PRACTICAL LESSON 3
STATISTICAL HYPOTHESES TESTING

OBJECTIVE
1. To study the general hypothesis testing scheme.
2. To learn to use parametric and nonparametric tests to test hypotheses about the equality of general means, variances and distribution laws of random variables

The student must study the material on the topic and be able to answer the following questions:

1. Statistical hypothesis.
2. General formulation of the hypothesis testing problem.
3. Comparison of the mean values of two normally distributed general populations. Student's criterion.
4. Testing hypotheses for variances. Fisher's criterion.
5. Testing hypotheses about distribution laws. Pearson's criterion.
6. Nonparametric tests.

A **statistical hypothesis** is any assumption about the distribution of observed random variables, any statement about the general population that is tested on a sample.

Statistical hypotheses are assumptions about the form of an unknown distribution or about the parameters of known distributions.

When testing any statistical hypothesis, a decision is never made with absolute certainty, there is always a risk of making a wrong decision. It is in the control and assessment of this risk that the essence of testing statistical hypotheses lies.

A statistical hypothesis is denoted by the letter *H*. We assume that there are 2 not intersecting hypotheses: **H_0 - null hypothesis (basic hypothesis) and H_1 - alternative hypothesis (competing hypothesis).**

The purpose of testing a statistical hypothesis is to accept (consider as true) either a null hypothesis H_0 or an alternative hypothesis H_1 on the basis of sample data.

For this a statistical tests (criteria) are used, which can be parametric and nonparametric.

A criterion is called **parametric** if it is based on a specific type of distribution of the general population (as a rule, normal) or uses the parameters of this population (mean, variance, etc.). The most commonly used parametric test is the study of differences in mean values based on the Student's *t*-test. A parametric *F*-test for equality of variances can be used to compare the variation (variances) in two populations.

A criterion is called **nonparametric** if it is not based on an assumption about the type of distribution of the general population and does not use the parameters of this population. Therefore, for nonparametric tests, it is also proposed to use a term «distribution-free test».

THE MAIN STAGES OF STATISTICAL HYPOTHESES TESTING:

1. A null hypothesis is formulated about the absence of differences between the general parameters, about the absence of a significant difference between the actual distribution and some theoretical (specified), etc. The essence of the null hypothesis: there is no difference between the general parameters and the differences in sample characteristics are random.

2. Competing hypothesis is formulated (the differences in sample characteristics are not random and the difference between the general parameters is statistically significant).

3. A degree of risk is accepted to reject the proposed null hypothesis, if it is true. This risk α is called the **significance level**.

4. To test the validity of the null hypothesis, a criterion is used. **A statistical criterion is a random variable *K*, which serves to test H_0.** A criterion depends upon the level of significance α and the number of degrees of freedom *f* (which is determined by sample size), $K = K(\alpha, f)$. The distribution functions of *K* are known and tabulated (are given in tables). **The actual (observed, experimental) value of the criterion K_{obs} is determined by results of an experiment (investigation).**

5. Using the table, the **critical value $K_{crit}(\alpha, f)$ is determined**, which is unlikely to be exceeded if the null hypothesis is true.

6. K_{obs} and $K_{crit}(\alpha, f)$ are compared.

In the case H_0 is true $K_{obs} < K_{crit}(\alpha, f)$, so

- **if $K_{obs} < K_{crit}(\alpha, f)$ null hypothesis H_0 is accepted;**
- **if $K_{obs} > K_{crit}(\alpha, f)$ H_0 is rejected and competing hypothesis H_1 is accepted.**

7. The **inference** is made about statistically significance (or insignificance) of corresponding general parameters at the chosen level of significance.

TESTING A HYPOTHESIS ABOUT POPULATION MEANS (EXPECTATION)

Two independent size n_1 and n_2 samples are compared, taken from normally distributed general populations with expectations μ_1 and μ_2, correspondingly. It is also supposed that the general variances are equal. The sample mean values are estimated $\bar{x}_1$ and $\bar{x}_2$, and the sample corrected variances S_1^2 and S_2^2 are also estimated. The significance level is given (chosen).

1. Null hypothesis is formulated, H_0: $\mu_1 = \mu_2$.
2. Competitive hypothesis is formulated, H_1: $\mu_1 \neq \mu_2$.
3. Student's t-criterion for comparison of means can be used to test the null hypothesis. t-criterion is calculated by the formula:

$$t_{\text{obs}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(n_1-1)S_1^2 + (n_2-1)S_2^2}} \cdot \sqrt{\frac{n_1+n_2-2}{\frac{1}{n_1}+\frac{1}{n_{21}}}}.$$

It is shown that in the case the null hypothesis is true the random variable t_{obs} is distributed according to t-distribution of Student with number $f = n_1 + n_2 - 2$ degrees of freedom.

4. $t_{\text{crit}}(\alpha, f = n_1 + n_2 - 2)$ is found from Student distribution table.
5. t_{obs} and t_{crit} are compared.

If $|t_{\text{obs}}| < t_{\text{crit}}(\alpha, f)$ the null hypothesis H_0 is accepted (the difference of general expectations is statistically insignificant).

If $|t_{\text{obs}}| > t_{\text{crit}}(\alpha, f)$ the null hypothesis H_0 is rejected and the competing H_1 is accepted (the difference of general expectations is statistically significant).

Problem 3.1.

Two independent small samples, size $n_1 = 5$ and $n_2 = 6$, are extracted from the normally distributed population X_1 and X_2. The sample means are calculated: $\bar{x}_1 = 3{,}3;\ \bar{x}_2 = 2{,}48$. It is known that general variances are approximately equal, i.e. $\sigma^2_{x_1} = \sigma^2_{x_2}$. At the significance level $\alpha \le 0{,}05$ test the null hypothesis:

$H_0:\ M(X_1) = M(X_2)$, if $t_{obs} = 3{,}27$.

Solution:

Number of degrees of freedom $f = n_1 + n_2 - 2 = 5 + 6 - 2 = 9$

From the table 1 "**Critical points of two-tailed Student's *t*-criterion**"

Number of degrees of freedom f	Levels of significance α, % (two-sided test)		
	5	1	0,1
1	12,71	63,66	64,60
…	…	…	…
8	2,31	3,36	5,04
9	**2,26**	3,25	4,78
10	2,23	3,17	4,59
…	…	…	…

$t_{crit}(\alpha \le 0{,}05, f = 9) = 2{,}26$.

$t_{obs} > t_{crit}(\alpha, f)$, hence H_0 is rejected.

Conclusion: $M(X_1) \ne M(X_2)$, at significance level of $\alpha \le 0{,}05$ the general expectations differ statistically significant.

TESTING A HYPOTHESIS ABOUT POPULATION VARIANCES

Let general populations X_1 and X_2 are normally distributed. Two independent size n_1 and n_2 samples are taken out of these populations and sample corrected variances S_1^2 and S_2^2 are estimated. At a given significance level it is necessary to test the null hypothesis about the equality of general variances σ_1^2 and σ_2^2.

1. H_0: $\sigma_1^2 = \sigma_2^2$.
   H_1: $\sigma_1^2 \ne \sigma_2^2$.
2. A random magnitude

$$F_{obs} = \frac{S^2_{greater}}{S^2_{smaller}},$$

which is equal to the ratio of greater variance to smaller variance is used as statistical criterion.

It is shown that in case the null hypothesis is true random variable F is distributed according to Fisher-Snedecor distribution with degrees of freedom $f_1 = n_1 - 1$ and $f_2 = n_2 - 1$, where n_1 – size of a sample with greater variance, n_2 – size of a sample with smaller variance.

Using F-distribution table $F_{crit}(\alpha, f_1, f_2)$ is found.

3. $F_{obs} = \dfrac{S^2_{greater}}{S^2_{smaller}}$ and $F_{crit}(\alpha, f_1, f_2)$ are compared.

If $F_{obs} < F_{crit}(\alpha, f_1, f_2)$ the null hypothesis H_0 is accepted (the difference of general variances σ_1^2 and σ_2^2 is statistically insignificant, $\sigma_1^2 = \sigma_2^2$)

If $F_{obs} > F_{crit}(\alpha, f_1, f_2)$ the null hypothesis H_0 is rejected and the competing H_1 is accepted (the difference of general variances σ_1^2 and σ_2^2 is statistically significant, $\sigma_1^2 \ne \sigma_2^2$).

Problem 3.2. For two independent samples of sizes $n_1 = 12$ and $n_2 = 15$, extracted from normal general populations X_1 and X_2 , the corrected sample variances are calculated: $S_1^2 = 11{,}41$ and $S_2^2 = 6{,}52$. At the significance level of $\alpha \le 0{,}05$ test the null hypothesis about the equality of general variances.

Solution: Competing hypothesis are put forward:

null hypothesis H_0: $\sigma_1^2 = \sigma_2^2$, general variances do not differ statistically significant (in case H_0 is true $F_{obs} < F_{crit}(\alpha, f_1, f_2)$);

conquering hypothesis H_1: $\sigma_1^2 \ne \sigma_2^2$, general variances differ statistically significant ($F_{obs} > F_{crit}(\alpha, f_1, f_2)$).

$F_{observed}$ is calculated: $F_{obs} = \dfrac{S^2_{greater}}{S^2_{smaller}} = \dfrac{11{,}41}{6{,}52} = 1{,}75$

From Fisher-Snedecor F-criterion table F_{crit} is found for $f_{greater} = n_1 - 1 = 12 - 1 = 11$ and $f_{smaller} = n_2 - 1 = 15 - 1 = 14$ at $\alpha \le 0{,}05$:

Table. Critical points of two-tailed F-criterion of Fisher-Snedecor for different numbers of degrees of freedom (f_1 and f_2) and significance levels $\alpha = 5\%$ (upper line) and $\alpha = 1\%$ (lower line)

f_2	f_1 – number of degrees of freedom for greater variance											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	161 4052	200 4999	216 5403	225 5625	230 5764	234 5889	237 5928	239 5981	241 6022	242 6056	243 6082	244 6106
2	18,5 98,5	19,0 99,0	19,2 99,2	19,3 99,3	19,3 99,3	19,4 99,3	19,4 99,3	19,4 99,4	19,4 99,4	19,4 99,4	19,4 99,4	19,4 99,4
…	…	…	…	…	…	…	…	…	…	…	…	…
13	4,7 9,1	3,8 6,7	3,4 5,7	3,2 5,2	3,0 4,9	2,9 4,6	2,8 4,4	2,8 4,3	2,7 4,2	2,7 4,1	2,6 4,0	2,6 4,0
14	4,6 8,9	3,7 6,5	3,3 5,6	3,1 5,0	3,0 4,7	2,9 4,5	2,8 4,3	2,7 4,1	2,7 4,0	2,6 3,9	2,6 3,9	2,5 3,8
15	4,5 8,7	3,7 6,4	3,3 5,4	3,1 4,9	2,9 4,6	2,8 4,3	2,7 4,1	2,6 4,0	2,6 3,9	2,6 3,8	2,5 3,7	2,5 3,7
…	…	…	…	…	…	…	…	…	…	…	…	…

$F_{obs} = 1{,}75$; $F_{crit} = 2{,}6$;

$F_{obs} < F_{crit}$. There is no reason to reject the null hypothesis about the equality of general variances. General variances do not differ statistically significant, $\sigma_1^2 = \sigma_2^2$.

Problem 3.3. For two independent small samples of volumes $n_1 = 5$ and $n_2 = 6$, extracted from normal general populations X_1 and X_2, the sample means $\bar{x}_1 = 8{,}3$ and $\bar{x}_2 = 7{,}48$ and sample variances $S_1^2 = 0{,}25$ and $S_2^2 = 0{,}108$ are calculated. At the significance level $\alpha \le 0.05$, test the null hypothesis $H_0 : M(X_1) = M(X_2)$.

Solution. Since the sample variances are different, first the null hypothesis about the equality of general variances must be checked using the Fisher-Snedecor test.

$H_0: \sigma_1^2 = \sigma_2^2$, general variances do not differ statistically significant (in case H_0 is true $F_{obs} < F_{crit}(\alpha, f_1, f_2)$);

$H_1: \sigma_1^2 \neq \sigma_2^2$, general variances differ statistically significant ($F_{obs} > F_{crit}(\alpha, f_1, f_2)$).

The ratio of the larger corrected variance to the smaller one:

$$F_{набл} = \frac{S^2_{greater}}{S^2_{smaller}} = \frac{0{,}25}{0{,}108} = 2{,}31.$$

$$n_1 = n_{greater} = 5; \quad f_{greater} = n_1 - 1 = 5 - 1 = 4.$$

$$n_2 = n_{smaller} = 6; \quad f_{smaller} = n_2 - 1 = 6 - 1 = 5.$$

Table 3. Critical points of two-tailed F-criterion of Fisher-Snedecor for different numbers of degrees of freedom (f_1 and f_2) and significance levels $\alpha = 5\%$ (upper line) and $\alpha = 1\%$ (lower line)

f_2	f_1 – number of degrees of freedom for greater variance											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	161 4052	200 4999	216 5403	225 5625	230 5764	234 5889	237 5928	239 5981	241 6022	242 6056	243 6082	244 6106
…	…	…	…	…	…	…	…	…	…	…	…	…
4	7,7 21,2	6,9 18,0	6,6 16,7	6,4 16,0	6,3 15,5	6,2 15,2	6,1 15,0	6,0 14,8	6,0 14,7	6,0 14,5	5,9 14,5	5,9 14,4
5	6,6 16,3	5,8 13,3	5,4 12,1	5,2 11,4	5,1 11,0	5,0 10,7	4,9 10,5	4,8 10,3	4,8 10,2	4,7 10,1	4.7 10,0	4,7 9,9
6	6,0 13,7	5,1 10,9	4,8 9,8	4,5 9,2	4,4 8,8	4,3 8,5	4,2 8,3	4,2 8,1	4,1 8,0	4,1 7,9	4,0 7,8	4,0 7.7
…	…	…	…	…	…	…	…	…	…	…	…	…

The table of critical points of F-criterion gives $F_{crit}(\alpha \le 0{,}05, f_1 = 4, f_2 = 5) = 5{,}2$

$F_{obs} < F_{crit}(\alpha, f_1, f_2)$, the null hypothesis is accepted $H_0: \sigma_1^2 = \sigma_2^2$, general variances do not differ statistically significant.

Then general expectations can be compared. Now,

$H_0:\ M(X_1) = M(X_2)$ (in case it is true $t_{obs} < t_{crit}(\alpha, f)$);

$H_1:\ M(X_1) \neq M(X_2)$ ($t_{obs} > t_{crit}(\alpha, f)$).

$t_{observed}$ is calculated:

$$t_{obs} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}} \cdot \sqrt{\frac{n_1 + n_2 - 2}{\frac{1}{n_1} + \frac{1}{n_2}}} =$$

$$= \frac{8{,}3 - 7{,}48}{\sqrt{(5-1)0{,}25 + (6-1)0{,}108}} \cdot \sqrt{\frac{5 + 6 - 2}{\frac{1}{5} + \frac{1}{6}}} = 3{,}27$$

From the table of critical points of Student's distribution at significance level of $\alpha \le 0{,}05$ for $f = n_1 + n_2 = 5 + 6 - 2 = 9$

Number of degrees of freedom f	Levels of significance α, % (two-sided test)		
	5	1	0,1
1	12,71	63,66	64,60
…	…	…	…
8	2,31	3,36	5,04
9	**2,26**	3,25	4,78
10	2,23	3,17	4,59
…	…	…	…

$t_{crit}(\alpha \le 0{,}05, f = 9) = 2{,}26.$

Since $t_{obs} > t_{crit}(\alpha, f)$ the null hypothesis is not true, $H_0: M(X_1) = M(X_2)$ is rejected and $H_1: M(X_1) \ne M(X_2)$ is accepted. General means (mathematical expectations) differ statistically significant.

Problem 3.4. Two independent small samples, size $n_1 = 11$ and $n_2 = 14$, are extracted from the normally distributed population X_1 и X_2. At the significance level $\alpha \le 0{,}05$ test the null hypothesis: $H_0: M(X_1) = M(X_2)$.

First sample:

x_{1i}	5	6	8	12	16	18	19	20
n	1	1	1	3	2	1	1	1

Second sample:

x_{2i}	6	8	9	13	14	16	19	20	21
n	1	1	1	2	3	3	1	1	1

1. Calculate the mean values for the first and second samples
2. Calculate corrected sample variances
3. Test the hypothesis about the equality of general variances at the significance level $\alpha \le 0{,}05$
4. If this hypothesis is true then test the hypothesis about the equality of expectations for the two general populations at the significance level $\alpha \le 0{,}05$.

TESTING A HYPOTHESIS ABOUT THE LAW OF DISTRIBUTION. PEARSON GOODNESS-OF-FIT TEST.

In many practical problems, random variable law of distribution is not known in advance, and it is necessary to choose a model that is consistent with the experimental results.

Null hypothesis H_0 is formulated,

H_0: an unknown function of distribution $F(x)$ of the random variable X studied coincides with some theoretical law (e.g., normal), $F(x) = F_{theor}(x)$.

Competitive hypothesis is formulated,

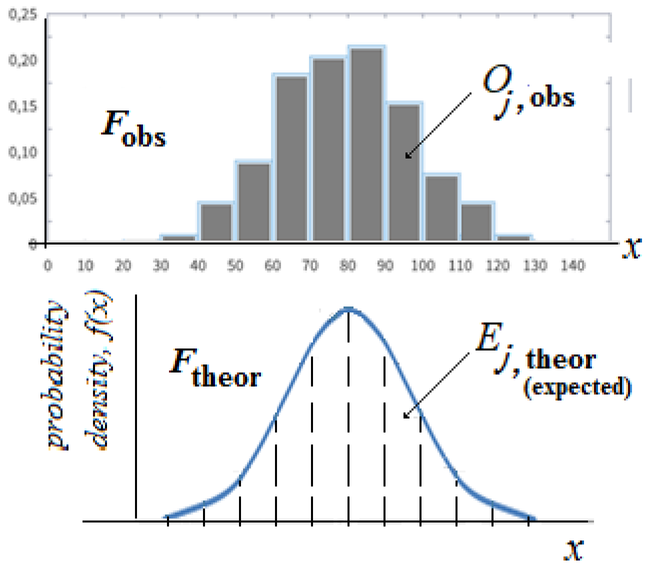
H_1: $F(x) \neq F_{theor}(x)$.

The significance level is given (e.g., $\alpha \leq 0{,}05$).

Pearson goodness-of-fit test is the one which can be used to test coincidence of distribution functions.

To use the test, the data must be presented as a variational series and grouped into classes (bins of equal size $\Delta_1 = \Delta_2 = \ldots = \Delta_i = \ldots = \Delta_k$) with the corresponding observed and theoretical frequencies of falling into a corresponding bin.

For instance, experimental results are grouped into classes, the histogram of distribution is drawn. Do these results correspond to a certain known (theoretical) distribution?



X	Δ_1	Δ_2	…	Δ_i	…	Δ_k
O_i, observed (empiric, experimental) frequency	O_1	O_2	…	O_i	…	O_k
E_i, expected (theoretic) frequency	E_1	E_2	…	E_i	…	E_k

As a statistical test (criterion) a random variable χ^2_{obs} ("chi square observed") is used, which is calculated by formula:

$$\chi^2_{obs} = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i},$$

with k being the number of classes.

In case the null hypothesis is true this random variable is distributed according Pearson's distribution with f degrees of freedom.

The number of degrees of freedom $f = k - 1$, if the hypothesis is simple and $f = k - c - 1$ if the hypothesis is complex(*).

The critical value is found from chi-square distribution table for a given significance level and the number of degrees of freedom, $\chi^2_{crit}(\alpha, f)$.

The observed and critical values are compared.

The result $\chi^2_{obs} < \chi^2_{crit}$ indicates that the null hypothesis is true.

The result $\chi^2_{obs} > \chi^2_{crit}$ means that H_0 should be rejected and H_1 should be accepted.

Population parameters are not used to calculate the chi-square test, so this is a nonparametric test.

(*) A hypothesis is **simple** if the experimental distribution is compared with just one certain distribution (with definite parameters). Then $f = k - 1$.

A **hypothesis is complex** if the experimental distribution is compared with a set of distributions defined by some number c of parameters. E.g., normal distribution is defined by two parameters μ and σ^2, so $c = 2$, and the number of degrees of freedom for the test of goodness-of-fit of some experimental distribution with some indefinite normal distribution

$$f = k - c - 1 = k - 2 - 1 = k - 3.$$

Problem 3.5. At significance level of $\alpha \le 0{,}05$ test the hypothesis that a general population is distributed according the normal law if the empirical (observed) and theoretical frequencies are known:

empirical frequencies	6	13	38	74	106	85	30	14
theoretical frequencies	3	14	42	82	99	76	37	13

Solution:

$$H_0: \; F(x) = F_{theoretical}(x) \Rightarrow (\chi^2_{obs} < \chi^2_{crit})$$

$$H_1: \; F(x) \ne F_{theoretical}(x) \Rightarrow (\chi^2_{obs} > \chi^2_{crit})$$

$$\chi^2_{\text{obs}} = \sum_{j=1}^{k} \frac{(n_{j,\text{obs}} - n_{j,\text{theor}})^2}{n_{j,\text{theor}}} = \frac{(6-3)^2}{3} + \frac{(13-14)^2}{14} + \frac{(38-42)^2}{42} + \frac{(74-82)^2}{82} +$$

$$+ \frac{(106-99)^2}{99} + \frac{(85-76)^2}{76} + \frac{(30-37)^2}{37} + \frac{(14-13)^2}{13} = 7{,}19$$

Table 2. Critical points of χ^2-criterion of Pearson (goodness-of-fit criterion), for different significance levels (α) and numbers of degrees of freedom (*f*)

Number of degrees of freedom *f*	Significance level, %				
	10	5	2	1	0,1
1	2,71	3,84	5,41	6,64	10,83
…	…	…	…	…	…
4	7,78	9,49	11,67	13,28	18,46
5	9,24	11,07	13,39	15,09	20,52
…	…	…	…	…	…
7	12,02	14,07	16,62	18,48	24,32
8	13,36	15,51	18,17	20,09	26,12
…	…	…	…	…	…

From the table for χ^2- distribution of Pearson at a given level of significance $\alpha \leq 0{,}05$ and for a given number of degrees of freedom $f = k - 3 = 8 - 3 = 5$ the critical value of χ^2_{crit} is obtained: $\chi^2_{\text{crit}} = \chi^2_{\text{crit}}(\alpha; f) = \chi^2_{\text{crit}}(0{,}05; 5) = 11{,}07$

$\chi^2_{obs} < \chi^2_{crit}$, the null hypothesis isaccepted: the general population is distributed according the normal law.

The case when χ^2- criterion is used to analyze (four-field) contingency tables is of special interest.

Problem 3.6: The study of the effect of smoking on the risk of developing arterial hypertension is being conducted. For this, two groups of subjects were selected - the first included 70 people who smoke at least 1 pack of cigarettes daily, the second - 80 nonsmokers of the same age. In the first group, 40 people had high blood pressure. In the second, arterial hypertension was observed in 32 people. Accordingly, 30 people in the group of smokers had normal blood pressure (70 - 40 = 30) and in the group of nonsmokers - 48 (80 - 32 = 48).
The hypotheses are put forward:
H_0: smoking has no effect on arterial pressure (no statistically significant differences between observed and expected values) $\Rightarrow (\chi^2_{obs} < \chi^2_{crit})$
H_1: smoking produces effect on arterial pressure (there is statistically significant difference between observed and expected values)$\Rightarrow (\chi^2_{obs} > \chi^2_{crit})$.
The problem is to test the null hypothesis (weather H_0 is true or false).

Solution:
A four-field contingency table can be filled for the initial data:

	Arterial hypertension present (1)	**No arterial hypertension (0)**	**Total**
Smoking (1)	40	30	70
Non-smokers (0)	32	48	80
Total	72	78	150

Each line here corresponds to a specific group of subjects (smokers/non-smokers). Columns show the number of persons with arterial hypertension or normal blood pressure.

The task for the researcher is: are there statistically significant differences between the frequency of people with arterial hypertension among smokers and non-smokers?

To answer this question the Pearson chi-square test should be calculated and compared with the critical value.

The above shown table can be represented as:

	Arterial hypertension present (1)	**No arterial hypertension (0)**	**Total**
Smoking (1)	$A = 40$	$B = 30$	$A + B = 70$
Non-smokers (0)	$C = 32$	$D = 48$	$C + D = 80$
Total	$A + C = 72$	$B + D = 78$	$A + B + C + D = 150$

The expected number of observations for each of the cells of the contingency table (provided that the null hypothesis of the absence of a relationship is true) is calculated by multiplying the sums of the rows and columns and then dividing the resulting product by the total number of observations. The table of expected values is presented below:

	Arterial hypertension present (1)	**No arterial hypertension (0)**	**Total**
Smoking (1)	$(A + B)*(A + C)/(A + B + C + D) = 70*72/150 = 33,6$	$(A + B)*(B + D)/(A + B + C + D) = 70*78/150 = 36,4$	$A + B = 70$
Non-smokers (0)	$(A + C)*(C + D)/(A + B + C + D) = 80*72/150 = 38,4$	$(B + D)*(C + D)/(A + B + C + D) = 80*78/150 = 41,6$	$C + D = 80$
Total	$A + C = 72$	$B + D = 78$	$A + B + C + D = 150$

Then chi-square criterion (observed) can be calculated:

$$\chi^2_{\text{obs}} = \sum_{j=1}^{k} \frac{(n_{j,\text{obs}} - n_{j,\text{theor}})^2}{n_{j,\text{theor}}} = \frac{(40-33{,}6)^2}{33{,}6} + \frac{(30-36{,}4)^2}{36{,}4} + \frac{(32-38{,}4)^2}{38{,}4} + \frac{(48-41{,}6)^2}{41{,}6} =$$

$= 4{,}396$

The number of degrees of freedom f is calculated as:

f = (number of lines – 1)*(number of columns – 1);

for a four-field contingency table $f = (2-1)*(2-1) = 1$.

Table 2. Critical points of χ^2-criterion of Pearson (goodness-of-fit criterion), for different significance levels (α) and numbers of degrees of freedom (*f*)

Number of degrees of freedom, f	Significance level, %				
	10	5	2	1	0,1
1	2,71	3,84	5,41	6,64	10,83
…	…	…	…	…	…

From the table for χ^2- distribution of Pearson at a given level of significance $\alpha \le 0{,}05$ and for a given number of degrees of freedom $f = 1$ the critical value of χ^2_{crit} is obtained: $\chi^2_{\text{crit}} = \chi^2_{\text{crit}}(\alpha; f) = \chi^2_{\text{crit}}(0{,}05; 1) = 3{,}84$.

$\chi^2_{obs} > \chi^2_{crit}$, the null hypothesis is rejected and H_1 is accepted: smoking influences upon the frequency of arterial hypertension.

Problem 3.7. A hypothetical immunological experiment is performed to establish the effectiveness of suppressing the development of a microbial disease when the appropriate antibodies are introduced into the body. In total, 111 mice were involved in the experiment, which were divided into two groups, including 57 and 54 animals, respectively. The first group of mice was injected with pathogenic bacteria, followed by the introduction of blood serum containing antibodies against these bacteria. Animals from the second group served as control - they received only bacterial injections. After some time of incubation, it turned out that 38 mice died and 73 survived. Of the dead, 13 belonged to the first group, and 25 to the second (control). The null hypothesis tested in this experiment can be formulated as follows: the administration of serum with antibodies does not have any effect on the survival of mice. In other words, we argue that the observed differences in the survival rate of mice (77.2% in the first group versus 53.7% in the second group) are completely random and not associated with the action of antibodies.

Solution:

The data obtained in the experiment can be presented in the form of a table:

Group	Perished	Survived	Total
Bacteria + serum	13	44	**57**
Bacteria only	25	29	**54**
Total	**38**	**73**	**111**

To test the null hypothesis formulated above, we need to know what the situation would be if antibodies did not really have any effect on the survival of the mice. In other words, you need to calculate the expected frequencies for the corresponding cells in the contingency table. How to do it? In the experiment, a total of 38 mice died, which is 34.2% of the total number of animals involved. If the administration of antibodies does not affect the survival of the mice, the same percentage of mortality should be observed in both experimental groups, namely 34.2%. Calculating how much is 34.2% of 57 and 54, we get 19.5 and 18.5. These are the expected mortality rates in our experimental groups. The expected survival rates are calculated in a similar way: since 73 mice survived in total, or 65.8% of their total number, the expected survival rates will be 37.5 and 35.5. Let's compose a new contingency table, now with the expected frequencies:

Group	**Perished (expected)**	**Survived (expected)**	**Total**
Bacteria + serum	19,5	37,5	**57**
Bacteria only	18,5	35,5	**54**
Total	**38**	**73**	**111**

Chi-square observed can be calculated:

$$\chi^2_{\text{obs}} = \sum_{j=1}^{k} \frac{(n_{j,\text{obs}} - n_{j,theor})^2}{n_{j,theor}} = \frac{(13-19{,}5)^2}{19{,}5} + \frac{(44-37{,}5n_{j,\text{theor}})^2}{37{,}5} + \frac{(25-18{,}5)^2}{18{,}5} +$$

$$+ \frac{(29-35{,}5)^2}{35{,}5} = 6{,}79$$

Number of degrees of freedom $f = (2 - 1)*(2 - 1) = 1$.

$\chi^2_{\text{crit}}(\alpha; f) = \chi^2_{\text{crit}}(0{,}05; 1) = 3{,}84$.

$\chi^2_{\text{crit}} > \chi^2_{\text{obs}}$

Conclusion: the null hypothesis H_0 is rejected, the alternating H_1 is accepted: the administration of serum with antibodies does have effect on the survival of mice, the observed differences in the survival rate of mice (77.2% in the first group versus 53.7% in the second group) are not random and are associated with the action of antibodies.

G-TEST OF SIGNS.

Another example of nonparametric test is G-test of signs. It makes it possible to establish how much the characteristic values change unidirectionally when the related homogeneous sample is re-measured. The sample size to use test of signs should range from 5 to 300.

Null and competitive hypotheses are formulated,

H_0: the impact has no effect

H_1: the impact is effective.

The level of significance is chosen.

In order to calculate the G-test of signs, it is necessary to fulfill the following algorithm:

1. Create a table of the values of the two samples (an example is given):

x_i(before impact)	12	14	9	6	12	11	13	10	8
y_i(after impact)	10	12	12	9	14	13	13	14	10
$y_i - x_i$	-2	-2	+3	+3	+2	+2	0	+4	+2

2. Calculate the differences (shifts) in pairs, $y_i - x_i$ (third row in the table).

3. Calculate the number of zero shifts (1 zero shift in the example) and exclude zero shifts from consideration.

4. Count the total number of values (no zero shifts). Consider this number as sample size n (in the example $n = 8$).

5. Calculate the number G of "negative" and "positive" shifts. Consider "typical" those shifts the quantity of which is greater and "atypical" those shifts whose number is smaller (in the example given typical shift is positive, and atypical – negative, $G_+ = 6$; $G_- = 2$).

6. **The number of atypical shifts is considered to be the observed (empirical) value of G-test** ($G_{obs} = G_- = 2$).

7. Get the critical value $G_{crit} = G_{crit}(\alpha, n)$ from the table (for the example given $G_{crit}(0{,}05; 8) = 1$).

8. Compare G_{obs} and G_{crit}.

If H_0 is true, $G_{obs} > G_{crit}$.

In case $G_{obs} < G_{crit}$ the null hypothesis is rejected and competitive H_1 is accepted.

In our example $G_{obs} > G_{crit}$ (2 > 1), so H_0 is accepted (the impact is of no effect, the typical shift is not valid).

The critical value of G-criterion of signs can also be calculated by formula for a given sample size as a whole part of A:

$$G_{crit} = [A] = \left[\frac{N-1}{2} - k\sqrt{N+1} \right],$$

with N – sample size and $k = 0{,}98$ for $\alpha \le 0{,}05$.

Problem 3.8.

The action of a factor on arterial pressure was investigated in 100 animals. The increase of pressure after the factor administration was observed 48 times (positive difference took place $n_+ = 48$), the decrease – 44 times (negative differences $n_- = 44$). At significance level of $\alpha \le 0{,}05$ test the null hypothesis H_0 : the impact of a factor is statistically insignificant.

Solution: positive shifts are typical, negative – atypical: $n_+ > n_-$.

$n_- = 44$ is taken as $G_{observed}$.

$G_{critical}$ is calculated as

$$G_{crit} = [A] = \left[\frac{N-1}{2} - k\sqrt{N+1} \right] = \left[\frac{100-1}{2} - 0{,}98\sqrt{100+1} \right] = [39{,}6] = 39 .$$

In case the null hypothesis H_0 (the impact of a factor is statistically insignificant) is true, $G_{obs} > G_{crit}$.

The results obtained: $G_{obs} = 44; \quad G_{crit} = 39 \Rightarrow H_0$ is accepted: the impact of a factor is statistically insignificant.