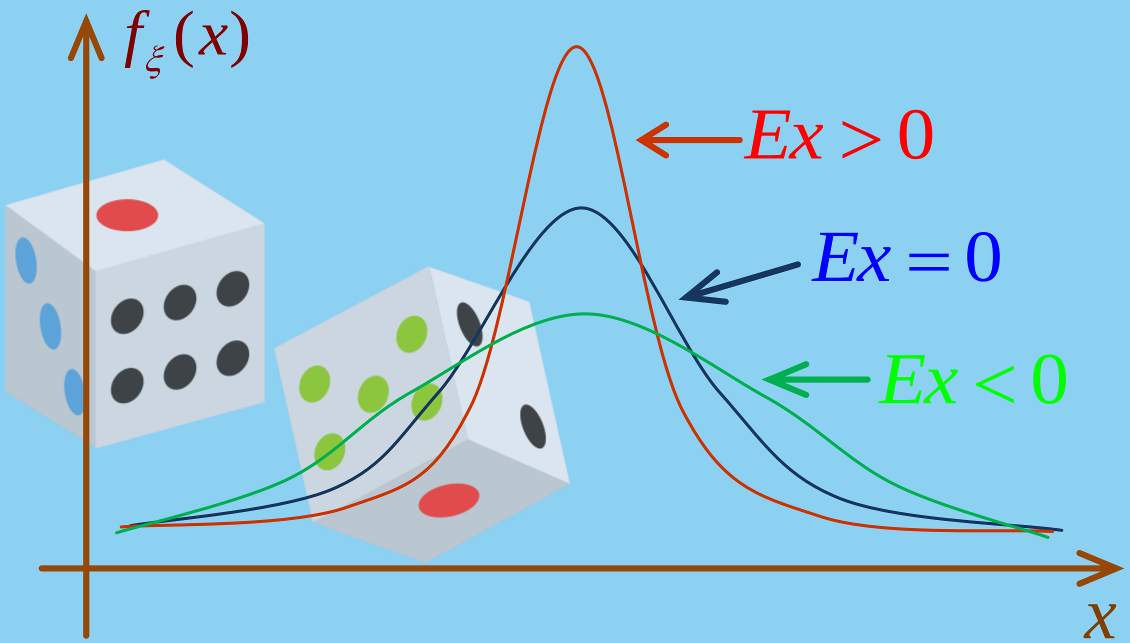
$(\Omega, \mathcal{F}, P)$

Vasyl Kushnirchuk

$P(A) = \frac{m}{n}$

PROBABILITY THEORY



Міністерство освіти і науки України
Чернівецький національний університет
імені Юрія Федьковича

Vasyl Kushnirchuk

PROBABILITY THEORY

Study guide

Василь Кушнірчук

ТЕОРІЯ ЙМОВІРНОСТЕЙ

Навчальний посібник

ЧНУ РУТА

Чернівці
Чернівецький національний університет
імені Юрія Федьковича
2024

УДК 519.21(075.8)
К 964

Рекомендовано Вченою радою
факультету математики та інформатики
Чернівецького національного університету імені Юрія Федьковича
(протокол № 12 від 25 червня 2024 р.)

Рецензенти :

Малик І.В., доктор фіз-мат.наук, доцент
кафедри математичних проблем управління
і кібернетики Чернівецького національного
університету імені Юрія Федьковича;

Лукашів Т.О., кандидат фіз-мат.наук, доцент
кафедри математичних проблем управління
і кібернетики Чернівецького національного
університету імені Юрія Федьковича.

Кушнірчук В.Й.
К 964 Теорія ймовірностей =Probability theory : навч. посібник / В.Й. Кушнірчук. – Чернівці : Чернівец. нац. ун-т ім. Ю. Федьковича, 2024. – 116 с.
ISBN 978-966-423-873-8

The textbook outlines the basic concepts of the course "Probability Theory", which the author teaches to students of the Faculty of Mathematics and Informatics and the Faculty of Economics. In addition to the theoretical material, the textbook contains practical problems for each chapter.

The textbook will be useful for students and anyone who wants to get acquainted with the basic concepts of probability theory.

УДК 519.21(075.8)

ISBN 978-966-423-873-8

INTRODUCTION

Probability theory is a science whose subject is the establishment of regularities in mass random phenomena (events).

The history of the development of probability theory is ancient. It began with research in gambling. They began to deal with it more seriously in the 16th and 17th centuries. Famous names include Jakob Bernoulli, Pierre-Simon Laplace, Carl Friedrich Gauss among the scientists who obtained important results in the theory of probabilities. Naturalists who achieved great success in the study of physical phenomena, in particular Augustin Louis Cauchy, were seriously engaged in the theory of probabilities.

Among Ukrainian mathematicians, Borys Hnedenko, Iosif Hykhman, Anatoliy Skorokhod, Mykhailo Yadrenko, Volodymyr Korolyuk made a significant contribution to the development of probability theory.

CHAPTER 1
RANDOM EVENTS

§1. A stochastic experiment. Random event. Frequency of random event

Let's imagine such an experiment: let the dice be rolled many times under the same conditions. In each toss, the number of points that will fall on the upper face of the cube is the result of the experiment. This result can be numbers 1, 2, 3, 4, 5, 6.

If such an experiment is carried out many times under the same conditions, then such an experiment is called stochastic. Sometimes they say probabilistic. The outcome of each individual toss is impossible to predict, but many outcomes can be described. We will describe the elementary results of the experiment, which are no longer broken down (interpreted) into smaller (more elementary) ones. With each elementary result, we will associate the events – "a one was rolled", "a two was rolled",..., "a six was rolled".

The set of these elementary events is called **the space of elementary events** and is denoted by Ω.

A condition is imposed on the space of elementary events – only one element of the space can appear as a result of the experiment. More complex events are formed from elementary events. For example, "an even number is on the upper face", "a number is greater than 3 on the upper face", etc. Each of the elementary events may happen or may not happen, but one of them will definitely take place (happen). We will call such **events random events**.

When tossing a coin, for example, elementary events "a coat of arms fell out" and "a figure fell out". We denote them, respectively, by C={"a coat of arms fell out"} and F={"a figure fell out"}, then the space Ω={C, F}.

One should not think that the space of elementary events is necessarily finite. Consider the following experiment: a coin is tossed until a coat of arms falls on the upper face. The space of elementary events in this experiment Ω = {*C*, *FC*, *FFC*, *FFFC*, *FFFFC*,...}. The number of elementary events is counted here.

And when measuring length, the space of elementary events is a segment, assuming that any real number can be the result of the measurement. There is a continuous number of elementary events.

If the experiment is carried out en masse, say *n* several times, and the number of experiments in which some recorded event occurs is noted *m*, then the ratio $\frac{m}{n}$ is called the frequency of the event. If this event is denoted by *A*, and $m(A)$ is the number of experiments in which it occurred, then the frequency of the event in the stochastic experiment is calculated by the formula $\nu_n(A) = \frac{m(A)}{n}$. It turns out that the frequency of the event has the property of stability, which consists in the fact that there exists $\lim_{n\to\infty} \nu_n(A)$, denote it by $p(A)$. This means, for example, when tossing a coin, that the frequency of occurrence of *F* at very large ones differs little from $\frac{1}{2}$. Due to the stability of the frequency of the event, the idea arose whether the frequency could not be used as a characteristic of the event. You can, but for this you need to carry out a lot of complex and sometimes very expensive, destructive, etc. experiments. Therefore, it is necessary to look for other approaches to characterize random events.

Let the space of elementary events Ω be given.

Definition. A random event *in a stochastic experiment is any subset of the space of elementary events.*

Elementary events that are elements of this subset are called ***elementary events*** that cause a random event.

In the future, random events are denoted by *A*, *B*, *C*, etc. – these are subsets of the set Ω.

The event itself Ω (it is a subset of Ω) is a ***reliabl*e** (probable) ***event***. It is called so because it always happens in every experiment.

A subset is also the empty set $\varnothing$. We will call such an event an ***impossible event***, it is not caused by any of the elementary events.

Example. Let us consider a stochastic experiment with the tossing of a die. Here Ω={1, 2, 3, 4, 5, 6}. Let's consider the event *A* – a number that is divisible by 3 without a remainder has fallen on the upper face of the cube. Then *A* ={3, 6}. The elementary events {3}, {6} cause the event *A*. Consider the event *B* – an even number is drawn. *B* ={2, 4, 6} is a subset of Ω.

Events *A* and *B* have common elements (element 6). This means that they can happen simultaneously.

Definition. *Events that can occur simultaneously are called **joint events**.*

Let's describe another event *C* – a number of no more than two points fell on the upper edge: $C = \{1, 2\}$. Events *A* and *C* cannot occur simultaneously (they do not have the same elementary events).

Definition. *Events that cannot occur simultaneously are called **incompatible events.***

In the experiment with tossing a coin before the first appearance of the coat of arms, the space of elementary events

$$\Omega = \{C, FC, FFC, FFFC, FFFFC, ...\}.$$

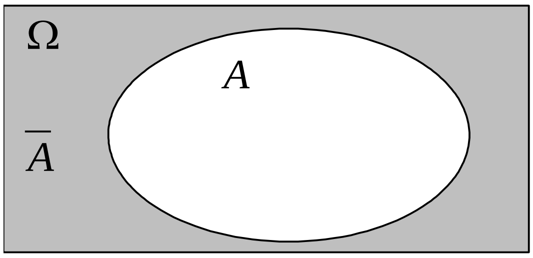
Let's consider the event *A* – the coat of arms appeared after an even number of tosses. Then

$$A = \{FC, FFFC, FFFFFC, ...\}.$$

Let the given event be *A*.

Definition. *The **opposite** of an event A is an event $\overline{A}$ that occurs when and only when the event A does not occur.*

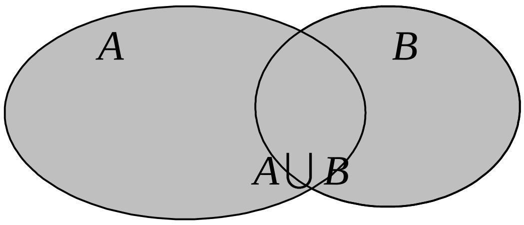
It is easy to understand that when an event A corresponds to some set, then $\overline{A}$ the addition of this set to the set corresponds to the event Ω. It is convenient to illustrate this with the help of Venn[*] diagrams:



Let events *A* and *B*.

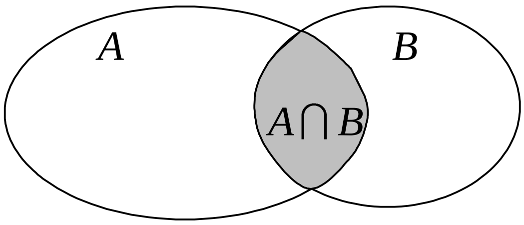
Definition. The sum of events *A and B is an event C that occurs when and only when at least one of the events A or B occurs.*

An event *A* is caused by a set of events (a set), and another set corre*sponds to the event B. The event C corresponds to the union of the sets A and* B. Therefore, it is natural to denote the sum of events (use to denote an event *C*) as follows: $C = A \bigcup B$.



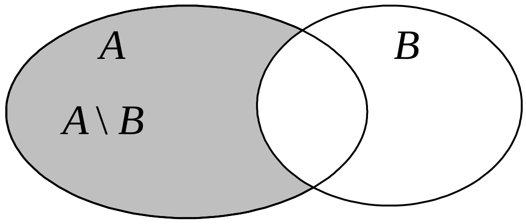
Definition. The product of events *A and B is an event C that occurs when and only when both event A and event B occur.*

Elementary events must occur, which cause the appearance of both events *A* and *B*, that is, the intersection of events, so it is natural to record $C = A \bigcap B$.



[*] *John Venn (1834-1923) is an English logician and philosopher*.

Definition. The difference between events *A and B is called the event C, which consists in the fact that the event A occurs and the event B does not occur. It is accepted to record* $C = A \setminus B = A \cap \overline{B}$.



Now we can write the following obvious relations:
$\overline{\Omega} = \varnothing$, $\overline{\varnothing} = \Omega$, $A \cup A = A$, $A \cap A = A$, $A \cup \Omega = \Omega$, $A \cap \Omega = A$, and others.

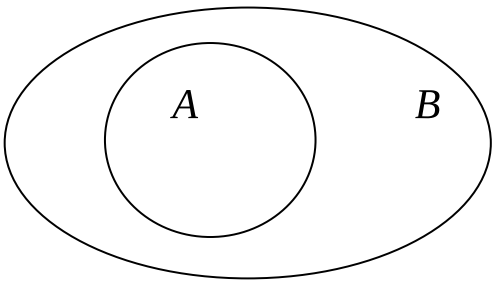
Since *A, B, C* are sets of elementary events, all the laws of set algebra apply to events, in particular:

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

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

$$\overline{A \cap B} = \overline{A} \cup \overline{B}, \ \overline{A \cup B} = \overline{A} \cap \overline{B}.$$

Definition. *An event A is said to cause an event B, if the execution of the event A leads to the execution of the event B.*



It is written like this: $A \subset B$. If an event *A* causes an event *B* ($A \subset B$) and an event *B* causes an event *A* ($B \subset A$), then the events *A* and *B* are considered the same (equal), i.e. $A = B$.

Obviously, what *A* causes Ω ($A \subset \Omega$) as well as

$$A \subset A \cup B, \ A \cap B \subset B, \ A \cap B \subset A.$$

§2. Classical definition of probability

Let the space of elementary events be given Ω. Suppose that it consists of no more than a counted number of elementary events (that is, these elementary events can be renumbered and there can be a finite number of them): $\Omega = \{\omega_1, \omega_2, ..., \omega_n, ...\}$. For some reasons, let each ω_i be matched with a real number p_i in such a way that

1) $p_i \geq 0$,

2) $\sum_{i=1}^{\infty} p_i = 1$.

Definition. *Let A some event, i.e. $A \subset \Omega$, then the* ***probability of the event*** *A is called the number $P(A)$, which by definition is equal to*

$$P(A) = \sum_{i:\omega_i \in A} p_i .$$

This definition can be called the **classical definition of probability**. Here they do not talk about how ω_i the number p_i is matched. If you choose these numbers differently, you can get different results.

Example 1. We will consider tossing a fair coin. With such a one-time experiment, a number or coat of arms may fall out on the upper face of the coin. In this case Ω={C, F} (that is $\omega_1 = C$, $\omega_2 = F$) is a finite set. For each of the elements ω_1 and ω_2, it is necessary to match the numbers with properties 1) and 2). If the sides of the coin are equal, then it is natural to choose these numbers as $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{2}$.

The following events can be considered in this experiment: $\varnothing$, {C}, {F}, {C, F}=Ω. It is obvious that $P(\varnothing) = 0$, because $\omega_1 = C$ and $\omega_2 = F$, are not included in $\varnothing$ and, $\sum_{i:\omega_i \in \varnothing} p_i = 0$. $P(\{C\}) = \frac{1}{2}$, $P(\{F\}) = \frac{1}{2}$, $P(\Omega) = 1$.

Example 2. Let's consider an experiment that consists in throwing a correct dice. The space of elementary events consists of six elementary events: 1 – "the number 1 fell on the upper face of the cube", etc. Ago $\Omega = \{1;\ 2;\ 3;\ 4;\ 5;\ 6\}$. Since all faces are equal, the cube is correct, we assign the same numbers with properties 1) and 2) to each elementary event. Ago $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = \frac{1}{6}$.

Different events can be considered in this experiment. There can be 2^6 a total of events. Consider the following event: *A* – a number appeared on the upper edge, which when divided by 3 gives a remainder of 1. Then $A = \{1;\ 4\}$, and the probability of this event according to the definition

$$P(A) = p_1 + p_4 = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

Example 3. Consider a "falsified" cube. In order for 6 to fall out more often, the center of gravity must be shifted down so that 1 is at the bottom and 6 is at the top. Let's assume that the center of gravity is shifted so that the numbers that are matched are: $\omega_6 \leftrightarrow p_6 = \frac{7}{36}$, $\omega_1 \leftrightarrow p_1 = \frac{5}{36}$ and the following numbers are matched to other elementary events. In such an experiment for the event A, from example 2, $P(A) = \frac{11}{36}$.

We will not consider how to choose these numbers p_i, you can conduct an experiment many times, analyze the frequencies and then choose the numbers p_i in a certain way.

Example 4. A coin is tossed until it lands with the coat of arms up. Here

$$\Omega = \{C, FC, FFC, FFFC, FFFFC, ..., FFF...FC, \ldots\}.$$

We are dealing with Ω, which contains a countable number of points (elementary consequences of ω_i). We match the numbers

$$\omega_1 \to p_1 = \frac{1}{2},\ \omega_2 \to p_2 = \left(\frac{1}{2}\right)^2,\ \omega_3 \to p_3 = \left(\frac{1}{2}\right)^3,\ \omega_4 \to p_4 = \left(\frac{1}{2}\right)^4,\ldots,$$

$$\omega_{n+1} \to p_{n+1} = \left(\frac{1}{2}\right)^{n+1},\ldots$$

to the elementary consequences. Let's check conditions 1) and 2). Condition 1) is fulfilled, because all $p_i \geq 0$. Condition 2) is also fulfilled:

$$\sum_i p_i = \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \ldots = \frac{\frac{1}{2}}{1-\frac{1}{2}} = 1.$$

Consider the event *A*, which consists in the fact that the number of rolls is divided by three without a remainder. Then it looks like

$$\Omega = \{FFC, FFFFFC, FFFFFFFFC, \ldots\}.$$

By definition, the probability of this event

$$P(A) = \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^6 + \left(\frac{1}{2}\right)^9 + \ldots = \frac{\left(\frac{1}{2}\right)^3}{1-\left(\frac{1}{2}\right)^3} = \frac{1\cdot 8}{8\cdot 7} = \frac{1}{7}.$$

In general, if the space of elementary events Ω is specified and the probability *P* (i.e., numbers $p_1,\ p_2,\ldots$) is introduced in some way, then it is said that a **probabilistic model** (Ω, P) is specified.

A partial case of the above is the case of finite Ω (finite number of elements in this set) and equally possible elementary results of the experiment. In this case, all p_i are the same, i.e. $p_i = p$. Therefore, the probability of any event *A* is: $P(A) = \sum_{i:\omega_i \in A} p_i = \sum_{i:\omega_i \in A} p = p \cdot m$, where m is the number of elementary events that cause the event *A* (which cause the event *A*), or the number of favorable occurrences of elementary events. But the sum over all i is $\sum_{i:\omega_i \in \Omega} p_i = p \sum_{i:\omega_i \in A} 1 = p \cdot n = 1$, where n is the number of all elementary events. From here we have $p = \frac{1}{n}$. We substitute in $P(A)$ and get

$P(A) = p \cdot m = \frac{m}{n}$, i.e

$$P(A) = \frac{m}{n}$$

– **the classical definition of probability**: the probability of an event A is equal to the ratio of the number of favorable elementary events to the total number of elementary events.

Example 5. From a carefully shuffled deck of cards (24 sheets), 3 cards are taken out at random. What is the probability that one queen and two aces are drawn?

◄ Thoroughly shuffled – means an equal chance of drawing a random card. Random means an equal opportunity to draw an arbitrary three cards. In this example, the total number of possible cases is the number of ways you can choose 3 cards out of 24:

$$n = C_{24}^{3} = \frac{24!}{3! \cdot 21!} = \frac{22 \cdot 23 \cdot 24}{1 \cdot 2 \cdot 3} = 4 \cdot 23 \cdot 22 .$$

The number of ways to choose one queen is 4, and the number of ways to choose two aces from four aces is C_4^2. Therefore, the number of favorable cases for the occurrence of the event

$$m = 4 \cdot C_4^2 = 4 \cdot \frac{4!}{2! \cdot 2!} = 4 \cdot \frac{3 \cdot 4}{2} = 24 .$$

Then

$$P(A) = \frac{m}{n} = \frac{24}{4 \cdot 22 \cdot 23} = \frac{3}{11 \cdot 23} = \frac{3}{253} . \blacktriangleright$$

§3. Geometric probability

If you shoot at a target, the question arises about the distance from the point of impact to the target. This distance can be an arbitrary number from some segment. Then the space of elementary events is continuous and it is impossible to find probabilities of events according to the classical definition.

Often, the space of elementary events is some part of a plane, or a straight line, or 3-dimensional space, and the events under consideration have such a property that they can be assigned a certain "measure".

Then the **probability of event** A is called measure A $(A \subset \Omega)$ divided by measure Ω, i.e. $P(A) = \dfrac{mes(A)}{mes(\Omega)}$. This is how **geometric probability** is determined.

Example (meeting problem).

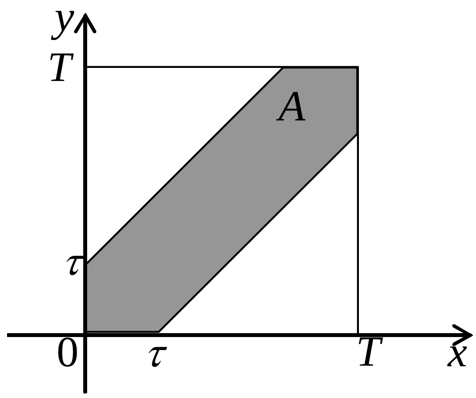
The two agreed to meet at some fixed time. The arrangement is as follows: the meeting can take place during time T (for example, one hour). The first to arrive waits for the second during time $\tau < T$. What is the probability that the meeting will take place if the arrival times of the meeting participants are equal (the arrival of the participants during time T is equally possible).

Let us denote: x – the moment of arrival of the first participant of the meeting (from the beginning of the stipulated time), y – the moment of arrival of the second participant of the meeting. It is clear that $0 \le x \le T$, $0 \le y \le T$. Then the elementary result of this experiment will be a pair of numbers (x, y). What is the space of elementary events? The space of elementary events consists of pairs of numbers (x, y) such that $0 \le x \le T$ and $0 \le y \le T$, i.e.

$$\Omega = \{(x, y):\ 0 \le x \le T,\ 0 \le y \le T\}.$$

Geometrically (on the plane) Ω are the points of the corresponding square. Each point determines the moment of arrival of meeting participants. Let's describe event *A* – the meeting will take place:

$$\Omega \supset A = \{(x, y) : 0 \le x \le T, 0 \le y \le T, \ |x - y| \le \tau\}.$$



Let's represent the set *A*, that is, the set of points of the square Ω that satisfy the inequality $|x - y| \le \tau$, which is equivalent to the following inequalities

$$-\tau \le x - y \le \tau \quad \Rightarrow \quad \begin{cases} x - y \le \tau, \\ x - y \ge -\tau, \end{cases} \quad \Rightarrow \quad \begin{cases} x - y \le \tau, \\ y - x \le \tau. \end{cases}$$

We will use the area as a measure on the plane. Therefore

$$mes\,\Omega = T^2, \ \ mes\,A = T^2 - (T - \tau)^2.$$

Then, by definition, the probability that the meeting will take place is equal to

$$P(A) = \frac{T^2 - (T - \tau)^2}{T^2} = 1 - \left(1 - \frac{\tau}{T}\right)^2.$$

§4. Axioms of probability theory

At the end of the 19th century, scientists approached the need to formulate a system of axioms of probability theory. Academician A.M.Kolmogorov* formulated the axioms in 1933.

Consider a stochastic experiment. We denote the corresponding space of elementary events by Ω. Let $\mathcal{F}$ be a class of events that satisfies the following conditions:

A$_1$) Ω as the event belongs to the class $\mathcal{F}$, i.e $\Omega \in \mathcal{F}$,

A$_2$) if A is an event from class $\mathcal{F}$, then $\overline{A}$ is an event from class $\mathcal{F}$,

A$_3$) if A_i $(i = 1, 2, \ldots)$ are events from the class $\mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Sets from the class $\mathcal{F}$ are called **random events**.

Let such a function of events $P(\cdot)$ (a mapping of the class $\mathcal{F}$) be given that satisfies the following conditions:

B$_1$) $P(\Omega) = 1$,

B$_2$) $P(A) \geq 0$ for any event $A \in \mathcal{F}$,

B$_3$) if A_i $(i = 1, 2, \ldots)$ are elements of class $\mathcal{F}$, such that $A_k \cap A_l = \varnothing$ $(k \neq l)$, then $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

If random event $A \in \mathcal{F}$, then $P(A)$ is called the **probability of random event** *A*.

---

* *Andrii Mykolayovych Kolmogorov (1903-1987) is a Soviet mathematician.*

The first three axioms ($A_1 - A_3$) actually determine the class of observed events (events that are allowed to be considered), the last three axioms ($B_1 - B_3$) determine the probability of these events.

Probability theory does not deal with how to construct a measure of events. It is assumed everywhere that $(\Omega, \mathcal{F}, P)$ is given. Then the **probability space** is said to be given.

This system of axioms is not contradictory, because there are objects, discussed above, on which they are fulfilled.

§5. Properties of probability

1⁰. *The probability of the opposite event is equal* $P(\overline{A}) = 1 - P(A)$.

◄ If A belongs to $\mathcal{F}$, then according to axiom A_2 belongs to $\mathcal{F}$. But $A \cup \overline{A} = \Omega$, $A \cap \overline{A} = \varnothing$, therefore according to axiom B_3

$$P(A \cup \overline{A}) = P(A) + P(\overline{A}) = P(\Omega) = 1.$$

From here we get the required equality. ►

Example 1. A dice is thrown. What is the probability that a unit is not dropped?

Let the event A be that a one was rolled when the dice was tossed. Then the event $\overline{A}$ is opposite to it – when the dice were thrown, no one came out. Therefore, according to property **1⁰**, $P(\overline{A}) = 1 - P(A) = 1 - \frac{1}{6} = \frac{5}{6}$.

Example 2. Find the probability that among five randomly selected numbers there is at least one of the three given numbers?

Let the event A be that there is at least one of the three given numbers among the five selected numbers. Let's find the probability of the opposite event $\overline{A}$ – among the five selected numbers, there is none of the three given numbers. Since five numbers can be chosen in different 10^5 ways, and five numbers that do not contain any of the three given numbers can be chosen in 7^5 different ways, then $P(\overline{A}) = \left(\frac{7}{10}\right)^5$. Therefore, from property $\mathbf{1^0}$,

$$P(A) = 1 - P(\overline{A}) = 1 - (0,7)^5.$$

$\mathbf{2^0}$. *The probability of an impossible event is zero* $P(\varnothing) = 0$.

◄ If $A = \Omega$, then $\overline{A} = \varnothing$, and from property $\mathbf{1^0}$, at $A = \Omega$, we obtain $P(\varnothing) = 1 - 1 = 0$. ►

$\mathbf{3^0}$. *Let events A and B be from the class* $\mathcal{F}$ *(*$A \in \mathcal{F}$, $B \in \mathcal{F}$*), in addition, event A causes event B (*$A \subset B$*), then the probability of the difference of events is equal to the difference of their probabilities*:

$$P(B \setminus A) = P(B) - P(A).$$

◄ It is clear that $B = A \bigcup (B \setminus A)$, it is important here that $A \bigcap (B \setminus A)$ is an impossible event, because $A \bigcap (B \setminus A) = A \bigcap (B \bigcap \overline{A}) = \varnothing$. Then, according to the axiom B_3, we have $P(B) = P(A) + P(B \setminus A)$. From here we get the required equality. ►

$\mathbf{4^0}$. (consequence of $\mathbf{3^0}$) *If* $A \in \mathcal{F}$, $B \in \mathcal{F}$*, and event A causes event B (*$A \subset B$*), then the probability of event A does not exceed the probability of event B:* $P(A) \le P(B)$.

◄ The proof follows from property $\mathbf{3^0}$, because $P(B \setminus A) \ge 0$ and then $P(B) - P(A) \ge 0$. ►

5⁰. *For any event A from the class $\mathcal{F}$, the probability $P(A)$ does not exceed 1*: $P(A) \le 1$.

◄ The proof follows from property **4⁰** for the event $B = \Omega$. ►

If we also take into account axiom B_2, then we get

$$0 \le P(A) \le 1.$$

6⁰. (Theorem of addition of probabilities). *Let $A \in \mathcal{F}$, $B \in \mathcal{F}$ – be random events. Then the probability of the sum of events $A \bigcup B$ is equal*

$$P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B).$$

◄ Let's write the sum of events $A \bigcup B$ as follows

$$A \bigcup B = (A \setminus (A \bigcap B)) \bigcup (B \setminus (A \bigcap B)) \bigcup (A \bigcap B).$$

In the right part, the terms are not compatible, that is, their sum is an impossible event (equal to $\varnothing$). Therefore, the probability of the sum of events, according to axiom B_3, is equal to t he sum of the probabilities of the terms. In addition, $A \bigcap B \subset A$ and $A \bigcap B \subset B$, so the probability of difference (in the first and second terms) is equal to the difference in probabilities. That is,

$$P(A \bigcup B) = P(A) - P(A \bigcap B) + P(B) - P(A \bigcap B) + P(A \bigcap B). \quad ►$$

Example 3. A tetrahedron whose faces are painted is tossed up (one face is yellow, the second is blue, the third is red, and all three colors are applied to the fourth) and the event is tracked in which color the face of the tetrahedron is painted, with which it will fall on the table. What is the probability that this face is yellow or blue?

◄ Consider the following events:

Y - on the face that fell on the table there is a yellow color;

B – the face that fell on the table is blue;

R – the face that fell on the table has a red color.

It is necessary to find $P(Y \bigcup B)$.

The probability of a tetrahedron falling on a certain face (specific) on the table, in the classical sense, is equal to 0,25. Let's count

$$P(Y) = \frac{2}{4} = \frac{1}{2} = P(B),\ P(Y \bigcap B) = \frac{1}{4}.$$

Then by property $\mathbf{6^0}$ we get

$$P(Ж \bigcup C) = P(Ж) + P(C) - P(Ж \bigcap C) = \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4}. \quad ►$$

$\mathbf{6^0}_{\mathbf{1.}}$ (Generalization $\mathbf{6^0}$). It is easy to get such a result

$$P(A \bigcup B \bigcup C) = P(A) + P(B) + P(C) -$$

$$-P(A \bigcap B) - P(A \bigcap C) - P(B \bigcap C) + P(A \bigcap B \bigcap C).$$

Remark 1. The theorem of adding probabilities in the case when events are incompatible is formulated as follows: the probability of the sum of events is equal to the sum of their probabilities:

$$P(A \bigcup B) = P(A) + P(B) \quad \text{if} \quad A \bigcap B = \varnothing.$$

Remark 2. If the probability of an event is zero, it does not mean that the event is impossible.

For example, we select a number on the segment [0, 1]. Consider the event: "the selected number is rational". This event is possible, but the measure of the corresponding set is zero and therefore the probability of the event is zero.

But if the event is impossible, then its probability is zero.

Therefore, the equality $P(A \bigcup B) = P(A) + P(B)$ is fulfilled not only in the case when $A \bigcap B = \varnothing$, but also when the probability of a possible event is zero (this follows from the previous remark 2).

§6. Conditional probabilities

Let the probability space $(\Omega, \mathcal{F}, P)$ be given. Suppose that some event $B \in \mathcal{F}$ has occurred. If some other event $A \in \mathcal{F}$ is given, then the question naturally arises: does the probability of event *A* depend on whether it is known whether event *B* has taken place, or whether nothing is known about event *B*?

Example. Two dice are rolled. The space of elementary events Ω in this experiment consists of 36 elementary events. $\mathcal{F}$ is the set of all subsets of the set Ω . Let the event *B*={ at least one of the two dice has one point}, i.e

$$B = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (3,1), (4,1), (5,1), (6,1)\}.$$

Let the event *A*={sum of points on two dice does not exceed four}. That is $A = \{(1,1),\ (1,2),\ (1,3),\ (2,1),\ (2,2),\ (3,1)\}$.

Let's calculate the probability of event *A*: $P(A)$=6/36=1/6. Suppose that information has been received that event *B* has occurred. Let's find the probability $P_{B \text{ took place}}(A)$ of event *A*, given that event *B* has occurred.

There are 5 beneficial consequences in the set *A*. But now the space of elementary events is not Ω, but the set *B*, because it is known that the event *B* occurred. And the total number of elementary events is 11. Therefore $P_{B \text{ took place}}(A) = 5/11$

Let's find $P(A \cap B)$. There are 5 common elements in *A* and *B*, and the entire space Ω contains 36 elementary consequences, that is, according to the classical definition of probability, $P(A \cap B) = 5/36$. Note that $P(B) = 11/36$. After analyzing the last three probabilities, we come to the conclusion that in this example

$$P_{B \text{ took place}}(A) = \frac{P(A \cap B)}{P(B)} = \frac{5/36}{11/36} = 5/11$$

Let's summarize this example.

Definition 1. *Let B be an event such that* $P(B) > 0$. *Then the* ***conditional probability*** *of event A, provided that event B has occurred (denoted by* $P_B(A)$ *or* $P(A|B)$*), is called the ratio of the probability of the product of events A and B to the probability of event B:*

$$P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Properties of conditional probability

1⁰. $P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$. $P(B|B) = 1$. $P(\varnothing|B) = 0$.

2⁰. $P(A \cap B) = P(B) \cdot P(A|B)$ – this is equivalent to the definition of conditional probability.

3⁰. If $P(A) > 0$ and $P(B) > 0$, then

$$P(A \cap B) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$$

– is **the theorem of multiplication of probabilities**: the probability of the product of events is equal to the product of their probabilities, one of the cofactors is a conditional probability.

Consider a generalization of the multiplication theorem. Let it be found

$$P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|A \cap B).$$

This is a formal notation, because it is necessary that all probabilities on the right side of the equality exist. So that it is not formal, taking into account that $A \cap B \subset A$ we get $P(A \cap B) \le P(A)$.

Therefore, it is sufficient to require that $P(A \cap B) > 0$ and then all conditional probabilities in the theorem of multiplication of probabilities exist.

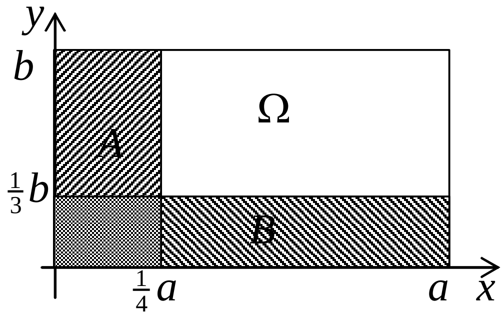
Definition 2. *Events A and B are called* ***independent*** *if*

$$P(A \cap B) = P(A) \cdot P(B).$$

It is easy to understand that events *A* and *B* are independent if and only if $P(A|B) = P(A)$ (if it is known that $P(B) > 0$) or $P(B|A) = P(B)$ (in the case when $P(A) > 0$).

The concepts of independence of events and incompatibility of events do not coincide. Events can be independent but compatible, or incompatible but dependent.

Example. Consider a rectangle, which is the space of elementary events. Two parts of this rectangle *A* and *B* have a common area as shown in the figure.



It is clear that $S_\Omega = a \cdot b$, $S_A = \frac{1}{4} a \cdot b$, $S_B = \frac{1}{3} a \cdot b$, $S_{A \cap B} = \frac{1}{12} a \cdot b$. Then $P(A \cap B) = \frac{1}{12}$, $P(A) = \frac{1}{4}$, $P(B) = \frac{1}{3}$, $P(A \cap B) = P(A) \cdot P(B)$. Therefore, the given events *A* and *B* are compatible but independent.

We will formulate and prove the main properties of independent events.

Statement 1. *If events A and B are independent, then events* $\overline{A}$ *and B are also independent, events A and* $\overline{B}$ *are independent, and events* $\overline{A}$ *and* $\overline{B}$ *are independent.*

Statement 2. *If events A and* B_1 *are independent, events A and* B_2 *are independent, events* B_1 *and* B_2 *are incompatible (i.e.* $B_1 \cap B_2 = \varnothing$*), then events A and* $B_1 \cup B_2$ *are independent.*

The condition of incompatibility of events B_1 and B_2 is essential and is not only of a technical nature for proof.

Definition 3. *Events* $A_1,\ A_2,...,\ A_n$ *are called* ***pairwise independent*** *if for any i, j such that* $1 \le i \le n,\ \ 1 \le j \le n,\ \ i \ne j$ *the probability of the product of events* A_i *and* A_j *is equal to the product of their probabilities*

$$P(A_i \cap A_j) = P(A_i) \cdot P(A_j).$$

Definition 4. *Events* $A_1,\ A_2,...,\ A_n$ *are called* ***independent in the aggregate*** *if for any set of indices* $i_1,...,i_k$, *such that* $1 \le i_1 < ... < i_k \le n$, $2 \le k \le n$, *the equality holds:*

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j}).$$

It is clear that from independence in the aggregate follows pairwise independence. The question arises whether these concepts are not equivalent? That is, does pairwise independence result in aggregate independence? The answer to this question is negative. Independence in the aggregate does not follow from pairwise independence (this is evidenced by the example of S.N.Bernshtein[*]).

---

[*] *Serhii Natanovich Bernshtein (1880-1968) was a Soviet mathematician.*

The example of Bernstein. Consider such a stochastic experiment. A tetrahedron whose faces are painted is tossed up (one face is yellow, the second is blue, the third is red, and the fourth is painted with all three colors) and the event is tracked in which color the face of the tetrahedron is painted, with which it will fall on the table. The following events are considered:

Y - on the face that fell on the table there is a yellow color;

B – the face that fell on the table is blue;

R – the face that fell on the table has a red color.

We will show that these events are pairwise independent, but not independent as a whole.

We consider all faces to be equal, and therefore the probability of a tetrahedron falling on a table (specifically) is 0.25, in the classical sense.

Let's calculate

$$P(Y) = \frac{2}{4} = \frac{1}{2} = P(B) = P(R).$$

Let's check the pairwise dependence of events:

$$P(Y \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(Y) \cdot P(B).$$

It follows that events Y and B are independent. Similarly,

$$P(Y \cap R) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(Y) \cdot P(R), \; P(B \cap R) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(B) \cdot P(R).$$

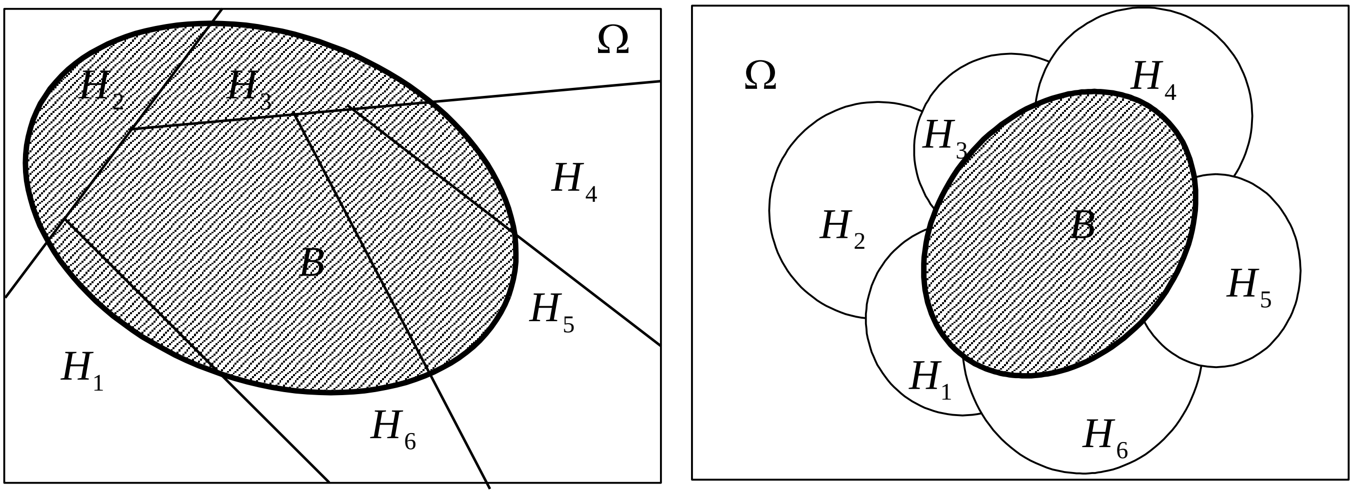
That is, the events Y and R and B, R are independent, and therefore the events Y, B, R are pairwise independent. Are they independent in aggregate? Because

$$P(Y \cap B \cap R) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(Y) \cdot P(B) \cdot P(R),$$

then the events Y, B, R are not independent as a whole.

§7. Formulas of full probability, Bayes

Let the probability space $(\Omega, \mathcal{F}, P)$ be given. A stochastic experiment is conducted in which some event *B* with $\mathcal{F}$ is observed. Let also be given some events $H_1, H_2, ..., H_n$, $H_i \in \mathcal{F}$, which are incompatible, that is, $H_i \bigcap H_j = \varnothing$, $i \neq j$. If *B* causes the sum of events $\bigcup_{i=1}^{n} H_i$, i.e. $B \subset \bigcup_{i=1}^{n} H_i$, then it is said that events $H_1, H_2, ..., H_n$ are **hypotheses** about event *B* (make up a system of hypotheses about event *B*). If $\bigcup_{i=1}^{n} H_i = \Omega$, then the system of hypotheses is said to be complete.



In the left figure, the system of hypotheses is complete.

Suppose that a set of hypotheses is given regarding the event *B* and the probability $P(H_i) > 0$, $i = \overline{1, n}$. Event *B* can be written as follows:

$$B = B \bigcap \left(\bigcup_{i=1}^{n} H_i \right) = \bigcup_{i=1}^{n} (B \bigcap H_i)$$

(the distributional property of probability is used here).

Since H_i are disjoint (incompatible events), then $(B \bigcap H_i)$ and $(B \bigcap H_j)$ are also incompatible, i.e.: $(B \bigcap H_i) \bigcap (B \bigcap H_j) = \varnothing$, $(i \neq j)$.

Therefore, the probability of the sum of events is equal to the sum of the probabilities of these events:

$$P(B) = P\left(\bigcup_{i=1}^{n}(B \cap H_i)\right) = \sum_{i=1}^{n} P(B \cap H_i), \qquad ...(1)$$

but

$$P(B \cap H_i) = P(H_i) \cdot P(B \mid H_i). \qquad ...(2)$$

Substituting equalities (2) into (1), we obtain:

$$P(B) = \sum_{i=1}^{n} P(H_i) \cdot P(B \mid H_i)$$

– is the **formula of total probability**.

Let us now additionally assume that $P(B) > 0$. Then, generally speaking, it would be possible to write:

$$P(B \cap H_i) = P(B) \cdot P(H_i \mid B). \qquad ...(3).$$

So, equating (2) and (3), we can write the following:

$$P(H_i) \cdot P(B \mid H_i) = P(B) \cdot P(H_i \mid B), \quad i = \overline{1,n}.$$

From here we determine $P(H_i \mid B)$:

$$P(H_i \mid B) = \frac{P(H_i) \cdot P(B \mid H_i)}{P(B)}, \quad i = \overline{1,n},$$

but the full probability formula can be used in the denominator. Then we get:

$$P(H_i \mid B) = \frac{P(H_i) \cdot P(B \mid H_i)}{\sum_{k=1}^{n} P(H_k) \cdot P(B \mid H_k)}, \quad i = \overline{1,n}$$

– **Bayes[*] formula**.

[*] *Thomas Bayes (1702-1761) is an English mathematician.*

The essence of these formulas is as follows: according to the formula of full probability – it is possible to calculate the full (unconditional) probability through conditional probabilities; Bayes' formula – if it is known that some event has occurred, then such additional information makes it possible to recalculate (reevaluate) the hypotheses and obtain the probabilities of the hypotheses, provided that some event has occurred.

We emphasize that $\sum_{i=1}^{n} P(H_i \mid B) = 1$. Formally, it is easy to get: $P(B) = \sum_{i=1}^{n} P(H_i) \cdot P(B \mid H_i)$ and dividing by $P(B)$ we get that $\sum_{i=1}^{n} P(H_i \mid B) = 1$. Although, it is important to emphasize that $\sum_{i=1}^{n} P(H_i) \le 1$, and here equality is achieved only in the case of a complete system of hypotheses, that is, when $\bigcup_{i=1}^{n} H_i = \Omega$, and the left part will be strictly less than one when $\bigcup_{i=1}^{n} H_i \subset \Omega$.

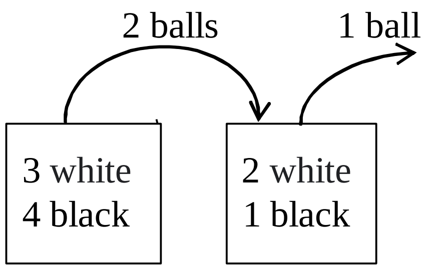
Example. There are white and black balls in two boxes. The first of them has three white and four black balls, the second has two white and one black. From the first box, two balls are randomly transferred to the second, after that, one ball is taken out of the second. The questions are: a) what is the probability that a white ball is drawn? b) what is the probability that two white balls were transferred from the first box to the second, if the ball removed from the second box turned out to be white?

◄ Let's denote the event B – a white ball is taken out of the second box. We will put forward the following hypotheses:

H_1 – two white balls were transferred from the first box to the second;

H_2 – one white and one black ball was transferred from the first box to the second;

H_3 – two black balls were transferred from the first box to the second.



These hypotheses are incompatible and form a complete group of events (hypothesis). Taking into account the notations adopted in point a) it is necessary to calculate $P(B)$, and in point b) $P(H_1 \mid B)$.

$$P(H_1) = \frac{C_3^2}{C_7^2} = \frac{3 \cdot 2}{7 \cdot 6} = \frac{1}{7},\ P(H_2) = \frac{C_3^1 \cdot C_4^1}{C_7^2} = \frac{2 \cdot 3 \cdot 4}{7 \cdot 6} = \frac{4}{7},$$

$$P(H_3) = \frac{C_4^2}{C_7^2} = \frac{4 \cdot 3}{7 \cdot 6} = \frac{2}{7},\ P(B/H_1) = \frac{4}{5},\ P(B/H_2) = \frac{3}{5},\ P(B/H_3) = \frac{2}{5},$$

$$P(B) = \frac{1}{7} \cdot \frac{4}{5} + \frac{4}{7} \cdot \frac{3}{5} + \frac{2}{7} \cdot \frac{2}{5} = \frac{20}{35} = \frac{4}{7},\ P(H_1/B) = \frac{1}{5}.$$

Note that the probability of the first hypothesis was equal to $P(H_1) = \frac{1}{7}$, and as a result of the experiment, when it became known that a white ball was taken out of the second box, it was reassessed and $P(H_1/B) = \frac{1}{5}$ was obtained.

The problem can be solved by putting forward another system of hypotheses:

H_4 – the ball that was removed from the second box was initially in the first box;

H_2 – the ball that was removed from the second box was originally in the second box.

These hypotheses are also incompatible and form a complete group of events. Then

$$P(H_1) = \frac{2}{5},\ P(H_2) = \frac{3}{5},\ P(B/H_1) = \frac{3}{7},\ P(B/H_2) = \frac{2}{3},$$

$$P(B) = \frac{2}{5} \cdot \frac{3}{7} + \frac{3}{5} \cdot \frac{2}{3} = \frac{4}{7}.\ \blacktriangleright$$

§8. Scheme of independent tests. Bernoulli's law (binomial law)

Suppose that the same experiment is carried out many times under the same conditions. The results of previous experiments do not affect the results of subsequent experiments and, naturally, the results of subsequent experiments do not affect the results of previous ones (this is the independence of tests). In each experiment, some event *A* is observed (event *A* is observed). The probability of the occurrence of this event in each of the experiments is known and is equal to $P(A)=p$. The question is what is the probability that in a series of experiments an event *A* occurs (will be observed) a certain number of times?

We will assume that a series of *n* experiments is conducted. Let's ask the question: what is the probability that in a series of *n* experiments event *A* will occur *m* times, where, of course, $0\le m\le n$? The probability that event *A* will occur in *m* experiments is denoted by $P_n(m)$.

The space of elementary events Ω can be presented in the form

$$\Omega=\left\{\underbrace{AAAAA...A}_{n},\;\underbrace{A\bar{A}AAA...A}_{n},\;\underbrace{AA\bar{A}A\bar{A}...A}_{n},\;\underbrace{\bar{A}AAA\bar{A}...A}_{n},...,\;\underbrace{\bar{A}\bar{A}\bar{A}\bar{A}\bar{A}...\bar{A}}_{n}\right\}$$

There are 2^n total points in the space Ω, because there are *n* positions, and each position can be *A* or $\bar{A}$. Each such point must be assigned a probability. Let's denote $P(\bar{A})=1-P(A)=1-p=q$. Then, since all experiments are independent, each point (event) is matched with products: point $AAAAA...A$ should be matched with $ppppp...p$, point $AA\bar{A}A\bar{A}...A$ should be matched with $ppqpq...p$, point $\bar{A}\bar{A}\bar{A}\bar{A}\bar{A}...\bar{A}$, respectively, with $qqqqq...q$, etc. It is easy to see that the sum of the matched probabilities is one, indeed, this sum is actually

$$1=(p+q)^n=\sum_{m=0}^{n}C_n^m p^m q^{n-m}.$$

Let's highlight those elementary results of the experiment in which the event A occurred m times, that is, those points (events) from the set Ω, in which m places are occupied by event A, and the remaining $n-m$ places are occupied by event $\bar{A}$. The probabilities corresponding to such events (points) are equal to $p^m q^{n-m}$. And there are C_n^m such points, that is, we get:

$$P_n(m)=C_n^m p^m q^{n-m}$$

– **Bernoulli's formula** for $m=0,1,2,...,n$.

In practice, the Bernoulli* formula can be used for fairly small values of n (due to the complexity of the calculations).

p.8.1. The most probable number

Suppose that a certain number of experiments are conducted. At the same time, a certain event can appear (happen) a certain number of times. The probability $P_n(m)$ is calculated by Bernoulli's formula. The question arises: for which value of m is the greatest probability $P_n(m)$. In fact, it is necessary to investigate the function $P_n(m)$ as a function of the argument m: we are interested in which m value of the function $P_n(m)$ will be the largest. Such m is called the most probable number.

Derivative methods cannot be used to find the maximum point of a function, because the function is not continuous. Therefore, consider the relationship:

$$\frac{P_n(m+1)}{P_n(m)}=\frac{C_n^{m+1}p^{m+1}q^{n-m-1}}{C_n^m p^m q^{n-m}}=\frac{(n-m)\cdot p}{(m+1)\cdot q}.$$

---

* *Jakob Bernoulli (1655-1705) is a Swiss mathematician.*

Let's investigate at what values of m this ratio is greater than unity (then the function $P_n(m)$ increases), and when the ratio is less than unity (then the function $P_n(m)$ decreases). We get:

1) $\frac{n-m}{m+1}\cdot\frac{p}{q}>1$, from here $np-mp>mq+q$, therefore $np-q>m(q+p)$ and $m<np-q$. That is, if $m<np-q$, then $P_n(m+1)>P_n(m)$, and therefore the function $P_n(m)$ increases;

2) similarly, $\frac{n-m}{m+1}\cdot\frac{p}{q}<1$, then $np-mp<mq+q$, hence $np-q<m(q+p)$ and $m>np-q$. That is, if $m>np-q$, then $P_n(m+1)<P_n(m)$, and therefore the function $P_n(m)$ decreases;

3) $\frac{n-m}{m+1}\cdot\frac{p}{q}=1$, hence $m=np-q$, that is, the probability takes the same values at m and $m+1$: $P_n(m+1)=P_n(m)$.

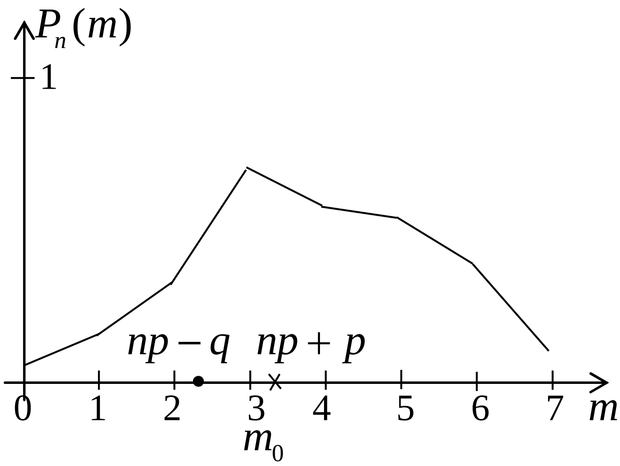
But m can only be an integer. So, the conclusion is as follows: if $m=np-q$ is an integer, then there are two most probable numbers:

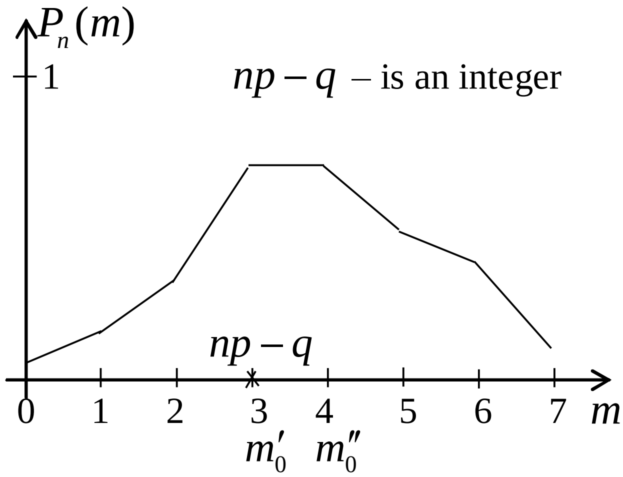
$$m_0'=np-q \text{ and } m_0''=m_0'+1=np+p$$

(because $np-q+1=np+(1-q)=np+p$);
if $m=np-q$ is not an integer, then there is one most probable number

$$m_0=[np-q]+1, \text{ or } m_0=[np+p],$$

here [.] is a whole part of the argument.





Algorithm for finding the most probable number:

1) calculate np; if np is an integer, then $m_0 = np$;

2) if np is not an integer, then calculate $np-q$;

3) if $np-q$ is not an integer, then the most probable number $m_0 = [np-q]+1$;

4) if $np-q$ is an integer, then there are two most probable numbers

$$m_0' = np-q \text{ and } m_0'' = np+p.$$

The most probable number m_0 (if it is unique) is between the numbers $np-q$ and $np+p$, that is, it can be found by the following formula

$$np-q \le m_0 < np+p.$$

p.8.2. Binomial distribution

Consider the scheme of independent tests. A set of numbers $\{P_n(k),\ k=0,\ 1,...,\ n\}$ is called a **binomial probability law** or **binomial probability distribution**.

Closely related to the independent trials design is another probability distribution called the geometric probability distribution.

Consider such a stochastic experiment. Let us be interested in a certain event *A*, which appears in each experiment with probability *p*. We will conduct a series of experiments until the first occurrence of event *A*. Let $B(m)$ denote the event, which consists in the fact that *m* experiments must be conducted before the occurrence of event *A*. The space of elementary events here is as follows:

$$\Omega = \left\{ A,\quad \bar{A}A,\quad \bar{A}\bar{A}A,\quad \bar{A}\bar{A}\bar{A}A,...,\quad \underbrace{\bar{A}\bar{A}...\bar{A}}_{m}A,... \right\}.$$

It is clear that the probability of the event $B(m)$

$$P(B(m)) = q^m p,\quad m = 0,1,2,....$$

The set of numbers $\{q^m p, \quad m = 0,1,2,...\}$ is called a **geometric probability distribution.**

Note that the sum of such numbers is equal to one. This probability distribution has the property of no aftereffect.

We denote by $B(\geq k)$ the event, which consists in the fact that at least k experiments must be conducted before the first occurrence of event A. Let's write the following property:

$$P\left(B(m+n) \mid B(\geq n)\right) = P(B(m))$$

This property is called the property of the absence of an aftereffect in the geometric distribution. We make sure that the last formula is correct.

◄ By definition of conditional probability

$$P\left(B(m+n) \mid B(\geq n)\right) = \frac{P(B(n+m) \cap B(\geq n))}{P(B(\geq n))},$$

but $B(\geq n) \supset B(n+m)$, so $B(n+m) \cap B(\geq n) = B(n+m)$. In addition,

$$B(\geq n) = B(n) \cup B(n+1) \cup B(n+2) \cup ...$$

and the events written on the right are incompatible, so the probability of the sum of events is equal to the sum of the probabilities of these events:

$$P(B(\geq n)) = \sum_{i=n}^{\infty} P(B(i)) = \sum_{i=n}^{\infty} p \cdot q^i = q^n \cdot p \cdot \sum_{i=0}^{\infty} q^i = q^n \cdot p \cdot \frac{1}{1-q} = q^n .$$

So

$$P\left(B(m+n) \mid B(\geq n)\right) = \frac{P(B(n+m))}{P(B(\geq n))} = \frac{q^{n+m} \cdot p}{q^n} = q^m \cdot p = P(B(m)). \ ►$$

It can also be proved that the only probability distribution that has the property of no aftereffect is the geometric distribution.

§9. Limit theorems in the scheme of independent tests

Bernoulli's formula is not convenient to use for large ones due to the complexity of the calculations. In order to simplify the calculations, the accuracy of the calculations must be sacrificed (the probability does not change).

p.9.1 The local Moivre*-Laplace** theorem

Theorem 1 (local Moivre-Laplac theorem).

Let the parameters p, q, $p \geq 0$, $q \geq 0$, $p+q=1$ *be used in the scheme of independent tests. If* $x_{n,m} = \dfrac{m-np}{\sqrt{npq}}$ *is uniformly bounded by a constant C, that is,* $|x_{n,m}| \leq C$ *and the limit* $\lim\limits_{n\to\infty}\sqrt{npq} = \infty$, *then*

$$\lim_{n\to\infty}\frac{\sqrt{2\pi}\sqrt{npq}P_n(m)}{e^{-\frac{1}{2}x_{n,m}^2}} = 1.$$

In the formulation of the theorem, instead of uniform boundedness $x_{n,m}$, it was possible to write: there exists a constant C such that $|x_{n,m}| \leq C$. The condition $\lim\limits_{n\to\infty}\sqrt{npq} = \infty$ is clear when p and q are constant. But experiments are possible in which p and q change as n increases. This theorem is also true for such cases, so this condition is in the theorem. The conclusion of the theorem is equivalent to the fact that the numerator and denominator are equivalent, that is,

$$P_n(m) \approx \frac{1}{\sqrt{2\pi}\sqrt{npq}}e^{-\frac{1}{2}x_{n,m}^2}.$$

---

* *Abraham De Moivre (1667-1754) is an English mathematician of French origin.*

** *Pierre-Simon Laplace (1749-1827) is a French mathematician.*

That is, the theorem proves the following approximate equality:

$$\sqrt{npq}P_n(m) \approx \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_{n,m}^2}.$$

Or, from here

$$P_n(m) = C_n^m p^m q^{n-m} \approx \frac{1}{\sqrt{npq}} \varphi(x_{n,m}),$$

where

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

– is a tabulated function.

So, the algorithm for calculating $P_n(m)$ for large *n* is as follows:

1. Calculate $x_{n,m}$.
2. Find $\varphi(x_{n,m})$ according to the tables.
3. Calculate $P_n(m)$ according to the formula

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \varphi(x_{n,m}).$$

The question arises: at what *n* can the theorem be used? If you write out the remaining terms and carry out all the calculations carefully, you can make sure that the theorem is applicable even when $\sqrt{npq} \geq 5$ (even when $\sqrt{npq} \geq 4$, although the error is approximately 25%).

The formulated theorem is called local, because it estimates the probability $P_n(m)$ that exactly *m* times the event occurred.

p.9.2. Moivre-Laplace integral theorem (a consequence of the local theorem)

Theorem 2 (Moivre-Laplace integral theorem). *Under the conditions of Theorem 1, for any a and b* $(-\infty < a < b < \infty)$, *the probability that* $a \le x_{n,m} \le b$ *is approximately equal to*

$$P_n\{a \le x_{n,m} \le b\} \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}x^2} dx .$$

Let's analyze how to apply Theorem 2. Suppose that we need to calculate the following probability: $P_n\{k_1 \le m \le k_2\}$. It is clear that

$$P_n\{k_1 \le m \le k_2\} = P_n\{k_1 - np \le m - np \le k_2 - np\} =$$

$$= P_n\left\{\frac{k_1 - np}{\sqrt{npq}} \le x_{n,m} \le \frac{k_2 - np}{\sqrt{npq}}\right\}.$$

We denote $a = \frac{k_1 - np}{\sqrt{npq}}$, $b = \frac{k_2 - np}{\sqrt{npq}}$ and now we can apply Theorem 2. Thus, to find the probability $P_n\{k_1 \le m \le k_2\}$ we need

1. calculate $a = \frac{k_1 - np}{\sqrt{npq}}$, $b = \frac{k_2 - np}{\sqrt{npq}}$;
2. apply Theorem 2.

Example 1. Let 100 experiments be conducted $(n = 100)$. Let the probability of a certain event be $p = 0{,}5$ in each experiment. Then $q = 0{,}5$. What is the probability that in a series of 100 experiments this event will occur 60 times?

◀ From the condition of the problem $\sqrt{npq} = 10 \cdot 0{,}5 = 5$. Therefore, theorem 1 can be applied. Under the condition of the problem, $m = 60$. Let's calculate $x_{n,m}$:

$$x_{n,m} = \frac{60 - 100 \cdot 0{,}5}{\sqrt{100 \cdot 0{,}5 \cdot 0{,}5}} = \frac{10}{5} = 2.$$

Then by Theorem 1

$$P_{100}(60) \approx \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{5} \cdot e^{-2} = \frac{1}{5} \cdot \varphi(2) = \frac{1}{5} \cdot 0{,}054 = 0{,}0108. \ ►$$

Example 2. Find the probability that in the same experiment the event will occur at least 40 times and not more than 60 times?

$$◀ \qquad P\{40 \le m \le 60\} = P\left\{\frac{40 - np}{\sqrt{npq}} \le \frac{m - np}{\sqrt{npq}} \le \frac{60 - np}{\sqrt{npq}}\right\} =$$

$$= P\left\{-2 \le \frac{m - np}{\sqrt{npq}} \le 2\right\} \approx \frac{1}{\sqrt{2\pi}} \int_{-2}^{2} e^{-\frac{x^2}{2}} dx.$$

Such an integral is calculated according to the tables. The tables are different, but the tables of the following function

$$\Phi_0(x) \approx \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt, \quad x \ge 0$$

are most often given.

Then

$$P\{40 \le m \le 60\} \approx 2\Phi_0(2) \approx 2 \cdot 0{,}4772 = 0{,}9544. \quad ►$$

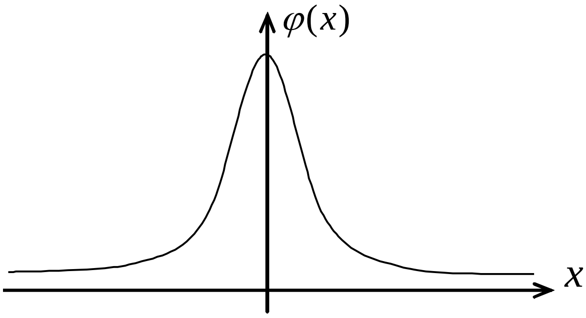
Sometimes there are tables of functions

$$\Phi_1(x) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt, \text{ for } x \le 0 \text{ or } x \ge 0.$$

All these functions can be simply obtained from each other, if we consider that the graph of the function

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

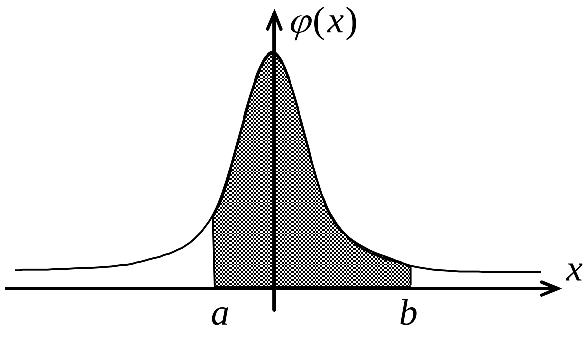
has the form as in the figure



and that the Poisson* integral

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}.$$

For example, at $a < 0$ and $b > 0$ and the tables of the function $\Phi_1(x)$ at $x \le 0$, we have that the required area is equal to $1 - \Phi_1(a) - \Phi_1(-b)$.



---

* *Simeon-Denis Poisson (1781-1840) is a French physicist and mathematician.*

p.9.3 Poisson's theorem (theorem about rare events)

Theorem 3. *Suppose that the condition of Theorem 2 that* $npq \to \infty$ *is not fulfilled when* $n \to \infty$, *but there is a constant C such that* $a_n = np < C.$. *Then, when* $n \to \infty$, *the probability*

$$P_n(m) \approx \frac{(a_n)^m}{m!} e^{-a_n}.$$

The name of the theorem is related to the fact that here the probability p of the occurrence of the event is close to zero (going to zero). There is also a case where the probability q of the non-occurrence of the event approaches zero, but then the probability of the occurrence of the event p approaches one, and it is possible to move to the opposite event, the probability of which will already approach zero.

This theorem should be used when $p < 0,02$ and sufficiently large *n*.

The set of numbers

$$\left\{ \frac{\lambda^m}{m!} e^{-\lambda} \ , \ m = 0, \ 1, \ 2, ... \right\}$$

for λ>0 forms a probability distribution called the **Poisson distribution**.

In fact, the theorem states that for small *p*, the limiting distribution for the binomial probability distribution is, for $n \to \infty$, the Poisson distribution with parameter $\lambda = a_n = np$.

Tables for different λ and *m* have been compiled for the function $\frac{\lambda^m}{m!} e^{-\lambda}$. The function $\sum_{m=0}^{x} \frac{\lambda^m}{m!} e^{-\lambda}$ is often tabulated. Using these tables, you can easily solve problems about finding $P_\lambda(m)$.

CHAPTER 2
RANDOM VARIABLES

§1. The concept of a random variable

The following concepts were considered above: stochastic experiment, frequency, space of elementary events, distribution, random event. Let's consider a new concept of a random variable, which will be defined with the help of the already introduced concept of a random event.

A random variable is a variable that can take on different values in different cases. Examples of random variables can be the number of points that fell out in one toss of a dice, the number of defective products among *n* products taken at random, the number of hits on the target with *n* shots, the time of the device's trouble-free operation, the flight range of the rocket, etc. The random variable ξ is a number that corresponds to each possible outcome of the experiment. Since the results of the experiment are described by elementary events, the random variable can be considered as a function $\xi = \xi(\omega)$ on the space of elementary events Ω .

Before formulating the definition of a random variable, let's consider a few examples.

Example 1. Consider the following stochastic experiment – a coin is tossed twice. The space of elementary events has the form

$$\Omega = \{CC,\ CF,\ FC,\ FF\}.$$

Let ξ be the number of appearances of the coat of arms. The value ξ is a function $\xi = \xi(\omega)$ of an elementary event (an elementary result of an experiment). The table of values of the function $\xi(\omega)$ has the following form:

ω	CC	CF	FC	FF
$\xi(\omega)$	2	1	1	0

Example 2. A dice is thrown. In this case, the space of elementary events has the following form:

$$\Omega = \{\omega_1,\ \omega_2,\ \omega_3,\ \omega_4,\ \omega_5,\ \omega_6\},$$

where ω_i means that *i* points have fallen. The random variable ξ – the number of points dropped – is a function of an elementary event, and $\xi(\omega) = i$, if $\omega = \omega_i$.

Example 3. A coin is tossed until a coat of arms appears. In this case

$$\Omega = \left\{C,\ \ FC,\ \ FFC,\ \ FFFC,...,\ \ \underbrace{FF...FF}_{n-1}C,...\right\} =$$

$$= \{\omega_1,\ \omega_2,\ \omega_3,\ \omega_4,\ ...,\ \omega_n,\ ...\}.$$

Let ξ be the number of coin flips. Then the quantity ξ is a function of the elementary event, and $\xi(\omega) = n$ if $\omega = \omega_n$ $(n = 1,\ 2,...)$.

These examples confirm that random variables can be interpreted as functions on the space of elementary outcomes of a stochastic experiment. We are dealing with the function $\xi(\omega) : \Omega \to \mathbb{R}$. This is a reflection that assigns points ω from the space of elementary events Ω with a number. Considering the $(\Omega, \mathcal{F}, P)$ probability space, only those events that are included in $\mathcal{F}$ can be studied, because only those events are assigned with the number *P*. The same function $\xi(\omega)$ under consideration can be matched with the following events $\{\xi(\omega)<0\},\ldots,\{\xi(\omega)<s\}$,…, which are random events. These events can be studied if they are included in the class of events $\mathcal{F}$. In the future, we will consider reflections $\xi(\omega)$ such that for any $x\in\mathbb{R}$, the event $\{\xi(\omega)<x\}$ is an element of the class of events $\mathcal{F}$, that is, $\{\xi(\omega)<x\}\in\mathcal{F}$.

Definition 1. *The reflection* $\xi(\omega):\ \Omega\to\mathbb{R}$ *is called a* ***random variable*** *if, for any real x, the event* $\{\omega:\xi(\omega)<x\}$ *belongs to the class of events* $\mathcal{F}$ *.*

It can be proved that if $\xi(\omega)$ and $\eta(\omega)$ are two random variables, then $\xi(\omega)\pm\eta(\omega)$ are also random variables, and $a\cdot\xi(\omega)$ is also a random variable. Therefore, a linear combination of random variables is also a random variable. In addition, $\xi(\omega)\cdot\eta(\omega)$ is a random variable, $\dfrac{\xi(\omega)}{\eta(\omega)}$ is a random variable, if the probability is $P\{\eta(\omega)=0\}=0$.

Let $\xi(\omega)$ be a random variable, i.e., for any $x\in\mathbb{R}$, the event $\{\omega:\xi(\omega)<x\}$ is an element of the class of events $\mathcal{F}$, then the probability of this event is determined. This probability $P\{\omega:\xi(\omega)<x\}=P\{\xi(\omega)<x\}$.is denoted by $F_\xi(x)$.

Definition 2. *Function*

$$F_\xi(x)=P\{\omega:\xi(\omega)<x\}$$

is called ***the distribution function of the random variable*** $\xi(\omega)$.

§2. Properties of the distribution function of a random variable

1⁰. The distribution function $F_\xi(x)$ is defined on the entire numerical axis (because the definition requires it).

2⁰. For any real *x*, the distribution function is $0 \le F_\xi(x) \le 1$ (because it is a probability).

3⁰. The distribution function $F_\xi(x)$ is non-decreasing.

4⁰. $\lim\limits_{x\to-\infty} F_\xi(x) = 0$, $\lim\limits_{x\to+\infty} F_\xi(x) = 1$.

5⁰. The distribution function of a random variable is continuous on the left, i.e

$$\lim_{x\to x_0-0} F_\xi(x) = F_\xi(x_0).$$

6⁰. For any real *a* and *b* connected by the relation $a < b$, the equality holds

$$P\{\omega : a \le \xi(\omega) < b\} = F_\xi(b) - F_\xi(a).$$

7⁰. The probability that the random variable $\xi(\omega)$ takes a value that does not exceed x_0 is equal to the value of the distribution function at the point x_0 on the right, i.e.

$$P\{\omega : \xi(\omega) \le x_0\} = F_\xi(x_0 + 0).$$

8⁰. The probability that the random variable $\xi(\omega)$ takes a fixed value x_0 is equal to the jump of the distribution function at the point x_0, i.e.

$$P\{\omega : \xi(\omega) = x_0\} = F_\xi(x_0 + 0) - F_\xi(x_0).$$

If the distribution function is continuous, then the probability of a random variable taking any particular value is zero.

Theorem (Riesz's*). *Let some function* $F(x)$ *be defined on the entire numerical axis, non-decreasing, continuous on the left and such that* $\lim\limits_{x\to-\infty} F(x)=0$ *and* $\lim\limits_{x\to+\infty} F(x)=1$. *Then there is such a probability space* $(\Omega, \mathcal{F}, P)$ *and such a random variable* $\xi(\omega)$ *on it that the distribution function of the random variable* $\xi(\omega)$ *will be the function* $F(x)$, *i.e.* $F_\xi(x)=F(x)$.

In fact, this theorem states that the defining properties of the distribution function of a random variable are monotonicity, continuity on the left, and normalization (equal to zero on $-\infty$ and one on $+\infty$).

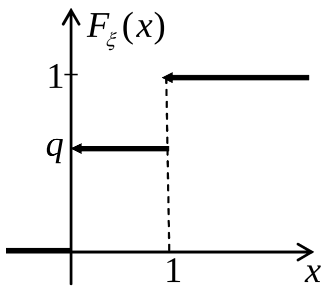
Example 1. Consider a random variable $\xi(\omega)$ that takes two values 0 and 1 with probabilities *q* and *p*, respectively (such a random variable is called an event indicator – the number of occurrences of an event in one trial). It is convenient to record this in the following table:

$\xi(\omega)$	0	1
P	q	p

$$p>0,\quad q>0,\quad p+q=1.$$

Then the distribution function of the random variable $\xi(\omega)$ and its graph have the following form:

$$F_\xi(x)=\begin{cases}0, & x\le 0,\\ q, & 0<x\le 1,\\ 1, & x>1.\end{cases}$$



At the point $x=0$, a jump in the value of the function by the amount *q* is $P\{\xi(\omega)=0\}$, and at the point $x=1$, a jump by the amount *p* and, accordingly, $P\{\xi(\omega)=1\}=p$.

---

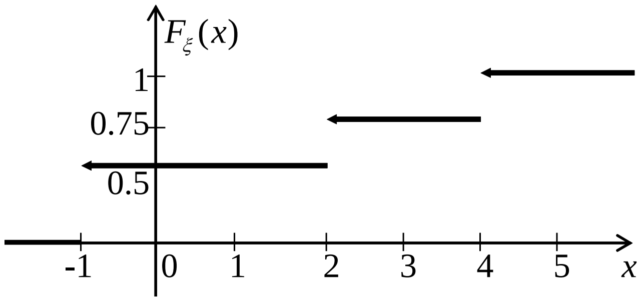
* *Frigyes Riesz (1880-1956) is a Hungarian mathematician.*

Example 2. Let the random variable $\xi(\omega)$ take three values -1, 2 and 4 with the following probabilities: $P\{\xi(\omega)=-1\}=0,5$, $P\{\xi(\omega)=2\}=0,25$ i $P\{\xi(\omega)=4\}=0,25$:

$\xi(\omega)$	-1	2	4
P	0,5	0,25	0,25

Then the distribution function of the random variable $\xi(\omega)$ and its graph have the following form:

$$F_\xi(x)=\begin{cases}0, & x\le -1,\\ 0,5, & -1<x\le 2,\\ 0,75, & 2<x\le 4,\\ 1, & x>4.\end{cases}$$



Example 3. Let the random variable $\xi(\omega)$ take n values $x_1, x_2,..., x_n$ with probabilities: $P\{\xi(\omega)=x_i\}=p_i$, $i=1, 2,..., n$.

$\xi(\omega)$	x_1	x_2	x_3	…	x_n
P	p_1	p_2	p_3	…	p_n

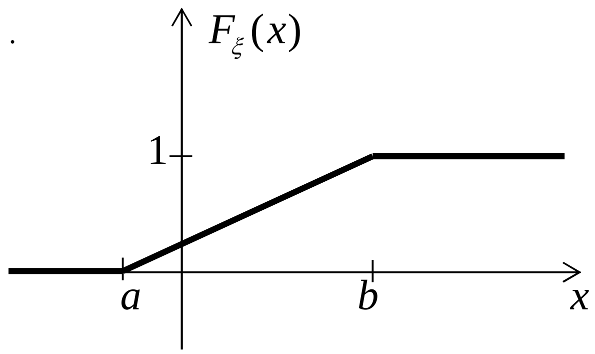
Here $x_1<x_2<...<x_n$, $p_1+p_2+...+p_n=1$. Let's write $F_\xi(x)$ analytically.

$$F_\xi(x)=\begin{cases}0, & x\le x_1,\\ \sum\limits_{i:\,x_i<x} p_i, & x_1<x\le x_n,\\ 1, & x>x_n.\end{cases}$$

Example 4. Let's write the following function $F_\xi(x)$ for $-\infty < a < b < +\infty$:

$$F_\xi(x) = \begin{cases} 0, & x \le a, \\ \dfrac{x-a}{b-a}, & a < x \le b, \\ 1, & x > b. \end{cases} \qquad \text{... (1)}$$

This function is continuous, non-decreasing, equals zero at $-\infty$ and equals one at $+\infty$. Therefore, according to the Riesz's theorem, it is a function of the distribution of some random variable.



If the distribution function of the random variable $\xi(\omega)$ has the form (1) then we say that this random variable is **uniformly distributed** over the interval $[a,\ b]$.

Example 5. If the distribution function of a random variable $\xi(\omega)$ has the form

$$F_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(t-a)^2}{2\sigma^2}} dt,$$

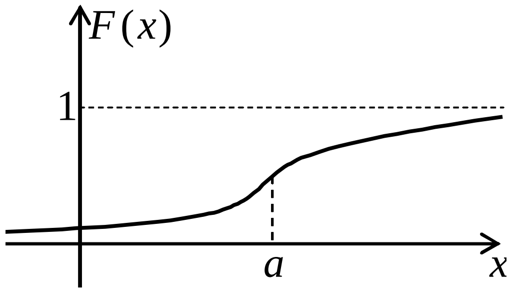
then we say that the random variable $\xi(\omega)$ is **normally distributed** with parameters a and σ $(a, \sigma \in \mathbb{R},\ \sigma > 0)$.

The fact that a random variable is normally distributed with parameters a and σ is written as follows: $N(a,\sigma)$. Such a random variable is often said to be **Gaussian* distributed** with parameters a and σ .

The question arises: can the function

$$F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}e^{\frac{-(t-a)^2}{2\sigma^2}}dt$$

be a distribution function? The graph of this function is shown in the figure



It is continuous, non-decreasing, and since at $x=+\infty$ it is a Poisson function, then $F(x)$ is a distribution function of some random variable.

For $N(0;1)$, such a function was already considered

$$F_{\xi}(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{\frac{-t^2}{2}}dt$$

and called the **Laplace function**.

---

* *Johann Carl Friedrich Gauss (1777-1855) is a German mathematician, astronomer, surveyor and physicist.*

§3. Classification of random variables

Let some random value $\xi(\omega): \Omega \to \mathbb{R}$ be given on the probability space $(\Omega, \mathcal{F}, P)$.

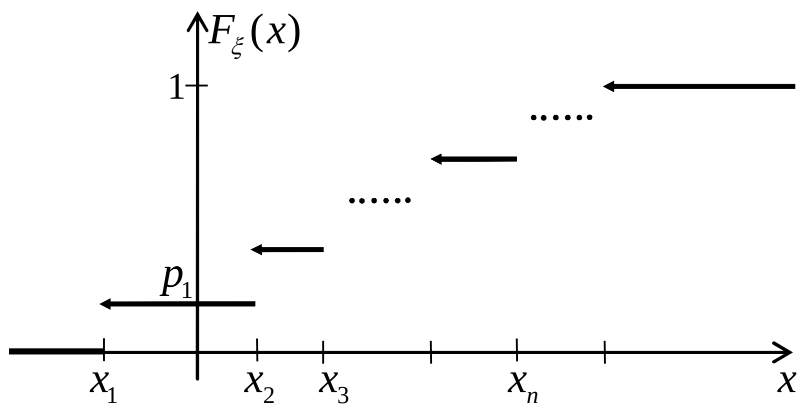
If a random variable $\xi(\omega)$ takes no more than a countable number of values, then such a random variable is called a **discrete random variable**.

These values it takes are denoted by $x_1, x_2, ..., x_n, ...$. The corresponding probabilities $p_1, p_2, ..., p_n, ...$, where

$$p_i = P\{\xi(\omega) = x_i\}, \ \ i = 1, 2, 3, ..., \ \ \sum_i p_i = 1,$$

must also be given. That is, the following table is actually set:

$\xi(\omega)$	x_1	x_2	$\dots$	x_n	$\dots$
P	p_1	p_2	$\dots$	p_n	$\dots$



Thus, the random variable is fully specified and its distribution function can be written:

$$F_\xi(x) = \sum_{i: x_i < x} p_i \ .$$

Therefore, the distribution function of a discrete random variable has no more than the number of growth points (x_0 is called **a growth point** for $F_\xi(x)$ if the inequality

$$F_\xi(x_0+\varepsilon)-F_\xi(x_0-\varepsilon)>0$$

holds for any $\varepsilon>0$).

If the distribution function of a random variable is absolutely continuous, then the corresponding random variable is called **absolutely continuous**, or simply **continuous**.

In the theory of probabilities, only continuous and discrete random variables are studied.

A completely continuous distribution function can be represented as

$$F_\xi(x)=\int_{-\infty}^{x} f_\xi(u)\,du\,, \qquad \text{... (2)}$$

where $f_\xi(u)\geq 0$ is a piecewise continuous function.

Therefore, sometimes the definition of a continuous random variable is given as follows: a random variable is called **continuous** if it can be written in the form (2).

A piecewise continuous function $f_\xi(x)$ is called the **distribution density** of a random variable.

Therefore, a random variable is said to be continuous if it has a density distribution.

p 3.1. Properties of distribution density

1⁰. $f_\xi(x) \geq 0, \ \forall x \in \mathbb{R}$. If this were not so, then the distribution function would not be monotonically nondecreasing.

2⁰. $\int\limits_{-\infty}^{+\infty} f_\xi(u)\,du = 1$.

◀ It must be so, because

$$\int\limits_{-\infty}^{+\infty} f_\xi(u)\,du = \lim_{x\to\infty}\int\limits_{-\infty}^{x} f_\xi(u)\,du = \lim_{x\to\infty} F_\xi(x) = 1$$

by property **4⁰** of the distribution function. ▶

3⁰. $P\{a \leq \xi < b\} = \int\limits_{a}^{b} f_\xi(u)\,du$.

◀ This property follows from property **6⁰** of the distribution function:

$$P\{a \leq \xi < b\} = F_\xi(b) - F_\xi(a) = \int\limits_{-\infty}^{b} f_\xi(u)\,du - \int\limits_{-\infty}^{a} f_\xi(u)\,du = \int\limits_{a}^{b} f_\xi(u)\,du. \ ▶$$

If the distribution function $F_\xi(x)$ is differentiable at the point *x*, then

$$F_\xi'(x) = f_\xi(x).$$

Properties **1⁰ – 3⁰** of the distribution density are often used in problem solving.

Example 1. Let $a > 0$ and

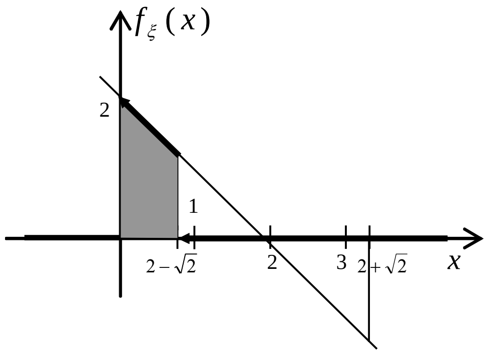
$$f_\xi(x) = \begin{cases} 0, & x \leq 0; \\ -x+2, & 0 < x \leq a; \\ 0, & x > a. \end{cases}$$

Find the constant *a*.

◀ Let's use property **2⁰** of the distribution density:

$$1=\int_{-\infty}^{+\infty} f_\xi(u)du=\int_{-\infty}^{0} f_\xi(u)du+\int_{0}^{a} f_\xi(u)du+\int_{a}^{+\infty} f_\xi(u)du=$$

$$=\int_0^a(-u+2)du=-\frac{a^2}{2}+2a.$$



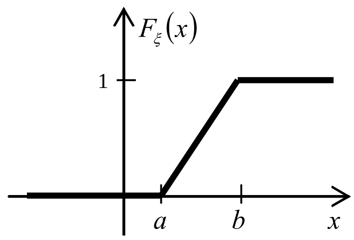
Hence $a^2-4a+2=0, a_{1,2}=2\pm\sqrt{2}$. What to choose a? It is clear from property **1⁰** that $a\neq 2+\sqrt{2}$, because $f_\xi(x)$ at $a=2+\sqrt{2}$ will be less than zero: $f_\xi(2+\sqrt{2})<0$. Therefore, $a=2-\sqrt{2}$. The area of the shaded part is equal to one (according to property **2⁰**). Answer: $a=2-\sqrt{2}$. ▶

If information about a random variable is given in the form of a distribution function or distribution density of a continuous random variable or a table of values and corresponding probabilities of a discrete random variable, then the distribution law of the random variable is said to be given.

Consider examples of density distributions.

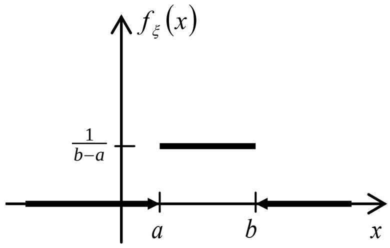
Example 2. For a **uniform distribution** on $[a,b]$, the distribution function $F_\xi(x)$ was written

$$F_\xi(x)=\begin{cases}0, & x<a;\\ \dfrac{x-a}{b-a}, & a\le x\le b;\\ 1, & x>b.\end{cases}$$



Then we get the distribution density

$$f_\xi(x)=\begin{cases}0, & x<a;\\ \dfrac{1}{b-a}, & a\le x\le b;\\ 0, & x>b.\end{cases}$$

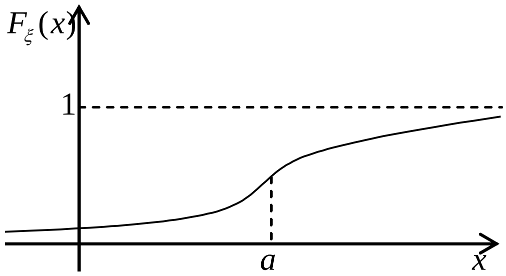


It can also be written differently:

$$f_\xi(x)=\begin{cases}0, & x\notin[a,b];\\ \dfrac{1}{b-a}, & x\in[a,b].\end{cases}$$

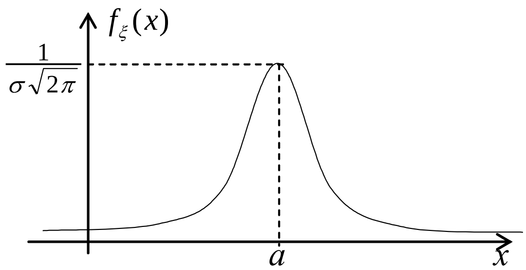
Example 3. For a **normal distribution** with parameters a and σ the distribution function has the following form:

$$F_\xi(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{(u-a)^2}{2\sigma^2}}du$$



Therefore, for $N(a,\ \sigma)$, the distribution density is written as follows:

$$f_\xi(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$$



Example 4. If the distribution function $F_\xi(x)$ of the random variable ξ has the form

$$F_\xi(x) = \begin{cases} 0, & x \le 0; \\ 1 - e^{-\lambda x}, & x > 0. \end{cases},$$

then it is said that the random variable ξ is **distributed according to the exponential law** (exponential law) with the parameter λ.

Then you can write down the distribution density for a random variable distributed according to the exponential law:

$$f_\xi(x) = \begin{cases} 0, & x \le 0; \\ \lambda e^{-\lambda x}, & x > 0. \end{cases}.$$

It turns out that a random variable distributed according to an exponential law has the property of no aftereffect (a discrete random variable distributed according to a geometric law had this property).

Example 5. A random variable is said to be **distributed according to Cauchy's* law** if

$$f_\xi(x) = \begin{cases} 0, & x < 0; \\ \dfrac{2}{\pi}\dfrac{1}{1+x^2}, & x \ge 0, \end{cases}$$

or

$$f_\xi(x) = \frac{1}{\pi}\frac{1}{1+x^2}, \quad -\infty < x < \infty.$$

Then the distribution function is written as follows:

$$F_\xi(x) = \begin{cases} 0, & x \le 0; \\ \dfrac{2}{\pi}\operatorname{arctg} x, & x > 0. \end{cases}$$

---

* *Augustin Louis Cauchy, 1789-1857 – French mathematician.*

§4. Vector random variables

Let the probability space $(\Omega, \mathcal{F}, P)$ and n random variables on it $\xi_1(\omega),\ \xi_2(\omega),...,\ \xi_n(\omega)$ be given. Since these are random values in the probability space, then for any $i = 1,\ 2,...,\ n$ $\{\omega : \xi_i(\omega) < x\} \in \mathcal{F}$. The random variables

$$\xi_1(\omega),\ \ \xi_2(\omega),...,\ \ \xi_n(\omega)$$

can be used to form a vector

$$\left(\xi_1(\omega),\ \ \xi_2(\omega),...,\ \ \xi_n(\omega)\right),$$

which is called a **random vector** – it is a vector whose components are random variables.

Since $\xi_i(\omega)$ are random variables, then ω-set

$$\left\{\omega:\ \xi_1(\omega) < x_1,\ \ \xi_2(\omega) < x_2,...,\ \ \xi_n(\omega) < x_n\right\},$$

where $(x_1,\ \ x_2,...,\ \ x_n) \in \mathbb{R}^n$ $(x_1,\ \ x_2,...,\ \ x_n$ – are real numbers), belongs to $\mathcal{F}$. By this ω-set we understand the intersection of the following sets:

$$\left\{\omega:\ \xi_1(\omega) < x_1,\ \ \xi_2(\omega) < x_2,...,\ \ \xi_n(\omega) < x_n\right\} =$$

$$= \left\{\omega:\ \xi_1(\omega) < x_1\right\} \bigcap \left\{\omega:\ \xi_2(\omega) < x_2\right\} \bigcap ... \bigcap \left\{\omega:\ \xi_n(\omega) < x_n\right\}$$

belonging to $\mathcal{F}$.

This means that there is a (defined) probability of such an event:

$$P\left\{\omega:\ \xi_1(\omega) < x_1,\ \ \xi_2(\omega) < x_2,...,\ \ \xi_n(\omega) < x_n\right\} = F_{\xi_1\xi_2...\xi_n}(x_1, x_2,...,\ \ x_n)$$

and this probability is called the **compatible distribution function** of the system of random variables $\xi_1,\ \ \xi_2,...,\ \ \xi_n$.

Properties of the compatible distribution function

1⁰. $F_{\xi_1\xi_2...\xi_n}(x_1,\ \ x_2,...,\ \ x_n)$ is defined for $(x_1,\ x_2,...,\ x_n)\in\mathbb{R}^n$.

2⁰. $0\le F_{\xi_1\xi_2...\xi_n}(x_1,\ x_2,...,\ x_n)\le 1$.

3⁰. $F_{\xi_1\xi_3\xi_2}(x_1,\ x_3,\ x_2)=F_{\xi_1\xi_2\xi_3}(x_1,\ x_2,\ x_3)$**.**

4⁰. $\lim\limits_{x_1\to+\infty} F_{\xi_1\xi_2...\xi_n}(x_1,\ x_2,...,\ x_n)=F_{\xi_2\xi_3...\xi_n}(x_2,\ x_3,...,\ x_n)$.

It is often written like this:

$$F_{\xi_1\xi_2}(\infty,x_2)=F_{\xi_2}(x_2),\ F_{\xi_1\xi_2}(x_1,\infty)=F_{\xi_1}(x_1).$$

It is clear that $F_{\xi_1\xi_2}(\infty,\infty)=1$.

An equality holds for three variables, for example, the following:

$$F_{\xi_1\xi_2\xi_3}(x_1,\infty,x_3)=F_{\xi_1\xi_3}(x_1,x_3).$$

So, if the compatible distribution function of random variables $\xi_1,\xi_2,...,\xi_n$, is known, it is easy to find the distribution functions of each of the random variables $\xi_1,\xi_2,...,\xi_n$, separately, for example:

$$F_{\xi_1\xi_2\ldots\xi_n}(x_1,\infty,\ldots,\infty)=F_{\xi_1}(x_1).$$

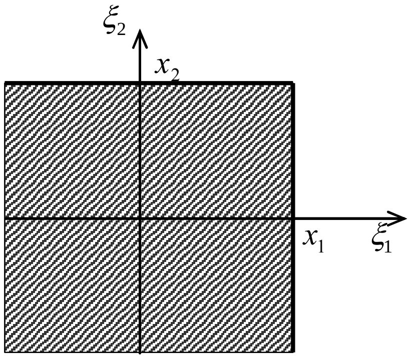
The question arises whether it is possible to find a compatible distribution function based on the known distribution functions $F_{\xi_1}(x_1)$, $F_{\xi_2}(x_2)$,…, $F_{\xi_n}(x_n)$? The answer is negative.

4′⁰. $\lim\limits_{x_i\to-\infty} F_{\xi_1\xi_2\ldots\xi_n}(x_1,\ \ x_2,\ldots,\ \ x_n)=0$.

The event that $\xi_1(\omega) < x_1$ and $\xi_2(\omega) < x_2$ that is,

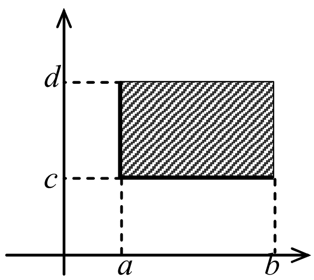
$$\{\omega : \xi_1(\omega) < x_1,\ \xi_2(\omega) < x_2\}$$

can be interpreted as the event that the random vector $\left(\xi_1(\omega),\quad \xi_2(\omega)\right)$ (that



is, a random point on the plane) is in the corner with the vertex at the point with coordinates (x_1, x_2)) (that is, in the negative orthogonal shifted to the point with coordinates (x_1, x_2)).

5⁰. Probability $P\left\{a \le \xi_1(\omega) < b,\quad c \le \xi_2(\omega) < d\right\} =$



$$= P\left\{\xi_1 < b,\quad \xi_2 < d\right\} - P\left\{\xi_1 < a,\quad \xi_2 < d\right\} -$$

$$-P\left\{\xi_1 < b,\quad \xi_2 < c\right\} + P\left\{\xi_1 < a,\quad \xi_2 < c\right\} =$$

$$= F_{\xi_1\xi_2}(b,\quad d) - F_{\xi_1\xi_2}(a,\quad d) - F_{\xi_1\xi_2}(b,\quad c) + F_{\xi_1\xi_2}(a,\quad c).$$

Definition. *Two random variables* $\xi_1(\omega)$, $\xi_2(\omega)$ *are called* ***independent*** *if*

$$F_{\xi_1\xi_2}(x_1, x_2) = F_{\xi_1}(x_1) \cdot F_{\xi_2}(x_2),$$

that is, if the compatible distribution function is equal to the product of the distribution functions.

§5. Numerical characteristics of random variables

In the previous section, several complete, comprehensive characteristics of random variables – the so-called distribution laws – were considered. Such characteristics were: for discrete random variables, the distribution function and the distribution series; for continuous random variables – distribution function and distribution density.

Each distribution law is some function, and the definition of such a function completely describes a random variable from a probabilistic point of view.

But in many practical issues, there is no need to characterize the random variable completely, comprehensively. It is often enough to specify only individual numerical parameters that to a certain extent characterize the essential properties of the distribution of a random variable: for example, some average value around which the possible values of the random variable are grouped; some number that characterizes the degree of dispersion of the values of a random variable around the average, etc. Such characteristics, which make it possible to express the most significant features of the distribution in a concise form, are called numerical characteristics of a random variable.

In the theory of probabilities and mathematical statistics, many different numerical characteristics are used, which have different purposes and different areas of application. In this section, we will consider only some of them, which are most often used.

Consider a scalar random variable ξ on the probability space $(\Omega, \mathcal{F}, P)$ with the distribution function $F_\xi(x)$.

Definition 1. The mathematical expectation *of a random variable* ξ *is called the number* $E\xi$, *which is equal to*

$$E\xi = \int_{-\infty}^{+\infty} x dF_{\xi}(x),$$

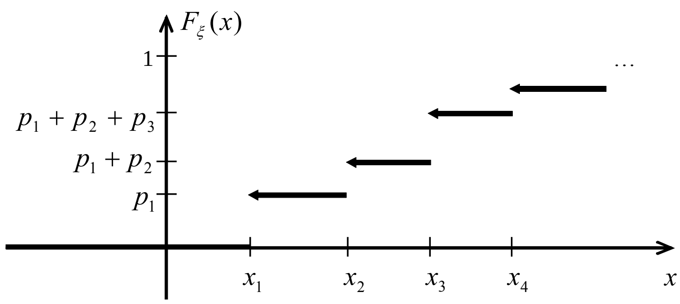
if such an integral absolutely coincides, that is, if an Stilltjes integral exists

$$\int_{-\infty}^{+\infty} |x| dF_{\xi}(x) < \infty.$$

If the integral does not coincide absolutely, then it is said that the mathematical expectation does not exist.

Assume that ξ is a discrete random variable. Then the distribution function is piecewise constant.

ξ	x_1	x_2	x_3	…	x_k	…
P	p_1	p_1	p_3	…	p_k	…



If the piecewise constant function $F(x)$ changes its values only at points $c_1, c_2, ..., c_k, ...$ and the function $f(x)$ is continuous on $[a, b]$, then the Stieltjes integral $\int_{-\infty}^{+\infty} f(x) dF(x)$ is the following sum: $\sum_{k=1}^{\infty} f(c_k)\left[F(c_k+0) - F(c_k-0)\right]$. Using $F(x) = F_{\xi}(x)$ and $f(x) = x$ in this sum, we get

$$\sum_{k=1}^{\infty} x_k \left[F_{\xi}(x_k+0) - F_{\xi}(x_k-0)\right].$$

But the jump of the distribution function at the point x_k, according to property of the distribution function, is equal to

$$F_\xi(x_k+0)-F_\xi(x_k-0)=F_\xi(x_k+0)-F_\xi(x_k)=P\{\xi=x_k\}=p_k\,.$$

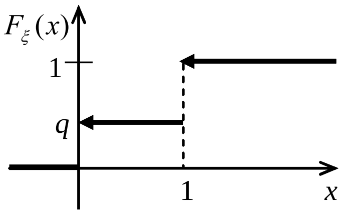
So, in the case of a **discrete random variable, the mathematical expectation** is equal to

$$E\xi=\sum_k x_k p_k\,,$$

if this series absolutely coincides.

Example 1. A random variable ξ takes the value

ξ	0	1
P	q	p



Such a random variable is also called an event indicator. Then the mathematical expectation of this random variable is

$$E\xi=0\cdot q+1\cdot p=p.$$

Example 2. The random variable ξ takes the values $0,1,2,...,n$ with probabilities

$$P\{\xi=k\}=C_n^k p^k q^{n-k},\ p\geq 0,\ q\geq 0,\ p+q=1$$

(Bernoulli distribution, binomial distribution).

$$E\xi=\sum_{k=0}^{n}x_k p_k=\sum_{k=0}^{n}kC_n^k p^k q^{n-k}=\sum_{k=1}^{n}kC_n^k p^k q^{n-k}=np\sum_{k=1}^{n}\frac{k}{n}C_n^k p^{k-1}q^{n-k}=$$

$$=np\sum_{k=1}^{n}\frac{k}{n}\frac{n!}{k!(n-k)!}p^{k-1}q^{n-k}=np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!}p^{k-1}q^{(n-1)-(k-1)}=$$

$$= np\sum_{m=0}^{n-1}\frac{(n-1)!}{m!(n-1-m)!}p^m q^{n-1-m} = np(p+q)^{n-1} = np\,.$$

Mathematical expectation of a random variable distributed according to Bernoulli's law

$$E\xi = np\,.$$

Example 3. Let the random variable ξ be distributed according to the Poisson law with the parameter λ, i.e.

$$P\{\xi = k\} = \frac{\lambda^k}{k!}e^{-\lambda}\,.$$

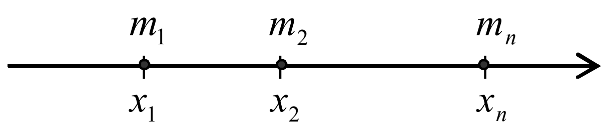
Then

$$E\xi = \sum_{k=0}^{\infty} k\cdot\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k\cdot k}{k!} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!} = \lambda e^{-\lambda}e^{\lambda} = \lambda.$$

For a random variable distributed according to the Poisson law with parameter λ, the mathematical expectation

$$E\xi = \lambda\,.$$

What is the meaning of mathematical expectation for a discrete random variable? Let the masses $m_1, m_2, ..., m_n$ be concentrated at the points



$x_1, x_2, ..., x_n$, respectively.

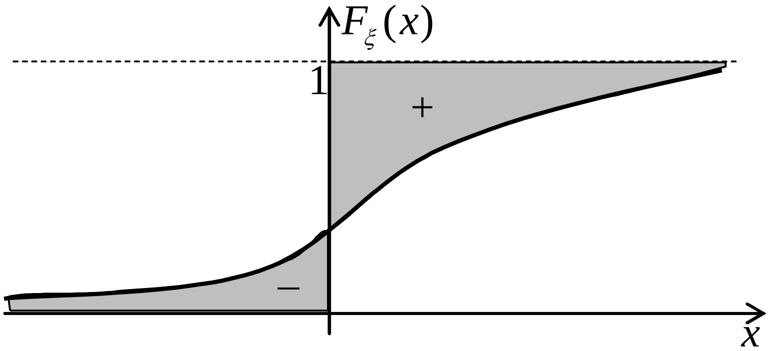
The center of gravity (mass) of such a system is equal to

$$\frac{x_1 m_1 + x_2 m_2 + ... + x_n m_n}{m_1 + m_2 + ... + m_n} = x_1\frac{m_1}{m} + x_2\frac{m_2}{m} + ... + x_n\frac{m_n}{m}\,,$$

where $m = m_1 + m_2 + ... + m_n$.

If the points of the number axis are concentrated masses of p_i, then the mathematical expectation is the center of mass of such a system. That is, the mathematical expectation is a value (some kind of average) around which the values of a random variable are grouped.

Using the definition of the Stiltjes integral, we can give a simple geometric interpretation of the concept of mathematical expectation: the mathematical expectation is equal to the difference of the areas of the figures bounded by the ordinate axis, the line $y=1$ and the curve $y=F_\xi(x)$ on the



interval $(0,+\infty)$ and between the abscissa axis, curve $y=F_\xi(x)$ and the ordinate axis on the interval $(-\infty,0)$. Note that the geometric illustration allows us to write the mathematical expectation in the following form:

$$E\xi=-\int_{-\infty}^{0}F_\xi(x)dx+\int_{0}^{+\infty}(1-F_\xi(x))dx.$$

If the random variable ξ is continuous with the distribution density $f_\xi(x)$, then the Stieltjes integral is equal to

$$\int_{-\infty}^{+\infty}xdF_\xi(x)=\int_{-\infty}^{+\infty}xf_\xi(x)dx.$$

So, for **random variables of absolutely continuous type, the mathematical expectation** is equal to:

$$E\xi=\int_{-\infty}^{+\infty}xf_\xi(x)dx.$$

This is again the center of mass if we consider a rod with density $f_\xi(x)$.

Example 4. Let the random variable ξ be uniformly distributed over the segment $[a, b]$. That is,

$$f_\xi(x) = \begin{cases} 0, & x \notin [a, b]; \\ \dfrac{1}{b-a}, & x \in [a, b]. \end{cases}$$

Then

$$E\xi = \int_{-\infty}^{+\infty} x f_\xi(x) dx = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \frac{x^2}{2}\bigg|_a^b = \frac{b^2 - a^2}{(b-a)2} = \frac{a+b}{2}.$$

For a random variable uniformly distributed on the segment $[a, b]$, the mathematical expectation is

$$E\xi = \frac{a+b}{2}.$$

Example 5. Let the random variable be normally distributed with parameters $N(a,\sigma)$. Let's find $E\xi$. The distribution density $f_\xi(x)$ of this absolutely continuous random variable is equal to

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Then

$$E\xi = \int_{-\infty}^{+\infty} x dF_\xi(x) = \int_{-\infty}^{+\infty} x f_\xi(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{(x-a)^2}{2\sigma^2}} dx =$$

we will replace $\dfrac{x-a}{\sigma} = t$, from which we get $x = \sigma t + a$, $dx = \sigma\, dt$ and so on

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma t + a) e^{-\frac{t^2}{2}} dt = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t e^{-\frac{t^2}{2}} dt + \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = a.$$

Here, the integral of the odd function in the symmetric limits is zero and the Poisson integral is $\sqrt{2\pi}$.

So, for a normally distributed random variable with parameters $N(a,\sigma)$

$$E\xi = a\,.$$

Example 6. Let the random variable be distributed according to Cauchy's law. The density distribution of this absolutely continuous random variable is

$$f_\xi(x) = \frac{1}{\pi}\frac{1}{1+x^2} \quad -\infty < x < +\infty.$$

Consider

$$\int_{-\infty}^{+\infty} |x|\, dF_\xi(x) = \int_{-\infty}^{+\infty} |x|\, f_\xi(x)dx = \frac{1}{\pi}\int_{-\infty}^{+\infty} |x| \frac{1}{1+x^2}dx = \frac{2}{\pi}\int_{0}^{+\infty} x\frac{1}{1+x^2}dx =$$

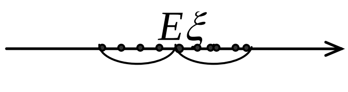
$$= \frac{1}{\pi}\int_{0}^{+\infty} \frac{d(1+x^2)}{1+x^2}dx = \frac{1}{\pi}\ln(1+x^2)\Big|_0^{+\infty} = \frac{1}{\pi}\lim_{x\to+\infty}\ln(1+x^2) = +\infty\,.$$

This means that a random variable distributed according to the Cauchy law does not have a mathematical expectation.

Definition 2. Variance *of a random variable* ξ *is the number* $D\xi$*, which is equal to the mathematical expectation of the square of the deviation of the random variable from its mathematical expectation:*

$$D\xi = E(\xi - E\xi)^2\,.$$

Variance means dispersion and may be interpreted as a measure of dispersion of the values of a random variable around its mathematical expectation.



Consider two random variables

ξ_1	-1	1
P	0,5	0,5

ξ_2	-5	5
P	0,5	0,5

It is obvious that $E\xi_1 = 0$ and $E\xi_2 = 0$, but the degree of dispersion (deviation) of these random variables around $E\xi_1$ and $E\xi_2$ is different.

If the random variable is discrete, then it turns out that

$$D\xi = \sum_k x_k^2 p_k - (E\xi)^2 = \sum_k x_k^2 p_k - \left(\sum_k x_k p_k \right)^2.$$

For a random variable of absolutely continuous type, we have:

$$D\xi = \int_{-\infty}^{\infty} x^2 f_\xi(x)dx - \left(\int_{-\infty}^{\infty} x f_\xi(x)dx \right)^2.$$

Definition 3. The moment of the k-th order *of the random variable* ξ *relative to the point* $x = a$ *is called the number*

$$\nu_k(a) = E(\xi - a)^k \ , \ k = 1, 2, 3, \ldots$$

If $a = 0$, *then* $\nu_k(0) = E(\xi)^k = \nu_k$ *is called the* ***initial moment of the k-th order*** *of the random variable* ξ.

If $a = E\xi$, then the number $\nu_k(E\xi) = E(\xi - E\xi)^k = \mu_k$ is called the **central moment of the *k*-th order** of the random variable ξ.

For the central moment of the *k*-th order, we can write the following:

$$\mu_k = E(\xi - E\xi)^k = E\left(\sum_{i=0}^{k} C_k^i \xi^i (-1)^{k-i} (E\xi)^{k-i} \right),$$

that is, the central moments are expressed through the initial ones.

§2. Properties of numerical characteristics

p.2.1. Properties of mathematical expectation

1⁰. *If the probability* $P\{\xi = C\} = 1$, *then* $E\xi = C$ *(more narrowly, this property is formulated as follows: the mathematical expectation of a constant is equal to a constant).*

◄ We can assume that this random variable is discrete:

ξ	C	$*$
P	1	0

here * is an arbitrary number and then $E\xi = C \cdot 1 + * \cdot 0 = C$. ►

2⁰. *If the probability* $P\{\xi \geq 0\} = 1$, *then* $E\xi \geq 0$. *(If* ξ *takes non-negative values, then the mathematical expectation is non-negative).*

◄ For discrete random variables $\sum x_k p_k \geq 0$ when $x_k \geq 0$ and if the series coincides.

For absolutely continuous random variables

$\int_{-\infty}^{\infty} x f_\xi(x) dx \geq 0$, because $f_\xi(x) \geq 0$ at $x \geq 0$ and $f_\xi(x) = 0$ at $x < 0$. ►

3⁰. *The mathematical expectation of the sum of random variables is equal to the sum of the mathematical expectations of random variables.*

For ξ_1, ξ_2, $E\xi_1$, $E\xi_2$, $\eta = \xi_1 + \xi_2$ we have

$$E\eta = E(\xi_1 + \xi_2) = E\xi_1 + E\xi_2.$$

In general: the mathematical expectation of the sum of a finite number of random variables is equal to the sum of their mathematical expectations

$$E\left(\sum_{i=1}^{n}\xi_i\right)=\sum_{i=1}^{n}E\xi_i .$$

4⁰. *If random variables ξ_1 and ξ_2 are* ***independent****, $E\xi_1$ and $E\xi_2$ exist, then the mathematical expectation of the product of random variables is equal to the product of mathematical expectations:*

$$E(\xi_1\cdot\xi_2)=E\xi_1\cdot E\xi_2 .$$

The importance of the condition of independence of random variables ξ_1 and ξ_2 demonstrated by the following example.

Example. Let the random variable ξ_1 be distributed according to the law

ξ_1	-2	2
P	0,5	0,5

and the random variable $\xi_2=\xi_1^3$, that is, its distribution law is as follows:

ξ_2	-8	8
P	0,5	0,5

$E\xi_1=0$, $E\xi_2=0$ ξ_1 and ξ_2 are dependent. Consider $\xi_1\cdot\xi_2=\xi_1^4$:

$\xi_1\cdot\xi_2$	16
P	1

Then $E(\xi_1\cdot\xi_2)=E\left(\xi_1^4\right)=16\neq E\xi_1\cdot E\xi_2$.

5⁰. *The constant multiplier can be taken as a sign of mathematical expectation:*

$$E(a\xi)=a\cdot E\xi .$$

This statement is obvious, since whatever ξ is, the constant a and the quantity ξ can be considered as independent quantities.

Using $\mathbf{5^0}$, let's summarize $\mathbf{3^0}$:

$\mathbf{3'^0}$. $E(a\xi_1 + b\xi_2) = a \cdot E\xi_1 + b \cdot E\xi_2$.

$\mathbf{6^0}$. *If the random variables* ξ_1 *and* ξ_2 *are such that* $\xi_1 \le \xi_2$, *then* $E\xi_1 \le E\xi_2$.

◀ Indeed, $\xi_2 - \xi_1$ is a nonnegative random variable. Therefore, according to property $\mathbf{2^0}$, $E(\xi_2 - \xi_1) \ge 0$. Using property $\mathbf{3'^0}$, we obtain $E\xi_2 \ge E\xi_1$. ▶

$\mathbf{7^0}$. *It is easy to understand that*

$$|E\xi| \le E|\xi|.$$

This statement follows from the inequalities $-|\xi| \le \xi \le |\xi|$ and property $\mathbf{6^0}$.

$\mathbf{8^0}$. *Let* ξ *be a nonnegative random variable* $\xi \ge 0$, *there is a mathematical expectation* $E\xi$ *and* $a > 0$ *a positive real number. Then the probability*

$$P\{\xi \ge a\} \le \frac{E\xi}{a}.$$

This inequality is called **Chebyshev's first inequality or Markov's inequality.**

$\mathbf{9^0}$. *Let* ξ *be a random variable,* $\eta = \varphi(\xi)$ *be a function of* ξ. *Then the mathematical expectation*

$$E\eta = \int_{-\infty}^{+\infty} \varphi(x)dF_\xi(x).$$

So, in addition to the definition of $D\xi = E(\xi - E\xi)^2$, according to property **9⁰**, you can also obtain the formula

$$D\xi = E\xi^2 - (E\xi)^2.$$

This formula can be obtained in another way (using the properties of mathematical expectation):

$$D\xi = E(\xi - E\xi)^2 = E\left(\xi^2 - 2\xi \cdot E\xi + (E\xi)^2\right) =$$

$$= E\xi^2 - 2E\xi \cdot E\xi + (E\xi)^2 = E\xi^2 - \left(E\xi\right)^2.$$

p.2.2. Variance properties

1^0. *For any random value ξ $D\xi \geq 0$. Moreover, $D\xi = 0$ if and only if there exists such a constant C that $P\{\xi = C\} = 1$ (that is, the variance of the constant is zero, it is possible to prove the opposite: if $D\xi = 0$, then ξ is a constant with probability one).*

2^0. $D(a\xi + b) = a^2 \cdot D\xi$.

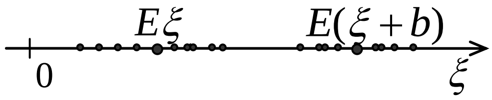
◄ According to the definition of variance

$$D(a\xi + b) = E\left(a\xi + b - E(a\xi + b)\right)^2 =$$

$$= E(a\xi + b - a \cdot E\xi - b)^2 = E\left(a(\xi - E\xi)\right)^2 =$$

$$= a^2 \cdot E(\xi - E\xi)^2 = a^2 \cdot D\xi. \quad ►$$

This property indicates that the variance $a\xi + b$ does not depend on *b*: because the random variable ξ is shifted by *b* and the random variable $\xi + b$ is obtained, and the variance may be interpreted as a measure of dispersion relative to the mathematical expectation (corresponding).



Definition 4. *The number* $\sigma\xi = \sigma_\xi = \sqrt{D\xi}$ *is called the* **standart deviation** *of the random variable* ξ.

Then you can write a property similar to $\mathbf{2^0}$.

$\mathbf{2'^0}$. $\sigma_{a\xi+b} = |a|\sigma_\xi$.

$\mathbf{3^0}$. *The variance of the sum or difference of independent random variables is equal to the sum of the variances of the random variables: if* ξ, η *are independent, then*

$$D(\xi \pm \eta) = D\xi + D\eta .$$

◄ According to the definition of variance

$$D(\xi \pm \eta) = E\left(\xi \pm \eta - E(\xi \pm \eta)\right)^2 = E\left((\xi - E\xi) \pm (\eta - E\eta)\right)^2 =$$

$$= E(\xi - E\xi)^2 + E(\eta - E\eta)^2 \pm 2E(\xi - E\xi)(\eta - E\eta) =$$

if ξ, η are independent, then $\xi - E\xi$ and $\eta - E\eta$ are independent, and then, according to property $\mathbf{4^0}$ of mathematical expectation, we obtain:

$$= D\xi + D\eta \pm 2 \cdot \underbrace{E(\xi - E\xi)}_{=0} \cdot \underbrace{E(\eta - E\eta)}_{=0} = D\xi + D\eta .$$ ►

The variance of the sum of pairwise independent random variables (pairwise independence is sufficient and not necessarily independence in the aggregate) is equal to the sum of the variances of the random variables.

Formulated properties $\mathbf{1^0}$–$\mathbf{3^0}$ are an important tool for calculating variances of random variables.

4^0. *For any $\varepsilon > 0$, the probability that the random variable ξ differs from its mathematical expectation by at least ε does not exceed the variance of the random variable ξ divided by ε^2:*

$$P\left\{|\xi - E\xi| \geq \varepsilon\right\} \leq \frac{D\xi}{\varepsilon^2}.$$

This inequality is called **Chebyshov's inequality** (**the second Chebyshov's inequality**).

◄ Consider the random variable $(\xi - E\xi)^2$. It is clear that

$$\eta = (\xi - E\xi)^2 \geq 0,$$

but also $\varepsilon > 0$. Therefore, according to the Markov inequality (the first Chebyshov inequality), we obtain:

$$P\left\{\eta \geq \varepsilon^2\right\} \leq \frac{E\eta}{\varepsilon^2}.$$

Hence

$$P\left\{\eta \geq \varepsilon^2\right\} = P\left\{(\xi - E\xi)^2 \geq \varepsilon^2\right\} = P\left\{|\xi - E\xi| \geq \varepsilon\right\},$$

but

$$E\eta = E(\xi - E\xi)^2 = D\xi.$$

It follows that

$$0 \leq P\left\{|\xi - E\xi| \geq \varepsilon\right\} \leq \frac{D\xi}{\varepsilon^2}. \text{ ►}$$

Chebyshov's inequality is used when constructing various estimates, when proving theorems. In addition, it can be used to roughly estimate the probability of deviation of a random variable from its mathematical expectation (for any random variable).

Example 1. Let the random variable be distributed according to the binomial law (Bernoulli's law), i.e.:

$$P\{\xi = k\} = C_n^k p^k q^{n-k},\ k = \overline{0,n} \text{ and } p+q=1,\ p>0,\ q>0.$$

$E\xi = np$ has already been calculated. Let's calculate $D\xi$.

Let's compare the random variable ξ with a sequence of independent trials. Let there be *n* such trials. Let us consider random variables ξ_i:

$$\xi_i = \begin{cases} 1, & \text{if event } A \text{ occurred in the } i\text{-th trial,} \\ 0, & \text{if event } A \text{ did not occur in the } i\text{-th trial,} \end{cases} \quad i = 1,2,...,n.$$

Random values ξ_i are independent, equally distributed.

ξ_i	0	1
P	q	p

Then for $i = 1,2,...,n$

$$E\xi_i = 0 \cdot q + 1 \cdot p = p,$$

$$D\xi_i = E\xi_i^2 - (E\xi_i)^2 = 0^2 \cdot q + 1^2 \cdot p - p^2 = p - p^2 = p(1-p) = p \cdot q.$$

It is easy to understand that the random variable $\xi = \xi_1 + \xi_2 + \ldots + \xi_n$. And now we apply the properties of mathematical expectation and variance for the sum of random variables:

$$E\xi = E\xi_1 + E\xi_2 + ... + E\xi_n = np$$

and

$$D\xi = D\xi_1 + D\xi_2 + ... + D\xi_n = nD\xi_i = npq.$$

So, for the binomial distribution (Bernoulli's distribution)

$$D\xi = npq\,,\ \ \sigma_\xi = \sqrt{npq}\,.$$

Example 2. Let the random variable ξ be uniformly distributed on $[a,b]$. This means that

$$f_\xi(x) = \begin{cases} 0, & x \notin [a,b]; \\ \dfrac{1}{b-a}, & x \in [a,b]. \end{cases}$$

In addition, we know that $E\xi = \dfrac{a+b}{2}$. Let's find $D\xi$.

$$D\xi = E\xi^2 - (E\xi)^2 = \int_a^b x^2 \frac{1}{b-a}dx - \left(\frac{a+b}{2}\right)^2 = \frac{1}{b-a}\frac{b^3-a^3}{3} - \frac{(a+b)^2}{4} =$$

$$= \frac{b^2+ab+a^2}{3} - \frac{a^2+2ab+b^2}{4} = \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}.$$

So, when $b > a$

$$D\xi = \frac{(b-a)^2}{12} \quad \sigma_\xi = \frac{b-a}{2\sqrt{3}}$$

for a random variable uniformly distributed on $[a,b]$.

The uniform distribution is determined by the numbers *a* and *b*. On the other hand, knowing $E\xi$ and $D\xi$, these numbers (*a* and *b*) can be found. Therefore, the uniform distribution law is completely determined by two moments (the first initial and the second central). If, in general, random variables are determined by a function or density, then for a uniform distribution law, a random variable is determined by two moments.

Example 3. Let the random variable ξ be normally distributed with parameters $N(a;\sigma)$, $\sigma > 0$. $E\xi = a$ was found. Let's calculate $D\xi$.

$$D\xi = E\xi^2 - (E\xi)^2 = \int_{-\infty}^{\infty} x^2 f_{\xi}(x)dx - a^2 = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} x^2 e^{-\frac{(x-a)^2}{2\sigma^2}}dx - a^2 =$$

replace variables $\frac{x-a}{\sigma} = t$, then $x = \sigma t + a$, $dx = \sigma \cdot dt$ and obtain:

$$= \frac{\sigma}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma^2 t^2 + 2a\sigma t + a^2)e^{-\frac{t^2}{2}}dt - a^2 =$$

$$= \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t^2 e^{\frac{-t^2}{2}}dt + \frac{2a\sigma}{\sqrt{2\pi}}\underbrace{\int_{-\infty}^{\infty} te^{-\frac{t^2}{2}}dt}_{\substack{=0,\ \text{integrand} \\ \text{function is odd}}} + \frac{a^2}{\sqrt{2\pi}}\underbrace{\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}dt}_{=\sqrt{2\pi}} - a^2 =$$

$$= \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}}dt = \frac{\sigma^2}{2\sqrt{2\pi}}\int_{-\infty}^{\infty} \underbrace{t}_{u} \cdot \underbrace{e^{-\frac{t^2}{2}}dt^2}_{dv} =$$

$$= \frac{\sigma^2}{\sqrt{2\pi}}\left(-te^{-\frac{t^2}{2}}\Bigg|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}dt\right) = \frac{\sigma^2}{\sqrt{2\pi}}(0 + \sqrt{2\pi}) = \sigma^2.$$

So, for a normally distributed random variable with parameters $N(a;\sigma)$

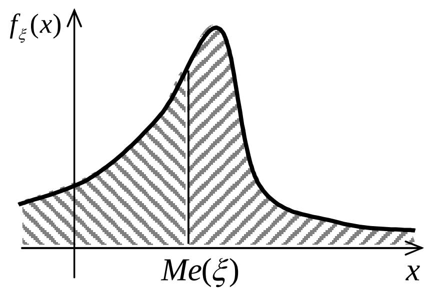
$$D\xi = \sigma^2, \quad \sigma_{\xi} = \sigma.$$

That is, the normal law of distribution is completely determined by two parameters - mathematical expectation and dispersion. There are such random variables that are determined by only one parameter (for example, Poisson's distribution is determined by mathematical expectation).

p.2.3. Characteristics of the position and form of distribution random variable

Let ξ be some random variable with distribution function $F_{\xi}(x)$.

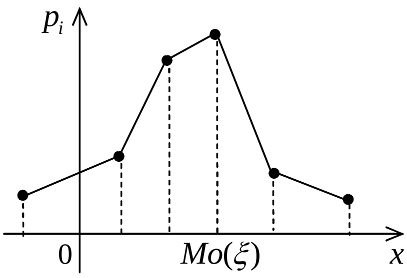
Definition 5. The median of the distribution *of the random variable* ξ *is the number* $Me(\xi)$ *such that* $P\{\xi < Me(\xi)\} = P\{\xi \geq Me(\xi)\} = \frac{1}{2}$. *Thus, the median characterizes the "central value" of a random variable in the sense that the probability of accepting a value less than the median and the probability of accepting a value greater than the median are equal to each other.*

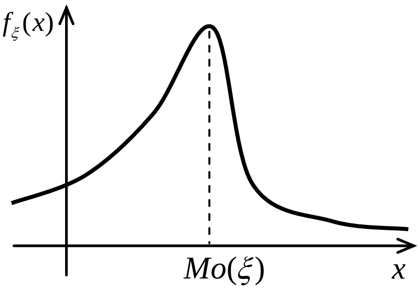


Definition 6. The mode of distribution *of a discrete random variable* ξ *is called its most probable value.*

Let the continuous random variable ξ have the distribution density $f_{\xi}(x)$.

Definition 7. *Each real value of x at which the distribution density reaches a maximum is called the mode of the distribution of a continuous random variable* ξ.





It is customary to denote the mode of distribution of a random variable ξ by $Mo(\xi)$. A distribution with a single mode is called **unimodal** (for example, a normal distribution), in the opposite case – **polymodal** (for example, a uniform distribution on a segment).

Using the definitions of the initial and central moments, as well as the mathematical expectation and variance and their properties, it is easy to obtain that the initial moment of the first order

$$\nu_1 = E\xi,$$

the central moment of the second order

$$\mu_2 = D\xi,$$

and also to express the central points through in

$$\mu_0 = 1,$$

$$\mu_1 = 0,$$

$$\mu_2 = \nu_2 - \nu_1^2,$$

$$\mu_3 = E(\xi - \nu_1)^3 = E(\xi^3 - 3\xi^2\nu_1 + 3\xi\nu_1^2 - \nu_1^3) =$$

$$= \nu_3 - 3\nu_2\nu_1 + 3\nu_1\nu_1^2 - \nu_1^3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3,$$

$$\mu_4 = E(\xi - \nu_1)^4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 4\nu_1\nu_1^3 + \nu_1^4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4.$$

Moments of higher orders are used for a more detailed description of the distributions.

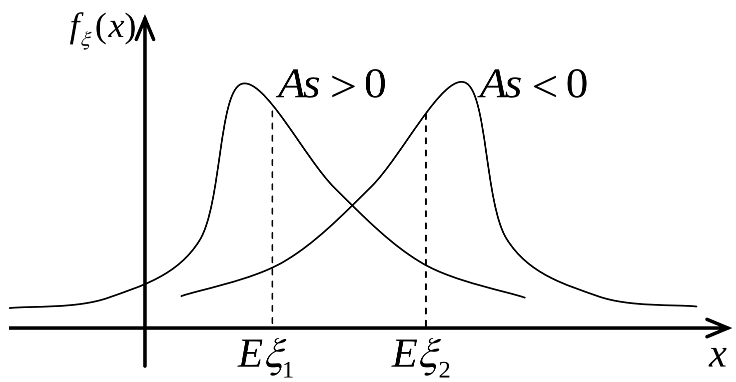
The third central moment serves to characterize the **asymmetry** (or "skewness") of the distribution. If the distribution is symmetrical with respect to the mathematical expectation (or, in a mechanical interpretation, the mass is distributed symmetrically with respect to the center of gravity), then all central moments of odd order (if they exist) are equal to zero.

Therefore, it is natural to choose one of the odd moments as a characteristic of distribution asymmetry. The simplest of them is the third central point. It has the dimension of a cube of a random variable; to obtain a dimensionless characteristic, the third moment of μ_3 is divided by the cube of the root mean square deviation.

Definition 8. *The coefficient of asymmetry, or simply the* ***asymmetry*** *of the random variable* ξ*, is the number As (if it exists), which is equal to*

$$As = \frac{\mu_3}{\sigma^3}.$$

The figure shows two asymmetric distributions – one with positive and the other with negative asymmetry.

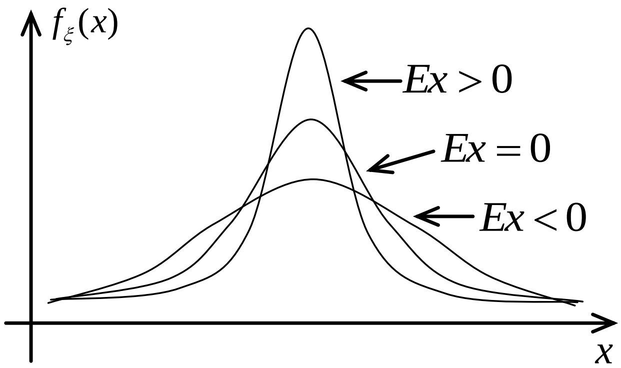


The fourth central moment serves to characterize the so-called "steepness", i.e., the sharp top of the distribution.

Definition 9. *The* **excess** *of a random variable* ξ *is the number* Ex *(if it exists), which is equal to*

$$Ex = \frac{\mu_4}{\sigma^4} - 3.$$

The number 3 is subtracted from the share of $\frac{\mu_4}{\sigma^4}$ because for a normal distribution, as it is easy to establish, $\frac{\mu_4}{\sigma^4} = 3$.



Thus, for a normally distributed random variable, the kurtosis is zero. If the probability distribution of the random variable ξ is unimodal and the density of the distribution $f_\xi(x)$ has a sharper peak than the density of the distribution of a normal random variable with the same variance, then the random variable ξ has positive kurtosis ($Ex > 0$), if $f_\xi(x)$ is less sharp and is more smoothed compared to the density of the corresponding normal distribution, then $Ex < 0$.

p.2.4. Correlation of two random variables

Let two random variables ξ and η be given, for which there are mathematical expectations and variances $E\xi$, $D\xi$, $E\eta$, $D\eta$ and $D\xi \neq 0$, $D\eta \neq 0$.

Definition 1. *The random variable* $\overset{\circ}{\xi} = \xi - E\xi$ *is called the* ***centered random variable*** *for* ξ.

It is characterized by the fact that its mathematical expectation $E\overset{\circ}{\xi} = 0$. Similarly, you can enter a centered random variable for η: $\overset{\circ}{\eta} = \eta - E\eta$ and, accordingly, $E\overset{\circ}{\eta} = 0$.

If $\overset{\circ}{\xi}$ and $E\xi$ are known, then you can go to ξ and vice versa.

Definition 2. *The random variable* $\overline{\xi} = \dfrac{\overset{\circ}{\xi}}{\sqrt{D\xi}} = \dfrac{\overset{\circ}{\xi}}{\sigma_\xi}$ *is called a* ***standardized*** *(standard)* ***random variable****.*

It is characterized by the fact that $E\overline{\xi} = 0$.

Let's find the variance of a standardized random variable

$$D\overline{\xi} = E\left(\overline{\xi} - E\overline{\xi}\right)^2 = E\overline{\xi}^2 = E\left(\frac{\xi - E\xi}{\sigma_\xi}\right)^2 =$$

$$= \frac{1}{\sigma_\xi^2} E\left(\xi - E\xi\right)^2 = \frac{D\xi}{\sigma_\xi^2} = \frac{\sigma_\xi^2}{\sigma_\xi^2} = 1,$$

i.e. $D\overline{\xi} = 1$ – the variance of the standardized random variable is equal to one.

Similarly, you can enter $\overline{\eta} = \dfrac{\overset{\circ}{\eta}}{\sigma_\eta}$ and, accordingly, $E\overline{\eta} = 0$, $D\overline{\eta} = 1$.

You can always go from any random variables to those in which the mathematical expectation is zero and the variance is one.

Definition 3. *The* **covariance** *of the random variables ξ and η is called the mathematical expectation of the product of the corresponding centered random variables $\overset{\circ}{\xi} \cdot \overset{\circ}{\eta}$:*

$$\operatorname{cov}(\xi, \eta) = E\big((\xi - E\xi) \cdot (\eta - E\eta)\big) = E\left(\overset{\circ}{\xi} \cdot \overset{\circ}{\eta}\right).$$

If ξ and η are independent, then $\overset{\circ}{\xi}$ and $\overset{\circ}{\eta}$ are also independent, which means that $E\left(\overset{\circ}{\xi} \cdot \overset{\circ}{\eta}\right) = E\overset{\circ}{\xi} \cdot E\overset{\circ}{\eta} = 0$, that is, $\operatorname{cov}(\xi, \eta) = 0$ for independent random variables. Therefore, covariance can be regarded as a measure (characteristic) of the dependence of two random variables.

Definition 4. *Covariance of standardized random variables is called the* **correlation coefficient** *of two random variables ξ and η:*

$$\operatorname{cov}\left(\overline{\xi}, \overline{\eta}\right) = \rho(\xi, \eta).$$

That is, the correlation coefficient

$$\rho(\xi, \eta) = E\left(\left(\overline{\xi} - E\overline{\xi}\right) \cdot \left(\overline{\eta} - E\overline{\eta}\right)\right) = E\left(\overline{\xi} \cdot \overline{\eta}\right) = E\left(\frac{\xi - E\xi}{\sigma_\xi} \cdot \frac{\eta - E\eta}{\sigma_\eta}\right) =$$

$$= \frac{1}{\sigma_\xi \cdot \sigma_\eta} E\big((\xi - E\xi) \cdot (\eta - E\eta)\big) = \frac{1}{\sigma_\xi \cdot \sigma_\eta} \cdot \operatorname{cov}(\xi, \eta).$$

How does the correlation coefficient characterize the dependence between random variables?

1. If ξ and η are independent, then $\rho(\xi,\eta)=0$. The converse statement is, generally speaking, false (if $\rho(\xi,\eta)=0$, then ξ and η are not necessarily independent).

2. The correlation coefficient for any random variables ξ and η does not exceed one by absolute value: $|\rho(\xi,\eta)|\le 1$.

3. $|\rho(\xi,\eta)|=1$ if and only if there are real numbers A and B such that $\eta = A\xi + B$ (that is, when the relationship between the random variables is linear. Such a relationship, linear, will later turn out to be the closest relationship between the random variables values). More generally, this characteristic is formulated as follows: $|\rho(\xi,\eta)|=1$ if and only if there exist such real numbers A and B that $P\{\eta = A\xi + B\}=1$.

If $\rho(\xi,\eta)>0$, then the relationship between ξ and η is called direct (as ξ increases, η increases and vice versa). If $\rho(\xi,\eta)<0$, it is called inverse (in the sense that as ξ increases, the random variable η decreases and vice versa).

Example 1. For a random variable $\eta = 3\xi - 4$ coefficient correlation $\rho(\xi,\eta)=1$, because the relationship is linear.

Example 2. If $\eta = -27\xi + 17$, then $\rho(\xi,\eta) = -1$.

EXERCISES

Topic 1. "Elements of Combinatorics"

1. There are n roads leading from city A to city B, and m roads from city B to city C. How many ways can you travel along the route A–B–C?
2. 7 roads lead to the top of the mountain. How many ways can a tourist go up and down the mountain? And if the ascent and descent are carried out in different ways?
3. 16 teams are participating in the draw of the national football championship. In how many ways can gold, silver and bronze awards be distributed?
4. How many three-digit numbers can be written using the digits 0, 1, 2, 3, 4?
5. How many three-digit numbers can be written using the digits 0, 1, 2, 3, 4, if each digit is used no more than once?
6. How many five-digit numbers are there that are multiples of 5?
7. 10 subjects are studied in the class. There are 6 lessons on Monday, and all lessons are different. How many ways can you make a schedule for Monday?
8. n points are chosen on one side of the triangle, m points on the other. Each vertex at the base of the triangle is connected by straight lines to points chosen on the opposite side. How many parts will the triangle be divided into by straight lines?
9. n chess players participate in the tournament. How many games will be played in a chess tournament if every two players play each other once?

10. How many diagonals can be drawn in a convex polygon with *n* sides?
11. *n* straight lines are drawn on the plane, and no two of them are parallel and no three intersect at the same point. How many points of intersection are formed in this case?
12. If you turn a sheet of paper 180˚, the numbers 0, 1, 8 do not change, 6 and 9 merge into one another, and the other numbers lose their meaning. How many seven-digit numbers are there whose value does not change when a sheet of paper is turned 180˚?
13. In how many ways can a commission consisting of four people be chosen from nine people?
14. At how many points do the diagonals of a convex *n*-gon intersect, if every three of them do not intersect at one point?
15. There are *n* light bulbs in the room. How many different ways of lighting a room are there?
16. In how many ways can four students be seated in 25 seats?
17. In how many ways can n guests be placed at a round table?
18. How many different words can be formed by rearranging the letters in the word "mathematics"?

Topic 2. "The space of elementary events. Operations on events"

19. A coin is tossed twice. Describe the space of elementary events. Describe the following events:
A – at least once the coat of arms will fall out;
B – the coat of arms will fall out on the second toss.

20. The dice are thrown twice. Describe the space of elementary events. Describe the events
A – the sum of points scored is equal to 8;
B – at most, a 6 will fall out once.

21. A coin is tossed, and then a dice is tossed. Describe the space of elementary events.

22. Construct a set of elementary events in such an experiment: a coin is tossed and a coat of arms is recorded; tossing continues until the coat of arms falls twice.

23. Let the experiment consist in measuring two values that take values from the segment $[0,1]$. Describe the space of elementary events.

24. A and B are subsets of the plane $\Omega = \mathbb{R}^2$, moreover
$$A = \{(x, y) \,:\, x + y \leq 10\},\ B = \{(x, y) \,:\, y \leq 2x + 5\}.$$
Describe sets $A \bigcap B$, $A \bigcup B$, $A \setminus B$, $\bar{A}$, $\bar{B}$, $\bar{A} \cap \bar{B}$.

25. Two points are chosen at random on the segment $[a, b]$. Let x and y be the coordinates of these points. Draw on the plane Oxy the regions corresponding to the events Ω, A, B, $A \bigcap B$, $A \setminus B$, $A \bigcup B$, where the event A is that the second point is closer to the left end of the segment than the first to the right, and B – is the distance between the points less than half the length of the segment.

26. From the numbers 1, 2, 3, 4, 5, first choose one, and then choose the second from the four remaining numbers. Describe the space of elementary events. Describe the events

A – an odd number is chosen first;
B – the second chosen odd number;
C – an odd number was chosen both times.

27. The coin is tossed until it lands on one side twice in a row. Describe the space of elementary events. Describe the events
A – the experiment will end before the sixth toss;
B – an even number of tosses will be made.

28. Three players *a*, *b*, *c* hold a chess tournament according to the following scheme: in the first round *a* and *b* play, player *c* is free, in the second round the winner of the first round and player *c* play, and the player who lost in the first round is free, in the winner of the second round and the player who rested in the second round play in the third round, and the winner of the second round is free. The tournament continues until one of the players wins two games in a row (he is declared the winner of the tournament). There are none. Describe the space of elementary events. Describe the events
A – the winner of the tournament will be player *a*;
B – the winner of the tournament will be player *b*;
C - the winner of the tournament will be player *c*.

29. Two different balls are placed in two boxes. Describe the space of elementary events. Describe the event – there is an empty box.

30. Two identical balls are placed in two boxes. Describe the space of elementary events. Describe the event – there is an empty box.

31. Three shots were fired at the target. Let A_i be the event, which consists in the fact that at the *i*-th shot there is a hit ($i = 1, 2, 3$). Express the following events in terms of A_i events:
A – there are three hits;
B – there was no hit;
C – only one hit;
D – at least two hits.

Topic 3. "Classical definition of probability"

32. A coin is tossed twice. Calculate the probabilities of events
A – at least once the coat of arms will fall out;
B – the coat of arms will fall out on the second toss.

33. The dice are thrown twice. Calculate the probabilities of events
A – the sum of points scored is equal to 8;
B – at most, a 6 will fall out once.

34. Two different balls are placed in two boxes. Find the probability that there is an empty box. And if the balls are the same?

35. There are a white and b black balls in the box. One ball is taken out of the box. She turned out to be white. This ball is put aside. After that, another ball is taken from the box. Find the probability that this ball is also white.

36. There are a white and b black balls in the box ($a \geq 2$). Two balls are taken out of the box at once. Find the probability that both balls are white.

37. The dice are thrown twice. Find the probability that the same number of points will appear both times.

38. Three players are playing cards. Each of them was dealt 10 cards and two are put aside in a "buy-in". One of the players sees that he has 6 diamond cards and 4 non-diamond cards in his hands. He discards two cards from these four and takes two from a "buy-in". Find the probability that he buys two diamond cards.

39. From a box containing n renumbered balls, all balls are taken out at random one by one. Find the probability that the numbers of the drawn balls will be in order.

40. What is the probability that a child who does not know how to read will form the word "*sport*" from the letters *o*, *p*, *r*, *s*, *t* of the split alphabet?

41. What is the probability that the word "*pineapple*" will be formed from the letters *a*, *e*, *e*, *i*, *l*, *n*, *p*, *p*, *p*?

42. What is the probability that the word "*statistics*" will be formed from the letters *a*, *c*, *i*, *i*, *s*, *s*, *s*, *t*, *t*, *t* ?

43. Find the probability that among n numbers selected at random
a) there is no number 5;
b) there is no number 2;
c) there are no numbers 5 and 2;
d) there is no number 5 or no number 2.

44. A dice is thrown 6 times. Calculate the probability that all six heads will fall out.

45. There are 5 passengers in the elevator; the elevator stops on the 9th floor. What is the probability that all passengers exit the elevator on different floors?

46. Calculate the probability that the birthdays of 12 people will be in different months of the year.

47. Calculate the probability that for the given thirty people out of 12 months of the year, 6 months have two birthdays and 6 have three birthdays.

48. There are five segments whose lengths are 1, 3, 5, 7, and 9 units, respectively. Determine the probability that a triangle can be constructed from the three selected segments.

49. 16 teams participate in the hockey tournament, from which two groups of 8 teams are randomly formed. There are 5 extra-class teams among the participants of the tournament. Find the probabilities of events
A – all extra-class teams will be in the same group;
B – two extra-class teams will fall into one of the groups, and three – into another.

50. A full deck of cards (52 cards) is randomly divided into two equal parts of 26 cards. Find the probabilities of events

A – two aces will appear in each of the parts;
B – in one of the parts there will be no aces, and in the other – all four;
C – in one of the parts there will be one ace, and in the other – three.

51. There are M defective parts in a batch consisting of N parts. n parts were chosen at random $(n \leq N)$. What is the probability that among them m defective $(m \leq M)$?

52. A coin is tossed until it lands on one side twice in a row. Describe the space of elementary events. Find the probabilities of events
A – the experiment will end before the sixth toss;
B – an even number of tosses will be made, and the probability that the experiment will continue indefinitely.

53. Three players a, b, c hold a chess tournament according to the following scheme: in the first round a and b play, player c is free, in the second round the winner of the first round and player c play, and the player who lost in the first round is free, in the winner of the second round and the player who rested in the second round play in the third round, and the winner of the second round is free. The tournament continues until one of the players wins two games in a row (he is declared the winner of the tournament). There are none. Describe the space of elementary events.
A – the winner of the tournament will be player a;
B – the winner of the tournament will be player b;
C – the winner of the tournament will be player c;
N – the winner will not be revealed until the n-th round;
D – the tournament will never end.

54. There are n pairs of shoes. From them, $2r$ shoes are randomly selected $(2r < n)$. What is the probability that among the selected shoes
a) there are no pairs;
b) there is exactly one complete pair;
c) are there exactly two complete pairs?

Topic 4. "Geometric probability"

55. In a circle of radius R a dot is thrown at random. What is the probability that the distance from this point to the center of the circle does not exceed r?

56. Two ships must approach the same berth. The appearance of ships are independent events that are equally possible during the day. Find the probability that one of the ships will have to wait for the release of the berth, if the first ship's parking time is one hour, and the second one is two hours.

57. Parallel straight lines are drawn on the plane, the distance between which is equal to $2a$. A circle of radius r $(r < a)$ is thrown at random on the plane. What is the probability that the circle does not cross any of the lines?

58. What is the probability that the sum of two randomly taken positive numbers, each of which does not exceed one, does not exceed one, and their product does not exceed $\frac{2}{9}$?

59. Parallel straight lines are drawn on the plane, the distance between which is equal to $2a$. A needle with a length of $2l$ $(l < a)$ is randomly thrown onto this plane. What is the probability that the needle will cross one of the lines? (Buffon's problem, 1777)

60. A random point $M(\xi, \eta)$ is chosen at random in a square with vertices (0; 0), (0; 1), (1; 1), (0; 1). Find the probability that the roots of the equation $x^2 + \xi x + \eta = 0$ are real.

61. A rod of length l was broken into three parts, choosing the place of the break at random. Find the probability that a triangle can be formed from the three parts obtained.

62. A point is randomly thrown into a circle of radius R. Find the probability that it will be outside the square inscribed in this circle.

63. Two real numbers x and y are chosen at random so that $|x| \le 3,\ |y| \le 5$. What is the probability that the fraction $\frac{x}{y}$ is positive

64. Two real numbers x and y are chosen at random so that $|x| \le 1,\ |y| \le 1$. What is the probability that $|x| < |y|$?

65. Two real numbers x and y are chosen at random so that $|x| \le 1,\ 0 \le y \le 1$. What is the probability that $x^2 < y$?

66. A point is randomly thrown into a square with vertices (0,0), (0,1), (1,0), (1,1). Let (ξ, η) be its coordinates. Find the probability $P\{|\xi - \eta| < z\}$ for $0 < z < 1$.

67. A point is randomly thrown into a square with vertices (0,0), (0,1), (1,0), (1,1). Let (ξ, η) be its coordinates. Find the probability $P\{\min(\xi, \eta) < z\}$ for $0 < z < 1$.

68. A point is randomly thrown into a square with vertices (0,0), (0,1), (1,0), (1,1). Let (ξ, η) be its coordinates. Find the probability $P\{\max(\xi, \eta) < z\}$ for $0 < z < 1$.

Topic 5. "Basic theorems of probability theory"

69. There are a white and b red balls in the box. Two balls are taken out of the box. What is the probability that these balls *a*) different colors; *b*) both white?

70. A coin is tossed twice. Describe the following events: *A* – the coat of arms fell out on the first toss; *B* – the coat of arms fell out on the second toss. Calculate probabilities $P(B)$, $P(A \cap B)$, $P(B | A)$.

71. A coin is tossed three times. Describe the following events: *A* – coat of arms fell out twice; *B* – at least one coat of arms fell out. Calculate probabilities $P(A)$, $P(B)$, $P(A \cap B)$, $P(A / B)$.

72. The student came to the exam knowing the answers to only 20 of the 25 questions in the program. The examiner asked the student 3 questions. Find the probability that the student knows the answers to all these questions.

73. Two dice are thrown. What is the probability that there will be at most one six if the sum of the points is known to be eight?

74. Throw three dice. What is the probability that at least one of them will get one point, if all three dice have different faces?

75. There are 12 red, 8 green and 10 blue balls in a box. Two balls are drawn at random. What is the probability that balls of different colors were drawn, if it is known that no blue ball was drawn?

76. Let the probability of sinking a ship for one torpedo equal to 0.3. What is the probability that four torpedoes will sink a ship if one torpedo hit is enough to sink it?

77. Six dice are thrown. Find the probability that all six heads will fall.

78. The wardrobe gave out numbers to four people who handed in their hats to the wardrobe at the same time. After that, she mixed up all the hats

and hung them up at random. Find the probabilities of the following events: A – the wardrobe will give each of the four people his own hat; B – exactly three people will receive their hats; C – exactly two people will receive their hats; D – exactly one person will receive his hat; E – none of the persons will receive their hat.

79. There are 9 new tennis balls in the box. Three balls are taken for the game, after the game they are put back in the box. Balls that have been played with and those that have not been played with are no different. What is the probability that after three games there are no unplayed balls (new balls) left in the box?

80. Two coins are tossed. Events are considered A – falling coat of arms on the first coin; B – falling coat of arms on the second coin. Find the probability of an event $C = A \bigcup B$.

81. Four cards are taken out of a deck of cards (36 sheets). Events under review A – there will be at least one diamond among these cards; B – there will be at least one king among these cards. Find the probability of the event $C = A \bigcup B$.

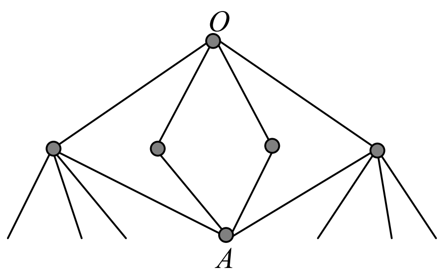
82. Rockets are fired at the target. The probability of each missile hitting the target is p. Hitting individual missiles are independent events. Each missile that hits the target destroys it with probability p_1. Shoot until the target is destroyed or until all the rocket ammunition is used. There are n rockets. What is the probability that a) not all ammunition will be used; b) after destroying the target, at least two missiles will remain; c) will no more than two rockets be used?

Topic 6. "Formulas of Total Probability, Bayes"

83. There are 5 rifles in the shooting range, each of which has a probability of hitting the target equal to 0.5; 0.6; 0.7; 0.8 and 0.9. Determine the probability of hitting the target with one shot if the rifle is selected at random.

84. There are two batches of products of 12 and 10 pieces, and in each batch one product is defective. The product taken at random from the first batch is transferred to the second. After that, one product is selected from the second batch. Determine the probability of choosing a defective product from the second batch.

85. The traveler leaves point *O* and at each intersection chooses one of the paths at random. What is the probability that the traveler will get to point *A*?



86. There are three identical boxes. In the first *a* white and *b* black balls; in the second *c* – white and *d* black balls; in the third – only white balls. One box is chosen at random and one ball is taken out of it. What is the probability that this ball is white?

87. Two dice are taken from a complete set of dominoes. Determine the probability that the second can be attached to the first.

88. In the first box there are 10 balls, 8 of them are white; in the second box there are 20 balls, of which 4 are white. One ball was taken at random from each box, and then one ball was chosen at random from these balls. Find the probability that a white ball is chosen.

89. There are 15 tennis balls in a box, 9 of them are new. For the first game, three balls are taken at random, which are returned to the box after the game. For the second game, three balls are also taken at random. Find the probability that all the balls taken for the second game are new.

90. Two machines produce the same parts, which are placed on a common conveyor. The productivity of the first machine is twice the productivity of the second machine. The first machine makes 60% of parts of excellent quality, and the second – 84%. A part was taken from the conveyor, which turned out to be of excellent quality. Find the probability that this part was made by the first machine.

91. There is one ball (white or black) in the box. A white ball is thrown into the box. Then a ball is taken out at random from the box. She turned out to be white. What is the probability that there is a white ball left in the box? (*Gartner's task*).

92. In a group of 10 students who came for an exam, three students are excellently prepared, four are well prepared, two are mediocre and one is poorly prepared. There are 20 questions in the exam tickets. An excellently prepared student can answer all 20 questions, a well-prepared student can answer 16 questions, an average student can answer 10 questions, and a poorly prepared student can answer 5 questions. A student called at random answered three arbitrarily asked questions. Find the probability that this student is prepared *a*) excellent; *b*) bad.

93. One lord, fed up with his seer with his false predictions, decided to execute him. But being a good lord, he decided to give the seer one last chance. He was ordered to divide four balls into two boxes: two black and two white (without leaving an empty box). The executioner will choose one of the boxes at random and take one bullet from it. If this ball is black, then the seer will be executed, if it is white, then his life will be saved. How should the seer place the balls in the boxes to ensure he has the greatest chance of being saved? (*Joke task*).

94. There are two boxes: in the first a white and b black balls; in the second c – white and d black balls. A box is chosen at random and one ball is taken out of it. She turned out to be white. Find the probability that the next ball drawn from the same box will also be white.

Topic 7. "Scheme of independent tests: Bernoulli's formula, the most probable number"

Indication 1. If n independent trials are conducted under the same conditions, and each trial can have k consequences A_1, A_2,..., A_k (mutually exclusive) with probabilities p_1, p_2,..., p_k, $\sum_{j=1}^{k} p_j = 1$, then the probability that event A_1 will appear in m_1 test, event A_2 in m_2 tests, etc., event A_k in m_k $\left(\sum_{i=1}^{k} m_i = n\right)$ is expressed by the formula

$$P_n(m_1, m_2, ..., m_k) = \frac{n!}{m_1! \cdot m_2! \cdot ... \cdot m_k!} p_1^{m_1} \cdot p_2^{m_2} \cdot ... \cdot p_k^{m_k} \text{ or}$$

$$P_n(m_1, m_2, ..., m_k) = n! \prod_{j=1}^{k} \frac{1}{m_j!} p_j^{m_j} .$$

Instructions 2. If n independent tests are conducted under the same conditions, in the first of which the probability of success is p_1, in the second – p_2, etc., in the n-th – p_n, then the probability that success will occur k times is equal to the coefficient at z^k of the function

$$\varphi_n(z) = \prod_{i=1}^{n} (p_i z + q_i) = (p_1 z + q_1)(p_2 z + q_2)...(p_n z + q_n),$$

which is called the generating function of the probability. The sum of all coefficients of the generating probability function is equal to one.

95. What is more likely to win against an equal opponent: a) three parties out of four, or five out of eight; b) at least three batches out of four, or at least five batches out of eight;

96. Firing at the target with three projectiles is conducted. Projectiles hit the target independently of each other. For each projectile, the probability of hitting the target is 0,4. If one projectile hits the target, it hits it (disables it) with a probability of 0,3; if two shells - with a probability of 0,7; if three shells – with a probability of 0,9. Find the total probability of the target being impressed.

97. The battery fired 14 shots at the target, the probability of hitting which is equal to 0,2. Find the most probability number of hits and the probability of this number of hits.

98. The probability of hitting the target with each shot from the cannon is 0,8. How many shots must be fired so that the most probability number of hits is 20?

99. The probability of at least one occurrence of an event in four independent trials is 0,59. What is the probability of an event occurring in one trial if the probability is the same in each trial?

100. Event *C* will occur if event *B* occurs at least three times. Determine the probability of occurrence of event *C*, if the probability of occurrence of event *B* in one trial is equal to 0,3 and
a) 5 independent trials are conducted;
b) 7 independent tests.

101. One mathematician, who smokes, carries two boxes of matches with him. Every time he wants to get a match, he chooses one of the boxes at random. Find the probability that when a mathematician chooses an empty box for the first time, there will be *r* matches in another box ($r = 0, 1, 2, ..., n$, where *n* is the number of matches that were in each box initially). (*Banach's problem*).

102. A person belonging to a certain population group is a brunette with a probability of 0,2, a brown-haired person with a probability of 0,3, a

blond with a probability of 0,4, and a redhead with a probability of 0,1. A group of six people is chosen at random. Find the probabilities of the following events:

A – there are at least four blondes in the group;

B – there is at least one redhead in the group;

C – the group has the same number of blondes and brown-haired people.

103. What is the probability of hitting the target at least twice, if the probability of hitting is 0,2 and 10 independent shots are fired? Find the probability, at most, of two hits, given that one hit has occurred.

104. In a match for the world chess championship, the game goes to six victories of one of the participants (draws are not taken into account, a win – 1 point, a loss or a draw – 0 points). Assuming that the participants in the match are equal in strength, find the probability that at the end of the match the loser will score *k* points ($k = 0, 1, 2, ..., 5$).

105. Four independent shots are fired at some target. The probability of hitting with different shots is equal to $p_1 = 0,1$; $p_2 = 0,2$; $p_3 = 0,3$; $p_4 = 0,4$. Find the probabilities $P_4(1)$, $P_4(2)$, $P_4(3)$, $P_4(4)$, $P_4(0)$, $P_4(\geq 1)$, $P_4(\geq 2)$.

106. The factory manufactures products, each of which passes four types of tests. The product passes the first test with a probability of 0,9; the second – 0,95; the third – 0,8; the fourth – 0,85. Find the probability that the product will be successful

a) all four tests;

b) two tests;

c) at least two tests.

Topic 8. "Scheme of independent tests: Laplace and Poisson formulas"

107. The probability that the mileage of the car without repair will be at least 20000 km is equal to 0,75. There are 550 cars in the garage. Determine the probability that such machines will turn out to be:
a) 400;
b) not less than 400.

108. The probability that the buyer needs shoes of the 41st size is 0,2. Find the probability that among 100 buyers need shoes of size 41:
a) 25 persons;
b) from 10 to 30 buyers;
c) no more than 30 buyers;
d) not less than 35 people.

109. The probability of a positive result in each of the experiments is equal to 0,9. How many experiments should be conducted so that with a probability of 0.98 it can be expected that at least 150 experiments will give a positive result?

110. The technical control department checks 900 parts for standardization. The probability that the part is standard is 0,9. Find with a probability of 0,9544 the boundaries that contain the number of m standard parts among the tested.

111. On average, 85% of first grade products come off the conveyor. How many products should be taken so that with a probability of 0,997 the deviation of the frequency of products of the first grade in them from 0,85 in absolute value does not exceed 0,01?

112. There are 500 students at the faculty. What is the probability that September 1 is the birthday of k students of this faculty at the same time? Calculate the specified probability for the values $k = 0, 1, 2, 3$.

113. The plant sent 5000 good-quality products to the base. The probability of damage to each product in transit is 0,0002. What is the probability that among 5000 products in transit:
a) exactly three products will be damaged;
b) exactly one product;
c) no more than three products;
d) more than three products?

114. The collection on the theory of probabilities was printed in a circulation of 100000 copies. The probability of incorrect brochures is 0,0001. Find the probability that the circulation has:
a) exactly 5 defective books;
b) not less than 5 and not more than 8 defective books;
c) more than three defective books?

115. The device consists of three independently working elements. The probabilities of failure-free operation of the elements are equal to $p_1 = 0,7$, $p_2 = 0,8$, $p_3 = 0,9$. Find the probability that during time T
a) all elements will work flawlessly;
b) two elements will work without fail;
c) one element will work without fail;
d) none of the elements will work flawlessly.

116. A gun is fired at a moving target. At the first shot, the probability of hitting is $p_1 = 0,8$. With each subsequent one, the probability decreases by a factor of two. Four shots were fired. Determine the probability that there was at least one hit.

Topic 9. "Random variables: the distribution law of the random variable, the distribution function of the random variable "

117. There are 5 white and 25 black balls in a box. One ball is taken out. The random variable ξ is the number of white balls that were taken out of the box. Construct the distribution series and the distribution function $F_{\xi}(x)$ of the random variable ξ.

118. Three coins are tossed. It is necessary: a) to set a random value ξ, which is equal to the number of "digits" that will fall out; b) construct the distribution series and the distribution function $F_{\xi}(x)$ of the value ξ, if the probability of the “coat of arms” falling is 0.5.

119. Two shooters shoot one shot at one target. The probability of hitting for the first shooter with one shot is 0.5, for the second - 0.4. The random variable k is the number of hits on the target. Find the distribution law and the distribution function ξ and calculate the probabilities of events $\{\xi < 2\}$, $\{\xi \geq 1\}$, $\{1 \leq \xi < 2\}$.

120. Successive reliability tests of five devices are carried out. Each subsequent device is tested only if the previous one was reliable. Construct a series of the distribution of the random variable ξ – the number of tested devices, if the probability of passing the test for each of them is 0.9.

121. There are 7 balls in the box, 4 of which are white. 3 balls are taken out at random from the box. Write down the law of distribution of the number of white balls among the chosen ones.

122. Out of 25 graduate students, only 5 defended with "excellence". 3 theses were chosen at random. Find the law of distribution of the random variable ξ – the number of works rated "excellent" among the selected. Calculate the probability of the event $\{\xi > 0\}$.

123. A worker maintains 4 machines that work independently. The probability that the machine will not require worker service within an hour is equal to 0.7 for the first machine, 0.75 for the second, 0.8 for the third, and 0.9 for the fourth. Construct a law for the distribution of a random number of machines that will not require the maintenance of a worker.

124. The subscriber forgot the last digit of the phone number, but remembers that it is even. Construct a series (law) and the distribution function of the random variable ξ – the number of telephone number dials made by him, if he dials the last digit at random, and does not dial the dialed digit in the future.

125. A dice is thrown *n* times. Find the distribution function of the number of sixes.

126. The probability that the clock needs additional adjustment is 0.2. Draw up a law for the distribution of the number of watches that need additional adjustment among three randomly selected ones. Construct a distribution function.

Topic 10. "Continuous random variables: distribution laws, distribution function and density"

127. The random variable is given by the distribution function

$$F_\xi(x)=\begin{cases}0, & x\le 2,\\ (x-2)^2, & 2<x\le 3,\\ 1, & x>3.\end{cases}$$

Find:

a) distribution density $f_\xi(x)$;

b) the probability of the value ξ falling into the interval $(1;\ 2{,}5)$;

c) the probability of the value ξ falling into the interval $(2{,}5;\ 3{,}5)$.

128. (*Cauchy distribution*). The distribution function of the random variable ξ is given by the formula $F_\xi(x)=A+B\text{arctg}x$ $(-\infty<x<\infty)$. Find:

a) constants *A* and *B*;

b) distribution density $f_\xi(x)$;

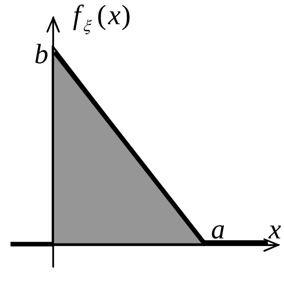
c) the probability that the value ξ will fall on the segment $[-1;\ 1]$.

129. The random variable ξ is distributed according to the "right triangle law" on the interval $[0;\ a]$.

a) Write down the expression for the distribution density $f_\xi(x)$.

b) Find the distribution function $F_\xi(x)$.

c) Find the probability of a random hit values of ξ on the section from $\frac{a}{2}$ to a.



130. The distribution curve of the random variable ξ is the upper arc of an ellipse with semi-axes *a* and *b*. The quantity *a* is known. Determine the value of *b* and construct the distribution function $F_\xi(x)$.

131. The random variable ξ has the distribution density $f_\xi(x) = \dfrac{A}{1+x^2}$ (*Cauchy's law*). Find the coefficient *A* and the distribution function $F_\xi(x)$.

132. The random variable ξ is given by the distribution function $F_\xi(x) = \dfrac{1}{2} + \dfrac{1}{\pi}\operatorname{arctg}\dfrac{x}{2}$ $(-\infty < x < \infty)$. Find a possible value of x_1 that satisfies the condition: with a probability of 0,25, the random variable ξ as a result of the tests will acquire a value greater than x_1.

133. The density of the distribution of the random variable ξ is given by the formula

$$f_\xi(x) = \begin{cases} 0, & x \le 1, \\ x - 0,5, & 1 < x \le 2, \\ 0, & x > 2. \end{cases}$$

Construct the distribution function $F_\xi(x)$ and its graph.

134. The random variable ξ has the distribution density

$$f_\xi(x) = \begin{cases} 0, & x \le 0, \\ 0,5\sin x, & 0 < x \le \pi, \\ 0, & x > \pi. \end{cases}$$

a) Construct the distribution function $F_\xi(x)$.

b) Find the probability that, as a result of the test, the value ξ will take a value from the interval $\left(0;\ \dfrac{\pi}{4}\right)$.

135. The density of the distribution of the random variable ξ is given as follows:

$$f_\xi(x)=\begin{cases}3x^2, & 0<x\le 1,\\ 0, & x\le 0,\ x>1.\end{cases}$$

a) Draw a graph of the distribution function $F_\xi(x)$.

b) Find the median and mode of the value ξ.

136. The density of the distribution of the random variable ξ is given as follows:

$$f_\xi(x)=\begin{cases}ax, & 0<x\le 2,\\ 0, & x\le 0,\ x>2.\end{cases}$$

a) Find the coefficient a and the distribution function $F_\xi(x)$.

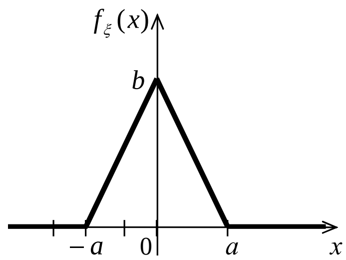
b) Find the probability $P\left\{\left|\xi-\frac{4}{3}\right|<0,5\right\}$.

137. A continuous random variable has a distribution function:

$$F_\xi(x)=\begin{cases}0, & x\le 0,\\ ax^2, & 0<x\le 1,\\ 1, & x>1.\end{cases}$$

Find the coefficient a, $f_\xi(x)$, $P(-0,25<\xi<0,5)$. Draw the graphs of $F_\xi(x)$ and $f_\xi(x)$.

138. The random variable ξ obeys Simpson's law ("the law of the isosceles triangle") on the segment $[-a;\ a]$.



a) Write an expression for the distribution density $f_\xi(x)$.

b) Draw the graph of the distribution function $F_\xi(x)$.

c) Find the probability of a random value ξ falling into the interval $\left(-\frac{a}{2};\ a\right)$.

Topic 11. "Numerical characteristics of discrete random variables"

139. Find the mathematical expectation and variance of the I_A indicator of event *A*, the probability of which is *p*.

140. Find the mathematical expectation and standard deviation of a random variable given by such a distribution law

ξ	3	5	7	9
P	0,4	0,3	0,2	0,1

141. Three independent shots are fired at the target, the probability of hitting with each shot is 0,4. The random variable ξ is the number of hits. Determine the mathematical expectation, variance, standard deviation and asymmetry of a random variable ξ.

142. There are 3 non-standard parts in a batch of 10 parts. Two details are selected at random. Find the mathematical expectation of a discrete random variable ξ – the number of non-standard parts among the selected, the variance $D\xi$ and the standard deviation σ_ξ.

143. The mathematical expectation and the variance of the random variable ξ are equal to 2 and 10, respectively. Find the mathematical expectation and the variance of the random variable $2\xi + 5$.

144. The discrete random variable ξ takes three possible values: $\xi_1 = 4$ with a probability $p_1 = 0{,}5$; $\xi_2 = 6$ with probability $p_2 = 0{,}3$ and ξ_3 with probability p_3. Find the value of ξ_3 and probability p_3 if the mathematical expectation $E\xi = 8$.

145. Throw *n* dice. The random variable ξ is the sum of the points that will fall on all the dice. Determine the mathematical expectation $E\xi$, the variance $D\xi$ and the standard deviation σ_ξ.

146. Prove that the mathematical expectation of deviation $\xi - E\xi$ is zero.

147. The discrete random variable ξ has only two possible values x_1 and x_2, and $x_2 > x_1$. The probability that ξ takes the value x_1 is 0,6. Find the distribution law of the quantity ξ if $E\xi = 1,4$ and $D\xi = 0,24$.

148. Prove that the mathematical expectation of a discrete random variable lies between its smallest and largest possible values.

149. Two shooters shoot two shots at the target independently of each other. The probability of hitting with each shot for the first shooter is 0,3, and for the second – 0,4. Draw up the law of distribution of the total number of hits on the target, find the mathematical expectation and variance of this random variable.

150. The sniper starts shooting at the target before the first hit, having four cartridges. The probability of hitting the target with the first shot is 0,8, and with each subsequent shot it is halved. Draw up the law of distribution of the number of fired shots and find the mathematical expectation of this random variable.

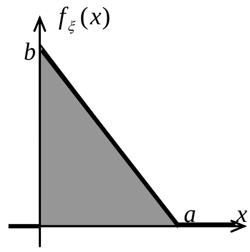
151. There are 8 standard parts in a batch of ten parts. Two parts are randomly selected at the same time. Draw up the law of distribution of the number of standard parts among the selected and find the mathematical expectation and variance of this random variable.

152. Among 10 pairs of shoes, 2 need repair. Draw up a law for the distribution of the number of pairs of shoes that need repair among three pairs taken at the same time. Find the mathematical expectation and variance of this random variable.

153. A student can take the exam no more than three times. Draw the law of the distribution of a random variable – the number of attempts to pass the exam, if the probability of passing it is 0,75 and continues to increase by 0,1 in each subsequent attempt. Find the mathematical expectation and variance of this random variable.

Topic 12. "Numerical characteristics of continuous random variables"

154. The random variable ξ is distributed according to the "right triangle law" on the interval $[0;\ a]$. Find: the mathematical expectation $E\xi$, the variance $D\xi$, the standard deviation sk and the third central moment μ_3 of this random variable.



155. The edge of the cube is measured approximately, and $a \le x \le b$. Considering the length of the edge of the cube as a random variable ξ distributed uniformly in the interval $[a;\ b]$. Find the mathematical expectation $E\xi^3$ of the volume of the cube.

156. (Exponential distribution.) The random variable ξ has the distribution density

$$f_\xi(t) = \begin{cases} 0, & t \le 0, \\ \lambda e^{-\lambda t}, & t > 0. \end{cases}$$

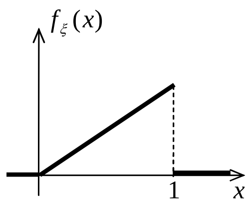
Plot the distribution function $F_\xi(t)$ and find $E\xi$ and $D\xi$.

157. The random variable ξ has the distribution density $f_\xi(x) = \frac{1}{2}\sin x$ in the interval $[0;\ \pi]$. Outside this interval $f_\xi(x) = 0$. Find the mathematical expectation of the random variable $\eta = \varphi(\xi) = \xi^2$ (without first finding the distribution density of the random variable η).

158. The random variables ξ and η are independent. Find the variance of the random variable $z = 3\xi + 2\eta$, if $D\xi = 5$, $D\eta = 6$.

159. The continuous random variable ξ obeys the distribution law with density $f_\xi(x) = A \cdot e^{-|x|}$. Determine the coefficient *A*, the mathematical expectation $E\xi$, the variance $D\xi$ and the standard deviation σ_ξ.

160. The random variable ξ has the following distribution law: Determine the mathematical expectation $E\xi$, the variance $D\xi$ and the standar deviation σ_ξ.



161. The distribution curve of the random variable ξ is the upper arc of an ellipse with semi-axes *a* and *b*. The quantity *a* is known. Determine the mathematical expectation $E\xi$, variance $D\xi$ and the standar deviation σ_ξ of this random variable.

162. The density of the distribution of the random variable ξ is given by the formula:

$$f_\xi(x) = \begin{cases} 0, & x \le 0, \\ ax, & 0 < x \le 2, \\ 0, & x > 2. \end{cases}$$

Find the constant *a*, the variance of the random variable ξ. Calculate the probability that the deviation of the random variable ξ from its mathematical expectation will be no more than 0.5.

163. The density of the distribution of the random variable ξ is given by the formula:

$$f_\xi(x) = \begin{cases} 0, & x \le 0, \\ 3x^2, & 0 < x \le 1, \\ 0, & x > 1. \end{cases}$$

Find the mode, median and mathematical expectation of the value ξ.

164. Prove that if ξ and η are independent random variables, then

$$D(\xi \cdot \eta) = D\xi \cdot D\eta + (E\xi)^2 D\eta + (E\eta)^2 D\xi .$$

165. The given density of the distribution of the random variable ξ:

$$f_\xi(x) = \begin{cases} 0, & x \le 0, \\ \frac{3}{2}x^2, & 0 < x \le 1, \\ \frac{3}{2}(2-x)^2, & 1 < x \le 2, \\ 0, & x > 2. \end{cases}$$

Find: a) initial and central moments of the first four orders; b) asymmetry and excess of this random variable.

166. The random variable ξ has an arcsine law with probability distribution density

$$f_\xi(x) = \begin{cases} \frac{1}{\pi\sqrt{a^2 - x^2}}, & |x| < a, \\ 0, & |x| > a. \end{cases}$$

Find: a) distribution function; b) $E\xi$ and $D\xi$; c) mode and median.

167. The random variable ξ is uniformly distributed. $E\xi = 4$, $D\xi = 3$. Find the density of the distribution of the random variable ξ.

168. Let ξ be a random variable uniformly distributed over the interval $[-1, 1]$. Find the distribution of the random variable $\eta = |\xi|$.

169. Let ξ be a random variable uniformly distributed over the interval $(0, 1)$. Find the distribution function of the random variable $\eta = \ln \frac{1}{\xi}$. Calculate $E\eta$.

170. The random variable ξ has the distribution density

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Find: a) $E\xi$ and $D\xi$; b) the probability that the random variable will take a value that belongs to the interval (α, β); c) the mode and median of the random variable ξ.

171. (*Rule of* 3σ). The random variable ξ – measurement error – is distributed according to the normal law. Find the probability that ks will take a value between -3σ and 3σ, where σ is the standard deviation of ξ (it is assumed that there are no systematic errors).

172. Prove that if the random variable ξ is distributed according to the normal law, then the linear function $\eta = A\xi + B$ $(A \neq 0)$ also has a normal distribution.

173. The random variable ξ has a normal distribution with mathematical expectation 0 and variance σ^2. Find $E\xi^k$, $k = 1, 2, ...$

Appendix

The value of the function $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$ Table 1

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	0,3970	0,3965	0,3961	0,3956	0,3951	0,3945	0,3939	0,3932	0,3925	0,3918
0,2	0,3910	0,3902	0,3894	0,3885	0,3876	0,3867	0,3857	0,3847	0,3836	0,3825
0,3	0,3814	0,3802	0,3790	0,3778	0,3765	0,3752	0,3739	0,3726	0,3712	0,3698
0,4	0,3683	0,3668	0,3652	0,3637	0,3621	0,3605	0,3589	0,3572	0,3555	0,3538
0,5	0,3521	0,3503	0,3485	0,3467	0,3448	0,3429	0,3410	0,3391	0,3372	0,3352
0,6	0,3332	0,3312	0,3292	0,3271	0,3251	0,3230	0,3209	0,3187	0,3166	0,3144
0,7	0,3123	0,3101	0,3079	0,3056	0,3034	0,3011	0,2989	0,2966	0,2943	0,2920
0,8	0,2897	0,2874	0,2850	0,2827	0,2803	0,2780	0,2756	0,2732	0,2709	0,2685
0,9	0,2661	0,2637	0,2613	0,2589	0,2565	0,2541	0,2516	0,2492	0,2468	0,2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	0,2179	0,2155	0,2131	0,2107	0,2083	0,2059	0,2036	0,2012	0,1989	0,1965
1,2	0,1942	0,1919	0,1895	0,1872	0,1849	0,1826	0,1804	0,1781	0,1758	0,1736
1,3	0,1714	0,1691	0,1669	0,1647	0,1626	0,1604	0,1582	0,1561	0,1539	0,1518
1,4	0,1497	0,1476	0,1456	0,1435	0,1415	0,1394	0,1374	0,1354	0,1334	0,1315
1,5	0,1295	0,1276	0,1257	0,1238	0,1219	0,1200	0,1182	0,1163	0,1145	0,1127
1,6	0,1109	0,1092	0,1074	0,1057	0,1040	0,1023	0,1006	0,0989	0,0973	0,0957
1,7	0,0940	0,0925	0,0909	0,0893	0,0878	0,0863	0,0848	0,0833	0,0818	0,0804
1,8	0,0790	0,0775	0,0761	0,0748	0,0734	0,0721	0,0707	0,0694	0,0681	0,0669
1,9	0,0656	0,0644	0,0632	0,0620	0,0608	0,0596	0,0584	0,0573	0,0562	0,0551
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0,0440	0,0431	0,0422	0,0413	0,0404	0,0395	0,0387	0,0379	0,0371	0,0363
2,2	0,0353	0,0347	0,0339	0,0332	0,0325	0,0317	0,0310	0,0303	0,0297	0,0290
2,3	0,0283	0,0277	0,0270	0,0264	0,0258	0,0252	0,0246	0,0241	0,0235	0,0229
2,4	0,0224	0,0219	0,0213	0,0208	0,0203	0,0198	0,0194	0,0189	0,0184	0,0180
2,5	0,0175	0,0171	0,0167	0,0163	0,0158	0,0154	0,0151	0,0147	0,0143	0,0139
2,6	0,0136	0,0132	0,0129	0,0126	0,0122	0,0119	0,0116	0,0113	0,0110	0,0107
2,7	0,0104	0,0101	0,0099	0,0096	0,0093	0,0091	0,0088	0,0086	0,0084	0,0081
2,8	0,0079	0,0077	0,0075	0,0073	0,0071	0,0069	0,0067	0,0065	0,0063	0,0061
2,9	0,0060	0,0058	0,0056	0,0055	0,0053	0,0051	0,0050	0,0048	0,0047	0,0046
3,0	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026	0,0025	0,0025
3,2	0,0024	0,0023	0,0022	0,0022	0,0021	0,0020	0,0020	0,0019	0,0018	0,0018
3,3	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014	0,0013	0,0013
3,4	0,0012	0,0012	0,0012	0,0011	0,0011	0,0010	0,0010	0,0010	0,0009	0,0009
3,5	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007	0,0007	0,0007	0,0006
3,6	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005	0,0004
3,7	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003	0,0003	0,0003	0,0003
3,8	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002	0,0002
3,9	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0001

The value of the function $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$ Table 2

x	**Φ(*x*)**	***x***	**Φ(*x*)**	***x***	**Φ(*x*)**	***x***	**Φ(*x*)**	***x***	**Φ(*x*)**	***x***	**Φ(*x*)**
0,00	0,0000	**0,43**	0,1664	**0,86**	0,3051	**1,29**	0,4015	**1,72**	0,4573	**2,30**	0,4893
0,01	0,0040	**0,44**	0,1700	**0,87**	0,3078	**1,30**	0,4032	**1,73**	0,4582	**2,32**	0,4898
0,02	0,0080	**0,45**	0,1736	**0,88**	0,3106	**1,31**	0,4049	**1,74**	0,4591	**2,34**	0,4904
0,03	0,0120	**0,46**	0,1772	**0,89**	0,3133	**1,32**	0,4066	**1,75**	0,4599	**2,36**	0,4909
0,04	0,0160	**0,47**	0,1808	**0,90**	0,3159	**1,33**	0,4082	**1,76**	0,4608	**2,38**	0,4913
0,05	0,0199	**0,48**	0,1844	**0,91**	0,3186	**1,34**	0,4099	**1,77**	0,4616	**2,40**	0,4918
0,06	0,0239	**0,49**	0,1879	**0,92**	0,3212	**1,35**	0,4115	**1,78**	0,4625	**2,42**	0,4922
0,07	0,0279	**0,50**	0,1915	**0,93**	0,3238	**1,36**	0,4131	**1,79**	0,4633	**2,44**	0,4927
0,08	0,0319	**0,51**	0,1950	**0,94**	0,3264	**1,37**	0,4147	**1,80**	0,4641	**2,46**	0,4931
0,09	0,0359	**0,52**	0,1985	**0,95**	0,3289	**1,38**	0,4162	**1,81**	0,4649	**2,48**	0,4934
0,10	0,0398	**0,53**	0,2019	**0,96**	0,3315	**1,39**	0,4177	**1,82**	0,4656	**2,50**	0,4938
0,11	0,0438	**0,54**	0,2054	**0,97**	0,3340	**1,40**	0,4192	**1,83**	0,4664	**2,52**	0,4941
0,12	0,0478	**0,55**	0,2088	**0,98**	0,3365	**1,41**	0,4207	**1,84**	0,4671	**2,54**	0,4945
0,13	0,0517	**0,56**	0,2123	**0,99**	0,3389	**1,42**	0,4222	**1,85**	0,4678	**2,56**	0,4948
0,14	0,0557	**0,57**	0,2157	**1,00**	0,3413	**1,43**	0,4236	**1,86**	0,4686	**2,58**	0,4951
0,15	0,0596	**0,58**	0,2190	**1,01**	0,3438	**1,44**	0,4251	**1,87**	0,4693	**2,60**	0,4953
0,16	0,0636	**0,59**	0,2224	**1,02**	0,3461	**1,45**	0,4265	**1,88**	0,4699	**2,62**	0,4956
0,17	0,0675	**0,60**	0,2257	**1,03**	0,3485	**1,46**	0,4279	**1,89**	0,4706	**2,64**	0,4959
0,18	0,0714	**0,61**	0,2291	**1,04**	0,3508	**1,47**	0,4292	**1,90**	0,4713	**2,66**	0,4961
0,19	0,0753	**0,62**	0,2324	**1,05**	0,3531	**1,48**	0,4306	**1,91**	0,4719	**2,68**	0,4963
0,20	0,0793	**0,63**	0,2357	**1,06**	0,3554	**1,49**	0,4319	**1,92**	0,4726	**2,70**	0,4965
0,21	0,0832	**0,64**	0,2389	**1,07**	0,3577	**1,50**	0,4332	**1,93**	0,4732	**2,72**	0,4967
0,22	0,0871	**0,65**	0,2422	**1,08**	0,3599	**1,51**	0,4345	**1,94**	0,4738	**2,74**	0,4969
0,23	0,0910	**0,66**	0,2454	**1,09**	0,3621	**1,52**	0,4357	**1,95**	0,4744	**2,76**	0,4971
0,24	0,0948	**0,67**	0,2486	**1,10**	0,3643	**1,53**	0,4370	**1,96**	0,4750	**2,78**	0,4973
0,25	0,0987	**0,68**	0,2517	**1,11**	0,3665	**1,54**	0,4382	**1,97**	0,4756	**2,80**	0,4974
0,26	0,1026	**0,69**	0,2549	**1,12**	0,3686	**1,55**	0,4394	**1,98**	0,4761	**2,82**	0,4976
0,27	0,1064	**0,70**	0,2580	**1,13**	0,3708	**1,56**	0,4406	**1,99**	0,4767	**2,84**	0,4977
0,28	0,1103	**0,71**	0,2611	**1,14**	0,3729	**1,57**	0,4418	**2,00**	0,4772	**2,86**	0,4979
0,29	0,1141	**0,72**	0,2642	**1,15**	0,3749	**1,58**	0,4429	**2,02**	0,4783	**2,88**	0,4980
0,30	0,1179	**0,73**	0,2673	**1,16**	0,3770	**1,59**	0,4441	**2,04**	0,4793	**2,90**	0,4981
0,31	0,1217	**0,74**	0,2703	**1,17**	0,3790	**1,60**	0,4452	**2,06**	0,4803	**2,92**	0,4982
0,32	0,1255	**0,75**	0,2734	**1,18**	0,3810	**1,61**	0,4463	**2,08**	0,4812	**2,94**	0,4984
0,33	0,1293	**0,76**	0,2764	**1,19**	0,3830	**1,62**	0,4474	**2,10**	0,4821	**2,96**	0,4985
0,34	0,1331	**0,77**	0,2794	**1,20**	0,3849	**1,63**	0,4484	**2,12**	0,4830	**2,98**	0,4986
0,35	0,1368	**0,78**	0,2823	**1,21**	0,3869	**1,64**	0,4495	**2,14**	0,4838	**3,00**	0,49865
0,36	0,1406	**0,79**	0,2852	**1,22**	0,3883	**1,65**	0,4505	**2,16**	0,4846	**3,20**	0,49931
0,37	0,1443	**0,80**	0,2881	**1,23**	0,3907	**1,66**	0,4515	**2,18**	0,4854	**3,40**	0,49966
0,38	0,1480	**0,81**	0,2910	**1,24**	0,3925	**1,67**	0,4525	**2,20**	0,4861	**3,60**	0,499841
0,39	0,1517	**0,82**	0,2939	**1,25**	0,3944	**1,68**	0,4535	**2,22**	0,4868	**3,80**	0,499928
0,40	0,1554	**0,83**	0,2967	**1,26**	0,3962	**1,69**	0,4545	**2,24**	0,4875	**4,00**	0,499968
0,41	0,1591	**0,84**	0,2995	**1,27**	0,3980	**1,70**	0,4554	**2,26**	0,4881	**4,50**	0,499997
0,42	0,1628	**0,85**	0,3023	**1,28**	0,3997	**1,71**	0,4564	**2,28**	0,4887	**5,00**	0,499997

Table 3

Probability $P_\lambda(\xi = m) = \frac{\lambda^m}{m!} e^{-\lambda}$ of Poisson distribution

m \ λ	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0	0,90484	0,81873	0,74082	0,67032	0,60653	0,54881	0,49659	0,44933	0,40657
1	0,09048	0,16375	0,22225	0,26813	0,30327	0,32929	0,34761	0,35946	0,36591
2	0,00452	0,01637	0,03334	0,05363	0,07582	0,09879	0,12166	0,14379	0,16466
3	0,00015	0,00109	0,00333	0,00715	0,01264	0,01976	0,02839	0,03834	0,04940
4		0,00005	0,00025	0,00072	0,00158	0,00296	0,00497	0,00767	0,01111
5			0,00002	0,00006	0,00016	0,00036	0,00070	0,00123	0,00200
6					0,00001	0,00004	0,00008	0,00016	0,00030
7								0,00002	0,00004

m \ λ	1	2	3	4	5	6	7	8	9	10
0	0,36788	0,13534	0,04979	0,01832	0,00674	0,00248	0,00091	0,00034	0,00012	0,00005
1	0,36788	0,27067	0,14936	0,07326	0,03369	0,01487	0,00638	0,00268	0,00111	0,00045
2	0,18394	0,27067	0,22404	0,14653	0,08422	0,04462	0,02234	0,01073	0,00500	0,00227
3	0,06131	0,18045	0,22404	0,19537	0,14037	0,08924	0,05213	0,02863	0,01499	0,00757
4	0,01533	0,09022	0,16803	0,19537	0,17547	0,13385	0,09123	0,05725	0,03374	0,01892
5	0,00307	0,03609	0,10082	0,15629	0,17547	0,16062	0,12772	0,09160	0,06073	0,03783
6	0,00051	0,01203	0,05041	0,10420	0,14622	0,16062	0,14900	0,12214	0,09109	0,06306
7	0,00007	0,00344	0,02160	0,05954	0,10444	0,13768	0,14900	0,13959	0,11712	0,09008
8		0,00086	0,00810	0,2977	0,06528	0,10326	0,13038	0,13959	0,13176	0,11260
9		0,00019	0,00270	0,01323	0,03627	0,06884	0,10140	0,12408	0,13176	0,12511
10		0,00004	0,00081	0,00529	0,01813	0,04130	0,07098	0,09926	0,11858	0,12511
11			0,00022	0,00192	0,00824	0,02253	0,04517	0,07219	0,09702	0,11374
12			0,00006	0,00064	0,00343	0,01126	0,02635	0,04813	0,07277	0,09478
13				0,00020	0,00132	0,00520	0,01419	0,02962	0,05038	0,07291
14				0,00006	0,00047	0,00223	0,00709	0,01692	0,03238	0,05208
15				0,00002	0,00016	0,00089	0,00331	0,00903	0,01943	0,3472
16					0,00005	0,00033	0,00145	0,00451	0,01093	0,2170
17					0,00001	0,00012	0,00060	0,00212	0,00579	0,01276
18						0,00004	0,00023	0,00094	0,00289	0,00709
19						0,00001	0,00009	0,00040	0,00137	0,00373
20							0,00003	0,00016	0,00062	0,00187
21								0,00006	0,00026	0,00089
22								0,00002	0,00011	0,00040
23									0,00004	0,00018
24									0,00002	0,00007
25										0,00003
26										0,00001

CONTENTS

Навчальне видання

Василь Йосипович **Кушнірчук**

ТЕОРІЯ ЙМОВІРНОСТЕЙ

Навчальний посібник

(англійською мовою)

Відповідальний за випуск ***Черевко І.М.***

Електронне видання

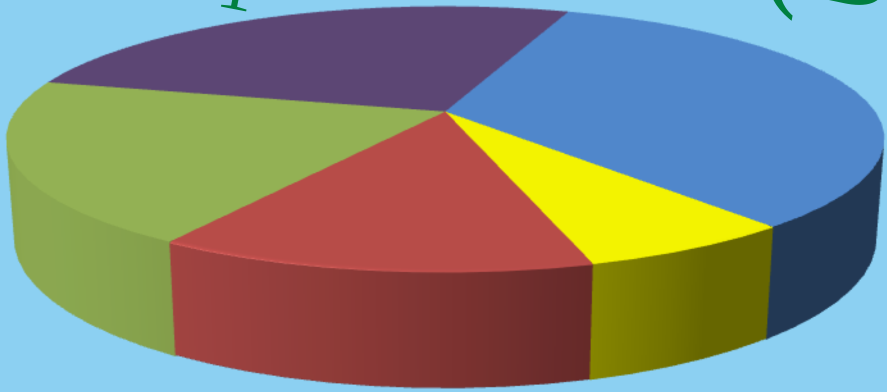
Підписано до друку 26.06.2024.
Умов.-друк. арк. 6,82. Обл.-вид. арк. 6,34.
Зам. Н-061.
Видавництво Чернівецького національного університету.
58002, Чернівці, вул. Коцюбинського, 2.
e-mail: ruta@chnu.edu.ua

Свідоцтво суб’єкта видавничої справи ДК № 891 від 08.04.2002.

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

$$P(A) = \sum_{i:\omega_i \in A} p_i$$

$$P(B) = \sum_{i=1}^{n} P(H_i) \cdot P(B \mid H_i)$$



$$D\xi = E(\xi - E\xi)^2$$

$$E\xi = \int_{-\infty}^{+\infty} x\,dF_\xi(x)$$