

**MINISTRY OF AGRICULTURAL POLICY
AND FOOD OF UKRAINE**

Poltava State Agrarian Academy

Department of Selection,
Seed Studies and Genetics

**BASICS OF SCIENTIFIC
RESEARCH IN AGRONOMY**

Methodological instructive regulations for an laboratory work of the
students of the Agronomy Faculty of education and qualification
level - "Bachelor"

Poltava – 2019

UDC 78.00106

The methodological instructive regulations for an laboratory work of the students of the Agronomy Faculty of education and qualification level "Bachelor" are given in.

Authors: PhD of Agriculture, associate professor **Marenych M.M.**,
PhD of Agriculture, associate professor **Yurchenko S.O.**

The reviewer: Doctor of Agricultural Sciences, Professor, Head of the Department of Plant Growing **Shevnikov M.Y.**

Considered at the meeting of the Department of Selection, Seed Studies and Genetics dated February 25, 2019, Record №8.

Recommended for publication by the Methodological Board of the Agronomy Faculty dated February 26, 2019, Record № 6.

METHODOLOGICAL INSTRUCTIVE REGULATIONS TO FULFIL A LABORATORY RESEARCH AND INDIVIDUAL TASKS

PREFACE

In experimental practice, including the agronomic one, it is impossible to speak about the truth of the trial or the establishment of a certain regularity without having conducted an appropriate major experiment. The logical completion of the trial is a mathematical processing of the experimental results.

The methodological instructive regulations are aimed at forming knowledge and practical skills, planning, designing, conducting field experiments, analysing the results, making calculations including modern software as well used by students of the Agronomy Faculty. For this purpose, each work, which provides an example of calculations of the biometric parameters, the truth of the trial, contains the nature of calculations and their solution by means of modern software. This will greatly facilitate the statistical processing of the experimental results and will help complete the diploma paper in a proper way.

Topic 1. TERMS, DEFINITIONS AND REQUIREMENTS TO THEIR USAGE

Block is a part of the repetition, a compact group of several sections of the experiment

Variant of the experiment is a plant, variety, conditions of cultivation, an agrotechnical measure under study

Variability (diversity) is a property of conditional units (plants, yields) to differ in homogeneous populations in the repeated plots of the field experiment

Variation set is a series of data that shows the values of the character that varies in ascending or descending order, and their corresponding frequency

Probability (P) is a degree of objective possibility of the event, the ratio of contributing cases to the total number of possible cases

Exclusion is a part of the record plot which is excluded from the record due to an accidental damage or error

Sample variance is a figure of the variation, variability of the character

Experimental plot is an elementary part of the field experiment, a part of the study area which has a certain size, shape and is meant to place a separate variant

Record plot is a part of the area of the experimental plot intended for the yield records without lateral and finite protective strips

Reliability of the experiment is a correctly planned, implemented scheme and method for holding the experiment, the correct choice of an object, conditions of running the experiment, and a method for statistical processing of the results

Protective strips are the boundary parts of the plots that are not subject to the record and prevent the influence of neighbouring variants, damage to the record part from accidental factors; they are also a turn space for machines and implements

Significance (importance) is a measure of an objective possibility (risk) to make a wrong conclusion while evaluating the results of the experiment. They use a 5% level of significance where the risk of making a mistake is 5% and a 1% level where this risk is equal to 1%

Control (standard) is one or several variants the experimental variants are compared to

Coefficient of variation is a relative figure of the variability of the character, it is the ratio of a standard deviation to the mean in percentage terms

The least significant difference (LSD) is a figure that shows the limit of possible chance differences in the experiment.

Experimental error is a measure of the difference between the results of the screening tests and the true value of the measured quantity

Repetition is a part of the area of the experimental plot with a full set of the

experimental variants. The repetition taken separately is, in fact, the experiment reduced in scope, and it allows to do all possible comparisons between the variants

Replication is the number of homonymous plots of each variant in the field experiment; in time terms, it is the number of years of the test

Test sowing is a broadcast sowing of one culture to increase soil fertility of the plot chosen for the experiment layout

Reconnaissance (test) crop is a broadcast sowing of one culture preceding the layout of the field experiment, it is conducted to evaluate the degree of homogeneity (by the way of fractured records of the yield) of soil fertility

Typology (representativeness) of the experiment is the compliance of the conditions of its operation with the soil, climatic and agronomical conditions of the agricultural production of this zone

The number of degrees of freedom ($v=n-1$) is the number of freely varying values; in the simplest case, it is equal to the number of all the observations minus 1

SHORT SYMBOL INDICATOR

μ is an assembly average

σ^2 is the variance of the general population

σ is a standard deviation of the general population

s^2 is a sample variance, a mean square

s is a sample standard deviation, a mean square deviation

V is a coefficient of variation

s_x is a mean error of the sample average

d is the difference between the sample means

l is the number of variants

n is the replication, a sample scope

N is the total number of observations

v is the number of degrees of freedom

P is probability

t_{05}, t_{01} is a conventional true value of Student's criteria

F is the value of Fisher's criteria

χ^2 is the value of Pearson's chi-squared test

H_0 is a null hypothesis

C is a corrective factor

C_y, C_p, C_z are the sums of squared deviations for different sources of the variation in the dispersion analysis

LSD is the least significant difference

r is a correlation coefficient

s_r is a linear correlation coefficient's error

b_{yx} is a regression coefficient

\div is a confidence interval

Σ is the sum, the summation sign

Topic 2. PLANNING A SCIENTIFIC RESEARCH

Purpose of the class: To learn how to plan full and incomplete factorial schemes of the experiments, to plan sampling records.

Scientific research by means of the field experiment involve three stages: planning, conducting and processing the results.

Planning is the definition of a task and objects of the research, the development of a scheme of the experiment, the choice of the optimal land of the field experiment. The mistakes made in planning cannot be corrected in the future even with the use of the most advanced computer technologies.

The preparation of research should be preceded by the choice of a topic, the definition of tasks and objects of the research, the study of the current state of the question, the advancement of a working hypothesis, the development of a scheme and methodology of the experiment.

The next stage is the study of literature, the advancement of a hypothesis or hypotheses – a starting point of the scheme of the future research and work program. The schemes of experiments, the main elements of methodology and technique of the experiment, the observations and records are indicated in the program. The main issue of planning is the development of a rational field experiment. When planning one-factor experiments, it should be remembered that the options may differ in quality: the study and evaluation of varieties, sowing methods, soil cultivation, etc., and also in quantity: the dose of fertilisers, watering and sowing rate, etc. If the variants differ in quality, then it is important to use the principle of a single difference, to choose a standard correctly, to determine other conditions of the experiment. In addition, it is necessary to correctly set a unit of variation and the number of gradations for the experiments with quantitative gradations. As a rule, 5-8 factor doses (gradations) under study are enough.

A feature of a multiple factor experiment is the ability to determine the effect of factors, the nature and range of their interaction under simultaneous application. In the multiple factor experiment, we must adhere to the principle of factoriality which provides for testing all possible combinations of factors and their gradations. The effect and interaction of both quantitative and qualitative factors and their gradations are studied in the factorial experiments. Zero gradation for a quantitative factor is its absence, for a qualitative one, this is a control variant – a standard scheme of soil cultivation or a variety.

The complete scheme of the multiple factor experiment is calculated by the M^n formula where the base is the number of gradations, the degree stands for the number of factors. If the number of gradations in the factors is different, then the total number of the variants is calculated as a product of gradations of all the factors. To make the scheme represent a full factorial experiment (FFE), one constructs a matrix. It is almost impossible to conduct an FFE with a large scheme in terms of production. Therefore, one should reduce the number of variants, but so that you do not lose the necessary information. In this case, an incomplete factorial scheme (IFS) is planned. To do this, one uses the methods of a conditional factor, written cubes, designing the schemes from fragments of the cube. One takes the initial and average gradation from several ones.

An important element of planning the experiment is the frequency of records and observations. In this case, it is necessary to determine which observations and records should be conducted, in which terms, which is the optimal volume of samples. Purposefulness is the most important requirement of observation. The sample size is calculated by the formula:

$$n = t^2 \frac{V^2}{\Delta\%},$$

where t is the value of Student's criterion, V is the variation coefficient, $\Delta\%$ is a tolerate relative error.

$$V = \frac{s}{\bar{x}} 100, \quad s = \frac{X \max - X \min}{6},$$

where s is a standard deviation, \bar{x} is an arithmetic mean.

The arithmetic mean can be calculated approximatively:

$$\bar{x} = \frac{X \max + X \min}{2},$$

Records and observations over growing conditions:

1. Soil characteristic
- type, forecrop, number of fertilisers;

- agrochemical characteristic of the soil: the content of humus, active forms of phosphorus, potassium, acidity;
 - the number of fertilisers applied to the culture under study.
2. Monthly or decade factors of weather conditions compared to the average perennial ones.
 3. Sowing phytosanitary condition; the damage by pests, affection by diseases, quantitative and species composition of weeds per 1 m², efficiency of herbicides, etc.

Records, observations and analyses of plants: the number of sprouts, size of the protective zone, phenological observations, plant stand, yield.

The results of the research are analysed by statistical methods using a computer technology and an appropriate software. The economic efficiency of the experiment can be determined by drawing up a technological map.

The sample volume to determine figures of the qualitative variability can be calculated by the formula

$$N_{0,95} = t_{0,95} \frac{pq}{s_p^2},$$

where p is the fraction of the feature's presence, q is the fraction of the absence, s_p^2 is the error of the fraction chosen by the researcher for the field observations -0.05 ... 0.1; $p = n / N$, where n is the number of individuals with this feature, N is the total number of objects, $q = 1 - p$

Questions for self instruction:

1. *Special aspects of multiple factor experiments*
2. *Determine the number of variants in the multiple factor experiment*
3. *Make a scheme of a full factorial experiment*
4. *Choose the experimental move*
5. *Make a scheme of an incomplete factorial experiment*
6. *Determine the sample size for the quantitative and qualitative variability of the characters*

Topic 3. METHODS FOR PLACING THE VARIANTS IN THE EXPERIMENT

Purpose of the class: To learn how to place the variants of the experiment depending on the goal and task of the research, to master the randomisation technique for the variants of the experiment

Standard placement of variants – a standard is placed over one (iamb method) or two (dactyl method) research variants. Thus, each variant is compared to its own control.

These methods have significant disadvantages – awkwardness and unreasonable use of the land area and drawbacks related thereto.

Systematic placement of variants – when the order of alternating variants in the repetition corresponds to a certain system.

Randomised placement of variants is a random placement within each repetition. Randomisation is carried out by means of a blind draw or a table of random numbers. This method is divided into:

- *a method of full randomisation* – an approximate randomisation of conditions of the experiment. It is effective with a small number of factors when one cannot take into account the territorial variability of soil fertility;

- *a method of randomised repetitions* – the most common in the world practice. In each repetition, the variants are distributed randomly. It is important for each repetition to have the highest possible level of soil fertility. Using the method, it is reasonable to have no more than 20 variants. If there are more than 8 variants, it is desirable to have one or two sections of the standard in each repetition.

The Latin square. The plot of square shape is divided into the horizontal and vertical directions into as many rows and columns as many variants are there in the experiment. Any row and column must have a full set of variants. The number of repetitions must be equal to the number of variants. If the number of variants is more than 7, then the method of *the Latin rectangle* is used. In this case, the number of variants should be a multiple of the number of repetitions. For example, you can place an experiment with 6, 9, 12, 18 variants in a 3-time repetition.

The method of split areas is mainly used for multiple factor experiments. In this case, the plots of the first order are divided into the plots of the second order and so on. The main plots and subplots are randomised. The main plots are randomised independently in each repetition, and the subplots are randomised separately for each large plot or even an entire repetition.

Questions for self instruction:

1. *Advantages and disadvantages of the standard method for placing the variants in the experiment*

2. *Advantages and disadvantages of the systematic method for placing the variants in the experiment*

3. *Methods of the randomised placement of the variants*

4. *The Latin square method*

5. *The Latin rectangle method*

6. *Perform a randomisation of the field experiment using a table of random numbers by means of the method of complete randomisation, randomised repetitions, the Latin square, the Latin rectangle*

Topic 4. YIELD RECORDS

Purpose of the class: To study the methods of yield records of the main and secondary production of agricultural crops

Yield records are one of the main works of the researcher. 1-2 days before the harvest, it is necessary to carefully examine the whole experiment, to restore the borders of all the plots, to remove the labels and foreign objects from the area, having allocated exclusions to the record plots. The reason for the exclusion of the entire plots may be damage to the massif during a natural disaster, damage to crops caused by cattle or a theft, decrease in number of tilled crops during the inter-row cultivation, the researcher's mistake when laying out or conducting the experiment. The plot is completely discarded when the exclusions occupy 50% of its area because one can decrease the area of the record plots not more than by 30-40%.

The harvest on all the experimental plots in the experiment or within the repetition should be collected in one day using one harvesting unit.

Yield records of the spiked cereals of a solid planting.

When planning the area of the record plot for the combining harvesting, it should be taken into account that the higher the yield of the crop is, the less the record area of the plot may be and vice versa. After threshing several plots, the bags of grain are weighed directly in the field, on the threshing floor or in another room. Once the grain is weighed, an average sample of 1-2 kg is taken from each bag to determine the moisture, litter and qualitative figures.

1. The bunker weight of the crop from the plot is converted to a hectare area using the coefficient over the area (C_a) which is found by the formula:

$$C_a = \frac{10000m^2}{A},$$

where 10000 m² stand for the area of 1 ha, m²; A is the area of the record plot, m².

Having multiplied the weight of the crop from the plot, determined in kilograms, by the conversion coefficient over the area and having divided the results by 100 to convert kilograms to centners, the bunker yield in c / ha is obtained.

2. The bunker yield is converted to 100% seed purity in a manner of multiplying by the percentage of pure grain and dividing it by 100. The percentage of purity is determined on the basis of analysing the grain samples of 500 g in a double replication.

3. The yield of clean harvested grains is converted to a standard 14% moisture by the formula:

$$Y = \frac{A(100 - B)}{100 - 14},$$

where Y is the yield of the pure grain at a standard moisture, c/ha; A is the yield of the pure grain at a field moisture, c/ha; B is the moisture content of grain at the harvest time, %; a standard moisture for grain crops. In this formula $\frac{100 - B}{100 - 14}$, the ratio is a conversion coefficient of 14% grain moisture.

The moisture of grain is determined by the formula:

$$B = \frac{b \times 100}{C},$$

where B is the grain moisture, %; b is the weight of evaporated water from a sample bottle with grain, g; C is the weight of the grain quantity in a sample bottle before desiccation, g.

Yield records of the secondary production of close-growing grain crops are determined by the ratio of the straw to grain. The harvest of pulse crops has usually a separate collection. If it is impossible to keep yield records of the close-growing grain crops from the whole record plot, the method of trial bundles is used.

The yield of grain and straw from the plot is determined by the formula:

$$Y = B \frac{D}{C},$$

where Y is the yield of air-dry grain (straw), kg; B is the weight of grain (straw) in a bundle after desiccation, kg; D is the weight of all the products from the plot including the weight of bundles at the time of their selection, kg; C is the weight of a trial bundle at the time of selection, kg.

The yield from the sample plots is calculated in total from each plot with a recalculation according to the formula:

$$Y = \frac{v \times 100}{p},$$

where Y is the yield, c/ha; v is the total weight of the yield of an appropriate production, kg; p is the total space of sample areas on the plot, m²; 100 is the number on dividing 10000 (1 ha area in m²) by 100 (to convert kg to centners).

Corn is hand-harvested, the corncobs from the whole record plot are snapped off and weighed.

1. The weight of the corncobs from the plot per 1 hectare is recalculated using the same conversion factor.

2. The yield factor of grain from the corncobs is determined by the ratio where the numerator will be the weight of grain from 20 peeled typical corncobs selected during the harvest, and the denominator is the weight of the corncobs before peeling.

Sunflower is hand-harvested or collected by mechanical means. If the record plot is 100 m² or less, the harvest is hand-picked in two stages: during the browning, sunflowers are cut down and stuck on the stalk pointing seeds down (to prevent damage caused by birds). Once the sunflowers are dried down, they are threshed, the seeds are weighed, the samples to determine the purity, moisture, and qualitative factors are taken from them.

Yield records of the root crops of sugar and fodder beets, carrots, potato tubers. If there is a lot of tare on the root crops or tubers, one takes special samples, weighing 20-30 kg, from the plot out of the weight harvest. Once the samples are weighed, they are washed, dried a little and weighed again.

Hay yield of annual and perennial grasses can be accounted in two ways: continuously throughout or by means of trial bundles. In a continuous way, the dried mass is weighed from the whole record area, and the method of trial bundles means that a trial bundle of at least 4-5 kg is taken out of the mowed hay of herbage, then all the green weight is scaled from the record area including the trial bundle.

The yield of the green weight or hay in pastures is accounted every time before grazing the livestock using the following methods:

1. One weighs the mowed mass in a variant from three record plots (2 m²) that are allocated on the plot in new places every time before grazing.
2. One weighs the mowed mass from the half of the record area, and on the second half, the mowed mass is also weighed before the next grazing.
3. The whole record area of the variant is divided into equal parts, multiples of grazings, a new plot is allocated to keep the yield records every time before grazing.

Yield records of the legume seeds and grasses. The yield is determined after the combining thresh of the windrows from the record area for plot harvesting in the stationary areas covered with tarpaulin. Once the collected yield is converted to the hectare area and 100% purity, the yields are converted to a standard 13% moisture, the perennial and annual grasses – to a 15% respectively.

Yield records of the textile crops (fiber flax and hemp) can be carried out using both a continuous weighing of the whole record area and trial bundles. In both cases, the yield of seeds and straw is determined. The

yield of the straw is converted to a 16% moisture and the seeds – to a 12% respectively.

To determine the fiber yield, the sample weight should be within 18-20 kg after threshing.

Questions for self instruction:

- 1. Yield records of the cereal and pulse crops*
- 2. Yield records of the secondary production*
- 3. Yield records of the root crops*
- 4. Yield records of the grasses*
- 5. Determination of the yield of textile crops.*
- 6. Methods for determination of the moisture content.*

Topic 5. DESIGNING THE PROGRAM AND RESEARCH METHODS FOR THE DIPLOMA PAPER. REGULATIONS

Purpose: To master the methodology for designing a research program, the procedure of reporting documentation.

The diploma paper is performed on the basis of the detailed study of scientific literature on specialty (monographs, articles, abstracts of theses, periodical literature including foreign, normative literature, etc.) and, most importantly, on the results of the independent research and experimental work.

The structure of the diploma paper: the title sheet; the content; the introduction; the literature review on the subject under study; the experimental part; measures on labour protection; the ecological expertise; conclusions and recommendations for production; the reference list; annexes.

Introduction. (3 pages). The introduction describes the main tasks of agriculture, outlines the current state of the problem (issue) the work is devoted to, features the novelty and relevance of this paper and substantiates its significance for production.

1. Literature review. (10-12 pages). This section covers the state of knowledge of the issue the topic of the diploma paper is devoted to.

When presenting the literature review, it is obligatory to have a reference to the sources included in the section "List of references".

2. Experimental part. (22-30 pages). This section includes the following subsections: the characteristics of the place of experiments; weather and climatic conditions during the research; research methodology of the

experimental work; results of the experimental work; economic evaluation of the work results and labour protection.

2.1. Characteristics of the place of experiments. A farm unit where the experimental work is performed as well as its geographic location are characterised in this subsection, the economic activity, planting acreage, yield, the quantity and quality of the commercial yield, the level of profitability, the degree of mechanisation of a certain crop production are described here.

2.2. Soil and climatic conditions during the research. The soil and climatic conditions (monthly precipitation and average monthly temperature that are analysed in comparison to the average long-term data) are characterised at full length in this subsection. The meteorological data should be taken at the nearest meteorological station. The soils (the type of soil, its texture, water and physical properties), local topography are described here.

2.3. Research methodology. The variants of the experiment, its procedure (the field experiment scheme, plot size, repetitions, the whole area of the experiment) are clearly represented here.

Concomitant observations, analyses and records, their procedures are described. The varieties studied in the experiment as well as the technology of the crop cultivation are definitely characterised.

2.4. Results of the experimental work. The data of the experimental research (the obtained data is presented in the form of tables, drawings, graphs, photographs, diagrams) is described in this section.

The analysis of the experimental research must be critical, accompanied by the analysis of the phenomena, factors, revealed regularities, elaboration of the measures on labour protection. The statistical data processing is carried out.

3. Economic evaluation of the results of scientific research. (4-5 pages). The economic effectiveness of the research results recommended for the implementation is analysed in this section. The main figures of the economic efficiency are the following: a raise in the gross production, a reduction of the production cost, a raise in the net income, an increase of the level of the output profitability. The students perform this part under the supervision of the management faculty counselors.

4. Measures on labour protection. (3-4 pages) When writing this section, the student is consulting the teacher who is reading the appropriate course. The analysis of the safety status, workplace sanitation and health,

fire-prevention measures is carried out here. Suggestions on safety arrangements and better working conditions are made here.

5. Environmental expertise (4-5 pages). This work is carried out under the supervision of the counselor of the Department of Ecology and Botany.

Conclusions and suggestions on the production (3-5 pages). This section is a summary part of the paper. It is considered to be done on the basis of the agronomic, economic and mathematical analyses. It covers the results of the research conducted and recommendations for the work. The conclusions are given as separate paragraphs within one graph for each of the sections.

Once this section is completed, the student puts one's own signature and the date of execution of the work.

The list of references includes 25-50 names of the used sources, various normative documents that were used while working on the diploma paper. The list is compiled in the alphabetical order (domestic authors first, foreign authors after) and numbered (Annex 2).

Annexes. This subsection includes instructions, illustrations, tables of the additional digital data and the result of the dispersion analysis.

Topic 6: STATISTICAL FIGURES OF QUANTITATIVE AND QUALITATIVE VARIABILITY

Purpose of the class: To master the methodology of determining the figures of quantitative and qualitative variability and to learn how to make calculations with the help of personal computer software

Statistical figures of the quantitative variability

The quantitative characters include factors that can be characterised in terms of quantity: yield, height and weight of plants, protein and gluten content, etc.

The main statistical characteristics include the mean, the variance (s^2), a standard deviation (s), the coefficient of variation (V), a relative error of the mean (s_x), a confidence interval for the mean ($\bar{x} \pm t_{sx}$). The mean for small and large samples is respectively equal to:

$$\bar{x} = \frac{\sum X}{n}$$

$$\bar{x} = \frac{\sum fX}{n}$$

$$s^2 = \frac{\sum X^2 - (\sum X)^2 : n}{n - 1}$$

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

$$V = \frac{s}{\bar{x}} 100 \qquad s_x = \frac{s}{\sqrt{n}} = \sqrt{\frac{s^2}{n}} \qquad s_{\bar{x}\%} = \frac{s_{\bar{x}}}{\bar{x}} 100$$

The statistical figures of the quantitative variability can be calculated with the help of modern PCs' software. The program of Excel' or STATISTICA is used for this purpose.

Example: To determine the figures of the variability of the stalk length of the second generation's hybrids of winter wheat:

The calculations in Excel' will be as follows:

75,4	Column	
85,3		
69,8	Mean	71,75
56,3	standard	3,616482
78,3	Median	75,15
74,9	Mode	78,3
85,4	Standard deviation	11,43632
59,6	Sample variance	130,7894
78,3	Excess	-1,20013
54,2	Asymmetry	-0,48484
	Interval	31,2
	Minimum	54,2
	Maximum	85,4
	Sum	717,5
	Count	10
	The biggest (1)	85,4
	The least (1)	54,2
	Reliability level (95,0%)	8,181058

The left column of the table contains the source data. Similar data is provided by STATISTICA statistics package. As seen from the example above, the software provides a larger list of statistical parameters that can also describe the variation series.

The main parameter of variability is the variance, but if one wants to evaluate the variability with the coefficient of variation, then it is figured with the help of a calculator.

Statistical characteristics of the qualitative variability

The figures of the qualitative variability include those that cannot be measured in quantity: illness, colour, foetus form, presence or absence of a feature. If the population under study is represented by two gradations (with a feature being present or absent), then such a feature is called alternative.

Statistical characteristics are calculated according to the formulas where N is a sample scope, k is the number of feature gradations, t is a theoretical value of Student's criterion. To determine V , it should be taken into account that the maximum variability is equal to 0.5 (50%) for two gradations of the feature, to 0.33 (33%) for three, to 0.25 (25%) for four, to 0.20 (20%) for five, to 0.167 (16.7%) for six. Given that $k = 2$

$$s = \sqrt{pq} \qquad s = \sqrt[k]{p_1 \times p_2 \times \dots \times p_k}$$

Given that $k > 2$ $p_1 = n_1/N$; $p_2 = n_2/N$; $p_k = n_k/N$. The fractional error is:

$$\lg s = \frac{\lg p_1 + \lg p_2 + \dots + \lg p_k}{k} \qquad V = \frac{s}{s_{\max}} 100 \qquad sp = \frac{s}{\sqrt{N}} = \sqrt{\frac{pq}{N}}$$

The confidence interval for the fraction is $p \pm ts_p$

Topic 3: DISPERSION ANALYSIS

Purpose of the class: To master the method for running a dispersion analysis with the help of software

The dispersion analysis was developed and introduced into the research practice by an English scientist R. A. Fisher who also formulated the law of distribution of the variance ratio:

$$F = \frac{s_1^2}{s_2^2}$$

where s_1^2 is the mean square of the sample means, s_2^2 is the mean square of the objects.

The point of the dispersion analysis is to decompose the total sum of the deviation squares and the total number of the degrees of freedom on the components corresponding to the structure of the experiment and to evaluate the significance of the action and interaction of the factors under study.

The theory of research planning and the statistical analysis are based on the principles of randomisation. In this case, the dispersion analysis gives the correct error of the experiment. If the experiment is not randomised, then you can get a biased experimental error, and the criteria of the

significance lose their validity. Therefore, they cannot be used as arguments for the statistical proof of effects.

One-factor experiment

Example: To perform a dispersion analysis of winter wheat yields in a one-factor field experiment. Null hypothesis: all the deviations in the experiment are random.

The experiment data is written down into the table:

The yield of winter wheat in the experiment

Variants	Yield, dt/ha				V sums	Mean
1 st	41,5	47,0	44,0	43,0	175,5	43,88
2	49,6	54,2	49,0	51,0	203,8	50,95
3	58,6	57,2	60,0	61,3	237,1	59,28
4	42,1	41,2	47,6	41,3	172,2	43,05
5	56,5	57,0	58,0	51,2	222,7	55,68
P sums	248,3	256,6	258,6	247,8	1011,3	50,57

Then the calculations are made in the following sequence:

The total number of observations $N = \sum n = 4 \cdot 5 = 20$

The corrective factor $C = (\sum X)^2 : N = 1011,3^2 : 20 = 51136,385$

$C_Y = \sum X^2 - C = (41,5^2 + 44^2 + \dots + 51,2^2) - 51136,385 = 911,15$

$C_V = \sum V^2 : n - C = (175,5^2 + 203,8^2 + 237,1^2 + 172,2^2 + 222,7^2) : 4 - 51136,385 = 813,42$

$C_P = \sum P^2 : l - C = (248,3^2 + 256,6^2 + 258,6^2 + 247,8^2) : 5 - 51136,385 = 18,67$

$C_Z = C_Y - C_V - C_P = 911,15 - 813,42 - 18,67 = 79,06$

After the calculation, one should fill in the table:

Results of the dispersion analysis

Dispersion	Sum of squares	Degree of freedom	Mean square	F_f	F_{05}
Total dispersion	911,15	19	—	—	—
That of variants	813,42	4	203,36	30,86	3,26
That of repetitions	18,67	3	6,22	0,94	3,49
Error dispersion	79,06	12	6,59	—	—

The theoretical value of Fisher's criterion is found in Table 2 of Annexes on account of four degrees of freedom for the dispersion of variants and 12 degrees for the error dispersion. If $F_f > F_{05}$, then it means there are significant differences in the variants in the experiment, and the null hypothesis is rejected.

Then one calculates the experimental error and the error of the means' difference:

$$s_x = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{6,59}{4}} = 1,28dt$$

$$s_d = \sqrt{\frac{2s^2}{n}} = \sqrt{\frac{2 \times 6,59}{4}} = 1,82dt$$

The least significant difference for a 5% (or a 1%) level of significance in absolute and relative figures:

$$LSD_{05} = t_{05} s_d = 2,18 \cdot 1,82 = 3,97 dt$$

$$LSD_{05} = \frac{t_{05} s_d}{\bar{x}} 100 = \frac{3,97}{50,57} 100 = 7,85\%$$

The values of t_{05} are taken from the table of values of Student's criterion for 12 degrees of freedom of residue (error). The results of the statistical processing are written down into the table:

Variant	Yield	Contrast with a standard		Group
		dt	%	
1	43,88	–	–	
2	50,95	7,07	16,11	I
3	59,28	15,4	35,1	I
4	43,05	–0,83	–1,89	II
5	55,68	11,8	26,89	I

The deviation of the average yield from the standard with a positive sign bigger than LSD_{05} is referred to the first group; that with a positive or negative sign which does not go beyond LSD is referred to the 2nd group, and the deviation with a negative sign bigger than the absolute value of LSD is referred to the 3rd group. The variance analysis performed with the Excel' software is as follows:

One-factor dispersion analysis

results

Groups	Count	Sum	Mean	Dispersion
Series 1	4	175,5	43,875	5,39583
Series 2	4	203,8	50,95	5,39666
Series 3	4	237,1	59,275	3,12916
Series 4	4	172,2	43,05	9,36333
Series 5	4	222,7	55,675	9,28916

Dispersion analysis

Source of variation	SS	df	MS	F	P-value	critical F
Between groups	813,423	4	203,3557	31,2142	4,11E-07	3,055568
Within groups	97,7225	15	6,514833			
Altogether	911,1455	19				

The calculations made in the STATISTICA program are as follows:

Summary of all the Effects; design: (noname.sta)1-VAR1

Effect	df Effect	MS Effect	df Error	F	p-level
1	4	203,3557	6,514833	31,21427	,000000

LSD calculations should be made using a calculator, although one can enter the necessary formulas in the spreadsheet. If required, in STATISTICA, one can also calculate the mean value of the variants, LSD (by LSD test), Scheffe and Duncan's criteria of significance which are also LSD evaluations. In this case, the program marks significant deviations of the variants from each other in red, and the most significant difference in these criteria is placed in the upper line.

Dispersion analysis of a two-factor experiment

Purpose of the class: To learn how to carry out a dispersion analysis of a two- and multiple factor field experiment

EXAMPLE 3: *The effect of three irrigation gradations (0 – without irrigation, 1 – moderate, 2 – strong) and four nitrogen doses (0 – without nitrogen, 1 – 30 kg / ha, 2 – 60 kg / ha, 3 – 90 kg / ha) on the yield of winter wheat is studied in a two-factor experiment set in four randomised repetitions. Conduct a dispersion analysis of the experimental results.*

As in the example of a one-factor experiment, the experimental data is written down into the table, and then the calculations are made:

$$N = l_a \cdot l_b \cdot n = 3 \cdot 4 \cdot 4 = 48$$

$$C = (\sum X)^2 : N = 2003^2 : 48 = 83583,52$$

$$C_Y = \sum X^2 - C = (29^2 + 30^2 + 35^2 + \dots + 58^2) - 83583,52 = 4534$$

$$C_p = \sum P^2 : l - C = (510^2 + 502^2 + 497^2 + 494^2) : 3 \cdot 4 - 83583,52 = 12,23$$

$$C_v = \sum V^2 : n - C = (129^2 + 128^2 + \dots + 237^2) : 4 - 83583,52 = 4235,23$$

$$C_z = C_Y - C_p - C_v = 4534 - 12,23 - 4235,23 = 286,54$$

The effect of irrigation and nitrogen doses on the wheat yield

A irrigation	B nitrogen doses	Repetitions				V sums	Mean
		I	II	III	IV		
0	0	29	30	35	35	129	32,25
	1	30	30	30	38	128	32
	2	28	30	28	29	115	28,75
	3	30	29	28	28	115	28,75
1	0	42	39	38	31	150	37,5
	1	50	49	43	47	186	46,5
	2	49	48	50	47	194	48,5
	3	54	52	50	49	205	51,25
2	0	40	41	41	47	169	42,25
	1	52	45	48	43	188	47,0
	2	48	48	46	45	187	46,75
	3	58	61	60	58	237	59,25
P sums		510	502	497	494	2003	41,73

The following table is compiled after:

Determination of the main effects and interactions

A irrigation	Nitrogen dose, N				A sums
	0	1	2	3	
0	129	128	115	115	487
1	150	186	194	205	735
2	169	188	187	237	781
B sums	448	502	496	557	2003

$$C_A = \sum A^2 : l_{Bn} - C = (487^2 + 735^2 + 781^2) : 4 \cdot 4 - 83583,52 = 3126,17$$

When $(l_A - 1) = (3 - 1) = 2$ degrees of freedom

$$C_B = \sum B^2 : l_{An} - C = (448^2 + 502^2 + 496^2 + 557^2) : 3 \cdot 4 - 83583,52 = 497,56$$

When $(l_B - 1) = (4 - 1) = 3$ degrees of freedom

$$C_{AB} = C_v - C_A - C_B = 611,5$$

When $(l_A - 1)(l_B - 1) = (3 - 1)(4 - 1) = 6$ degrees of freedom

Results of the dispersion analysis

Dispersion	Sum of squares	Degree of freedom	Mean square	F _f	F ₀₅
Total	4534	47	—	—	—
Repetitions	12,23	3	—	—	—
Irrigation A	3126	2	1563	179,86	3,30
Nitrogen B	497,56	3	165,85	19,09	2,90
A-B interaction	611,5	6	101,92	11,73	2,40
Residue (error)	286,71	33	8,69	—	—

$$s_d = \sqrt{\frac{2s^2}{n}} = \sqrt{\frac{2 \times 8,69}{4}} = 2,08dt$$

$$LSD_{05} = t_{05} \cdot s_d = 2 \cdot 2,08 = 4,16 dt.$$

The evaluation of the significance of the main effects and interactions upon LSD_{05}

For A factor:

$$s_d = \sqrt{\frac{2s^2}{nl_b}} = \sqrt{\frac{2 \times 8,69}{4 \times 4}} = 1,04$$

$$LSD_{05} = 2 \cdot 1,04 = 2,08 dt$$

For B factor:

$$s_d = \sqrt{\frac{2s^2}{nl_A}} = \sqrt{\frac{2 \times 8,69}{4 \times 3}} = 1,20dt$$

$$LSD_{05} = 2 \cdot 1,20 = 2,40 dt$$

The summary table is compiled after:

The effect of irrigation and nitrogen doses on the winter wheat yield

A irrigation	B nitrogen doses				Mean upon A factor ($LSD=2,08$)
	0	30	60	90	
Without irrigation	32,25	32	28,75	28,75	30,44
Moderate	37,5	46,5	48,5	51,25	45,94
Major	42,25	47,0	46,75	59,25	48,81
Medium upon B factor ($LSD=2,4$)	37,33	41,83	41,33	46,42	41,73

The dispersion analysis carried out in Excel' will be as follows:

RESULTS	Count	Sum	Mean	Dispersion
Row 1	12	510	42,5	118,4545
Row 2	12	502	41,83333	109,2424
Row 3	12	497	41,41667	100,2652

Row 4	12	497	41,41667	85,17424		
Column 1	4	129	32,25	10,25		
Column 2	4	128	32	16		
Column 3	4	115	28,75	0,916667		
Column 4	4	115	28,75	0,916667		
Column 5	4	150	37,5	21,66667		
Column 6	4	189	47,25	9,583333		
Column 7	4	194	48,5	1,666667		
Column 8	4	205	51,25	4,916667		
Column 9	4	169	42,25	10,25		
Column 10	4	188	47	15,33333		
Column 11	4	187	46,75	2,25		
Column 12	4	237	59,25	2,25		
Dispersion analysis						
<i>Variation source</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F critical</i>
Rows	9,4166	3	3,13888	0,37182	0,773841	2,891568
Columns	4265,91	11	387,810	45,9386	8,59E-17	2,093252
Error	278,583	33	8,44191			
Altogether	4553,91	47				

A similar result can be obtained using the STATISTICA 5.0 ('99) program. However, the latter has a significant advantage since it helps to run an analysis of the experiment with a large number of factors and to immediately evaluate it according to the same criteria as a one-factor experiment, that is, to determine the least significant difference. In addition, this program allows to evaluate both the effect of each factor separately and the interaction effects in more detail.

Tasks for the independent work:

1. *The essence of the dispersion analysis.*
2. *The determination of the least significant difference.*
3. *Checking the calculations of the dispersion analysis.*
4. *The interpretation of the results of the statistical processing of the experimental results.*
5. *Conduct a dispersion analysis of a one-factor experiment.*
6. *Conduct a dispersion analysis of a multiple factor field experiment using the STATISTICA statistic package*

Topic: CORRELATION AND REGRESSION

Purpose of the class: To master the methodology of correlation and regression analysis

In agronomic experiments, it is often necessary to deal with correlation links. Correlation can be linear and curvilinear according to the form, direct and inverse according to the direction, simple when the dependence between two characters is studied, and plural when three or more characters are studied. To evaluate the strength of the correlation, one uses its coefficient.

The linear correlation dependence is expressed by the equation of the straight line $Y = a + bx$ which is called the Y regression equation for X. The linear regression is a dependence when, at any value of X, its same increments cause the same function changes. If given the same increment of the argument, the function has different variations, then the regression is called curvilinear.

The value of the correlation coefficient lies in the range from -1 to +1, it is a dimensionless number which is calculated by the formula:

$$r = \frac{\sum (X - \bar{x})(Y - \bar{y})}{\sqrt{\sum (X - \bar{x})^2 \sum (Y - \bar{y})^2}} = \frac{\sum XY - (\sum X \sum Y) : n}{\sqrt{(\sum X^2 - (\sum X)^2 : n)(\sum Y^2 - (\sum Y)^2 : n)}}$$

It is believed that the correlation between the characters is weak given that $r < 0,3$, is average given that $r = 0,3 \dots 0,7$ and is strong given that $r > 0,7$. Although the degree of the characters' association is most fully characterised by the determination coefficient $d = r^2$ which shows the proportion (%) of changes that depend on the factor under study.

To evaluate the reliability of the correlation coefficient, one calculates its error and criteria of significance:

$$sr = \sqrt{\frac{1 - r^2}{n - 2}}$$

$$t_r = \frac{r}{s_r}$$

If $t_f > t_{\text{teor}}$, then the correlative relationship is significant and conversely. The correlation coefficients close to 1 are always more accurate than those that are close to zero. However, the correlation coefficient does not let establish quantitative changes of the resulting character in relation to the factorial one per unit of measurement. Therefore, a regression analysis is used.

The coefficient of the linear regression is called a number that indicates the direction and amount of the average change of Y character (function) in relation to the change

$$b_{yx} = \frac{\sum(X - \bar{x})(Y - \bar{y})}{\sum(X - \bar{x})^2}; \quad b_{xy} = \frac{\sum(X - \bar{x})(Y - \bar{y})}{\sum(Y - \bar{y})^2};$$

of X character (argument) per unit of measurement.

$$b_{yx} \cdot b_{xy} = r^2 = d$$

the error's regression coefficient:

$$s_{byx} = s_r \sqrt{\frac{\sum(Y - \bar{y})^2}{\sum(X - \bar{x})^2}}; s_{bxy} = s_r \sqrt{\frac{\sum(X - \bar{x})^2}{\sum(Y - \bar{y})^2}}$$

$$t_b = b / s_b.$$

Example: To perform a correlation and regression analysis of the experiment where the relationship between the weight of 1000 grains and the yield of winter wheat is studied.

The calculation of the complementary quantities to determine the correlation and regression

Pair number	Values of the characters		X ²	Y ²	XY
	X, g	Y, dt/ha			
1	38,1	48	1451,61	2304	1828,8
2	43,3	58,2	1874,89	3387,24	2520,06
3	47,3	57,4	2237,29	3294,76	2715,02
4	42,1	50,7	1772,41	2570,49	2134,47
5	45,2	58,2	2043,04	3387,24	2630,64
6	47,4	56,3	2246,76	3169,69	2668,62
7	45,3	50	2052,09	2500	2265
8	44,3	51,6	1962,49	2662,56	2285,88
Sum	ΣX = 353	ΣY = 430,4	ΣX² = 15640,58	ΣY² = 23275,98	ΣXY = 19048,49

$$\bar{x} = \sum X : n = 44,13$$

$$\bar{y} = \sum Y : n = 53,8$$

$$\sum (X - \bar{x})^2 = \sum X^2 - (\sum X)^2 : n = 15640,38 - 15576,125 = 64,46$$

$$\sum (X - \bar{x})(Y - \bar{y}) = \sum XY - (\sum X \sum Y) : n = 19048,49 - 18991,4 = 57,09$$

$$\sum (Y - \bar{y})^2 = \sum Y^2 - (\sum Y)^2 : n = 23275,98 - 23155,52 = 120,46$$

The correlation and regression coefficients:

$$r = \frac{\sum (X - \bar{x})(Y - \bar{y})}{\sqrt{\sum (X - \bar{x})^2 \sum (Y - \bar{y})^2}} = \frac{57,09}{\sqrt{64,46 \times 120,46}} = 0,648 \quad b_{yx} = \frac{\sum (X - \bar{x})(Y - \bar{y})}{\sum (X - \bar{x})^2} = \frac{57,09}{64,46} = 0,89 dt / ha$$

$$sr = \sqrt{\frac{1 - r^2}{n - 2}} = 0,097 \quad s_{byx} = s_r \sqrt{\frac{\sum (Y - \bar{y})^2}{\sum (X - \bar{x})^2}} = 0,097 \sqrt{\frac{120,46}{64,46}} = 0,133 dt / ha$$

The regression equation:

$$Y = \bar{y} + b_{yx}(X - \bar{x}) = 53,8 + 0,89(X - 44,13) = 0,89X + 14,52$$

The significance criteria $t_r = r/s_r = 6,68$; $v = n - 2 = 8 - 2 = 6$;

$$s_{yx} = s_r \sqrt{\sum (Y - \bar{y})^2} = 0,097 \sqrt{120,46} = 1,06 dt / ha$$

$$t_{05} = 2,45$$

The confidence intervals $\pm t_{05} s_r = 0,65 \pm 2,45 \cdot 0,097 = 0,65 \pm 0,24$
(0,45 ÷ 0,89)

$$b_{yx} \pm t_{05} s_{byx} = 0,89 \pm 2,45 \cdot 0,0133 = 0,89 \pm 0,133 (0,56 \div 1,22)$$

Since $t_{if} > t_{teor}$ and the confidence intervals do not include zero value, the correlation and regression are significant and the null hypothesis is rejected. The regression equation defines the theoretical Y values for the extreme X values:

$$Y_{x=38,1} = 0,89 \cdot 38,1 - 0,04 = 33,91$$

$$Y_{x=47,4} = 0,89 \cdot 47,4 - 0,04 = 42,19$$

The found points (38,1; 33,91) and (47,4; 42,19) are plotted on the graph and, by connecting them directly, one obtains the theoretical regression line of Y against X. On account of the determination coefficient $d = 0,42$, about 42% of Y change in this example is due to the change of X argument. To establish a confidence zone, it is necessary to lay off the value

of one (68% - zone) or two (95% - zone) errors – $\pm s_{yx}$ or $\pm 2s_{yx}$ up and down from the theoretical regression line and connect the dots with dashed lines, what is called the regression confidence zone.

One can also run this analysis using the software mentioned above. In Excel' it will look as follows:

DRAWING THE
CONCLUSIONS

<i>Regression statistics</i>	
Plural R	0,64790
R-square	0,41977
Standard R-square	0,32307
Standard error	3,41304
Observation	8

Dispersion analysis

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>F significance</i>
Regression	1	50,56657	50,5665	4,34088	0,08233
Residue	6	69,89343	11,6489		
Altogether	7	120,46			

	<i>Coefficients</i>	<i>Standard error</i>	<i>t-statistics</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95 %</i>	<i>Upper 95 %</i>
Y-intercept	14,7169	18,79731	0,78293	0,46344	-31,278	60,7123	-31,27	60,7123
X 1 variable	0,8857	0,425123	2,08347	0,08233	-0,1545	1,92597	-0,155	1,92597

However, it is better to construct the regression plot in the STATISTICA program (fig.1). In the figure, the graph is represented as it is shown by the program, but the inscriptions can be made in Cyrillic characters. The latter program is better for the statistical data processing: it allows to significantly expand the computation and perform the processing of various kinds, but compared to Excel' it is somewhat harder to use.

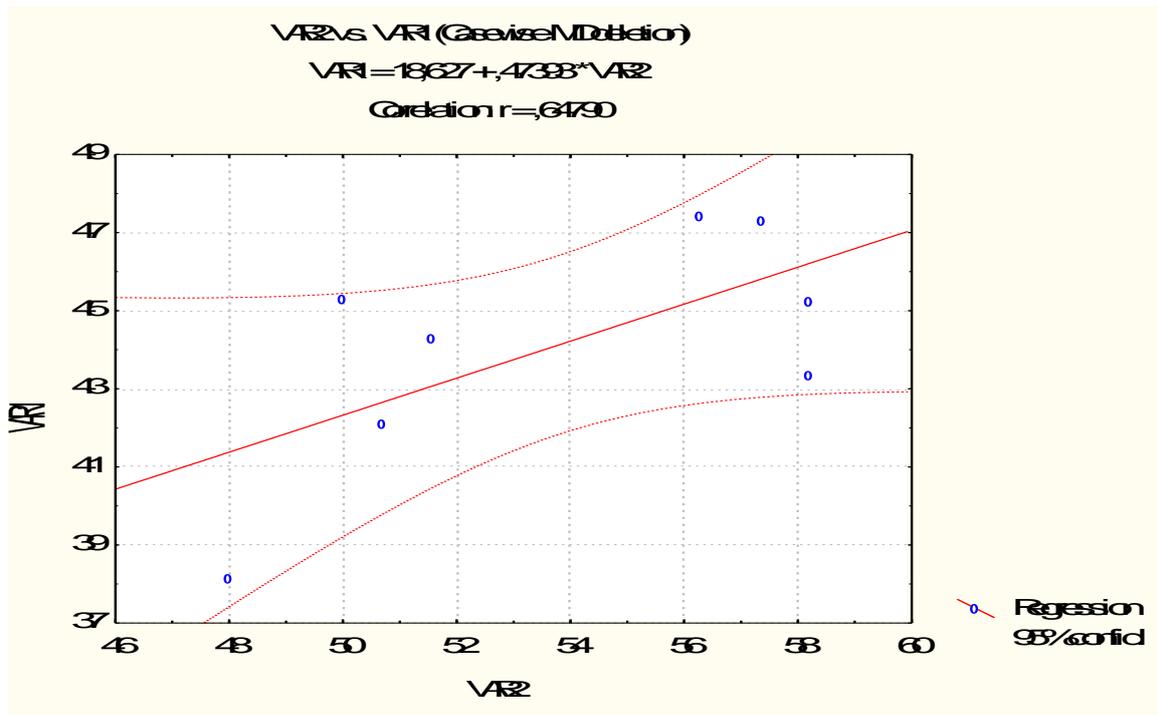


Fig.1. The example of the graphical depiction of the correlation analysis results

Tasks for the independent work:

- 1. The application of correlation and regression analyses in agronomic research.*
- 2. The determination of a correlative relationship and classification of correlations.*
- 3. The determination of regression.*
- 4. The properties of the correlation and regression coefficients.*
- 5. Conduct a correlation and regression analysis of the experiment using the software.*
- 6. The interpretation of the experimental results.*

Topic 9. DETERMINATION OF THE HERITABLE COEFFICIENT

Purpose of the class: *To learn how to calculate the values of the coefficients of the croppers' characteristic inheritance by means of the correlation and dispersion analysis*

The heritable coefficient (h^2) is of utmost importance for the rational planning of the parental pair selection for mating and forecasting the efficiency of selection. The general phenotypic variability of the quantitative characteristics (s^2_f) can be decomposed into hereditary - (s^2_g) (genetic) and modification (paratypic) - (s^2_p) what is due to the environmental factors:

$$s^2_f = s^2_g + s^2_p$$

Supposing that there is no bond between the genotype and paratype and setting s^2_f to 1, then

$$\frac{s^2_g}{s^2_f} + \frac{s^2_p}{s^2_f} = 1(100\%)$$

a fraction of the genetic variation in total variability is called a heritable coefficient:

$$h^2 = \frac{s^2_g}{s^2_f}$$

To calculate the heritable coefficient, there are also some known methods using the correlation, regression and dispersion analyses.

Tasks for the independent work:

- 1. The determination of the components of the phenotypic demonstration of the character.*
- 2. Methods for determining the heritable coefficient.*
- 3. Rules of applying the analysis model to determine the influence of parental forms.*
- 4. The determination of the heritable coefficient in the selection.*
- 5. Determine the coefficients of the croppers' characteristic inheritance using the software by means of the dispersion or regression analysis.*
- 6. The interpretation of the experimental results.*

LIST
of the questions for the test in "Basics of scientific research in agronomy" for extramural students of the Faculty of Agronomy

1. The definition and essence of the basic methods for agronomic research
2. The classification of agronomic experiments
3. Requirements to the field experiment
4. The essence of the "experimental error" concept
5. The types of errors in the experiment
6. The variation of soil fertility of the experimental plot, methods for determination and leveling
7. The concept of the field experiment technique
8. The influence of the elements in the field experiment technique on the experimental error
9. Basic elements of the field experiment technique
10. Methods for laying out the experimental variants
11. Planning a scientific experiment
12. The classification of the field experiments
13. Observations and records in the field experiment
14. Methods for sampling and observation
15. The technique of laying out the field experiment
16. Requirements to the field work on the experimental plot
17. Methods for the harvest records
18. The analysis of research data
19. Methods for research on the vegetable crops
20. Methods for holding the experiments with fruit crops
21. Peculiarities of research on the hayfields and pastures
22. Trial and demonstration experiments
23. Documentation and reporting in the field experiment
24. The final thesis as a form of a scientific report and its requirements
25. General and sample population
26. Statistical figures of the quantitative variability
27. Statistical figures of the qualitative variability
28. Characteristics of the normal Student's distribution
29. Probability, confidence interval, equal significance
30. The essence of the null hypothesis
31. The evaluation of differences using LSD and Student's criteria
32. χ^2 and the application rules

33. The estimation of differences between dispersions according to Fisher's criterion
34. The basics and essence of the dispersion analysis
35. One-factor dispersion analysis
36. Multiple factor dispersion analysis
37. The ratio of the theoretical and actual value of Fisher's criterion.
Formulating the conclusions
38. Types of correlation and regression analysis
39. Coefficients of the correlation and regression, methods for determination
40. The determination of the heritability estimate by means of the correlation, regression, dispersion analyses

Task 1

Using the data of Annex B, determine the statistical parameters of the sample – the mean, the variance, a standard deviation, an average error, the coefficient of variation.

Task 2

Determine the least significant difference according to Student's criterion. Check the accuracy of the differences at a 5% and a 1% level of significance. The output data is in the table of Annex B by the reference number of Annex B.

Task 3

Run a dispersion analysis of a one-factor field experiment conducted by the Latin square method of 4 x 4. The numbers of variants are determined using Annex D. The control variant is the first one.

The yield of winter wheat, t/ha

Variant	Repetition				Sum	Mean
	I	II	III	IV		
1 st						
2						
3						
4						

Task 4

Calculate the coefficients of the pair correlation, regression and determination between the length of the head and the number of spikelets in the head using four samples of Annex E by the reference number in Annex B to form the sample.

Annex A

t-criterion value at a 5, 1, 0,1% significance level

Number of the degrees of freedom	Significance level		
	0,05	0,01	0,001
1.	12,71	63,66	-
2.	4,3	9,93	31,6
3.	3,18	5,84	12,94
4.	2,78	4,6	8,61
5.	2,57	4,03	6,86
6.	2,45	3,71	5,96
7.	2,37	3,5	5,41
8.	2,31	3,36	5,04
9.	2,26	3,25	4,78
10.	2,23	3,17	4,59
11.	2,2	3,11	4,44
12.	2,18	3,06	4,32
13.	2,16	3,01	4,22
14.	2,15	2,98	4,14
15.	2,13	2,95	4,07
16.	2,12	2,92	4,02
17.	2,11	2,9	3,97
18.	2,1	2,88	3,92
19.	2,09	2,86	3,88
20.	2,09	2,85	3,85
21.	2,08	2,83	3,82
22.	2,07	2,82	3,79
23.	2,07	2,81	3,77
24.	2,06	2,8	3,75
25.	2,06	2,79	3,73
26.	2,06	2,78	3,71
27.	2,05	2,77	3,69
28.	2,05	2,76	3,67
29.	2,05	2,76	3,66
30.	2,04	2,75	3,65
50.	2,01	2,68	3,5
100.	1,98	2,63	3,39
∞	1,96	2,58	3,29

Annex B

F-criterion value at a 5% significance level (a 95% probability)

Degrees of freedom for the minor dispersion (denominator)	Degrees of freedom for the major dispersion (numerator)													
	1	2	3	4	5	6	7	8	9	10	12	24	50	100
1.	161	200	216	225	230	234	237	239	241	242	244	249	252	253
2.	18,51	19,00	19,16	19,25	19,3	19,33	19,36	19,37	19,38	19,39	19,41	19,45	19,47	19,49
3.	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,74	8,64	8,58	8,56
4.	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,77	5,70	5,66
5.	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,68	4,53	4,44	4,40
6.	5,99	5,14	4,76	4,53	4,39	4,27	4,21	4,15	4,10	4,06	4,00	3,84	3,75	3,71
7.	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,57	3,41	3,32	3,28
8.	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,28	3,12	3,03	2,98
9.	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,07	2,90	2,80	2,76
10.	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,91	2,74	2,64	2,59
11.	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,79	2,61	2,50	2,45
12.	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,69	2,50	2,40	2,35
13.	4,64	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,60	2,42	2,32	2,26
14.	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,53	2,35	2,24	2,19
15.	4,54	3,60	3,29	3,06	2,9	2,79	2,70	2,64	2,59	2,55	2,48	2,29	2,18	2,12
16.	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,24	2,13	2,07
17.	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,38	2,19	2,08	2,02
18.	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,15	2,04	1,98
19.	4,38	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,38	2,31	2,11	2,00	1,94
20.	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,28	2,08	1,96	1,90
21.	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,05	1,93	1,87
22.	4,30	3,44	3,05	2,82	2,66	2,55	2,47	2,40	2,35	2,30	2,23	2,03	1,91	1,84
23.	4,28	3,42	3,03	2,80	2,64	2,53	2,45	2,38	2,32	2,28	2,20	2,00	1,88	1,82
24.	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,18	1,98	1,86	1,80
25.	4,24	3,38	2,99	2,76	2,60	2,49	2,41	2,34	2,25	2,24	2,16	1,96	1,84	1,77
26.	4,22	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	1,95	1,82	1,76
28.	4,20	3,34	2,95	2,71	2,56	2,44	2,36	2,29	2,24	2,19	2,12	1,91	1,78	1,72
30.	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,12	2,09	1,89	1,76	1,69
40.	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	2,00	1,79	1,66	1,59
50.	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,95	1,74	1,60	1,52
100.	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,85	1,63	1,48	1,39

F-criterion value at a 1% significance level (a 99% probability)

Degrees of freedom for the minor dispersion (denominator)	Degrees of freedom for the major dispersion (numerator)													
	1	2	3	4	5	6	7	8	9	10	12	24	50	100
1.	4052	4999	5403	5625	5764	5889	5928	5981	6022	6056	6106	6234	6302	6334
2.	98,49	99,01	99,17	99,25	99,30	99,33	99,34	99,36	99,38	99,40	99,42	99,46	99,48	99,49
3.	34,12	30,81	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,05	26,60	26,35	26,23
4.	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,54	14,37	13,93	13,69	13,57
5.	16,26	13,27	12,06	11,39	10,97	10,67	10,45	10,27	10,15	10,05	9,89	9,47	9,24	9,13
6.	13,74	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,31	7,09	6,99
7.	12,25	9,55	8,45	7,85	7,46	7,19	7,00	6,84	6,71	6,62	6,47	6,07	5,85	5,75
8.	11,26	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,67	5,28	5,06	4,96
9.	10,56	8,02	6,99	6,42	6,06	5,80	5,62	5,47	5,35	5,26	5,11	4,73	4,51	4,41
10.	10,04	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,71	4,33	4,12	4,01
11.	9,85	7,20	6,22	5,67	5,32	5,07	4,88	4,74	4,63	4,54	4,40	4,02	3,80	3,70
12.	9,33	6,93	5,95	5,41	5,06	4,82	4,65	4,50	4,39	4,30	4,16	3,78	3,56	3,46
13.	9,07	6,70	5,74	5,20	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,59	3,37	3,27
14.	8,86	6,51	5,56	5,03	4,69	4,46	4,28	4,14	4,03	3,94	3,80	3,43	3,21	3,11
15.	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,29	3,07	2,97
16.	8,53	6,23	5,29	4,77	4,44	4,20	3,89	3,78	3,69	3,61	3,45	3,18	2,96	2,86
17.	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,45	3,08	2,86	2,76
18.	8,28	6,01	5,09	4,58	4,25	4,01	3,85	3,71	3,60	3,51	3,37	3,00	2,78	2,68
19.	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,68	3,52	3,43	3,30	2,92	2,70	2,63
20.	8,10	5,85	4,94	4,43	4,10	3,87	3,71	3,56	3,45	3,37	3,23	2,86	2,63	2,53
21.	8,02	5,78	4,87	4,37	4,04	3,81	3,65	3,51	3,40	3,31	3,17	2,80	2,58	2,47
22.	7,94	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,75	2,53	2,42
23.	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,70	2,48	2,37
24.	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,25	3,17	3,03	2,66	2,44	2,33
25.	7,77	5,57	4,68	4,18	3,86	3,63	3,46	3,32	3,21	3,13	2,99	2,62	2,40	2,29
26.	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,17	3,09	2,96	2,58	2,36	2,25
28.	7,64	5,45	4,57	4,07	3,76	3,53	3,36	3,23	3,11	3,03	2,90	2,52	2,30	2,18
30.	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,06	2,98	2,84	2,47	2,24	2,13
40.	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,88	2,80	2,66	2,29	2,05	1,94
50.	7,17	5,06	4,20	3,72	3,41	3,18	3,02	2,88	2,78	2,70	2,56	2,18	1,94	1,81
100.	6,90	4,82	3,98	3,51	3,20	2,99	2,82	2,69	2,59	2,51	2,36	1,98	1,73	1,59

The reference number to fulfil the task and test

Next to last reference number	Last reference number									
	1	2	3	4	5	6	7	8	9	10
1	1, 10, 20, 30, 40	2, 11, 21, 31, 39	3, 12, 22, 32, 38	4, 13, 23, 33, 37	5, 14, 24, 34, 36	6, 15, 25, 27, 35	7, 16, 26, 28, 36	8, 17, 20, 29, 37	9, 18, 21, 30, 38	10, 19, 22, 31, 39
2	2, 9, 19, 29, 39	3, 8, 18, 28, 38	4, 7, 17, 27, 37	5, 6, 16, 26, 36	6, 15, 20, 27, 37	7, 14, 19, 28, 38	8, 13, 18, 29, 39	9, 12, 19, 30, 40	10, 11, 18, 31, 39	11, 12, 19, 32, 38
3	3, 10, 19, 27, 31	4, 11, 18, 28, 32	5, 12, 20, 29, 33	6, 13, 21, 30, 34	7, 14, 17, 31, 35	1, 8, 15, 18, 32, 36	9, 16, 19, 33, 37	10, 15, 20, 34, 38	1, 16, 21, 35, 39	2, 17, 22, 36, 40
4	4, 9, 15, 27, 31	5, 8, 16, 28, 32	6, 9, 17, 29, 33	7, 10, 18, 30, 34	8, 11, 19, 31, 35	9, 12, 20, 32, 36	10, 13, 21, 33, 37	11, 14, 22, 34, 38	1, 15, 24, 35, 39	2, 16, 25, 36, 40
5	5, 10, 15, 30, 40	6, 11, 16, 31, 39	7, 12, 17, 32, 38	8, 13, 18, 32, 37	9, 14, 19, 33, 36	10, 15, 20, 34, 35	1, 9, 19, 33, 36	2, 10, 20, 34, 37	3, 11, 21, 35, 38	4, 12, 22, 36, 39
6	1, 9, 20, 27, 35	2, 8, 19, 28, 36	3, 7, 20, 28, 37	4, 8, 21, 29, 38	5, 9, 20, 30, 39	6, 10, 21, 31, 40	7, 11, 22, 32, 39	8, 12, 23, 33, 40	9, 13, 24, 34, 39	10, 14, 25, 35, 40
7	2, 8, 15, 29, 35	3, 9, 16, 30, 36	4, 10, 17, 31, 37	5, 11, 18, 32, 38	6, 12, 19, 33, 39	7, 13, 20, 34, 40	8, 14, 21, 35, 39	9, 15, 22, 36, 40	10, 16, 23, 37, 38	1, 17, 24, 27, 39
8	3, 10, 21, 30, 40	4, 11, 20, 29, 39	5, 12, 19, 28, 38	6, 13, 20, 27, 37	7, 14, 21, 26, 36	8, 15, 22, 30, 35	9, 16, 23, 31, 36	10, 17, 24, 32, 37	2, 18, 25, 33, 38	1, 19, 24, 34, 39
9	4, 11, 21, 32, 38	5, 12, 22, 33, 39	6, 13, 23, 34, 40	7, 13, 24, 35, 39	8, 14, 25, 36, 40	9, 15, 26, 30, 39	10, 16, 20, 31, 38	1, 17, 21, 31, 37	2, 19, 22, 32, 38	3, 20, 23, 33, 39
0	5, 9, 17, 28, 32	1, 5, 18, 27, 33	2, 6, 19, 28, 34	3, 7, 20, 29, 35	6, 19, 26, 30, 36	7, 20, 25, 29, 35	8, 19, 24, 30, 36	9, 20, 25, 31, 37	10, 19, 26, 33, 38	11, 20, 25, 35, 40

Reference numbers to fulfil the task 2-4 of the test

Next to last reference number	Last reference number									
	1	2	3	4	5	6	7	8	9	0
1	1, 8, 14, 19	5, 7, 16, 20	2, 10, 13, 21	3, 9, 15, 23	1, 7, 18, 22,	4, 11, 15, 24	2, 9, 14, 25	1, 8, 15, 19	6, 10, 16, 20	5, 12, 17, 21
2	4, 12, 18, 19	3, 7, 13, 23	6, 9, 14, 24	1, 10, 16, 25	5, 18, 13, 20	2, 7, 14, 19	3, 9, 17, 23	6, 11, 16, 22	4, 8, 17, 21	1, 11, 18, 24
3	2, 7, 15, 25	6, 10, 17, 19	1, 8, 15, 20	6, 9, 18, 21	3, 7, 13, 22	5, 11, 16, 23	2, 10, 17, 24	4, 7, 14, 25	1, 12, 16, 24	3, 7, 13, 20
4	5, 9, 18, 19	11, 3, 17, 20	6, 8, 15, 21	1, 9, 18, 21	4, 10, 15, 23	2, 7, 18, 24	3, 8, 16, 25	4, 12, 13, 21	6, 7, 15, 23	2, 11, 18, 19
5	2, 10, 16, 20	1, 7, 17, 23	6, 11, 16, 20	3, 8, 13, 19	5, 8, 16, 20	4, 9, 15, 21	1, 11, 14, 22	3, 10, 17, 23	6, 9, 14, 24	4, 8, 14, 25
6	6, 12, 18, 21	3, 7, 16, 21	1, 9, 13, 19	4, 1, 18, 20	2, 8, 15, 21	1, 10, 16, 23	5, 7, 14, 22	4, 9, 13, 25	6, 12, 15, 24	3, 11, 16, 20
7	2, 10, 18, 22	6, 11, 17, 20	5, 8, 13, 25	3, 10, 16, 20	6, 9, 17, 19	4, 7, 16, 23	2, 9, 18, 22	3, 8, 17, 24	1, 7, 16, 25	4, 9, 18, 22
8	4, 12, 14, 23	1, 7, 13, 19	2, 9, 17, 23	5, 10, 13, 21	1, 12, 15, 22	6, 9, 13, 25	2, 11, 17, 24	4, 8, 14, 19	3, 10, 18, 20	5, 7, 13, 21
9	2, 12, 13, 24	5, 10, 18, 25	6, 12, 13, 24	4, 7, 18, 20	5, 12, 15, 21	2, 12, 14, 23	3, 7, 16, 25	6, 11, 15, 20	5, 11, 17, 19	3, 12, 15, 23
0	5, 11, 17, 25	4, 8, 18, 25	1, 9, 15, 21	5, 10, 17, 22	2, 8, 14, 23	3, 10, 13, 20	5, 12, 15, 19	6, 18, 17, 21	1, 11, 14, 24	4, 10, 16, 24

Annex F**Sample (spike length, cm)**

Sample number	1	2	3	4	5	6	7	8	9	10
1	4,8	5,0	5,1	4,7	5,8	4,8	6,0	4,9	5,3	5,2
2	5,2	5,8	6,2	6,8	5,2	5,3	4,8	4,9	5,3	5,4
3	5,4	4,8	5,1	5,8	4,8	5,0	5,3	6,1	6,0	6,0
4	6,0	5,8	5,4	4,8	5,0	4,9	5,1	5,3	5,3	4,7
5	4,7	4,9	5,1	5,3	5,2	5,5	5,6	5,4	5,5	4,9
6	4,9	5,1	5,2	5,4	5,5	5,2	5,3	5,4	5,2	5,6
7	5,6	5,6	5,3	5,3	5,0	4,9	5,5	5,2	5,0	5,8
8	5,8	5,2	5,5	5,1	5,0	5,2	5,8	5,4	5,4	5,5
9	5,5	6,2	5,8	6,1	6,5	6,4	6,3	6,4	6,0	5,9
10	5,9	6,0	5,9	6,0	6,1	6,2	6,4	6,3	6,3	6,2
11	6,2	6,0	5,4	5,5	5,2	5,2	4,8	5,8	5,4	5,3
12	5,3	6,3	5,8	5,7	5,6	5,8	5,9	6,0	6,0	5,9
13	5,9	6,1	5,2	5,9	5,8	6,3	5,8	5,8	5,7	5,6
14	5,6	6,1	6,2	6,4	6,3	6,4	6,3	6,3	6,1	5,8
15	5,8	6,3	6,4	6,6	6,7	6,8	7,1	7,2	6,8	6,9
16	6,9	7,1	7,3	7,2	7,0	6,9	5,8	5,7	7,2	6,9
17	6,9	6,9	7,1	7,3	7,5	7,5	7,6	7,8	7,9	7,8
18	7,1	7,8	7,3	6,3	6,4	6,5	6,6	6,8	6,9	7,3
19	7,3	7,5	7,6	7,7	7,8	7,9	7,3	7,3	7,4	7,8
20	7,7	7,7	7,3	7,5	7,3	7,2	7,1	7,2	7,3	7,2
21	7,2	7,2	7,3	7,7	7,8	7,7	7,3	7,5	7,5	7,8
22	7,5	8,0	7,8	7,7	7,2	7,3	8,1	8,2	7,6	7,7
23	8,2	8,0	7,9	7,7	7,2	7,3	7,0	7,1	7,3	7,0
24	7,0	8,0	7,3	7,4	7,5	7,6	7,6	7,7	7,7	7,8
25	7,8	7,3	7,4	7,5	7,7	8,2	8,3	8,4	8,0	8,2

Winter wheat yield (the reference for calculations)

Variant	Yield in the repetitions, t/ha			
	1	2	3	4
1.	2,31	2,28	2,34	2,29
2.	2,33	2,37	2,29	2,35
3.	2,32	2,34	2,39	2,40
4.	2,27	2,30	2,26	2,32
5.	2,44	2,51	2,43	2,49
6.	2,30	2,42	2,35	2,38
7.	2,74	2,67	2,71	2,65
8.	2,68	2,71	2,69	2,77
9.	2,59	2,57	2,64	2,67
10.	2,62	2,68	2,59	2,65
11.	2,78	2,73	2,70	2,69
12.	2,81	2,77	2,85	2,79
13.	3,02	3,07	2,95	2,98
14.	2,96	2,95	3,00	3,03
15.	2,99	2,87	2,93	2,96
16.	3,07	3,02	3,11	2,98
17.	2,86	3,03	2,93	2,90
18.	3,01	3,08	3,12	3,03
19.	3,19	3,26	3,21	3,23
20.	3,28	3,20	3,16	3,22
21.	3,18	3,27	3,21	3,25
22.	3,11	3,18	3,17	3,22
23.	3,07	3,15	3,06	3,11
24.	3,12	3,23	3,19	3,26
25.	3,10	3,17	3,08	3,21

the spike length (X) and the number of spikelets in a spike (Y) of the wheat plants

Sample number	Figures	Plants				
		1	2	3	4	5
1	X	7,3	8,6	7,8	12,3	9,0
	Y	14	17	15	23	18
2	X	6,9	7,8	8,0	9,3	8,7
	Y	13	15	15	17	17
3	X	10,1	8,5	11,2	6,9	8,9
	Y	21	16	21	13	19
4	X	9,4	9,0	10,5	13,1	7,7
	Y	18	19	22	25	14
5	X	8,6	7,8	10,2	8,7	9,6
	Y	18	15	19	16	21
6	X	7,5	9,5	8,1	10,7	12,2
	Y	15	20	17	20	23
7	X	9,9	8,3	8,8	7,1	12,1
	Y	18	16	18	14	25
8	X	10,7	9,4	9,8	8,6	8,1
	Y	22	19	19	17	16
9	X	7,1	9,6	8,3	10,7	8,5
	Y	14	20	16	22	17
10	X	11,7	7,0	7,8	12,4	8,1
	Y	23	14	15	25	16
11	X	10,5	6,8	7,7	8,1	9,6
	Y	20	14	15	16	18
12	X	9,8	10,3	12,1	11,5	10,3
	Y	19	19	24	23	20
13	X	7,9	8,3	8,3	7,3	9,1
	Y	16	17	16	15	17
14	X	7,0	8,7	9,9	10,6	7,1
	Y	15	17	19	21	15
15	X	8,8	9,2	10,4	9,6	8,0
	Y	18	18	21	19	16
16	X	7,1	7,4	7,8	8,1	8,0
	Y	14	15	15	17	15
17	X	9,7	10,6	11,5	13,2	10,1
	Y	20	21	22	25	19
18	X	10,2	11,1	8,0	7,7	8,1
	Y	20	21	16	16	17
19	X	12,0	9,6	7,1	7,7	7,8
	Y	23	19	13	15	16

Continuation of Annex H

20	X	10,5	9,2	8,6	10,6	8,8
	Y	21	18	19	20	20
21	X	9,1	8,6	8,3	9,7	11,0
	Y	18	17	16	20	22
22	X	7,5	7,0	8,7	9,1	10,1
	Y	14	13	17	18	19
23	X	8,3	10,6	9,5	7,0	8,1
	Y	16	21	18	15	17
24	X	9,3	6,8	7,7	7,9	8,9
	Y	18	11	15	16	17
25	X	10,8	9,4	9,0	7,1	8,8
	Y	22	19	17	13	15

List of references

1. Dospekhov B.A. The procedure of the field experiment (with the basics of the statistical data processing). — M.: Agropromizdat, 1985. — 351 p.
2. Lakin H.F. Biometry. — M.: Vyshcha shkola, 1980. — 296 p.
3. Moyseychenko V.F. Basics of scientific research in agronomy. — K.: Vyshcha shkola — 1994.— 334 p.
4. Molostov A.S. Procedure of the field experiment. — M.: Kolos., 1966.— 239 p.
5. Moyseychenko V.F. Basics of scientific research in fruit-farming, vegetable, viticulture and the storage technique of fruit and vegetable products. — K.: NYK VO, 1992. — 346 p.
6. Volf V.H. Statistical processing of the experimental data. — M.: Kolos. — 1956. — 256 p.
7. Computer methods in agriculture and biology: The manual (O.M. Tsarenko, Yu.A.Zlobin, V.H.Skliar, S.M.Panchenko).- Sumy: “Univer-sytetska knyha” publishing office 2000.- 203 p.
8. Pidipryhora V.S., Pysarenko P.V. The practicum on the basics of scientific research in agronomy. — Poltava. — 2003. — 138 p.