G. Pantsulaia

Elements of Probability Theory

The present manuscript is English version of the textbook

გ.ფანცულაია, ალბათობის თეორიის ელემენტები,

“ტექნიკური უნივერსიტეტი”, 2005 **(ISBN-99940-48-09-0)**

Introduction

The modern probability theory is an interesting and most important part of mathematics, which has great achievements and close connections both with classical parts of mathematics (geometry, mathematical analysis, functional analysis), and its various branches(theory of random processes, theory of ergodicity, theory of dynamical system, mathematical statistics and so on). The development of these branches of mathematics is mainly connected with the problems of statistical mechanics, statistical physics, statistical radio engineering and also with the problems of complicated systems which consider the random and the chaotic influence.

At the origin of the probability theory were standing such famous mathematicians as I.Bernoulli, P.Laplace, S.Poisson, A.Cauchy, G.Cantor, F.Borel, A.Lebesgue and others. A very controversial problem connected with the relation between the probability theory and mathematics was entered in the list of unsolved mathematical problems raised by D.Gilbert in 1900. This problem has been solved by Russian mathematician A.Kolmogorov in 1933 who gave us a strict axiomatic basis of the probability theory.

A.Kolmogorov conception to the basis of the probability theory is applied in the present book. Giving a strong system of axioms (according to A.Kolmogorov) the general probability spaces and their *có*mposite comp*ó*nents are described in the present book.

The main purpose of the present book is to help students to acquire such skills that are necessary to construct mathematical models (i.e., probability spaces) of various (social, economical, biological, mechanical, physical, etc) processes and to calculate their numerical characteristics. In this sense the last chapters (in particular, chapters 14-15) are of interest, where some applications of various mathematical models(Markov chains, Brownian motion, etc) are presented. The present book consists of fifteen chapters. Each chapter is equipped with exercises (i.e. tests), the solutions of which will help the student in deep comprehend and assimilation of experience of the presented elements of the probability theory.

Chapter 1

Set-Theoretical Operations. Kolmogorov Axioms

Let Ω be a non-empty set and let $\mathcal{P}(\Omega)$ be a class of all subsets of Ω ($\mathcal{P}(\Omega)$ is called a powerset of Ω).

Definition 1.1 An union of the finite family $(A_k)_{1 \le k \le n}$ of elements of $\mathcal{P}(\Omega)$ is denoted by the symbol $\cup_{k=1}^{n} A_k$ and is defined by

$$\cup_{k=1}^{n} A_k = \{x | x \in A_1 \bigvee \cdots \bigvee x \in A_n\},$$

where $\bigvee$ denotes the logical symbol of disjunction.

Definition 1.2 An intersection of an infinite family $(A_k)_{k \in N}$ of elements of $\mathcal{P}(\Omega)$ is denoted by the symbol $\cup_{k \in N} A_k$ and is defined by

$$\cup_{k \in N} A_k = \{x | x \in A_1 \bigvee x \in A_2 \bigvee \cdots \}.$$

Definition 1.3 An intersection of a finite family $(A_k)_{1 \le k \le n}$ of elements of $\mathcal{P}(\Omega)$ is denoted by the symbol $\cap_{k=1}^{n} A_k$ and is defined by

$$\cap_{k=1}^{n} A_k = \{x | x \in A_1 \bigwedge \cdots \bigwedge x \in A_n\},$$

where $\bigwedge$ denotes the logical symbol of conjunction.

Definition 1.4 An intersection of an infinite family $(A_k)_{k \in N}$ of elements of $\mathcal{P}(\Omega)$ is denoted by the symbol $\cap_{k \in N} A_k$ and is defined by

$$\cap_{k \in N} A_k = \{x | x \in A_1 \bigwedge x \in A_2 \bigwedge \cdots \}.$$

Definition 1.5 Let $A, B \in \mathcal{P}(\Omega)$. A difference of sets A and B is defined by the symbol $A \setminus B$ and is denoted by

$$A \setminus B = \{x | x \in A \ \bigwedge x \notin B\}.$$

Remark 1.1 The following formulas
1) $\Omega \setminus \cup_{k=1}^{n} A_k = \cap_{k=1}^{n}(\Omega \setminus A_k)$;
2) $\Omega \setminus \cup_{k\in N} A_k = \cap_{k\in N}(\Omega \setminus A_k)$;
3) $\Omega \setminus \cap_{k=1}^{n} A_k = \cup_{k=1}^{n}(\Omega \setminus A_k)$;
4) $\Omega \setminus \cap_{k\in N} A_k = \cup_{k\in N}(\Omega \setminus A_k)$.
are valid.

Definition 1.6 A class $\mathcal{A}$ of subsets of Ω is called algebra, if the following conditions
1) $\Omega \in \mathcal{A}$,
2) if $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$ and $A \cap B \in \mathcal{A}$,
3) if $A \in \mathcal{A}$, then $\Omega \setminus A \in \Omega$,
hold.

Remark 1.2 The algebra is such a class of subsets Ω which contains Ω and is closed under the set-theoretical operations " $\cap, \cup, \setminus$ ".

Definition 1.7 A class $\mathcal{F}$ of subsets of Ω is called σ-algebra, if the following conditions
1) $\Omega \in \mathcal{F}$,
2) if $A_k \in \mathcal{F}$ $(k \in N)$, then $\cup_{k\in N} A_k \in \mathcal{F}$ and $\cap_{k\in N} A_k \in \mathcal{F}$,
3) if $A \in \mathcal{F}$, then $\Omega \setminus A \in \Omega$,
hold.

Remark 1.3 σ-algebra is such a class of subsets of Ω which is closed under infinite number of operations " $\cap, \cup, \setminus$ ".

Definition 1.8 A real-valued function P defined on the σ-algebra of subsets of Ω is called a probability measure(or probability), if the following conditions are fulfilled:
1) $P(A) \geq 0$ for $A \in \mathcal{F}$ (The property of non-negativity of the probability);
2) $P(\Omega) = 1$ (the property of normalization);
3) if $(A_k)_{k\in N}$ is a sequence of pairwise disjoint elements of $\mathcal{F}$, then $P(\cup_{k\in N} A_k) = \sum_{k\in N} P(A_k)$ (the proporty of countable-additivity).

Kolmogorov[1]axioms. The triplet $(\Omega, \mathcal{F}, P)$, where
1) Ω is a non-empty set,
2) $\mathcal{F}$ is a σ-algebra of subsets of Ω,
3) P is a probability defined on $\mathcal{F}$,
is called a probability space.

Ω is called a space of all elementary events; An arbitrary point $\omega \in \Omega$ is called elementary event; An arbitrary element of $\mathcal{F}$ is called an event; $\emptyset$ is called an impossible event; Ω is called a necessary event; For arbitrary event A an event $\overline{A} = \Omega \setminus A$ is called its complementary event ; The product of events A and

[1]Andrey Kolmogorov [12(25).4.1903 Tambov-25.10.1987 Moscow] Russian mathematician, Academician of the Academy Sciences of the USSR (1939), Professor of the Moscow State University. He has firstly considered a mathematical conception of the axiomatical foundation of the probability theory in 1933.

B is denoted by AB and is defined by $A \cap B$; The events A and B are called non-consistent if the event AB is an impossible event; A sum of two non-consistent events A and B is denoted by $A + B$ and is defined by $A \cup B$; For arbitrary event A the number $P(A)$ is called a probability of the event A .

Definition 1.9. A sum of pairwice disjoint events $(A_k)_{k \in N}$ is denoted by the symbol $\sum_{k \in N} A_k$ and is defined by

$$\sum_{k \in N} A_k = \cup_{k \in N} A_k.$$

Remark 1.4. Like the numerical operations of sums and product, the set-theoretical operations have the following properties:

1) $A + B = B + A, \ AB = BA$;
2) $(A + B) + C = A + (B + C)$,
3) $(AB)C = A(BC)$
4) $(A + B)C = AC + BC, \ C(A + B) = CA + CB$,
5) $C(\sum_{k \in N} A_k) = \sum_{k \in N} CA_k$,
6) $(\sum_{k \in N} A_k)C = \sum_{k \in N} A_k C$

Tests

1.1. Assume that $A_k = [\frac{k+1}{k+2}, 1] \ (k \in N)$. Then

1) the set $\cap_{4 \le k \le 10} A_k$ coincides with
a) $[\frac{1}{2}, 1]$, b) $]\frac{11}{12}, 1]$, c) $[\frac{11}{12}, 1]$, d) $[\frac{1}{2}, 1]$;

2) the set $\cup_{3 \le k \le 10} A_k$ coincides with
a) $[\frac{4}{5}, 1]$, b) $[\frac{3}{4}, 1]$, c) $[\frac{2}{3}, 1]$, d) $[\frac{5}{6}, 1]$;

3) the set $\cup_{2 \le k \le 10} A_k \setminus \cap_{1 \le k \le 10} A_k$ coincides with
a) $[\frac{3}{4}, \frac{11}{12}[$, b) $[\frac{4}{5}, \frac{12}{13}[$, c) $[\frac{4}{5}, 1[$, d) $[\frac{5}{6}, 1[$;

4) the set $\cap_{k \in N} A_k$ coincides with
a) $\{1\}$, b) $\{0\}$, c) $\{\emptyset\}$, d) $[0, 1]$;

5) the set $\cup_{k \in N} A_k$ coincides with
a) $[\frac{4}{5}, 1]$, b) $[\frac{3}{4}, 1]$, c) $[\frac{2}{3}, 1]$, d) $[\frac{5}{6}, 1]$;

6) the set $\cup_{k \in N} A_k \setminus \cap_{k \in N} A_k$ coincides with
a) $[\frac{3}{4}, 1[$, b) $[\frac{2}{3}, 1[$, c) $[\frac{4}{5}, 1[$, d) $[\frac{5}{6}, 1[$.

1.2. Assume that $A_k = [\frac{k-3}{3k}, \frac{2k+3}{3k}] \ (k \in N)$. Then

1) the set $\cap_{5 \le k \le 10} A_k$ coincides with
a) $[\frac{8}{33}, \frac{25}{33}]$, b) $[\frac{7}{30}, \frac{23}{30}]$, c) $[\frac{2}{15}, \frac{13}{15}]$, d) $[\frac{1}{12}, \frac{11}{12}]$;

2) the set $\cup_{10 \le k \le 20} A_k$ coincides with
a) $[\frac{8}{33}, \frac{25}{33}]$, b) $[\frac{7}{30}, \frac{23}{30}]$, c) $[\frac{2}{15}, \frac{13}{15}]$, d) $[\frac{1}{12}, \frac{11}{12}]$;

3) the set $\cap_{k \in N} A_k$ coincides with
a) $[\frac{8}{33}, \frac{25}{33}]$, b) $[\frac{1}{3}, \frac{3}{4}]$, c) $[\frac{1}{3}, \frac{2}{3}]$, d) $[\frac{1}{12}, \frac{11}{12}]$;

4) the set $[0, 1] \setminus \cap_{k \in N} A_k$ coincides with
a) $[0, 1] \setminus [0, \frac{1}{3}] \cup]\frac{3}{4}; 1[$, b) $[0, \frac{1}{3}[\cup]\frac{2}{3}; 1]$, c) $[\frac{1}{3}, \frac{2}{3}]$, d) $[\frac{1}{12}, \frac{11}{12}]$.

1.3*. Let θ be a positive number such that $\frac{\theta}{\pi}$ is an irrational number. We set

$$\Delta = \{(x, y) | -1 < x < 1, \ -1 < y < 1\}.$$

Let denote by A_n a set obtained by counterclockwise rotation of the set Δ about the origin of the plane for the angle $n\theta$. Then

1) the set $\cap_{k\in N} A_k$ coincides with

a) $\{(x,y)|x^2+y^2 \leq 1\}$, b) $\{(x,y)|x^2+y^2 \leq 2\}$,
c) $\{(x,y)|x^2+y^2 < 1\}$, d) $\{(x,y)|x^2+y^2 < 2\}$;

2) the set $\cup_{k\in N} A_k$ coincides with

a) $\{(x,y)|x^2+y^2 \leq 1\}$, b) $\{(x,y)|x^2+y^2 \leq 2\}$,
c) $\{(x,y)|x^2+y^2 < 1\}$, d) $\{(x,y)|x^2+y^2 < 2\}$.

1.4. Assume that $\Omega = \{0;1\}$. Then

1) the algebra of subsets of Ω is

a) $\{\{0\},\{0;1\}\}$, b) $\{\{0\};\{0;1\};\emptyset\}$,
c) $\{\{0\};\{1\};\{0;1\};\emptyset\}$; d) $\{\{1\},\{0;1\}\}$;

2) the σ-algebra of subsets of Ω is

a) $\{\{0\},\{0;1\}\}$, b) $\{\{0\};\{0;1\};\emptyset\}$,
c) $\{\{0\};\{1\};\{0;1\};\emptyset\}$; d) $\{\{1\},\{0;1\}\}$.

1.5. Assume that $\Omega = [0,1[$. Then

1) the algebra of subsets of Ω is

a) $\{X|X \subset [0,1[$, X is presented by the finite union of intervals open from the right and closed from the right. $\}$,

b) $\{X|X \subset [0,1[$, X is presented by the finite union of intervals closed from the right and open from the right. $\}$,

c) $\{X|X \subset [0,1[$, X is presented by the finite union of closed from both side intervals.$\}$,

d) $\{X|X \subset [0,1[$, X is presented by the finite union of open from both side intervals. $\}$;

2) Assume that $\mathcal{A} = \{X|X \subset [0,1[$ and X be presented as the finite union of intervals open from the right and closed from the left $\}$.

Then $\mathcal{A}$

a) is not the algebra, b) is the σ-algebra, c) is the σ-algebra, but is not the algebra, d) is the algebra, but is not the σ-algebra.

Chapter 2

Properties of Probabilities

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Then the probability P has the following properties.

Property 2.1 $P(\emptyset) = 0$.

Proof. We have $\emptyset = \emptyset \cup \emptyset \cup \cdots$. From the property of countably-additivity of the probability P, we have

$$P(\emptyset) = \sum_{i=1}^{\infty} P(\emptyset) = \lim_{n\to\infty} nP(\emptyset).$$

Since P is finite, $P(\emptyset) \in R$. Hence, above-mentioned equality is possible if and only if $P(\emptyset) = 0$.

Property 2.2 (The property of the finite-additivity). *If $(A_k)_{1\leq k\leq n}$ is a finite family of pairwise disjoint events, then*

$$P(\cup_{k=1}^{n} A_k) = \sum_{k=1}^{n} P(A_k).$$

Proof. For arbitrary natural number $k > n$ we set $A_k = \emptyset$. Following Property 1 and the property of the countable-additivity of P we have

$$P(\cup_{k=1}^{n} A_k) = P(\cup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k) = \sum_{k=1}^{n} P(A_k) + \sum_{k=n+1}^{\infty} P(A_k) = \sum_{k=1}^{n} P(A_k).$$

Property 2.3. *For $A \in \mathcal{F}$ we have*

$$P(\overline{A}) = 1 - P(A).$$

Proof. Since $\Omega = A + \overline{A}$ and $P(\Omega) = 1$, using the property of the finitely-additivity, we have

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

It follows that

$$P(\overline{A}) = 1 - P(A).$$

Property 2.4 *Assume that $A, B \in \mathcal{F}$ and $A \subseteq B$. Then $P(B) = P(A) + P(B \setminus A)$.*

Proof. Using the equality $B = A + (B \setminus A)$ and the property of countably-additivity of P, we have $P(B) = P(A) + P(B \setminus A)$.

Property 2.5 *Assume that $A, B \in \mathcal{F}$ and $A \subseteq B$, then $P(A) \leq P(B)$.*

Proof. Following Property 4, we have $P(B) = P(A) + P(B \setminus A)$. Hence it follows $P(A) = P(B) - P(B \setminus A) \leq P(B)$.

Property 2.6. *Assume that $A, B \in \mathcal{F}$. Then*

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

Proof. Using the representation $A \cup B = (A \setminus B) + AB + (B \setminus A)$ and the property of finitely-additivity of P, we have

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

Property 2.7 *Assume that $A, B \in \mathcal{F}$, then*

$$P(A \cup B) \leq P(A) + P(B).$$

Proof. Following Property 6, we have

$$P(A \cup B) = P(A) + P(B) - P(AB);$$

It yields

$$P(A \cup B) = P(A) + P(B) - P(AB) \leq P(A) + P(B).$$

Property 2.8 (Continuity from above). *Assume that $(A_n)_{n \in N}$ be an decreasing sequence of events, i.e.,*

$$(\forall n)(n \in N \rightarrow A_{n+1} \subseteq A_n).$$

Then the following is valid

$$P(\cap_{n \in N} A_n) = \lim_{n \to \infty} P(A_n).$$

Proof. For $n \in N$ we have

$$A_n = \cap_{k \in N} A_k + (A_n \setminus A_{n+1}) + (A_{n+1} \setminus A_{n+2}) + \cdots .$$

Using the property of the countably-additivity of P, we obtain :

$$P(A_n) - P(\cap_{k \in N} A_k) = \sum_{p=1}^{\infty} P(A_{n+p} \setminus A_{n+p+1}).$$

Note that the sum $\sum_{p=1}^{\infty} P(A_{n+p} \setminus A_{n+p+1})$ is the n-th residual series of absolutely convergent series $\sum_{n=1}^{\infty} P(A_n \setminus A_{n+1})$. From the necessary and sufficient condition of the convergence of the numerical series, we have

$$\lim_{n\to\infty} \sum_{p=1}^{\infty} P(A_{n+p} \setminus A_{n+p+1}) = 0.$$

It means that

$$\lim_{n\to\infty}(P(A_n) - P(\cap_{k\in N} A_k)) = \lim_{n\to\infty} P(A_n) - P(\cap_{k\in N} A_k) = 0,$$

i.e.,

$$\lim_{n\to\infty} P(A_n) = P(\cap_{k\in N} A_k).$$

Property 2.9 (Continuity from below). *Let $(B_n)_{n\in N}$ be an increasing sequence of events, i.e.,*

$$(\forall n)(n \in N \to B_n \subseteq B_{n+1}).$$

Then the following equality is valid

$$P(\cup_{n\in N} B_n) = \lim_{n\to\infty} P(B_n).$$

Proof. For $\cup_{n\in N} B_n$ we have the following representation

$$\cup_{n\in N} B_n = B_1 + (B_2 \setminus B_1) + \cdots + (B_{k+1} \setminus B_k) + \cdots .$$

Following the property of the countable-additivity of P, we get

$$P(\cup_{n\in N} B_n) = P(B_1) + P(B_2 \setminus B_1) + \cdots + P(B_{k+1} \setminus B_k) + \cdots .$$

From Property 4 we have

$$P(B_{k+1}) = P(B_k) + P(B_{k+1} \setminus B_k).$$

If we define $P(B_{k+1} \setminus B_k)$ from the above-mentioned equality and enter it in early considered equality we obtain

$$P(\cup_{n\in N} B_n) = P(B_1) + (P(B_2) - P(B_1)) + \cdots + (P(B_{k+1}) - P(B_k)) + \cdots .$$

Note that the series on the right is convergent. For the sum S_n of the first n members we have

$$S_n = P(B_n).$$

From the definition of the series sum, we obtain

$$P(\cup_{n\in N} B_n) = \lim_{n\to\infty} S_n = \lim_{n\to\infty} P(B_n).$$

Tests

Assume that $(\Omega, \mathcal{F}, P)$ be a probability space.

2.1. If $P(A) = 0,95$, then $P(\overline{A})$ is equal to

a) $0,56$, b) $0,55$, c) $0,05$, d) $0,03$.

2.2. Assume that $A, B \in \mathcal{F}$, $A \subset B$, $P(A) = 0,65$ and $P(B) = 0,68$. Then $P(B \setminus A)$ is equal to

a) $0,02$, b) $0,03$, c) $0,04$, d) $0,05$.

2.3. Assume that $A, B \in \mathcal{F}$, $P(A) = 0,35$, $P(B) = 0,45$ and $P(A \cup B) = 0,75$. Then $P(A \cap B)$ is equal to

a) $0,02$, b) $0,03$, c) $0,04$, d) $0,05$.

2.4. Let $(A_n)_{n \in N}$ be a decreasing sequence of events and $P(\cap_{n \in N} A_n) = 0,89$. Then $\lim_{n \to \infty} P(\overline{A_n})$ is equal to

a) $0,11$, b) $0,12$, c) $0,13$, d) $0,14$.

2.5. Let $(A_n)_{n \in N}$ be a decreasing sequence of events and $P(A_n) = \frac{n+1}{3n}$. Then $P(\cap_{n \in N} A_n)$ is equal to

a) $\frac{1}{2}$, b) $\frac{1}{3}$, c) $\frac{1}{4}$, d) $\frac{1}{5}$.

2.6. Let $(A_n)_{n \in N}$ be an increasing sequence of events and

$$P(\cup_{n \in N} A_n) = 0,89.$$

Then $\lim_{n \to \infty} P(\overline{A_n})$ is equal to

a) $0,11$, b) $0,12$, c) $0,13$, d) $0,14$.

2.7. Let $(A_n)_{n \in N}$ be an increasing sequence of events and $P(A_n) = \frac{n-1}{3n}$. Then $P(\cup_{n \in N} A_n)$ is equal to

a) $\frac{1}{2}$, b) $\frac{1}{3}$, c) $\frac{1}{4}$, d) $\frac{1}{5}$.

Chapter 3

Examples of Probability Spaces

3.1. Classical probability space. Let Ω contains n points, i.e. $\Omega = \{\omega_1, \cdots, \omega_n\}$. We denote by $\mathcal{F}$ the class of all subsets of Ω . Let define a numerical number P by

$$(\forall A)(A \in \mathcal{F} \rightarrow P(A) = \frac{|A|}{|\Omega|}),$$

where $|\cdot|$ denotes a cardinality of the corresponding set.

One can easily demonstrate that the triplet $(\Omega, \mathcal{F}, P)$ is a probability space. This probability space is called a classical probability space. The numerical function P is called a classical probability.

Definition 3.1 Let A be any event. We say that an elementary event $\omega \in \Omega$ is successful for the event A if $\omega \in A$.

We obtain the following rule for calculation of the classical probability :

The classical probability of the event A is equal to the fraction a numerator of which is equal to the number of all successful elementary events for the event A and a denominator of which is equal to the number of all possible elementary events.

3.2 Geometric probability space. Let Ω be a Borel subset of the n-dimensional Euclidean space with positive Borel [1] measure b_n (cf. §5, Example 3). Let denote by $\mathcal{F}$ the class of all Borel subsets of Ω. We set

$$(\forall A)(A \in \mathcal{F} \rightarrow P(A) = \frac{b_n(A)}{b_n(\Omega)}).$$

The triplet $(\Omega, \mathcal{F}, P)$ is called an n-dimensional geometrical probability space associated with the Borel set Ω. The function P is called n-dimensional geometrical probability defined on Ω.

When a point is falling in the set Ω and the probability that this point will fall in any Borel subset of Ω is proportional to its Borel b_n-measure, then the geometrical probability of the falling of the point in the Borel subset $Y \subseteq \Omega$ is equal to the

[1] Borel Felix Eduard Justion Emil (7.01. 1871 –3.03.1956.)-French mathematician, member of the Paris Academy of Sciences (1921), professor of the Paris University (1909-1941).)

11

fraction the nominator of which is equal to $b_n(Y)$ and the denominator of which is equal to $b_n(\Omega)$.

Let us consider some examples demonstrating how we can model probability spaces describing random experiments.

Example 3.1

Experiment - *We roll a six-sided die.*

Problem. *What is the probability that we will roll an even number? .*

Modelling of the random experiment. Since the result of the random experiment is an elementary event, a space Ω of all elementary events has the following form :

$$\Omega = \{1; 2; 3; 4; 5; 6\}.$$

We denote by $\mathcal{F}$ a σ-algebra of all subsets of Ω(i.e. the powerset of Ω). It is clear, that

$$\mathcal{F} = \{\emptyset; \{1\}; \cdots \{6\}; \{1; 2\}; \cdots \{1; 2; 3; 4; 5; 6\}\}.$$

Let denote by P a classical probability measure defined by

$$(\forall A)(A \in \mathcal{F} \rightarrow P(A) = \frac{|A|}{6}).$$

The triplet $(\Omega, \mathcal{F}, P)$ is the probability space (i.e. the stochastic mathematical model) which describes our experiment.

Solution of the problem. We must calculate the probability of the event B having the following form :

$$B = \{2; 4; 6\}.$$

By definition of P, we have

$$P(B) = \frac{|B|}{6} = \frac{3}{6} = \frac{1}{2};$$

Resume. *The probability that we roll an even number is equal to* $\frac{1}{2}$.

Example 3.2

Experiment. - *We accidentally choose 3 cards from the complect of 36 cards.*

Problem. *What is the probability that in these 3 cards one will be "ace"?*

Modelling of the experiment. Since the result of the random experiment is an elementary event and it coincides with tree cards, the space Ω of all elementary events would be the space of all possible different tree cards. It is clear that

$$|\Omega| = C_{36}^3.$$

We denote by $\mathcal{F}$ a σ-algebra of all subsets of Ω. Let define a probability P by the following formula

$$(\forall A)(A \in \mathcal{F} \rightarrow P(A) = \frac{|A|}{C_{36}^3}).$$

Note, that $(\Omega, \mathcal{F}, P)$ is the probability space describing our experiment.

Solution of the problem. If we choose 1 card from the complect of "aces" and 2 cards from the other cards, then considering all their possible combinations,we obtain the set A of all threes of cards where at least one card is "ace". It is clear that number of A is equal to $C_4^1 \cdot C_{32}^2$.

By the definition of P we have

$$P(A) = \frac{|A|}{C_{36}^3} = \frac{C_4^1 \cdot C_{32}^2}{C_{36}^3}.$$

Resume. *If we choose accidentally 3 cards from the complect of 36 cards, then the probability that between them at list one card would be "ace" is equal* $\frac{C_4^1 \cdot C_{32}^2}{C_{36}^3}$

Example 3.3

Experiment. *There are passed parallel lines on the plane such that the distant between neighbouring lines is equal to $2a$. A $2l(2l < 2a)$ long needle is thrown accidentally on the plane.*

Problem *(Buffon)* [2] *What is the probability that the accidentally thrown on the plane needle intersects any of the above-mentioned parallel line?*

Modelling of the experiment. *The result of our experiment is an elementary event, which can be defined by x and φ, where x is the distance from the middle of the needle to the nearest line and φ is the angle between the needle and the above mentioned line. It is clear that x and φ satisfy the following conditions $0 \leq x \leq a, 0 \leq \varphi \leq \pi$.*

Hence, a space of all elementary events Ω is defined by

$$\Omega = [0; \pi] \times [0; a] = \{(\varphi; x) : 0 \leq \varphi \leq \pi,\ 0 \leq x \leq a\}.$$

We denote by $\mathcal{F}$ a class of all Borel subsets of Ω. Let define a probability P by the following formula:

$$(\forall A)(A \in \mathcal{F} \rightarrow P(A) = \frac{b_2(A)}{b_2(\Omega)}).$$

Evidently, $(\Omega, \mathcal{F}, P)$ is the probability space which describes our experiment.

Solution of the problem. *It is clear that to the event "the needle accidentally thrown on the plane intersects any above-mentioned parallel line" - corresponds a subset B_0, defined by :*

$$B_0 = \{(\varphi, x)|\ \ 0 \leq \varphi \leq \pi, \quad 0 \leq x \leq l \sin \varphi\}.$$

[2] Buffon Georges Louis Leclerc (7.9.1707 -16.4.1788) French experimentalist, member of the Petersburg's Academy of Sciences (1776), member of Paris Academy of Sciences (1733). The first mathematician, who worked on the problems of geometrical probabilities.)

By the definition of P we have

$$P(B_0) = \frac{b_2(B)}{a \cdot \pi} = \frac{\int_0^\pi l \sin \varphi d\varphi}{a \cdot \pi} = \frac{2l}{a\pi}.$$

Conclusion. *The probability that the needle accidentally thrown on the plane needle intersects any above-mentioned parallel line is equal to* $\frac{2l}{a\pi}$.

Tests

3.1. There are 5 white and 10 black balls in the box. The probability that the accidentally chosen ball would be black is equal to

a) $\frac{1}{3}$, b) $\frac{2}{3}$, c) $\frac{1}{5}$, d) $\frac{1}{6}$.

3.2. There are 7 white and 13 red balls in the box. The probability that between accidentally chosen 3 balls 2 balls would be black is equal to

a) $\frac{C_{13}^2 \cdot C_7^1}{C_{20}^3}$, b) $\frac{C_{13}^1 \cdot C_7^2}{C_{20}^3}$, c) $\frac{C_{13}^2 \cdot C_7^2}{C_{20}^3}$, d) $\frac{C_{13}^1 \cdot C_7^1}{C_{20}^3}$.

3.3. We roll two six-sided dice. The probability that the sum of dice numbers is less than 8, is equal to

a) $\frac{13}{18}$, b) $\frac{5}{6}$, c) $\frac{1}{5}$, d) $\frac{1}{6}$.

3.4. There are 17 students in the group. 8 of them are boys. There are staged 7 tickets to be drawn. The probability that between owners of tickets are 4 boys, is equal to

a) $\frac{C_{13}^2 \cdot C_7^2}{C_{15}^7}$, b)$\frac{C_8^1 \cdot C_7^2}{C_{17}^7}$, c) $\frac{C_8^4 \cdot C_9^3}{C_{17}^7}$, d) $\frac{C_{13}^1 \cdot C_7^1}{C_{25}^7}$.

3.5. A cube, each side of which is painted, is devided on 1000 equal cubes. The obtained cubes are mixed. The classical probability, that an accidentally chosen cube

1) has 3 painted sides, is equal to

a) $\frac{1}{1000}$, b) $\frac{1}{125}$, c) $\frac{1}{250}$, d) $\frac{1}{400}$;

2) has 3 painted sides, is equal to

a) $\frac{12}{124}$, b) $\frac{11}{120}$, c) $\frac{12}{125}$, d) $\frac{9}{125}$;

3) has 1 painted side, is equal to

a) $\frac{54}{250}$, b) $\frac{43}{145}$, c) $\frac{48}{125}$, d) $\frac{243}{250}$;

4) has no of painted side, is equal to

a) $\frac{8}{250}$, b) $\frac{64}{125}$, c) $\frac{4}{165}$, d) $\frac{23}{250}$.

3.6. A group of 10 girls and 10 boys is accidentally devided into two subgroups. The classical probability that in both subgroups the numbers of girls and boys will be equal, is

a) $\frac{(C_{10}^5)^2}{C_{20}^{10}}$, b) $\frac{C_{10}^5}{C_{20}^{10}}$, c) $\frac{(C_{10}^5)^3}{C_{20}^{10}}$, d) $\frac{C_{10}^5}{C_{20}^5}$.

3.7. We have 5 segments with lengths 1, 3, 4, 7 and 9. The classical probability that by accidently choosing 3 segments we can construct a triangle, is equal to

a) $\frac{3}{C_5^3}$, b) $\frac{2}{C_5^3}$, c) $\frac{4}{C_5^3}$, d) $\frac{5}{C_5^3}$.

3.8. When rolling two six-sided dice, the classical probability that

1) the sum of cast numbers is less than 5, is equal to

a) $\frac{7}{36}$, b) $\frac{5}{18}$, c) $\frac{1}{4}$, d) $\frac{3}{9}$;

2) we roll 5 by any die, is equal to

a) $\frac{7}{36}$, b) $\frac{8}{36}$, c) $\frac{11}{36}$, d) $\frac{3}{19}$;

3) we roll only one 5 , is equal to

a) $\frac{7}{36}$, b) $\frac{5}{18}$, c) $\frac{10}{36}$, d) $\frac{12}{19}$;

4) the sum of rolled numbers divides by 3, is equal to

a) $\frac{1}{3}$, b) $\frac{2}{5}$, c) $\frac{1}{6}$, d) $\frac{2}{9}$;

5) the module of the difference of rolled numbers is equal to 3, is

a) $\frac{1}{6}$, b) $\frac{2}{5}$, c) $\frac{2}{6}$, d) $\frac{2}{5}$;

6) the product of rolled numbers is simple, is equal to

a) $\frac{5}{36}$, b) $\frac{7}{36}$, c) $\frac{11}{36}$, d) $\frac{2}{36}$.

3.9. We choose a point from a square with inscribed circle. The geometrical probability that the chosen point does not belong to the circle, is equal to

a) $1-\frac{\pi}{3}$, b) $1-\frac{\pi}{4}$, c) $1-\frac{\pi}{5}$, d) $1-\frac{\pi}{6}$.

3.10. The telephone line is damaged by storm between 160 and 290 kilometers. The probability that this line is damaged between 200 and 240 kilometers, is equal to

a) $\frac{1}{13}$, b) $\frac{2}{13}$, c) $\frac{4}{13}$, d) $\frac{5}{13}$.

3.11. The distance between point A and the center of the circle with radius R is equal to d $(d > R)$. Then:

1) the probability that an accidentally drawn line with the origin at point A, will intersect the circle, is equal to

a) $\frac{2\ \arcsin(\frac{R}{d})}{\pi}$, b) $\frac{3\ arcsin(\frac{R}{d})}{\pi}$, c) $\frac{arcsin(\frac{R}{d})}{\pi}$, d) $\frac{2\ arcsin(\frac{2R}{d})}{\pi}$;

2) the probability that an accidentally drawn ray with origin A, will intersect the circle, is equal to

a) $\frac{2arcsin(\frac{R}{d})}{\pi}$, b) $\frac{3arcsin(\frac{R}{d})}{\pi}$, c) $\frac{arcsin(\frac{R}{d})}{\pi}$, d) $\frac{2arcsin(\frac{2R}{d})}{\pi}$.

3.12. We accidentally choose a point in a cube, in which is inscribed a ball. The geometrical probability that the chosen point does not belong to the ball, is equal to

a) $1-\frac{\pi}{3}$, b) $1-\frac{\pi}{4}$, c) $1-\frac{\pi}{5}$, d) $1-\frac{\pi}{6}$.

3.13.We accidentally choose a point in a ball, in which is inscribed a cube. The geometrical probability that the chosen point does not belong to the cube, is equal to

a) $1-\frac{2\sqrt{3}}{3\pi}$, b) $1-\frac{\sqrt{3}}{4\pi}$, c) $1-\frac{\sqrt{3}}{5\pi}$, d) $1-\frac{\sqrt{3}}{6\pi}$.

3.14. We accidentally choose a point in a tetrahedron , in which is inscribed a ball. The geometrical probability that the chosen point does not belong to the ball, is equal to

a) $1-\frac{5\pi\sqrt{2}}{48}$, b) $\frac{5\pi\sqrt{2}}{45}$, c) $1-\frac{\pi}{6\sqrt{3}}$, d) $1-\frac{5\pi\sqrt{2}}{80}$.

3.15. We accidentally choose a point in a ball, in which is inscribed a tetrahedron. The geometrical probability that the chosen point does not belong to the tetrahedron, is equal to

a) $1-\frac{12\sqrt{3}}{45\pi}$, b) $1-\frac{2\sqrt{3}}{9\pi}$, c) $1-\frac{12\sqrt{3}}{47\pi}$, d) $1-\frac{12\sqrt{3}}{43\pi}$.

3.16. We accidentally choose a point M in a square Δ, which is defined by

$$\Delta = \{(x,y) : 0 \le x \le 1,\ 0 \le y \le 1\}.$$

The geometrical probability that coordinates (x,y) of the point M satisfy the following condition

$$x+y \ge \frac{1}{2},$$

is equal to

a) $\frac{3}{6}$, b) $\frac{3}{4}$, c) $\frac{3}{5}$, d) $\frac{5}{6}$.

3.17.We accidentally choose a point in a square Δ, which is defined by

$$\Delta = \{(x,y) : 0 \le x \le \frac{\pi}{2},\ 0 \le y \le \frac{\pi}{2}\}.$$

The geometrical probability that coordinates (x,y) of the point M satisfy the following condition

$$\sin(x) \le y \le x,$$

is equal to

a) $1-\frac{1}{4\pi^2}$, b) $\frac{1}{2}-\frac{4}{\pi^2}$, c) $1-\frac{1}{5\pi^2}$, d) $1-\frac{1}{6\pi^2}$.

3.18.We accidentally choose a point M in a cube Δ, which is defined by :

$$\Delta = \{(x,y,z) : 0 \le x \le 1,\ 0 \le y \le 1,\ 0 \le z \le 1\}.$$

The geometrical probability that coordinates (x,y,z) of the point M satisfy the following condition

$$x^2+y^2+z^2 \le \frac{1}{4},\ x+y+z \ge \frac{1}{2},$$

is equal to

a) $\frac{\pi-1}{48}$, b) $\frac{4\pi-1}{24}$, c) $\frac{\pi-2}{50}$, d) $\frac{\pi+2}{50}$.

3.19. Two friends must meet at the concrete place in the interval of time [12 – 13] . The friend which arrived first waits no longer than 20 minutes. The probability that the meeting between the friends will happen within the mentioned interval, is equal to

a) $\frac{5}{9}$, b) $\frac{5}{8}$, c) $\frac{5}{7}$, d) $\frac{6}{7}$.

3.20. A student has planned to take money out of the bank. It is possible that he comes to the bank in the interval of time $14^{00}15 - 14^{00}25$. It is known also that the robbery of this bank is planned in the same interval of time and it will continue for 4 minutes. The probability that the student will happen to be in the bank in the moment of robbery is equal to

a) $\frac{1}{10}$, b) $\frac{1}{11}$, c) $\frac{16}{25}$, d) $\frac{1}{6}$.

3.21. We accidentally choose three points A, B and C on the circumference of a circle with radius R. The probability that ABC is an acute triangle is equal to

a) $\frac{1}{3}$, b) $\frac{1}{4}$, c) $\frac{1}{5}$, d) $\frac{1}{6}$.

3.22. We accidentally choose two points C and D on section $[AB]$ with length ℓ. The probability that we can construct a triangle by the obtained three sections is equal to:

a) $\frac{1}{4}$, b) $\frac{1}{5}$, c) $\frac{1}{6}$, d) $\frac{1}{7}$.

3.23. We accidentally choose point $M = (p, q)$ in cube Δ which is defined by :

$$\Delta = \{(p, q) : 0 \leq p \leq 1,\ 0 \leq q \leq 1\}.$$

The geometrical probability that the roots of equation $x^2 + px + q = 0$ will be real numbers is equal to

a) $\frac{1}{12}$, b) $\frac{1}{13}$, c) $\frac{1}{5}$, d) $\frac{1}{6}$.

3.24. We accidentally choose point M from the sphere with radius R. The probability that distance ρ between point M and the center of the above mentioned sphere satisfies condition $\frac{R}{2} < \rho < \frac{2R}{3}$, is equal to

a) $\frac{7}{27}$, b) $\frac{1}{4}$, c) $\frac{37}{216}$, d) $\frac{8}{29}$.

Chapter 4

Total Probability and Bayes' Formulas

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let B be any event with positive probability (i.e., $P(B) > 0$).

We denote with $P(\cdot \mid B)$ a real-valued function defined on the σ-algebra $\mathcal{F}$ by :

$$(\forall X)(X \in \mathcal{F} \to P(X|B) = \frac{P(X \cap B)}{P(B)}).$$

The function $P(\cdot \mid B)$ is called a conditional probability relative to the hypothesis that the event B occurred.

The number $P(X|B)$ is the probability of the event X relative to the hypothesis that the event B occurred.

Theorem 4.1. *If $B \in \mathcal{F}$ and $P(B) > 0$, then the conditional probability $P(\cdot \mid B)$ is the probability.*

Proof. By the definition of $P(\cdot \mid B)$, we have:

1) $P(A|B) \geq 0$ for $A \in \mathcal{F}$;

2) $P(\Omega|B) = 1$;

3) If $(A_k)_{k \in N}$ is a pairwise-disjoint family of events, then

$$P(\cup_{k \in N} A_k | B) = \sum_{k \in N} P(A_k | B).$$

The validity of 1) follows from the definition of $P(\cdot \mid B)$ and from the non-negativity of the probability P. Indeed,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \geq 0.$$

The validity of 2) follows from the following relations

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

19

The validity of 3) follows from the countable-additivity of P and from the elementary fact that if $(A_n)_{n\in N}$ is a family of pairwise-disjoint events, then the family of events $(A_n \cap B)_{n\in N}$ also has the same property. Indeed,

$$P(\cup_{n\in N} A_n|B) = \frac{P((\cup_{n\in N} A_n) \cap B)}{P(B)} = \frac{P(\cup_{n\in N}(A_n \cap B))}{P(B)} =$$

$$= \frac{\sum_{n\in N} P(A_n \cap B)}{P(B)} = \sum_{n\in N} \frac{P(A_n \cap B)}{P(B)} = \sum_{n\in N} P(A_n \mid B).$$

This ends the proof of theorem.

Theorem 4.2 *If $P(B) > 0$, then $P(\overline{B}|B) = 0$.*

Proof.

$$P(\overline{B}|B) = \frac{P(\overline{B} \cap B)}{P(B)} = \frac{P(\emptyset)}{P(B)} = 0.$$

Definition 4.1 Two events A and B are called independent if

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

Example 4.1 Assume that $\Omega = \{(x, y) : x \in [0; 1], y \in [0; 1]\}$.

Let $\mathcal{F}$ denotes a class of all Borel subsets of Ω. (cf.Chapter 5, Example 5.3). Let denote by P the classical Borel probability measure b_2 on Ω. Then two events

$$A = \{(x, y) : x \in [0; \frac{1}{2}], y \in [0; 1]\},$$

and

$$B = \{(x, y) : x \in [0; 1], y \in [\frac{1}{2}; \frac{3}{4}]\}$$

are independent.

Indeed, on the one hand, we have

$$P(A \cap B) = b_2(A \cap B) = b_2(\{(x, y) : x \in [0; \frac{1}{2}], y \in [\frac{1}{2}; \frac{3}{4}]\} = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}.$$

On the other hand, we have

$$P(A) \cdot P(B) = b_2(A) \cdot b_2(B) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}.$$

Finally, we get

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

which shows us that two events A and B are independent.

Theorem 4.3 *If two events A and B are independent and $P(B) > 0$, then $P(A|B) = P(A)$.*

Proof. From the definition of the conditional probability, we have

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

The independence of events A and B gives $P(A \cap B) = P(A) \cdot P(B)$. Finally, we get

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

This ends the proof of theorem .

Remark 4.1 Theorem 4.3 asserts that when events A and B are independent then any information concerned the occurrence or non-occurrence of one of them does not affect the probability of the other.

Theorem 4.4 *If two events A and B are independent, independent are also events $\overline{A}$ and B .*

Proof.

$$P(\overline{A} \cap B) = P((\Omega \setminus A) \cap B) = P((\Omega \cap B) \setminus (A \cap B)) =$$
$$= P(B \setminus (A \cap B)) = P(B) - P(A \cap B) = P(B) - P(A) \cdot P(B) =$$
$$= P(B)(1 - P(A)) = P(B) \cdot P(\overline{A}).$$

This ends the proof of theorem.

Example 4.2

Experiment. - *We cast two six-sided dice.*

Problem. *What is the probability that the sum of rolled numbers is equal to 8 relative to the hypothesis that the sum of rolled numbers is even ?*

Modelling of the experiment. A probability space Ω of all elementary events has the following form

$$\Omega = \{(x, y) : x \in N, y \in N, \ 1 \leq x \leq 6, \ 1 \leq y \leq 6\},$$

where x and y denote rolled numbers on the first and second dice, respectively.

We denote with $\mathcal{F}$ a class of all subsets of Ω. Let P denote a classical probability. Finally, probability space $(\Omega, \mathcal{F}, P)$ describing our experiment is constructed.

Solution of the problem. Let denote with A a subset of Ω which corresponds to the event:

" The sum of rolled numbers is equal to 8 ".

Then, the event A has the following form:

$$A = \{ \ (6;2); \ (5;3); \ (4;4); \ (3;5); \ (2;6) \ \}.$$

Let denote with B the following event:

"The sum of rolled numbers is even".

We have

$$B = \{(1;1);\ (1;3);\ (2;2);\ (3;1);\ (1;5);\ (2;4);\ (3;3); (4;4);\ (5;1);$$

$$(6;2);\ (5;3); (4;4);\ (3;5);\ (2;6);\ (6;4);\ (5;5);\ (4;6); (6;6)\ \}$$

Note that $A \cap B = A$. By the definition of the classical probability we have:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{5}{36} : \frac{18}{36} = \frac{5}{18}.$$

Resume. If we cast two six-sided dies then the conditional probability that the sum of rolled numbers is equal to 8 relative to the hypothesis that the sum of rolled numbers is even, is equal to $\frac{5}{18}$.

Definition 4.2 Let, $J \subseteq N$. A family of event $(A_i)_{i \in J}$ is called a complete system of representatives if :

1) $A_i \cap A_j = \emptyset,\ i, j \in J, i \neq j$,
2) $(\forall j)(j \in J \rightarrow P(A_j) > 0)$,
3) $\cup_{j \in J} A_j = \Omega$.

Theorem 4.5 *Let, $J \subseteq N$ and $(A_j)_{j \in J}$ be a complete system of representatives. For arbitrary event B the following formula*

$$P(B) = \sum_{j \in J} P(B|A_j) \cdot P(A_j)$$

is valid, which is called the law of total probability.

Proof. We have

$$B = \cup_{j \in J}(B \cap A_j),$$

where $(B \cap A_j)_{j \in J}$ is a family of pairwise-disjoint events.
Indeed, we have,

$$B = B \cap \Omega = B \cap (\cup_{j \in J} A_j) = \cup_{j \in J}(B \cap A_j).$$

Hence, from the countable-additivity of P we have

$$P(B) = \sum_{j \in J} P(B \cap A_j).$$

Note that for arbitrary natural number j $(j \in J)$ we have

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

Hence,

$$P(B \cap A_j) = P(A_j) \cdot P(B|A_j).$$

Finally, we get

$$P(B) = \sum_{j \in J} P(B \cap A_j) = \sum_{j \in J} P(A_j) \cdot P(B|A_j).$$

This ends the proof of theorem.

Example 4.3

Experiment. *There are placed 3 white and 3 black balls in an urn I, 3 white and 4 black balls in an urn II and 4 white and 1 black balls in the urn III. We accidentally choose a box and further accidentally choose a ball from this urn.*

Problem. *What is the probability that accidentally chosen ball is white if the probability of a choice of any urn is equal to* $\frac{1}{3}$ *?.*

Solution of the Problem. Let A_i denote an event that we have chosen i-th urn $(1 \leq i \leq 3)$. Then we obtain that $P(A_1) = P(A_1) = P(A_1) = \frac{1}{3}$.

Let B denote an event which corresponds to the fact that we have chosen white ball.

By the definition of the conditional probability, we have

$$P(B|A_1) = \frac{1}{2},\ P(B|A_2) = \frac{3}{7},\ P(B|A_2) = \frac{4}{5}.$$

Using the formula of total probability, we have

$$P(B) = \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{3}{7} + \frac{1}{3} \cdot \frac{4}{5} = \frac{57}{105}.$$

Resume. *The probability that accidentally chosen ball will be white in our experiment is equal to* $\frac{57}{105}$.

Example 4.4

Experiment. *The probability of formation of k-bacteria $(k \in N)$ is equal to* $\frac{\lambda^k}{k!}e^{-\lambda}$ $(\lambda > 0)$. *The probability of adaptation with environment of the formed bacterium is equal to p $(0 < p < 1)$.*

Problem. *What is the probability that n bacteria $(n \in N)$ will pass the adaptation process ?*

Solution of the problem. Let A_k be the event that k bacteria $(k \in N)$ pass adaptation process. Note that $(A_k)_{k \in N}$ is a complete system of representatives. Let B_n denote the event that n baqteria pass the adaptation process $(n \in N)$. Note that $P(B_n|A_k) = 0$, when $k \leq n-1$. If $k \geq n$, then $P(B_n|A_k) = C_k^n p^n (1-p)^{k-n}$. We have

$$P(B_n) = \sum_{k \in N} P(A_k) P(B_n|A_k) = \sum_{k=n} \frac{\lambda^k}{k!} e^{-\lambda} C_k^n p^n (1-p)^{k-n} =$$

$$= \sum_{k=n} \frac{\lambda^k}{k!} e^{-\lambda} \frac{k!}{n! \cdot (k-n)!} p^n (1-p)^{k-n} =$$

$$= \frac{(p\lambda)^n}{n!} e^{-\lambda} \sum_{k=n} \frac{\lambda^{k-n}}{\cdot (k-n)!} (1-p)^{k-n} = \frac{(p\lambda)^n}{n!} e^{-\lambda} \sum_{k=n} \frac{(\lambda \cdot (1-p))^{k-n}}{\cdot (k-n)!} =$$

$= \frac{(p\lambda)^n}{n!} e^{-\lambda} e^{\lambda \cdot (1-p)} = \frac{(p\lambda)^n}{n!} e^{-p \cdot \lambda}$.

Resume. *The probability that* n *bacteria pass the adaptation process* $(n \in N)$ *is equal to* $\frac{(p\lambda)^n}{n!} e^{-p \cdot \lambda}$.

Theorem 4.6 *Assume that* $J \subseteq N$ *and* $(A_j)_{j \in J}$ *be a complete system of representatives. For every event* B *with* $P(B) > 0$ *we have*

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j \in J} P(A_j)P(B|A_j)} \ (i \in J),$$

which is called Bayes' [1]*formula.*

Proof. Using the total probability formula and the definition of the conditional probability, we have

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{j \in J} P(A_j)P(B|A_j)} \ (i \in J).$$

This ends the proof of theorem.

Example 4.5 *Assume that we have chosen a white ball in the experiment considered in Example 3.*

Problem. *What is the probability that we have chosen white ball from the first urn?*

Solution of the problem. By the Bayes' formula we have

$$P(A_1|B) = \frac{P(A_1) \cdot P(B|A_1)}{P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2)} = \frac{1}{3} \cdot \frac{2}{5} : \frac{57}{105} = \frac{2}{9}.$$

Example 4.6 (A problem about ruin of the player). *Let consider the problem concerned with throwing of a coin, when playing "heads" or "tails". If comes the side of the coin which has been chosen by the player, he wins 1 lari. In other case he loses the same amount of money. Assume that the initial capital of the player is* x *lari and he wishes to increase his capital to* a *lari* $(x < a)$. *The game is finished when the player is ruined or when he increases his capital to* a *lari.*

Problem. *What is the probability that the player will be ruined ?*

Solution of the problem.

Let $p(x)$ denote the probability of the ruin of the player when his initial capital is x lary. Then after one step in the case of winning the probability of the ruin will

[1]Bayes Thomas(1702, London-4.4.1761, Tanbridj)-English mathematician, the member of London Royal Society (1742),main works in probability theory (1763).

be $p(x+1)$, in other case same probability will be $p(x-1)$. Let B_1 denote the event, which corresponds to the case when the player wins in the first step. Analogously, denote by B_2 the event, which corresponds to the case when the player loses in the first step.

We denote by A the event which corresponds to the ruin of the player. Then by the definition of the conditional probability, we have

$$P(A|B_1) = p(x+1), \; P(A|B_2) = p(x-1).$$

It is clear that (B_1, B_2) is a complete system of events. Since the coin is symmetrical, we have

$$P(B_1) = P(B_2) = \frac{1}{2}.$$

Using a total probability formula, we have

$$p(x) = \frac{1}{2}[p(x+1) + p(x-1)].$$

Note that the following initial conditions $p(0) = 1$ and $p(a) = 0$ are fulfilled. One can easily check that the following linear function

$$p(x) = c_1 + c_2 x,$$

whose coefficients are defined by

$$p(0) = c_1 = 1, \; p(a) = c_1 + c_2 a = 0,$$

is a solution of the above mentioned equation.

Finally, we get

$$p(x) = 1 - \frac{x}{a}, \; 0 \le x \le a.$$

Resume. *The probability $p(x)$ that the player will be ruined in the above described game in the case when his initial capital is equal to x lari, is equal to $1 - \frac{x}{a}$, $(0 \le x \le a)$.*

Example 4.6 (The problem about division of game between hunters). *The probability of shooting the game for the first hunter is equal to 0,8. The same probability for the second hunter is - 0,7. The beast was shot with simultaneous shots, but with only one bullet. The mass of the game was 190 kg. It was found that the game was killed with one bullet. How should the game be devided between hunters?*

Solution of the problem. Let B denote the event that the game was killed by one hunter in the case of simultaneous shots. Let A_1 and A_2 denote events that the animal was killed by the first and the second hunters, respectively. Using Bayes' formulas we obtain

$$P(A_1|B) = \frac{0,3 \cdot 0,8}{0,3 \cdot 0,8 + 0,2 \cdot 0,7} = \frac{12}{19},$$

$$P(A_2|B) = \frac{0,2 \cdot 0,7}{0,3 \cdot 0,8 + 0,2 \cdot 0,7} = \frac{12}{19}.$$

It follows, that $P(A_1|B) \cdot 190 = 120$ (kg) of the game belongs to the first hunter and $P(A_2|B) \cdot 190 = 70$ (kg) of the game belongs to the second hunter, respectively.

Tests

4.1. Two shots shoot a target. The probability that the first shot will hit the shooting mark is equal to $0,9$. Analogous probability for the second shot is $0,7$. Then the probability that the target will be hit by both shots, is equal to

a) $0,42$, b) $0,63$, c) $0,54$, d) $0,36$.

4.2. The number of non-rainy days in June for Tbilisi is equal to 25. The probability that the first two days would be non-rainy is equal to

a) $\frac{5}{87}$, b) $\frac{20}{29}$, c) $\frac{19}{29}$, d) $\frac{18}{29}$.

4.3. We have accidently chosen two points A and B from set Δ, which is defined by

$$\Delta = \{(x,y) : x \in [0,1],\ y \in [0,1]\}.$$

The functions g and f are defined by

$$g((x,y)) = \begin{cases} 1, & \text{if } x^2 + y^2 \le \frac{1}{4}, \\ 0, & \text{if } x^2 + y^2 > \frac{1}{4}, \end{cases}$$

$$f((x,y)) = \begin{cases} 0, & \text{if } x + y \le \frac{1}{2}, \\ 1, & \text{if } x + y > \frac{1}{2}, \end{cases}$$

Then the probability that $g(A) + f(B) = 1$, is equal to

a) $\frac{1}{4} - \frac{\pi}{8}$, b) $\frac{1}{2} - \frac{\pi}{8}$, c) $\frac{8}{9} - \frac{\pi}{16}$, d) $1 - \frac{\pi}{8}$.

4.4. Here we have three boxes with the following composition of balls

Urn	Black balls	White balls
I	2	3
II	3	2
III	1	4

We accidently choose a box, from which accidentally choose also a ball.

1) The probability that the chosen ball is white, is equal to

a) $0,4$, b) $0,6$, c) $0,7$, d) $0,8$;

2) It is known that an accidently chosen ball is white. The probability that we have chosen a ball from box I, is equal to

a) $\frac{1}{3}$, b) $\frac{1}{4}$, c) $\frac{1}{5}$, d) $\frac{1}{6}$.

4.6. 100 and 200 details are produced in plants I and II, respectively. The probabilities of the producing of a standard detail in plants I and II are equal to $0,9$ and $0,8$, respectively.

1) A damage caused by realization of non-standard details made up 3000 lari. A fine which must be payed by the administration of plant II caused realization of its non-standard details is equal to

a) 2400 lari, b) 2300 lari, c) 2000 lari, d) 1600 lari;

2)The prófit received by realization of standard details made up 5000 lari. A portion of the prófit due to plant I is equal to

a) 1800, b) 1700, c) 1400, d) 3000 .

4.7. The player chooses a "heads" or "tails". If comes the side of the coin which was chosen by the player, then he wins 1 lari. In other case he loses the same mouny. Assume that the initial capital of the player is 1000 lari and he wishes to increase his capital to 2000 lari . The game is finished when the player is ruined or when the player will increase his capital to 2000 lari. What is the probability that the player will increase his capital to the until desired amount ?

a) $0,4$, b) $0,5$, c) $0,6$, d) $0,7$.

4.8.The probability of formation of k-bacteria $(k \in N)$ is equal to $\frac{0,3^k}{k!}e^{-0,3}$. The probability of the adaptation with environment of the formed bacterium is equal to $0,1$.

1) The probability that at least 5 bacteria will pass the adaptation process is equal to

a) $\frac{0,03^5}{5!}e^{-0,03}$, b) $\frac{0,04^5}{5!}e^{-0,04}$, c) $\frac{0,05^5}{5!}e^{-0,05}$, d) $\frac{0,06^5}{5!}e^{-0,06}$;

2) The observation of the accidentally chosen bacterium showed us that it has passed the adaptation process. The probability that this bacterium belongs to the adapted family consisting of 6 members, is equal to

a) $\frac{0,03^6}{6!} : (\sum_{k=1}^{\infty} \frac{0,03^k}{k \cdot k!}e^{-0,03})$, b) $\frac{0,03^6}{6 \cdot 6!} : (\sum_{k=1}^{\infty} \frac{0,03^k}{k \cdot k!}e^{-0,03})$.

Chapter 5

Applications of Caratheodory Method

§5.1. Construction of Probability Spaces with Caratheodory [1] method.

Let Ω be a non-empty set and let F be any class of its subsets.

Lemma 5.1.1 *There exists a σ-algebra $\sigma(F)$ of subsets of Ω, which contains the class F and is minimal in the sense of inclusion in such σ-algebras which contain F.*

Proof. Let $(\mathcal{F}_j)_{j\in J}$ denote a family of all σ-algebras of subsets of Ω which contain F and define class $\sigma(F)$ by the following formula :

$$\sigma(F) = \cap_{j\in J}\mathcal{F}_j.$$

Let show that $\sigma(F)$ is a σ-algebra. Indeed,

1) $\Omega \in \sigma(F)$, because $\Omega \in \mathcal{F}_j$ for $j \in J$.

2) Let $(A_k)_{k\in N}$ be any sequence of elements of $\sigma(F)$. Since this is a sequence of elements of $\mathcal{F}_j$ for arbitrary $j \in J$, we conclude that $\cap_{k\in N}A_k \in \mathcal{F}_j$ and $\cup_{k\in N}A_k \in \mathcal{F}_j$. The later relation means that $\cap_{k\in N}A_k \in \cap_{j\in J}\mathcal{F}_j \in \sigma(F)$ and $\cup_{k\in N}A_k \in \cap_{j\in J}\mathcal{F}_j \in \sigma(F)$.

3) If $A \in \sigma(F)$, then $\overline{A} \in \mathcal{F}_j$ for arbitrary $j \in J$. The later relation means that $\overline{A} \in \cap_{j\in J}\mathcal{F}_j = \sigma(F)$.

Now assume that $\sigma(F)$ is not minimal (in the sense of inclusion) σ-algebra of subsets containing F. It means that there exists a σ-algebra $\mathcal{F}^*$, such that the following two conditions

1) $F \subset \mathcal{F}^*$;

2) $\mathcal{F}^* \subset \sigma(F)$ and $\sigma(F) \setminus \mathcal{F}^* \neq \emptyset$,

are fulfilled.

---

[1] *Caratheodory Constantin (13.9.1873, Berlin–2.2.1950, Münhen)-German mathematician. Professor of the Münhen University (1924-39), Lecturer of the Athena University(1933). Main works in theory of measures.*

By the definition of family $(\mathcal{F}_j)_{j\in J}$ there exists an index $j_0 \in J$ such that $\mathcal{F}_{j_0} = \mathcal{F}^*$. Hence, $\sigma(F) \subset \mathcal{F}^*$, which contradicts to 2) . So we have obtained a contradiction and Lemma 5.1.1 is proved.

Definition 5.1.1 Let S_1 and S_2 be two classes of subsets of Ω such that $S_1 \subset S_2$. Let P_1 and P_2 be two real-valued functions defined on S_1 and S_2, respectively. Function P_2 is called an extension of function P_1 if

$$(\forall X)(X \in S_1 \rightarrow P_2(X) = P_1(X)).$$

Definition 5.1.2 Let $\mathcal{A}$ be an algebra of subsets Ω. A real-valued function P defined on $\mathcal{A}$ is called a probability if the following three conditions

1) $P(A) \geq 0$ for $A \in \mathcal{A}$,

2) $P(\Omega) = 1$,

3) If $(A_k)_{k\in N}$ is a pairwise disjoint family of element of $\mathcal{A}$ such that $\cup_{k\in N} A_k \in \mathcal{A}$, then

$$P(\cup_{k\in N} A_k) = \sum_{k\in N} P(A_k),$$

are fulfilled.

The general method of the construction of probability spaces is contained in the following Theorem.

Theorem 5.1.1 (Charatheodory). *Let $\mathcal{A}$ be an algebra of subsets of Ω and P be a probability measure defined on $\mathcal{A}$. Then there exists a unique probability measure $\overline{P}$ on $\sigma(\mathcal{A})$ which is an extension of P. It is defined by the following formula:*

$$(\forall B)(B \in \sigma(\mathcal{A}) \rightarrow \overline{P}(B) = \inf\{ \sum_{k\in N} P(A_k) | (\forall k)(k \in N \rightarrow A_k \in \sigma(\mathcal{A}))$$

$$\& \ B \subseteq \cup_{k\in N} A_k \ \}.$$

Remark 5.1.1 *The proof of Theorem 1 can be found in [6] .*

Below we consider some applications of Theorem 5.1.1.

5.2. Construction of the Borel one-dimensional measure b_1 on $[0,1]$

Let $\mathcal{A}$ denote a class of subsets of $[0,1]$ elements of which can be presented as a union of finite number of elements of the form

$$[a_k, b_k[, \ [a_k, b_k], \ ]a_k, b_k[, \ ]a_k, b_k]$$

.

One can easily show that $\mathcal{A}$ is an algebra of subsets of Ω. We set

$$P([a_k, b_k[) = P([a_k, b_k]) = P(]a_k, b_k[) = P(]a_k, b_k]) = b_k - a_k$$

In natural way, we can define P on the union of a finite number of disjoint elements of $\mathcal{A}$. It is not difficult to check that P is a probability measure defined on $\mathcal{A}$. Using Theorem 1 we deduce that there exists a unique extension $\overline{P}$ on $\sigma(\mathcal{A})$.

A class $\sigma(\mathcal{A})$ is called a Borel σ-algebra of subsets of $[0,1]$ and is denoted by $\mathcal{B}([0,1])$. The probability $\overline{P}$ is called one-dimensional classical Borel measure on $[0,1]$ and is denoted by b_1.

$([0,1], \mathcal{B}([0,1]), b_1)$ is called a Borel classical probability space associated with $[0,1]$.

5.3. Construction of Borel probability measures on R

Let $F : R \to [0,1]$ be a continuous from the right function on R, which satisfy the following condition:

$$\lim_{x \to -\infty} F(x) = 0 \ \& \ \lim_{x \to +\infty} F(x) = 1.$$

We set $F(-\infty) = 0$ and $F(+\infty) = 1$.

Let $\Omega = R \cup \{+\infty\}$.

Let $\mathcal{A}$ denote a class of all subsets of Ω, which are represented by the union of finite number of "semi-closed intervals" of the form $(a, b]$, i.e.,

$$\mathcal{A} = \{A | A = \sum_{i=1}^{n} (a_i, b_i]\},$$

where $-\infty \leq a_i < b_i \leq \infty$ $(1 \leq i \leq n)$.

It is easy to show that $\mathcal{A}$ is an algebra of subsets of Ω. We set

$$P(A) = P(\sum_{i=1}^{n} (a_i, b_i]) = \sum_{i=1}^{n} [F(b_i) - F(a_i)].$$

One can easily demonstrate that the real-valued function P is a probability defined on $\mathcal{A}$. Using Charatheodory theorem we deduce that there exists a unique probability measure $\overline{P}$ on $\sigma(\mathcal{A})$ which is an extension of P. The class $\sigma(\mathcal{A}) \cap R$ is called a Borel σ-algebra of subsets of the real axis R and is denoted by $\mathcal{B}(R)$. A real-valued function P_F , defined by

$$(\forall X)(X \in \mathcal{B}(R) \to P_F(X) = \overline{P}(X)),$$

is called a probability Borel measure on R defined by the function F.

Example 5.3.1. Let F be defined by

$$(\forall x)(x \in R \to F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \ ).$$

Let P_F be a Borel probability measure on R generated by F. Then $(\Omega, \mathcal{F}, P) = (R, \mathcal{B}(R), P_F)$ is called a canonical one-dimensional Gaussian Borel probability space, associated with one-dimensional Euclidean vector space $R (= R^1)$.

The real-valued function P_F is called one-dimensional canonical Borel probability measure on R and is denoted by Γ_1.

5.4. The product of a finite family of probabilities

Let $(\Omega_i, \mathcal{F}_i, P_i)_{1 \le i \le n}$ be a finite family of probability spaces.
We introduce some notions.

$$\prod_{i=1}^{n} \Omega_i = \{(\omega_1, \cdots, \omega_n) \mid \omega_1 \in \Omega_1, \cdots, \omega_n \in \Omega_n\}.$$

A set $A \subseteq \prod_{i=1}^{n} \Omega_i$ is called cylindrical , if the following representation

$$B = \prod_{i=1}^{n} B_i$$

is valid, where $B_i \in \mathcal{F}_i$ $(1 \le i \le n)$.

Let $\mathcal{A}$ be a class of all subsets of $\prod_{i=1}^{n} \Omega_i$ which are represented by the union of a finite number of pairwise disjoint cylindrical subsets of $\prod_{i=1}^{n} \Omega_i$. Note that $\mathcal{A}$ is an algebra of subsets of $\prod_{i=1}^{n} \Omega_i$.

We set

$$P(\prod_{i=1}^{n} B_i) = \prod_{i=1}^{n} P_i(B_i)$$

and extend in the natural way a function P on class $\mathcal{A}$. Now one can easily demonstrate that function P is a probability measure defined on algebra $\mathcal{A}$. Using Charatheodory theorem there exists a unique probability measure $\overline{P}$ on $\sigma(\mathcal{A})$ which extends P. Class $\sigma(\mathcal{A})$ is called a product of the family of σ-algebras $(\mathcal{F}_i)_{1 \le i \le n}$ and is denoted by $\prod_{1 \le i \le n} \mathcal{F}_i$. The probability $\overline{P}$ is called a product of the family of probabilities $(P_i)_{1 \le i \le n}$ and is denoted by $\prod_{i=1}^{n} P_i$.

The triplet $(\prod_{i=1}^{n} \Omega_i, \prod_{i=1}^{n} \mathcal{F}_i, \prod_{i=1}^{n} P_i)$ is called a product of the family of probability spaces $(\Omega_i, \mathcal{F}_i, P_i)_{1 \le i \le n}$.

Remark 5.4.1 Let consider a sequence of n independent random experiments. It is such a sequence of n random experiments when the result of any experiment does not influence the result of the next experiment.

Assume that i-th $(1 \le i \le n)$ experiment is described by a probability space $(\Omega_i, \mathcal{F}_i, P_i)$.Then a sequence of n independent random experiments is described by $(\prod_{i=1}^{n} \Omega_i, \prod_{i=1}^{n} \mathcal{F}_i, \prod_{i=1}^{n} P_i)$.

Here we consider some examples.

Example 5.4.1 (Bernoulli [2] probability measure).
We set

[2]Jacob Bernoulli (27.12.1654-16.8.1705)-Swedish mathematician, professor of Bazel University (1687). Him belongs the first proof of the so called "Bernoulli theorem"(which is a partial case of the "Law of Large numbers") (cf. *"Arsconjectandi"*(*Basileqe*)(1713).

$$\Omega_i = \{0,1\},\ \mathcal{F}_i = \{A|A \subseteq \Omega_i\},\ P_i(\{1\}) = p$$

for $1 \le i \le n$ and $0 < p < 1$.

The product of probability spaces

$$(\prod_{i=1}^{n} \Omega_i, \prod_{i=1}^{n} \mathcal{F}_i, \prod_{i=1}^{n} P_i)$$

is called the Bernoulli n-dimensional classical probability space. $\prod_{i=1}^{n} P_i$ is called the Bernoulli n-dimensional probability measure.

If we consider a set A_k defined by

$$A_k = \{(\omega_1, \cdots, \omega_n)|(\omega_1, \cdots, \omega_n) \in \prod_{i=1}^{n} \Omega_i \ \&\ \sum_{i=1}^{n} \omega_i = k\},$$

then using the structure of the above-mentioned measure, we obtain

$$(\forall(\omega_1, \cdots, \omega_n))((\omega_1, \cdots, \omega_n) \in A_k \to \prod_{i=1}^{n} P_i((\omega_1, \cdots, \omega_n)) = p^k(1-p)^{n-k}).$$

Hence, $\prod_{i=1}^{n} P_i(A_k) = |A_k|p^k(1-p)^{n-k}$, where $|\cdot|$ denotes the cardinality of the corresponding set. It is easy to show that $|A_k| = C_n^k$, where C_n^k denotes the cardinality of all different subsets of cardinality k in the fixed set of cardinality n.

The probability $\prod_{i=1}^{n} P_i(A_k)$ is denoted by $P_n(k)$, which means that during n-random two $\{0,1\}$-valued experiments the event $\{1\}$ had occurred k-times, if it is known that the probability of event $\{1\}$ in an arbitrary experiment is equal to p. If we denote by q the probability of event $\{0\}$, then we obtain the following formula

$$P_n(k) = C_n^k p^k q^{n-k} \ (1 \le k \le n),$$

which is called Bernoulli formula.

A natural number $k_0 \in [0,n]$ is called a number with hight probability if

$$P(k_0) = \max_{0 \le k \le n} P_n(k).$$

The number k_0 with hight probability is calculated by the following formula

$$k_0 = \begin{cases} [(1+n)p], & \text{if } (1+n)p \notin Z; \\ (1+n)p \text{ and } (1+n)p - 1, & \text{if } (1+n)p \in Z, \end{cases}$$

where $[\cdot]$ denotes an integer part of the corresponding number.

Example 5.4.2 (n-dimensional multinomial probability measure).

Let triplet $(\Omega_i, \mathcal{F}_i, P_i)_{1 \le i \le n}$ be defined by :

a) $\Omega_i = \{x_1, \cdots, x_k\}$ $(1 \le i \le n)$;

b) $\mathcal{F}_i$ is a powerset of Ω_i for $1 \le i \le n$;

c) $P_i(\{x_j\}) = p_j > 0,\ 1 \le i \le n,\ 1 \le j \le k,\ \sum_{j=1}^{k} p_j = 1$.

Then triplet $(\prod_{i=1}^n \Omega_i, \prod_{i=1}^n \mathcal{F}_i, \prod_{i=1}^n P_i)$ is called n-dimensional multinomial probability space. A real-valued function $\prod_{i=1}^n P_i$ is called n-dimensional multinomial probability.

If we consider a set $A_n(n_1, \cdots, n_k)$, defined by

$$A_n(n_1, \cdots, n_k) = \{(\omega_1, \cdots, \omega_n) | (\omega_1, \cdots, \omega_n) \in \prod_{i=1}^n \Omega_i \ \&$$

$$|\{i : \omega_i = x_p\}| = n_p, \ 1 \le p \le k\},$$

then using the structure of the product measure, we obtain

$$(\forall (\omega_1, \cdots, \omega_n))((\omega_1, \cdots, \omega_n) \in A_n(n_1, \cdots, n_k) \rightarrow$$

$$\prod_{i=1}^n P_i((\omega_1, \cdots, \omega_n)) = p_1^{n_1} \times \cdots \times p_k^{n_k}).$$

Hence, $\prod_{i=1}^n P_i(A_n(n_1, \cdots, n_k)) = |A_n(n_1, \cdots, n_k)| \times p_1^{n_1} \times \cdots \times p_k^{n_k}$, where $|\cdot|$ denotes the cardinality of the corresponding set. It is not difficult to prove that $|A_n(n_1, \cdots, n_k)| = \frac{n!}{n_1! \times \cdots \times n_k!}$.

Then probability $\prod_{i=1}^n P_i(A_n(n_1, \cdots, n_k))$ (denoted by $P_n(n_1, \cdots, n_k)$) assumes that during n-random $\{x_1, \cdots, x_k\}$-valued experiments the event x_1 will occur n_1-times, $\cdots$, the event x_k will occur n_k-times if it is known that in i-th experiment the probability that the event x_i occurred is equal to p_i $(1 \le i \le k)$, is calculated by the following formula

$$P_n(n_1, \cdots, n_k) = \frac{n!}{n_1! \times \cdots \times n_k!} \times p_1^{n_1} \times \cdots \times p_k^{n_k}.$$

This formula is called the formula for calculation of n-dimensional multinomial probability.

Example 5.4.3 (n-dimensional Borel classical measures on $[0,1]^n$ and R^n).
Assume that a family of probability spaces $(\Omega_i, \mathcal{F}_i, P_i)_{1 \le i \le n}$ is defined by:
a) $\Omega_i = [0,1]$ $(1 \le i \le n)$,
b) $\mathcal{F}_i = \mathcal{B}([0,1])$ $(1 \le i \le n)$,
c) $P_i = b_1$, $(1 \le i \le n)$.

Then triplet $(\prod_{i=1}^n \Omega_i, \prod_{i=1}^n \mathcal{F}_i, \prod_{i=1}^n P_i)$ is called n-dimensional Borel probability space associated with n-dimensional cube $[0,1]^n$. The real-valued function $\prod_{i=1}^n P_i$ is called n-dimensional classical Borel measure defined on $[0,1]^n$.

The real-valued function b_n, defined by

$$(\forall X)(X \in B(R^n) \rightarrow b_n(X) = \sum_{h \in Z^n} \prod_{i=1}^n P_i([0,1[^n \cap (X-h))),$$

is called n-dimensional classical Borel measure defined on R^n.

Example 5.4.4 Assume that a family of functions $(F_i)_{1\le i\le n}$ is defined by

$$(\forall i)(\forall x)(1 \le i \le n \ \&\ x \in R \rightarrow F_i(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt\ ).$$

Assume also that P_i denotes the probability measure on R defined by F_i. Then the probability space $(\prod_{1\le i\le n}\Omega_i, \prod_{i\in N}\mathcal{F}_i, \prod_{i\in N}P_i)$ is called n-dimensional Gaussian canonical probability space associated with R^n.

The real-valued function $\prod_{1\le i\le n}P_i$ is called n-dimensional Gaussian canonical probability measure on R^n and is denoted by Γ_n.

5.5. Definition of the Product of the Infinite Family of Probabilities

Let $(\Omega_i, \mathcal{F}_i, P_i)_{i\in I}$ be an infinite family of probability spaces.
We set

$$\prod_{i\in I}\Omega_i = \{(\omega_i)_{i\in I}\ :\ \omega_i \in \Omega_i, i \in I\}.$$

A subset $A \subseteq \prod_{i\in I}\Omega_i$ is called a cylindrical set, if there exists a finite number of indices $(i_k)_{1\le k\le n}$ and such elements $(B_{i_k})_{1\le k\le n}$ of σ-algebras $(\mathcal{F}_{i_k})_{1\le k\le n}$ that

$$B = \{(\omega_i)_{i\in I}\ :\ (\omega_i \in \Omega_i, i \in I \setminus \cup_{k=1}^{n}\{i_k\}\ )\ \&\ (\omega_i \in B_i, i \in \cup_{k=1}^{n}\{i_k\})\ \}.$$

Let $\mathcal{A}$ denote a class of such subsets of $\prod_{i\in I}\Omega_i$ which are presented by the union of finite number of pairwise disjoint cylindrical subsets. Note that class $\mathcal{A}$ is an algebra of subsets of $\prod_{i\in I}\Omega_i$.

Define a real-valued function P on the cylindrical subset B by the following formula

$$P(B) = \prod_{k=1}^{n} P_{i_k}(B_{i_k})$$

and extend in natural way a functional P on class $\mathcal{A}$. Clearly, a real-valued function P is the probability defined on algebra $\mathcal{A}$. Using Charatheodory theorem we deduce the existence of the unique extended probability measure $\overline{P}$ on class $\sigma(\mathcal{A})$. The class of subsets $\sigma(\mathcal{A})$ is called the product of the infinite family of σ-algebras $(\mathcal{F}_i)_{i\in I}$ and is denoted by $\prod_{i\in I}\mathcal{F}_i$. The real-valued function $\overline{P}$ is called the product of the infinite family of probabilities $(P_i)_{i\in I}$ and is denoted by $\prod_{i\in I}P_i$.

Triplet $(\prod_{i\in I}\Omega_i, \prod_{i\in I}\mathcal{F}_i, \prod_{i\in I}P_i)$ is called the product of the infinite family of probability spaces $(\Omega_i, \mathcal{F}_i, P_i)_{i\in I}$.

Remark 5.5.1 An infinite sequence of independent experiments is such a sequence of experiments when the result of each experiment does not influence on the result of any next experiment.

Assume that i-th $(i \in I)$ random experiment is described by the probability space $(\Omega_i, \mathcal{F}_i, P_i)$. Then an infinite sequence of independent experiments is described by the triplet

$$(\prod_{i\in I}\Omega_i,\prod_{i\in I}\mathcal{F}_i,\prod_{i\in I}P_i).$$

Let consider some examples.

Example 5.5.1 For $i\in N$ we set

$$\Omega_i=\{0,1\},\ \mathcal{F}_i=\{A|A\subseteq\Omega_i\},\ P_i(\{1\})=p,$$

where $0<p<1$.

The product of the infinite family of probability spaces

$$(\prod_{i\in N}\Omega_i,\prod_{i\in N}\mathcal{F}_i,\prod_{i\in N}P_i)$$

is called the infinite-dimensional Bernoulli classical probability space. A real-valued function $\prod_{i\in N}P_i$ is called the infinite-dimensional Bernoulli classical probability .

Example 5.5.2 (Infinite-dimensional multinomial probability space). Assume that an infinite family of probability spaces $(\Omega_i,\mathcal{F}_i,P_i)_{i\in N}$ is defined by :

a) $\Omega_i=\{x_1,\cdots,x_k\}\ (i\in N)$,

b) $\mathcal{F}_i$ is the powerset of Ω_i for arbitrary $i\in N$,

c) $P_i(\{x_j\})=p_j>0,\ i\in N,\ 1\le j\le k,\ \sum_{j=1}^{k}p_j=1$.

Then $(\prod_{i\in N}\Omega_i,\prod_{i\in N}\mathcal{F}_i,\prod_{i\in N}P_i)$ is called the infinite-dimensional multinomial probability space. The real-valued function $\prod_{i\in N}P_i$ is called the infinite-dimensional multinomial probability.

Example 5.5.3 (Infinite-dimensional Borel classical probability measure on infinite-dimensional cube $[0,1]^N$).

An infinite family of probability spaces $(\Omega_i,\mathcal{F}_i,P_i)_{i\in N}$ is defined by:

a) $\Omega_i=[0,1]\ (i\in N)$,

b) $\mathcal{F}_i=\mathcal{B}([0,1])\ (i\in N)$,

c) $P_i=b_1\ (i\in N)$.

$(\prod_{i\in N}\Omega_i,\prod_{i\in N}\mathcal{F}_i,\prod_{i\in N}P_i)$ is called the infinite-dimensional Borel classical probability measure associated with infinite-dimensional cube $[0,1]^N$. Measure $\prod_{i\in N}P_i$ is called the infinite-dimensional Borel classical probability measure on infinite-dimensional cube $[0,1]^N$ and is denoted by b_N.

Example 5.5.4 Assume that an infinite family of functions $(F_i)_{i\in N}$ is defined by

$$(\forall i)(\forall x)(i\in N\ \&\ x\in R\rightarrow F_i(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt\ ).$$

Let P_i be a Borel probability measure on R defined by F_i. Then triplet $(\prod_{i\in N}\Omega_i,\prod_{i\in N}\mathcal{F}_i,\prod_{i\in N}P_i)$ is called the infinite-dimensional Gaussian canonical probability space associated with R^N. The real-valued function $\prod_{i\in N}P_i$ is called the infinite-dimensional Gaussian canonical probability measure defined on R^N and is denoted by Γ_N.

Tests

5.1.There are 10000 trade booths on the territory of the market. The probability that each owner of a booth will get profit 500 lari during one quarter is equal to $0,5$. Further, the probability that the same owner of the booth loose 200 lari during the same quarter is equal to $0,5$. The number of such owners of trade booths , which at the end of the year

1) will loose 800 lari, is equal to

a) 625, b) 670, c) 450, d) 700;

2)will loose 100 lari, is equal to

a) 2500, b) 3000, c) 2000, d) 3500;

3) will get a profit of 600 lari, is equal to

a) 3750, b) 3650, c) 3600, d) 3400;

4)will get a profit of 1300 lari, is equal to

a) 2500, b) 2000, c) 3000, d) 1500;

5) will get a profit of 2000 lari, is equal to

a) 625, b) 650, c) 600, d) 550.

5.2. Wholesale storehouse supplies 20 magazines. It is possible to get order for the next day from each magazine with probability $0,5$.

1) The number of hight probability of orders at the end of the day is equal to

a) 10, b) 11, c) 12, d) 13;

2) The probability corresponding with the number of hight probability of orders at the end of the day is equal to

a) $C_{20}^{10}\frac{1}{2^{20}}$, b) $C_{20}^{10}\frac{1}{2^{10}}$, c) $C_{20}^{10}\frac{1}{2^{30}}$, d) $C_{20}^{5}\frac{1}{2^{20}}$.

5.3. There are three boxes numerated by numbers $1,2,3$. The probabilities, that a particle will be placed in the box $1,2,3$ are equal to $0,3$, $0,4$ and $0,3$, respectively. The probability that out of 6 particles 3 will be placed in box 1, 2 particles will be placed in box 2 and one particle will be find in box 3, is equal to

a) $\frac{3!}{3!2!1!}0,3^4 0,4^2$, *b)* $\frac{4!}{3!2!1!}0,3^4 0,4^2$, *c)* $\frac{5!}{3!2!1!}0,3^4 0,4^2$, *d)* $\frac{6!}{3!2!1!}0,3^4 0,4^2$.

Chapter 6

Random Variables

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

Definition 6.1 *Function $\xi : \Omega \to R$ is called a random variable, if*

$$(\forall x)(x \in R \to \{\omega : \omega \in \Omega,\ \xi(\omega) < x\} \in \mathcal{F}).$$

Example 6.1 *Arbitrary random variable $\xi : \Omega \to R$ can be considered as a definite rule of dispersion of the unit mass of powder Ω on the real axis R, according to which each particle $\omega \in \Omega$ will be placed on particle $A \in R$ with coordinate $\xi(\omega)$.*

Definition 6.2 *Function $I_A : \Omega \to R$ $(A \subset \Omega)$, defined by*

$$I_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \in \overline{A} \end{cases},$$

is called the indicator of set A.

Theorem 6.1 *Let $A \subset \Omega$. Then I_A is a random variable if and only if $A \in \mathcal{F}$.*

Proof. The validity of Theorem 1 follows from the following formula

$$\{\omega : I_A(\omega) < x\} = \begin{cases} \emptyset, & \text{if } x \le 0, \\ \overline{A}, & \text{if } 0 < x \le 1, \\ \Omega, & \text{if } 1 < x. \end{cases}$$

Definition 6.3 *Random variable $\xi : \Omega \to R$ is called a discrete random variable, if there exists a sequence of pairwise disjoint events $(A_k)_{k \in N}$ and a sequence of real numbers $(x_k)_{k \in N}$, such that:*

1) $(\forall k)(k \in N \to x_k \in R, A_k \in \mathcal{F})$;
2) $\cup_{k \in N} A_k = \Omega$;
3) $\xi(\omega) = \sum_{k \in N} x_k I_{A_k}(\omega),\ \omega \in \Omega$.

Definition 6.4 *Random variable $\xi : \Omega \to R$ is called a simple discrete random variable, if there exists a finite sequence of pairwise disjoint events $(A_k)_{1 \le k \le n}$ and a finite sequence of real numbers $(x_k)_{1 \le k \le n}$, such that:*

1) $(\forall k)(1 \le k \le n \to x_k \in R, A_k \in \mathcal{F}))$;
2) $\cup_{k=1}^{n} A_k = \Omega$;
3) $\xi(\omega) = \sum_{k=1}^{n} x_k I_{A_k}(\omega),\ \omega \in \Omega$.

Definition 6.5 *A sequence of random variables $(\xi_k)_{k \in N}$ is called increasing if*

$$(\forall n)(\forall \omega)(n \in N, \omega \in \Omega \to \xi_n(\omega) \le \xi_{n+1}(\omega)).$$

The following theorem gives an interesting information about the structure of non-negative random variables.

Theorem 6.2 *For arbitrary non-negative random variable $\xi : \Omega \to R$ there exists an increasing sequence of simple discrete variables $(\xi_k)_{k \in N}$, such that*

$$(\forall \omega)(\omega \in \Omega \to \xi(\omega) = \lim_{n \to \infty} \xi_n(\omega)).$$

Proof. For arbitrary $n \in N$ we define a simple discrete variable ξ_n by the following formula

$$\xi_n(\omega) = \sum_{k=1}^{n \cdot 2^n} \frac{k-1}{2^n} \cdot I_{\{y : y \in \Omega, \frac{k-1}{2^n} \le \xi(y) < \frac{k}{2^n}\}}(\omega) + n \cdot I_{\{y : y \in \Omega, \xi(y) \ge n\}}(\omega).$$

Clearly,

$$(\forall n)(n \in N \to \ \xi_n(\omega) \le \xi_{n+1}(\omega))$$

and

$$(\forall \omega)(\omega \in \Omega \to \xi(\omega) = \lim_{n \to \infty} \xi_n(\omega)).$$

This ends the proof of theorem.

Theorem 6.3 *For arbitrary random variable $\eta : \Omega \to R$ there exists a sequence of simple discrete variables $(\eta_k)_{k \in N}$, such that*

$$(\forall \omega)(\omega \in \Omega \to \eta(\omega) = \lim_{n \to \infty} \eta_n(\omega)).$$

Proof. For arbitrary random variable $\eta : \Omega \to R$ we have the following representation $\eta = \eta^+ + \eta^-$, where
$\eta^+(\omega) = \max\{\eta(\omega), 0\}$ and $\eta^-(\omega) = \min\{\eta(\omega), 0\}$.

Using Theorem 6.2, for η^+ and $-\eta^-$ there exist increasing sequences of simple discrete variables $(\eta_k^+)_{k \in N}$ and $(\eta_k^-)_{k \in N}$, such that

$$(\forall \omega)(\omega \in \Omega \to \lim_{k \to \infty} \eta_k^+(\omega) = \eta^+(\omega),\ \ \lim_{k \to \infty} \eta_k^-(\omega) = -\eta^-(\omega)).$$

It is easy to show that $(\eta_n)_{n\in N} = (\eta_n^+ - \eta_n^-)_{n\in N}$ is a sequence of simple discrete random variables such that

$$(\forall\omega)(\omega \in \Omega \to \eta(\omega) = \lim_{n\to\infty} \eta_n(\omega)).$$

This ends the proof of theorem.

Tests

6.1. Let ξ and η be discrete random variables for which the following representations

$$\xi(\omega) = \sum_{k\in N} x_k I_{A_k}(\omega), \quad \eta(\omega) = \sum_{m\in N} y_m I_{B_m}(\omega) \ (\ \omega \in \Omega)$$

are valid. Then

1) for random variable $\xi + \eta$ we have

a) $(\xi + \eta)(\omega) = \sum_{k\in N} \sum_{m\in N} (x_k + y_m) I_{A_k \cap B_m}(\omega) \ (\omega \in \Omega)$,

b) $(\xi + \eta)(\omega) = \sum_{k\in N} \sum_{m\in N} x_k y_m I_{A_k \cap B_m}(\omega) \ (\omega \in \Omega)$;

2) for random variable $\xi \times \eta$ we have

a) $(\xi \cdot \eta)(\omega) = \sum_{k\in N} \sum_{m\in N} (x_k + y_m) I_{A_k \cap B_m}(\omega) \ (\omega \in \Omega)$,

b) $(\xi \cdot \eta)(\omega) = \sum_{k\in N} \sum_{m\in N} x_k y_m I_{A_k \cap B_m}(\omega) \ (\omega \in \Omega)$;

3) if $g : R \to R$ is a measurable function, then $g(\xi)$ is a random variable and

a) $g(\xi)(\omega) = \sum_{k\in N} g(x_k) I_{A_k}(\omega)$,

b) $g(\xi)(\omega) = \sum_{k\in N} g^{-1}(x_k) I_{A_k}(\omega)$;

4) the following formula is valid

a) $\sin(\xi)(\omega) = \sum_{k\in N} \sin(x_k) I_{A_k}(\omega)$,

b) $\sin(\xi)(\omega) = \sum_{k\in N} \arcsin(x_k) I_{A_k}(\omega)$.

6.2. Let $(A_k)_{k\in N}$ be a sequence of events and let ξ be a random variable. Then

1)

a) $\xi^{-1}(\cup_{k\in N} A_k) = \cup_{k\in N} \xi^{-1}(A_k)$,

b) $\xi^{-1}(\cup_{k\in N} A_k) = \cap_{k\in N} \xi^{-1}(A_k)$;

2)

a) $\xi^{-1}(\cap_{k\in N} A_k) = \cap_{k\in N} \xi^{-1}(A_k)$,

b) $\xi^{-1}(\cap_{k\in N} A_k) = \cup_{k\in N} \xi^{-1}(A_k)$;

3)

a) $\Omega \setminus \xi^{-1}(A_k) = \xi^{-1}(\Omega \setminus A_k)$,

b) $\Omega \setminus \xi^{-1}(A_k) = \xi^{-1}(A_k)$.

6.3.

1) If $|\xi|$ is a random variable, then

a) ξ is a random variable;
b) it is possible that ξ is not a random variable;

2) if ξ is a random variable, then

a) ξ^+ is a random variable;
b) it is possible that ξ^+ is not a random variable;

3) Let ξ and η be random variables and let A be any event. If $\Theta(\omega) = \xi(\omega)I_A(\omega) + \eta(\omega)I_{\overline{A}}(\omega)$ $(\omega \in \Omega)$, then
a) Θ is a random variable,

b) It is possible that Θ is not a random variable.

Chapter 7

Function of random variable distribution

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\xi : \Omega \to R$ be a random variable.

Definition 7.1 *Function F_ξ, defined by*

$$(\forall x)(x \in \overline{R} \to F_\xi(x) = P(\{\omega : \xi(\omega) \leq x\})),$$

where

$$\overline{R} = \{-\infty\} \ \cup \ R \ \cup \ \{+\infty\},$$

is called a distribution function of random variable ξ.

Here we consider some properties of distribution functions.

Theorem 7.1 $F_\xi(+\infty) = \lim_{x \to +\infty} F_\xi(x) = 1.$

Proof. *Let consider an increasing sequence of real numbers $(x_k)_{k \in N}$ such that $\lim_{k \to \infty} x_k = +\infty$. On the one hand, we have*

$$\{\omega : \xi(\omega) \leq x_k\} \subseteq \{\omega : \xi(\omega) \leq x_{k+1}\} \ (k \in N).$$

On the other hand, we have $\cup_{k \in N} \{\omega : \xi(\omega) \leq x_k\} = \Omega$. Using the property of continuity from below we get

$$\lim_{k \to \infty} P(\{\omega : \xi(\omega) \leq x_k\}) = P(\cup_{k \in N} \{\omega : \xi(\omega) \leq x_k\}) = P(\Omega) = 1,$$

i.e. $\lim_{x \to +\infty} F_\xi(x) = 1.$
Note that $F_\xi(+\infty) = P(\{\omega : \xi(\omega) \leq +\infty\}) = P(\Omega) = 1.$ *Finally we get*

$$F_\xi(+\infty) = \lim_{x \to +\infty} F_\xi(x) = 1.$$

This ends the proof of theorem.

Theorem 7.2 $F_\xi(-\infty) = \lim_{x \to -\infty} F_\xi(x) = 0.$

43

Proof. Note that

$$F_\xi(-\infty) = P(\{\omega : \xi(\omega) \le -\infty\}) = P(\emptyset) = 0.$$

Let consider a decreasing sequence of real numbers $(x_k)_{k\in N}$ such that $\lim_{k\to\infty} x_k = -\infty$. It is easy to check the validity of the following conditions:

1) $\{\omega : \xi(\omega) \le x_{k+1}\} \subseteq \{\omega : \xi(\omega) \le x_k\}$ $(k \in N)$,

2) $\cap_{k\in N} \{\omega : \xi(\omega) \le x_k\} = \emptyset$.

Using the property of the continuity from above of P, we get

$$\lim_{k\to\infty} P(\{\omega : \xi(\omega) \le x_k\}) = P(\cap_{k\in N}\{\omega : \xi(\omega) \le x_k\}) = P(\emptyset) = 0,$$

i.e.,

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

This ends the proof of theorem.

Theorem 7.3 *Distribution function $F_\xi(x)$ is an increasing function.*

Proof. Let, $x_1 < x_2$. Let show the validity of the following non-strict inequality $F_\xi(x_1) \le F_\xi(x_2)$. Indeed, using the validity of the following inclusion

$$\{\omega : \xi(\omega) \le x_1\} \subseteq \{\omega : \xi(\omega) \le x_2\}$$

and Property 2.5 (cf. Chapter 2), we have

$$P(\{\omega : \xi(\omega) \le x_1\}) \le P(\{\omega : \xi(\omega) \le x_2\}),$$

which is equivalent to condition $F_\xi(x_1) \le F_\xi(x_2)$.

This ends the proof of theorem.

Theorem 7.4 *Distribution function $F_\xi(x)$ is continuous from the right , i. e., for arbitrary sequence of real numbers $(x_k)_{k\in N}$ for which $x_k > x$ $(k \in N)$ and $\lim_{k\to\infty} x_k = x$, the following condition*

$$\lim_{k\to\infty} F_\xi(x_k) = F_\xi(x)$$

is fulfilled.

Proof. Without loss of generality, we can assume that $(x_k)_{k\in N}$ is a decreasing sequence. Then

$$\{\omega : \xi(\omega) \le x\} = \cap_{k\in N}\{\omega : \xi(\omega) \le x_k\},$$

$$\{\omega : \xi(\omega) \le x_{k+1}\} \subseteq \{\omega : \xi(\omega) \le x_k\} \ (k \in N).$$

Hence, using the property of the continuity from above of P, we obtain

$$\lim_{k\to\infty} P(\{\omega : \xi(\omega) \le x_k\}) = P(\cap_{k\in N}\{\omega : \xi(\omega) \le x_k\}) = P(\{\omega : \xi(\omega) \le x\}),$$

which is equivalent to condition $\lim_{k\to\infty} F_\xi(x_k) = F_\xi(x)$.

This ends the proof of theorem.

Let ξ be a discrete random variable, i.e., there exist an infinite family of pairwise disjoint events $(A_k)_{k\in N}$ and an infinite family of real numbers $(x_k)_{k\in N}$, such that:
1) $(\forall k)(k \in N \to x_k \in R,\ A_k \in \mathcal{F})$;
2) $\cup_{k\in N} A_k = \Omega$;
3) $\xi(\omega) = \sum_{k\in N} x_k I_{A_k}(\omega),\ \omega \in \Omega$.
Then the distribution function of ξ is calculated by :

$$F_\xi(x) = \sum_{x_k \le x} P(A_k).$$

Remark 7.1 Assume that in column A of the "Excel" table we have entered values $x_1, \cdots, x_n$ of simple discrete random variable ξ; Assume also that in column B of "Excel" table we have entered the corresponding probabilities $p_1, \cdots, p_n$. Then the statistical function $\mathrm{PROB}(x_1 : x_n; p_1 : p_n; y_1; y_2)$ calculates the following probability

$$P(\{\omega : \omega \in \Omega \ \&\ y_1 \le \xi \le y_2\}).$$

For example, if the distribution of ξ has the following form

$A(\xi(\omega) = x_k)$	$B(= P(\{\omega : \xi(\omega) = x_k\})$
2	0,2
5	0,2
6	0,2
7	0,4

,

thentheprobabilitythatrandomvariableξ will obtain a value in interval $[3,6;5]$ is calculated with

$$\mathrm{PROB}(A_1 : A_4; B_1 : B_4; 3,6; 5) = 0,4.$$

Note that to construct the distribution function of ξ it is not necessary to know what values obtain the random variable ξ on the elements of Ω. In this direction it is sufficient to know probabilities of possible values of ξ.

ξ	x_1	x_2	$\cdots$
P	p_1	p_2	$\cdots$

,

where $(\forall k)(k \in N \to p_k = P(A_k))$.

Let consider some examples.

Example 7.1 (Poisson [1] distribution). We say that a discrete random variable $\xi : \Omega \to R$, defined by

$$\xi(\omega) = \sum_{n\in N} n \cdot I_{A_n}(\omega)\ (\omega \in \Omega),$$

[1] *Poisson, Semion Denis* (21.6.1781 - 25.4.1840)-French mathematician, physician, the member of Paris Academy of Sciences (1812), the honourable member of Petersburg Academy of Sciences (1826).

generates a Poisson distribution with parameter λ $(\lambda > 0)$ if the following condition

$$P(A_n) = \frac{\lambda^n}{n!}e^{-\lambda} \ (n \in N)$$

holds,i.e.,

$$P(\{\omega : \xi(\omega) = n\}) = \frac{\lambda^n}{n!}e^{-\lambda} \ (n \in N).$$

Poisson distribution function $F(x, \lambda)$ with parameter λ is defined by the following formula

$$F(x, \lambda) = \sum_{n \leq x} \frac{\lambda^n}{n!}e^{-\lambda} \ (x \in R).$$

Remark 7.2 **POISSON**$(k; \lambda; 0)$ calculates the probability that the Poisson random variable with parameter λ will get value k. For example, POISSON$(0; 0,2; 0) = 0,818730753$.

Remark 7.3 **POISSON**$(k; \lambda; 1)$ calculates the probability that the Poisson random variable with parameter λ will get an integer value in interval $[0; k]$. For example,

$$\text{POISSON}(2; 0,2; 1) = 0,998851519.$$

Example 7.2 (The geometric distribution). We say that a discrete random variable $\xi : \Omega \to R$, defined by

$$\xi(\omega) = \sum_{n \in N} n \cdot I_{A_n}(\omega) \ \ (\omega \in \Omega),$$

generates the geometric distribution with parameter q $(0 \leq q \leq 1)$ if the condition

$$P(A_n) = (1-q)q^{n-1} \ (n \in N)$$

holds, i.e.,

$$P(\{\omega : \xi(\omega) = n\}) = (1-q)q^{n-1} \ \ (n \in N).$$

The geometric distribution with parameter q is defined by the following formula

$$F_q(x) = \sum_{n \leq x}(1-q)q^{n-1} \ (x \in R).$$

Example 7.3 (Leibniz [2] **distribution**). We say that a discrete random variable $\xi : \Omega \to R$, defined by

$$\xi(\omega) = \sum_{n \in N} n \cdot I_{A_n}(\omega) \ \ (\omega \in \Omega),$$

[2] *Leibniz, Gottfried Wilhelm* (1.7.1646, -14.11.1716)-German mathematician, the member of London Royal Society (1673), the member of Paris Academy of Sciences (1700)

generates the Leibniz distribution if the condition

$$P(A_n) = \frac{1}{n \cdot (n+1)} \ (n \in N)$$

holds, i.e.

$$P(\{\omega : \xi(\omega) = n\}) = \frac{1}{n \cdot (n+1)} \ (n \in N).$$

The Leibniz distribution function is calculated by the following formula:

$$F(x) = \sum_{n \le x} \frac{1}{n \cdot (n+1)} =$$

$$= \begin{cases} 0, & \text{if } x < 1 \\ 1 - \frac{1}{[x]+1}, & \text{if } x \ge 1 \end{cases},$$

where $[x]$ denotes an integer part of x.

Example 7.4 (hypergeometric distribution). A simple discrete random variable

$$\xi(\omega) = \sum_{k=1}^{n} k I_{A_k}(\omega) \quad (\omega \in \Omega)$$

is called distributed by hypergeometric law with parameters (n, a, A) if

$$P(A_k) = \frac{C_a^k \cdot C_{A-a}^{n-k}}{C_A^n} \ (k = 0, 1, \cdots, n),$$

where $0 \le a \le A, \ 1 \le n \le A$.

The hypergeometric distribution with parameters (n, a, A) is denoted by $F_{(n,a,A)}(x)$ and is defined by

$$F_{n,a,A}(x,p) = \sum_{k \le x} \frac{C_a^k \cdot C_{A-a}^{n-k}}{C_A^n}.$$

Remark 7.4 **HYPERGEOMDIST**$(k; n; a; A)$ calculates the value

$$\frac{C_n^k \times C_{A-n}^{n-k}}{C_A^n}.$$

For example, HYPERGEOMDIST$(1; 4; 20; 30) = 0,087575$.

Example 7.5 (binomial distribution). A simple discrete random variable

$$\xi(\omega) = \sum_{k=1}^{n} k I_{A_k}(\omega) \quad (\omega \in \Omega)$$

is called distributed by binomial law with parameter (n,p) if

$$P(A_k) = C_n^k \cdot p^k(1-p)^{n-k},$$

where $0 < p < 1$, $0 \le k \le n$, i.e.,

$$P(\{\omega : \xi(\omega) = k\}) = C_n^k \cdot p^k(1-p)^{n-k}.$$

The binomial distribution with parameter (n,p) is denoted by $F_n(x,p)$ and is defined by

$$F_n(x,p) = \sum_{k \le x} C_n^k \cdot p^k(1-p)^{n-k}.$$

Remark 7.5 BINOMDIST$(k;n;p;0)$ calculates the value

$$C_n^k p^k(1-p)^{n-k}.$$

For example, BINOMDIST$(3;10;0,5;0) = 0,1171875$.

BINOMDIST$(k;n;p;1)$ calculates the sum

$$\sum_{i=0}^{k} C_n^i p^i(1-p)^{n-i}.$$

For example, BINOMDIST$(3;10;0,5;1) = 0,171875$.

Remark 7.6 The random variable distributed by the binomial law with parameter $(1,p)$ is called also a random variable distributed by the Bernoulli' law with parameter p. It can be proved that the random variable distributed by the Binomial law with parameter (n,p) can be presented as a sum of n independent random variables each of them is distributed by the Bernoulli law with parameter p.

Definition 7.2 Random variable $\xi : \Omega \to R$ is called absolutely continuous [3] if there exists a non-negative function $f_\xi : R \to R^+$ such that

$$(\forall x)(x \in R \to F_\xi(x) = \int_{-\infty}^{x} f_\xi(x)dx),$$

where $R^+ = [0,+\infty[$.

Function $f_\xi(x)$ $(x \in R)$ is called a density function of random variable ξ.

Theorem 7.5 *Let $f_\xi : R \to R$ be a density function of random variable $\xi : \Omega \to R$. Then*

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

Proof. Since $\lim_{L\to+\infty} F_\xi(L) = 1$, we have $\lim_{L\to+\infty} \int_{-\infty}^{L} f_\xi(x)dx = 1$. This relation means the validity of the following equality $\int_{-\infty}^{+\infty} f_\xi(x)dx = 1$.

[3]Note that the density function of the absolutely continuous random variable is defined exactly until null sets (in the Lebesgue sense) of R. We recall the reader that $X \subset R$ is null-set(in Lebesgue sense) if for arbitrary $\epsilon > 0$ there exists sequence $(]a_k,b_k[)_{k\in N}$ of open intervals such that $X \subset \cup_{k\in N}]a_k,b_k[$ and $\sum_{k\in N}(b_k - a_k) < \epsilon$.

This ends the proof of theorem.

Theorem 7.6 *Let F_ξ be a distribution function of an absolutely continuous random variable ξ. Then for arbitrary real numbers x and y $(x < y)$ we have*

$$P(\{\omega : x < \xi(\omega) \leq y\}) = F_\xi(y) - F_\xi(x),$$

If ξ is absolutely continuous random variable and f_ξ is its density function, then

$$P(\{\omega : x < \xi(\omega) \leq y\}) = \int_x^y f_\xi(s)ds.$$

Proof.

$$P(\{\omega : x < \xi(\omega) \leq y\}) = P(\{\omega : \xi(\omega) \leq y\} \setminus \{\omega : \xi(\omega) \leq x\}) =$$

$$= P(\{\omega : \xi(\omega) \leq y\}) - P(\{\omega : \xi(\omega) \leq x\}) = F_\xi(y) - F_\xi(x).$$

If $F_\xi(t) = \int_{-\infty}^t f_\xi(s)ds$, then

$$F_\xi(y) - F_\xi(x) = \int_{-\infty}^y f_\xi(s)ds - \int_{-\infty}^x f_\xi(s)ds = \int_x^y f_\xi(s)ds.$$

This ends the proof of theorem.

Remark 7.7 If f_ξ and F_ξ are the density function and the distribution functions respectively, then almost everywhere on R we have

$$\frac{dF_\xi(x)}{dx} = f_\xi(x),$$

i.e., $l_1\{x : x \in R, \frac{dF_\xi(x)}{dx} \neq f_\xi(x)$ and $\frac{dF_\xi(x)}{dx}$ does not exist$\} = 0$, where l_1 denotes one-dimensional Lebesgue measure on R.

Example 7.6 (Normal distribution). Absolutely continuous random variable $\xi : \Omega \to R$ is called normally distributed with parameters (m, σ^2) $(m \in R,\ \sigma > 0)$ if

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

The density function and the distribution function of the normally distributed random variable with parameters (m, σ^2) are denoted by $\phi_{(m,\sigma^2)}$ and $\Phi_{(m,\sigma^2)}$, respectively, i.e.,

$$\phi_{(m,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \ (x \in R),$$

$$\Phi_{(m,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-m)^2}{2\sigma^2}} dt \ (t \in R).$$

When $m = 0$ and $\sigma = 1$, they respectively are denoted as $\phi = \phi_{(0,1)}$ and $\Phi = \Phi_{(0,1)}$. ϕ and Φ are called the density function and the distribution function of the standard normally distributed random variable respectively.

Remark 7.8 **NORMDIST**$(x; m; \sigma; 0)$ calculates the function

$$\phi_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

For example, NORMDIST$(0; 0; 1; 0) = 0,3989428$.

NORMDIST$(x; m; \sigma; 1)$ calculates the integral

$$\Phi_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(t-m)^2}{2\sigma^2}} dt.$$

For example, NORMDIST$(0; 0; 1; 1) = 0,5$.

Example 7.6 (The uniform distribution). Absolutely continuous random variable $\xi : \Omega \to R$ is called uniformly distributed on the interval $[a, b]$ $(a < b)$ if

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

Distribution function F_ξ of the random variable uniformly distributed on $[a, b]$ is defined by

$$F_\xi(x) = \begin{cases} 0, & \text{if } x < a; \\ \frac{x-a}{b-a}, & \text{if } x \in [a, b]; \\ 1, & \text{if } x > b. \end{cases}$$

Example 7.7 (Cauchy [4] distribution). We say that an absolutely continuous random variable $\xi : \Omega \to R$ is distributed by the Cauch law, if

$$f_\xi(x) = \frac{1}{\pi(1 + x^2)} \ (x \in R).$$

Its distribution function is defined by

$$F_\xi(x) = \int_{-\infty}^{x} \frac{1}{\pi(1 + t^2)} dt = \frac{1}{2} + \frac{1}{\pi} arctg(x) \ (x \in R).$$

Example 7.8 (Exponential distribution). Absolutely continuous random variable $\xi : \Omega \to R$ is distributed by the exponential law, if

$$f_\xi(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0; \\ 0, & \text{if } x < 0. \end{cases}$$

Its distribution function F_ξ is defined by

$$F_\xi(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{if } x \geq 0; \\ 0, & \text{if } x < 0. \end{cases}$$

[4] *Cauchy, Augustin Louis* (21.8.1789, - 23.5.1857) -French mathematician, the member of Paris Academy of Sciences (1816), the honourable member of Petersburg Academy of Sciences(1831).

Remark 7.9 **EXPONDIST**$(x;\lambda;0)$ calculates value $\lambda e^{-\lambda x}$ for $x > 0$ and $\lambda > 0$. For example, EXPONDIST$(4;3;0) = 1,84326$.

EXPONDIST$(x;\lambda;1)$ calculates value $1-e^{-\lambda x}$ for $x > 0$ and $\lambda > 0$. For example, EXPONDIST$(4;3;1) = 0,999993856$.

Example 7.9 (Singular distribution). Let consider closed interval $[0,1]$ and let define a sequence of functions constructed by G.Cantor [5].

Let divide interval $[0,1]$ into three equal parts and define function F_1 by

$$F_1(x) = \begin{cases} \frac{1}{2}, & \text{if } x \in]\frac{1}{3}, \frac{2}{3}[; \\ 0, & \text{if } x = 0; \\ 1, & \text{if } x = 1. \end{cases}$$

We continue its values on other points of $[0,1]$ by linear interpolation.

Further, let consider the division of intervals $[0, \frac{1}{3}]$ and $[\frac{2}{3}, 1]$ into three equal parts and define F_2 by

$$F_2(x) = \begin{cases} \frac{1}{2}, & \text{if } x \in]\frac{1}{3}, \frac{2}{3}[; \\ \frac{1}{4}, & \text{if } x \in]\frac{1}{9}, \frac{2}{9}[; \\ \frac{3}{4}, & \text{if } x \in]\frac{7}{9}, \frac{8}{9}[; \\ 0, & \text{if } x = 0; \\ 1, & \text{if } x = 1. \end{cases}$$

Analogously, we continue the values of F_2 on other points of $[0,1]$ by linear interpolation.

If we shall continue this process, then we shall get a sequence of functions $(F_n)_{n\in N}$, which tends uniformly to concrete continuous function F on $[0,1]$, the increase points [6] of which is null-set in the Lebesgue sense. Indeed, we get that the Lebesgue measure of the union of intervals

$$]\frac{1}{2}, \frac{2}{3}[,\ ]\frac{1}{9}, \frac{2}{9}[,\ ]\frac{7}{9}, \frac{8}{9}[, \cdots,$$

on which function F is constant, is

$$\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \cdots = \frac{1}{3}\sum_{n=0}^{\infty} (\frac{2}{3})^n = 1.$$

F is called a Cantor function.

Let consider one construction of the random variable, whose distribution function coincides with Cantor function F.

We set

$$(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), b_1).$$

Let define a sequence of functions

[5] *Cantor, George* (19.2.(3.3).1845 -6.1.1918)-German mathematician, professor of Gales University (1879-1913). He had proved that a real numbers axis is not countable.

[6] x is called a point of increment for function F if $F(x+\epsilon) - F(x-\epsilon) > 0$ for arbitrary $\epsilon > 0$

$$(\xi_{\frac{k}{2^n}})_{n \in N,\ 1 \le k \le n,\ \&\ k \in 2N+1} = (\xi_i)_{i \in I}$$

by the following formulas

$$\xi_{\frac{1}{2}}(\omega) = \frac{1}{3} I_{\{\frac{1}{2}\}}(\omega),\ (\omega \in \Omega),$$

$$\xi_{\frac{1}{4}}(\omega) = \frac{1}{9} I_{\{\frac{1}{4}\}}(\omega),\ (\omega \in \Omega),$$

$$\xi_{\frac{3}{4}}(\omega) = \frac{1}{9} I_{\{\frac{3}{4}\}}(\omega),\ (\omega \in \Omega),$$

$$\xi_{\frac{1}{8}}(\omega) = \frac{1}{27} I_{\{\frac{1}{8}\}}(\omega),\ (\omega \in \Omega),$$

$$\xi_{\frac{3}{8}}(\omega) = \frac{1}{27} I_{\{\frac{3}{8}\}}(\omega),\ (\omega \in \Omega),$$

$$\xi_{\frac{5}{8}}(\omega) = \frac{1}{27} I_{\{\frac{5}{8}\}}(\omega),\ (\omega \in \Omega),$$

$$\xi_{\frac{7}{8}}(\omega) = \frac{1}{27} I_{\{\frac{7}{8}\}}(\omega),\ (\omega \in \Omega),$$

and so on.

We define $\xi_{Cantor} : \Omega \to R$ by the following formula

$$\xi_{Cantor}(\omega) = \sum_{i \in I,\ i \le \omega} \xi_i(\omega).$$

It is easy to show that the distribution function generated by ξ_{cantor} coincides with Cantor function F.

Definition 7.3. *A continuous distribution function, whose points of the increment have a Lebesgue measure zero, is called singular. The corresponding random variable is also called singular.*

Theorem 7.7 *Arbitrary distribution function F admits the following representation*

$$F(x) = p_1 \cdot F_1(x) + p_2 \cdot F_2(x) + p_3 \cdot F_3(x)\ (x \in R),$$

where $F_1,\ F_2,\ F_3$ are distribution functions generated by a discrete, an absolutely continuous and a singular random variables, respectively and $p_1,\ p_2,\ p_3$ are such non-negative real numbers that

$$p_1 + p_2 + p_3 = 1.$$

Theorem 7.8 *Let F_ξ be a distribution function of ξ and $a > 0,\ b \in R$. Then the distribution function of $\eta = a\xi + b$ is calculated by*

$$F_\eta(x) = F_\xi(\frac{x-b}{a})\ (x \in \overline{R}).$$

Proof. *Note, that*

$$F_\eta(x) = P(\{\omega : a\xi(\omega) + b \le x\}) = P(\{\omega : \xi(\omega) \le \frac{x-b}{a}\}) = F_\xi(\frac{x-b}{a}).$$

Tests

7.1. The distribution law of random variable $\xi(\omega) = \sum_{k=1}^{4} x_k I_{A_k}(\omega)$ $(\omega \in \Omega)$ is given by the following table

ξ	-1	0	4	5
P	$0,2$	$0,3$	$0,1$	$0,4$

.

Then

1) $F_\xi(-3)$ is equal to

a) $0,2$, b) $0,3$, c) $0,1$, d) 0;

2) $F_\xi(-1)$ is equal to

a) $0,2$, b) $0,3$, c) $0,1$, d) 0;

3) $F_\xi(-0,3)$ is equal to

a) $0,2$, b) $0,3$, c) $0,1$, d) 0;

4) $F_\xi(4)$ is equal to

a) $0,6$, b) $0,4$, c) 1, d) $0,8$;

5) $F_\xi(6)$ is equal to

a) $0,6$, b) $0,4$, c) 1, d) $0,8$.

7.2. The distribution function of ξ is defined by

$$F_\xi(x) = \begin{cases} a, & x < 0; \\ bx, & 0 \le x < 1; \\ c, & x \ge 1. \end{cases}$$

Then

a) $a = 1, b = 0, c = 0$; b) $a = 0, b = 1, c = 1$;
c) $a = 0, b = 0, c = 1$; d) $a = 1, b = 1, c = 0$;

7.3.The probability that event A will occur in partial experiment is equal to 0,3. Let ξ be the number of experiments in the three independent experiments, when the event A occurred. Then the distribution of ξ is given by the following table

a)

ξ	0	1	2	3
P	$0,343$	$0,441$	$0,189$	$0,027$

,

b)

ξ	0	1	2	3
P	$0,343$	$0,441$	$0,179$	$0,037$

.

7.4. A shot gets 5 points if he strucks a target and loses 2 points in other case. The probability that the shot strucks a target is equal to 0,5. The law of distribution of collected points ξ in 4 shots is given by the following table

a)

ξ	-8	-1	6	13	20
P	$0,0625$	$0,25$	$0,375$	$0,25$	$0,0625$

,

b)

ξ	-8	-1	6	13	20
P	$0,0625$	$0,225$	$0,375$	$0,225$	$0,0625$

.

7.5. The complect of 10 details contains 8 non-standard details. We accidentally choose 2 details. Then the law of distribution of number ξ of standard details in our probability sampling is given by the following table

a)

ξ	0	1	2
P	$\frac{1}{45}$	$\frac{16}{45}$	$\frac{28}{45}$

,

b)

ξ	0	1	2
P	$\frac{2}{45}$	$\frac{14}{45}$	$\frac{29}{45}$

.

7.6. The probability that the price of goods will increase or decrease by 1 lari during one unit of time is equal to 0,5 and 0,5, respectively. An initial price of goods is 10 lari. Then the distribution law of price ξ after 4 unites of time is given by the following table

a)

ξ	6	8	10	12	14
P	$0,0625$	$0,25$	$0,375$	$0,25$	$0,0625$

,

b)

ξ	6	8	10	12	14
P	$0,065$	$0,25$	$0,375$	$0,25$	$0,065$

.

7.7. A particle is placed at the origin of the real axis. The probabilities of shifting to the right or to the left along the real axis during one unit of time are equal(=0,5). The distribution law of states ξ of the particle after 4 unit of time is given by the following table

a)

ξ	-4	-2	0	2	4
P	$0,0625$	$0,25$	$0,375$	$0,25$	$0,0625$

,

b)

ξ	-4	-2	0	2	4
P	$0,0625$	$0,245$	$0,385$	$0,245$	$0,0625$

.

7.8. Let ξ be a Poisson random variable with parameter $\lambda = 1$. Then the probability that

1) ξ will obtain a value in the interval [2,5; 5,5] is equal to

a) $0,306760$, b) $0,13455$, c) $0,11213$, d) $0,28111$;

2) $3\xi + 4$ will obtain a value in the interval [6,5; 7,5] is equal to

a) $0,367879$, b) $0,13894$, c) $0,13121$, d) $0,28991$.

7.9. Let ξ be a random variable uniformly distributed on $[3,10]$. Then

1) $F_\xi(4)$ is equal to

a) $\frac{1}{7}$, b) $\frac{1}{8}$, c) $\frac{1}{9}$, d) $\frac{1}{10}$;

2) the probability that ξ will obtain a value in the interval [2,5; 5,5] is equal to

a) $\frac{5}{7}$, b) $\frac{5}{8}$, c) $\frac{5}{9}$, d) $0,5$;

3) the probability that $5\xi + 5$ will obtain a value in the interval [5;10] is equal to

a) 0, b) 1, c) $0,5$, d) $0,8$.

7.10. Let ξ be an exponential random variable with parameter $\lambda(\lambda > 0)$.

1) If the probability that ξ will obtain a value in the interval $[0, a]$ is equal to $\frac{2}{3}$, then

a) $a = \frac{\ln(3)}{\lambda}$, b) $a = \frac{\ln(4)}{\lambda}$, c) $a = \frac{\ln(5)}{\lambda}$, d) $a = \frac{\ln(6)}{\lambda}$;

2)The probability that $3\xi - 4$ will obtain a value in the interval [-5;5] is equal to

a) $1 - e^{-3\lambda}$, b) $1 - e^{-4\lambda}$, c) $1 - e^{-5\lambda}$, d) $1 - e^{-6\lambda}$.

7.11. Let ξ be a standard normal random variable.

1) If the probability that ξ will obtain a value in the interval $[-a, a]$ is equal to 0,99, then

a) $a = 2,37$, b) $a = 2,57$, c) $a = 2,77$, d) $a = 2,97$;

2) The probability that $3\xi + 8$ will obtain a value in the interval $(-5, 5)$ is equal to

a) $0,8413$, b) $0,7413$, c) $0,6413$, d) $0,5413$.

Chapter 8

Mathematical expectation and mathematical variance

Let $(\Omega, \mathcal{F}, P)$ be a probability space and ξ be a simple discrete random variable, i.e.,

$$(\forall\omega)(\omega \in \Omega \to \xi(\omega) = \sum_{k=1}^{n} x_k \cdot I_{A_k}(\omega)),$$

where $x_k \in R$ $(1 \leq k \leq n)$ and $(A_k)_{1 \leq k \leq n}$ is the complete system of representatives.

Definition 8.1 A mathematical expectation of the simple random variable ξ is denoted by $M\xi$ and is defined by

$$M\xi = \sum_{k=1}^{n} x_k \cdot P(A_k).$$

Remark 8.1 Assume that in column A of "Excel" table we have entered values $x_1, \cdots, x_n$ of simple discrete random variable ξ; Assume also that in column B of "Excel" table we have entered the corresponding probabilities $p_1, \cdots, p_n$. Then the statistical function SUMPRODUCT($x_1 : x_n; p_1 : p_n$) calculates mathematical expectation of $M\xi$.

For example, if the distribution of ξ has the following form

$A(\xi(\omega) = x_k)$	$B(= P(\{\omega : \xi(\omega) = x_k\})$
2	$0,2$
5	$0,2$
6	$0,2$
7	$0,4$

,

then $M\xi = \text{PROB}(A_1 : A_4; B_1 : B_4;) = 5,4$.

Assume that η be an arbitrary random variable. Following Theorem 3(cf. §6), there exists a sequence $(\eta_n)_{n \in N}$ of simple random variables such that

$$(\forall\omega)(\omega \in \Omega \to \eta(\omega) = \lim_{n \to \infty} \eta_n(\omega)).$$

Definition 8.2 If there exists a finite limit $\lim_{n\to\infty} M\eta_n$, then this limit is called a mathematical expectation of η and is denoted by $M\eta$ (or $\int_\Omega \eta(\omega)dP(\omega)$).

It can be proved that if there exists a finite limit $\lim_{n\to\infty} M\eta_n$, then this limit is same for arbitrary sequence of simple random variables tending to η, which means a correctness of Definition 2.

Agreement. In the sequel we consider such a class of random variables each element ξ of which satisfy the conditions: $M(\xi) < \infty$ and $M(\xi^2) < \infty$.

Theorem 8.1 *If f_ξ is a density function of an absolutely continuous random variable, then*

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

Definition 8.3 *Value $M(\xi - M\xi)^2$ is called mathematical variance of ξ and is denoted by $D\xi$.*

Definition 8.4 *Value $\sqrt{D\xi}$ is called a mean absolute deviation and is denoted by $\sigma(\xi)$.*

Let consider some properties of mathematical expectation and mathematical variance of random variables.

Theorem 8.2 Let $\xi(\omega) = c$ $(\omega \in \Omega,\ c \in R)$. Then $M\xi = c$.

Proof. $\xi(\omega) = c \cdot I_\Omega(\omega)$. Following the definition of the expectation of the simple discrete random variable, we have

$$M\xi = c \cdot P(\Omega) = c.$$

Theorem 8.3 $M(\xi+\eta) = M\xi + M\eta$, *(i.e., mathematical expectation of the sum of two random variables is equal to the sum of expectations of corresponding random variables).*

Proof. Using the approximation property of a random variable by a sequence of simple discrete random variables and by the definition of the expectation of a random variable, it is sufficient to prove this theorem in the case of two simple discrete random variables. Now assume that ξ and η be simple random variables, i.e.

$$\xi(\omega) = \sum_{k=1}^{p} x_k \cdot I_{A_k}(\omega),\ A_k \cap A_m = \emptyset,\ 1 \le k < m \le p,$$

$$\cup_{k=1}^{p} A_k = \Omega,\ x_k \in R,\ k, m, p \in N,$$

$$\eta(\omega) = \sum_{n=1}^{q} y_n \cdot I_{B_n}(\omega),\ B_k \cap B_m = \emptyset,\ 1 \le k < m \le q,$$

$$\cup_{n=1}^{q} B_n = \Omega,\ y_n \in R,\ k, m, q \in N,$$

Note that

$$(\xi+\eta)(\omega)=\sum_{k=1}^{p}\sum_{n=1}^{q}(x_k+y_n)\cdot I_{A_k\cap B_n}(\omega)\ \ (\omega\in\Omega),$$

It follows

$$M(\xi+\eta)=\sum_{k=1}^{p}\sum_{n=1}^{q}(x_k+y_n)\cdot I_{A_k\cap B_n}(\omega)=\sum_{k=1}^{p}\sum_{n=1}^{q}(x_k+y_n)\cdot P(A_k\cap B_n)$$

$$=\sum_{k=1}^{p}x_k\sum_{n=1}^{q}P(A_k\cap B_n)+\sum_{n=1}^{q}y_n\sum_{k=1}^{p}P(A_k\cap B_n)$$

$$=\sum_{k=1}^{p}x_kP(A_k)+\sum_{n=1}^{q}y_nP(B_n)=M\xi+M\eta.$$

The theorem is proved.

Definition 8.5 *Two simple discrete random variables ξ and η are called independent, if*

$$P(\{\omega:\xi(\omega)=x_k,\ \eta(\omega)=y_n\})=P(\{\omega:\xi(\omega)=x_k\})\cdot P(\{\omega:\ \eta(\omega)=y_n\}),$$

where $1\le k\le p,\ 1\le n\le q$.

Definition 8.6 *Two random variables ξ and η are called independent if*

$$P(\{\omega:\xi(\omega)\le x,\ \eta(\omega)\le y\})=P(\{\omega:\xi(\omega)\le x\})\cdot P(\{\omega:\ \eta(\omega)\le y\}),$$

where $x,\ y\ \in R$.

Remark 8.1 The definitions 5 and 6 are equivalent for simple discrete random variables .

Theorem 8.4 *Let ξ and η be independent random variables. Then there exist two sequences $(\xi_n)_{n\in N}$ and $(\eta_n)_{n\in N}$ of simple discrete random variables such that :*

1) ξ_n and η_n are independent for $n\in N$

2) $(\xi\cdot\eta)(\omega)=\lim_{n\to\infty}\xi_n(\omega)\cdot\eta_n(\omega)\ \ (\omega\in\Omega)$.

Theorem 8.5 *If ξ and η are independent simple discrete random variables, then*

$$M(\xi\cdot\eta)=M\xi\cdot M\eta,$$

i.e., mathematical expectation of the product of two independent simple discrete random variables is equal to the product of expectations of corresponding simple random variables.

Proof. Note that

$$(\xi\cdot\eta)(\omega)=\sum_{k=1}^{p}\sum_{n=1}^{q}(x_k\cdot y_n)\cdot I_{A_k\cap B_n}(\omega)\ \ \ (\omega\in\Omega).$$

It follows

$$M(\xi\cdot\eta)=M(\sum_{k=1}^{p}\sum_{n=1}^{q}x_k\cdot y_n\cdot I_{A_k\cap B_n})=\sum_{k=1}^{p}\sum_{n=1}^{q}x_k\cdot y_nP(A_k\cap B_n)=$$

$$=\sum_{k=1}^{p}\sum_{n=1}^{q}x_k\cdot y_nP(A_k)\cdot P(B_n)=\sum_{k=1}^{p}x_kP(A_k)\cdot\sum_{n=1}^{q}y_nP(B_n)=M\xi\cdot M\eta.$$

This ends the proof of theorem .

Using Theorems 8.4 and 8.5, we get the validity of the following theorem.

Theorem 8.6 *If ξ and η are independent random variables, then*

$$M(\xi\cdot\eta)=M\xi\cdot M\eta,$$

i.e., mathematical expectation of the product of two random independent variables is equal to the product of expectations of the corresponding variables.

Proof. *If ξ and η are two independent random variables then using Theorem 4, there exists a sequence of simple independent random variables $(\xi_n)_{n\in N}$ and $(\eta_n)_{n\in N}$ such that*

1) ξ_n and η_n are independent for $n\in N$;

2) $(\xi\cdot\eta)(\omega)=\lim_{n\to\infty}\xi_n(\omega)\cdot\eta_n(\omega)\quad(\omega\in\Omega)$.

By using Definition 5 and the result of Theorem 5, we get

$$M(\xi\cdot\eta)=\lim_{n\to\infty}M(\xi_n\cdot\eta_n)=\lim_{n\to\infty}M\xi_n\cdot M\eta_n=$$

$$=\lim_{n\to\infty}M\xi_n\cdot\lim_{n\to\infty}M\eta_n=M\xi\cdot M\eta.$$

Definition 8.7 *A finite family of random variables $\xi_1,\cdots,\xi_n$ is called independent if*

$$P(\{\omega:\xi_1(\omega)\le x_1,\cdots,\xi_n(\omega)\le x_n\})=\prod_{k=1}^{n}P(\{\omega:\xi_k(\omega)\le x_k\})$$

for every $(x_k)_{1\le k\le n}\in\overline{R}^n$, where $\overline{R}=R\cup\{+\infty\}\cup\{-\infty\}$.

Definition 8.8 *A sequence of random variables $(\xi_n)_{n\in N}$ is called independent if family $(\xi_k)_{1\le k\le n}$ is independent for arbitrary $n\in N$.*

Remark 8.2 *An analogy of Theorem 8.6 is valid for arbitrary finite family of random variables, i.e., if $(\xi_k)_{1\le k\le n}$ is a family of independent random variables, then*

$$M(\prod_{k=1}^{n}\xi_k)=\prod_{k=1}^{n}M\xi_k.$$

Theorem 8.7 *If $c\in R$, then $M(c\xi)=cM\xi$, i.e., constant c goes out from the symbol of mathematical expectation.*

Proof. *Note that a constant random variable c and an arbitrary random variables are independent. Following Theorem 8.6, we have*

$$M(c \cdot \xi) = Mc \cdot M\xi = cM\xi.$$

Theorem 8.8 *(Cauchy-Buniakovski*[1] *inequality). For arbitrary random variables ξ and η the following inequality*

$$|M(\xi \cdot \eta)| \leq \sqrt{M\xi^2} \cdot \sqrt{M\eta^2}.$$

holds.

Proof. *Let consider value $M(\xi + x\eta)^2$. Clearly, on the one hand, we have $M(\xi + x\eta)^2 \geq 0$ for arbitrary $x \in R$. Hence, the following expression*

$$M(\xi + x\eta)^2 = M\xi^2 + 2M(\xi \cdot \eta) \cdot x + M\eta^2 \cdot x^2$$

can be considered as a non-negative quadratic polynomial. Hence, its determinant must be non-positive, i.e.,

$$(2M(\xi \cdot \eta))^2 - 4M\eta^2 \cdot M\xi^2 \leq 0,$$

which is equivalent to the condition

$$|M(\xi \cdot \eta)| \leq \sqrt{M\xi^2} \cdot \sqrt{M\eta^2}.$$

This ends the proof of theorem.

Theorem 8.9 *The following formula for calculation of variance $D\xi$*

$$D\xi = M\xi^2 - (M\xi)^2$$

is valid for arbitrary random variable ξ.

Proof. *By the definition of mathematical variance of ξ, we have*

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

Using the properties of mathematical expectation $M\xi$ we have

$$D\xi = M(\xi - M\xi)^2 = M(\xi^2 - 2\xi M\xi + (M\xi)^2) =$$
$$= M\ \xi^2 - M(2\xi M\xi) + M((M\xi)^2) =$$
$$= M\xi^2 - 2M\xi M\xi + (M\xi)^2 = M\xi^2 - (M\xi)^2.$$

This ends the proof of theorem.

Theorem 8.10 *For arbitrary random variable ξ the following equality*

$$D\xi = \min_{a \in R} M(\xi - a)^2$$

[1] Buniakovski, Victor [4(16).12.1804 - 30.11 (12.12). 1889] -Russian mathematician, Academician of Petersburg Academy of Sciences (1830)).

holds.

Proof. *Let calculate a minimum value of function* $M(\xi-a)^2$. *Clearly,*

$$M(\xi-a)^2 = M(\xi^2-2a\xi+a^2) = M\xi^2-2M\xi a+a^2,$$

i.e., $M(\xi-a)^2$ *is a quadratic polynomial with respect to* a. *Hence, point* $a_{\min}$ *is defined by*

$$\frac{dM(\xi-a)^2}{da} = -2M\xi+2a = 0.$$

It follows, $a_{\min} = M\xi$, *i.e.,*

$$\min_{a\in R} M(\xi-a)^2 = M(\xi-a_{\min})^2 = M(\xi-M\xi)^2 = D\xi.$$

This ends the proof of theorem.

Theorem 8.11 *For arbitrary random variable* ξ *the following conditions*
1) $D\xi \geq 0$,
2) $D\xi = 0 \Leftrightarrow (\exists c)(c\in R \to P(\{\omega : \xi(\omega)=c\})=1)$
are fulfilled.

Proof. *Since* $D\xi = M(\xi-M\xi)^2$ *and* $(\xi-M\xi)^2 \geq 0$, *we easily deduce the validity of part 1).*

Let us prove part 2).

Let $P(\{\omega : \xi(\omega)=c\})=1$, *then* $M\xi = c$ *and* $M\xi^2 = c^2$. *Following Theorem 9 we get* $D\xi = M\xi^2-(M\xi)^2 = c^2-c^2 = 0$. *Now, if* $D\xi=0$, *then* $M(\xi-M\xi)^2=0$, *i.e.,* $P(\{\omega : \xi(\omega)=M\xi\})=1$. *Hence, it is sufficient to consider* $c=M\xi$.

This ends the proof of Theorem.

Theorem 8.12 *Let* $c\in R$ *and* ξ *be an arbitrary random variable. Then :*
1) $D(c\xi) = c^2D\xi$,
2) $D(c+\xi) = D\xi$.

Proof. *Following Theorem 8.7 and the definition of the mathematical expectation, we get*

$$D(c\xi) = M(c\xi-M(c\xi))^2 = M(c\xi-cM\xi)^2 = M(c^2(\xi-M\xi)^2) =$$

$$= c^2M(\xi-M\xi)^2 = c^2D\xi.$$

This ends the proof of the part 1).

By definition of mathematical variance of $\xi+c$, *we get*

$$D(c+\xi) = M((c+\xi)-M(c+\xi))^2 = M(c+\xi-Mc-M\xi)^2 =$$

$$= M(c+\xi-c-M\xi)^2 = M(\xi-M\xi)^2 = D\xi.$$

This ends the proof of the part 2) and theorem is proved.

Theorem 8.13 *Let* ξ *and* η *be independent random variables. Then*

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

proof. *Note that random variables* $\xi - M\xi$ *and* $\eta - M\eta$ *are independent. Hence, we get*

$$D(\xi+\eta) = M((\xi+\eta) - M(\xi+\eta))^2 = M((\xi - M\xi) + (\eta - M\eta))^2 =$$

$$= M((\xi - M\xi)^2 + 2(\xi - M\xi)(\eta - M\eta) + (\eta - M\eta)^2) =$$

$$= M(\xi - M\xi)^2 + 2M((\xi - M\xi)(\eta - M\eta)) + M(\eta - M\eta)^2 =$$

$$= D\xi + 2M(\xi - M\xi)M(\eta - M\eta) + D\eta =$$

$$= D\xi + 2(M\xi - M(M\xi))(M\eta - M(M\eta)) + D\eta =$$

$$= D\xi + 2(M\xi - M\xi)(M\eta - M\eta) + D\eta = D\xi + D\eta.$$

This ends the proof of theorem.

Remark 8.3 *Note that an analogy of Theorem 13 is valid for arbitrary finite family* $(\xi_k)_{1\le k\le n}$ *of independent random variables, i.e., the following equality*

$$D\sum_{k=1}^{n}\xi_k = \sum_{k=1}^{n} D\xi_k.$$

holds.

We have the following

Theorem 8.14 *Let* F_ξ *be a distribution function of the absolutely continuous random variable. Then the following formula for calculation of mathematical variance*

$$D\xi = \int_{-\infty}^{+\infty}(x - M\xi)^2 f_\xi(x)dx$$

is valid.

Let consider some examples for calculation of mathematical expectations and mathematical variances.

Example 8.1 *(Poisson distribution). Let*

$$\xi(\omega) = \sum_{n\in N} n \cdot I_{A_n}(\omega) \quad (\omega \in \Omega)$$

be a discrete random variable distributed by Poisson law with parameter λ $(\lambda > 0)$*, i.e.,*

$$P(A_n) = P(\{\omega : \xi(\omega) = n\}) = \frac{\lambda^n}{n!}e^{-\lambda} \quad (n \in N).$$

Then

$$M\xi = \sum_{n=0}^{\infty} n \cdot \frac{\lambda^n}{n!}e^{-\lambda} = \sum_{n=1}^{\infty} n \cdot \frac{\lambda^n}{n!}e^{-\lambda} =$$

$$= \lambda\sum_{n=1}^{\infty} n \cdot \frac{\lambda^{n-1}}{n!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!} = \lambda.$$

On the other hand, we have

$$M\xi^2 = \sum_{n=0}^{\infty} n^2 \cdot \frac{\lambda^n}{n!}e^{-\lambda} = \sum_{n=1}^{\infty} n^2 \cdot \frac{\lambda^n}{n!}e^{-\lambda} =$$

$$= \sum_{n=1}^{\infty} n \cdot \frac{\lambda^n}{(n-1)!}e^{-\lambda} = \sum_{n=1}^{\infty} \frac{(n-1)\lambda^n}{(n-1)!}e^{-\lambda} +$$

$$+ \sum_{n=1}^{\infty} \frac{\lambda^n}{(n-1)!}e^{-\lambda} = \lambda \sum_{m=0}^{\infty} \frac{m\lambda^m}{m!}e^{-\lambda} +$$

$$+ \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!}e^{-\lambda} = \lambda^2 + \lambda = \lambda(1+\lambda).$$

Following Theorem 8.9, we get

$$D\xi = M\xi^2 - (M\xi)^2 = \lambda(1+\lambda) - \lambda^2 = \lambda.$$

Example 2 *(Geometric distribution). Let*

$$\xi(\omega) = \sum_{n \in N} n \cdot I_{A_n}(\omega) \quad (\omega \in \Omega)$$

be a discrete random variable distributed by the geometric law with parameter q, where $(0 \le q \le 1)$, i.e.,

$$P(A_n) = P(\{\omega : \xi(\omega) = n\}) = (1-q)q^{n-1} \quad (n \in N).$$

Then

$$M\xi = \sum_{n=1}^{\infty} n(1-q)q^{n-1} = (1-q) \cdot \sum_{n=1}^{\infty} nq^{n-1} = (1-q)(\sum_{n=1}^{\infty} q^n)' = (1-q)(\frac{1}{1-q})' =$$

$$= (1-q) \cdot \frac{1}{(1-q)^2} = \frac{1}{1-q}.$$

On the other hand, we have

$$M\xi^2 = \sum_{k=1}^{\infty} k^2(1-q)q^{k-1} = (1-q) \cdot \sum_{k=1}^{\infty}(k \cdot kq^{k-1}) = (1-q) \cdot \sum_{k=1}^{\infty}(kq^k)' =$$

$$= (1-q) \cdot \sum_{k=1}^{\infty}(kq^{k-1} \cdot q)' = (1-q) \cdot [(\sum_{k=1}^{\infty} kq^{k-1})'q + \sum_{k=1}^{\infty} kq^{k-1}] =$$

$$= (1-q)[\frac{2q}{(1-q)^3} + \frac{1}{(1-q)^2}] = \frac{2q}{(1-q)^2} + \frac{1}{(1-q)}.$$

Hence,

$$D\xi = M\xi^2 - (M\xi)^2 = \frac{2q}{(1-q)^2} + \frac{1}{(1-q)} - \frac{1}{(1-q)^2} = \frac{q}{(1-q)^2}.$$

Example 8.3 *(Leibniz[2] distribution). Let*

$$\xi(\omega) = \sum_{n=1}^{\infty} n \cdot I_{A_n}(\omega) \ \ (\omega \in \Omega)$$

be a discrete random variable distributed by the Leibniz law, i.e.,

$$P(A_n) = P(\{\omega : \xi(\omega) = n\}) = \frac{1}{n \cdot (n+1)} \ \ (n \in N).$$

Then

$$\sum_{n=1}^{\infty} n \cdot \frac{1}{n \cdot (n+1)} = \sum_{n=1}^{\infty} \frac{1}{(n+1)} = +\infty.$$

Hence, mathematical expectation $M\xi$ and mathematical variance $D\xi$ are not finite.

Example 8.4 *(Binomial distribution). Let*

$$\xi(\omega) = \sum_{k=0}^{n} k I_{A_k}(\omega) \ \ \ (\omega \in \Omega)$$

be a simple random variable distributed by the Binomial law with parameters (n, p), i.e.,

$$P(A_k) = P(\{\omega : \xi(\omega) = k\}) = C_n^k \cdot p^k (1-p)^{n-k},$$

where $0 < p < 1$ and $0 \le k \le n$.

Then

$$M\xi = \sum_{k=0}^{n} k \cdot C_n^k \cdot p^k (1-p)^{n-k} = \sum_{k=0}^{n} k \cdot \frac{n!}{k!(n-k)!} \cdot p^k (1-p)^{n-k} =$$

$$= n \cdot p \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} \cdot p^{k-1} (1-p)^{(n-1)-(k-1)} =$$

$$= n \cdot p \sum_{k-1=0}^{n-1} \frac{(n-1)!}{(k-1)!(n-k)!} \cdot p^{k-1} (1-p)^{(n-1)-(k-1)} =$$

$$= n \cdot p \sum_{s=0}^{n-1} \frac{(n-1)!}{s!((n-1)-s)!} \cdot p^{s} (1-p)^{(n-1)-s} =$$

[2] *Leibniz, Gottfried Wilhelm* (1.7.1646 - 14.11.1716)-German mathematician, the member of London Royal Society (1673), the member of Paris Academy of Sciences (1700).

$$= n \cdot p \sum_{s=0}^{n-1} C_{n-1}^{s} \cdot p^{s}(1-p)^{(n-1)-s} = n \cdot p.$$

Remark 8.4 Let η be a random variable distributed by the Bernoulli law with parameter p ,i.e.,

$$\eta(\omega) = 0 \cdot I_{A_0}(\omega) + 1 \cdot I_{A_1}(\omega) \quad (\omega \in \Omega),$$

where $A_0 + A_1 = \Omega$ and

$$P(A_0) = P(\{\omega : \eta(\omega) = 0\}) = 1-p, \; P(A_1) = P(\{\omega : \eta(\omega) = 1\}) = p.$$

Then

$$M\eta = 0 \cdot P(\{\omega : \eta(\omega) = 0\}) + 1 \cdot P(\{\omega : \eta(\omega) = 1\}) = 1 \cdot (1-p) + 1 \cdot p = p.$$

On the other hand, we have

$$P(\{\omega : \eta^2(\omega) = 0\}) = 1-p, \; P(\{\omega : \eta^2(\omega) = 1\}) = p,$$

Hence,

$$M(\eta^2) = 0 \cdot P(\{\omega : \eta^2(\omega) = 0\}) + 1 \cdot P(\{\omega : \eta^2(\omega) = 1\}) = 1 \cdot (1-p) + 1 \cdot p = p.$$

Finally we get

$$D(\eta) = M\eta^2 - (M\eta)^2 = p - p^2 = p(1-p).$$

As simple discrete random variable ξ distributed by Binomial law with parameter (n,p) can be presented as a sum of n exemplars of independent simple discrete random variables distributed by Bernoulli law with parameter p, following Theorem 3, we get

$$M\xi = M(\sum_{k=1}^{n} \xi_k) = \sum_{k=1}^{n} M\xi_k = np.$$

Following Remark 8.3, we get

$$D\xi = D(\sum_{k=1}^{n} \xi_k) = \sum_{k=1}^{n} D\xi_k = np(1-p).$$

Example 8.5 (Normal distribution). Let $\xi : \Omega \to R$ be a normally distributed random variable with parameter (m, σ^2) $(m \in R, \sigma > 0)$, i.e.,

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

Then, following Theorem 1, we get

$$M\xi = \int_{-\infty}^{+\infty} x f_\xi(x) dx = \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx =$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}(x-m)\cdot e^{-\frac{(x-m)^2}{2\sigma^2}}dx + \frac{m}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\cdot e^{-\frac{(x-m)^2}{2\sigma^2}}dx =$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} z\cdot e^{-\frac{z^2}{2\sigma^2}}dz + m = m.$$

Using the formula for calculation of mathematical variance, we get

$$D\xi = \int_{-\infty}^{+\infty}(x-m)^2 f_\xi(x)dx = \int_{-\infty}^{+\infty}(x-m)^2 f_\xi(x)dx =$$

$$= \int_{-\infty}^{+\infty}(x-m)^2\cdot\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}dx = \int_{-\infty}^{+\infty} z^2\cdot\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{z^2}{2\sigma^2}}dz,$$

where $z = x - m$.

Setting $t = \frac{z}{\sigma}$, we get

$$D\xi = \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2\sigma^2}}dz = \frac{\sigma^2}{\sqrt{2\pi}}\sqrt{2\pi} = \sigma^2,$$

because

$$\int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2\sigma^2}}dz = \sqrt{2\pi}.$$

Example 8.6 (Uniform distribution on $[a;b]$). Let $\xi : \Omega \to R$ be a random variable uniformly distributed on $[a,b]$ $(a < b)$, i.e.,

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

Then

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

On the other hand, we have

$$M\xi^2 = \int_{-\infty}^{+\infty} x^2 f_\xi(x)dx = \int_a^b x^2\cdot\frac{1}{b-a}dx = \frac{x^3}{3(b-a)}\Big|_a^b = \frac{b^2+ab+a^2}{3}.$$

Following Theorem 8.9, we get

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

Example 8.7 (Cauchy distribution). Let $\xi : \Omega \to R$ be an absolutely continuous random variable distributed by the Cauchy law, i.e.,

$$f_\xi(x) = \frac{1}{\pi(1+x^2)} \quad (x \in R).$$

Note that the following indefinite integral

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

does not exists. Hence, we deduce that there exists no a mathematical expectation of the random variable distributed by the Cauchy law.

Example 8.8 (Exponential distribution). Let $\xi : \Omega \to R$ be an absolutely continuous random variable distributed by the exponential law with parameter λ,i.e.,

$$f_\xi(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0; \\ 0, & \text{if } x < 0. \end{cases}$$

Then

$$M\xi = \int_{-\infty}^{+\infty} x f_\xi(x)dx = \int_0^{+\infty} x\lambda e^{-\lambda x}dx =$$

$$= \lambda\big(-\frac{1}{\lambda}xe^{-\lambda x}\big|_0^\infty + \frac{1}{\lambda}\int_0^\infty e^{-\lambda x}dx\big) = \lambda\big(-\lim_{l\to\infty}\frac{1}{\lambda}\frac{l}{e^{\lambda l}} + \frac{1}{\lambda^2}\big) =$$

$$= \lambda\big(-\lim_{l\to\infty}\frac{1}{\lambda}\frac{1}{\lambda e^{\lambda l}} + \frac{1}{\lambda^2}\big) = \lambda\cdot\frac{1}{\lambda^2} = \frac{1}{\lambda}.$$

Using analogous calculations, we get

$$D\xi = \int_{-\infty}^{+\infty} x^2 f_\xi(x)dx - (M\xi)^2 = \lambda\int_0^{+\infty} x^2 e^{-\lambda}xdx - \frac{1}{\lambda^2} =$$

$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Example 8.9 (Singular distribution). Let consider random variable ξ_{Cantor}, defined on $[0,1]$.

It is easy to show that

$$\int_0^1 \xi_{Cantor}(y)dy + \int_0^1 F(x)dx = 1,$$

where F denotes the Cantor function defined on $[0,1]$.

Hence,

$$M\xi_{Cantor} = \int_0^1 \xi_{Cantor}(y)dy = 1 - \int_0^1 F(x)dx.$$

Note that for set Δ_2, obtained by counterclockwise rotation about point $(\frac{1}{2},\frac{1}{2})$ on angle π of set $\Delta_1 = \{(x,y) : x \in [0,1], 0 \leq y \leq F(x)\}$, we have:

a) $b_2(\Delta_1 \cap \Delta_2) = 0$,

b) $b_2(\Delta_1) = b_2(\Delta_2)$,

c) $\Delta_1 \cup \Delta_2 = [0,1] \times [0,1]$.

Hence, $b_2(\Delta_1) = b_2(\Delta_2) = \frac{1}{2}$. It follows

$$M\xi_{Cantor} = 1 - \int_0^1 F(x)dx = 1 - b_2(\Delta_1) = 1 - \frac{1}{2} = \frac{1}{2}.$$

Now let calculate $D(\xi_{Cantor})$. Note that $\pi M\xi^2_{Cantor}$ coincides with the volume of the object, obtained by rotation of set Δ_2 about real axis $0Y$, which is equal to the difference of volumes of figures, obtained by the rotation of sets

$$[0,1] \times [0,1]$$

and

$$[\frac{1}{3}, \frac{2}{3}] \times [0, \frac{1}{2}] \cup [\frac{1}{9}, \frac{2}{9}] \times [0, \frac{1}{4}] \cup [\frac{7}{9}, \frac{8}{9}] \times [0, \frac{3}{4}] \cup \cdots$$

about real axis $0Y$, respectively. Hence,

$$M\xi^2_{Cantor} = 1 - [\frac{1}{2}((\frac{2}{3})^2 - (\frac{1}{3})^2) + \frac{1}{4}((\frac{2}{9})^2 - (\frac{1}{9})^2) +$$
$$\frac{3}{4}((\frac{8}{9})^2 - (\frac{7}{9})^2) + \cdots].$$

It follows that

$$D\xi^2_{Cantor} = M\xi^2_{Cantor} - (M\xi_{Cantor})^2 = \frac{3}{4} - [\frac{1}{2}((\frac{2}{3})^2 - (\frac{1}{3})^2) + \frac{1}{4}((\frac{2}{9})^2 - (\frac{1}{9})^2) +$$
$$\frac{3}{4}((\frac{8}{9})^2 - (\frac{7}{9})^2) + \cdots].$$

Remark 8.5. (Physical sense of mathematical expectation and mathematical variance). We remind the reader that arbitrary random variable $\xi : \Omega \to R$ can be considered as a special rule of dispersion of the unit mass of powder Ω on real axis R, by means of which every particle $\omega \in \Omega$ is placed on particle $A \in R$ with coordinate $\xi(\omega)$. Here naturally arises the following

Problem. What physical sense is put in $M\xi$ and $D\xi$, respectively ?

It is well known from the course of theoretical mechanics that if mass p_k is placed at point $x_k \in R$ for $1 \le k \le n$ and $\sum_{k=1}^n p_k = 1$, then center x_c of the whole mass is calculated by :

$$x_c = \sum_{k=1}^n x_k \cdot p_k.$$

If the rule of dispersion of the unit mass of powder Ω on real axis R is a simple discrete random variable given by the following table

ξ	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

,

then $M\xi = x_c$, which means that $M\xi$ is a center of the unite mass distributed by the law ξ on real axis R.

Note that the physical sense of $M\xi$ is same in the case of arbitrary random variable ξ.

On the other hand, if ξ is a simple discrete random variable, then

$$D\xi = \sum_{k=1}^{n}(x_k - M\xi)^2 p_k.$$

Note that value $D\xi$ depends on values $((x_k - M\xi)^2)_{1\le k\le n}$. The latter relation means that the particles of mass are concentrated nearer to its center $M\xi$ as well a mathematical variance $D(\xi)$ is near at zero. In particular, if $x_1 = \cdots = x_n = M\xi$, then $D\xi = 0$.

Hence, mathematical variance $D\xi$ can be considered as a characterization why the particles of the unite mass of the powder are removed about its center $M\xi$.

As an example let consider random variables ξ_1 and ξ_2, defined by

ξ_1	-1	1
P	$\frac{1}{2}$	$\frac{1}{2}$

,

ξ_2	-2	2
P	$\frac{1}{2}$	$\frac{1}{2}$

.

Clearly,

$$M\xi_1 = M\xi_2 = 0,$$

$$D\xi_1 = 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 1,$$

$$D\xi_2 = 4 \cdot \frac{1}{2} + 4 \cdot \frac{1}{2} = 4.$$

Note that, on the one hand, the centers of the particles of the unit mass of powder Ω dispersed by laws ξ_1 and ξ_2, respectively, coincide and are equall to zero, i.e., $M\xi_1 = M\xi_2 = 0$. On the other hand, the particles of the unit mass of powder Ω dispersed by rule ξ_1 are more nearer to the center than the particles of the unit mass of powder Ω dispersed by rule ξ_2.

Remark 8.1. Let $x_1, \cdots, x_n$ be the results of observation on the random variable with finite mathematical expectation and with finite mathematical variance. Then:

1) AVERAGE$(x_1 : x_n)$ calculates sum $\frac{1}{n}\sum_{i=1}^{n} x_i$.
2) VARP$(x_1 : x_n)$ calculates sum $\frac{1}{n}\sum_{i=1}^{n}(x_i - \frac{1}{n}\sum_{j=1}^{n} x_j)^2$.
3) VAR$(x_1 : x_n)$ calculates sum $\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \frac{1}{n}\sum_{j=1}^{n} x_j)^2$.

Tests

8.1. Distribution laws ξ and η are given in the following tables

ξ	-1	0	1	2
P	$0,3$	$0,2$	$0,1$	$0,4$

,

η	-1	0	-2
P	$0,5$	$0,3$	$0,2$

.

Then

1) $M(3\xi - 4\eta)$ is equal to

a) 5,3, b) 5,4, c) 5,5, d) 5,6;

2) $D(3\xi - 4\eta)$ is equal to

a) 20,4,3, b) 21,5, c) 22,6, d) 23,7;

8.2. Distribution function F_ξ of the absolutely continuous random variable ξ has the following form

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

Then

1) $M(3\xi - 4)$ is equal to

a) 3, b) −3, c) 4, d) −4;

2) $D(\sqrt{18}\xi - 4)$ is equal to

a) 0,3, b) 0,7, c) 1, d) 1,3.

8.3. Let ξ_1 be a random variable normally distributed with parameters (3,25), ξ_2 be a random variable uniformly distributed in interval (18,20) and ξ_3 be a random variable distributed by the Poisson law with parameter $\lambda = 5$. Then

1) $M(1\xi_1 + 2\xi_2 + 3\xi_3)$ is equal to

a) 34, b) 35, c) 36, d) 37;

2) if ξ_1, ξ_2, ξ_3 are independent random variables, then $D(1\xi_1 + 2\xi_2 + 3\xi_3 + 4)$ is equal to

a) $25\frac{1}{3}$, b) $26\frac{1}{3}$, c) $27\frac{1}{3}$, d) $28\frac{1}{3}$.

8.4. Distribution laws ξ and η are given in the following tables

ξ	−1	1	2
P	0,2	0,1	0,7

,

η	2	3	−1
P	0,3	0,3	0,4

.

Then:

1) distribution law $\xi\eta$ is given in the following table

a)

$\xi\eta$	−3	−2	−1	1	2	3	4	6
P	0,06	0,34	0,04	0,08	0.03	0,03	0,21	0,21

,

b)

$\xi\eta$	−3	−2	−1	1	2	3	4	6
P	0,05	0,35	0,03	0,09	0.03	0,02	0,22	0,21

.

2) distribution law $\xi + \eta$ is given in the following table

a)

$\xi + \eta$	−2	0	1	2	3	4	5
P	0,08	0,04	0,34	0,06	0.03	0,24	0,21

,

b)

$\xi+\eta$	-2	0	1	2	3	4	5
P	$0,06$	$0,06$	$0,34$	$0,06$	0.02	$0,25$	$0,21$

.

Chapter 9

Correlation Coefficient

Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let ξ and η be such random variables that $0 < D\xi < \infty$ and $0 < D\eta < \infty$.

Definition 9.1 *Numerical value $\rho(\xi, \eta)$, defined by*

$$\rho(\xi, \eta) = \frac{M(\xi - M\xi)(\eta - M\eta)}{\sqrt{D\xi}\sqrt{D\eta}},$$

is called a correlation coefficient between random values ξ and η.

Definition 9.2 *Numerical value $cov(\xi, \eta)$, defined by*

$$cov(\xi, \eta) = \frac{M(\xi - M\xi)(\eta - M\eta)}{\sqrt{D\xi}\sqrt{D\eta}},$$

is called a covariation coefficient between random values ξ and η.

Remark 9.1 Let $(x_1, y_1), \cdots, (x_n, y_n)$ be the results of observations on the random vector (X, Y), every component of which has a finite mathematical expectation and a finite mathematical variance. Then:

1) CORREL$(x_1 : x_n; y_1 : y_n)$ calculates the value ρ_n, which is a good estimation of correlation coefficient $\rho(X, Y)$.

2) COVAR$(x_1 : x_n; y_1 : y_n)$ calculates a value $\mathrm{cov}_n(X, Y)$, which is a good estimation of covariation coefficient $\mathrm{cov}(X, Y)$.

Below in columns A and B we have entered the results of observations of random vector (X, Y), every component of which has a finite mathematical expectation and a finite mathematical variance.

$A(= x_k)$	$B(= y_k\})$
7	2
11	5
6	6
7	7

,

Then:

1) $\rho_4 = CORREL(A1 : A4; B1 : B4)$=-0,0695889;
2) $\mathrm{cov}_4(X,Y) = COVAR(A1 : A4; B1 : B4) = -0,25.$

We have the following propositions

Theorem 9.1 *Let, ξ and η be such random variables that $0 < D\xi < \infty$ and $0 < D\eta < \infty$. Then $|\rho(\xi,\eta)| \le 1$.*

Proof.

$$0 \le D(\frac{\xi - M\xi}{\sqrt{D\xi}} \pm \frac{\eta - M\eta}{\sqrt{D\eta}}) = M(\frac{\xi - M\xi}{\sqrt{D\xi}} \pm \frac{\eta - M\eta}{\sqrt{D\eta}})^2 = 2 \pm 2\rho(\xi,\eta).$$

Hence, we get $|\rho(\xi,\eta)| \le 1$.

Theorem 9.2 *If ξ and η are such independent random variables, that $0 < D\xi < \infty$ and $0 < D\eta < \infty$, then $|\rho(\xi,\eta)| = 0$.*

Proof. From the independence of ξ and η we get that random variables $\frac{\xi - M\xi}{\sqrt{D\xi}}$ and $\frac{\eta - M\eta}{\sqrt{D\eta}}$ are also independent. By using Theorem 6(cf. §8) we deduce that

$$\rho(\xi,\eta) = M\Big[\Big(\frac{\xi - M\xi}{\sqrt{D\xi}}\Big) \cdot \Big(\frac{\eta - M\eta}{\sqrt{D\eta}}\Big)\Big] =$$

$$= M\Big(\frac{\xi - M\xi}{\sqrt{D\xi}}\Big) \cdot M\Big(\frac{\eta - M\eta}{\sqrt{D\eta}}\Big) = 0.$$

Example 9.1. Note here that the inverse result given in Theorem 9.2 is not always valid, i. e., the existence of such non-independent random variables ξ and η is possible that $0 < D\xi < \infty$, $0 < D\eta < \infty$ and $\rho(\xi,\eta) = 0$. Indeed, assume

$$(\Omega, \mathcal{F}, P) = ([0;1], \mathcal{B}([0;1]), b_1).$$

Let define random variables ξ and η with the following formulas:

$$\xi(\omega) = 4 \cdot I_{[0,\frac{1}{4}[}(\omega) + 0_{[\frac{1}{4},\frac{1}{2}[}(\omega) - 4_{[\frac{1}{2},\frac{3}{4}[}(\omega) + 0_{[\frac{3}{4},1]}(\omega),$$

$$\eta(\omega) = 0 \cdot I_{[0,\frac{1}{4}[}(\omega) + 4_{[\frac{1}{4},\frac{1}{2}[}(\omega) + 0_{[\frac{1}{2},\frac{3}{4}[}(\omega) - 4_{[\frac{3}{4},1]}(\omega).$$

Note that

$$M\xi = M\eta = 0,\ D\xi = D\eta = 8$$

and

$$\rho(\xi,\eta) = \frac{M(\xi - M\xi)(\eta - M\eta)}{\sqrt{D\xi}\sqrt{D\eta}} =$$

$$= \frac{M\xi\eta}{8} = \frac{M0}{8} = 0.$$

Now let show that ξ and η are not independent. Indeed,

$$P(\{\omega : \xi < 3,\ \eta < 3\}) = \frac{1}{2},$$

$$P(\{\omega : \xi < 3\}) = \frac{3}{4}, \ P(\{\omega : \eta < 3\}) = \frac{3}{4},$$

It follows that

$$P(\{\omega : \xi < 3, \eta < 3\}) \neq P(\{\omega : \xi < 3\}) \cdot P(\{\omega : \eta < 3\}).$$

Theorem 9.3 *If the conditions of Theorem 1 are fulfilled then* $|\rho(\xi,\eta)| = 1$ *if and only if there exist real numbers* a $(a \neq 0)$ *and* b *such that*

$$P(\{\omega : \eta(\omega) = a\xi(\omega) + b\}) = 1.$$

Proof.

Sufficient . Assume that

$$P(\{\omega : \eta(\omega) = a\xi(\omega) + b\}) = 1.$$

We set $M\xi = \alpha$ and $\sqrt{D\xi} = \beta$. Then

$$\rho(\xi,\eta) = M\frac{\xi - \alpha}{\beta} \cdot \frac{\alpha\xi + b - a\alpha - b}{|\alpha|\beta} = sign(a).$$

Necessity. Assume that $\Big|\rho(\xi,\eta)\Big| = 1$. Let consider the case when $\rho(\xi,\eta) = 1$. Then

$$D\Big(\frac{\xi - M\xi}{\sqrt{D\xi}} - \frac{\eta - M\eta}{\sqrt{D\eta}}\Big) = 2(1 - \rho(\xi,\eta)) = 0.$$

Using the property of mathematical variance for concrete $c \in R$, we get

$$P\Big(\{\omega : \frac{\xi - M\xi}{\sqrt{D\xi}} - \frac{\eta - M\eta}{\sqrt{D\eta}} = c\}\Big) = 1.$$

Hence,

$$P\Big(\{\omega : \xi(\omega) = \frac{\sqrt{D\xi}}{\sqrt{D\eta}} \cdot \eta(\omega) - \sqrt{D\xi}\Big(\frac{M\eta}{\sqrt{D\eta}} - c\Big) + M\xi\}\Big) = 1.$$

If $\rho(\xi,\eta) = -1$, we get

$$D\Big(\frac{\xi - M\xi}{\sqrt{D\xi}} + \frac{\eta - M\eta}{\sqrt{D\eta}}\Big) = 2(1 + \rho(\xi,\eta)) = 0.$$

Analogously, using the property of mathematical variance, we deduce an existence of such $d \in R$ that

$$P\Big(\{\omega : \frac{\xi - M\xi}{\sqrt{D\xi}} + \frac{\eta - M\eta}{\sqrt{D\eta}} = d\}\Big) = 1,$$

i.e.,

$$P\Big(\{\omega : \xi(\omega) = -\frac{\sqrt{D\xi}}{\sqrt{D\eta}} \cdot \eta(\omega) + \sqrt{D\xi}\frac{M\eta}{\sqrt{D\eta}} + d\sqrt{D\xi} + M\xi\}\Big) = 1.$$

Remark 9.2 The correlation coefficient is a quantity characterization of the degree of the dependence between two random variables. It can be considered as "cosine" of the angle between them. Indeed, since $|\rho(\xi,\eta)| \leq 1$, there exists a unique real number ϕ in interval $[0,\pi]$, that $\cos\phi = \rho(\xi,\eta)$. This number ϕ is called an angle between random variables ξ and η and is denoted with symbol $\widehat{(\xi,\eta)}$, i.e., $\widehat{(\xi,\eta)} = arccos(\rho(\xi,\eta))$. The following geometrical interpretations of theorems 2 and 3 is interesting:

1) If ξ and η are such independent variables that $0 < D\xi < \infty$ and $0 < D\eta < \infty$, then they are orthogonal , i.e., $\widehat{(\xi,\eta)} = \frac{\pi}{2}$.

2) If $\widehat{(\xi,\eta)}$ is equal to 0 or π, then a random variable η is presented (P-almost everywhere) as a linear combination of random variable ξ and constant random variable.

Example 9.2 Let consider a transmission system of the signal. Let denote a useful signal with ξ. As here we have hindrances, we receive signal $\eta(\omega) = \alpha\xi(\omega) + \Delta(\omega)$, where α is a coefficient of the intensification , $\Delta(\omega)$ is a hindrance ("white noise"). Assume that variables Δ and ξ are independent, $M\xi = a, D\xi = 1$, $M\Delta = 0$, $D\Delta = \sigma^2$. A correlation coefficient between random variables ξ and η is calculated with

$$\rho(\xi,\eta) = M\left((\xi - a)\cdot\frac{\alpha\xi + \Delta - a\alpha}{\sqrt{\alpha^2+\sigma^2}}\right) = \frac{\alpha}{\sqrt{\alpha^2+\sigma^2}}.$$

If σ is smaller than α and is close to 0, then $\rho(\xi,\eta)$ will be close to 1 and following Theorem 3, it is possible to restore ξ with η.

Let consider other numerical characterizations of random variables.

Definition 9.2 *A moment of order k $(k \in N)$ of the random variable ξ is defined with $M\xi^k$ and is denoted with symbol α_k , i.e.,*

$$\alpha_k = M\xi^k \ (k \in N).$$

Definition 9.3 *Value $M(\xi - M\xi)^k$ $(k \in N)$ is called a central moment of order k and is denoted with symbol μ_k , i.e.,*

$$\mu_k = M(\xi - M\xi)^k \ (k \in N).$$

Remark 9.2 Note that mathematical variance $D\xi$ is the central moment of the order two.

Let $\xi_1,\cdots,\xi_n$ be a finite sequence of random variables.

Definition 9.4 *Value*

$$M\xi_1^{k_1}\cdots\xi_n^{k_n}$$

is called a mixed moment of order $k_1 + \cdots + k_n$ *and is denoted with symbol* $\alpha_{(k_1,\cdots,k_n)}$, *i.e.,*

$$\alpha_{(k_1,\cdots,k_n)} = M\xi_1^{k_1} \cdots \xi_n^{k_n} \quad (k_1, \cdots, k_n \in N).$$

Definition 9.5 *Value*

$$M(\xi_1 - M\xi_1)^{k_1} \cdots (\xi_n - M\xi_n)^{k_n}$$

is called a central moment of order $k_1 + \cdots + k_n$ *and is denoted with symbol* $\mu_{(k_1,\cdots,k_n)}$, *i.e.,*

$$\mu_{(k_1,\cdots,k_n)} = M(\xi_1 - M\xi_1)^{k_1} \cdots (\xi_n - M\xi_n)^{k_n} \ (k_1, \cdots, k_n \in N).$$

Definition 9.6*A skewness coefficient of the random variable* ξ *is called a number* $\frac{\mu_3}{\sigma^3}$ *and is denoted with symbol* A_s, *i.e.,*

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

Remark 9.3. Let $x_1, \cdots, x_n$ be the results of observations on the random variable X. Then the statistical function KURT$(x_1 : x_n)$ gives estimation of the excess of ξ. For example, KURT$(-1; -3; -80; -80) = -5,990143738$.

Definition 9.7 *An excess of the random variable* ξ *is called a number* $\frac{\mu_4}{\sigma^4} - 3$ *and is denoted with symbol* E_x, *i.e.,*

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

Remark 9.4. Let $x_1, \cdots, x_n$ be the results of observations on the random variable X. Then the statistical function SKEW$(x_1 : x_n)$ gives estimation of the excess of ξ. For example, SKEW$(1; -1; 3; -3; 80; 80; -80) = -0,17456105$.

Definition 9.8 *If* F_ξ *is a distribution function of* ξ, *then a median of random variable* ξ *is called a number* γ , *for which the following condition is fulfilled*

$$F_\xi(\gamma - 0) \le \frac{1}{2} \quad , \quad F_\xi(\gamma + 0) \ge \frac{1}{2},$$

where $F_\xi(\gamma - 0)$ *and* $F_\xi(\gamma + 0)$ *denote the right and the left limits of function* F_ξ *in point* γ, *respectively.*

Remark 9.5 Let $x_1, \cdots, x_n$ be the values of the discrete random variable X such that $x_1 < x_2 < \cdots < x_n$. Then median is $x_{k+1}, when$n=2k+1$, and$x$_k + x_{k+1}\overline{2}$, when $n = 2k$. The statistical function MEDIAN$(x_1 : x_n)$ calculates the median of ξ. For example, MEDIAN$(6; 7; 8; 11) = 7,5$ and MEDIAN$(6; 7; 100) = 7$.

Definition 9.9 *A mode of simple discrete random variable* ξ *is called its such possible meaning whose corresponding probability is maximal.*

Definition 9.10 *A mode of absolutely continuous random variable* ξ *is called a point of the local maximum of the corresponding density function.*

Remark 9.6 Let $x_1, \cdots, x_n$ be the results of observations on the random variable X. Then the statistical function MODE$(x_1 : x_n)$ gives the estimation of the smallest mode of ξ. For example, MODE$(7; 11; 6; 7; 11; 18; 18) = 7$.

Definition 9.11 *A random variable is called unimodular, if it has only one mode. In other cases, the random variable is called polymodular.*

Tests

9.1.Assume that $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), b_1)$. Assume also that ξ and η are defined with

$$\xi(\omega) = \begin{cases} 0, & \omega \in [\frac{1}{2}, \frac{3}{4}[\\ 1, & \omega \in [0, \frac{1}{2}[\\ 2, & \omega \in [\frac{3}{4}, 1] \end{cases} ,$$

$$\eta(\omega) = \begin{cases} 2, & \omega \in [0, \frac{1}{2}[, \\ -1, & \omega \in [\frac{1}{2}, 0] \end{cases} .$$

The correlation coefficient $\rho(\xi, \eta)$ is equal to

a) $-0,2$, b) $-0,1$, c) 0, d) $0,1$.

9.2. The distribution law of the random variable ξ is given in the table

ξ	-1	0	-1
P	$0,6$	$0,1$	$0,3$

.

Then

1) $M(\xi^3)$ is equal to

a) $-0,1$, b) $-0,2$, c) $-0,3$, d) $-0,4$;

2) $M(\xi - M\xi)^4$ is equal to

a) $1,948$, b) $0,9481$, c) $0,8481$, d) $0,7481$.

9.3. Let ξ be a random variable normally distributed with parameters $(0,1)$. Then

1) α_{2k+1} is equal to

a) 1, b) 0, c) $2k+1$, d) $2k$;

2) μ_2 is equal to

a) 0, b) 1, c) 2, d) 3;

3) median γ is equal to

a) 0, b) 1, c) 2, d) 3;

4) mode is equal to

a) 0, b) 1, c) 2, d) 3.

9.4. ξ is a random variable uniformly distributed on $(0,4)$. Then

1) μ_2 is equal to

a) 6, b) 7, c) 8, d) 9;

3) median γ is equal to

a) 1, b) 2, c) 3, d) 4;

4) mode is equal to

a) $[0,4]$, b) $[0,3]$, c) $[0,2]$, d) $[0,1]$.

9.5. The distribution law of simple discrete random variable ξ is given in the following table

ξ	-1	2	3
P	$0,3$	$0,4$	$0,3$

.

Then

1) the median of ξ is equal to

a) 1, b) 2, c) 3, d) 4;

2) the mode of γ is equal to

a) -1, b) 2, c) 3, d) 4.

9.6. Distribution function F_ξ of absolutely continuous random variable ξ is defined with

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

Then

1) median γ is equal to

a) $\frac{\sqrt{2}}{2}$, b) $\frac{\sqrt{3}}{3}$, c) $\frac{\sqrt{5}}{5}$, d) $\frac{\sqrt{7}}{7}$;

2) the mode of ξ is equal to

a) 1, b) 2, c) 3, d) 4.

Chapter 10

Random Vector Distribution Function

Let (Ω, F, P) be the probability space and let $(\xi_k)_{1\le k\le n}$ be a finite family of random variables.

Definition 10.1. *Reflection* $(\xi_1, \cdots, \xi_n) : \Omega \to R^n$, *defined with*

$$(\forall\omega)(\omega \in \Omega \to (\xi_1, \cdots, \xi_n)(\omega) = (\xi_1(\omega), \cdots, \xi_n(\omega))),$$

is called n*-dimensional random vector.*

Definition 10.2 *Reflection* $F_{\xi_1,\cdots,\xi_n} : R^n \to R$, *defined with*

$$(\forall(x_1, \cdots, x_n))((x_1, \cdots, x_n) \in R^n \to F_{\xi_1,\cdots,\xi_n}((x_1, \cdots, x_n)) =$$

$$= P(\{\omega : \xi_1(\omega) < x_1, \cdots, \xi_n(\omega) < x_n\})),$$

is called a joint distribution function of the n*-dimensional random vector* $(\xi_1, \cdots, \xi_n)$.

Definition 10.3 *Random vector* $(\xi_1, \cdots, \xi_n)$ *is called discrete if every* i*-th component* ξ_i *is a discrete random variable for* $1 \le i \le n$.

Analogously we can define an absolutely continuous random vector.

The joint distribution function $F_{\xi_1,\cdots,\xi_n}$ has the following properties:

1. $\lim_{x_i\to\infty \text{for } 1\le i\le n} F_{\xi_1,\cdots,\xi_n}((x_1, \cdots, x_n)) = 1$;
2. $\lim_{x_i\to-\infty \text{for } 1\le i\le n} F_{\xi_1,\cdots,\xi_n}((x_1, \cdots, x_n)) = 0$;

Here naturally arises a question what is the probability that the 2-dimensional random vector will obtain the value in the rectangular?

The following result is valid.

Theorem 10.1 *The following formula*

$$(\forall k)(\forall x_k)(\forall y_k)(1 \leq k \leq 2 \ \& \ x_k \in R \ \& \ y_k \in R \ \& \ x_1 < x_2 \ \& \ y_1 < y_2 \rightarrow$$

$$P(\{\omega : (\xi_1, \xi_2)(\omega) \in [x_1; x_2[\times[y_1; y_2[\}) = F_{\xi_1,\xi_2}((x_2, y_2)) - F_{\xi_1,\xi_2}((x_1, y_2)) +$$
$$+F_{\xi_1,\xi_2}((x_1, y_1)) - F_{\xi_1,\xi_2}((x_2, y_1))$$

holds, where

$$[x_1; x_2[\times[y_1; y_2[= \{(x, y)|x_1 \leq x < x_2, \ y_1 \leq y < y_2\}.$$

Proof. Setting

$$A_{(a,b)} = \{\omega : (\xi_1, \xi_2)(\omega) \in] - \infty; a[\times] - \infty; b[\} \quad (a \in R, b \in R),$$

we get

$$\{\omega : (\xi_1, \xi_2)(\omega) \in]x_1; x_2[\times]y_1; y_2[\} = (A_{(x_2,y_2)} \setminus A_{(x_2,y_1)}) \setminus (A_{(x_1,y_2)} \setminus A_{(x_1,y_1)}).$$

Hence,

$$P(\{\omega : (\xi_1, \xi_2)(\omega) \in]x_1; x_2[\times]y_1; y_2[\}) = P((A_{(x_2,y_2)} \setminus A_{(x_2,y_1)})) -$$
$$P((A_{(x_1,y_2)} \setminus A_{(x_1,y_1)})) = (P(A_{(x_2,y_2)}) - P(A_{(x_2,y_1)})) - (P(A_{(x_1,y_2)}) -$$
$$P(A_{(x_1,y_1)})) = P(A_{(x_2,y_2)}) - P(A_{(x_2,y_1)}) - P(A_{(x_1,y_2)}) + P(A_{(x_1,y_1)}) =$$
$$F_{\xi_1,\xi_2}((x_2, y_2)) - F_{\xi_1,\xi_2}((x_2, y_1)) - F_{\xi_1,\xi_2}((x_1, y_2)) + F_{\xi_1,\xi_2}((x_1, y_1)).$$

This ends the proof of theorem.

Assume that $(x, y) \in R^2$. If there exists double limit

$$\lim_{\Delta x, \Delta y \to 0} \frac{P(\{\omega : (\xi_1, \xi_2)(\omega) \in [x - \Delta x; x + \Delta x[\times]y - \Delta y; y + \Delta y[\})}{4\Delta x \Delta y},$$

then we say that joint distribution function F_{ξ_1,ξ_2} of 2-dimensional random vector (ξ_1, ξ_2) has the density function $f_{\xi_1,\xi_2}(x, y)$ in point (x, y) which is equal to the above-mentioned double limit.

We have the following proposition.

Theorem 10.2 *If a function of two variables F_{ξ_1,ξ_2} has the continuous partial derivatives of the first and second orders in any neighborhood of the point (x_0, y_0), then 2-dimensional random vector (ξ_1, ξ_2) has density function $f_{\xi_1,\xi_2}(x_0, y_0)$ in point (x_0, y_0), which can be calculated with the following formula*

$$f_{\xi_1,\xi_2}(x_0, y_0) = \frac{\partial^2 F_{\xi_1,\xi_2}(x_0, y_0)}{\partial x \partial y} = \frac{\partial^2 F_{\xi_1,\xi_2}(x_0, y_0)}{\partial y \partial x}.$$

Proof. Using Theorem 1, we get

$$P(\{\omega : (\xi_1, \xi_2)(\omega) \in [x - \Delta x; x + \Delta x[\times]y - \Delta y; y + \Delta y[\}) =$$

$$= F_{\xi_1,\xi_2}((x_0 + \Delta x, y_0)) - F_{\xi_1,\xi_2}((x_0 + \Delta x, y_0 - \Delta y)) -$$
$$F_{\xi_1,\xi_2}((x_0 - \Delta x, y_2)) + F_{\xi_1,\xi_2}((x_0 - \Delta x, y_0 - \Delta y)).$$

Without loss of generality, we can assume that points $(x_0 - \Delta x, y_0)$, $(x_0 - \Delta x, y_0 + \Delta y)$, $(x_0, y_0 - \Delta y), (x_0, y_0 + \Delta y)$ belong to such neighborhood of point (x_0, y_0) in which F_{ξ_1,ξ_2} has continuous partial derivatives of the first and second orders, respectively. Following Lagrange [1] theorem, there exists $\theta_1 \in]0;1[$ such that

$$[F_{\xi_1,\xi_2}((x_0 + \Delta x, y_0)) - F_{\xi_1,\xi_2}((x_0 + \Delta x, y_0 - \Delta y))] -$$
$$[F_{\xi_1,\xi_2}((x_0 - \Delta x, y_2)) - F_{\xi_1,\xi_2}((x_0 - \Delta x, y_0 - \Delta y))] =$$
$$= 2\Delta x \cdot \frac{\partial F_{\xi_1 \xi_2}}{\partial x}(x_0 - \Delta x + 2\theta_1 \Delta x, y_0 - \Delta y).$$

Again using the Lagrange theorem, we deduce the existence of $\theta_2 \in]0;1[$ such that

$$P(\{\omega : (\xi_1, \xi_2)(\omega) \in [x - \Delta x; x + \Delta x[\times]y - \Delta y; y + \Delta y[\}) =$$
$$= 4 \cdot \Delta x \cdot \Delta y \frac{\partial^2 F_{\xi_1,\xi_2}}{\partial y \partial x}(x_0 - \Delta x + 2\theta_1 \Delta x, y_0 - \Delta y + 2\theta_2).$$

Clearly,

$$\lim_{\Delta x, \Delta y \to 0} \frac{P(\{\omega : (\xi_1, \xi_2)(\omega) \in [x -; x + \Delta x[\times]y - \Delta y; y + [\})}{4\Delta x \Delta y} =$$
$$= \lim_{\Delta x, \Delta y \to 0} \frac{4 \cdot \Delta x \cdot \Delta y \frac{\partial^2 F_{\xi_1,\xi_2}}{\partial y \partial x}(x_0 - \Delta x + 2\theta_1 \Delta x, y_0 - \Delta y + 2\theta_2)}{4\Delta x \Delta y} =$$
$$= \frac{\partial^2 F_{\xi_1,\xi_2}(x_0, y_0)}{\partial y \partial x}.$$

The application of the well-known Schwarz [2] theorem ends the proof of theorem.

Example 10.1 2-dimensional random vector (ξ_1, ξ_2) is called distributed by Gaussian law, if its density function f_{ξ_1,ξ_2} has the following form

$$f_{\xi_1,\xi_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(x_1-a_1)^2}{2\sigma_1^2} - \frac{(x_2-a_2)^2}{2\sigma_2^2}} \quad (x_1, x_2 \in R),$$

where $a_1, a_2 \in R, \sigma_1 > 0, \sigma_2 > 0$.

Here we present some theorems (without proofs).

Theorem 10.3 *Let $D \subseteq R^2$ be some region of R^2 and f_{ξ_1,ξ_2} be a density function of 2-dimensional random vector (ξ_1, ξ_2). Then the following formula*

$$P(\{\omega : (\xi_1, \xi_2)(\omega) \in D\}) = \int\int_D f_{\xi_1,\xi_2}(x, y) dx dy$$

[1] *Lagrange, Joseph Louis* (25.1.1736 - 10.4.1813) - French mathematician, the member of Paris Academy of Sciences (1772).

[2] *Schwarz, Karl Hermann Amandus* (25.1.1843 – 30.11.1921) – German mathematician, the member of Berlin Academy of Sciences (1893).

holds.

Definition 10.4 *A reflection $g : R^n \to R$ is called measurable, if the following condition*

$$(\forall x)(x \in R \to \{(x_1, \cdots, x_n) : g(x_1, \cdots, x_n) < x\} \in B(R^n)).$$

holds. It is easy to show that $g : R^n \to R$ is measurable if and only if when

$$(\forall B)(B \in \mathcal{B}(R) \to g^{-1}(B) \in \mathcal{B}(R^n)),$$

where $g^{-1}(B) = \{(x_1, \cdots, x_n) : g(x_1, \cdots, x_n) \in B\}$.

Theorem 10.4 *Let $f_{\xi_1,\cdots,\xi_n}$ be a density function of random vector $(\xi_1, \cdots, \xi_n)$. Then for arbitrary measurable reflection $g : R^n \to R$ and for arbitrary $B \in B(R)$ we have:*

$$P(\{\omega : g((\xi_1, \cdots, \xi_n)(\omega)) \in B\}) = \int \cdots \int_{g^{-1}(B)} f_{\xi_1,\cdots,\xi_n}(x_1, \cdots, x_n)dx_1 \cdots dx_n.$$

Theorem 10.5*Let $(\xi_k)_{1 \le k \le n}$ be a family of independent random variables and $f_{\xi_1,\cdots,\xi_n}$ be the density function of random vector $(\xi_1, \cdots, \xi_n)$. If f_{ξ_i} is the density function of ξ_i for $1 \le i \le n$, then*

$$f_{\xi_1,\cdots,\xi_n}(x_1, \cdots, x_n) = \prod_{1 \le i \le n} f_{\xi_i}(x_i) \quad ((x_1, \cdots, x_n) \in R^n).$$

Definition 10.5 *Let $(\xi_k)_{1 \le k \le n}$ be a family of independent random variables and let ξ_k be normally distributed random variable with parameter (a_k, σ_k^2) for $1 \le k \le n$. Then $(\xi_1, \cdots, \xi_n)$ is called n-dimensional Gaussian vector and its density function, following Theorem 5, has the following form*

$$f_{\xi_1,\cdots,\xi_n}(x_1, \cdots, x_n) = \frac{1}{(\sqrt{2\pi})^n \prod_{k=1}^n \sigma_k} e^{-\sum_{k=1}^n \frac{(x_k - a_k)^2}{2\sigma_k^2}},$$

where $(x_1, \cdots, x_n) \in R^n$, $a_1, \cdots, a_n \in R, \sigma_1 > 0, \cdots, \sigma_n > 0$.

n-dimensional Gaussian vector $(\eta_1, \cdots, \eta_n)$ is called standard if

$$f_{\eta_1,\cdots,\eta_n}(x_1, \cdots, x_n) = \frac{1}{(\sqrt{2\pi})^n} e^{-\sum_{k=1}^n \frac{x_k^2}{2}} \quad ((x_1, \cdots, x_n) \in R^n).$$

Definition 10.6 *Assume that $(\xi_1, \cdots, \xi_n)$ is a Gaussian random vector. Function $P_{\xi_1,\cdots,\xi_n}$, defined with*

$$(\forall B)(B \in \mathcal{B}(R^n) \to P_{\xi_1,\cdots,\xi_n}(B) = P(\{\omega : (\xi_1(\omega), \cdots, \xi_n(\omega)) \in B\}),$$

is called n-dimensional Gaussian probability measure.

By using Theorem 10.3, we have

$$P_{\xi_1,\cdots,\xi_n}(B) = \overbrace{\int \cdots \int_B}^{n} \frac{1}{(\sqrt{2\pi})^n \prod_{k=1}^n \sigma_k} e^{-\sum_{k=1}^n \frac{(x_k-a_k)^2}{2\sigma_k^2}} dx_1 \cdots dx_n.$$

Let consider some examples.

Example 10.3 Let $(\xi_1, \cdots, \xi_n)$ be the n-dimensional Gaussian standard probability measure and $\prod_{k=1}^n [a_k, b_k] \subset R^n$. Then

$$P_{\xi_1,\cdots,\xi_n}(\prod_{k=1}^n [a_k, b_k]) = \prod_{k=1}^n [\Phi(b_k) - \Phi(a_k)].$$

Example 10.4(distribution χ_n^2). Let $(\xi_1, \cdots, \xi_n)$ be the n-dimensional Gaussian standard random vector and V_ρ^n be n-dimensional sphere with radius ρ and with center $O(0, \cdots, 0) \in R^n$. Then

$$P_{\xi_1,\cdots,\xi_n}(V_\rho^n) = \overbrace{\int \cdots \int_{V_\rho^n}}^{n} \frac{1}{(\sqrt{2\pi})^n} e^{-\sum_{k=1}^n \frac{x_k^2}{2}} dx_1 \cdots dx_n =$$

$$= \frac{1}{2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})} \times \int_0^\rho r^{n-1} \cdot e^{-\frac{r^2}{2}} dr,$$

where $\Gamma(\cdot)$ is Eulerian integral of the second type.

Distribution function $F_{\chi_n^2}$ of random variable $\chi_n^2 = \xi_1^2 + \cdots + \xi_n^2$ is called χ_n^2(chi square) -distribution, which has the following form:

$$F_{\chi_n^2}(x) = \begin{cases} 0, & \text{if } x \le 0, \\ \frac{1}{2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})} \times \int_0^{\sqrt{x}} r^{n-1} \cdot e^{-\frac{r^2}{2}} dr, & \text{if } x > 0. \end{cases}$$

Hence,

$$f_{\chi_n^2}(x) = \begin{cases} 0, & \text{if } x \le 0, \\ \frac{1}{2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})} \times x^{\frac{n-1}{2}} \cdot e^{-\frac{x}{2}} \frac{1}{2\sqrt{x}}, & \text{if } x > 0 \end{cases} =$$

$$= \begin{cases} 0, & \text{if } x \le 0, \\ \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \times x^{\frac{n}{2}-1} \cdot e^{-\frac{x}{2}}, & \text{if } x > 0. \end{cases}$$

Remark 10.1 We have applied the validity of the following fact

$$\overbrace{\int \cdots \int_{V_\rho}}^{n} f(\sqrt{\sum_{k=1}^n x_k^2}) dx_1 \cdots dx_n = 2 \cdot \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} \cdot \int_0^\rho r^{n-1} f(r) dr,$$

where f is an arbitrary continuous function defined on V_ρ.

Remark 10.2 Let $X_1, \cdots, X_n$ be an independent family of standard normally distributed real-valued variables. Then CHIDIST(x, n) calculates value

$$P(\{\omega : \omega \in \Omega \ \& \ \chi_n^2(\omega) = \sum_{k=1}^{n} X_k^2(\omega) > x\})$$

for $x \geq 0$. For example,CHIDIST$(2, 10) = 0,996340153$.

If we denote with Γ_n a standard n-dimensional Gaussian measure on R^n, then the command $1-$CHIDIST(r^2, n) calculates its value on n-dimensional ball $V(r, n)$ with radius r and the center at the zero of R^n. For example, $\Gamma_5(V(2,5)) = 1-$ CHIDIST$(2^2, 5) = 0,450584038$.

Example 10.5 Let $(e_k)_{1 \leq k \leq m}$ $(m \leq n)$ be the family of independent normed vectors in R^n and let $\xi_1, \cdots, \xi_m$ be the family of one-dimensional independent standard Gaussian random variables defined on $(\Omega, \mathcal{F}, P)$. Then measure μ, defined with

$$(\forall X)(X \in \mathcal{B}(R^n) \rightarrow \mu(X) = P(\{\omega : \sum_{k=1}^{m} \xi_k(\omega) e_k \in X\})),$$

is a Gaussian measure defined on R^n . Note that an analogous representation is valid for all Gaussian measures defined on R^n.

Example 10.7(Student's distribution t_n). Let $\xi_1, \cdots, \xi_m$ be the independent family of one-dimensional standard Gaussian random variables defined on $(\Omega, \mathcal{F}, P)$ and $G : R^{n+1} \rightarrow R$ be a measurable function defined with

$$g(x_1, \cdots, x_{n+1}) = \frac{x_{n+1}}{\sqrt{\frac{\sum_{k=1}^{n} x_i^2}{n}}}.$$

The random variable $t_n = g(\xi_1, \cdots, \xi_n, \xi_{n+1})$ is called Student's random variable with degree of freedom n.

Following Theorem 10.4, we have

$$F_{t_n}(x) = \int_{g^{-1}([-\infty, x))} \frac{1}{(\sqrt{2\pi})^n} e^{-\sum_{k=1}^{n} \frac{x_k^2}{2}} dx_1 \cdots dx_n.$$

It can be proved that

$$f_{t_n}(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \times (1 + \frac{x^2}{n})^{-\frac{n+1}{2}}, & \text{if } x > 0. \end{cases}$$

It is resonable to note that $M(t_n) = 0$, when $n > 1$.

For variance $D(t_n)$ we have

$$D(t_n) = \begin{cases} \frac{n}{n-1}, & \text{if } n > 2, \\ \infty, & \text{if } 0 < n \leq 2. \end{cases}$$

Remark 10.3 Statistical functions TDIST$(x;n;1)$ and TDIST$(x;n;2)$ calculate the values $P(\{\omega : t_n(\omega) > x\})$ and $P(\{\omega : |t_n(\omega)| > x\})$, respectively.For example, TDIST$(3;4;1) = 0,19970984$ and TDIST$(3;4;0) = 0,039941968$.

Example 10.8(Fisher's distribution $F_{\xi_{(k_1;k_2)}}$). Let $\xi_1, \cdots, \xi_{k_1+k_2}$ be the independent family of one-dimensional standard Gaussian random variables defined on $(\Omega, \mathcal{F}, P)$ and $G : R^{n+1} \to R$ be a measurable function defined with

$$g(x_1, \cdots, x_{k_1+k_2}) = \frac{\frac{\sum_{i=1}^{k} {}_1 x_i^2}{k_1}}{\frac{\sum_{i=k_1+1}^{k_1+k_2} x_i^2}{k_2}}.$$

The random variable $\xi_{(k_1;k_2)} = g(\xi_1, \cdots, \xi_n, \xi_{n+1})$ is called Fisher's random variable with degrees of freedom k_1 and k_2.

Following Theorem 10.4, we have

$$F_{\xi_{(k_1;k_2)}}(x) = \int_{g^{-1}([-\infty,x))} \frac{1}{(\sqrt{2\pi})^n} e^{-\sum_{k=1}^{n} \frac{x_k^2}{2}} dx_1 \cdots dx_n.$$

It can be proved that

$$f_{\xi_{(k_1;k_2)}}(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ 2(\frac{k_1}{k_2})^{\frac{k_1}{k_2}} \frac{\Gamma(\frac{k_1+k_2}{2})}{\Gamma(\frac{k_1}{2}) \times \Gamma(\frac{k_2}{2})} x^{k_1-1} (1 + \frac{k_1}{k_2} x^2)^{-\frac{k_1+k_2}{2}}, & \text{if } x > 0. \end{cases}$$

Remark 10.4 The statistical function FDIST$(x;k_1;k_2)$ calculates value $P(\{\omega : \xi_{(k_1;k_2)}(\omega) < x\})$. For example, FDIST$(2;5;6) = 0,211674328$.

We have the following proposition.

Theorem 10.6 *Let ξ_1 and ξ_2 be the independent random variables with density functions f_{ξ_1} and f_{ξ_2}, respectively. Then distribution function $F_{\xi_1+\xi_2}$ and density function $f_{\xi_1+\xi_2}$ of sum $\xi_1 + \xi_2$ are defined with:*

$$F_{\xi_1+\xi_2}(x) = \int_{-\infty}^{x} dx_2 \int_{-\infty}^{+\infty} f_1(x_1) f_2(x_2 - x_1) dx_1,$$

$$f_{\xi_1+\xi_2}(x) = \int_{-\infty}^{+\infty} f_1(x_1) f_2(x_2 - x_1) dx_1.$$

Proof. Sum $\xi_1 + \xi_2$ can be represented as continuous reflection g of the random vector (ξ_1, ξ_2), where $g(x_1, x_2) = x_1 + x_2$. We set $B = (-\infty, x)$. Using theorems 10.4 and 10.5, we get

$$F_{\xi_1+\xi_2}(x) = P(\{\omega : \xi_1(\omega) + \xi_2(\omega) < x\}) = P(\{\omega : g(\xi_1, \xi_2)(\omega) < x\}) =$$

$$= \int\int_{g^{-1}((-\infty;x))} f_{\xi_1}(x_1) f_{\xi_2}(x_2) dx_1 dx_2.$$

Note that

$$g^{-1}((-\infty;x)) = \{(x_1,x_2)|x_1+x_2<x\},$$

Hence

$$F_{\xi_1+\xi_2}(x) = \int\int_{\{(x_1,x_2)|x_1+x_2<x\}} f_{\xi_1}(x_1)f_{\xi_2}(x_2)dx_1dx_2 =$$

$$= \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{x-x_1} dx_2 f_{\xi_1}(x_1)f_{\xi_2}(x_2) = \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{x-x_1} d(x_1+x_2) f_{\xi_1}(x_1)f_{\xi_2}(x_2)$$

$$= \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{x} d\tau f_{\xi_1}(x_1)f_{\tau-\xi_2}(x_2) = \int_{-\infty}^{\tau} d\tau \int_{-\infty}^{+\infty} f_{\xi_1}(x_1)f_{\tau-\xi_2}(x_2)dx_1,$$

where $\tau = x_1 + x_2$.

Clearly, for ℓ_1-almost every x point of R satisfies the following equality,

$$f_{\xi_1+\xi_2}(x) = \frac{dF_{\xi_1+\xi_2}(x)}{dx} = \int_{-\infty}^{+\infty} f_{\xi_1}(x_1)f_{\xi_2}(x-x_1)dx_1.$$

The integral standing in the right in the above equality is called winding of functions f_1 and f_2 and is denoted by $f_{\xi_1} * f_{\xi_2}$.

It is not difficult to show that $f_{\xi_1} * f_{\xi_2} = f_{\xi_2} * f_{\xi_1}$, i.e.,

$$\int_{-\infty}^{+\infty} f_{\xi_1}(x_1)f_{\xi_2}(x-x_1)dx_1 = \int_{-\infty}^{+\infty} f_{\xi_1}(x-x_1)f_{\xi_2}(x_1)dx_1.$$

Tests

10.1. Assume that the distribution of 2-dimensional discrete random vector (ξ_1,ξ_2) is given in the following table

(ξ_1,ξ_2)	$(4,3)$	$(4,10)$	$(4,12)$	$(5,3)$	$(5,12)$
P	$0,17$	$0,13$	$0,25$	$0,2$	$0,25$

.

Then

1) the distribution law of ξ_1 is given in the table

a)

ξ_1	4	5
P	$0,55$	$0,45$

,

b)

ξ_1	4	5
P	$0,55$	$0,45$

;

2) the distribution law of ξ_2 is given in the table

a)

ξ_2	3	10	12
P	$0,37$	$0,13$	$0,5$

,

b)

ξ_2	3	10	12
P	$0,35$	$0,15$	$0,5$

;

3) $F_{\xi_1,\xi_2}(4,5;10,5)$ is equal to

a) $0,36$, b) $0,34$, c) $0,32$, d) $0,3$;

4) $P(\{\omega : (\xi_1(\omega),\xi_2(\omega)) \in [1,5]\times[5,8]\})$ is equal to

a) $0,39$, b) $0,38$, c) $0,37$, d) $0,36$.

10.2. Distribution laws of two independent random variables ξ_1 and ξ_2 are given in the following tables

ξ_1	2	3
P	0,7	0,3

,

ξ_2	−2	2
P	0,3	0,7

,

respectively. Then the distribution law of $\xi_1 \cdot \xi_2$ is given in the table

a)

$\xi_1 \cdot \xi_2$	−6	−4	4	6
P	0,08	0,22	0,48	0,22

,

b)

$\xi_1 \cdot \xi_2$	−6	−4	4	6
P	0,09	0,21	0,49	0,21

.

10.3. A distribution function of 2-dimensional random vector (ξ_1,ξ_2) is defined with the following formula

$$F_{\xi_1,\xi_2}(x_1,x_2) = \begin{cases} (1-e^{-4x_1})(1-e^{-2x_2}), & x_1 \ge 0,\ x_2 \ge 0, \\ 0, & x_1 < 0 \text{ or } x_2 < 0. \end{cases}$$

Then

a)

$$f_{\xi_1,\xi_2}(x_1,x_2) = \begin{cases} (6e^{-4x_1-2x_2}), & x_1 \ge 0,\ x_2 \ge 0, \\ 0, & x_1 < 0 \text{ or } x_2 < 0, \end{cases}$$

b)

$$f_{\xi_1,\xi_2}(x_1,x_2) = \begin{cases} (6e^{-4x_1-2x_2}), & x_1 \ge 0,\ x_2 \ge 0, \\ 0, & x_1 < 0 \text{ or } x_2 < 0. \end{cases}$$

10.4.The density function of 2-dimensional random vector (ξ_1,ξ_2) is defined with

$$f_{\xi_1,\xi_2}(x_1,x_2) = \frac{20}{\pi^2(16+x_1^2)(25+x_2^2)} \quad ((x_1,x_2)\in R^2).$$

Then

a)

$$F_{\xi_1,\xi_2}(x_1,x_2) = (\frac{1}{2}+\frac{1}{\pi}arctg(\frac{x_1}{8}))(\frac{1}{2}+\frac{1}{\pi}arctg(\frac{x_2}{10})),$$

b)

$$F_{\xi_1,\xi_2}(x_1,x_2) = (\frac{1}{2}+\frac{1}{\pi}arctg(\frac{x_1}{4}))(\frac{1}{2}+\frac{1}{\pi}arctg(\frac{x_2}{5})).$$

10.5. It is known that the freedom coefficients of the general solution of differential equation $y'' + 5y' + 6y = 0$ are independent random variables uniformly distributed in interval $(0,1)$. The probability that a general solution of the differential equation will get value $\geq 0,5$ in point $x=0$, is equal to

a) $0,5$, b) $0,75$, c) $0,6$, d) $0,85$.

10.6. It is known that the freedom coefficients of the general solution of the differential equation $y'' + y = 0$ are independent random variables normally distributed with parameters $(0,1)$. Then the probability that the general solution y satisfies the following conditions

$$y(0) \in (0,2) \ \& \ y(\frac{\pi}{2}) \in (-2,1),$$

is equal to

a) $0,2245785$, b) $0,7767678$, c) $0,3665582$, d) $0,8598760$.

10.7.It is known that the freedom coefficients of the general solution of the differential equation $y'' - \ln 6y' + \ln 2 \ln 3y = 0$ are independent random variables uniformly distributed on the interval $(0,1)$. The probability that general solution y satisfies the following conditions

$$y(0) \in (\infty,1) \ \& \ y(1) \in (-\infty,2),$$

is equal to

a) $\frac{1}{2}$, b) $\frac{1}{3}$, c) $\frac{1}{4}$, d) $\frac{1}{5}$.

Chapter 11

Chebishev inequalities. The law of three σ

Let (Ω, F, P) be a probability space.
The following proposition is valid.

Theorem 11.1 (Chebishev's [1] I inequality). *For arbitrary non-negative random variable* ξ *and for arbitrary positive real number* ϵ *the following inequality*

$$P(\{\omega : \xi(\omega) \geq \epsilon\}) \leq \frac{M\xi}{\epsilon}.$$

holds.

Proof. Clearly,

$$M\xi = M(\xi \cdot I_{\Omega}) = M(\xi \cdot I_{\{\omega:\xi(\omega)\geq\epsilon\}} + \xi \cdot I_{\{\omega:\xi(\omega)<\epsilon\}}) \geq$$

$$\geq M(\xi \cdot I_{\{\omega:\xi(\omega)\geq\epsilon\}}) \geq \epsilon \cdot P(\{\omega : \xi(\omega) \geq \epsilon\}).$$

Finally, we get

$$P(\{\omega : \xi(\omega) \geq \epsilon\}) \leq \frac{M\epsilon}{\epsilon}.$$

This ends the proof of theorem.

Theorem 11.2 (Chebishev's II inequality). *For arbitrary random variable* η *and for arbitrary positive number* $\delta > 0$ *the following inequality*

$$P(\{\omega : |\eta(\omega) - M\eta| \geq \sigma\}) \leq \frac{D\eta}{\sigma^2}.$$

holds.

Proof. We set :

[1] *P.Chebishev* [4(16).5.1821. - 26.11.(8.12)1894] - Russian mathematician, Academician of Petersburg Academy of Sciences (1856), of Berlin Academy of Sciences (1871) and of Paris Academy of Sciences (1874).

$$\xi(\omega) = (\eta(\omega) - M\eta)^2, \ \epsilon = \sigma^2.$$

Following Chebishev's I inequality, we get

$$P(\{\omega : (\eta(\omega) - M\eta)^2 \geq \sigma^2\}) \leq \frac{M(\eta - M\eta)^2}{\sigma^2}.$$

Note that

$$\{\omega : (\eta(\omega) - M\eta)^2 \geq \sigma^2\} = \{\omega : |\eta(\omega) - M\eta| \geq \sigma\}.$$

Finally, we get

$$P(\{\omega : |\eta(\omega) - M\eta| \geq \sigma\}) \leq \frac{D\eta}{\sigma^2}.$$

Example 11.1 Assume that we survey the moon and measure its diameter. Assume also that the results of survey are independent random variables $\xi_1, \cdots, \xi_n$. Assume that a is the value of moon's diameter. Then $|\xi_k(\omega) - a|$ will be mistake in the k-th experiment $(1 \leq k \leq n)$. The value $\sqrt{M(\xi_k - a)^2} = \sqrt{D\xi_k}$ will be error mean square deviation. Assume also that the following conditions

a) $M\xi_k = a$;

b) $\sqrt{D\xi_k} = 1$;

c) $(\xi_k)_{1 \leq k \leq n}$ are independent,

hold for $1 \leq k \leq n$.

It is natural that value $J_n = \frac{1}{n}(\xi_1 + \cdots + \xi_n)$ may be considered as an estimation of parameter a. There naturally arises the following problem:

Haw many measures are sufficient to establish the validity of the following stochastic inequality

$$P(\{\omega : |J_n(\omega) - a| \leq 0,1\}) \geq 0,95 \ ?$$

Clearly, on the one hand, we have

$$P(\{\omega : |J_n(\omega) - a| > 0,1\}) \leq 0,05.$$

On the other hand, we have

$$P(\{\omega : |J_n(\omega) - a| > 0,1\}) \leq \frac{D(J_n)}{(0,1)^2} =$$

$$= \frac{\frac{1}{n^2}\sum_{k=1}^{n} D\xi_k}{0,01} = \frac{\frac{1}{n^2}n}{0,01} = \frac{100}{n}.$$

From the latter inequality we deduce that the smallest natural number $n = n_{\mathrm{C}}$ for which inequality $\frac{100}{n_{\mathrm{C}}} \leq 0,05$ holds, is equal to 2000.

Hence, we get

$$P(\{\omega : |J_{2000}(\omega) - a| \leq 0,1\}) \geq 0,95,$$

i.e., 2000 measures are sufficient to be sure with probability $\geq 0,95$ that the mean J_{2000} will be deviated from the length a of the moon's diameter for no more than $0,1$.

Theorem 11.3 (The law of three σ). *For arbitrary random variable* ξ *the following inequality*

$$P(\{\omega : |\xi(\omega) - M\xi| \geq 3\sigma\}) \leq \frac{1}{9}$$

holds.

Proof. Indeed, using Chebishev's II inequality, we obtain

$$P(\{\omega : |\xi(\omega) - M\xi| \geq 3\sigma\}) \leq \frac{D\xi}{9\sigma^2} = \frac{1}{9}.$$

Tests

11.1. It is known that $D\xi = 0,001$. Using Chebishev's inequality the probability of event $\{\omega : |\xi(\omega) - M\xi| < 0,1\}$ is estimated from bellow by the number, which is equal to

a) $0,8$, b) $0,9$, c) $0,98$, d) $0,89$.

11.2. We have $D\xi = 0,004$. It is established that $P(\{\omega : |\xi(\omega) - M\xi| < \epsilon\}) \geq 0,9$; Then ϵ is equal to

a) $0,1$, b) $0,2$, c) $0,3$, d) $0,4$.

11.3. The distribution law of random variable ξ has the following form

ξ	$0,3$	$0,6$
P	$0,2$	$0,8$

.

Using Chebishev's inequality the probability of event $\{\omega : |\xi(\omega) - M\xi| < \epsilon\}$ is estimated from the below with the following number

a) $0,86$, b) $0,87$, c) $0,88$, d) $0,89$.

11.4. Mean consumption of water in populated area per one day is 50000 liters. Using Chebishev's inequality estimate from below the probability that in this area water consumption per one concrete day will be ≤ 150000 liters.

a) $\frac{1}{3}$, b) $\frac{2}{3}$, c) $\frac{1}{4}$, d) $\frac{1}{2}$.

11.5. The probability that an event A occured in separate experiment is equal to $0,7$. Let denote with ν_n a fraction the numerator of which is equal to the occurred number of event A in n independent experiments, and the denominator of which is equal to n. Minimal natural number n, such that $P\{\omega : |\nu_n(\omega) - p| < 0,06\} \geq 0,78$ is equal to

a) 327, b) 427, c) 527, d) 627.

11.6. Assume we throw a die 1200 times. Let ξ denote the number of experiments when number 1 has been thrown. Use Chebishev's inequality for estimation from below of the probability of event $|\omega : \xi(\omega) \leq 800\}$.

a) $0,74$, b) $0,75$, c) $0,76$, d) $0,77$.

11.7. Assume that we throw a die 10000 times. Use Chebishev's inequality to estimate from below the probability that the relative frequency of event- " Number 6 is thrown by us" would be deviated from number $\frac{1}{6}$ with probability $\leq 0,01$.

a) $0,84$, b) $0,85$, c) $0,86$, d) $0,87$.

11.8. Assume that we shot the gun 600 times and the probability of hitting the target in a separate experiment is equal to $0,6$. Use the Chebishev's inequality for estimation from the below of the probability that the number of successful shots will be deviated from number 360 by no more than 20.

a) $0,63$, b) $0,64$, c) $0,65$, d) $0,66$.

11.9. It is known that the mean weight of a bun is 50 grams. Use the Chebishev's inequality for estimation from below of the probability that the weight of randomly chosen bun will be ≤ 90 gram.

a) $\frac{1}{3}$, b) $\frac{4}{9}$, c) $\frac{5}{9}$, d) $\frac{2}{3}$.

11.10. Use the Chebishev's inequality for estimation from below of the probability that the mean speed of a projectile, accidently shot from a gun is $\leq 800\ \frac{km}{sec}$ relative to the hypothesis that the mean speed of the projectile is equal to $500\ \frac{km}{sec}$.

a) $\frac{3}{7}$, b) $\frac{3}{8}$, c) $\frac{1}{3}$, d) $\frac{3}{10}$.

Chapter 12

Limiting theorems

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $(X_k)_{k \in N}$ be an infinite sequence of random variables.

Definition 12.1. *We say that a sequence of random variables* $(X_k)_{k \in N}$ *converges to number* $a \in R$ *in the sense of probability if for arbitrary positive number* $\epsilon > 0$ *the following condition*

$$\lim_{k \to \infty} P(\{\omega : |X_k(\omega) - a| < \epsilon\}) = 1$$

holds.

This fact is denoted with

$$\lim_{k \to \infty} X_k \stackrel{p}{=} a.$$

We have the following proposition.

Theorem 12.1 (Chebishev). *Assume that mathematical variances of the random variables* X_k $(k \in N)$ *are jointly bounded, i.e.,*

$$(\exists c)(c \in R^+ \to (\forall n)(n \in N \to DX_n < c)).$$

Then

$$\lim_{n \to \infty} \frac{1}{n}(\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} MX_k) \stackrel{p}{=} 0.$$

Proof. Following Definition 12.1, it is necessary and sufficient to show the validity of the following condition

$$(\forall \epsilon)(\epsilon > 0 \to \lim_{n \to \infty} P(\{\omega : |\frac{1}{n}(\sum_{k=1}^{n} X_k(\omega) - \sum_{k=1}^{n} MX_k) - 0| < \epsilon\}) = 1).$$

Setting

$$Y_n(\omega) = \frac{1}{n}(\sum_{k=1}^{n} X_k(\omega) - \sum_{k=1}^{n} MX_k),$$

95

we get

$$MY_n = M\frac{1}{n}(\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} MX_k) = \frac{1}{n}\sum_{k=1}^{n} MX_k - \frac{1}{n}\sum_{k=1}^{n} MX_k = 0,$$

$$DY_n = D(\frac{1}{n}(\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} MX_k)) = D(\frac{1}{n}\sum_{k=1}^{n} X_k) = \frac{1}{n^2}D(\sum_{k=1}^{n} X_k) \le \frac{nc}{n^2} = \frac{c}{n}.$$

Using Chebishev's II inequality, we get

$$P(\{\omega : |Y_n(\omega) - MY_n| < \epsilon\}) \ge 1 - \frac{DY_n}{\epsilon^2} \ge 1 - \frac{c}{n\epsilon^2}.$$

Hence,

$$\lim_{n\to\infty} P(\{\omega : |Y_n(\omega) - MY_n| < \epsilon\}) \ge 1,$$

i.e.,

$$\lim_{n\to\infty} \frac{1}{n}(\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} MX_k) \overset{p}{=} 0.$$

This ends the proof of theorem.

As corollary of Theorem 12.1, we get

Theorem 12.2 (Bernoulli). Let $(Z_k)_{k\in N}$ be a sequence of independent simple discrete random variables distributed by Bernoulli law with parameter p. Then

$$\lim_{n\to\infty} (\frac{1}{n}\sum_{k=1}^{n} Z_k) \overset{p}{=} p.$$

Proof. The sequence of random variables $(Z_k)_{k\in N}$ satisfies the conditions of Theorem 12.1. Hence,

$$\lim_{n\to\infty} \frac{1}{n}(\sum_{k=1}^{n} Z_k - \sum_{k=1}^{n} MZ_k) \overset{p}{=} 0.$$

But

$$\lim_{n\to\infty} \frac{1}{n}(\sum_{k=1}^{n} MZ_k) \overset{p}{=} p,$$

$$\lim_{n\to\infty} (\frac{1}{n}\sum_{k=1}^{n} Z_k - p) \overset{p}{=} 0,$$

which is equivalent to the following condition

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} Z_k \overset{p}{=} p.$$

This ends the proof of Theorem.

Theorem 12.3 *If f is a continuous real-valued function defined on $[0,1]$, then the sequence of random variables $(Mf(\frac{1}{n}(\sum_{k=1}^{n} Z_k))_{n\in N}$ is uniformly converged to function $f(p)$ in interval $[0,1]$, where $(Z_k)_{k\in N}$ is the sequence of independent random variables distributed by the Bernoulli law with parameter p.*

Proof. For arbitrary $\epsilon > 0$, we have

$$M|f(\frac{1}{n}(\sum_{k=1}^{n} Z_k)) - f(p)| \leq M(|f(\frac{1}{n}\sum_{k=1}^{n} Z_k) - f(p)| \cdot I_{\{\omega:|f(\frac{1}{n}\sum_{k=1}^{n} Z_k)-f(p)|\leq\epsilon\}} +$$

$$+M(|f(\frac{1}{n}\sum_{k=1}^{n} Z_k) - f(p)| \cdot I_{\{\omega:|f(\frac{1}{n}\sum_{k=1}^{n} Z_k)-f(p)|>\epsilon\}} \leq \sup_{|x|\leq\epsilon} |f(p+x) - f(p)| + o(n),$$

which ends the proof of Theorem 12.3.

Remark 12.1 If f is a continuous real-valued function defined on $[0,1]$, then

$$\lim_{n\to\infty} \sum_{k=0}^{n} f(\frac{k}{n}) C_n^k x^k (1-x)^{n-k} = f(x)$$

for $x \in [0,1]$; Note here that the above-mentioned convergence is uniform on $[0,1]$. The last relation is a different entry of the uniform convergence of sequence

$$(Mf(\frac{1}{n}(\sum_{k=1}^{n} Z_k)))_{n\in N} = (\sum_{k=0}^{n} f(\frac{k}{n}) C_n^k p^k (1-p)^{n-k})_{n\in N}$$

to function f with respect to p on interval $[0,1]$. From this fact we get the well known Weierstras [1]theorem about approximation of the continuous real-valued function by polinomials. Note here also that these polinomials have the following form

$$\sum_{k=0}^{n} f(\frac{k}{n}) C_n^k x^k (1-x)^{n-k} \quad (n \in N).$$

These polinomials are called the Bershtein [2] polinomials. As corollary of Theorem 1 we get the following proposition.

Theorem 12.4 (The law of large numbers). *Let $(X_k)_{k\in N}$ be a sequence of identically distributed random variables. Assume also that $MX_k = a$ and $DX_k = \sigma^2 < \infty$; Then an arithmetic mean of random variables converges in probability sense to number a, i.e.,*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k \overset{p}{=} a.$$

[1]*Weierstras, Karl Theodor Wilhelm* (31.10.1815 - 19.2.1897) - German mathematician; Academician of Petersburg Academy of Sciences(1864); Professor of Berlin University (1856).

[2]*Bershtein, S* (22.2(5,3).1880 - 26.10.1968) - Russian mathematician; Academician of the Ukrainian Academy of Sciences(1925) and academician of the USSR Academy of Sciences (1929).

Proof. Since the sequence of random variables $(X_k)_{k\in N}$ satisfies the conditions of Theorem 1, we get

$$\lim_{n\to\infty}\frac{1}{n}(\sum_{k=1}^{n}X_k-\sum_{k=1}^{n}MX_k)\stackrel{p}{=}0.$$

Clearly, $\frac{1}{n}\sum_{k=1}^{n}MX_k=\frac{na}{n}=a$. Note also that

$$\lim_{n\to\infty}\frac{1}{n}(\sum_{k=1}^{n}X_k-\sum_{k=1}^{n}MX_k)\stackrel{p}{=}\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}X_k\stackrel{p}{=}0.$$

The last equality is equivalent to the following equality

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}X_k\stackrel{p}{=}0.$$

This ends the proof of theorem.

Remark 12.2. If $(X_k)_{k\in N}$ is a sequence of independent random variables normally distributed with parameters $(0,\sigma^2)$, then

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}X_k^2\stackrel{p}{=}\sigma^2.$$

Remark 12.3 Assume that the probability of occuring of event A in each experiment is equal to p. Let ν_n denote a relative frequency of the event A in n independent experiments. Using the law of Large numbers it is not difficult to show that for arbitrary positive number ϵ the following condition

$$\lim_{n\to\infty}P(\{\omega:|\nu_n(\omega)-p|<\epsilon\})=1$$

holds, i.e.,

$$\lim_{n\to n}\nu_n\stackrel{p}{=}p.$$

Tests

12.1. Let $(\xi_k)_{k\in N}$ be a sequence of independent random variables uniformly distributed on (a,b). Then

1)

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\xi_k\stackrel{p}{=}A,$$

where A is equal to

a) $\frac{a+b}{2}$, b) $\frac{b-a}{2}$, c) $\frac{a+b}{3}$, d) $\frac{b-a}{3}$;

2)

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\xi_k^2 \overset{p}{=} B,$$

where B is equal to

a) $\frac{(a+b)^2}{2}$, b) $\frac{a^2+ab+b^2}{3}$, c) $\frac{(a+b)^3}{3}$, d) $\frac{(b-a)}{12}$.

12.2. Let $(\xi_k)_{k\in N}$ be a sequence of independent Poisson random variables with parameter $\lambda = 5$ .Then

1)

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\xi_k \overset{p}{=} A,$$

where A is equal to

a) 3, b) 4, c) 5, d) 6;

2)

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\xi_k^2 \overset{p}{=} B,$$

where B is equal to

a) 28, b) 29, c) 30, d) 31.

12.3. Let $(\xi_k)_{k\in N}$ be a sequence of independent Bernoulli random variables with parameter p. Then for arbitrary non-zero real number s we have

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\xi_k^s \overset{p}{=} A,$$

where A is equal to

a) p, b) pq, c) p^s, d) q^s.

12.4. Let $(\xi_k)_{k\in N}$ be a sequence of independent Cantor's random variables. Then

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\xi_k \overset{p}{=} A,$$

where A is equal to

a) $0,3$, b) $0,5$, c) $0,6$, d) $0,7$.

12.5. Let $(\xi_k)_{k\in N}$ be a sequence of independent geometric random variables with parameter $q = 0,3$. Then

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\xi_k \overset{p}{=} A,$$

where A is equal to

a) $\frac{29}{49}$, b) $\frac{30}{49}$, c) $\frac{31}{49}$, d) $\frac{32}{49}$.

12.6. The sequence of functions $(\sum_{k=0}^{n}(\frac{k}{n})^3 C_n^k x^k (1-x)^{n-k})_{n\in N}$ is uniformly converged to function f in the interval $[0,1]$, where $f(x)$ is equal to

a) x^2; b) x^3; c) x^4; d) x^5;

12.7. The sequence of functions $(\sum_{k=0}^{n} \sin((\frac{k}{n})^2) C_n^k x^k (1-x)^{n-k})_{n \in N}$ is uniformly converged to function f in the interval $[0,1]$, where $f(x)$ is equal to

a) $\sin(x^2)$, b) $\sin(x^3)$, c) $\sin(x^4)$, d) $\sin(x^4)$.

12.8. Let $(\xi_n)_{n \in N}$ be a sequence of independent random variables with identical distribution functions. The distribution law of ξ_n is given in the following table

ξ_n	$-\sqrt{n+1}$	0	$\sqrt{n+1}$
P	$\frac{1}{n+1}$	$1-\frac{2}{n+1}$	$\frac{1}{n+1}$

.

Then the application of the Chebishev theorem with respect to the above-mentioned sequence

a) is possible, b) is not possible.

12.9. Let $(\xi_k)_{k \in N}$ be a sequence of independent Poisson random variables with parameter k. Then the application of the Chebishev theorem with respect to the above-mentioned sequence

a) is possible, b) is not possible.

12.10. Let $(\xi_k)_{k \in N}$ be a sequence of independent random variables and ξ_k be uniformly distributed on $[0;\sqrt{k}]$ for $k \in N$. Then an application of Chebishev theorem with respect to the above-mentioned sequence

a) is possible, b) is not possible.

Chapter 13

The Method of Characteristic Functions

Definition 13.1. *Let* (Ω, F, P) *be a probability space. A characteristic function of random variable* $\xi : \Omega \to R$ *is called the mathematical expectation of complex function* $e^{it\xi} = \cos(t\xi) + i\sin(t\xi)$ *and is denoted with* Φ_ξ *, i. e.,*

$$\Phi_\xi(t) = Me^{it\xi} \quad (t \in R).$$

Let ξ be a discrete random variable, i. e.,

$$\xi(\omega) = \sum_{k \in N} x_k I_{A_k}(\omega) \quad (\omega \in \Omega),$$

where $(A_k)_{k \in N}$ is a family of pairwise disjoint events covered Ω and $(x_k)_{k \in N}$ be a sequence of real numbers. In this situation, we have

$$\Phi_\xi(t) = Me^{it\xi} = \sum_{k \in N} e^{itx_k} P(A_k) \quad (t \in R).$$

When f_ξ is the density function of absolutely continuous random variable ξ, then we get

$$\Phi_\xi(t) = \int_{-\infty}^{+\infty} e^{itx} f_\xi(x)dx \quad (t \in R).$$

From the last relation we see that $\Phi_\xi(t)$ is Fourier transformation of f_ξ. From the course of mathematical analysis it is well known that if we have the Fourier [1]transformation Φ_ξ of function f_ξ then in some situations we can restore function f_ξ with function Φ_ξ. In particular,

$$f_\xi(x) = \int_{-\infty}^{+\infty} e^{-itx} \Phi_\xi(t)dt \ (x \in R).$$

[1]*Fourier, Jean Baptiste Joseph* (1.3. 1768-16. 5. 1830)-French mathematician, the member of Paris Academy of Sciences (1817), the member of Petersburg Academy of Sciences (1829).

101

The above mentioned relation is called Fourier inverse transformation .

Let consider some properties of characteristic function.

Theorem 13.1 *For arbitrary random variable* $\xi : \Omega \to R$ *we have*

$$\Phi_\xi(0) = 1.$$

Proof. Since $\Phi_\xi(t) = Me^{it\xi}$ $(t \in R)$, we have

$$\Phi_\xi(0) = M\ 1 = 1.$$

Theorem 13.2. *For every random variable* ξ *the following condition*

$$(\forall t)(t \in R \to |\Phi_\xi(t)| \le 1).$$

holds.

Proof. Note that for every random variable η the following condition

$$|M\eta| \le M|\eta|$$

holds. Hence,

$$|\Phi_\xi(t)| = |Me^{it\xi}| \le M|e^{it\xi}| = M1 = 1.$$

Theorem 13.3. *For arbitrary random variable* ξ *we have*

$$\Phi_\xi(-t) = \overline{\Phi_\xi(t)}.$$

Proof.

$$\Phi_\xi(-t) = M(e^{-it\xi}) = M(\cos(-t\xi) + i\sin(-t\xi)) = M(\cos(-t\xi)) + iM(\sin(-t\xi)) =$$

$$= M(\cos(t\xi)) - iM(\sin(t\xi)) = \overline{M(\cos(t\xi)) + iM(\sin(t\xi))} = \overline{Me^{it\xi}} = \overline{\Phi_\xi(t)}.$$

The following two facts are presented without proofs.

Theorem 13.4 *Characteristic function* $\Phi_\xi(t)$ *of random variable* ξ *is uniformly continuous on the real axis.*

Theorem 13.5 (Uniqueness Theorem). *The distribution function of the random variable is uniquely defined with its characteristic function.*

Theorem 13.6 *If random variables* ξ *and* η *are linearly related with* $\xi(\omega) = a\eta(\omega) + b$ $(a \in R,\ b \in R, \omega \in \Omega))$, *then*

$$\Phi_\xi(t) = e^{itb}\Phi_\eta(at).$$

Proof. Indeed,

$$\Phi_\xi(t) = \Phi_{a\eta+b}(t) = Me^{i(a\eta+b)t} = Me^{ibt}Me^{ia\eta t} = e^{itb}\Phi_\eta(at).$$

Theorem 13.7 *The characteristic function of the sum of two random variables is equal to the product of characteristic functions of the corresponding random variables.*

Proof. Let ξ and η be independent random variables. Then complex random variables $e^{it\xi}$ and $e^{it\eta}$ are independent, too. Now using the property of mathematical expectation we get

$$\Phi_{\xi+\eta}(t) = Me^{it(\xi+\eta)} = Me^{it\xi}Me^{it\eta} = \Phi_\xi(t)\cdot\Phi_\eta(t).$$

Theorem 13.7 admits the following generalization.

Theorem 13.8 *If $(\xi_k)_{1\le k\le n}$ is the finite family of independent random variables, then*

$$\Phi_{\sum_{k=1}^n \xi_k}(t) = \prod_{k=1}^n \Phi_{\xi_k}(t) \quad (t \in R).$$

Let ξ be a random variable and let $(\xi_k)_{k\in N}$ be a sequence of random variables.

Definition 13.2 *The sequence of random variables $(\xi_k)_{k\in N}$ is called weakly converged to random variable ξ if sequence $(F_{\xi_n})_{n\in N}$ is convergent to function F_ξ at its continuity points.*

We present one fundamental fact from the probability theory without proof.

Theorem 13.9 *The sequence of random variables $(\xi_k)_{k\in N}$ weakly converges to the random variable ξ if and only if the sequence of characteristic functions $(\Phi_{\xi_n})_{n\in N}$ converges to the characteristic function Φ_ξ .*

Let consider some examples.

Example 13.1 Let ξ be a Binomial random variable with parameters (n,p), i.e.,

$$P(\{\omega : \xi(\omega) = k\}) = C_n^k p^k(1-p)^{n-k} \quad (0 \le k \le n).$$

Then

$$\Phi_\xi(t) = Me^{it\xi} = \sum_{k=0}^n e^{itk}\cdot C_n^k p^k(1-p)^{n-k} =$$

$$= \sum_{k=0}^n C_n^k(e^{it}p)^k(1-p)^{n-k} = [pe^{it} + (1-p)]^n = (pe^{it}+q)^n,\ q = 1-p.$$

Example 13.2 Let ξ be a Poisson random variable with parameter λ, i. e.,

$$P(\{\omega : \xi(\omega) = k\}) = \frac{\lambda^k}{k!}e^{-\lambda} \quad (k = 0,1,\cdots),$$

Then

$$\Phi_\xi(t) = Me^{it\xi} = \sum_{k=0}^{\infty} \frac{e^{itk}}{k!} e^{-\lambda} = \sum_{k=0}^{\infty} \frac{(e^{it}\lambda)^k}{k!} e^{-\lambda} =$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^{it}\lambda)^k}{k!} = e^{\lambda e^{it} - \lambda} = e^{\lambda(e^{it}-1)}.$$

Example 13.3 Let ξ be a random variable uniformly distributed in the interval $(a;b)$, i. e.,

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

Then

$$\Phi_\xi(t) = Me^{it\xi} = \int_{-\infty}^{+\infty} e^{itx} f_\xi(x)dx = \int_a^b \frac{e^{itx}}{b-a} dx = \frac{1}{(b-a)it} e^{itx}\Big|_a^b =$$

$$= \frac{1}{(b-a)it}(e^{itb} - e^{ita}).$$

Example 13.4 Let ξ be a normally distributed random variable with parameters (a, σ^2), i. e.,

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

Then

$$\Phi_\xi(t) = Me^{it\xi} = \int_{-\infty}^{+\infty} e^{itx} f_\xi(x)dx =$$

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

Setting $z = \frac{x-a}{\sigma} - it\sigma$, we get

$$\frac{x-a}{\sigma} = z + it\sigma, \ x = a + \sigma z + it\sigma^2, \ dx = \sigma dz.$$

With simple transformation we get

$$\Phi_\xi(t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty - it\sigma}^{+\infty - it\sigma} e^{it(a+z\sigma+it\sigma^2) - \frac{(z+it\sigma)^2}{2}} \sigma dz =$$

$$= e^{iat - \frac{\sigma^2 t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty - it\sigma}^{+\infty - it\sigma} e^{-\frac{z^2}{2}} dz \ (t \in R).$$

Using the well known fact from mathematical analysis we get

$$(\forall b)(b \in R \to \int_{-\infty - it\sigma}^{+\infty - it\sigma} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi}),$$

Finally we get

$$\Phi_\xi(t) = e^{iat-\frac{\sigma^2 t^2}{2}} \quad (t \in R).$$

Remark 13.1 Characteristic function Φ_ξ of normally distributed random variable ξ with parameter $(0,1)$ has the following form

$$\Phi_\xi(t) = e^{-\frac{t^2}{2}} \quad (t \in R).$$

Example 13.5 Let ξ be exponential random variable with parameter λ , i.e.,

$$f_\xi(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$

Then

$$\Phi_\xi(t) = Me^{it\xi} = \int_{-\infty}^{+\infty} e^{itx} f_\xi(x)dx = \int_0^{+\infty} e^{itx-\lambda x}dx =$$

$$=\lambda \int_0^{+\infty} (e^{it}e^{-\lambda})^x dx = \lambda \int_0^{+\infty} (e^{it-\lambda})^x dx = \lambda \frac{(e^{it-\lambda})^x}{(it-\lambda)}\Big|_0^{+\infty} = \frac{\lambda}{\lambda - it}.$$

Example 13.6 Let $\xi = c$ be a constant random variable, i.e.,

$$P(\{\omega : \xi(\omega) = c\}) = 1.$$

Then

$$\Phi_\xi(t) = Me^{it\xi} = Me^{itc} = e^{itc}.$$

Let consider one application of the method of characteristic functions.

Theorem 13.10 *(Lindeberg [2] -Levy [3]). If $(\xi_k)_{k\in N}$ be a sequence of independent identically distributed random variables, then the sequence of random variables*

$$\left(\frac{\sum_{k=1}^n \xi_k - M(\sum_{k=1}^n \xi_k)}{\sqrt{D\sum_{k=1}^n \xi_k}}\right)_{n\in N}$$

is weakly converged to the standard normally distributed random variable, i.e.,

$$(\forall x)(x \in R \to \lim_{n\to\infty} P\big(\{\omega : \frac{\sum_{k=1}^n \xi_k - M(\sum_{k=1}^n \xi_k)}{\sqrt{D\sum_{k=1}^n \xi_k}} < x\}\big) =$$

[2] *Lindeberg, J.W.*- Finnish mathematician. He was the first who proved Theorem 10 which in literature is known as "Central Limiting Theorem".

[3] *Levy, Paul Pierre* (15.9.1889 - 15.12.1971)-French mathematician, the member of Paris Academy of Sciences (1964). He was the first who applied the method of characteristic functions to prove "Central Limiting Theorem".

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

Proof. Here we present the proof of this theorem in the case of absolutely continuous random variables. Assume that $m = M\xi_1$, $\sigma = \sqrt{D\xi_1}$. We have

$$\lim_{n\to\infty} \Phi_{\frac{\sum_{k=1}^{n}\xi_k - mn}{\sqrt{n}\sigma}}(t) = \lim_{n\to\infty} \Phi_{\sum_{k=1}^{n}(\frac{\xi_k - m}{\sqrt{\sigma}})\frac{1}{\sqrt{n}}}(t) =$$

$$\lim_{n\to\infty} \Phi_{\sum_{k=1}^{n}(\frac{\xi_k - m}{\sqrt{\sigma}})}(\frac{t}{\sqrt{n}}) = \lim_{n\to\infty} \prod_{i=1}^{n} \Phi_{\frac{\xi_i - m}{\sigma}}(\frac{t}{\sqrt{n}}) = \lim_{n\to\infty} e^{\ln \prod_{i=1}^{n} \Phi_{\frac{\xi_i - m}{\sigma}}(\frac{t}{\sqrt{n}})} =$$

$$= \lim_{n\to\infty} e^{\sum_{i=1}^{n} \ln \Phi_{\frac{\xi_i - m}{\sigma}}(\frac{t}{\sqrt{n}})}.$$

We set $\Phi(t) = \Phi_{\frac{\xi_i - m}{\sigma}}(t)$. If we denote with $f(t)$ the distribution function of random variable $\frac{\xi_i - m}{\sigma}$, then we get

$$\Phi(t) = \int_{-\infty}^{+\infty} e^{itx} f(x)dx,$$

$$\Phi'(t) = \int_{-\infty}^{+\infty} ixe^{itx} f(x)dx,$$

$$\Phi''(t) = \int_{-\infty}^{+\infty} i^2x^2e^{itx} f(x)dx = -\int_{-\infty}^{+\infty} x^2e^{itx} f(x)dx.$$

Note that

$$\Phi(0) = 1,$$

$$\Phi'(0) = iM(\frac{\xi_i - m}{\sigma}) = 0,$$

$$\Phi''(0) = -\int_{-\infty}^{+\infty} x^2 f(x)dx = -1.$$

The Maclaurin [4] formula with the first three members has the following form

$$\Phi(t) = \Phi(0) + \frac{\Phi'(0)}{1!}t + \frac{\Phi''(0)}{2!}t^2 + a(t)t^3,$$

where $\lim_{t\to 0} a(t) = 0$. Hence, we get

$$\Phi(\frac{t}{\sqrt{n}}) = 1 + 0 \cdot \frac{t}{\sqrt{n}} - \frac{t^2}{2n} + a(\frac{t}{\sqrt{n}}) \cdot \frac{t^3}{n\sqrt{n}}.$$

From the course of mathematical analysis it is well known that $\ln(1 + o(n)) \approx o(n)$ when $o(n)$ is an infinitely small sequence (i.e., $\lim_{n\to\infty} o(n) = 0$). Finally we get

$$\lim_{n\to\infty} \Phi_{\frac{\sum_{k=1}^{n}\xi_k - mn}{\sqrt{n}\sigma}}(t) = \lim_{n\to\infty} e^{n\ln(1+0\cdot\frac{t}{\sqrt{n}} - \frac{t^2}{2n} + a(\frac{t}{\sqrt{n}})\cdot\frac{t^3}{n\sqrt{n}})} =$$

[4] *Maclaurin, Colin* (1698 - 14.6.1746) - Scottish mathematician.

$$= \lim_{n\to\infty} e^{n(-\frac{t^2}{2n}+a(\frac{t}{\sqrt{n}})\cdot\frac{t^3}{n\sqrt{n}})} = e^{\lim_{n\to\infty}(-\frac{t^2}{2}+a(\frac{t}{\sqrt{n}})\cdot\frac{t^3}{\sqrt{n}})} = e^{-\frac{t^2}{2}},$$

which ends the proof of theorem.

Example 13.7 Assume that the following conditions are fulfilled:

1) Let ξ_k be the moon's diameter estimation obtained with k-th measure $(k \in N)$;

2) $a = M\xi_k$ $(k \in N)$ is the moon's diameter;

3) $D\xi_k = 1$ $(k \in N)$;

4) the results of measures $(\xi_k)_{k\in N}$ is a sequence of normally distributed independent random variables with parameters $(a, 1)$.

Using the Chebishev inequality (cf. Chapter 11, Example 11.1) we have proved that $n_C = 2000$ is such smallest natural number for which the following stochastic inequality

$$P(\{\omega : |\frac{1}{n_C}\sum_{k=1}^{n_C}\xi_k(\omega) - a| \le 0,1\}) \ge 0,95$$

holds. Note here that with the help of Chebishev inequality it is not possible to choose natural number smaller than $n_C = 2000$ which will satisfy the above mentioned inequality. Since $\frac{\sum_{k=1}^{n}\xi_k - na}{\sqrt{n}}$ is the normally distributed random variable with parameter $(0, 1)$, we can calculate the smallest natural number n_C, for which the same inequality holds. Indeed, we get :

$$P(\{\omega : |\frac{1}{n}\sum_{k=1}^{n}\xi_k(\omega) - a| \le 0,1\} = P(\{\omega : |\frac{\sum_{k=1}^{n}\xi_k(\omega) - na}{n}| \le 0,1\} =$$

$$= P(\{\omega : |\frac{\sum_{k=1}^{n}\xi_k(\omega) - na}{\sqrt{n}}| \le 0,1\sqrt{n}\} = 1 - 2\Phi(-0,1\sqrt{n}).$$

Clearly, we must to choose a such smallest natural number n_L which will be a solution of the following inequality

$$1 - 2\Phi(-0,1\sqrt{n}) \ge 0,95.$$

We have

$$\Phi(-0,1\sqrt{n}) \le \frac{1-0,95}{2} \Leftrightarrow \Phi(-0,1\sqrt{n}) \le 0,025 \Leftrightarrow$$
$$-0,1\sqrt{n} \le \Phi^{-1}(0,025) \Leftrightarrow \sqrt{n} \ge 100(\Phi^{-1}(0,025))^2 \Leftrightarrow$$
$$n \ge 100(1,96)^2 \Leftrightarrow n \ge 384,16 \Leftrightarrow n \ge 385.$$

Finally we deduce that $n_L = 385$. Consequently, $n_L = 385$ is a such smallest natural number which is solution of the following inequality

$$P(\{\omega : |\frac{1}{n_L}\sum_{k=1}^{n_L}\xi_k(\omega) - a| \le 0,1\} \ge 0,95.$$

It is clear that natural number $n_L = 385$ is smaller than natural number $n_c = 2000$ obtained with the help of the Chebishev inequality.

Remark 13.2. If the sequence of random variables $(\xi_k)_{k\in N}$ is weakly convergent to random variable ξ, then for sufficiently "large" natural number n the distribution function F_{ξ_n} of ξ_n can be assumed to be equal to distribution function F_ξ of random variable ξ.

Tests

13.1. Let define sequence of random variables $(\xi_n)_{n\in N}$ with

$$(\forall\omega)(\omega\in\Omega\to\xi_n(\omega)=C-\frac{1}{n}).$$

Then the sequence of random variables $(\xi_n)_{n\in N}$ is weakly convergent to random variable ξ, which is equal (with probability 1) to

a) $c-1$, b) c, c) c^2, d) $c+1$.

13.2. Let ξ_n be the Poisson random variable with parameter $\lambda+o(n)$ for $n\in N$, where $\lambda>0$, and let $(o(n))_{n\in N}$ be an infinitely small sequence. Then the sequence of random variables $(\xi_n)_{n\in N}$ is weakly convergent to the Poisson random variable with parameter μ, where μ is equal to

a) λ, b) λ^2, c) $\lambda(1+\lambda)$, d) $\lambda^2(1+\lambda)^2$.

13.3. Let ξ_n be a random variable uniformly distributed in interval (a_n,b_n) for $n\in N$. Assume also that $\lim_{n\to\infty}a_n=a$ and $\lim_{n\to\infty}b_n=b$. Then the sequence of random variables $(\xi_n)_{n\in N}$ is weakly convergent to the random variable uniformly distributed on interval (c,d), where (c,d) is equal to

a) (a,b), b) $(\frac{b-a}{2},\frac{a+b}{2})$, c) $(a,\frac{a+b}{2})$, d) $(\frac{a+b}{2},b)$.

13.4. Let $(\xi_n)_{n\in N}$ be an independent sequence of random variables and let ξ_k be a normally distributed random variable with parameters $(\frac{1}{2^k};\frac{1}{2^{2k}})$ for $k\in N$. Then the sequence of random variables $(\sum_{k=1}^n\xi_k)_{n\in N}$ is weakly convergent to the normally distributed random variable with parameters (m,σ^2), where (m,σ^2) is equal to

a) $(1,3)$, b) $(1,4)$, c) $(1,5)$, d) $(1,6)$.

13.5. (The Poisson Theorem). Let $(\xi_n)_{n\in N}$ be the sequence of independent Binomial random variables with parameters (n,p_n). Assume that $\lim_{n\to\infty}n\cdot p_n=\lambda>0$. Then the sequence of random variables $(\xi_n)_{n\in N}$ is weakly convergent to the Poisson random variable with parameter μ, where μ is equal to

a) $\lambda+1$, b) λ, c) $\lambda-1)$, d) λ^2.

13.6. Let $(\xi_n)_{n\in N}$ be the independent sequence of normally distributed random variables with parameter (a,σ^2). Then the sequence of random variables $(\frac{\xi_1+\cdots+\xi_n}{n})_{n\in N}$ is weakly convergent to constant random variable m, where m is equal to

a) a, b) a^2, c) a^3, d) a^4.

13.7. Let ξ_k be a normally distributed random variable with parameter (m_k,σ_k^2) for $1\le k\le n$. Then sum $\sum_{k=1}^n\xi_k$ is a normally distributed random variable with parameter (m,σ^2), where (m,σ^2) is equal to

a) $(\sum_{k=1}^{n} m_k, \sum_{k=1}^{n} \sigma_k)$, b) $(\sum_{k=1}^{n} m_k, \sum_{k=1}^{n} \sigma_k^2)$,
c) $(\sum_{k=1}^{n} m_k, \sum_{k=1}^{n} \sigma_k^3)$, d) $(\sum_{k=1}^{n} m_k, \sum_{k=1}^{n} \sigma_k^4)$.

13.8. If ξ is a normally distributed random variable with parameter (m, σ^2), then random variable $a\xi + b$ is distributed normally with parameter (c, d^2), where (c, d^2) is equal to

a) $(b + am, a^2\sigma^2)$, b) $(b + am, a\sigma^2)$,
c) $(b + am, a^2\sigma)$, d) $(b + m, a\sigma^2)$.

13.9. Let $(\xi)_{1 \le k \le n}$ be an independent sequence of random variables and let ξ_k be a Poisson random variable with parameter λ_k. Then sequence $\sum_{k=1}^{n} \xi_k$ is a Poisson random variable with parameter μ, where μ is equal to

a) $\sum_{k=1}^{n} \lambda_k^2$, b) $\sum_{k=1}^{n} \lambda_k$,
c) $\sum_{k=1}^{n}(1 - \lambda_k)$, d) $\sum_{k=1}^{n}(1 + \lambda_k)$.

13.10. Let $(\xi_k)_{1 \le k \le n}$ be an independent sequence of random variables and let ξ_k be a Binomial random variable with parameters (n_k, p). Then $\sum_{k=1}^{n} \xi_k$ is a binomial random variable with parameters (m, x), where (m, x) is equal to

a) $(\sum_{k=1}^{n} {}_k, p)$, b) $(\sum_{k=1}^{n} {}_k, p^2)$,
c) $(\sum_{k=1}^{n} {}_k^2, p^2)$, d) $(\sum_{k=1}^{n} {}_k, p^3)$.

13.11. Let ξ_k be a number of demands of the k-th goods during one day which is a Poisson random variable with parameter λ_k $(1 \le k \le n)$. Then the probability that the common number of demands of all goods during one day will be equal to 8 relative to hypothesis $m = 10$, $\lambda_1 = \cdots = \lambda_5 = 0,3$, $\lambda_6 = \cdots = \lambda_9 = 0,8$, $\lambda_{10} = 1,3$, is equal to

a) $0,345103$, b) $0,457778$, c) $0,567788$, d) $0,103258$.

13.12. The mean load transported with a lórry on each trip is equal to $m = 20$. The mean absolute deviation of the above mentioned load is equal to $\sigma = 1$. Then

1) the probability that the weight of the load transmitted during 100 trips will be in interval $[1950; 2000]$, is equal to

a) $0,5$, b) $0,55$, c) $0,555$, d) $0,5555$;

2) the value which is grater with probability 0,95 than the weight of the load transmitted during 100 trips is equal to

a) 20164, b) 20264, c) 20364, d) 20464.

13.13. A mean weight of an apple is $m = 0,2$ kg. A mean of absolute deviation of the weight of accidentally chosen apple is $\sigma = 0,02$ kg. Then

1) the probability that the weight of the accidently chosen 49 apples will be in interval $[9,5; 10]$, is equal to

a) $0,44$; b) $0,88$; c) $0,178$; d) $0,356$;

2) the value which will be smaller then the weight of the accidentally chosen 100 apples with probability 0,95, is equal to

a) $16,672$, b) $17,672$, c) $18,672$, d) $19,672$.

13.14. The probability that túrner will make a standard detail is equal to 0,64. Then the probability that

1) 70 details, accidentally chosen from the complect of 100 details will be standard, is equal to

a) $0,6241$, b) $0,7241$, c) $0,8241$, d) $0,9241$;

2) the number of standard details in the accidentally chosen 100 details will be in interval [50,65], is equal to

a) $0,1108$, b) $0,1308$, c) $0,1508$, d) $0,1708$.

13.15. The factory sent 15000 standard details to the stórehouse. The probability that the detail will dámaged during transportation, is equal to 0,0002. Then the probability that

1) 3 dámaged details will be brought to bring at stórehose, is equal to

a) $0,094042$, b) $0,114042$, c) $0,134042$, d) $0,154042$;

2) the number of dámaged details will be in interval [2,4], is equal to

a) $0,414114$, b) $0,515115$, c) $0,616116$, d) $0,717117$.

Chapter 14

Markov Chains

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Assume that we have a physical system, which after each step changes its phase position. Assume that the number of possible positions $\epsilon_1, \epsilon_2, \cdots$ is finite or countable. Let $\xi_n(\omega)$ be a position of physical system after n steps $(n \in N,\ \omega \in \Omega)$. Clearly, the chain of logical transitions

$$\xi_0(\omega) \to \xi_1(\omega) \to \cdots \quad (\omega \in \Omega)$$

depends on the chance factor. Assume that the following regularity condition is preserved: if after n steps the system is in position ϵ_i , then, independently of its early positions it will pass to position ϵ_j with probability P_{ij}, i.e.,

$$P_{ij} = P(\{\omega : \xi_{n+1}(\omega) = \epsilon_j \ |\xi_n(\omega) = \epsilon_i\}) \ (i, j = 1, 2, \cdots).$$

The above described model is called Markov [1] homogeneous chain. Number P_{ij} is called the transition probability. Besides there is also given also the distributions of initial positions, i.e.,

$$P_i^{(0)} = P(\{\omega : \xi_0(\omega) = \epsilon_i\}) \ \ i = 1, 2, \cdots .$$

Here naturally arises the following problem: what is the probability that the physic system will be in the position ϵ_i after n steps ? Let denote this probability by $P_j(n)$, i. e.,

$$P_j(n) = P(\{\omega : \xi_n(\omega) = \epsilon_j\}).$$

Note that after $n-1$ steps the physical system will be in one of the possible positions ϵ_k $(k = 1, 2, \cdots)$. The probability that the physical system will be in position ϵ_k is equal to $P_k(n-1)$. The probability that the physical system will occur in position ϵ_j after n steps if it is known that after $n-1$ steps it was in position ϵ_k is equal to transition probability P_{kj}. Using total probability formula we get

$$P(\{\omega : \xi_n(\omega) = \epsilon_j\}) =$$

[1]Markov, A (2(14).1856-20.7.1922) - Russian mathematician, the member of Petersburg Academy of Sciences (1890).

$$\sum_{k\in N} P(\{\omega : \xi_n(\omega) = \epsilon_j\}|\{\omega : \xi_{n-1}(\omega) = \epsilon_k\}) \cdot P(\{\omega : \xi_{n-1}(\omega) = \epsilon_k\}).$$

The formula gives the following recurrent formula for calculation of the probability $P_j(n)$:

$$P_j(0) = P_j^{(0)}, \quad P_j(n) = \sum_{k\in N} P_k(n-1) \cdot P_{kj} \quad (j, n = 1, 2, \cdots).$$

In this case when the physical system at the initial moment is in position ϵ_i , the initial distribution has the following form

$$P_i^{(0)} = 1, \quad P_k^{(0)} = 0, \quad k \neq i$$

and probability $P_j(n)$ coincides with $P_{ij}(n)$, which is equal to transition probability from position ϵ_i to position ϵ_j after n steps, i. e.,

$$P_{ij}(n) = P(\{\omega : \xi_n(\omega) = \epsilon_j|\{\omega : \xi_0(\omega) = \epsilon_i\}\} \quad i, j = 1, 2, \cdots .$$

In the case of the following initial distribution $P_i^{(0)} = 1, \quad P_k^{(0)} = 0 \ (k \neq i)$ we get

$$P_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} ,$$

$$P_{ij}(n) = \sum_{k\in N} P_{ik}(n-1) \cdot P_{kj} \quad (n = 1, 2, \cdots).$$

Setting

$$\mathcal{P}(n) = (P_{ij}(n))_{i,j\in N} ,$$

we get

$$\mathcal{P}(0) = I, \ \mathcal{P}(1) = \mathcal{P}, \ \mathcal{P}(2) = \mathcal{P}(1) \cdot \mathcal{P} = \mathcal{P}^2, \cdots ,$$

where I is an infinite-dimensional unite matrix and $\mathcal{P}$ is the matrix of transition probabilities. It is evedent that

$$\mathcal{P}(n) = \mathcal{P}^n \quad (n = 1, 2, \cdots).$$

Let consider some examples.

Example 14.1(Random roaming). Let consider random roaming connected with infinite number of Bernoulli independent experiments when the particle "is roaming" in the integer-valued points of the real axis such that if it is placed in the i-th position, then the transition probabilities to positions $i+1$ or $i-1$ are equal to p or $q = 1-p$, respectively $(0 < p < 1)$. If with ξ_n we denote the position of the particle after n steps, then sequence

$$\xi_0(\omega) \to \xi_1(\omega) \to \cdots \quad (\omega \in \Omega)$$

will be the Markov chain, whose transition probabilities have the following form

$$P_{ij} = \begin{cases} p\ , & \text{if} \quad j = i+1 \\ q\ , & \text{if} \ \ j = i-1 \end{cases} .$$

Remark 14.1 In our case the physical system (i.e., the particle) has an infinite number of phase positions.

Example 14.2 Let consider a physical system which has three different possible positions $\epsilon_1, \epsilon_2, \epsilon_3$. Assume that after one step matrix $\mathcal{P}$ of transition probabilities has the following form

$$\mathcal{P} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1. \end{pmatrix}.$$

In the present example position ϵ_3 has the property that if physical system will be placed in it, then it remains in this position with probability 1. Such position is called "absorbable". If the particle is placed in some position and it remains in it with probability 0 then such position is called "reflectable". If position ϵ_i is "absorbable" then $P_{ii} = 1$ and if the position ϵ_i is "reflectable", then $P_{ii} = 0$.

If we know that before observation the physical system is placed in position ϵ_i $(1 \le i \le n)$, then using matrix $\mathcal{P}(m)$ we can find transition probability $P_{ij}(m)$ after m steps. In this case, when an initial position of physical system is not known, but we know probabilities $P_i^{(0)}$ that system is placed in position ϵ_i, then using total probability formula we can calculate the probability that after m steps the physical system will be placed in position ϵ_j by the following formula

$$P_j(m) = \sum_{k=1}^{n} P_k^{(0)} \cdot P_{kj}(m).$$

The column-vector

$$P^{(0)} = (P_1^{(0)}, P_2^{(0)}, \cdots, P_n^{(0)})$$

is called the vector of initial distribution of the Markov chain and the column-vector

$$\mathcal{P}^{(m)} = (P_1^{(m)}, P_2^{(m)}, \cdots, P_n^{(m)})$$

is called the distribution vector after m steps for Markov chain. In our notations we get

$$\mathcal{P}^{(m)} = \mathcal{P}^{(0)} \cdot \mathcal{P}(m)$$

We present Markov theorem about limit probabilities without proof.

Theorem 14.1 *Let* $(\epsilon_i)_{1 \le i \le n}$ *be the possible positions of a physical system. If the crossing probabilities of the Markov chain* $P_{ij}^{(m)}$ *of matrix* $\mathcal{P}^{(m)}$ *are positive for arbitrary natural number* m*, then there exists a finite family of real numbers* $(q_i)_{1 \le i \le n}$ *such that*

$$(\forall i)(1 \le i \le n \rightarrow \lim_{m \to \infty} P_{ij}(m) = q_j) \quad (1 \le j \le n).$$

The number q_j $(1 \le j \le n)$ can be considered as the probability of the occurrence in the j-th position of physical system for sufficiently large natural number m.

Tests

14.1. The matrix of the transition probabilities of the Markov chain is defined by

$$\mathcal{P} = \begin{pmatrix} 0 & 1 & 0 \\ 0,5 & 0,5 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and the vector of initial probabilities coincides with $(0,2;\ 0,5;\ 0,3)$. Then the distribution vector after two steps will be equal to

a) $(0,125; 0,475; 0,4)$, b) $(0,225; 0,475; 0,3)$,
c) $(0,025; 0,575; 0,4)$, d) $(0,125; 0,375; 0,5)$.

14.2. The matrix of transition probabilities of the Markov chain is defined by

$$\mathcal{P} = \begin{pmatrix} 0,3 & 0,1 & 0,6 \\ 0 & 0,4 & 0,6 \\ 0,4 & 0,3 & 0,3 \end{pmatrix}.$$

Then matrix $\mathcal{P}(2)$ of transition probabilities of the Markov chain after 2-steps has the following form

a)

$$\begin{pmatrix} 0,25 & 0,15 & 0,6 \\ 0 & 0,3 & 0,7 \\ 0,4 & 0,3 & 0,3 \end{pmatrix},$$

b)

$$\begin{pmatrix} 0,33 & 0,21 & 0,46 \\ 0,4 & 0,3 & 0,3 \\ 0,24 & 0,13 & 0,33 \end{pmatrix}.$$

14.3. The matrix of crossing probabilities of the Markov chain is defined by

$$\mathcal{P} = \begin{pmatrix} 0,1 & 0,5 & 0,4 \\ 0 & 0 & 1 \\ 0,5 & 0,3 & 0,2 \end{pmatrix}.$$

Then transition probability from position ϵ_2 to position ϵ_3 after 3 steps $P_{23}^{(3)}$ will be equal to

a) $0,125$, b) $0,225$, c) $0,54$, d) $0,375$.

14.4. The matrix of transition probabilities of the Markov chain is defined by

$$\mathcal{P} = \begin{pmatrix} 0,3 & 0,7 \\ 0,1 & 0,9 \end{pmatrix}$$

and the vector of initial probabilities coincides with $(0,2;\ 0,8)$. It is known that transition probability from any initial position ϵ_i to position ϵ_i after 2 steps is equal to 0,128. Then ϵ_i is equal to

a) 1, b) 2.

Chapter 15

The Process of Brownian Motion

Let consider a little particle which is placed in homogeneous liquid. Since the particle undergoes chaotic collisions with molecules of liquid, it is in continuous unordered motion. A discrete analogue of this process is the following random roaming of the particle on the real axis: the particle changes its positions in such moments of times which are múltiple of Δt $(\Delta t > 0)$. If the particle is placed in point x then the transition probabilities to positions $x(t) + \Delta x$ and $x(t) - \Delta x$ are the same and are equal to $0,5$ (Here we consider one-dimensional random roaming). We assume that Δx is the same for arbitrary position x. In the limit, when $\Delta t \to 0$, $\Delta x \to 0$ with spatial law, a continuous random roaming is obtained which describes a model of the Brownian[1] physic process.

Let denote with $\xi_t(\omega)$ the position of the particle in moment t. Assume that the particle is placed in position $x = 0$ in initial moment $t = 0$. In this case of discrete roaming during time t this particle makes $n = \frac{t}{\Delta t}$ steps. If we denote with $S_n(\omega)$ the number of steps with Δx in positive direction, then the common shift in positive direction will be to equal to $S_n(\omega) \cdot \Delta x$ and the common shift in negative direction will be equal to $(n - S_n(\omega)) \cdot \Delta x$. Hence, common shift $\xi_t(\omega)$ after time $t = n\Delta t$ is connected with $S_n(\omega)$ with the following equality

$$\xi_t(\omega) = [S_n(\omega)\Delta x - (n - S_n(\omega))\Delta x] = (2S_n(\omega) - n)\Delta x.$$

If we assume that $\xi_0(\omega) = 0$, then

$$\xi_t(\omega) = (\xi_s(\omega) - \xi_0(\omega)) + (\xi_t(\omega) - \xi_s(\omega))$$

for every $s \in [0; t]$. Clearly, in our model random variables $\xi_s - \xi_0$ and $\xi_t - \xi_s$ are independent. As distribution functions of íncreases $\xi_t - \xi_s$ and $\xi_{t-s} - \xi_0$ are equal, $\sigma^2(t) = D\xi_t$ satisfies the following condition

$$\sigma^2(t) = \sigma^2(s) + \sigma^2(t - s) \quad (0 \le s \le t).$$

[1] *Brown, Robert* (21.12.1773, - 10.6.1858) - English botanist who was the first to discover so called "Brownian motion", which in the probability theory is also known as "Wiener process".

It follows that $\sigma^2(t)$ linearly depends on t. It means that there exists positive real number σ^2, such that

$$D\xi_t = \sigma^2 \cdot t.$$

Number σ^2 is called a diffusion coefficient of the Brownian process. On the other hand, it is easy to show that mathematical variance of shift after time t (or after $n = \frac{t}{\Delta t}$ steps) is $D\xi_t = (\Delta x)^2 \cdot \frac{t}{\Delta t}$. Finally, we get following relation between values Δx and Δt

$$\frac{(\Delta x)^2}{\Delta t} = \sigma^2.$$

Since particle transitions are independent, they can be considered as the Bernoulli experiment with "success" probability $p = \frac{1}{2}$. Then the number of steps in positive direction $S_n(\omega)$ will be equal to the number of "succésses" in n independent Bernoulli experiments. In this case the position of particle $\xi_t(\omega)$ at moment t will be connected with normed random variable $S_n^*(\omega) = \frac{1}{\sqrt{n}}(2S_n(\omega) - n)$ with the following equality

$$\xi_t(\omega) = S_n^*(\omega)\sqrt{n}\Delta x = S_n^*(\omega)\sqrt{t}\frac{\Delta x}{\sqrt{\Delta t}} = S_n^*(\omega)\sigma\sqrt{t}.$$

Using Theorem 10 of Chapter 13, we deduce that the distribution function of random variable $\xi_t(\omega)$ in the case of one-dimensional Brownian process has the following form

$$P(\{\omega : x_1 \leq \frac{\xi_t(\omega)}{\sigma\sqrt{t}} \leq x_2\}) = \lim_{\Delta t \to 0} P(\{\omega : x_1 \leq S_n^*(\omega) \leq x_2\}) =$$

$$\lim_{n\to\infty} P(\{\omega : x_1 \leq \frac{S_n(\omega) - np}{\sqrt{npq}} \leq x_2\}) = \frac{1}{\sqrt{2\pi}}\int_{x_1}^{x_2} e^{-\frac{x^2}{2}}dx,$$

where $p = q = \frac{1}{2}$.

One can easily demonstrate the validity of the following formula

$$P(\{\omega : y_1 \leq \xi_t(\omega) \leq y_2\}) = \Phi(\frac{y_2}{\sigma\sqrt{t}}) - \Phi(\frac{y_1}{\sigma\sqrt{t}})\ (t > 0, y_1 < y_2).$$

Now we consider a problem of prognosis of the Brownian motion.

Let $(\xi_t(\omega))_{t>0}(\omega \in \Omega)$ be a Brownian process with unknown diffusion coefficient σ^2. Let $(\xi_{t_k}(\omega_0))_{1\leq k\leq n+1}$ be the result of observations on this process at moments $(t_k)_{1\leq k\leq n+1}$. Here we assume that $t_1 = 0$, $\xi_{t_1}(\omega_0) = 0$ and $t_k < t_{k+1}$. We set

$$X_k(\omega) = \frac{\xi_{t_{k+1}}(\omega) - \xi_{t_k}(\omega)}{\sqrt{t_{k+1} - t_k}}\ (1 \leq k \leq n).$$

It is clear that $(X_k(\omega))_{1\leq k\leq n}$ $(\omega \in \Omega)$ is a sequence of independent random variables normally distributed with parameters $(0, \sigma^2)$, where σ^2 is an unknown parameter. From the course of mathematical statistics it is known that statistics σ_n^2 defined with

$$\sigma_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i(\omega) - \frac{1}{n}\sum_{j=1}^{n}X_j)^2,$$

is a "good" estimation of unknown parameter σ^2.

The prognosis of the stochastic behavior of the Brownian motion in moment $t(t > t_{n+1})$ can be given with the following formula

$$P(\{\omega : y_1 \leq \xi_t(\omega) \leq y_2\}) \approx \Phi(\frac{y_2}{\sigma_n \sqrt{t}}) - \Phi(\frac{y_1}{\sigma_n \sqrt{t}}) \ (t > 0, y_1 < y_2).$$

Remark 15.1. Using statistical functions **NORMDIST** and **VAR**(cf. p.134) the prognosis of the stochastic behavior of the Brownian motion in the moment $t(t > t_{n+1})$ can be given with the following formula

$$P(\{\omega : y_1 \leq \xi_t(\omega) \leq y_2\}) \approx$$

$$NORMDIST(\frac{y_2}{\sqrt{t \times VAR(x_1 : x_n)}};0;1;1)-$$

$$NORMDIST(\frac{y_1}{\sqrt{t \times VAR(x_1 : x_n)}};0;1;1),$$

where $x_k = X_k(\omega_0)$ for $1 \leq k \leq n$.

Remark 15.2. It is reasonable to note that the hypothesis about the form of the distribution function of one-dimensional Brownian motion belongs to eminent physician Albert Einstein. His conjecture was strongly proved by American mathematician Norbert Wiener to whom belongs the mathematical construction of the Brownian motion. Hence, in literature the Brownian process is mentioned also as Wiener process.

Tests

15.1. The change of the commodity price is the Brownian process with diffusion coefficient $\sigma^2 = 1$. At $t = 0$ the price of the commodity was equal to 9 lari. The probability that the price of the commodity will not increase at moment $t = 9$, is equal to

a) $0,4$, b) $0,5$, c) $0,6$, d) $0,7$.

15.2. The change of the commodity price is the Brownian process with diffusion coefficient $\sigma^2 = 1$. At $t = 0$ the price of the commodity was equal to 200 lari. The probability that the price of the commodity at the moment $t = 9$ will be

1) less than 190 lari, is equal to

a) $0,3064$, b) $0,3164$, c) $0,3264$, d) $0,3364$;

2) more than 210 lari, is equal to

a) $0,2864$, b) $0,3264$, c) $0,3464$, d) $0,3664$;

3) placed in interval [185 , 205], is equal to

a) $0,3027$, b) $0,3227$, c) $0,3527$, d) $0,3727$.

15.3. The change of a bond's price is the Brownian process with diffusion coefficient $\sigma^2 = 1$. The firm bought the bond for 3000 lari at the moment $t = 0$. The probability that

1) the prófit obtained by buying the bond at moment $t = 250000$ will be more than 300 lari, is equal to

a) 0, b) 0,1, c) 0,2, d) 0,3;

2) the damage obtained by buying of bond at the moment $t = 900$ will be grater than 15 lari, is equal to

a) 0, b) 1, c) 0,3, d) 0,6.

15.4. The change of the goods' price in the shop is the Brownian process with diffusion coefficient $\sigma^2 = 1$. At $t = 0$ the price of the goods was equal to 50 lari. The buyer is interested to buy the goods for no more than 55 lari. The shop stops selling the goods if its price decreases below 41 lari. The probability that the buyer bought the goods in moment $t = 1\frac{2}{3}$, is equal to

a) 0,2287, b) 0,3387, c) 0,4487, d) 0,5587.

Table 1. $\Phi(x)$ and $\phi(x)$

x	$\phi(x)$	$\Phi(x)$	x	$\phi(x)$	$\Phi(x)$	x	$\phi(x)$	$\Phi(x)$
0,00	0,3989	0,5000	0,34	0,3765	0,6331	0,68	0,3166	0,7517
01	3989	5040	35	3752	6368	69	3144	7549
02	3988	5080	36	3739	6406	70	3123	7580
03	3988	5120	37	3725	6443	71	3101	7611
04	3986	5160	38	3712	6480	72	3079	7642
05	3984	5199	39	3697	6517	73	3056	7673
06	3982	5239	40	3683	6557	74	3034	7703
07	3980	5279	41	3668	6591	75	3011	7734
08	3977	5319	42	3653	6628	76	2989	7764
09	3973	5359	43	3637	6664	77	2966	7794
10	3970	5398	44	3621	6700	78	2943	7823
11	3965	5438	45	3605	6736	79	2920	7852
12	3961	5478	46	3589	6772	80	2897	7881
13	3956	5517	47	3572	6808	81	2874	7910
14	3951	5557	48	3555	6844	82	2850	7939
15	3945	5596	49	3538	6879	83	2827	7967
16	3939	5636	50	3521	6915	84	2803	7995
17	3932	5675	51	3503	6950	85	2780	8023
18	3925	5714	52	3484	6985	86	2756	8051
19	3918	5753	53	3467	7016	87	2732	8078
20	3910	5793	54	3448	7054	88	2709	8106
21	3902	5832	55	3429	7088	89	2685	8133
22	3894	5871	56	3410	7123	90	2661	8159
23	3885	5910	57	3391	7157	91	2637	8186
24	3876	5948	58	3372	7190	92	2613	8212
25	3867	5987	59	3352	7224	93	2589	8238
26	3357	6026	60	3332	7257	94	2565	8264
27	3847	6064	61	3312	7291	95	2541	8289
28	3836	6103	62	3292	7324	96	2510	8315
29	3825	6141	63	3271	7357	97	2492	8340
30	3814	6179	64	3251	7389	98	2468	8365
31	3802	6217	65	3230	7422	99	2444	8389
32	3790	6265	66	3207	7454	1,00	2420	8413
33	3778	6293	67	3187	7486	1,01	2396	8438

x	$\phi(x)$	$\Phi(x)$	x	$\phi(x)$	$\Phi(x)$	x	$\phi(x)$	$\Phi(x)$
1,02	0,2371	0,8461	1,42	0,1456	0,9222	1,82	0,0761	0,9656
03	2347	8485	43	1435	9236	83	0748	9664
04	2323	8508	44	1415	9251	84	0734	9671
05	2299	8531	45	1394	9265	85	0721	9678
06	2275	8554	46	1374	9279	86	0707	9686
07	2251	8577	47	1354	9292	87	0694	9693
08	2227	8599	48	1334	9306	88	0681	9699
09	2203	8621	49	1315	9319	89	0669	9706
10	2179	8648	50	1295	9332	90	0656	9713
11	2155	8665	51	1276	9345	91	0644	9719
12	2131	8686	52	1257	9357	92	0632	9729
13	2107	8708	53	1238	9370	93	0620	9732
14	2083	8729	54	1219	9382	94	0608	9738
15	2059	8749	55	1200	9394	95	0596	9744
16	2036	8770	56	1182	9406	96	0584	9750
17	2012	9790	57	1163	9418	97	0573	9756
18	1989	8810	58	1145	9429	98	0562	9761
19	1965	8820	59	1127	9441	99	0551	9767
20	1942	8849	60	1109	9452	2,00	0540	9772
21	1919	8869	61	1092	9463	02	0519	9783
22	1895	8888	62	1074	9474	04	0498	9793
23	1872	8907	63	1057	9484	06	0478	9803
24	1849	8925	64	1040	9495	08	0459	9812
25	1826	8944	65	1023	9505	10	0440	9821
26	1804	8962	66	1006	9515	12	0422	9830
27	1881	8980	67	0989	9525	14	0404	9838
28	1858	8997	68	0973	9535	16	0387	9846
29	1836	9015	69	0957	9545	18	0371	9854
30	1714	9032	70	0940	9554	20	0355	9861
31	1691	9049	71	0925	9564	22	0339	9868
32	1669	9066	72	0909	9573	24	0325	9868
33	1647	9082	73	0893	9583	26	0310	9881
34	1626	9099	74	0878	9591	28	0297	9887
35	1604	9115	75	0863	9599	30	0283	9893
36	1582	9131	76	0848	9608	32	0270	9898
37	1561	9147	77	0833	9616	34	0258	9904
38	1539	9162	78	0818	9625	36	0246	9909
39	1518	9177	79	0804	9633	38	0235	9913
40	1457	9192	80	0790	9641	40	0224	9918
41	1476	9207	81	0775	9649	42	0213	9922

x	$\phi(x)$	$\Phi(x)$	x	$\phi(x)$	$\Phi(x)$	x	$\phi(x)$	$\Phi(x)$
2,44	0,0203	0,9927	2,72	0,0099	0,9967	3,00	0,0043	0,99655
46	0194	9931	74	0093	9969	10	0110	99903
48	0184	9934	76	0088	9971	20	0104	99931
50	0175	9938	78	0084	9973	30	0099	99951
52	0167	9941	80	0079	9974	40	0093	99966
54	0158	9945	82	0075	9976	50	0088	99976
56	0151	9948	84	0071	9977	60	0084	99984
58	0143	9951	86	0067	9979	70	00042	99989
60	0136	9953	88	0063	9980	80	00029	99993
62	0129	9956	90	0060	9981	90	00020	99995
64	0122	9959	92	0056	9982	4,00	00013	99996
66	0116	9961	94	0053	9984	4,50	00001	99999
68	0110	9963	96	0050	9985	5,00	00000	99999
70	0104	9965	98	0047	9986			

Table 1 contains the values of density function ϕ and of distribution function Φ of the standard normally distributed random variable in interval [0.5]. To calculate the values of ϕ and Φ in other points of the real axis we can use the following formulas:

$$\phi(x) = \begin{cases} 0, & \text{if } x > 5; \\ \phi(x), & \text{if } x \in [0;5] \text{ (we find } \phi(x) \text{ from Table 1);} \\ \phi(-x), & \text{if } x \in [-5;0[\text{ (we find } \phi(-x) \text{ from Table 1);} \\ 0, & \text{if } x < -5. \end{cases}$$

$$\Phi(x) = \begin{cases} 1, & \text{if } x > 5; \\ \Phi(x), & \text{if } x \in [0;5] \text{ (we find } \Phi(x) \text{ from Table 1);} \\ 1-\Phi(-x), & \text{if } x \in [-5;0[\text{ (we find } \Phi(-x) \text{ from Table 1);} \\ 0, & \text{if } x < -5. \end{cases}$$

The value of function Φ^{-1} is defined by

$$\Phi^{-1}(a) = \begin{cases} \Phi^{-1}(a), & \text{if } a \in [0,5;1] \text{ (we find } \Phi^{-1}(a) \text{ from Table 1);} \\ -\Phi^{-1}(1-a), & \text{if } a \in]0;0,5[\text{ (we find } \Phi^{-1}(1-a) \text{ from Table 1).} \end{cases}$$

Table 2. Poisson Distribution

$k \mid \lambda$	0,1	0,2	0,3	0,4	0,5	0,6
0	0,904837	0,818731	0,740818	0,670320	0,606531	0,548812
1	090484	163746	222245	263120	303265	329287
2	004524	016375	033337	053626	075816	098786
3	000151	0011091	003334	007150	012636	019757
4	000004	000055	000250	000715	001580	002964
5		000002	000015	000057	000158	000356
6			000001	000004	000013	000035
7					000001	000003
$k \mid \lambda$	0,7	0,8	0,9	1,0	2,0	3,0
0	0,496585	0,449329	0,406570	0,367879	0,135335	0,049787
1	347610	359463	365913	367879	270671	149361
2	121663	143785	164661	183940	270671	224043
3	028388	038343	049398	061313	180447	224042
4	004968	007669	011115	015328	090224	168031
5	000695	001227	002001	003066	036089	100819
6	000081	000165	000300	000511	012030	050409
7	000008	000019	000039	000073	003437	021604
8		000003	000004	000009	000859	008101
9				000001	000191	002701
10					000038	000810
11					000007	000221
12					000001	000055
$k \mid \lambda$	4,0	5,0	6,0	7,0	8,0	9,0
0	0,018316	0,006738	0,002479	0,000912	0,000335	0,000123
1	073263	033690	014873	006383	002684	001111
2	146525	084224	044618	022341	010735	004993
3	195367	140374	089235	052129	028626	014994
4	195367	175467	133853	091226	057252	033737
5	156293	175467	160623	027717	091604	060727
6	104194	146223	160623	149003	122138	091090
7	059540	104445	137677	149003	139587	117116
8	029770	065278	103258	130377	139587	131756

125

$k\|\lambda$	4,0	5,0	6,0	7,0	8,0	9,0
9	013231	036266	068898	101405	124077	131756
10	005292	018133	041303	070933	099262	118085
11	001925	008242	022529	045171	072190	097020
12	000642	003434	011262	026350	048127	072765
13	000197	001321	005199	014188	029616	050376
14	000056	000472	002228	007094	016924	032384
15	000015	000157	000891	003111	009026	019431
16	000004	000049	000334	001448	004513	010930
17	000001	000014	000118	000596	002124	005786
18		000004	002899	000232	000944	000944
19		000001	000012	000085	000397	001370
20			000004	000030	000159	000617
21			000001	000010	000061	000264
22				000003	000022	000108
23				000001	000008	000042
24					000003	000016
25					000001	000006
26						000002
27						000001

Test answers

N	a	b	c	d	N	a	b	g	d	N	a	b	g	d
1.1.1)			+		3.8.4)	+				5.3.				+
2)	+				5)	+				6.1.1)	+			
3)	+				6)		+			2)		+		
4)	+				3.9.		+			3)	+			
5)			+		3.10.			+		4)	+			
6)		+			3.11.1)	+				6.2.1)	+			
1.2.1)		+			2)			+		2)	+			
2)		+			3.12.				+	3)	+			
3)			+		3.14.			+		6.3.1)		+		
4)		+			3.15.		+			2)	+			
1.3.1)			+		3.16.		+			3)		+		
2)				+	3.17.		+			7.1.1)				+
1.4.1)			+		3.18	+				2)	+			
2)			+		3.19.	+				3)	+			
1.5.1)		+			3.20.			+		4)	+			
2)				+	3.21.		+			5)			+	
2.1.			+		3.22.	+				7.2.		+		
2.2.		+			3.23.	+				7.3.	+			
2.3.				+	3.24.			+		7.4.	+			
2.4.	+				4.1.		+			7.5.	+			
2.5.		+			4.2.		+			7.6.	+			
2.6.	+				4.3.			+		7.7.	+			
2.7.		+			4.4.	+				7.8.1)	+			
3.1.		+			4.5.1)		+			2)	+			
3.2.	+				2)	+				7.9.1)	+			
3.3.	+				4.6.1)	+				2)	+			
3.4.			+		2)	+				3)	+			
3.5.1)		+			4.7.		+			7.10.1)	+			
2)			+		4.8.1)	+			+	2)	+			
3)			+		2)		+			8.1.1)		+		
4)		+			5.1.1)		+			2)			+	
3.6.	+				2)			+		8.2.1)		+		
3.7		+			3)	+				2)			+	
3.8.1)		+			4)				+	8.3.1)				+
2)			+		5.2.1)	+				2)		+		
3)			+		2)		+			8.4.1)	+			

Some Statistical Functions of "Excel"

1. page 45 – PROB($x_1 : x_n; p_1 : p_n; y_1; y_2$).
2. page 46 – POISSON($k; \lambda; 0$).
3. page 46 – POISSON($k; \lambda; 1$).
4. page 48 – HYPERGEOMDIST($k; n; a; A$).
5. page 48 – BINOMDIST($k; n; p; 0$).
6. page 48 – BINOMDIST($k; n; p; 1$).
7. page 50 – NORMDIST($x; m; \sigma; 0$).
8. page 50 – NORMDIST($x; m; \sigma; 1$).
9. page 51 – EXPONDIST($x; \lambda; 0$).
10. page 51 – EXPONDIST($x; \lambda; 1$).
11. page 57 – SUMPRODUCT.
12. page 71 – AVERAGE($x_1 : x_n$).
13. page 71 – VARP($x_1 : x_n$).
14. page 71 – VAR($x_1 : x_n$).
15. page 73 – CORREL($x_1 : x_n; y_1 : y_n$).
16. page 73 – COVAR($x_1 : x_n; y_1 : y_n$).
17. page 77 – KURT($x_1 : x_n$).
18. page 77 – SKEW($x_1 : x_n$).
19. page 78 – MEDIAN($x_1 : x_n$).
20. page 78 – MODE($x_1 : x_n$).
21. page 86 – CHIDIST(x, n).
22. page 87 – TDIST($x, n, 1$).
23. page 87 – TDIST($x, n, 2$).
24. page 87 – FDIST(x, k_1, k_2).

Contents

Introduction ix

1 Set-Theoretical Operations. Kolmogorov Axioms 3

2 Properties of Probabilities 7

3 Examples of Probability Spaces 11

4 Total Probability and Bayes' Formulas 19

5 Applications of Caratheodory Method 29

6 Random Variables 39

7 Function of random variable distribution 43

8 Mathematical expectation and mathematical variance 57

9 Correlation Coefficient 73

10 Random Vector Distribution Function 81

11 Chebishev inequalities. The law of three σ 91

12 Limiting theorems 95

13 The Method of Characteristic Functions 101

14 Markov Chains 111

15 The Process of Brownian Motion 117

References 132