Ministry of Education of the Republic of Belarus
Educational Establishment “Vitebsk State University
named after P.M. Masherov”
Department of Psychology

QUALITATIVE AND QUANTITATIVE RESEARCH METHODS IN PSYCHOLOGY

Course of lections

Vitebsk
VSU named after P.M. Masherov
2022

UDC 159.9.072:001.891(075.8)
LBC 88в6я73+72.52я73
Q19

Published by the decision of the Research and Methodology Council of the Educational Establishment "Vitebsk State University named after P.M. Masherov". Minutes № 3 d/d 03.03.2022.

Compiled by: Associate Professor of Psychology Department of the VSU named after P.M. Masherov, PhD in Psychology **N.V. Kukhtova;** Lecturer of Psychology Department of the VSU named after P.M. Masherov **M.M. Morozhanova**

R e v i e w e r s :
Head of the Department of Psychology and Pedagogy
of the Vitebsk State Order of Peoples' Friendship Medical University,
PhD in Medical Sciences, Assistant Professor *A.L. Tserkovsky;*
Head of the Department of Social Pedagogical Work
of the VSU named after P.M. Masherov,
PhD in Pedagogy, Assistant Professor *E.L. Mihailova*

Q19 **Qualitative and quantitative research methods in psychology :** course of lections / compl. by: N.V. Kukhtova, M.M. Morozhanova. – Vitebsk : VSU named after P.M. Masherov, 2022. – 48 p.

The course of lectures "Qualitative and quantitative research methods in psychology" consists of lectures, questions and tasks for independent work and control on each topic. Contains a detailed overview of the main methods of qualitative and quantitative research in psychology. The course of lectures "Qualitative and quantitative research methods in psychology" is addressed to undergraduates of specialty 1-23 80 03 "Psychology".

UDC 159.9.072:001.891(075.8)
LBC 88в6я73+72.52я73

© VSU named after P.M. Masherov, 2022

CONTENT

INTRODUCTION

Goals and objectives of the academic discipline:

The main purpose of this course is to draw on the basic knowledge of psychology already available to undergraduates, to familiarize them with the current state and prospects for the development of the leading branches of psychological practice.

The objectives of the study of the discipline are: to characterize the modern trends in the development of psychology; critically analyze the main theoretical approaches of modern psychological science; develop undergraduates' skills of independent and conscious determination of the methodology and methods of scientific research; to acquaint undergraduates with the most promising areas of application of psychological knowledge in practice; to learn how to identify psychological problems, to reveal the unity of theory and practice in the context of problematic issues; to acquire the skills of an independent search for approaches to solving both theoretical and purely applied problems; to master the skills of determining the degree of relevance of a particular psychological problem in the context of modern psychological science and psychological practice in various branches of psychology; to learn to see interdisciplinary issues in the structure of psychological problems and to set tasks with the involvement of specialists from other sciences; to learn methods of translating theoretical knowledge into a form that ensures the practical use of this knowledge.

2. The place of the academic discipline in the system of specialist training:

In the master's degree system, the academic discipline "Qualitative and quantitative Research methods in Psychology" is included in the module "Applied Problems of Behavior Psychology" (a component of a higher education institution). The discipline "Qualitative and quantitative research methods in psychology" is based on a wide use of knowledge of the disciplines "Psychodiagnostics", "Theory and practice of practical activity of a psychologist", "Labor Psychology", "Social Psychology", "Management Psychology". The discipline "Qualitative and quantitative research methods in psychology" is a brief systematic presentation of the knowledge accumulated by modern psychology and related sciences about the variety of theoretical directions, methodological approaches to the study of personality and human behavior, as well as the types and directions of psychological assistance and methods of their statistical processing. The current social situation increases the requirements for the level of professional training of a specialist psychologist, who must be able to navigate in a variety of theoretical areas, plan psychological research and their own practical activities, use the achievements of other sciences that study people and society. The program is designed to provide the basic psychological competence of future masters in theoretical and practical aspects. The program focuses the attention of undergraduates on psychological problems that are directly related to their future professional activities. Under this discipline "of

Qualitative and quantitative research methods in psychology" discusses trends in the development of modern psychological practice, the problem of correlation between psychological practice and the fundamental psychological knowledge, the variety of technologies used in psychological practice, perspective directions of research of problems of psychological practice, the role of psychological practice in the social life of the modern Belarusian society.

3. Requirements for the development of the academic discipline.

As a result of studying the academic discipline, the master's student should know:

- the main trends in the development of psychological practice in the Republic of Belarus, the countries of the near and far abroad;
- general and particular problems of modern psychological practice;
- current problem fields and promising areas of psychological practice;

be able to:

- identify areas of application of fundamental psychological knowledge in various areas of psychological practice;
- identify and analyze technologies used in psychological practice;
- identify and analyze the needs for the main types of psychological services of modern Belarusian society;

to own:

- skills of primary assessment of the psychological state of an adult, assessment of the degree of his maturity in socio-psychological terms,
- ways to stop behavior that deviates from the norm;
- research skills;
- oral and written communication skills;
- techniques of management of educational-cognitive, scientific-research activities.

In the process of mastering the discipline "Current directions of psychological Practice", the master's student forms and demonstrates the following **competencies:**

Universal competencies:

UK-2. Be able to use qualitative and quantitative research methods in psychological practice.

MODULE 1
APPLICATION OF MATHEMATICAL STATISTICS IN PSYCHOLOGY

Lecture 1. History of mathematical methods in psychology

Lecture plan

1. Stages of development of mathematical methods in psychology.
2. History of the use of mathematical tools in psychology in the late 19th - early 20th centuries.
3. Use of mathematical tools in psychology in the early - mid 20th century.
4. Modern use of mathematical tools in psychology.

Basic concepts: science, psychology, mathematical methods

1. Stages of development of mathematical methods in psychology.

1. Introduction. The process of mathematization of psychology began from the moment it was separated into an experimental discipline. This process goes through a number of stages.

The first is the use of mathematical methods for analyzing and processing the results of experimental research, as well as the derivation of simple laws (late XIX century – early XX century). This is the time of developing the law of learning, the psychophysical law, and the method of factor analysis.

The second (40–50s) - creating models of mental processes and human behavior using previously developed mathematical tools.

The third (60s to the present) is the separation of mathematical psychology into a separate discipline, the main purpose of which is to develop a mathematical apparatus for modeling mental processes and analyzing data from a psychological experiment.

The fourth stage has not yet arrived. This period should be characterized by the formation of theoretical psychology and the withering away - mathematical.

2. History of the use of mathematical tools in psychology in the late 19th - early 20th centuries

Psychology in its scientific development inevitably had to pass and has passed the path of mathematization, although not in all countries and not in full. The exact date of the beginning of the path of mathematization, perhaps, does not know any science. However, for psychology, April 18, 1822 can be taken as the conditional date of the beginning of this way. It was then at the Royal German scientific society, that Johann Friedrich Herbart read a report "On the Possibility and Necessity of Applying Mathematics to Psychology". The main idea of the report was that if psychology wants to be a science, like physics, mathematics must and can be applied in it. Two years after this report, J. F. Herbart published the book "An Attempt to Found the Science of Psychology on Experience, Metaphysics, and Mathematics". In 1850 in Leipzig the second fundamental book by Moritz-Wilhelm Drobisch “Foundations of Mathematical Psychology”. Thus, this psychological discipline also has an exact date of appearance in science. However, some modern psychologists who write in the field of mathematical psychology associate the beginning of its development with the American journal, which appeared in 1963. Thus, mathematical psychology, or more precisely, mathematized psychology, developed for a century before the Americans. And the beginning of this process of mathematization of psychology was laid by I. F. Herbart and M.-V. Drobish.

It should be said that in terms of innovations, Drobish's mathematical psychology is inferior to that made by his teacher Herbart. In modern psychology mathematical modeling is a product of the second half of the 20th century. In the preface to Nechaev's translation of Herbart's psychology, the Russian professor A. I. Vvedensky, famous for his "psychology without any metaphysics", spoke very disparagingly of Herbart's attempt to apply mathematics in psychology. But natural scientists, in particular psychophysicists Theodor Fechner, and the famous Wilhelm Wundt, who worked in Leipzig, could not pass by the fundamental publications of J. F. Herbart and M.-V. Drobish. It was they who mathematically implemented in psychology Herbart's ideas about psychological values, thresholds of consciousness, reaction time of human consciousness, and implemented them using modern mathematics. The main methods of mathematics at that time-differential and integral calculus, equations of relatively simple dependencies were quite suitable for identifying and describing the simplest psychophysical laws and various human reactions. It was not for nothing that W. Wundt categorically denied the possibility of empirical psychology to investigate higher mental functions. They were, according to Wundt, in the field of a special, essentially metaphysical, psychology of peoples. English-speaking scientists began to create mathematical tools for studying complex multidimensional objects, including higher mental functions-intelligence, abilities, and personality. Among other results, it turned out that the height of descendants tends to return to the average height of ancestors, so the concept of *"regression"* appeared, and equations were obtained that express this relation. The coefficient proposed by the Frenchman A. Bravais was improved, which quantitatively expresses the ratio of two changing variables, i.e. *correlation.* Now this coefficient is one of the most important means of multidimensional data analysis, even the symbol has retained its abbreviation: small Latin "r" from English - relation.

3. Use of mathematical tools in psychology in the early – mid 20th century

The twentieth century introduced several unexpected features into the "relationship" between psychology and mathematics. So, if at the beginning of the century some aspects of this issue are widely discussed by various scientists, including A. Poincare, I. P. Pavlov, A. Einstein and others, then in the 30-40s they no longer cause such an acute interest. This can be judged at least by the fact that in psychological research of this period, very little attention is paid to quantitative methods, formalization of psychological phenomena is preferred to qualitative descriptions, etc.

Since the late 1920s, mathematical methods have increasingly penetrated into psychology and are creatively used in it. The psychological theory of measurement is being intensively developed. Stochastic models of learning in behavioral psychology are developed on the basis of the Markov chain apparatus. Created in the field of biology by Ronald Fischer, dispersion analysis is becoming the main mathematical method in genetic psychology. Mathematical models from automatic control theory and Shannon's information theory are widely used in engineering and general psychology. As a result, modern scientific psychology in many of its branches is mathematized in a significant way. At the same time, newly emerging mathematical innovations are often borrowed by psychologists for their own purposes. For example, the appearance of an algorithmic language for control problems, proposed by A. A. Lyapunov and G. A. Shestopal, was used by V. N. Pushkin to compile algorithms for the activity of a train dispatcher.

Just a few years pass, and interest in the use of mathematics in psychology flares up with new, unprecedented force. The reason for this was the emergence and rapid development of a number of technical sciences, primarily cybernetics. The latter contributed to the rapid improvement of many mathematical methods, which, in connection with the new problems that arose in psychology, could be used in it much more effectively than before.

But the main distinguishing feature of the interaction between psychology and mathematics of this period was the appeal of mathematics to psychology. This unexpected turn is explained by the fact that in the middle of the XX century, due to the development of electronic computing and advances in psychology, neurology and physiology, was a real opportunity to raise the problem of "improving" thought processes. Mathematical modeling in psychology is initially associated with efforts to find forms of describing mental phenomena using the categories of abstract algebra, topology, classical and vector analysis. For example, J. Piaget used a special logical–mathematical apparatus to describe the structure of intellectual acts, which correlates mental operations with mathematical ones. These operations can be linked into mobile integral structures that are characterized by their reversibility and associativity. K. Levin describes human behavior in terms of geometric topology and vector analysis. At the same time, the living space of an individual began to be interpreted by him as an integral field, within which the aspirations and intentions of a person arise and change, having like vectors, direction, magnitude and points of application. As one of the important characteristics of each point of the field, its valence (positive – negative motives) was considered. When describing the same field, K. Levin uses such abstract mathematical concepts as gradient, rotor, divergence, and others. Despite the fact that the constructed mathematical models, as well as later ones, could not fully explain the phenomena and processes of the corresponding field of psychology, they allowed us to establish a number of important facts and laws that currently belong to the "basics" of psychological science (stadiality of intelligence formation, the presence of a level of claims, the Zeigarnik effect, etc.).

4. Modern use of mathematical tools in psychology

Analyzing the relationship between mathematics and psychology, we can say that they are complex and ambiguous. On the one hand, many common psychological theories and concepts have proved their validity without any mathematical and statistical confirmation (for example, the theory of psychoanalysis of Z. Freud; the concept of humanistic psychology, etc.). On the other hand, the use of mathematical methods, of course, to a large extent objectifies the conclusions of psychologists, makes it possible to isolate from a huge set of disparate psychological data factors and variables that are important in one sense or another that have the greatest influence on the investigated resulting feature.

The use of statistical methods in experimental psychology began at the beginning of the last century and included several areas that brought deep qualitative transformations to psychological science. For example, F. Galton was able to make deep progress in the study of the relationship between heredity and external influence, using the ideas of regression and correlation analysis. To study the structure of intelligence, C. Spearman was the first to use factor analysis. The introduction of a number of statistical methods into the practice of a psychologist made it possible to significantly objectify the quality control of psychological tests used in professional selection.

Currently, the development of the apparatus for mathematical modeling of mental processes and phenomena has led to the formation of an integral scientific field –mathematical psychology, which, in particular, deals with the systemic problems of interaction between psychology, semiotics and information theory. In search of building computers that should model certain types of mental activity, their creators turned to two basic principles of psychology and physiology of higher nervous activity - associations and reflex. Thus, the founder of cybernetics, N. Wiener, wishing to construct a scheme of intelligent activity, turns to Locke's concept of the mechanism of associations and the teaching of I.P. Pavlov about conditioned reflexes. All this quite convincingly characterizes one of the general laws of modern scientific knowledge, which is that many problems and approaches to their solution often arise at the intersection of various sciences. At present psychological research is increasingly turning to mathematical methods, widely using them to confirm statistical hypotheses, find relationships, predict and build mathematical models of various psychological phenomena, processes and situations.

In psychology, mathematical methods are widely used. This is due to several factors:

1) mathematical methods make it possible to make the process of studying phenomena clearer and more structured;

2) mathematical methods are necessary for processing a large number of empirical data (their quantitative exponents), for their generalization and organization in the empirical model of dependence research;

3) for the functional purpose of these methods and the needs of psychological science, there are two groups of mathematical methods, the use of which in psychological research is the most frequent: the first - methods of mathematical modeling, the second - methods of mathematical statistics (or statistical methods).

Thus, the need to apply mathematical methods in psychological science is due to the fact that psychological research currently cannot be carried out only from descriptive phenomenological positions, but requires the identification of the most objective quantitative and structural characteristics of the facts and phenomena under study, allowing us to consider the conclusions obtained as sufficiently reliable. For this purpose, special procedures used to compare the properties of numbers and geometric objects with mental phenomena and processes according to certain rules.

The relationship between mathematics and psychology is complex and ambiguous. On the one hand, many widespread psychological theories and concepts have proved their validity without any mathematical and statistical confirmation (the theory of psychoanalysis Z. Freud; the concept of humanistic psychology, etc.). On the other hand, the use of mathematical methods, of course, to a large extent objectifies the conclusions of psychologists, allows us to isolate from a huge set of disparate psychological data important factors and variables in one sense or another that have the greatest impact on the resulting feature under study.

Questions and tasks for independent work

1. The subject of psychology and the range of phenomena studied by it.
2. The evolution of the subject of psychology during the XIX-XX centuries: periods of crises and integration.
3. Psychology and "normal science": the problem of scientific criteria.

Task

Read Chapter 5 **Psychological Measurement** in the textbook Research Methods in Psychology:

Price, P. C. Research Methods in Psychology. 2nd Canadian Edition / I-Chant A. Chiang, Rajiv S. Jhangiani, and Paul C. BC Campus, 2015.

Write your abstracts on the following points:

1. Define measurement.
2. Give several examples of measurement in psychology.
3. Distinguish the four levels of measurement.
4. Give examples of each, and explain why this distinction is important.

Lecture 2. Different approaches to the concept of probability. The algebra of probabilities. The formula of total probability

Lecture plan

1. Phenomenon and process. The law and the case. Determinism and uncertainty of natural and social phenomena and processes.
2. Probabilistic formalism of descriptive statistics. Randomness and probability. Event. Probability of events.

Basic concepts: phenomenon, process, law, probability, event

1. Phenomenon and process. The law and the case. Determinism and uncertainty of natural and social phenomena and processes

The world of events surrounding a person. An accidental phenomenon. Random event as the outcome of the experiment. A reliable, impossible, and random event. Probability theory as a mathematical science. The probability is mathematical. Notation used in probability theory. Classical scheme of probability theory. Geometric probability. Statistical probability. Subjective probability. The probability space.

The world that we can observe – both nature and society – characterized by causal relationships (connections) of events, phenomena and processes.

A **phenomenon** (plural, phenomena) is a general result that has been observed reliably in systematic empirical research. Phenomenon-a picture of a set of events, usually simultaneous. It is represented as a static description of the state of the population (scene).

Process - dynamics of a phenomenon, i.e. change, development of a phenomenon over time. It is a complete unfolding of events (scenes) that characterize the phenomenon in time (in a certain sense, a movie).

At the same time, two principles are found in cause-and-effect relationships-determinism and uncertainty. The understanding of these principles is contained in the procedure **" if ..., then ...",** by which the causal relationship is formalized. The answer at any time and under the given conditions of experience – **"if THIS, then UNAMBIGUOUSLY that"** - means *determinism* in the observed relationship. Conversely, the answer "**if THIS, then PROBABLY THAT"** means that the relationship is *uncertain*. Therefore, it is customary to distinguish between law and case.

A law is a formalized one-to-one relationship between observed events, phenomena, and processes that can be detected and confirmed experimentally; the immutability of experimental conditions and the independence of observers are important.

Case – a probabilistic relationship between observed events, phenomena, and processes, if the conditions of experience are unchanged.

There are two extreme philosophical concepts - the world is deterministic (subject to law) and the world is indeterminate (random). A pragmatic view of the nature of things is the recognition of the complex dependence of events, phenomena, and processes at various scales and levels of consideration, the dynamic asynchronous interweaving of many different causes that characterize and influence (micro-) dependencies. Probability theory and mathematical statistics provide researchers with tools that allow them to analyze and characterize this type of mass micro-behavior at the macro level, without going into details. For this purpose, statistical (probabilistic) models of phenomena and processes are constructed.

A model (probabilistic) is a qualitative and/or quantitative description of a phenomenon or process, at a higher or lower level of detail and / or abstraction, and claiming to be quite accurate. Often, the model discusses the nature of the inaccuracy and then points out ways to improve the accuracy of the model representations. Additional research may be conducted to detect unknown deterministic relationships between events; taking into account the frequency of repetition and/or mass observation of a random phenomenon and process, attempts may be made to improve accuracy by identifying stable features – statistics.

2. Probabilistic formalism of descriptive statistics. Randomness and probability. Event. Probability of events

Probability theory is the mathematical foundation of modern statistics. A number of experiments and problems are known, as well as their experimental analogues, which lead to classical probabilistic schemes. These are the toss of a symmetrical coin, dice, card games, games with urns containing various objects. As a result of solving these problems, the formal foundations of probability theory were laid, which are briefly discussed below.

Random event and probability. Statistical research involves complex natural, technological, or social mechanisms whose results and manifestations cannot be reliably predicted. Therefore, statistics, like probability theory, considers the studied phenomena as random. In experiments, random phenomena are represented by random events.

Example of product production: you need to formulate a statistical experiment to assess the quality of the production process. In this case, a random event may, for example, be the appearance of a certain number of defective products in the control batch of products. To describe the random phenomenon under study, we compare a certain quantitative measure to the random event that represents it, which would reflect the frequency of occurrence of this event (for example, how often 5 defective products are found in a batch of 100 products). This measure is called *probability* $\boldsymbol{P}$ in theory and is understood as a function of the space of elementary random events. This is a positive real number in the range from 0 to 1. Probability of an impossible event is considered to be 0, and the probability of a reliable event is 1.

Random variable and its description. When conducting experiments, random events are usually correlated with a certain value $\boldsymbol{X}$, which is usually called **random**. Depending on the nature of the phenomenon under study, the corresponding random variables can be discrete (for example, the number of defective products in a batch) or continuous (such as the value of blood pressure).

A description of a random variable is given by its **distribution function**. The distribution function for a particular value *x* is determined by the probability of the event that the random variable X will take a value less than this particular x: $F(x) = P(X < x)$.

The distribution function is a positive nondecreasing function that takes values from 0 to 1 (Fig. 1) For discrete random variables, you can specify the probability $P_i=P(x_i)$ of a particular value x_i appearing. For continuous random variables, this description is the first derivative of the distribution function, which is called the **probability density** of the values of the random variable *x* (sometimes it is simply called the distribution of the random variable):

$$F'(x)=\frac{\partial F}{\partial x}=f(x)$$

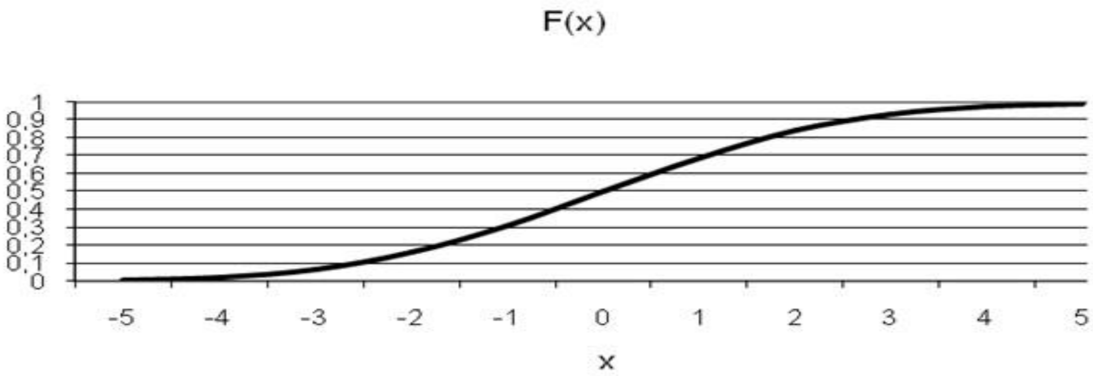

Fig. 1. Distribution function of a random variable

To understand the meaning of probability density, you can use the analogy with the mass and density of the body, representing the distribution function as the mass of the body, the volume of the body – the range of values of a random variable, the body density – the probability density. This is a positive function whose area under the curve relative to the axis of the random variable values is equal to one (Fig. 2).

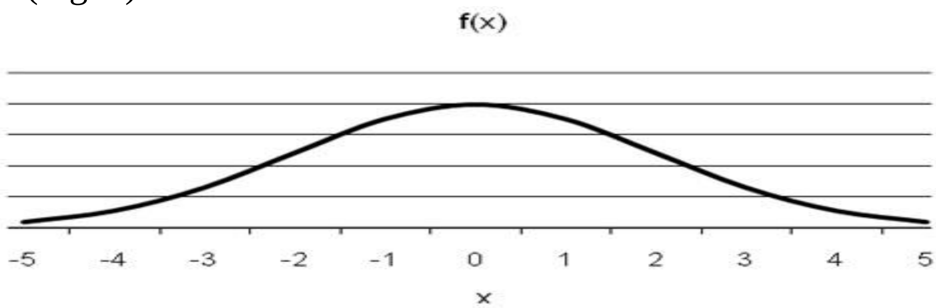

Fig. 2. The probability density of a random variable

Thus, the random variable is fully characterized by its distribution function (for discrete and continuous variables) or probability density (for continuous variables).

For a more compact (partial) description of a random variable, use numerical characteristics (or statistics) of random variables. The most important of them:

- **mean value *μ***.
- **dispersion *σ2***, or the **root mean square (standard) deviation *σ***, which is equal to the square root of the dispersion.

The *average value* is some central value of a random variable. It determines the position of the probability density curve along the *x*-axis of the random variable values. In particular, as the average value of μ increases, the probability density curve shifts to the right (Fig. 3).

The *standard deviation* characterizes the spread of the main mass of values of a random variable relative to the average.

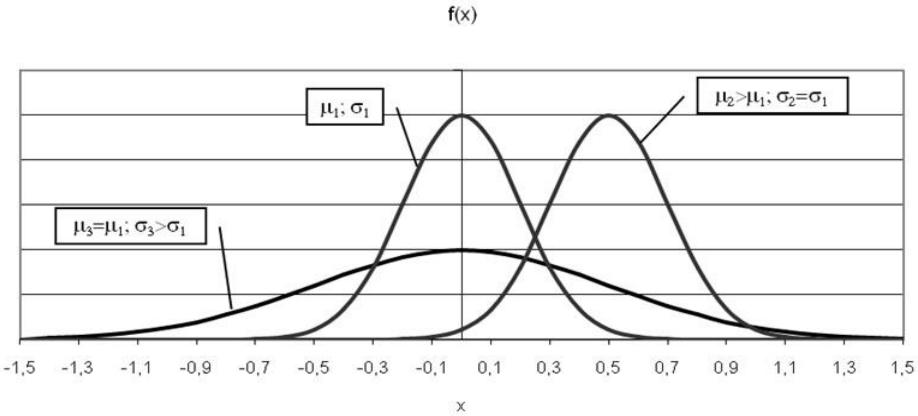

Fig. 3. The change curve of the probability density with increasing mean and variance

The results of observations are interpreted as realizations of the corresponding random variable. All possible (real or speculative) realizations of a random variable make up its so-called **General population**. The General population is rarely observed in the experiment, mainly due to limited resources (time, cost of the survey, capacity of information storage devices, etc.).

Statistical experience (experiment). The totality of observations from the point of view of practical statistics is data. It is known that arithmetic operations can be performed with classical functions as with functions. Probabilities can also be used as functions, and the laws of addition and multiplication of probabilities are formalized in probability theory. This is the basis of statistical calculations. Getting a specific implementation of a random variable is called statistical experience. Absolute immutability of the conditions of any experience cannot be ensured, so in practice it is only a question of ensuring relatively stable conditions. Meeting the requirements of experience stability is the subject of the data collection methodology. It is necessary to take into account the possibility that the data were obtained in conditions of non-constant conditions of experience; for subsequent processing and analysis of such data, adequate methods should be used.

Central limit theorem. The most common probability model in statistics is the normal (Gaussian) model. It reduces a large number of practical problems of probability theory and mathematical statistics. The General set of this model is a set of outcomes of statistical experience that reflects the additive interaction of a large number of independent random phenomena (processes). This probabilistic model belongs to the class of so-called **parametric models**, i.e. models whose analytical form is uniquely defined using a finite number of parameters. In the Gaussian case, the model parameters are the mathematical expectation μ (the average value of the General population) and the dispersion σ^2 - a measure of the spread of the values of a random variable relative to the mathematical expectation. The prevalence of the Gaussian model is explained by a fundamental property of nature that characterizes the probability distribution of the sum of a large number of independent random phenomena (processes). This property, formulated as a **Central limit theorem**, is shown in the fact that the sum of n equally distributed random variables tends to a normal (Gaussian) distribution with an infinite increase in n. (In practice, the effect of additive interaction of the order of $n=30$ independent random variables is sufficient.)

Along with parametric models, **nonparametric models** are also used in statistics. In nonparametric models, the probability distribution law of the General population is assumed to be unknown.

Questions and tasks for independent work

1. Concepts of variable, case.
2. Types of variables in relation to hypotheses.
3. Samples. The concepts of sampling and the general population.
4. The concept of distribution. A general idea of the law of large numbers.

Task

Using the literature (a list of basic and additional literature), make a comparative characteristic: how do nonparametric models differ from parametric models.

Lecture 3. The types of random events. Algebra of events

Lecture plan

1. Probability theory
2. The basic operations on events and the concept of the event algebra

Basic concepts: events, dependent events, independent events, random events, algebra of events, probability theory

1. Probability theory

Mathematical statistics is related to another mathematical science – probability theory and is based on its mathematical apparatus.

Probability theory is the science that studies the patterns generated by random events. Pedagogical phenomena are among the mass ones: they cover large groups of people, repeat from year to year, and occur continuously. Indicators (parameters, results) of the pedagogical process are probabilistic in nature: the same pedagogical influence can lead to different consequences (random events, random variables). Nevertheless, when conditions are repeatedly reproduced, certain consequences appear more often than others – this is the manifestation of so-called statistical regularities (which are studied by probability theory and mathematical statistics).

A random variable (**RV**) is a numerical characteristic that is measured over the course of an experiment and depends on a random outcome. RV implemented in the course of experience is itself random. Each RV sets the probability distribution. The main property of pedagogical processes and phenomena is their probabilistic nature (under these conditions, they can occur, be realized, but may not occur). For such phenomena, the concept of probability plays an essential role.

Probability (P) shows the degree to which a given event, phenomenon, or result can occur. The probability of an impossible event is zero, and the probability of a reliable event is one (100%). The probability of any event is between 0 and 1, depending on how random the event is.

If we are interested in event *A*, we can most likely observe and record the facts of its occurrence. The need for the concept of probability and its calculation will arise only when we do not observe this event every time, or realize that it may or may not happen. In both cases, it is useful to use the concept of the frequency of occurrence of an event *f(A)* - as the ratio of the number of cases of its occurrence (favorable outcomes) to the total number of observations. The frequency of occurrence of a random event depends not only on the degree of randomness of the event itself but also on the number (number) of observations of this random variable.

There are two types of RV samples: **dependent** and **independent.**

If the results of measuring a property in the objects of the first sample do not affect the results of measuring this property in the objects of the second sample, then such samples are considered **independent**.

In cases where the results of one sample affect the results of another sample, the samples are considered **dependent**. The classic way to get dependent dimensions is to measure the same property (or different properties) twice in members of the same group.

Event A does not depend on event B if the probability of event A does not depend on whether or not event B. Events A and B are independent if P(AB)=P(A)P(B). In practice, the independence of an event is established from the conditions of experience, the researcher's intuition, and practice.

RV can be discrete (we can number its possible values), for example, the roll of the dice =4,6,2, and continuous (its distribution function F(x) is continuous), for example, the service life of a light bulb.

Mathematical expectation – a numerical characteristic of RV, approximately equal to the average value of RV:

$$M_{(x)}=x_1p_1+x_2p_2+...+x_np_n$$

Thus, when constructing a theory, it is necessary to enter a number **P (A)** called the probability of an event **A**, which is implemented using one of the axioms, which is called the axiom of the existence of probability.

Next, we need to consider the basic properties of frequencies and Express these properties as statements about the properties of probabilities. These statements, together with the postulate of the existence of probability, form a system of axioms of probability theory.

The frequency **v(A)** can be considered as the result of measuring (estimating) the probability **P(A)** from experimental data. Thus, the equality **P (A)=p** means that for a large number of **N** experiments, **v (A)≈p**, and the error **|v(A)−p|** tends to decrease with increasing **N**.

Since **0≤n≤N**, the frequency **v(A)=**$\frac{n}{N}$ of occurrence of event A in a series of N experiments satisfies the condition **0≤v (A)≤1**. A similar condition must be met by the probability: **0≤P(A)≤1.**

Consider the probability values at the boundaries of the interval [0,1]. Let P (A)=0, then the event A is called impossible and is denoted by the symbol ∅. For an impossible event, its frequency is v(A)≈0 and tends to approach zero with an increasing number of N experiments. If P(A)=1, then event A is called valid and is denoted by the symbol E. The frequency of a reliable event v(E)≈1 and with an increase in the number of N experiments tends to approach 1.

2. Basic operations on events and the concept of the algebra of events

Let **A** be an event.

An event **A complement is an event** $\bar{A}$ that consists of the fact that event **A** did not occur. Operations on events can be given a simple geometric interpretation. Consider this interpretation of the complement operation.

Let the experiment consist in randomly throwing a point on a plane, and the set of conditions **U** is such that the outcome of each experiment is that the point falls into the area **E** of the plane.

By definition, $\bar{A}$ is an event that **A** did not occur. Therefore, in this interpretation, $\bar{A}$ is the non-occurrence of a point in the region **A.**

Union (or sum) of two events **A** and **B** is called the third event **C**, which consists in the fact that at least one of the events **A** or **B** occurred. To combine, we will use the notation **C=A∪B** or **C=A+B**. The "or" between two events can be used as an indication of the operation of combining two events. The union operation can also be interpreted geometrically. Let **A** be an event consisting in the fact that a point accidentally thrown on the plane falls into the area also denoted by **A**. Similarly, the event **B** is the contact point in the area **B**. The union operation is defined for an arbitrary number of events. For example, the event $\mathbf{C=A\cup B\cup ...}$ is when at least one of the events **A, B, ...** occurs. The event $\mathbf{D=\cup^n_{i=1} A_i}$ is when at least one of the events **A1, ..., An** occurs. Obviously the join operation is commutative by definition: $\mathbf{A\cup B=B\cup A}$, and associative, which also follows from the definition: $\mathbf{(A\cup B)\cup C=A\cup(B\cup C)}$.

Intersection of two events **A** and **B** is the third event **C**, which consists in the fact that both events **A** and **B** occurred. To denote the intersection operation, we will use the notation $\mathbf{C=A\cap B}$ or **C=AB**. The intersection operation, as well as the union operation, is defined for an arbitrary number of events. For example, the event $\mathbf{C=A\cap B\cap ...}$ is that all events **A, B, …** occur. The event $\mathbf{D=\cap^n_{i=1} A_i}$ is that all events **A1, ..., An** occur. By definition, the intersection operation is commutative, that is, the condition: $\mathbf{A\cap B=B\cap A}$ is satisfied, and it is also associative:

$$(A\cap B)\cap C=A\cap(B\cap C).$$

The union $\cup$ and intersection $\cap$ operations are mutually distributive. In particular, the union operation is distributive with respect to the intersection:

$$A\cup (B\cap C)=(A\cup B)\cap(A\cup C).$$

Note that if the " + " sign is used for the union operation, and the absence of a sign is used for the intersection, it takes a well-known form: A(B+C)=(AB)+(AC) - the Distributive Law of multiplication with respect to addition in the algebra of numbers. In contrast, the Distributive Law of addition with respect to multiplication has no analog in the algebra of numbers.

The considered operations on events are algebraic in nature. Therefore, in probability theory, the algebra of events is important, which is defined as follows.

An event system **F** is called an event algebra if, for any pair of events **A** and **B**, it follows from the conditions $\mathbf{A\in F}$, $\mathbf{B\in F}$ that the events $\bar{A}$, $\bar{B}$, $\mathbf{A\cap B}$, $\mathbf{A\cup B}$ are contained in **F**.

An event algebra is a system of events closed with respect to the operations of addition, intersection, and union.

Event **A** is called impossible if **P (A)=0**. We will use the $\emptyset$ symbol to indicate an impossible event. Event **A** is called valid if **P (A)=1**. A valid event is indicated by the symbol **E.** Obviously $\emptyset\cap\mathbf{E}=\emptyset$, $\emptyset$**=E.** Events **A** and $\bar{A}$ are called opposite events. There are equalities $\bar{\bar{A}}$**=A**, $\mathbf{A\cap\bar{A}=}$, $\mathbf{A+\bar{A}=E}$.

Events A and B are called incompatible if $\mathbf{A\cap B=\emptyset}$. Since ***overline AA=***$\emptyset$***,*** events **A** and $\bar{A}$ are incompatible. Events **A1, ... , An** form a complete group if $\mathbf{\cup^n_{i=1}A_i=E}$. This means that at least one of the events that form a complete group will appear as a result of the experience.

Events **A** and **B** are called independent if **P(A)** does not depend on whether event **B** occurred or not, and Vice versa, **P(B)** does not depend on whether event **A** occurred or not. If event **A** occurs whenever event **B** occurs, then **A** is called a consequence of event **B**, this is written as the relation **B⊂A** or **A⊃B**, which is read as *"from **B** follows **A**"* and **A***"is a consequence of **B**"*. The relation of the corollary can be given a geometric interpretation. If **A⊂B** and **B⊂A**, then the events **A** and **B** are called equivalent, this is written as **A=B**.

Event **C**, which consists in the fact that event **A** occurred and event **B** did not occur, is called the difference between events **A** and **B** and is denoted by **C=A B.**

Questions and tasks for independent work

1. The space of elementary events.
2. Random events.
3. The concept of probability.
4. Probability distribution.

Task

Using the literature, define the sample and the general population.

Give examples.

Lecture 4. Random variables and their characteristics. The law distribution of a random variable

Lecture plan

1. Random variables and their characteristics
2. Laws of distribution of random variables
3. Binomial distribution (Bernoulli distribution)

Basic concepts: random variable, discrete random variable, continuous random variable, distribution.

1. Random variables and their characteristics

A random variable is a quantity that, as a result of experience, can take on a particular value, it is not known in advance what it is.

A discrete (discontinuous) random variable is a random variable that takes separate values that can be renumbered.

A continuous random variable is a random variable where the data can take infinitely many values. For example, a random variable measuring the time taken for something to be done is continuous since there are an infinite number of possible times that can be taken.

The law of distribution of a random variable is any relation that establishes a relationship between the possible values of a random variable and the corresponding probabilities. The law of distribution can have different forms.

A series of distributions of a discrete random variable **X** is called a table, which lists the possible (different) values of this random variable **x1, x2,..., xn** with their corresponding probabilities **p1, p2,..., pn.**

x_i	x_1	x_2	**...**	x_n
p_i	p_1	p_2	**...**	p_n

The law of distribution of a random discrete quantity (**X**) is any relation that establishes a relationship between the possible values of a random variable (**x1, x2,..., xn**) and the corresponding probabilities (**p1,p2,..., pn**). In this case, the events (**x1,x2,...,xn**) form a complete group (i.e. the occurrence of one of them is a reliable event), which means

$$\sum_{i=1}^{n} p_i = 1$$

In this case, the random variable **X** is said to be subject to this distribution law.

2. Laws of distribution of random variables

Are random phenomena obeying any laws? Yes, but these laws are different from the physical laws we are used to. The values of random variables cannot be predicted even under known experimental conditions, we can only indicate the probability that a random variable will take a particular value. But knowing the probability distribution of random variables, we can draw conclusions about the events in which they participate. However, these conclusions are also probabilistic.

Let some random variable be discrete, i.e. it can only take fixed values of **Xi**. In this case, a series of probability values **P (Xi)** for all **i=1...n** acceptable values of this quantity is called its distribution law.

The law of distribution is a relation that establishes a relationship between the possible values of a random variable and the probabilities with which these values are accepted. The distribution law fully characterizes the random variable.

When building a mathematical model to test a statistical hypothesis, you must enter a mathematical assumption about the distribution law of a random variable (parametric way to build the model).

The nonparametric approach to describing a mathematical model (a random variable does not have a parametric distribution law) is less accurate, but has a broader scope.

Currently, there are a sufficient number of various packages of computer applications for these purposes. Among all probability distributions, there are those that are used in practice especially often, they are studied in detail and their properties are well known. Many of these distributions underlie entire fields of knowledge, such as reliability theory, quality control, game theory, and so on.

Binomial distribution (bernoulli distribution)

It occurs when the question is raised: how many times an event occurs in a series of a certain number of independent observations (experiments) performed under the same conditions.

For convenience and clarity, we assume that we know the value **p** – the probability that the user who entered the store will be a buyer and **(1–p)=q** – the probability that the user who entered the store will not be a customer.

If **X** is the number of customers out of the total number of **n** users, then the probability that there are **k** customers among n users is equal to

P(X=k)= $\frac{n!}{k!(n-k)!}p^k q^{n-k} = C_n^k p^k q^{n-k}$, where k=0,1,..., n.

Formula (1) is called the *Bernoulli formula*. For a large number of tests, the binomial distribution tends to be normal.

Poisson distribution

Plays an important role in a number of issues of physics, reliability theory, queuing theory, etc. Wherever a random number of events may occur over a certain period of time (radioactive decay, telephone calls, equipment failures, accidents, etc.).

Consider the most typical situation in which a Poisson distribution occurs. Let's assume that some events (purchases in the store) can occur at random times. Let's determine the number of occurrences of such events in the time interval from 0 to T. The random number of events that occurred during the time from 0 to T is distributed according to Poisson's law with the parameter l=aT, where a>0 is the problem parameter that reflects the average frequency of events. The probability of *k* purchases over a large time interval (for example, – day) is P(Z=k) = $\frac{\lambda^k}{k!}e^{-\lambda}$

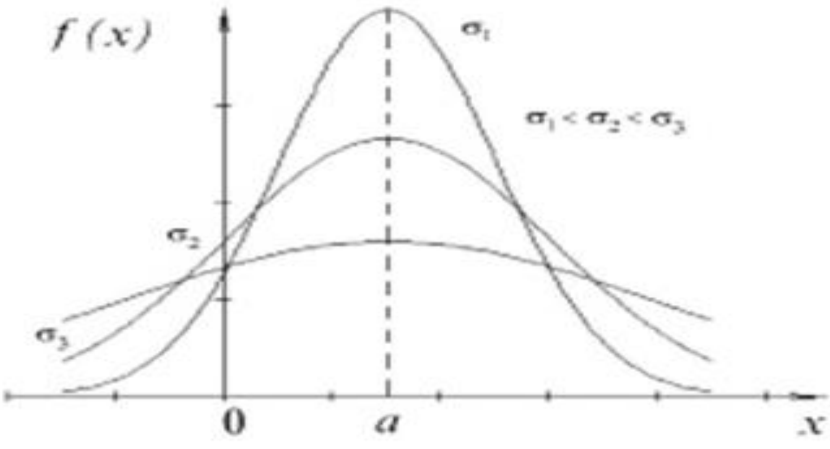

Fig. 4 - Poisson distribution

Normal (gaussian) distribution

The normal (Gaussian) distribution is Central to the theory and practice of probabilistic statistical research. It was discovered and studied by K. Gauss (1809) and P. Laplace, who came to the normal function in connection with the work on the theory of astronomical observation errors. A continuous random variable \(X\) is called distributed according to the normal law if its distribution density is equal to

$$f(x)=\frac{1}{\sqrt{2\pi\sigma}}\exp-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$$

$$-\infty<x<\infty,\ \sigma>0,\ -\infty<\mu<\infty$$

where **μ** coincides with the mathematical expectation of the value **X: μ=M(X)**, the parameter s coincides with the mean square deviation of the value **X: =s(X).**

The graph of the normal distribution function, as can be seen from the figure, has the form of a dome-shaped curve, called a Gaussian, the maximum point has coordinates

$$(a;\frac{1}{\sqrt{2\pi}})$$

This means that this ordinate decreases with increasing value of ***s*** (the curve "shrinks" to the Ox axis) and increases with decreasing value of ***s*** (the curve "stretches" in the positive direction of the OU axis). Changing the values of the parameter **μ** (with the constant value of ***s***) does not affect the shape of the curve, but only moves the curve along the Ox axis.

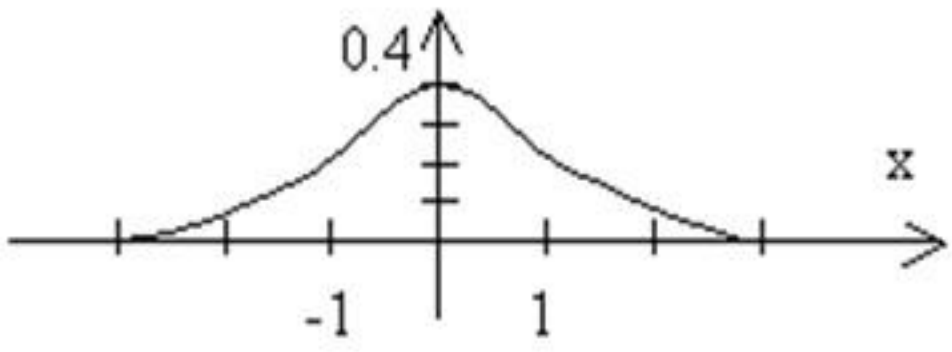

Fig. 5 - A normal distribution

A normal distribution with parameters **μ=0** and **s=1** is called a normalized distribution. The distribution function of a random variable in this case will have the form:

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

This curve at **μ=0**, **σ=1** has received the status of a standard, it is called a unit normal curve, that is, any collected data tends to be transformed so that the curve of their distribution is as close as possible to this standard curve.

The normalized curve was invented to solve problems in probability theory, but it turned out in practice that it perfectly approximates the frequency distribution for a large number of observations for a set of variables.

Uniform distribution

The uniform probability distribution is the simplest and can be either discrete or continuous. A discrete uniform distribution is a distribution for which the probability of each of the values of a random variable is the same, i.e.:

$$P(X)=\frac{1}{N}$$

where N is the number of possible values of the random variable.

The probability distribution of a continuous random variable X that takes all its values from the segment [a; b] is called uniform if its probability density on this segment is constant, but outside it is zero:

$$f(x)=\begin{cases} 0 & \text{if } x<a \\ \frac{1}{b-a} & \text{if } a\le x\le b \\ 0 & \text{if } x>b \end{cases}$$

Student's t-distribution (or simply the t-distribution)

One of the tasks for a researcher in psychology is to compare two distributions that may differ in terms of averages, variance, asymmetry, kurtosis and a combination of these parameters. The distributions may also differ in the frequencies of each bit interval. The differences between the distributions can be detected using the parametric t–Student criterion and the nonparametric Pearson and Kolmogorov-Smirnov criteria.

If the random variables *x1, x2, ... xn* are independent, and each of them has a standard normal distribution N(0,1), then the random variable has a distribution called the Student's distribution:

$$t_n=\frac{x_0}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}}$$

The Student's t –test is applied when the task is to compare the averages of two distributions. The criterion is based on an assessment of the common parts of two distributions.

The limitation of the criterion is that the distributions must be normal.

Questions for independent work

1. Equally possible events. Compatible and incompatible events. Pairwise incompatible events. Reliable and impossible events as categories related to a particular test.
2. A complete system of events. A complete system for a pair of incompatible events.
3. The concept of an event.
4. Random and non-random events.
5. Measures of the possibility of occurrence of the event.
6. The concept of the event system.
7. A random variable as a function on the space of elementary events.
8. Algebra of random variables. Numerical characteristics of a random variable.

Lecture 5. Mathematical expectation and variance of a discrete random variable

Lecture plan
1. Descriptive statistics
2. Graphical visualization of sample data
3. Sample mean value

Basic concepts: minimum, maximum, mean or average, mode, median, quartile, percentile, Chebyshev formula

1.Descriptive statistics

The first section of mathematical statistics – descriptive statistics-is intended for presenting data in a convenient way and describing information in terms of mathematical statistics and probability theory. The main quantity in statistical measurements is a unit of the statistical population (for example, any of the criteria for evaluating the quality of a teacher-researcher). A unit of a statistical population is characterized by a set of attributes or parameters. The values of each parameter or attribute can be different and generally form a series of random values **x1, x2,..., xn.**

A variable is a measurement parameter that can be controlled or manipulated in a study. Since the values of variables are not constant, you need to learn how to describe their variability. For this purpose, descriptive or descriptive statistics are invented: minimum, maximum, mean, dispersion, standard deviation, median, quartiles, and mode.

The relative value of a parameter is the ratio of the number of objects that have this metric to the sample size. Expressed as a relative number or as a percentage (percentage value).

The specific value of this attribute is a calculated value that shows the number of objects with this indicator that would be contained in a conditional sample consisting of 10, or 100, 1000, and so on.

Minimum and maximum are the minimum and maximum values of a variable.

Mean or average (estimate of the mean, sample mean) – the sum of the variable values divided by n (the number of variable values). If you have the values *X(1),..., X (N),* then the formula for the sample mean is:

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{x_1 + x_2 + ... + x_N}{N}$$

The sample mean is the point, the sum of the observation deviations from which is equal to 0. Formally, this is written as follows:

$$(\bar{x} - x_1) + (\bar{x} - x_2) + ... + (\bar{x} - x_n) = 0$$

To estimate the **degree of variation (deviation)** of an indicator from its average value, along with the maximum and minimum values, the concepts of dispersion and standard deviation are used.

Sample variance is a measure of the changeability of a variable. The term was first introduced by Fischer in 1918. The sample variance is calculated using the formula:

$$s^2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N}$$

where **`x** is the sample mean, **N** is the number of observations in the sample.

The variance varies from zero to infinity. An extreme value of 0 means that there is no variability when the variable values are constant.

Standard deviation, the mean square deviation is calculated as the square root of the variance. The higher the variance or standard deviation, the more widely the variable values are scattered relative to the mean.

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

The Median is the score found at the exact middle of the set of values. One way to compute the median is to list all scores in numerical order, and then locate the score in the center of the sample. For example, if there are 500 scores in the list, score #250 would be the median. If we order the 8 scores shown above, we would get: 15, 15, 15, 20, 20, 21, 25, 36

There are 8 scores and score #4 and #5 represent the halfway point. Since both of these scores are 20, the median is 20. If the two middle scores had different values, you would have to interpolate to determine the median.

Quartile are values that divide the two halves of the sample (split by the median) in half again (from the word quart-quarter).

Distinguish the top quartile, which is greater than the median and bisects the upper part of sample (values of variable is more than median) and lower quartile which is less than median and bisects the lower part of the sample. The lower quartile is often indicated by the symbol **25%**, which means that **25%** of the variable values are less than the lower quartile. The upper quartile is often indicated by the symbol **75%**, which means that **75%** of the variable values are less than the upper quartile.

The Mode is the most frequently occurring value in the set of scores. To determine the mode, you might again order the scores as shown above, and then count each one. The most frequently occurring value is the mode. In our example, the value 15 occurs three times and is the model. In some distributions there is more than one modal value. For instance, in a bimodal distribution there are two values that occur most frequently.

Notice that for the same set of 8 scores we got three different values (20.875, 20, and 15) for the median and mode respectively. If the distribution is truly normal (i.e., bell-shaped), the mean, median and mode are all equal to each other.

Asymmetrical distribution is a situation in which the values of variables occur at irregular frequencies and the mean, median, and mode occur at different points. An asymmetric distribution exhibits skewness.

Kurtosis is a measure of the steepness of the distribution curve. The kurtosis is equal to:

$$E_x = \frac{\sum (x_i - \bar{x})^4 / n}{\sigma_x^4} - 3$$

2. First look at the data. Graphical visualization of sample data. Scatter plot.

The quality of the data to be analyzed must be evaluated in advance. Data can be taken as a sample from a known General population or obtained by modeling. The model way of forming a sample belonging to a certain General population is to obtain data using a special computer program – **a random number generator**.

The *MS Excel* analysis package includes the utility **Sample**. Creates a sample from the General population, considering the input range (data) as a General population. If the population is too large for processing or plotting, you can use a representative sample.

To analyze data quality, it is useful to first perform a graphical visualization of the sample. A dotted graphical representation based on the sequence number, time, or other appropriate category of sample data is called a **scatter plot.** The experiment number is plotted on the horizontal axis, and the sample values of a random variable are plotted on the vertical axis. The scatter plot allows you to visually assess the area of localization (concentration) and the degree of data dispersion.

A quantitative description of the sample data

After building and filling in the tables with sample data, they begin to describe them numerically. Determine the volume – the amount of data and the range of changes in the random variable in the sample-the difference between the maximum and minimum values in the sample (span). To create a **histogram** – selective (statistical) way of probability density function the range of variation of a random variable is marked at intervals and starts the procedure of sorting data, which notes the **frequency** hit count data from the sample at an appropriate interval and builds the corresponding graphical image. There are absolute and relative frequencies. The latter are defined as the number of hits in the intervals divided by the sample size (the total amount of data). The sum of the relative frequencies in the histogram is 1, and the relative frequencies themselves can be expressed as a percentage. To construct a histogram, first decide how many bars or intervals, also called classes, represent the data. Many histograms consist of five to bars or classes for clarity. The number of bars needs to be chosen. The MS Excel analysis package uses the utility **histogram.**

Descriptive statistics are broken down into measures of central tendency and measures of variability. Measures of central tendency include *the mean, median and mode, while measures of variability include standard deviation, variance, minimum and maximum variables, and kurtosis and skewness.*

As a result of statistical analysis of the sample using descriptive statistics, we get point or interval estimates of the parameters of the General population.

Point estimates are represented by a single number. The following parameter estimates are point estimates.

The *sum* is calculated by summing all the sample data, taking into account the signs, and is denoted by $\sum X$.

The *minimum* and *maximum* values are denoted, respectively, as min X and max X.

The *interval (span)* of the sample is defined as the difference between the largest and smallest values min X - max X.

The *average (arithmetic)* value of the sample data is calculated using the formula:

$$\bar{X} = \frac{\sum X}{N}$$

The *standard error* of the mean is rarely used because of the complexity of using it in further calculations.

The *standard deviation* of the sample data (from the average) is calculated using the formula:

$$s = \sqrt{\frac{\sum (X - \bar{X})^2}{N-1}} = \sqrt{\frac{1}{N-1} \cdot \left(\sum X^2 - \frac{(\sum X)^2}{N} \right)}$$

The *variance* of the sample data is defined as S^2.

It should be emphasized that the point estimate is a random variable, since its value differs in different samples when a random phenomenon is observed.

Interval estimates are represented by a pair of numbers (the boundaries of a certain interval); interval estimates are given together with the probability or **level of reliability (confidence)** (the estimated value falls within the specified **confidence interval**).

Reliability level – the probability that the true value of the estimated statistics is in the confidence interval constructed (most often based on a point estimate). The reliability level is often set as a percentage.

There is often a need to group and / or rank data. From the histogram and frequency polygon for each class, the **frequency and cumulative frequency** are known, respectively. Standard groups formed in statistics are **percentiles (percentile ranks), deciles, quartiles**, and so on.

Percentile – a number indicating which percentage of data is below or above the specified value. To calculate the percentile, use the formula

$$PercentileRank = L\% + \left(\frac{score - LRL}{h} \cdot I\% \right)$$

where *L%* is the percentage of data lying below the (critical) interval*; I%* – the percentage of data that belong to a given interval; *LRL* – lower real border of the specified interval; *h* is the size (step) interval; *score* – the value for which is determined by a percentile. Every 10th percentile is called a decile, and every 25th quartile (the second quartile corresponds to the median). In *MS Excel*, the corresponding procedures are included in the **Rank and percentile** analysis package utility. Used to output a table containing the ordinal and percentage ranks for each value in the data set. This procedure can be used to analyze the relative position of data in a set.

3. Sample mean value

The sample mean is a statistic whose expected degree of variability from a sample from a given population is less than the variability of the sample data itself. It is also known that the sample mean is an unbiased estimate of the average of the General population (the mathematical expectation of a random variable). However, the sample average remains (as opposed to the mathematical expectation!) random variable. Thus, for the sample average, it makes sense to have their own statistics – the sample mean, variance, and so on.

The problem of determining confidence intervals for the sample average is a classic statistical problem. Three cases are typical:

- The General population is distributed according to a normal law with a known standard deviation – the parameters of the normal law are used to solve the problem.

For example, there is a random sample of length N=144 with an average value X=100 taken from a General population of length N=1000 with a known standard deviation σ=60. Then the 95% confidence interval for an unknown value of the average of the General population (mathematical expectation) is calculated as follows, i.e. with 95% confidence, it can be stated that the average of the General population lies in the range between 90.89 and 109.11.

- The General population is distributed according to the normal law, but the standard deviation is unknown, and the sample length n is less than 30 – the parameters of the student's t-distribution with (n-1) degree of freedom are used to solve the problem;

For example, there is a random sample of length n=144 with an average value X=100 and a sample standard deviation S=60. Then the 95% confidence interval for an unknown value **μ** of the General population average (mathematical expectation) is calculated as follows:

$$\mu = \bar{X} \pm t_{0.025} \cdot \frac{s}{\sqrt{n}} = 100 \pm 2.16 \cdot \frac{60}{\sqrt{14}} = 100 \pm 34.637$$

that is, with 95% reliability, we can say that the middle of the General population lies in the range between 65.363 and 134.637.

• The distribution of the General population is unknown, but its standard deviation is known – the *K-parameter* from Chebyshev's theorem is used to solve the problem (an analysis that does not depend on the form of the distribution).

For example, there is a random sample of length with an n=14 average value X=100 taken from a General population of length N=1000 with a known standard deviation σ=60. Then the 95% confidence interval for an unknown value of the General population average (mathematical expectation) is calculated as follows:
$\mu = \bar{X} \pm K \cdot \sigma_{\bar{x}} = 100 \pm K \cdot \frac{\sigma}{\sqrt{n}} = 100 \pm K \cdot 16.035$ where *K* is calculated using the Chebyshev formula $K = \sqrt{\frac{1}{1-0.95}} \cong 4.47$ As a result, we have, i.e. with 95% reliability, we can say that the middle of the General population, from which the sample is taken, lies in the range between 28.321 and 171.679.

Questions for independent work
1. Grouping of statistical data.
2. Geometric interpretation of statistical distributions of the sample.
3. Normal distribution. Lyapunov's Theorem. Indicative distribution.
4. The mathematical expectation of a discrete random variable.
5. The mathematical expectation of the sum of random variables.
6. The variance of the sum of independent random variables.
7. Chebyshev's inequality and the law of large numbers.
8. The Poisson Distribution.

Lecture 6. Combinatorics.
General rules of combinatorics. Basic formulas of combinatorics

Lecture plan
1. Sum rule.
2. Samples.
3. Placements.

Basic concepts: combinatorics, samples, placements

1. Sum rule

Combinatorics is the study of collections of objects. Specifically, counting objects, arrangement, derangement, etc. of objects along with their mathematical properties.

The entire variety of combinatorial formulas can be derived from two main statements concerning finite sets – **the sum rule** and **the product rule**.

If an event **e1** can be done in n1 ways and an event **e2** can be done in **n2** ways and **e1** and **e2** are mutually exclusive, then the number of ways of both events occurring is **n1 + n2**

Example 1. If there are X books on the first shelf and Y on the second, then you can select a book from the first or second shelf in X+Y ways.

Example 2. The student must complete practical work in mathematics. He was offered a choice of 17 topics in algebra and 13 topics in geometry. How many ways can he choose one topic for practical work? *Decision:* according to the sum rule, we get 17+13=30 options.

A tuples is a finite sequence of elements of a set that can be repeated.

If two events are not mutually exclusive (that is, we do them separately), then we apply **the product rule.**

Suppose a procedure can be accomplished with two disjoint subtasks. If there are **n1** ways of doing the first task and **n2** ways of doing the second, then there are **n1 · n2** ways of doing the overall procedure.

Example 1. If there are 5 books on the first shelf and 10 on the second, then you can choose one book from the first shelf and one from the second in 5*10=50 ways.

Example 2. The bookbinder must bind 12 different books in red, green, and brown bindings. How many ways can it do this? *Decision.* There are 12 books and 3 colors, so according to the product rule, 12*3=36 binding options are possible.

2. Samples.

If a subset is selected from a set of items, it is called a sample. There are **ordered** and **unordered samples**. In an ordered sample, the order in which its elements follow is essential. In other words, by changing the order of the elements, we get a different sample.

Example 1. From the digits 1, 2, 3, 4, 5, you can make the following three-digit numbers 123, 431, 524, ...and so on. These are ordered three-element samples, since 123 and 132 are different numbers.

Example 2. Of the 20 students in the class to choose two persons on duty. Any pair of attendants is an unordered two-element selection, since the order of their selection is not important.

3. Placements.

Placements of **n** elements by **k** elements (k < n) are combinations made up of **n** elements by **k** elements that differ either by the elements themselves or by the order of the elements.

The number of **placements without repetitions** from **n** to **k** (**n** different elements) is calculated by the formula:

$$A_n^k = \frac{n!}{(n-k)!} \quad \textbf{(1)}$$

Placements with repetitions of **n** elements by **k** are ordered k-element selections in which elements can **be repeated**. The number of repetitive placements is calculated using the formula:

$$\bar{A}_n^k = n^k \quad \textbf{(2)}$$

Example. Let's take the letters B, A, and C. Which placements of these letters, taken two at a time, can be obtained? How many such sets will be obtained if: 1) the letters in the set are not repeated; 2) the letters can be repeated?

Decision.

The following sets are obtained: BA, BC, AC, AB, CB, CA according to the formula **(1)**, we get:

$$A_3^2 = \frac{3!}{(3-2)!} = \frac{1 \cdot 2 \cdot 3}{1} = 6 \text{ sets.}$$

The resulting sets are: BB, BA, BC, AA, AB, AC, CC, CB, and CA. Using the formula **(2)**, we get:

$$\bar{A}_3^2 = 3^2 = 9 \text{ sets.}$$

<u>Example.</u> There are 6 traffic lights along the road. How many different combinations of their signals can there be if each traffic light has 3 states: "**r**ed", "**y**ellow", "**g**reen"?

Decision.

Let's write out several combinations: RRRYGG, GGGGGG, RYGRYG and so on. We can see that the composition of the sample changes and the order of elements is significant (because if, for example, in the sample RYGRYG swap R and Y, the situation on the road will be different).

Therefore, we apply the formula **(2)** and calculate the number of placements with repetitions from 3 to 6, we get $\bar{A}_3^6 = 3^6 = 729$ combinations.

Questions and tasks for independent work

1. Sets and operations on them. Algebra of sets.
2. Splitting a set into subsets. Tuples and Cartesian product of sets.
3. Mapping of multiples. The sum rules.
4. The concept of factorial.
5. Placements and permutations.
6. Combinations and binomial coefficients.
7. Pascal triangle.

Task

Pascal's Triangle was also known to the Chinese in the 11th century. The Chinese Mathematician, Jia Xian devised a triangular representation of the coefficients of the binomial theorem in the 11th century. Later, another Chinese Mathematician Yang Hui further studied Jui Xian's triangle, finding more properties in it. His study and discoveries found in the triangle further popularized the triangle. In China they call the triangle Yanghui's Triangle.

What line of Pascal's triangle is 1, 10, 45, 120, 210, 252, 210, 120, 45, 10, 1?

MODULE 2
QUALITATIVE AND QUANTITATIVE APPROACHES TO PSYCHOLOGICAL RESEARCH

Lecture 7. Experiment in psychological research

Lecture plan
1. Experiment in psychological research
2. Types of experiment
3. Requirements for conducting the experiment
4. Stages of the experiment

Basic concepts: experiment, lab experiment, field experiment, formative experiment

1. Experiment in psychological research.

An experiment is a method of empirical research based on the active and purposeful creation of controlled conditions that allow us to identify essential properties and connections in the object under study. Experiment as a method of scientific research has the following *main features*:

- a more active attitude towards the object than during observation, up to its change and transformation;
- multiple reproducibility of the studied object at the request of the researcher;
- ability to detect properties and relationships of phenomena that are not observed in natural conditions;
- the possibility of considering the phenomenon as if in "pure form" by isolating it from complicating circumstances or by changing the conditions of the experiment;
- ability to control the" behavior " of the research object and verify its results;
- an experiment is always guided by an idea, a concept, a hypothesis;

Despite the variety of experiments, they can be classified on various grounds. There are real and mental basic *types of experiments.* In turn, real experiments are divided into *natural,* when the object of research is in natural conditions, which can change at the will of the experimenter; *model,* when the real object of research and the conditions affecting it are replaced by a model; *social,* when experiments are aimed at studying certain social phenomena. In modern science, a thought experiment is widely used – a system of mental procedures performed on idealized objects. A thought experiment is a theoretical model of real experimental situations. At the same time, the researcher operates not with real objects and conditions of their existence, but with conceptual samples of these objects. It is necessary to emphasize the important role of social experiments at the present time, which contribute to the introduction of new forms of social organization and optimization of social management. The object of a social experiment, in the role of which a certain group of people acts, is one of the participants in the experiment, whose interests have to be considered, and the researcher himself is included in the situation he is studying. The content and procedures of social experiments are also determined by the legal and moral norms of society.

Each method of empirical knowledge performs special, *specific functions*. The main functions of experimental methods include, first, cognitive. It is expressed in the search, detection, comparison and experimental analysis of new phenomena, their properties and relationships. A scientific experiment, as a rule, provides new information about the objects under study.

Secondly, experimental methods also perform a verification function. Scientific experiment as the most important element of practice always acts as an objective criterion of the truth of empirical and theoretical knowledge.

Third, experimental methods also have an analytical function associated with the analysis of the results of the experiment.

The main focus of this analysis is to find out whether the results of the experiment can serve as a basis for revising previous theoretical concepts, or they can form the basis for further development of these concepts, or for putting forward new scientific hypotheses.

Fourth, experimental methods also perform a demonstration function. Experimental methods are inextricably linked to observation, measurement, and comparison. Therefore, many of the functions performed by observation, comparison, and measurement are also inherent in experimental methods.

2. Types of experiment

There are three types of experiments you need to know:

- *Lab Experiment*. A laboratory experiment is an experiment conducted under highly controlled conditions (not necessarily a laboratory), where accurate measurements are possible.
- *Field Experiment*. Field experiments are done in the everyday (i.e. real life) environment of the participants. The experimenter still manipulates the independent variable, but in a real-life setting (so cannot really control extraneous variables).
- *Formative experiment* (combines the procedures of research and training, education, training.) It can be carried out in the form of trainings, games, solving specific situations, etc. The method is not limited to registering results, but creates special conditions for the development of personal qualities of the subjects. It is widely used in educational psychology.

3. Requirements for conducting the experiment

In the experiment, various experimental schemes and plans are distinguished:

1. The procedure of repeated measurement (when the subjects perform tasks under different conditions: separately and in the presence of others.) When organizing an experimental plan with repeated measurement, the same subjects are presented with different conditions of the independent variable and the dependent variable is measured.

2. Independent subjects and the intergroup scheme (when different groups perform the same task under different conditions of the independent variable).

3. Scheme of matched pairs (these are identical monozygotic twins. They are naturally matched pairs, at least at the time of birth).

4. Experiment with one participant (when one person takes part in the experiment and it takes a lot of time or when you just need to get acquainted with the problem in order to formulate some hypotheses.

Despite the specifics and features of various types of experimental method, the following *general requirements* for conducting an experiment, regardless of its types, are distinguished:

- the presence of the object of research, allowing the possibility of describing the system of variables that determine its behavior; the possibility of quantitative and qualitative measurements; control of influencing factors, the state of the object and conditions during the experiment;
- the description of the object of experimental observation is carried out in the system of its components;
- definition and description of the conditions of existence of the object of research (industry, type of production, working conditions, form of interaction of employees, etc.);
- the presence of a clearly formulated experimental hypothesis about cause-and-effect relationships;
- subject definition of the concepts of the formulated hypothesis of the experiment;
- a reasonable selection of the independent variable and the dependent variable;
- description of the specific conditions of the research object's activity (time, place, socio-economic situation, etc.).

In terms of *structure*, the experiment is a complex formation containing several successive stages of its implementation, certain important measurement techniques, tools and procedures.

4. Stages of the experiment.

There are the following *stages of the experiment*.

The first (preparatory) stage in the most general form includes a detailed development of the methodological and methodological sections of the experiment program, as well as a work plan for its organization and implementation.

The second (implementation) stage of the experiment is essentially the main stage, during which the subject-practical experimental activity of the experimenter and the subjects is carried out.

The third (final) stage of the experiment includes procedures for analyzing, processing and generalizing the results of the experiment into a reliable and confirmed scientific fact.

The preparation of an experiment consists of two blocks of operations: the justification of the experiment and the construction of a program for its implementation.

The *rationale for an experiment* begins, as a rule, with an analysis of its necessity, i.e., the study of the objective social need for experimental research.

Then, by understanding the available resources for the implementation of a sociological experiment, the possibilities of its implementation are identified. Finally, the possible costs and damage from the test are compared with the expected positive results and a conclusion is made about the need or impossibility of conducting the experiment.

Questions and tasks for independent work

1. Areas of use and specifics of the method. Types of causal relationships in psychology.
2. Basic concepts of the experiment. Types of experiments.
3. Experimental schemes (plans).
4. The problem of experiment quality: basic concepts and evaluation criteria.

Task

Search the Internet for descriptions of the most influential psychological experiments in history.

Lecture 8. Mixed qualitative and quantitative methods

Lecture plan

1. Characteristics of mixed research
2. Advantages and disadvantages of mixed studies
3. Types of mixed research

Basic concepts: "mixed" research, qualitative and quantitative methods

1. Characteristics of mixed research

To implement "mixed" research that combines qualitative and quantitative methods, it is important to understand the nature, capabilities and limitations of each approach. For example, the main characteristics of traditional quantitative research are a focus on deduction, confirmation of theories / hypotheses, explanation, standardized data collection, and statistical analysis. The main characteristics of traditional qualitative research are inductive logic, continuous search, exploration, hypotheses / theories, continuity of data collection and analysis. The researcher acts as the main tool, where his subjectivity is the basis of cognitive capabilities. Thus, the planes for comparing qualitative and quantitative methods are based on the following elements: induction / deduction; subjectivity / objectivity and contextuality / the ability to make generalizations. The views of the proponents of quantitative methods are based on the general propositions of positivism, where observables and objects are equated in properties with physical phenomena [Ayer, 1959; Maxwell, Delaney, 2004; Popper, 1959; Schrag, 1992].

"Mixed methods research" (or, briefly, MMR) originated in the early twentieth century. Thus, modern social researchers are among three methodological and scientific paradigms: quantitative, qualitative and "mixed". "Mixed research", understood exclusively in a qualitative and quantitative way, refers, rather, to a narrower interpretation of this direction. In a broader context, "mixed" research involves a wide variety of non-trivial (different from classical) combinations of methods of different classes in one study (using online and offline methods, verbal and visual data, etc.). The fundamental point of "mixed" design in research is the condition that the combination of methods is a combination of the strengths of each of them. Thus, the main argument of mixed research is that the data obtained from a combination of methods will exceed the possibilities from the application of a single method. This strategy implies a special type of research, where quantitative and qualitative research approaches, methods, techniques and description languages are integrated [Burke, Onwuegbuzie, 2004].

The philosophical platform of the strategy of mixing methods is classical pragmatism, represented primarily by the ideas of Ch. Peirce, W. James, and J. Dewey. Adherents of the school of pragmatism believe that scientific methods should be applied to find the "best" answer out of all possible ones. Here, an attempt is made to consider various points of view, perspectives and views, including the optics of qualitative and quantitative research. In the broadest sense, pragmatism is a kind of conceptual framework that allows you to combine methods and research tasks. From a pragmatic point of view, research is one of the forms of action to achieve the goal, which is based on solving research questions [Morgan, 2013, c. 42–43].

2. Advantages and disadvantages of mixed studies

The advantage of mixing methods is the mutual enrichment of cognitive capabilities, data, and interpretations. Qualitative data (words, photos, and narratives) can be used to adequately understand numbers. In turn, numbers can be used to understand words, photos, and narratives. In general, the complex qualitative and quantitative design provides an opportunity to answer a wider range of research questions, and more reasonable generalizations of data are assumed. Among the most obvious disadvantages are procedural ones associated with increased resource costs: first, temporary (long – term), second, financial ("expensive" complex projects), and third, human (usually requires specialists who combine the approaches of both paradigms). The strategy of "mixing" includes many similar names. So, there is a mention of "combined" research [Thomas, 2003], integrative studies [Johnson, 2004], multi / multimethod studies [Hunter, Brewer, 2003; Morse, 2003], triangulated studies [Sandelowski, 2003], and mixed research [Johnson, 2004, et al.].

3. Types of mixed research

Mixed studies as a broader and more universal concept is divided into subtypes. First of all, there are mixed methods, where qualitative and quantitative approaches are integrated within one or several stages of the research process. Secondly, there are actually mixed methods (mixed methods), where it is assumed to include quantitative and qualitative stages in the study, the use of which is relatively autonomous.

R. Burke et al., in turn, divide mixed studies into five groups:

1. Selected based on the item being mixed (what is being mixed).
2. Compose studies where mixing occurs at different stages.
3. Describes the breadth of mixed studies.
4. Describes the reasons for mixing methods in the study.
5. The definition of the method of conducting the study is included, where the categorization is "bottom-up" or "top-down".

American sociologist D. Morgan notes that mixed research offers three options for integrating the results of qualitative and quantitative methods.

First, the variant of "intersecting" (convergent) data allows you to compare the results of different methods that answer the same question (**Qual = Quant**). As a rule, the qualitative and quantitative parts are realized simultaneously.

Secondly, the option of "additional coverage" implies the achievement of different research tasks by different methods (**Qual + Quant**).

The third principle of "sequential contributions" allows you to build a chain of continuity between data: the results of one method will serve as data for "input" to the next method (**Qual - > Quant**).

Combining the three cognitive capabilities of qualitative methods is used in a research design model called the "sequential priorities model" (Morgan, 2013). Specialists in mixed methods usually identify the dominant component in the project (qualitative or quantitative), which allows you to answer most of the questions of interest. According to D. Morgan's ideas, there are three ways to use the strengths of qualitative data in a quantitative project (for example, in a survey).

First, they use the intelligence capabilities of qualitative methods when it is not clear what the subject of the study is and what the research program should look like.

Secondly, the ability to find out the "meanings" of the studied phenomena, which is the strength of qualitative methods.

Third, the ability to identify specific ways to design tools and formulate questions.

As D. Morgan writes in his work, there are two main approaches to enriching qualitative projects with quantitative methods. First, to test hypotheses and theories that have been developed as a result of a quality project. Secondly, in order to generalize the obtained qualitative results beyond the limits of the studied case. Both approaches require the operationalization of qualitative results into measurable variables and testable relationships.

The principle of mixed research is that researchers should create projects that effectively answer their research questions. The author obviously needed to obtain as wide a profile as possible so his design included a survey with demographic, Likert-scaled, and open-ended questions; however, he also needed opinions, which are subjective, and best dealt with by qualitative methods. Questions of frequency may best be explored by quantitative methods, and perception and opinion by qualitative. If the questions deal with both of these, then mixed methods are likely to be preferred.

Questions and tasks for independent work

1. Methodology of qualitative and quantitative research.
2. Quantitative content analysis, qualitative content analysis.
3. Planning of quantitative research from the point of view of qualitative methodology.
4. Qualitative methods for obtaining "raw" data.

Task Based on the literature review (see Creswell et al., 2003), make diagrams of *general typologies of mixed-methods designs.*

Lecture 9. Qualitative methods in solving practical problems

Lecture plan

1. General characteristics of qualitative research
2. Methods of analysis of qualitative methods in solving practical problems
3. Qualitative research as an approach and epistemology in solving practical problems

Basic concepts: qualitative methods, qualitative research, "soft" research strategies

1. General characteristics of qualitative research

Dissemination of qualitative research is one of the current trends in methodology. Various studies of individual cases, qualitative analysis of documents and audiovisual materials, "field", included ethnographic observations, unstructured in-depth interviews are gaining an increasing number of supporters in sociology, anthropology, psychology, political science, pedagogy and other sciences.

Qualitative research is a field of research characterized by a variety of theoretical foundations, ethical guidelines, methodological solutions, principles and procedures used. This is a certain approach to generating knowledge. Qualitative research includes, for example, phenomenological, narrative research, discourse analysis, and other "soft" research strategies. Based on these features, it is possible to give the most complete definition of qualitative research. This is a comprehensive study based on the field form of work, involving the collection of detailed descriptions of human experiences and meanings, data processing using special text analysis procedures, their interpretation taking into account the socio-cultural context, the lack of strict standardization, attention to individual cases, reliance on the point of view of the studied people and the reflection of the researcher himself. The modern field of qualitative research is an extremely branched tree of directions, strategies and methods. There is no single, generally accepted classification of them.

The category of "quality" includes: case study, biographical research, life history, phenomenological, hermeneutical method, ethnomethodology, ethnography, phenomenography, "grounded theory", focus group method, Delphi method (expert evaluation method), structural analysis of events, qualitative content analysis, discourse analysis, conversation analysis, narrative analysis, feminist, dialogical, cultural research, heuristic research, action research, clinical research, etc.

Qualitative methods are complex research methods that reveal in detail not statistical quantitative patterns, but the real content of motivational and emotional aspects of information perception. And these methods in a certain way allow:

1. to overcome the subject-object dichotomy of the natural science research paradigm, using both subjective and objective knowledge;
2. to achieve integrity in the description and understanding of the phenomenon under study;
3. to obtain information corresponding to the categorical apparatus used by the individual in the perception and evaluation of information;
4. to achieve an in-depth understanding of the phenomena under study, without focusing on mass data collection;
5. present the results and give practical recommendations in an accessible, understandable form to the customer;
6. to get a rich amount of data faster and cheaper than using traditional quantitative methods.

The researchers identify the following methods of Qualitative data analysis of the “Big 6”

Qualitative content analysis. Content analysis is a research tool used to determine the presence of certain words, topics or concepts in some given qualitative data (for example, in the text of an interview). Using content analysis, researchers can quantify and analyze the presence, meanings, and relationships of certain words, topics, or concepts.

Narrative analysis - is an analysis in which the researcher collects and studies the stories of individuals about various specific life situations. The goal is to understand how people experience certain events, structure them into sequential sequences and give them subjective meaning.

Discourse analysis is a blanket term for a range of qualitative research approaches used in analyzing the use of language in social contexts. Researchers employ these techniques to understand the world by investigating the underlying meaning of what people say and how they say it, whether in face-to-face conversation, documents, non-verbal interaction, or images.

Thematic analysis is a method of "identifying, analyzing, and reporting patterns (themes) within data".

Grounded theory (GT) is a method that enables you to study a particular phenomenon or process and discover new theories that are based on the collection and analysis of real world data.

Interpretive phenomenological analysis (IPA) is a qualitative approach which aims to provide detailed examinations of personal lived experience.

2. Methods of analysis of qualitative methods in solving practical problems

Qualitative research is research that primarily deals with qualitative data and qualitative methods of analysis. **Qualitative methods of analysis** are special forms of text analysis that focus not on the calculation and statistical generalization of certain text units, but on their content division, generalization, comparison and interpretation. These approaches are aimed at solving, in fact, one common problem: how to categorize data and establish links between categories. This task is the core of qualitative analysis. Among the most typical procedures of qualitative analysis are the selection of phrases, patterns, themes, metaphors, their clustering, counting, comparison, generalization, comparison with existing constructs and theories, etc.

Qualitative data is any information collected by a researcher that is expressed not in numbers, but in words: some content that is extracted by the researcher from observation protocols, interviews, documents and audiovisual materials and formulated, encoded and transmitted to them verbally. In this sense, qualitative research is research that primarily or exclusively uses words as data and means of analysis.

3. Qualitative research as an approach and epistemology in solving practical problems

Summarizing the classifications, definitions and descriptions of qualitative research available in the literature, we can distinguish: 1) preference for the field form of work; 2) appeal to the social context; 3) pursuit for richness and holism of description; 4) interest in individual cases; 5) inductive approach to data; 6) flexibility and lack of rigid standardization; 7) interpretation of the person under study as an expert; 8) increased attention to language; 9) orientation to the study of meanings and experiences; 10) reliance on the researcher's reflexivity.

Appeal to the social context. Engaging and in-depth analysis of social, cultural, historical, ideological, political contexts, as well as the communicative context of the researcher's interaction with the people under study is another important feature of qualitative research. The studied phenomena are analyzed in the broad context of their manifestation.

The pursuit of wealth and holism description. The concept of thick description, introduced by the famous ethnographer K. Geertz in the analysis of various cultures, has now become one of the central concepts of qualitative methodology. It means the desire to obtain the most rich, detailed, multi-sided, systematic, holistic descriptions of the phenomenon under study. Getting rich descriptions is also seen as another way to solve the problem of validity and reliability of qualitative research.

Interest in individual cases. The emphasis on the study of exceptional phenomena and specific cases is also the "calling card" of qualitative research. We can say that this is their essence: qualitative research highlights what can be overlooked in statistical generalization. They perform a heuristic function and serve for a deeper and broader understanding of the phenomenon under study, its boundaries and possible variations.

An inductive approach to data. The collection and analysis of empirical data takes place outside of any pre-formulated generalizations, theories, hypotheses. Qualitative research usually ends rather than begins with hypotheses and theories. That is why they are often presented as the best way to pre-study something, discover, explore new areas, develop theories and hypotheses about the phenomenon being studied.

Flexibility and lack of rigid standardization. The lack of strict standardization allows you to remain more receptive to reality and free in relation to those procedures and ways of working with data. In qualitative research, the researcher himself is recognized as the main "measuring device". In the field work, the researcher observes, makes notes in an unstructured form, which then comprehends and translates into the form of a narrative. The interpretation of the researcher as a research tool, the flexibility of solutions and the reliance on a non – rigid research design is another important feature of this methodology.

Interpretation of the person under study as an expert. Proponents of this methodology consider the people studied more as experts (their life and experiences, their culture and community, the specifics of using a certain discourse, etc.), rather than as naive and passive subjects or test subjects. This is also related to one of the procedures that is sometimes used in qualitative research in order to verify the data and make them more reliable: discussion with the study people of the results and conclusions obtained in order to make some corrections, clarifications, additions to them.

Increased attention to the language. There are several features of the interpretation of language in qualitative research. First, *natural language* is seen as the medium that most authentically represents the reality of human existence. Secondly, the description and analysis uses *everyday language*, which makes it possible to describe the phenomenon more sensitively, clearly, vitally and naturally than the academic language. Third, when analyzing and writing reports, researchers strive to preserve the original language of the subjects, their way of describing, expressing, and authentic forms of categorization.

Orientation to the study of meanings and experiences. Qualitative research seeks to clarify the position of the other, preferring direct dialogue, taking into account the contexts. And this is considered one of the most complete and reliable ways to study the meanings that people attach to the objects and events of their own lives, and the experiences associated with them.

Reliance on the researcher's reflexivity. The researcher's appeal to his own subjectivity and reflection of his observations, experiences, personal experience, language, knowledge, assumptions, research position, ethnic, cultural and gender identity is an integral component of any qualitative research today. Proponents of qualitative methodology believe that without involving the researcher's reflexivity in the report, it is impossible to provide a truly valid and environmentally friendly social study.

Questions and tasks for independent work

1. Describe qualitative research as a special type of research organization.
2. What is the scientific status of qualitative research, the problems of scientific criteria.
3. Briefly describe individual qualitative strategies (ethnographic, phenomenological and its variations, grounded theory, discourse analysis approach, narrative and biographical approach, psychoanalytic approach).
4. Comparison of qualitative and quantitative methodology. The possibilities of their combination.

Task

Case studies: qualitative research in the media, qualitative research on the Internet; qualitative research in advertising; qualitative marketing research; qualitative research in organizational diagnostics; qualitative research in the field of health.

Lecture 10. Analysis and presentation of the results of qualitative research

Lecture plan

1. Data analysis
2. Phenomenological Analysis

3. Hermeneutic Analysis

4. Contextual issues and qualitative research

Basic concepts: qualitative data, content analysis, narrative analysis, discourse analysis, grounded theory

1. Data analysis

Qualitative data refers to non-numeric information such as interview transcripts, notes, video and audio recordings, images and text documents. *Qualitative data analysis can be divided into the following five categories:*

1. Content analysis. This refers to the process of categorizing verbal or behavioural data to classify, summarize and tabulate the data.

2. Narrative analysis. This method involves the reformulation of stories presented by respondents taking into account context of each case and different experiences of each respondent. In other words, narrative analysis is the revision of primary qualitative data by researcher.

3. **Discourse analysis**. A method of analysis of naturally occurring talk and all types of written text.

4. Framework analysis. This is more advanced method that consists of several stages such as familiarization, identifying a thematic framework, coding, charting, mapping and interpretation.

5. Grounded theory. This method of qualitative data analysis starts with an analysis of a single case to formulate a theory. Then, additional cases are examined to see if they contribute to the theory.

Qualitative data analysis can be conducted through the following three steps:

Step 1: Developing and Applying Codes. Coding can be explained as categorization of data. A 'code' can be a word or a short phrase that represents a theme or an idea. All codes need to be assigned meaningful titles. A wide range of non-quantifiable elements such as events, behaviours, activities, meanings etc. can be coded.

There are three types of coding:

Open coding. The initial organization of raw data to try to make sense of it.

Axial coding. Interconnecting and linking the categories of codes.

Selective coding. Formulating the story through connecting the categories.

Step 2: Identifying themes, patterns and relationships. Unlike quantitative methods, in qualitative data analysis there are no universally applicable techniques that can be applied to generate findings. Analytical and critical thinking skills of researcher plays significant role in data analysis in qualitative studies. Therefore, no qualitative study can be repeated to generate the same results.

Nevertheless, there is a set of techniques that you can use to identify common themes, patterns and relationships within responses of sample group members in relation to codes that have been specified in the previous stage.

Specifically, the most popular and effective methods of qualitative data interpretation include the following:

- ***Word and phrase repetitions*** – scanning primary data for words and phrases most commonly used by respondents, as well as, words and phrases used with unusual emotions;
- ***Primary and secondary data comparisons*** – comparing the findings of interview/focus group/observation/any other qualitative data collection method with the findings of literature review and discussing differences between them;
- ***Search for missing information*** – discussions about which aspects of the issue was not mentioned by respondents, although you expected them to be mentioned;
- ***Metaphors and analogues*** – comparing primary research findings to phenomena from a different area and discussing similarities and differences.

Step 3: **Summarizing the data**. At this last stage, the research results are linked to the hypotheses or goals and objectives of the study. When writing data analysis, notable quotes from the transcript are used to highlight the main topics in the conclusions and possible contradictions.

It is important to note that the process of qualitative data analysis described above is general, and different types of qualitative research may require several different methods of data analysis.

2. Phenomenological Analysis

The meaning of the phenomenon is conceptualized in the interior of the individual's awareness. Phenomenology is an approach to explore people's everyday life experience. It is used when the study is about the life experiences of a concept or phenomenon experienced by one or more individuals. A phenomenological researcher investigates subjective phenomena [Creswell,2009]. For example, a thoughtful understanding of the meaningful aspects of having a conversation or a talk together or the experience of interacting online and the kind of contact or closeness we experience through email, texting, or social networks may be of value to professional practitioners as well as to anyone involved in the conversational relations of everyday living. This type of research is used to study areas in which there is little knowledge.

Therefore, phenomenological research is a design of inquiry coming from philosophy and psychology in which the researcher describes the lived experiences of individuals about a phenomenon as described by participants [Creswell, 2014].

Phenomenological analysis is based on discussions and reflections of direct sense perception and experiences of the researched phenomenon. A starting point of the strategy is your ability to approach a project without a priori assumptions, definitions or theoretical frameworks. A key aspect of this method of analysis is phenomenological reduction. You need to eliminate any factors, which disturb your perceptions or which you consider as outside factors. As phenomenological research is based on either your own or other people's experiences and perceptions of the phenomenon, so too does phenomenological analysis focus on either your own or other people's experiences and perceptions of the phenomenon. Phenomenological analysis is a broad and loose name for various types of analysis based on the phenomenological orientation of the Philosophy of science. These orientations lay emphasis on experiences, interpretations and bodily sensations. You can combine phenomenological analysis with other modes of analysis. A combination of the rules of phenomenological analysis and hermeneutics is known as phenomenological hermeneutic analysis.

3. Hermeneutic Analysis

Hermeneutic analysis is a name for various methods of analysis, which are based on interpreting. The strategy forms an opposite to those research strategies which stress objectivity and independence of interpretations in the formation of knowledge. Hermeneutic research includes various different approaches. Methods of analysis may also vary, and in different disciplines discipline-specific methods exist for interpreting phenomena. Hermeneutic analysis enables you to elicit an in-depth understanding of meanings of, for example: human practices, culture, works of art and texts. Understanding is produced through systematic interpretation processes.

These processes are known as a *hermeneutic circle.* Interpretation of details affects the interpretation of the entire phenomenon; reviews of these interpretations produce a deepening understanding of the phenomenon. You can combine hermeneutic analysis with other methods of analysis that aim to interpret and understand meanings. A combination of the rules of hermeneutics and phenomenology forms *phenomenological hermeneutic analysis*. Hermeneutic analysis also forms the basis of various discipline-specific methods of analysis and close readings methods.

4. Contextual issues and qualitative research

In research, contextualization is a way to approach a research project or link it to the relevant research and to the research setting. Contextualization gives confidence and support to the research project as a whole. The contextualization of research takes various forms. The two main ways to contextualize research will be illustrated and discussed below.

First, we contextualize our research in relation to the existing literature and prior research.

Second, we contextualize our research by linking it to

(a) the specific context in which it was conducted, such as the institution and workplace (*micro-level*); and

(b) the location or general setting of the research, such as the geographical area and location (*macro-level*).

Based on these two main ways of contextualizing the research, another, broader contextualization of the research project is identified, namely, linking the research with other disciplines (research areas) and other contexts (locations). In all cases, research should be contextualized and based on theory and / or practice.

Questions and tasks for independent work

1. What are the principles of data analysis in different qualitative research strategies? Describe the principles of analysis in the phenomenological and hermeneutic traditions.
2. What are the levels of data analysis of qualitative research? Briefly describe them: the level of the source material, the level of description, the level of interpretation.
3. Briefly describe the methods of data analysis in the following approaches: the grounded theory of A. Strauss, Glazer and J. Corbin, Q-methodology, narrative analysis, discursive analysis.
4. What are the ways to solve the problem of validity and reliability of the obtained results of qualitative research?
5. What is the problem of integrating qualitative methodology?

Task

Using the literature, get acquainted with classical and modern qualitative research in the social sciences and humanities: their capabilities, limitations and perspectives.

BASIC AND ADDITIONAL LITERATURE

1. Matematicheskie metody psihologicheskogo issledovaniya: analiz i interpretaciya dannyh : uchebnoe posobie dlya studentov vysshih uchebnyh zavedenij, obuchayushchihsya po napravleniyu i po special'nostyam psihologii / A.D. Nasledov. – SPb.: Rech'. – 2012. – 392 s.
2. Busygina N.P. Kachestvennye i kolichestvennye metody issledovanij v psihologii. M.: YUrajt, 2015.
3. Sidorenko E.V. Metody matematicheskoj obrabotki v psihologii. – SPb.: OOO «Rech'», 2000. – 350 s.
4. Mitina O.V. Matematicheskie metody v psihologii: Praktikum. – Moskva : Aspekt Press, 2008. – 238 s.
5. The Sage handbook of Qualitative research in psychology / eds. C. Willig, W. StaintonRogers. – L.: Sage publications, 2008. – 631 p.
6. Howitt D., Cramer. D. Introduction to Statistics in Psychology. N.-Y.: FT, 2010.
7. Krichevec A.N., D'yachkov A.G., SHikin E.V. Matematika dlya psihologov. – M.: Flinta, 2015.Publications in electronic resources: http://www.qualitative-research.net/index.php/fqs/index http://www.apa.org/pubs/journals/qua/

Educational publication

QUALITATIVE AND QUANTITATIVE RESEARCH METHODS IN PSYCHOLOGY

Course of lections

Compiled by:
KUKHTOVA Natalia Valentinovna
MOROZHANOVA Maria Mikhailovna

Technical editor *G.V. Razboyeva*

Computer design *A.V. Tabanyukhova*

Signed to print 29.09.2022. Format 60x84 $^{1}/_{16}$. Offset paper.
Conv. printed sheets 2,79. Pub. sheet 2,78. Circulation 9 copy. Order 166.

Publiser and polygraphic processing – Educational Establishment
“Vitebsk State University named after P.M. Masherov”.

State Registration Certificate as publisher, printer and distributor of editions
№ 1/255 d/d 31.03.2014.

Printed on a risograph of Educational Establishment
“Vitebsk State University named after P.M. Masherov”.
210038, Vitebsk, Moskovsky Prospekt, 33.