2_i	0,339	0,071	0,001	0,024	0,286	0,721


0,1
0,09
0,08
0,07
0,06
0,05
0,04
0,03
0,02
0,01
0
12
16
20
24
28
Histogram
fi

CALCULATIONS

The sample mean: $\overline{x_s} = 20{,}4;$

The sample mean of square: $\overline{x_s^2} = 434{,}4;$

The variance: $S_x^2 = \overline{x_s^2} - \left(\overline{x_s}\right)^2 = 434{,}4 - 20{,}4^2 = 18{,}24$

The root-mean-square deviation:

$$S_x = \sqrt{S_x^2} = \sqrt{18{,}24} \approx 4{,}27$$

CALCULATIONS

Let's find the value:

$$t_1 = \frac{x'_1 - \bar{x}_s}{S_x} = \frac{12 - 20,4}{4,27} = -1,97$$

Let's use **TABLE 1**:

$$\varphi(t_1) = \varphi(-1,97) = \varphi(1,97) = 0,0573$$

CALCULATIONS

Let’s find the value of the differential function:

$$f_1 = \frac{\varphi(t_1)}{S_x} = \frac{0,0573}{4,27} = 0,013$$

Let’s find the “theoretical” frequency:

$$\tilde{m}_1 = h_1 n f_1 = 4 \cdot 40 \cdot 0,013 = 2,15$$

$$\sum \tilde{m}_i = 39,4 \approx 40$$

CALCULATIONS

Let's find the empirical value of chi-square:

$$\chi_s^2 = 0{,}721$$

Let's find the number of degrees of freedom:

$$k^* = k - r - 1 = 5 - 2 - 1 = 2.$$

CALCULATIONS

Let's use TABLE 4 for chi-square:

$$\chi^2_{0,05} = 6,0 \qquad \chi^2_{0,01} = 9,2$$

Since $\chi_s^2 = 0,721 < \chi^2_{0,05} = 6,0$

the hypothesis about the theoretical normal law is

correct for initial data.

Let's consider Romanovskiy's test for goodness of fit

We have:

$$k^* = k - r - 1$$

$$k^* = 7 - 2 - 1 = 4$$

$$t = \left| \frac{\chi^2 - k^*}{\sqrt{2k^*}} \right|$$

Let’s consider Romanovskit’s test for goodness of fit

We have:

$$k^* = k - r - 1$$

$$k^* = 7 - 2 - 1 = 4$$

$$t = \left| \frac{\chi^2 - k^*}{\sqrt{2k^*}} \right| = \left| \frac{0{,}95435 - 4}{\sqrt{8}} \right| = \left| -1{,}077 \right| = 1{,}077$$

Let’s find the value:

$$t = \left| \frac{\chi^2 - k^*}{\sqrt{2k^*}} \right| = 1{,}077$$

If $t \le 1{,}96$, then with the probability 95% we can accept the hypothesis H_0, that is, the distribution law corresponds to the theoretical law;

ROMANOVSKIY'S TEST

$$t = \left| \frac{\chi^2 - k^*}{\sqrt{2k^*}} \right|$$

The number of degrees of freedom:

$$k^* = k - r - 1$$

If $t \le 1{,}96$, then with the probability 95% we can accept the hypothesis H_0 , that is, the distribution law corresponds to the theoretical law.

ROMANOVSKIY'S TEST

$$t = \left| \frac{\chi^2 - k^*}{\sqrt{2k^*}} \right|$$

The number of degrees of freedom:

$$k^* = k - r - 1$$

If $t \geq 2{,}58$, then with the significance level $\alpha = 0{,}01$ we can reject H_0 .

ROMANOVSKIY'S TEST

$$t = \left| \frac{\chi^2 - k^*}{\sqrt{2k^*}} \right|$$

The number of degrees of freedom:

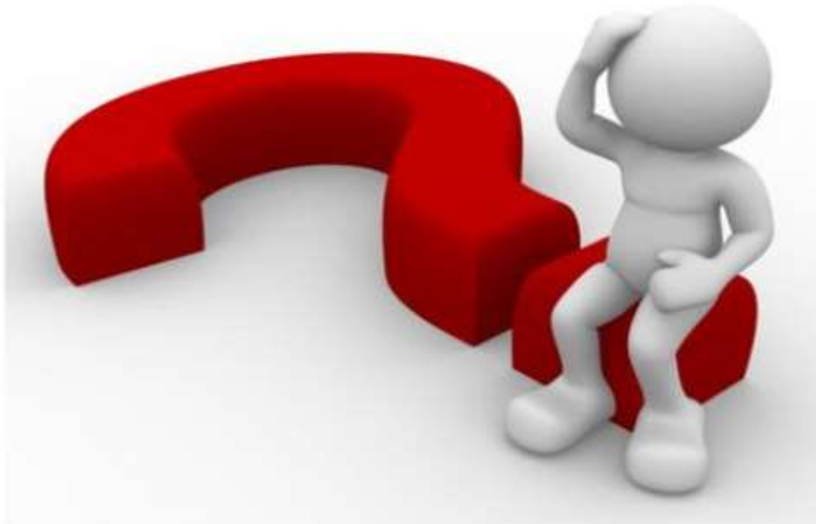
$$k^* = k - r - 1$$

If $t \in (1{,}96; 2{,}58)$ then we have an indeter-minateness relative to the distribution law, that is, we can't reject or can't accept H_0 .

S A M P L E (UL)

10,1 8,2 11,3 7,3 11,1 6,9 7,7 9,2 8,5 6,4 6,5 10,5 10,9 7,5 9,0 7,6 9,7 7,1 10,6 9,1 11,0 10,8 11,3 8,6 6,4 9,9 8,7 9,4 10,3 6,5 6,8 8,8 10,7 6,9 9,3 10,0 6,6 11,4 8,4 10,2 7,3 9,6 8,2 8,5 8,9 7,4 11,4 9,0 8,0 8,1 9,5 7,0 11,2 6,5 11,4 6,7 10,4 8,3 7,9 7,2 6,6 9,8 10,9 7,5 8,2

Thank you four your attention!!



Any questions?