МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Нижегородский государственный
университет им. Н.И. Лобачевского

А.Б. Колпаков

А.С. Рукомина

Краткий курс лекций по дисциплине «*Теория вероятностей и Математическая статистика*» Часть 2. Математическаястатистика. (*Short course of lectures on the discipline "Probability theory and Mathematical statistics"* Part 2. Mathematical statistics)

Учебно-методическоепособие

Рекомендовано методической комиссией института экономики и предпринимательства для студентов ННГУ, обучающихся по направлениям подготовки 38.03.01 «Экономика» и 38.03.02 «Менеджмент».

Нижний Новгород
2021

УДК 519.2(075.8)
ББК 22.171я73
К61
Колпаков А.Б., Рукомина А.С., Краткий курс лекций по дисциплине «Теория вероятностей и математическая статистика» Часть 2. Математическаястатистика.
(A.B. Kolpakov and A.S. Rukomina Short course of lectures on the discipline “Probability theory and Mathematical statistics” Part 2. Mathematical statistics): Учебно-методическое пособие. – Нижний Новгород: Нижегородский госуниверситет, 2021. — 73с.

Рецензент:

к.ф.-м.н., доцент В.И. Перова

В настоящем пособии представлено краткое изложение материла по математической статистике в соответствии с рабочими программами по дисциплине *«Теория вероятностей и математическая статистика»* разработанными для таких направлений подготовки специалистов как Экономика (38.03.01) и Менеджмент (38.03.02). Материал также может быть использован и при подготовке по направлению Управление персоналом (38.03.03) в процессе изучения дисциплины *«Математика».*

Пособие рассчитано как на русских студентов, так и на студентов-иностранцев, которые, как показывает опыт, далеко не всегда достаточно хорошо владеют русским языком для того чтобы без каких-либо затруднений понять смысл различных математических терминов и фраз. В связи с этим, каждый раздел предлагаемого курса дублирован на английском языке.

Ответственный за выпуск:
Председатель методической комиссии ИЭП ННГУ,

к.э.н., доцент Макарова С.Д.

УДК 519.2(075.8)
ББК 22.171я73
К61

Содержание (Contents)

Введение

К настоящему времени, опубликовано довольно большое количество учебных и различных методических пособий, посвященных дисциплине *Теория вероятностей и математическая статистика*. Однако, при работе над ними авторы, как правило, брали за основу конкретные рабочие программы, ориентированные на специфику подготовки специалистов в конкретных учебных заведениях. Настоящее пособие представляет собой сборник коротких лекций по *математической статистике*, написанный на базе учебного материала читаемого студентам Института экономики и предпринимательства Нижегородского государственного университета Н.И. Лобачевского. Предлагаемый учебный материал главным образом рассчитан на такие направлениям подготовки как **Менеджмент** (38.03.02) и **Экономика** (38.03.01). Кроме того, он может быть использован и слушателями, проходящими подготовку по такому направлению как **Управление персоналом** (38.03.03), в процессе изучения курса математики.

Пособие рассчитано и на студентов-иностранцев, которые, как показывает опыт, далеко не всегда настолько хорошо владеют русским языком, что могут без каких-либо затруднений понять смысл различных математических терминов и выражений. В таких случаях, часто оказывается полезным переход на английское произношение и написание соответствующих математических слов и словосочетаний.

Надеемся, что использование представленного материала существенно облегчит и ускорит процессы понимания и запоминания математических терминов. Сопоставление английского и русского написания одного и того же термина повысит грамотность иностранных слушателей в русском языке и ускорит процесс освоения.

Для русскоязычных студентов, пособие окажется полезным как с точки зрения освоения курса, так и совершенствования уровня подготовки по английскому языку.

Introduction

By the present time, quite a large number of textbooks and various methodological manuals dedicated to the discipline *of Probability Theory and Mathematical Statistics* have been published. It is worth noting that, when working on them, the authors usually used specific work programs, focused on the nature of training specialists in particular educational institutions. The present textbook is a compilation of brief lectures on *Mathematical statistics*, based on the teaching material, presented to the students of Nizhny Novgorod N. I. Lobachevsky State University's Institute of Economics and Entrepreneurship. The proposed study material is mainly intended for such specialties as **Management** (38.03.02) and **Economics** (38.03.01). Moreover, it can also be used by students majoring in **Human Resources Management** (38.03.03.03) in the process of studying the Mathematics course.

The textbook is also designed for foreign students, since experience shows that they do not always have such a good command of Russian, which allows them to understand the meaning of various mathematical terms and expressions easily. In such cases, the transition to English pronunciation and spelling of the corresponding mathematical words and phrases is often useful.

This material will hopefully make it easier and faster to understand and remember mathematical terms. Comparing English and Russian spellings of the same term will increase foreign students' literacy in Russian and accelerate the learning process.

This textbook will be useful for Russian-speaking students, both in terms of mastering the course and improving their English proficiency.

Раздел 1.

1.1. Математическая статистика. Основные понятия и задачи.

Математическая статистика – наука, занимающаяся разработкой математических методов сбора, систематизации, обработки и интерпретации статистических данных, а также использования их для различных научных или практических выводов. Знание методов математической статистики и умение ими оперировать являются необходимой предпосылкой для успешного эконометрического анализа.

Основная задача математической статистики состоит в получении информации о характере поведения некоторой случайной величины по относительно небольшому количеству ее значений – *выборке*, которую получают случайным образом из всего множества значений рассматриваемой случайной величины – *генеральной совокупности*.

Приведем основные характеристики и методы анализа статистических данных, которые получили широкое распространение в такой дисциплине как эконометрика.

1.1.1. Генеральная совокупность и выборка

Как известно, статистическая устойчивость результатов наблюдений имеет место при достаточно большом числе проводимых измерений. Однако, на практике, обычно работают лишь с небольшим числом наблюдений, по той причине, что исследование всей совокупности объектов во многих случаях оказывается невозможным и нецелесообразным.

Естественно, что характеристики случайной величины, определенные по более малому числу наблюдений, могут не совпадать с теми же величинами, вычисленными по б*о*льшему числу (в пределе – бесконечно большому) наблюдений, которые выполнены в тех же условиях. Поэтому, для оценки соответствующих различий, в математической статистике вводят абстрактное понятие – *генеральная совокупность*, представляющая собой множество всех теоретически возможных значений исследуемой случайной величины, реализуемых в данных

условиях, и *выборки* (т. е. выборочной совокупности), состоящей из ограниченного числа значений (наблюдений) и представляющей собой часть генеральной совокупности, отобранную с целью изучения. В соответствии с этим принято различать *выборочные характеристикислучайной величины*, найденные по ограниченному числу наблюдений и зависящие от этого числа, и соответствующие характеристики генеральной совокупности, не зависящие от числа наблюдений. Выборочные характеристики рассматриваются как приближенные оценки соответствующих генеральных характеристик.

Основная задача выборочного метода в математической статистике состоит *в исследовании свойств выборки и обобщении этих свойств с наибольшей надежностью на всю генеральную совокупность.*

Можно утверждать, что выборка будет *репрезентативной*, т. е. представительной – достаточно полно отражающей пропорции генеральной совокупности, если отбор объектов для исследования в процессе ее формирования будет носить случайный характер.

1.1.2. Способы представления статистических данных.

При анализе какого-либо экономического показателя X в фиксированный момент времени (либо без учета фактора времени) регистрируют и представляют в виде ряда его наблюдаемые выборочные значения: x_1, x_2, …, x_n, которые принято называть *вариантами*. При проведении статистического анализа, эти значения обычно упорядочивают по возрастанию или убыванию, т.е. производят *ранжирование* вариантов ряда. Далее, в случае большого объема выборки, производят *группировку* вариантов, разбивая их на отдельные интервалы. При этом, числа, показывающие сколько раз встречаются те или иные варианты в каждом конкретном интервале, называются *частотами*, а их отношение к общему числу наблюдений – *относительными частотами*/. Частоты принято еще называть *весами*.

Согласно определению, *ряд вариантов с соответствующими им весами, ранжированный в порядке возрастания или убывания, называется вариационным рядом.*

Принято различать два вида вариационных рядов: *дискретные* и *непрерывные*. Ряд называется *дискретным*, если любые его варианты отличаются на постоянную величину, и – *непрерывным (интервальным),* если варианты могут отличаться один от другого на сколь угодно малую величину. В процессе изучения вариационных рядов, кроме таких понятий как частота и относительная частота, используются понятия *накопленной частоты* и *накопленной относительной частоты*. Накопленная частота $n_i^{нак}$ показывает, сколько наблюдалось вариантов со значением признака, меньшим x. Отношение $n_i^{нак}$ к общему числу наблюдений n, т. е. $w_i^{нак} = n_i^{нак} / n$ называется накопленной *относительной частотой*.

Предположим, что имеется выборка, в которой количество различных вариантов x_i равно k (т.е., i = 1, 2, …, k) ($k \le n$, где n – общее число наблюдений). При этом $x_1 < x_2 < \ldots < x_k$. Если какое-либо значение x_i встретилось в выборке n_i раз, то число n_i это *абсолютная частота* значения x_i, а величина $\omega_i = n_i / n$ – *относительная частота* появления значения x_i. Тогда, наблюдаемые выборочные значения могут быть представлены в виде следующего *вариационного ряда*:

X	$x_1, x_2, \ldots x_k$
n_i	$n_1, n_2, \ldots n_k$
$\omega_i = \frac{n_i}{n}$	$\frac{n_1}{n}, \frac{n_2}{n}, \ldots \frac{n_k}{n}$

При этом, $\sum_{i=1}^{k} n_i = n$, $\sum_{i=1}^{k} (n_i / n) = 1$.

С помощью вариационного ряда можно также построить эмпирическую функцию распределения для рассматриваемой случайной величины.

***Определение*:.** *Эмпирической (выборочной) функцией распределения* $F_n(x)$ величины X называется статистическая вероятность (относительная частота) появления события, заключающегося в том, что X примет значение, меньше указанного x, т. е.:

$$F_n(x) = \omega(X < x) = w_i^{нак}$$

Другими словами, эмпирическая функция распределения $F_n(x)$, для данного значения x, представляет собой накопленную частоту $w_i^{нак} = n_i^{нак} / n$.

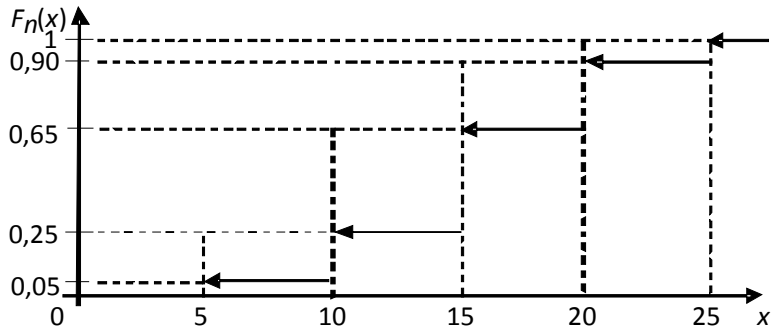
Пример 1. Анализируется прибыль X (%) предприятий некоторой отрасли. Обследованы n = 100 предприятий, данные по которым занесены в следующий вариационный ряд:

X,(%)	5	10	15	20	25
n_i	5	20	40	25	10
n_i / n	0,05	0,2	0,4	0,25	0,1

Необходимо определить эмпирическую функцию распределения $F_n(x)$ и построить ее график.

Решение.

$$F_n(x) = \begin{cases} 0, & x \le 5 \\ 0{,}05, & 5 < x \le 10 \\ 0{,}25, & 10 < x \le 15 \\ 0{,}65, & 15 < x \le 20 \\ 0{,}90, & 20 < x \le 25 \\ 1 & x > 25. \end{cases}$$



Как уже отмечалось выше, при большом объеме выборки, ее элементы могут быть сгруппированы в интервальный вариационный ряд. Для этого, n наблюдаемых значений выборки x_1, x_2, …, x_n разбивают на k непересекающихся интервалов равной ширины h (h – шаг разбиения). Согласно формуле Стерджеса, рекомендуемое число таких интервалов $k = 1 + 3{,}322 \lg n$, а $h = (x_{\max} - x_{\min}) / (1 + 3{,}322 \lg n)$, где $(x_{\max} - x_{\min})$ – размах выборки.

Пусть, например, n_i – количество наблюдаемых значений случайной величины X, попадающих в i-й интервал; $\omega_i = n_i / n$ – относительная частота попадания значений X в i-й интервал. Тогда интервальный вариационный ряд будет иметь вид:

$[x_{i-1}, x)$	$[x_0, x_1)$	$[x_1, x_2)$	…	$[x_{k-1}, x_k)$
n_i	n_1	n_2	…	n_k
n_i / n	n_1 / n	n_2 / n	…	n_k / n

Интервальный вариационный ряд наглядно может быть представлен в виде *гистограммы* – графика, где по оси абсцисс откладываются интервалы, на каждом из которых строятся прямоугольники с высотой и площадью, пропорциональной относительной частоте попадания случайной величины X в данный интервал. Так, на i-м интервале строится прямоугольник высотой n_i / nh. На основании гистограммы обычно выдвигают предположение о виде закона распределения исследуемой случайной величины.

Кроме гистограммы, для графического изображения вариационных рядов часто используются *полигон* и *кумулятивная кривая*.

Полигон представляет собой ломаную линию, в которой концы отрезков прямой имеют координаты (x_i, n_i), $i = 1,2,...m$. Он служит для изображения дискретного вариационного ряда.

Кумулятивная кривая (кумулята) представляет кривую накопленных частот. Для дискретного ряда – ломаная линия, которая соединяет точки $(x_i, n_i^{нак})$ или $(x_i, w_i^{нак}), i = 1,2,...m$. Для интервального ряда – ломаная начинается с точки, абсцисса которой равна началу первого интервала, а ордината равна накопленной частоте (или накопленной относительной частоте) равной нулю. Другие точки этой ломаной соответствуют концам интервалов.

1.2. Числовые характеристики выборки

Поскольку на практике обычно работают с выборкой, нас будут интересовать именно выборочные числовые характеристики, которые являются оценками соответствующих генеральных характеристик.

1.2.1 Средние величины и показатели вариации

Прежде всего, отметим, что средние величины, характеризующие значения признака, вокруг которого концентрируются наблюдения, принято различать на *аналитические* и *порядковые*.

1.2.1.1. Аналитические средние величины

Самым простым показателем вариации является разность между максимальным и минимальным выборочным значением, которую принято называть

размахом выборки: $R = x_{\max} - x_{\min}$. Однако, наибольший интерес представляют те величины, которые представляют меры рассеяния наблюдений вокруг средних величин.

Согласно определению, *средним значением дискретной случайной величины* X, *или ее математическим ожиданием* $M(X)$, *называется сумма произведений всех ее значений* x_i *на соответствующие этим значениям вероятности* $p_i\ (i = 1,...n)$ т. е.:

$$M(X) = \sum_{i=1}^{n} x_i p_i$$

Если в этой формуле положить все вероятности $p_i = 1/n$, то получим *выборочное среднее арифметическое* наблюдаемых значений выборки (задаваемой в виде не сгруппированного ряда) для рассматриваемой случайной величины X:

$$M_в(X) = \bar{x}_в = \frac{1}{n}\sum_{i=1}^{m} x_i \qquad (1)$$

Соответственно, для *выборочной дисперсии* $D_в(X)$ *и среднеквадратического отклонения* s будут иметь место следующие выражения:

$$D_в(X) = \frac{1}{n}\sum_{i=1}^{m}(x_i - \bar{x}_в)^2\ ; \qquad s = \sqrt{\frac{1}{n}\sum_{i=1}^{m}(x_i - \bar{x}_в)^2} \qquad (2)$$

Кроме того, используется такая характеристика как *среднее линейное (или среднее абсолютное) отклонение вариационного ряда*:

$$d = \frac{1}{n}\sum_{i=1}^{m}|\ x_i - \bar{x}\ | \qquad (3)$$

В том случае, если выборка задается в виде вариационного ряда, выражения (1)-(3) принимают вид:

$$M_в(X) = \bar{x}_в = \frac{1}{n}\sum_{i=1}^{k} n_i x_i\ , \qquad D_в(X) = \frac{1}{n}\sum_{i=1}^{k} n_i (x_i - x_в)^2$$

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{n} n_i (x_i - \bar{x}_в)^2} \qquad d = \frac{1}{n}\sum_{i=1}^{m}|\ x_i - \bar{x}\ | n_i$$

Используется также безразмерная характеристика – *коэффициент вариации*:

$$V_в = \frac{s}{\bar{x}} \cdot 100\ \%.(\ \bar{x} \neq 0\ )$$

1.2.1.2. Порядковые средние величины

Наиболее часто из этих величин используются *мода* и *медиана*. Согласно определению, *мода* M_o *представляет собой варианту, которая имеет наибольшую частоту.*

Например, для ряда

Варианта x_i	1	2	3	3	5	6
Частота n_i	2	3	6	8	25	10

мода равна 5.

Медианой M_e *называется варианта, приходящаяся на середину ранжированного ряда наблюдений.*

Можно сказать, что медиана представляет собой варианту, которая делит вариационный ряд на две части равные по числу вариант. Если число вариант нечетное, т. е. $n = 2k+1$, $(k = 0,1,2,...)$, то $M_e = x_{k+1}$. При четном значении числе n, т.е, когда $n = 2k$: $M_e = (x_k + x_{k+1})/2$.

Так, например, для ряда: 2, 3, 5, 6, 7, медиана равна 5; для ряда: 2, 3, 5, 6, 7, 9 медиана равна $(5+6)/2 = 5{,}5$.

1.2.1.3 Начальные и центральные моменты

Более общими характеристиками вариационного ряда, чем среднее арифметическое и дисперсия, являются такие характеристики как *моменты.*

Начальным моментом порядка k *случайной величины* X *принято называть математическое ожидание величины* X^k, т. е.

$$v_k = \frac{1}{n}\sum_{i=1}^{m} n_i x_i^k$$

Центральным моментом порядка k *случайной величины* X *принято называть математическое ожидание величины* $(X - M(X))^k$, т. е.

$$\mu_k = \frac{1}{n}\sum_{i=1}^{m} n_i (x_i - \bar{x})^k$$

Коэффициентом асимметрии вариационного ряда называется число

$$A = \frac{1}{ns^3}\sum_{i=1}^{m} n_i (x_i - \bar{x})^3$$

В том случае если коэффициент асимметрии равен нулю, то это говорит о том, что распределение имеет симметричную форму. При $A > 0$ ($A < 0$), соответственно, имеет место положительная (правосторонняя) или отрицательная (левосторонняя) асимметрия.

Эксцессом вариационного ряда называется число

$$E = \frac{1}{ns^4}\sum_{i=1}^{m} n_i (x_i - \bar{x})^4 - 3$$

Данная величина является показателем «крутости» вариационного ряда по сравнению с нормальным распределением. При $E > 0$ ($E < 0$), полигон вариационного ряда имеет более крутую (пологую) вершину по сравнению с нормальной кривой.

Section 1.
1.1. Mathematical statistics. Major concepts and problems.

Mathematical statistics is a science dealing with development of mathematical methods of collection, systematization, processing and interpretation of statistical data, their use for various scientific or practical findings. Ability to operate with mathematical statistics methods is a necessary prerequisite for successful econometric analysis.

Mathematical statistics is aimed at obtaining information about a random variable characteristic according to a relatively small number of its values, i.e. a *sample* that is obtained randomly from the whole set of values of the considered random variable, i.e. the *entire assembly*.

Below are the main characteristics and methods of statistical data analysis, which are widely used in a discipline called econometrics.

1.1.1. Entire assembly and a sample

The statistical stability of observation results is known to take place at sufficiently large number of conducted measurements. But in practice we usually work only with a small number of observations, due to the fact that it is impossible and unreasonable to study the whole set of objects in many cases.

The characteristics of a random variable determined from a smaller number of observations may naturally not coincide with those calculated from a larger number (infinitely larger in the limit) of observations made under the same conditions. For this reason, in order to estimate the corresponding differences, mathematical statistics introduces an abstract notion - the *entire assembly*, which is a set of all theoretically possible values of the studied random variable materialized under given conditions, and a *sample* (i.e. sampling population) consisting of a limited number of values (observations) and representing a part of the entire assembly selected for the purpose of study. Accordingly, it is customary to distinguish *sampling characteristics of a random variable* found from a limited number of observations and which depend on this number, and corresponding characteristic of the entire assembly, which do not depend on the number of observations. Sample characteristics are considered to be approximate estimates of the corresponding general characteristics.

Sampling method in mathematical statistics has the primary task of *investigating the properties of a sample and generalizing these properties to the whole entire assembly with the greatest reliability*.

We can state that the sample will be *representative*, i.e. it will be sufficiently reflecting the proportions of the entire assembly, if the selection of objects for study during its formation will be of random nature.

1.1.2. Ways of presenting statistical data.

During the analysis of any economic indicator X at a fixed point in time (or without taking into account the time aspect) its observed sample values are recorded and presented as a series: $x_1, x_2, \ldots, x_n$, which are usually called *variants*. When conducting statistical analysis, these values are usually arranged in ascending or descend-

ing order, i.e. the *ranking* of the variants in the series is made. Furthermore, in the case of a large sample size, *grouping* of variants, by dividing them into separate intervals, is made. In this case, the numbers showing how many times these or those variants occur in each specific interval are called frequencies, and their ratio to the total number of observations is called *relative frequencies*. Frequencies are also commonly referred to as *weights*.

The definition calls *a series of variants with their respective weights, ranked in ascending or descending order, a variation series.*

Two types of variation series: *discrete* and *continuous* are commonly used. A series is called *discrete* if any of its variants differ by a constant value, and it is called *continuous (interval)* if variants may differ from one another by an arbitrarily small value. While studying variation series, in addition to such concepts as frequency and relative frequency, the concepts of *cumulative frequency* and *cumulative relative frequency* are used. The cumulative frequency $n_i^{нак}$ shows how many variants with a characteristic value less than x were observed. The ratio of $n_i^{нак}$ to the total number of observations n, i.e. $w_i^{нак} = n_i^{нак} / n$ is called the *cumulative relative frequency*.

Let us assume that there is a sample in which the number of different variants x_i is equal to k (i.e., i = 1, 2, …, k) $k \le n$, in which $n-$is the total number of observations). In this case $x_1<x_2< \ldots <x_k$. If any value of x_i occurs $n_i$times in the sample, then the number $n_i$is the *absolute frequency* of the x_i value, and $\omega_i = n_i / n$value is the *relative frequency* of the x_i value occurrence. Then, the observed sample values can be represented as the following *variation series*:

X	$x_1, x_2, \ldots x_k$
n_i	$n_1, n_2, \ldots n_k$
$\omega_i = \frac{n_i}{n}$	$\frac{n_1}{n}, \frac{n_2}{n}, \ldots \frac{n_k}{n}$

Thereby, $\sum_{i=1}^{k} n_i = n$, $\sum_{i=1}^{k} (n_i / n) = 1$.

An empirical distribution function for the random variable in question can also be plotted using a variation series.

Definition: *The empirical (sample) distribution function $F_n(x)$* of the variable X is the statistical probability (relative frequency) of the event where X acquires a value less than the specified x, i.e:

$$F_n(x) = \omega(X < x) = w_i^{нак}$$

In other words, the empirical distribution function $F_n(x)$ is the cumulative frequency $w_i^{нак} = n_i^{нак} / n$ for a given value of x

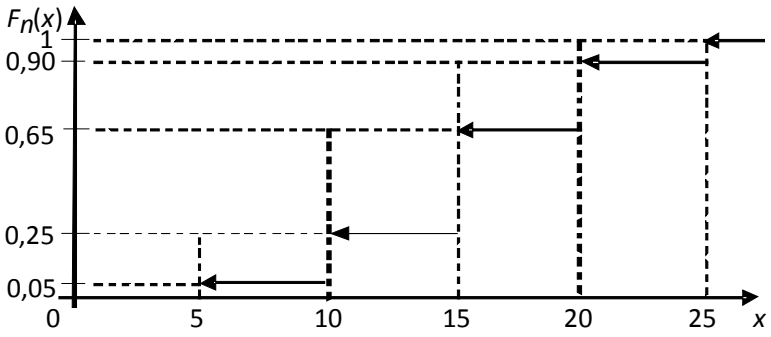
Example 1. Profit X (%) of enterprises of some industry is analyzed. $n = 100$ enterprises are surveyed, the data on which are recorded in the following variation series:

X,(%)	5	10	15	20	25
n_i	5	20	40	25	10
n_i / n	0.05	0.2	0.4	0.25	0.1

It is necessary to determine the empirical distribution function $F_n(x)$ and plot its graph.

Solution.

$$F_n(x) = \begin{cases} 0. & x \le 5 \\ 0.05, & 5 < x \le 10 \\ 0.25, & 10 < x \le 15 \\ 0.65, & 15 < x \le 20 \\ 0.90, & 20 < x \le 25 \\ 1 & x > 25. \end{cases}$$



As we mentioned above, if the sample size is large, its elements can be grouped into an interval variation series. For this purpose, n observed sample values $x_1, x_2, \ldots, x_n$ are divided into k non-intersecting intervals of equal width h (h is the step of division). According to Sturges' formula, the recommended number of such intervals is $k = 1 + 3.322 \lg n$, and $h = (x_{\max} - x_{\min}) / (1 + 3.322 \lg n)$, in which$(x_{\max} - x_{\min})$ is the sample spread.

Let us assume, that n_i is the number of observed values of a random variable X which fall into the i-th interval; $\omega_i = n_i / n$ is the relative frequency of values of X fall-

ing into the i-th interval. Then the interval variation series will have the following form:

$[x_{i-1}, x)$	$[x_0, x_1)$	$[x_1, x_2)$	…	$[x_{k-1}, x_k)$
n_i	n_1	n_2	…	n_k
n_i / n	n_1 / n	n_2 / n	…	n_k / n

An interval variation series can be visualized as a *histogram* - a graph, where intervals are plotted along the abscissa axis, each of which contains rectangles with the height and area proportional to the relative frequency of a random variable x occurring in a given interval. Thus, a rectangle with height n_i / nh is plotted on the i-th interval. The histogram is usually used as a basis for assuming a distribution law of the random variable in question.

In addition to the histogram, *polygon* and *cumulative curve* are often used for graphical representation of variation series.

A polygon is a broken line in which the ends of the line segments have coordinates (x_i, n_i), $i = 1,2,...m$. It serves to represent a discrete variation series.

The cumulative curve is a curve of cumulative frequencies. For a discrete series, it is a broken line that connects points $(x_i, n_i^{нак})$ or $(x_i, w_i^{нак}), i = 1,2,...m$. For an interval series, the broken line starts with a point the abscissa of which is equal to the beginning of the first interval, and the ordinate is equal to the cumulative frequency (or cumulative relative frequency) equal to zero. Other points of this broken line correspond to the ends of intervals.

1.2. Numerical characteristics of the sample

Since we are usually working with a sample in practice, we will be interested in the sample numerical characteristics, which are estimates of the corresponding general characteristics.

1.2.1 Averages and variation indicators

The first thing to note is that it is customary to distinguish between *analytical* and *ordinal* averages, which characterize the values of the attribute over which observations are concentrated.

1.2.1.1. Analytical averages

The simplest variation indicator is the difference between the maximum and the minimum sample value, which is usually called the *sampling spread*:

$R = x_{\max} - x_{\min}$. However, the most interesting quantities are those that represent measures of observation dispersion around the averages.

According to the definition, *the mean value of a discrete random variable* X*, or its mathematical expectation* $M(X)$*, is the sum of the products of all its values* x_i *by the probabilities* $(i = 1,...n)$ *corresponding to these values,* i.e*:*

$$M(X) = \sum_{i=1}^{n} x_i p_i$$

If we put all the probabilities of $p_i = 1/n$ in this formula, we obtain the *sample mean* of the observed values of the sample (given as an ungrouped series) for the random variable in question X:

$$M_в(X) = \bar{x}_в = \frac{1}{n}\sum_{i=1}^{m} x_i \qquad (1)$$

Accordingly, the following expressions will apply for the *sampling variance* $D_в(X)$ and the *standard deviation* s:

$$D_в(X) = \frac{1}{n}\sum_{i=1}^{m}(x_i - \bar{x}_в)^2; \qquad s = \sqrt{\frac{1}{n}\sum_{i=1}^{m}(x_i - \bar{x}_в)^2} \qquad (2)$$

In addition, the *linear mean (or average absolute) deviation of the variation series* is used as a characteristic:

$$d = \frac{1}{n}\sum_{i=1}^{m}| x_i - \bar{x} | \qquad (3)$$

If the sample is given in the form of a variation series, expressions (1)-(3) will take the following form:

$$M_в(X) = \bar{x}_в = \frac{1}{n}\sum_{i=1}^{k} n_i x_i, \qquad D_в(X) = \frac{1}{n}\sum_{i=1}^{k} n_i (x_i - x_в)^2$$

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{n} n_i (x_i - \bar{x}_в)^2} \qquad d = \frac{1}{n}\sum_{i=1}^{m}| x_i - \bar{x} | n_i$$

A dimensionless characteristic, *the variation coefficient*, is also used:

$$V_в = \frac{S}{\bar{x}} \cdot 100\ \%.(\ \bar{x} \neq 0\ )$$

1.2.1.2. Ordinal averages

Mode and *median* are the most commonly used values. According to the definition, mo*de* M_o *is the variant that has the highest frequency*.

For example, in a series

Variant x_i	1	2	3	3	5	6
Frequency n_i	2	3	6	8	25	10

the mode is 5.

The median M_e *is the variant falling in the middle of the ranked series of observations.*

We may say that the median is a variant which divides the variation series into two parts equal in the number of variants. If the number of variants is odd, i.e. $n = 2k + 1$, $(k = 0,1,2,...)$, then $M_e = x_{k+1}$. If n is even, i.e., when $n = 2k$: $M_e = (x_k + x_{k+1})/2$.

For example, in the series: 2 , 3 , 5 , 6 , 7 , the median is equal to 5 ; in the series: 2 , 3 , 5 , 6 , 7 , 9 the median is equal to $(5+6)/2 = 5.5$.

1.2.1.3 Initial and central moments

Moments are more general characteristics of a variation series compared to the arithmetic mean and variance.

The initial moment of order k *of a random variable* X *is usually called the mathematical expectation of the value* X^k, i.e.

$$v_k = \frac{1}{n}\sum_{i=1}^{m} n_i x_i^k$$

The central moment of order k *of a random variable* X *is usually called the mathematical expectation of the value* $(X - M(X))^k$, i. e.

$$\mu_k = \frac{1}{n}\sum_{i=1}^{m} n_i (x_i - \bar{x})^k$$

The asymmetry coefficient of a variation series is the number

$$A = \frac{1}{ns^3} \sum_{i=1}^{m} n_i (x_i - \bar{x})^3$$

If the asymmetry coefficient is zero, it means that the distribution has a symmetric form. If $A > 0$ ($A < 0$) the asymmetry is positive (right-handed) or negative (left-handed), respectively.

The variation series excess is the number

$$E = \frac{1}{ns^4} \sum_{i=1}^{m} n_i (x_i - \bar{x})^4 - 3$$

This value is an indicator of variation series "steepness" as compared to the normal distribution. If $E > 0$ ($E < 0$), the polygon of variation series has a steeper (hollow) peak compared to the normal curve

Раздел 2.

2.1. Оценки параметров и свойства выборочных оценок

Предположим, что имеется генеральная совокупность представляющая собой ряд значений какой-то интересующей нас случайной величины X с известным законом распределения зависящим от одного или нескольких параметров. Возникает задача оценки этих параметров. При этом, какие именно параметры подлежат оценке зависит от конкретного вида закона распределения. Так, например, если известно, что интересующая случайная величина X распределена по закону Пуассона, то требуется оценить параметр λ, которым это распределение полностью определяется; если X распределена по нормальному закону – требуется проводить оценку математического ожидания и среднего квадратического отклонения.

В распоряжении исследователя обычно только данные конкретной выборки полученные в результате проведенных n наблюдений (опытов): X_1, $X_2,\ldots,X_n$, по которым требуется определить неизвестный параметр θ.

Величины X_1, $X_2,\ldots,X_n$ являются случайными: X_1 – реализация первого наблюдения, X_2 – второго и т. д. Причем, случайные величины X_i, $(i=1,2,\ldots,n)$ имеют такое же распределение, что и случайная величина X.

Найти *статистическую* оценку (или просто оценку) $\widetilde{\theta}_n$ неизвестного параметра θ_n теоретического распределения – это значит найти функцию от наблюдаемых случайных величин, которая и дает приближенное значение оцениваемого параметра.

$$\widetilde{\theta}_n = \widetilde{\theta}_n(X_1, X_2,\ldots, X_n)$$

Функцию выборки принято называть *статистикой.*

Сам процесс нахождения оценок неизвестного генерального параметра θ, от которого зависит распределение случайной величины *X*, будем называть *оцениванием*. Следует отметить, что необходимые оценки всегда строятся на основе ограниченного набора данных. Это влечет за собой определенную вероятность погрешности в статистических выводах. Кроме того, от выборки к выборке, значения оценок могут изменяться. В связи с этим, при исследовании различных параметров генеральной совокупности на основе выборки, возможно получение лишь приближенных оценок интересующих параметров.

Цель любого оценивания – получение наиболее точного значения определяемой характеристики (наилучшей оценки).

Принято различать два вида оценок параметров распределения генеральной совокупности – *точечные* и *интервальные*.

2.1.1 Точечные оценки

Точечной оценкой $\tilde{\theta}$ *параметра* θ *называется числовое значение этого параметра, полученное по определенным правилам, по выборке объема n.* Как уже было отмечено выше, оценка $\tilde{\theta}$ является функцией выборки, отобранной для изучения: $\tilde{\theta} = \tilde{\theta}(X_1, X_2, ..., X_n)$. Следовательно, она может рассматриваться как случайная величина со своими числовыми характеристиками. При этом, качество оценки определяют, проверяя, обладает ли она свойствами *несмещенности, состоятельности и эффективности.*

Оценка $\tilde{\theta}$ *называется несмещенной, в том случае если, при любом объеме выборки, ее математическое ожидание равно оцениваемому параметру, т.е.,* $M(\tilde{\theta}) = \theta$.

При $M(\tilde{\theta}) \to \theta$, оценку $\tilde{\theta}$ принято называть *асимптотически несмещенной.*

В том случае, если равенство $M(\tilde{\theta}) = \theta$ не выполняется, оценка является *смещенной, а* разность $(M(\tilde{\theta}) - \theta)$ – *смещением* или *систематической ошибкой оценивания*. При этом, возможны два случая:

1). $M(\tilde{\theta}) > \theta$, то оценка $\tilde{\theta}$ дает систематическую ошибку в сторону завышения;

2). $M(\tilde{\theta}) < \theta$ – в сторону занижения.

Отметим, что требование несмещенности особое значение имеет при малом числе наблюдений (опытов).

Оценка $\tilde{\theta}$ *параметра* θ *называется состоятельной, если она сходится по вероятности к оцениваемому параметру;* т.е., для любого сколь угодно малого $\varepsilon > 0$ должно выполняться условие:

$$\lim_{n \to \infty} P(\left|\tilde{\theta} - \theta\right| < \varepsilon) = 1,$$

Последнее равенство означает, что с увеличением объема выборки мы все ближе подходим к истинному значению параметра θ.

Свойство состоятельности является обязательным для любого правила оценивания. (Несостоятельные оценки не используются !)

Эффективной принято называть статистическую оценку $\tilde{\theta}$*, которая (при заданном объеме выборки* n*) должна обладать наименьшим разбросом относительно оцениваемого параметра* θ*, т. е. должна иметь наименьшую возможную дисперсию среди всех возможных несмещенных оценок.*

Другими словами, оценка $\tilde{\theta}$ эффективна, если ее дисперсия минимальна.

Наиболее известными методами нахождения точечных оценок параметров генеральной совокупности являются *метод моментов*, *метод максимального правдоподобия*и *метод наименьших квадратов*.

2.1.2. Методы нахождения точечных оценок

2.1.2.1 Метод моментов

Данный метод состоит в приравнивании определенного количества *выборочных моментов распределения (начальных* v_k или центральных μ_k, *или тех и других) соответствующим теоретическим моментам распределения* ($\tilde{v}_k$ или $\tilde{\mu}_k$) *заданным по выборке.*

Здесь следует напомнить, что выборочные моменты $k-$гопорядка, случайной величины X, определяются по следующим формулам:

$$v_k = \frac{1}{n}\sum_{i=1}^{m} n_i x_i^k \qquad\qquad \mu_k = \frac{1}{n}\sum_{i=1}^{m} n_i (x_i - \bar{x})^k$$

Соответствующие им теоретические моменты определяются по формулам:

$\tilde{v}_k = \sum_{i=1}^{m} x_i^k p_i$ $\qquad$ $\tilde{\mu}_k = \sum_{i=1}^{m} (x_i - a)^k p_i -$ для дискретной случайной величины с функцией вероятностей: $p_i = \varphi(x_i, \theta)$;

$\tilde{v}_k = \int_{-\infty}^{\infty} x^k \varphi(x,\theta)dx$, $\qquad$ $\tilde{\mu}_k = \int_{-\infty}^{\infty} (x_i - a)^k \varphi(x,\theta)dx$ – для непрерывной случайной величины с плотностью вероятностей $\varphi(x,\theta)$, где $a = M(X)$.

Пример 1: Требуется найти оценку метода моментов для параметра λ закона Пуассона.

Решение: В данном случае, для нахождения единственного параметра λ, достаточно приравнять выборочный (эмпирический) $v_{k=1}$ и теоретический $\widetilde{v}_{k=1}$ начальные моменты первого порядка. Отметим, что, в данном случае, $\widetilde{v}_{k=1}$ представляет собой математическое ожидание $M(X)$ случайной величины X распределенной по закону Пуассона. Известно, что для величины с таким распределением, $M(X)=\lambda$. Что касается момента $v_{k=1}$, то, согласно формуле $v_k=\frac{1}{n}\sum_{i=1}^{m} n_i x_i^k$, он будет равен величине выборочной средней $v_k=\overline{x_в}$. Следовательно, оценка метода моментов параметра λ закона Пуассона, есть выборочная средняя $\overline{x_в}$, т.е., $\lambda=\overline{x_в}$.

Пример 2: Имеется случайная величина X распределенная по нормальному закону. Требуется найти оценки параметров распределения.

Решение:

Поскольку рассматриваемая величина распределена по нормальному закону, то вполне понятно, что в качестве параметров здесь выступают математическое ожидание и среднее квадратическое отклонение, т.к. именно они полностью определяют нормальное распределение. Поэтому, данная задача, по сути дела, сводится к тому, чтобы по выборке: $x_1,\ x_2,\ldots,x_n$ найти точечные оценки неизвестных параметров: $a=M(X)$ и $\sigma=\sqrt{D(X)}$

Согласно методу моментов, приравниваем их, соответственно, к выборочной средней ($v_1=M(X)-$ начальный момент первого порядка) и выборочной дисперсии ($\mu_2=D(X)-$ центральный момент второго порядка.

Получаем:

$$\begin{cases} M(X)=\overline{x_в} \\ D(X)=D_в \end{cases} \quad \text{или} \quad \begin{cases} a=\overline{x_в} \\ \sigma^2=D_в \end{cases}$$

Итак, искомые оценки параметров нормального распределения: $\widetilde{\theta}_1 = \overline{x_в}$ и $\widetilde{\theta}_2 = \sqrt{D_в}$.

Оценки метода моментов обычно состоятельны, однако по своей эффективности они не являются «наилучшими». Их эффективность часто оказывается значительно меньше единицы.

2.1.2.2 Метод максимального правдоподобия

Данный метод является основным методом получения оценок параметров генеральной совокупности по данным выборки.

Предположим, что имеется *непрерывная* случайная величина X , которая в результате проведенных n испытаний приняла значения x_1, x_2, …, x_n. Предположим также, что известен и вид закона распределения величины X, например, вид плотности распр*еделения f(x, θ), зависящей от неизвестного параметра θ.* Поскольку сам параметр θ не известен*, для него* требуется найти точечную оценку с помощью имеющейся выборки.

Функцией правдоподобия, построенной по выборке x_1, x_2, …, x_n, принято называть функцию аргумента θ *следующего вида:*

$$L(x_1, x_2, ..., x_n; \theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot ... \cdot f(x_n, \theta)$$

В том случае, если случайная величина X – дискретная, то функция правдоподобия имеет вид:

$$L(x_1, x_2, ..., x_n; \theta) = p(x_1, \theta) \cdot p(x_2, \theta) \cdot ... \cdot p(x_n, \theta)$$

где через $p(x_n, \theta)$ обозначена вероятность того, что в результате испытания X примет значение $x_i \ (i = 1,2,...,n)$.

Согласно методу максимального правдоподобия, в качестве оценки параметра θ принимается такое значение $\widetilde{\theta}_n$*, при котором функция правдоподобия L принимает максимальное значение.*

Нахождение оценки $\widetilde{\theta}_n$ значительно упрощается, если максимизировать не саму функцию L, а ее натуральный логарифм, т.е. $\ln(L)$. Это оказывается возможным по той причине, что максимум обеих функций достигается при одном и том же значении θ. Поэтому для отыскания оценки параметра θ (одного или нескольких) требуется решить уравнение (либо систему уравнений) правдоподобия, получаемое приравниванием производной (частных производных) нулю по параметру (параметрам) θ:

$$\frac{d\ln(L)}{d\theta} = 0 \text{ или } \frac{1}{L}\frac{dL}{d\theta} = 0$$

После этого, требуется выбрать то решение, при котором функция $\ln(L)$ принимает максимальное значение.

Пример 3: Требуется, с помощью метода максимального правдоподобия, найти оценку параметра λ в распределении Пуассона.

Решение:В данном случае. Закон распределения Пуассона принимает вид:

$$P_m(X = x_i) = \frac{\lambda^{x_i} \cdot e^{-\lambda}}{x_i!};$$

x_i – число появлений интересующего события в i-м опыте ($i = 1,2,....,n$), который состоит из m испытаний. С учетом того, что в рассматриваемом примере неизвестный параметр $\theta = \lambda$,

$$L(x_1, x_2, ..., x_n; \lambda) = p(x_1, \lambda) \cdot p(x_2, \lambda) \cdot ... \cdot p(x_n, \lambda) =$$

$$= \frac{\lambda^{x_1} e^{-\lambda}}{x_1!} \cdot \frac{\lambda^{x_2} e^{-\lambda}}{x_2!} ..... \frac{\lambda^{x_n} e^{-\lambda}}{x_n!} = \frac{\lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}}{x_1! x_1! .... x_n!}$$

Тогда логарифмическая функция:

$$\ln(L) = (\sum_{i=1}^{n} x_i)\ln(\lambda) - n\lambda - \ln(x_1! x_1! .... x_n!)$$

Первая производная от логарифмической функции по λ: $\frac{d}{d\lambda}\ln(L) = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n$

Приравняв первую производную нулю, получим уравнение правдоподобия:

$\frac{\sum_{i=1}^{n} x_i}{\lambda} - n = 0$, из которого найдем интересующее значение неизвестного параметра: $\lambda = \frac{\sum_{i=1}^{n} x_i}{n} = \overline{x_в}$. Вторая производная по λ будет иметь вид: $\frac{d^2}{d\lambda^2}\ln(L) = -\frac{\sum_{i=1}^{n} x_i}{\lambda^2}$. Поскольку она отрицательна, то точка $\lambda = \overline{x_в}$ – есть точка максимума и она подходит в качестве оценки наибольшего правдоподобия параметра λ в распределении Пуассона.

Пример 4. В результате n испытаний, величина X приняла значения x_1, x_2, …, x_n, Требуется, с помощью метода максимального правдоподобия, найти оценки параметров a и σ нормального распределения: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$

Решение: С учетом того, что в данном случае, $\theta_1 = a$ и $\theta_2 = \sigma$, функция правдоподобия будет иметь вид:

$$L = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_1-a)^2}{2\sigma^2}} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_2-a)^2}{2\sigma^2}} \cdots \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_n-a)^2}{2\sigma^2}} = \frac{1}{\sigma^n(\sqrt{2\pi})^n} e^{-\frac{\sum_{i=1}^{n}(x_1-a)^2}{2\sigma^2}}$$

Логарифмическая функция правдоподобия:

$$\ln(L) = -n\ln(\sigma) + \ln\left(\frac{1}{\left(\sqrt{2\pi}\right)^n}\right) - \frac{\sum_{i=1}^{n}(x_i - a)^2}{2\sigma^2}$$

Частные производные по a и по σ:

$$\frac{d\ln(L)}{da}=\frac{\sum_{i=1}^{n}x_i-na}{\sigma^2} \qquad \frac{d\ln(L)}{d\sigma}=-\frac{n}{\sigma}+\frac{\sum_{i=1}^{n}(x_i-a)^2}{\sigma^2}.$$

После приравнивания нулю частных производных и решения полученной системы из двух уравнений относительно a и σ^2, получим искомые оценки:

$$a=\frac{\sum_{i=1}^{n}x_i}{n}=\overline{x_в} \qquad \sigma^2=\frac{\sum_{i=1}^{n}(x_i-\overline{x_в})^2}{n}=D_в$$

Для широкого класса задач оценки метода максимального правдоподобия являются состоятельными и эффективными. В то же время, они могут быть смещенными. Недостатком метода является необходимость знания закона распределения случайной величины.

2.1.2.3 Метод наименьших квадратов

Данный метод является одним из наиболее простых приемов построения оценок. Он основан на минимизации суммы квадратов отклонений выборочных данных от искомой оценки θ. Другими словами, в методе наименьших квадратов требуется найти такое значение $\widetilde{\theta}$, при котором $\sum_{i=1}^{n}(X_i-\theta)^2\to\min$.

Пример 5 С помощью метода наименьших квадратов, найти оценку $\widetilde{\theta}_n$ для генеральной средней $\bar{x}_0$.

Решение: Найдем интересующую нас оценку $\widetilde{\theta}_n$ из условия минимизации функции: $F(\theta)=\sum_{i=1}^{n}(X_i-\widetilde{\theta}_n)^2\to\min$.

Используя необходимое условие экстремума, приравняем нулю первую производную $\frac{d}{d\widetilde{\theta}_n}F=\sum_{i=1}^{n}(X_i-\widetilde{\theta}_n)^2=0$, следовательно $\sum_{i=1}^{n}X_i-\widetilde{\theta}_n n=0$ и $\widetilde{\theta}_n=\frac{\sum_{i=1}^{n}X_i}{n}=\bar{x}$

т.е., искомая оценка генеральной средней $\bar{x}_0$ есть выборочная средняя $\bar{x}$.

Section 2.

2.1. Parameter estimates and properties of sample estimates

Let us assume that there is an entire assembly, which is a number of values of some random variable of interest with a known distribution law depending on one or more parameters. A task of estimating these parameters arises. In this case, the particular distribution law determines, which parameters are subject to estimation. For example, if it is known that a random variable of interest X is distributed according to Poisson's law, then we need to estimate the parameter λ that fully defines this distribution; if X is distributed according to the normal law, we need to estimate mathematical expectation and standard deviation.

Typically, the researcher possesses only the data of a particular sample obtained as a result of n observations (experiments): $X_1,\ X_2,\ldots,X_n$, according to which the unknown parameter θ needs to be determined.

The values $X_1,\ X_2,\ldots,X_n$ are random: $X_1-$is the implementation of the first observation, X_2- is the implementation of the second one, etc. Moreover, the random variables X_i, $(i=1,2,....,n)$ have the same distribution as the random variable X.

Finding a *statistical* estimate (or simply an estimate) $\widetilde{\theta}_n$ of an unknown parameter θ_n of a theoretical distribution means finding a function of the observed random variables that gives an approximate value of the estimated parameter.

$$\widetilde{\theta}_n=\widetilde{\theta}_n(X_1,X_2,...,X_n)$$

The sampling function is called a sample *statistics*.

The process of finding estimates of the unknown general parameter θ, on which the distribution of the random variable *X* depends, will be called *estimation*. It should be noted that the necessary estimates are always generated on the basis of a limited set of data. This entails a certain probability of error in statistical findings. Moreover, the values of the estimates may vary from sample to sample. Due to this, when studying different parameters of entire assembly on the basis of a sample, it is possible to obtain only approximate estimates of the parameters of interest.

Any estimation is aimed at getting the most exact value of a certain characteristic (the best estimation).

There are two types of estimates of entire assembly distribution parameters: *point and interval estimates*.

2.1.1 Point estimates

The point estimate $\widetilde{\theta}$ of the parameter θ is the numerical value of this parameter obtained by certain rules from a sample size n. As we mentioned above, the estimate $\widetilde{\theta}$ is a function of the sample selected for the study: $\widetilde{\theta} = \widetilde{\theta}(X_1, X_2, ..., X_n)$. Consequently, it can be regarded as a random variable with its own numerical characteristics. In this case, the quality of an estimate is determined by checking whether it has the properties of *unbiasedness, consistency, and efficiency.*

An estimate $\widetilde{\theta}$ is called unbiased if its mathematical expectation is equal to the estimated parameter at any sample size, i.e. $M(\widetilde{\theta}) = \theta$

When $M(\widetilde{\theta}) \to \theta$, the estimate $\widetilde{\theta}$ is called *asymptotically unbiased*.

If the equality $M(\widetilde{\theta}) = \theta$ is not fulfilled, the estimation is *biased*, and the difference ($M(\widetilde{\theta}) - \theta$) is a *bias* or a *systematic estimation error*. However, there are two possible cases:

1). $M(\widetilde{\theta}) > \theta$, then the estimate $\widetilde{\theta}$ gives a systematic error towards overestimation;

2). $M(\widetilde{\theta}) < \theta$, then the estimate gives a systematic error towards underestimation.

Note, that the requirement of unbiasedness is especially important in case of a small number of observations (experiments).

An estimate $\widetilde{\theta}$ of parameter θ is called consistent if it is convergent in probability to the estimated parameter; in other words, the condition must be satisfied for any arbitrarily small $\varepsilon > 0$:

$$\lim_{n \to \infty} P(\left|\widetilde{\theta} - \theta\right| < \varepsilon) = 1,$$

The last equality means that we are getting closer to the true value of θ parameter as the sample size increases.

The property of consistency is mandatory for any estimation rule. (Non-consistent estimates are not used !)

An effective estimate is a statistical estimate $\widetilde{\theta}$, which (at a given sample size n) must have the smallest scatter with respect to the estimated parameter θ, i.e. it must have the smallest possible variance among all possible unbiased estimates.

In other words, an estimate $\widetilde{\theta}$ is effective if its variance is minimal.

The best known methods of finding point estimates of entire assembly parameters are the *method of moments*, the *maximum likelihood method* and the *method of least squares*.

2.1.2. Methods of finding point estimates

2.1.2.1 Method of moments

This method consists in equating a *certain number of sampling moments* in the distribution (*initial* v_k or *central* μ_k, or *both*) *with the corresponding theoretical moments in the distribution* ($\tilde{v}_k$ or $\tilde{\mu}_k$) *given by the sample.*

It should be recalled here that the sample moments of the k – th order of the random variable X, are determined by the following formulas

$$v_k = \frac{1}{n}\sum_{i=1}^{m} n_i x_i^k \qquad\qquad \mu_k = \frac{1}{n}\sum_{i=1}^{m} n_i (x_i - \bar{x})^k$$

Their corresponding theoretical moments are determined by the formulas:

$\tilde{v}_k = \sum_{i=1}^{m} x_i^k p_i$ $\qquad$ $\tilde{\mu}_k = \sum_{i=1}^{m} (x_i - a)^k p_i$ – for a discrete random variable with a probability function: $p_i = \varphi(x_i, \theta)$;

$\tilde{v}_k = \int_{-\infty}^{\infty} x^k \varphi(x,\theta)dx$, $\qquad$ $\tilde{\mu}_k = \int_{-\infty}^{\infty} (x_i - a)^k \varphi(x,\theta)dx$ – for a continuous random variable with probability density $\varphi(x,\theta)$, in which $a = M(X)$.

Example 1: It is required to find an estimate of the method of moments for the parameter λ of Poisson's law.

Solution: In this case, in order to find a single parameter λ, it is sufficient to equate the sample (empirical) $v_{k=1}$ and the theoretical $\tilde{v}_{k=1}$ initial moments of the first order. It should be noted that, in this case, $\tilde{v}_{k=1}$ is the mathematical expectation $M(X)$ of a random variable X distributed according to Poisson's law. It is known that for a value with such a distribution, $M(X) = \lambda$. As for the moment $v_{k=1}$, according to the formula $v_k = \frac{1}{n}\sum_{i=1}^{m} n_i x_i^k$, it will be equal to the value of the sample mean $v_k = \overline{x_в}$. Con-

sequently, the estimation of the moments method of Poisson's law parameter λ, is the sample mean $\overline{x_в}$, i.e., $\lambda = \overline{x_в}$.

Example 2: We have a random variable X distributed according to the normal law. We need to find estimates of the distribution parameters.

Solution:

Since the considered value is distributed according to the normal law, it is quite clear that the parameters here are the mathematical expectation and the standard deviation, because they completely determine the normal distribution. Therefore, this task, in fact, is to find point estimates of the unknown parameters $a = M(X)$ and $\sigma = \sqrt{D(X)}$ from the sample: $x_1, x_2, \ldots, x_n$

According to the method of moments, we equate them, respectively, to the sample mean ($v_1 = M(X) -$ is the initial moment of the first order) and the sample variance ($\mu_2 = D(X) -$ is the central moment of the second order).

We get:

$$\begin{cases} M(X) = \overline{x_в} \\ D(X) = D_в \end{cases} \quad \text{or} \quad \begin{cases} a = \overline{x_в} \\ \sigma^2 = D_в \end{cases}$$

Thus, the required estimates of the normal distribution parameters: $\widetilde{\theta}_1 = \overline{x_в}$ and $\widetilde{\theta}_2 = \sqrt{D_в}$.

The estimates obtained using the method of moments are usually valid, but they are not the "best" in their efficiency. Their efficiency is often much less than 1.

2.1.2.2 Maximum likelihood method

This method is the main method of obtaining estimates of entire assembly parameters from sample data.

Let us assume that there is a *continuous* random variable X, which as a result of n tests took values $x_1, x_2, \ldots, x_n$. Let us also assume that the distribution law of the X variable is also known, e.g., the *distribution density f(x, θ) depending on the un-*

known parameter θ. Since the parameter θ itself is not known, we need to find a point estimate for it with the help of available sample.

A likelihood function plotted from a sample x_1, x_2, …, x_n„ *is usually called a function of the argument θ of the following form:*

$$L(x_1, x_2, ..., x_n; \theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot ... \cdot f(x_n, \theta)$$

If a random variable X is discrete, the likelihood function should be:

$$L(x_1, x_2, ..., x_n; \theta) = p(x_1, \theta) \cdot p(x_2, \theta) \cdot ... \cdot p(x_n, \theta)$$

where $p(x_n, \theta)$ denotes the probability that X will take the value of x_i $(i = 1,2,...,n)$ as a result of the test.

According to the maximum likelihood method, such a value of $\tilde{\theta}_n$ *at which the likelihood function L acquires the maximum value is taken as an estimate of θ parameter.*

Finding the estimate of $\tilde{\theta}_n$ is greatly simplified if we maximize not the function L itself, but its natural logarithm, i.e. $\ln(L)$. This is possible because the maximum of both functions is reached at the same value of θ. Therefore, in order to find an estimate of θ parameter (one or several parameters) it is necessary to solve the likelihood equation (or a system of equations) obtained by equating the derivative (partial derivatives) to zero by the parameter (parameters) θ:

$$\frac{d \ln(L)}{d\theta} = 0 \text{ or } \frac{1}{L}\frac{dL}{d\theta} = 0$$

After that, it is required to choose the solution at which the function $\ln(L)$ acquires the maximum value.

Example 3: It is required to find an estimate of λ parameter in the Poisson's distribution using the maximum likelihood method

Solution: In this case the Poisson's distribution law takes the following form:

$$P_m(X = x_i) = \frac{\lambda^{x_i} \cdot e^{-\lambda}}{x_i!};$$

$x_i -$ is the number of occurrences of a particular event in the i-th experiment ($i = 1,2,....,n$), which consists of m experiments. Considering that the unknown parameter in the example under consideration $\theta = \lambda$,

$$L(x_1, x_2, ..., x_n; \lambda) = p(x_1, \lambda) \cdot p(x_2, \lambda) \cdot ... \cdot p(x_n, \lambda) =$$

$$= \frac{\lambda^{x_1} e^{-\lambda}}{x_1!} \cdot \frac{\lambda^{x_2} e^{-\lambda}}{x_2!} ..... \frac{\lambda^{x_n} e^{-\lambda}}{x_n!} = \frac{\lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}}{x_1! x_1! .... x_n!}$$

Then the logarithmic function is:

$$\ln(L) = (\sum_{i=1}^{n} x_i) \ln(\lambda) - n\lambda - \ln(x_1! x_1! .... x_n!)$$

The first derivative of the logarithmic function λ: $\frac{d}{d\lambda} \ln(L) = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n$

By equating the first derivative to zero, we obtain the likelihood equation:

$\frac{\sum_{i=1}^{n} x_i}{\lambda} - n = 0$, with the help of which we obtain the value of the unknown parameter of interest: $\lambda = \frac{\sum_{i=1}^{n} x_i}{n} = \overline{x_в}$. The second derivative of λ will be: $\frac{d^2}{d\lambda^2} \ln(L) = -\frac{\sum_{i=1}^{n} x_i}{\lambda^2}$.

Since it is negative, the point $\lambda = \overline{x_в}$ is the maximum point and it is suitable as an estimate of the greatest likelihood of λ in the Poisson's distribution

Example 4. As a result of n experiments, the variable X acquired the values x_1, x_2, …, x_n, It is required to find estimates of a and σ parameters of the normal distribution using the maximum likelihood method: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$

Solution: Considering that in this case, $\theta_1 = a$ and $\theta_2 = \sigma$, the likelihood function will be:

$$L = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_1-a)^2}{2\sigma^2}} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_2-a)^2}{2\sigma^2}} \cdots \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_n-a)^2}{2\sigma^2}} = \frac{1}{\sigma^n(\sqrt{2\pi})^n} e^{-\frac{\sum_{i=1}^{n}(x_1-a)^2}{2\sigma^2}}$$

Logarithmic likelihood function:

$$\ln(L) = -n\ln(\sigma) + \ln\left(\frac{1}{\left(\sqrt{2\pi}\right)^n}\right) - \frac{\sum_{i=1}^{n}(x_i - a)^2}{2\sigma^2}$$

The partial derivatives of a and σ :

$$\frac{d\ln(L)}{da} = \frac{\sum_{i=1}^{n} x_i - na}{\sigma^2} \qquad \frac{d\ln(L)}{d\sigma} = -\frac{n}{\sigma} + \frac{\sum_{i=1}^{n}(x_i - a)^2}{\sigma^2}.$$

After equating the partial derivatives to zero and solving the resulting system of two equations with respect to a and σ^2, we obtain the desired estimates

$$a = \frac{\sum_{i=1}^{n} x_i}{n} = \overline{x_в} \qquad \sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x_в})^2}{n} = D_в$$

The estimates of the maximum likelihood method are consistent and efficient for a wide class of problems. At the same time, they can be biased. The disadvantage of the method is the need to know the distribution law of a random variable.

2.1.2.3 Least squares method

This method is one of the simplest methods of estimation. It is based on minimizing the sum of squares of the sample data deviations from the desired estimate θ. In other words, the method of least squares requires finding such a value of $\tilde{\theta}$ at which $\sum_{i=1}^{n}(X_i - \theta)^2 \to \min$.

Example 5: Using the least squares method, find an estimate $\tilde{\theta}_n$ for the general average $\bar{x}_0$.

Solution: Let's find the estimate $\widetilde{\theta}_n$ we are interested in from the minimization condition of the function: $F(\theta) = \sum_{i=1}^{n} (X_i - \widetilde{\theta}_n)^2 \to \min$.

Using the necessary extremum condition, we equate the first derivative to zero

$\frac{d}{d\widetilde{\theta}_n} F = \sum_{i=1}^{n} (X_i - \widetilde{\theta}_n)^2 = 0$, therefore $\sum_{i=1}^{n} X_i - \widetilde{\theta}_n n = 0$ and $\widetilde{\theta}_n = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{x}$

i.e., the desired estimate of the general average $\bar{x}_0$ is the sample average $\bar{x}$.

Раздел 3.

3.1. Интервальная оценка

Точечные оценки неизвестного параметра θ, которые рассматривались ранее, имеют значение только в качестве первоначальных результатов обработки наблюдений, поскольку не дают никакой информации об оценочной точности.

В связи с этим, наряду с точечными, принято рассматривать и оценки интервальные, которые позволяют получить информацию о *точности* и *надежности* оценивания неизвестного параметра, что особенно важно для выборок небольшого объема.

Согласно определению, интервальной называют оценку параметра θ, которая определяется двумя числами – концами интервала: $\tilde{\theta}_n^{(1)}$ и $\tilde{\theta}_n^{(2)}$ покрывающего оцениваемый параметр θ с заданной вероятностью γ.

Интервал $(\tilde{\theta}_n^{(1)}, \tilde{\theta}_n^{(2)})$ принято называть *доверительным*, а вероятность γ *– доверительной вероятностью* или *надежностью* оценки. Величина доверительного интервала, непосредственно характеризующая точность оценки, зависит от объема выборки n (уменьшается с ростом n) и надежности γ (увеличивается с приближением γ к единице).

Следует отметить, что довольно часто (но не всегда) доверительный интервал выбирают симметричным относительно несмещенной точечной оценки $\tilde{\theta}$. Т.е., выбирается интервал вида $(\tilde{\theta} - \varepsilon, \tilde{\theta} + \varepsilon)$.

Здесь, число $\varepsilon > 0$ такое, что $|\theta - \tilde{\theta}| < \varepsilon$, характеризует точность оценки.Оценка считается тем выше, чем меньше разность $|\theta - \tilde{\theta}|$.

Величина γ выбирается заранее равной $0{,}9$; $0{,}95$; $0{,}99$ или $0{,}999$, что зависит от конкретной решаемой задачи.

Итак, предположим что вероятность P, того что $|\theta - \tilde{\theta}| < \varepsilon$, равна γ, т.е., выполняется равенство: $P(|\theta - \tilde{\theta}| < \varepsilon) = \gamma$, или, что то же самое: $P(\tilde{\theta} - \varepsilon < \theta < \tilde{\theta} + \varepsilon) = \gamma$

Согласно последнему соотношению, вероятность того, что интервал $(\tilde{\theta} - \varepsilon, \tilde{\theta} + \varepsilon)$ покрывает неизвестный параметр θ, равна γ.

Поскольку в эконометрических задачах часто приходится находить доверительные интервалы параметров случайной величины, приведем примеры их построения. Для определенности, будем рассматривать случаи, когда выборка производится из генеральной совокупности имеющей нормальное распределение.

3.1.1. Построение доверительного интервала для оценки математического ожидания при известной дисперсии

Допустим, что имеется случайная величина X распределенная по нормальному закону. При этом, известно среднее квадратическое отклонение σ этого распределения, а величина $\gamma -$ задана.

Предположим, что ряд конкретных значений: $x_1,\ x_2,\dots x_n$, представляет собой выборку, которая получена в результате проведения n независимых наблюдений за величиной X. Для того, чтобы подчеркнуть случайный характер величин, перепишем их в виде: $X_1,\ X_2,\dots X_n$, а следовательно, под X_i будем понимать значение случайной величины X в i-м опыте. Отметим, что случайные величины $X_1,\ X_2,\dots X_n$ являются независимыми и закон распределения любой из них совпадает с законом распределения случайной величины X. При этом, выборочное среднее значение $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ также будет распределено по нормальному закону. Параметры его распределения имеют вид: $M(\overline{X}) = a$; $D(\overline{X}) = \frac{\sigma^2}{n}$.

Теперь, потребуем чтобы выполнялось соотношение: $P(|\ \overline{X} - a\ | < \varepsilon) = \gamma$, где γ – это заданная надежность.

Пользуясь формулой $P(|\ X - a\ | < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right)$ (где Φ – функция Лапласа), можно, заменив X на $\overline{X}$ и σ на $\sigma(\overline{X}) = \sqrt{D(\overline{X})} = \frac{\sigma}{\sqrt{n}}$, получить:

$$P(|\ \overline{X} - a\ | < \varepsilon) = 2\Phi\left(\frac{\varepsilon \cdot \sqrt{n}}{\sigma}\right) = 2\Phi(t) = \gamma,$$

где $t = (\varepsilon\sqrt{n})/\sigma$. Определив из последнего равенства $\varepsilon = t\sigma/\sqrt{n}$, получаем:

$$P(|\ \overline{X} - a\ | < t\sigma/\sqrt{n}) = 2\Phi(t) \text{ или}$$

$$P(\overline{X} - t\sigma/\sqrt{n} < a < \overline{X} + t\sigma/\sqrt{n}) = 2\Phi(t) = \gamma$$

В соответствии с определением доверительного интервала, получаем, что доверительный интервал для $a = M(X)$:

$$(\overline{X} - t\sigma/\sqrt{n} < a < \overline{X} + t\sigma/\sqrt{n}),$$

где t определяется из уравнения $2\Phi(t) = \gamma$ по таблицам функции Лапласа, при заданном значении γ.

(*Таблицы функции Лапласа представлены в конце любого пособия по теории вероятностей и математической статистике*)

Следовательно, можно утверждать, что полученный доверительный интервал покрывает неизвестный параметр a. Точность проведенной оценки: $\varepsilon = t\sigma/\sqrt{n}$.

Пример 1: Имеется случайная величина X распределенная по нормальному закону с известной доверительной вероятностью $\sigma = 20$. За этой величиной произведено пять независимых наблюдений, результаты которых: $x_1 = -25$, $x_2 = 34$, $x_3 = -20$, $x_4 = 10$. $x_5 = 21$. Найти оценку математического ожидания a и построить для него 95% - доверительный интервал.

Решение: Определяем среднее выборочное значение: $\overline{x}_в = \dfrac{-25+34-20+10+21}{5} = 4$. С учетом того, что $\gamma = 0{,}95$ и $\Phi(t) = 0{,}475$, по таблице функций Лапласа, получаем, что $t = 1{,}96$. Тогда $\varepsilon = \dfrac{t\sigma}{\sqrt{n}} = \dfrac{1{,}96 \cdot 20}{\sqrt{5}} \approx 17{,}5$. Доверительный интервал для $a = M(X)$: $(4 - 17{,}5; 4 + 17{,}5)$, т.е. $(13{,}5; 21{,}5)$.

Пример 2. Случайная величина X имеет нормальное распределение с известным средним квадратическим отклонением $\sigma = 3$. Требуется найти доверительные интервалы для неизвестного математического ожидания a по выборочным средним $\overline{x}$, если объем выборки $n = 36$ и задана надежность оценки $\gamma = 0{,}95$.

Решение: Согласно условию задачи: $2\Phi(t) = 0{,}95$, следовательно $\Phi(t) = 0{,}475$. По таблице функций Лапласа, получаем, что $t = 1{,}96$. Определим точность оценки:

$$\varepsilon = t\sigma/\sqrt{n} = (1{,}96 \cdot 3)/\sqrt{36 = 0{,}98}$$

Следовательно, доверительный интервал: $(\bar{x}-0{,}98;\bar{x}+0{,}98)$.

Например, если $\bar{x}=4{,}1$, то доверительный интервал имеет следующие доверительные границы:

$$\bar{x}-0{,}98=4{,}1-0{,}98=3{,}12;\ \bar{x}+0{,}98=4{,}1+0{,}98=5{,}08$$

Т.е., значение неизвестного параметра a, которое согласуется с данными выборки, удовлетворяет неравенству: $3{,}12<a<5{,}08$.

3.1.2. Построение доверительного интервала для оценки математического ожидания при неизвестной дисперсии

Предположим, что имеется случайная величина X, распределенная по нормальному закону. При этом, среднее квадратическое отклонение σ этого распределения является величиной неизвестной, а величина $\gamma-$ задана.

Требуется найти такое число ε, чтобы выполнялось соотношение:

$P(|\overline{X}-a|<\varepsilon)=\gamma$, или $P(\overline{X}-\varepsilon<a<\overline{X}+\varepsilon)=\gamma$.

Введем случайную величину: $T=\dfrac{\overline{X}-a}{(S/\sqrt{n})}$, где n – объем выборки; S – так называемое исправленное среднее квадратическое отклонение, которое вычисляется по выборке: $S=\sqrt{\dfrac{1}{n-1}\cdot\sum_{i=1}^{n}(X_i-\overline{X}_i)^2}$

Величина T имеет распределение Стьюдента с $(n-1)$ степенями свободы. Плотность этого распределения имеет вид:

$f(t,n-1)=\dfrac{\Gamma(n/2)}{\sqrt{\pi(n-1)}\cdot\Gamma\left(\dfrac{n-1}{2}\right)}\cdot\left(1+\dfrac{t^2}{n-1}\right)^{-\frac{n}{2}}$, где Γ – гамма функция (берется по справочным таблицам).

В левой части равенства $P(|\overline{X}-a|<\varepsilon)=\gamma$, перейдем от случайной величины $\overline{X}$ к случайной величине T:

$$P\left(\frac{|\overline{X}-a|}{\frac{S}{\sqrt{n}}}<\frac{\varepsilon}{\frac{S}{\sqrt{n}}}\right)=\gamma \quad \text{или} \quad P(|T|<\varepsilon/(S/\sqrt{n}))=\gamma \quad \text{или} \, P(|T|<t_\gamma)=\gamma,$$

где $t_\gamma=\dfrac{\varepsilon\cdot\sqrt{n}}{S}$.

Величина t_γ находится из равенства: $2\int_0^{t_\gamma}f(t,n-1)dt=\gamma$.

Пользуясь таблицей для распределения Стьюдента, находим значение t_γ в зависимости от доверительной вероятности γ и числа степеней свободы $(n-1)$. Далее, определив значение t_γ из равенства $t_\gamma = (\varepsilon \cdot \sqrt{n})/S$, находим значение $\varepsilon = t_\gamma \cdot \frac{S}{\sqrt{n}}$.

Следовательно, равенство $P(|\overline{X} - a| < \varepsilon) = \gamma$ принимает вид:

$$P\left(\overline{X} - t_\gamma \cdot \frac{S}{\sqrt{n}} < a < \overline{X} + t_\gamma \cdot \frac{S}{\sqrt{n}}\right) = \gamma$$

Это означает, что интервал $\left(\overline{X} - t_\gamma \cdot \frac{S}{\sqrt{n}}; \overline{X} + t_\gamma \cdot \frac{S}{\sqrt{n}}\right)$ покрывает значение $a = M(X)$ с вероятностью γ, т.е., является доверительным интервалом для математического ожидания рассматриваемой случайной величины X.

Пример 3. По условию примера 1, считая что случайная величина X распределенная по нормальному закону с неизвестной доверительной вероятностью σ, построить для неизвестного $a = M(X)$ доверительный интервал. Считать, что $\gamma = 0{,}95$.

Решение: Согласно проведенной выше оценке, $\overline{x} = 4$. Находим значение S:

$$S^2 = \frac{1}{4}((-25-4)^2 \cdot 1 + (34-4)^2 + (-20-4)^2 + (10-4)^2 + (21-4)^2) = 660{,}5$$

$S \approx 25{,}7$. По таблице для $\gamma = 0{,}95$ и $n - 1 = 4$, находим $t_\gamma = 2{,}78$. Следовательно, $\varepsilon = 2{,}78 \cdot \frac{25{,}7}{2{,}24} \approx 31{,}9$ и доверительный интервал такой: $(-27{,}9; 35{,}9)$.

3.1.2.1. Доверительный интервал для среднего квадратического отклонения нормального распределения

Предположим, что имеется случайная величина X, распределенная по нормальному закону. При этом, σ – неизвестная величина, а значение γ задано. Установлено, что если известно $a = M(X)$, то доверительный интервал для среднего квадратического отклонения σ имеет вид: $\left(\frac{\sqrt{n} \cdot S_0}{\chi_2}; \frac{\sqrt{n} \cdot S_0}{\chi_1}\right)$, где n – объем выборки; $S_0^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - a)^2$,

а $\chi^2_{\frac{1+\gamma}{2};n}$; $\chi^2_{\frac{1-\gamma}{2};n}$ это так называемые *квантили* χ^2 распределения с n степенями свободы. Они определяются по специальной таблице квантилей $\chi^2_{\alpha,n}$ распределения χ^2_n.

Если значение $a = M(X)$ не известно, то доверительный интервал для неизвестного σ имеет вид:
$\left(\frac{\sqrt{n-1} \cdot S}{\chi_2}; \frac{\sqrt{n-1} \cdot S_0}{\chi_1} \right)$, где n – объем выборки, $S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2}$ – исправленное среднее квадратическое отклонение; $\chi_1^2 = \chi^2_{\frac{1+\gamma}{2};n-1}$ и $\chi_2^2 = \chi^2_{\frac{1-\gamma}{2};n-1}$ – квантили определяемые по таблице $\chi^2_{\alpha,k}$ при $k = n-1$, а также, соответственно, при $\alpha = \frac{1+\gamma}{2}$ и $\alpha = \frac{1-\gamma}{2}$.

Пример 4. С целью оценки параметра нормально распределенной случайной величины была сделана выборка объемом в 30 и вычислено $S = 1{,}5$. Найти доверительный интервал, покрывающий σ с вероятностью $\gamma = 0{,}90$.

Решение: Согласно условию задачи, имеем: $n = 30$, $\gamma = 0{,}9$. Далее, по таблице $\chi^2_{\alpha,k}$ находим

$$\chi_1^2 = \chi^2_{\frac{1+0{,}9}{2};30-1} = \chi^2(0{,}95;29) = 17{,}7\,;\ \chi_2^2 = \chi^2_{\frac{1-0{,}9}{2};30-1} = \chi^2(0{,}05;29) = 42{,}6\,.$$

Следовательно, доверительный интервал имеет вид:
$\left(\frac{\sqrt{30-1} \cdot 1{,}5}{\sqrt{42{,}6}}; \frac{\sqrt{30-1} \cdot 1{,}5}{\sqrt{17{,}7}} \right)$, или $1{,}238 < \sigma < 1{,}920$.

Section 3.
3.1. Interval estimation

The point estimates of the unknown parameter θ, which were considered earlier, have significance only as initial results of observational processing, since they do not provide any information about the estimated accuracy.

In this connection, along with point estimates, it is customary to consider interval estimates, which allow us to obtain information about the *accuracy* and *reliability* of the unknown parameter estimation, which is especially important for small size samples.

According to definition, an interval estimation is an estimation of parameter θ, which is defined by two numbers - the ends of interval $\tilde{\theta}_n^{(1)}$ and $\tilde{\theta}_n^{(2)}$, covering estimated parameter θ with given probability γ.

The interval $(\tilde{\theta}_n^{(1)}, \tilde{\theta}_n^{(2)})$ is usually called a *confidence* interval, and the probability γ is called the *confidence probability* or *estimation reliability*. The confidence interval value, which is a direct characteristic of estimation accuracy, depends on the sample

size n (decreasing as n increases) and the reliability γ (increasing as γ approaches 1).

It is worth noting that quite often (but not always) the confidence interval is chosen symmetrically with respect to the unbiased point estimate $\widetilde{\theta}$. In other words, an interval of the form $(\widetilde{\theta} - \varepsilon, \widetilde{\theta} + \varepsilon)$ is chosen.

Here, the number $\varepsilon > 0$ is such that $|\theta - \widetilde{\theta}| < \varepsilon$ characterizes the estimation accuracy. The higher the estimate is, the smaller the difference $|\theta - \widetilde{\theta}|$.

The value of γ is chosen in advance equal to $0{,}9$; $0{,}95$; $0{,}99$ or $0{,}999$ which depends on the particular problem to be solved.

Thus, let us assume that the probability P, that $|\theta - \widetilde{\theta}| < \varepsilon$ is equal to γ, i.e., the following equation is true: $P(|\theta - \widetilde{\theta}| < \varepsilon) = \gamma$, or the same: $P(\widetilde{\theta} - \varepsilon < \theta < \widetilde{\theta} + \varepsilon) = \gamma$

According to the latter relation, the probability that the interval $(\widetilde{\theta} - \varepsilon, \widetilde{\theta} + \varepsilon)$ covers the unknown parameter θ, is equal to γ.

Since econometric problems often require finding confidence intervals for parameters of a random variable, let us provide examples of their plotting. For the purpose of certainty, we will consider the cases when the sample is taken from the entire assembly with normal distribution.

3.1.1. Plotting the confidence interval to estimate the mathematical expectation at a known variance

Let us assume that there is a random variable X distributed according to the normal law. In this case, the standard deviation σ of this distribution is known, and the value $\gamma -$ is defined.

Let us suppose that a series of particular values x_1, $x_2, \ldots x_n$, is a sample that is obtained by making n independent observations of the value X. In order to emphasize the random nature of the values, let us rewrite them in the form: X_1, $X_2, \ldots X_n$ and therefore, we shall refer to X_i as the value of the random variable X in the i^{th} experiment. Let us note that the random variables X_1, $X_2, \ldots X_n$ are independent and the distribution law of any of them coincides with the distribution law of the random variable X. In this case, the sample mean of $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ will also be distributed ac-

cording to the normal law. The parameters of its distribution are as follows: $M(\overline{X}) = a$; $D(\overline{X}) = \frac{\sigma^2}{n}$.

Now, we require that the following equation is fulfilled: $P(|\overline{X} - a| < \varepsilon) = \gamma$, in which γ is a given reliability.

With the formula $P(|X - a| < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right)$ (in which Φ is the Laplace function), by replacing X with $\overline{X}$ and σ with $\sigma(\overline{X}) = \sqrt{D(\overline{X})} = \frac{\sigma}{\sqrt{n}}$, we can get:

$$P(|\overline{X} - a| < \varepsilon) = 2\Phi\left(\frac{\varepsilon \cdot \sqrt{n}}{\sigma}\right) = 2\Phi(t) = \gamma,$$

In which $t = (\varepsilon\sqrt{n})/\sigma$. By determining $\varepsilon = t\sigma/\sqrt{n}$ from the last equality, we obtain:

$$P(|\overline{X} - a| < t\sigma/\sqrt{n}) = 2\Phi(t) \text{ or}$$

$$P(\overline{X} - t\sigma/\sqrt{n} < a < \overline{X} + t\sigma/\sqrt{n}) = 2\Phi(t) = \gamma$$

According to the confidence interval definition, we obtain the confidence interval for $a = M(X)$:

$$(\overline{X} - t\sigma/\sqrt{n} < a < \overline{X} + t\sigma/\sqrt{n}),$$

in which t is determined from the equation $2\Phi(t) = \gamma$ by Laplace function tables, at a given value γ.

(Laplace function tables can be found at the end of any textbook on probability theory and mathematical statistics.)

Therefore, we can say that the resulting confidence interval covers the unknown parameter a. The accuracy of the estimate: $\varepsilon = t\sigma/\sqrt{n}$.

Example 1: There is a random variable X distributed according to the normal law with a known confidence level $\sigma = 20$. There are five independent observations of this quantity, the results of which are: $x_1 = -25$, $x_2 = 34$, $x_3 = -20$, $x_4 = 10$. $x_5 = 21$. Find an estimate of the mathematical expectation a and plot the 95% - confidence interval for it.

Solution: Let's determine the sample mean value: $\bar{x}_в = \frac{-25+34-20+10+21}{5} = 4$.

Given that $\gamma = 0.95$ and $\Phi(t) = 0.475$, by the table of Laplace functions, we obtain that $t = 1.96$. Then $\varepsilon = \frac{t\sigma}{\sqrt{n}} = \frac{1.96 \cdot 20}{\sqrt{5}} \approx 17.5$. Confidence interval for $a = M(X)$: $(4 - 17.5; 4 + 17.5)$, i.e. $(13.5; 21.5)$.

Example 2. A random variable X has a normal distribution with known standard deviation $\sigma = 3$. Find the confidence intervals for the unknown mathematical expectation a over the sample mean $\bar{x}$, if the sample size $n = 36$ and the reliability of the estimate $\gamma = 0.95$ are given.

Solution: According to the task condition: $2\Phi(t) = 0.95$, therefore $\Phi(t) = 0.475$. According to the table of Laplace functions, we obtain $t = 1.96$. Let us determine the accuracy of the estimate:

$$\varepsilon = t\sigma / \sqrt{n} = (1.96 \cdot 3) / \sqrt{36} = 0.98$$

Thus, the confidence interval: $(\bar{x} - 0.98; \bar{x} + 0.98)$.

For example, if $\bar{x} = 4.1$, the confidence interval has the following confidence boundaries:

$$\bar{x} - 0.98 = 4.1 - 0.98 = 3.12;\ \bar{x} + 0.98 = 4.1 + 0.98 = 5.08$$

This means that the value of the unknown parameter a, which is consistent with the sample data, satisfies the inequality: $3.12 < a < 5.08$.

3.1.2. Plotting the confidence interval to estimate the mathematical expectation for an unknown variance

Let us assume that there is a random variable X distributed according to the normal law. In this case, the standard deviation σ of this distribution is an unknown variable, and the value of γ – is given.

Find such a number ε that the following equation is satisfied

$P(|\overline{X} - a| < \varepsilon) = \gamma$, or $P(\overline{X} - \varepsilon < a < \overline{X} + \varepsilon) = \gamma$.

Let us introduce a random variable: $T = \dfrac{\overline{X} - a}{(S/\sqrt{n})}$ in which n is the sample size; S is the so-called corrected standard deviation, which is calculated from the sample: $S = \sqrt{\dfrac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \overline{X}_i)^2}$

The value T has a Student's distribution with $(n-1)$ degrees of freedom. The density of this distribution is:

$f(t, n-1) = \dfrac{\Gamma(n/2)}{\sqrt{\pi(n-1)} \cdot \Gamma\left(\dfrac{n-1}{2}\right)} \cdot \left(1 + \dfrac{t^2}{n-1}\right)^{-\frac{n}{2}}$, in which Γ – is the gamma function (taken from the reference tables).

In the left side of the equality $P(|\overline{X} - a| < \varepsilon) = \gamma$, we change from a random variable $\overline{X}$ to a random variable T:

$$P\left(\frac{|\overline{X} - a|}{\dfrac{S}{\sqrt{n}}} < \frac{\varepsilon}{\dfrac{S}{\sqrt{n}}}\right) = \gamma \quad \text{or} \quad P(|T| < \varepsilon/(S/\sqrt{n})) = \gamma \text{ or } P(|T| < t_\gamma) = \gamma,$$

In which $t_\gamma = \dfrac{\varepsilon \cdot \sqrt{n}}{S}$.

The value of t_γ is found from the equality: $2\int_0^{t_\gamma} f(t, n-1)dt = \gamma$.

Using the Student's distribution table, we find the value of t_γ depending on the confidence probability γ and the number of degrees of freedom $(n-1)$. Then, by determining the value t_γ from the equality $t_\gamma = (\varepsilon \cdot \sqrt{n})/S$, we find the value $\varepsilon = t_\gamma \cdot \dfrac{S}{\sqrt{n}}$.

Consequently, the equality $P(|\overline{X} - a| < \varepsilon) = \gamma$ takes the form:

$$P\left(\overline{X} - t_\gamma \cdot \frac{S}{\sqrt{n}} < a < \overline{X} + t_\gamma \cdot \frac{S}{\sqrt{n}}\right) = \gamma$$

This means that the interval $\left(\overline{X} - t_\gamma \cdot \dfrac{S}{\sqrt{n}}; \overline{X} + t_\gamma \cdot \dfrac{S}{\sqrt{n}}\right)$ covers the value of $a = M(X)$ with probability γ, i.e., it is the confidence interval for the mathematical expectation of the considered random variable X.

Example 3. According to the condition of Example 1, assuming that the random variable X is normally distributed with an unknown confidence probability σ, plot a confidence interval for the unknown $a = M(X)$. Consider that $\gamma = 0.95$.

Solution: According to the above estimation, $\bar{x} = 4$. We find the value of S:

$$S^2 = \frac{1}{4}((-25-4)^2 \cdot 1 + (34-4)^2 + (-20-4)^2 + (10-4)^2 + (21-4)^2) = 660.5$$

$S \approx 25.7$. According to the table for $\gamma = 0.95$ and $n-1=4$, we find $t_\gamma = 2.78$. Consequently, $\varepsilon = 2.78 \cdot \frac{25.7}{2.24} \approx 31.9$ and the confidence interval is: $(-27{,}9;35{,}9)$.

3.1.2.1. Confidence interval for the standard deviation of the normal distribution

Let us assume that there is a random variable X distributed according to the normal law. In this case, σ is an unknown variable, and the value of γ is specified. It is found that if $a = M(X)$ is known, then the confidence interval for the standard deviation σ is as follows: $\left(\frac{\sqrt{n}\cdot S_0}{\chi_2};\frac{\sqrt{n}\cdot S_0}{\chi_1}\right)$, in which n is the sample size; $S_0^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - a)^2$,

and $\chi^2_{\frac{1+\gamma}{2};n}$; $\chi^2_{\frac{1-\gamma}{2};n}$ are the so-called quantiles of the χ^2 distribution with n degrees of freedom. They are determined by a special table of quantiles $\chi^2_{\alpha,n}$ of χ^2_n distribution.

If the value of $a = M(X)$ is not known, the confidence interval for the unknown x has the following form:

$\left(\frac{\sqrt{n-1}\cdot S}{\chi_2};\frac{\sqrt{n-1}\cdot S_0}{\chi_1}\right)$, in which n is the sample volume, $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$ - corrected standard deviation; $\chi_1^2 = \chi^2_{\frac{1+\gamma}{2};n-1}$ and $\chi_2^2 = \chi^2_{\frac{1-\gamma}{2};n-1}$ - quantiles determined from the table $\chi^2_{\alpha,k}$ at $k = n-1$, as well as, respectively, at $\alpha = \frac{1+\gamma}{2}$ and $\alpha = \frac{1-\gamma}{2}$.

Example 4. In order to estimate the parameter of a normally distributed random variable, a sample volume of 30 was taken and $S = 1{,}5$ was calculated. Find the confidence interval covering σ with probability $\gamma = 0.90$.

Solution: According to the condition of the task, we have: $n = 30$, $\gamma = 0.9$. Then, according to the table $\chi^2_{\alpha,k}$ we find

$$\chi_1^2 = \chi^2_{\frac{1+0.9}{2};30-1} = \chi^2(0.95;29) = 17.7\,;\ \chi_2^2 = \chi^2_{\frac{1-0.9}{2};30-1} = \chi^2(0.05;29) = 42.6\,.$$

Consequently, the confidence interval has the following form:

$\left(\frac{\sqrt{30-1}\cdot 1.5}{\sqrt{42.6}};\frac{\sqrt{30-1}\cdot 1.5}{\sqrt{17.7}}\right)$, or $1.238 < \sigma < 1.920$.

Раздел 4.

4.1. Статистические гипотезы и их проверка. Основные понятия

Статистической гипотезой принято называть любое предположение о виде или параметрах неизвестного закона распределения.

Та гипотеза, которая однозначно определяет закон распределения, называется *простой,* в противном случае – сложной.

Нулевой (H_0) называют выдвинутую гипотезу, которую необходимо проверить. Гипотезу (H_1) противоположную нулевой, называют *конкурирующей* (или *альтернативной*).

Статистическим критерием принято называть однозначно определенное правило, устанавливающее условия, при которых проверяемую гипотезу H_0 следует либо принять, либо отвергнуть.

Основу критерия представляет специально составленная характеристика выборки (статистика) $\tilde{\theta}_n = f(x_1, x_2, ..., x_n)$, точное или приближенное распределение которой известно.

Пусть имеется выборка объемом n: $X_1, X_2, ..., X_n$. Каждый критерий разбивает все множество возможных значений статистики на две непересекающиеся подмножества: так называемую критическую область (область отклонения гипотезы) и область принятия гипотезы.

Основной принцип проверки гипотезы состоит в том, что если наблюдаемые значения статистики критерия попадают в критическую область, то гипотезу отвергают. В противном случае – принимают. Этот принцип не дает логического доказательства или опровержения гипотезы. При его использовании возможны четыре случая:

– гипотеза H_0 верна и ее принимают согласно критерию;

– гипотеза H_0 не верна и ее отвергают согласно критерию;

– гипотеза H_0 верна, но ее отвергают согласно критерию (ошибка 1-го рода);

– гипотеза H_0 не верна, но ее принимают согласно критерию (ошибка 2-го рода).

Уровнем значимости α называют вероятность совершить ошибку первого рода (т.е. отвергнуть нулевую гипотезу H_0, когда на самом деле, она верна). С уменьшением α возрастает вероятность ошибки β второго рода (принять H_0,когда она не верна).

Мощностью критерия $(1-\beta)$ принято называть вероятность того, что нулевая гипотеза H_0 будет отвергнута, если верна конкурирующая H_1. Т.е. не допустить ошибку второго рода.

Обозначим через $P(\widetilde{\theta}_n \in W / H)$ – вероятность попадания статистики критерия $\widetilde{\theta}_n$ в критическую область W, если верна гипотеза H. Тогда требования к критической области аналитически можно записать в следующем виде:

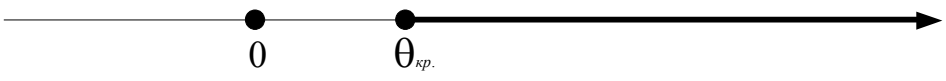
$$\begin{cases} P(\widetilde{\theta}_n \in W / H_0) = \alpha, \\ P(\widetilde{\theta}_n \in W / H_1) = 1 - \beta = \max, \end{cases}$$

где второе условие выражает требование максимума мощности критерия.

Согласно последнему условию, критическая область выбирается так, чтобы вероятность попадания в нее была минимальной (равной α), если верна нулевая гипотеза H_0, а в противоположном случае – максимальной.

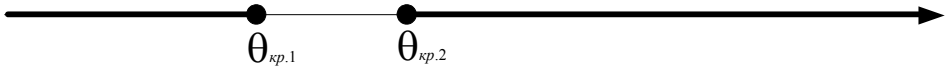
В зависимости от вида конкурирующей гипотезы H_1 выбирают

правостороннюю ($\widetilde{\theta}_n > \theta_{кр}$):



левостороннюю ($\widetilde{\theta}_n < \theta_{кр}$):

и двустороннюю ($\widetilde{\theta}_n < \theta_{кр.1}$ и $\widetilde{\theta}_n > \theta_{кр.2}$)



критические области.

При заданном уровне значимости α, границы критических областей определяют из следующих соотношений:

– для правосторонней критической области

$$P(\widetilde{\theta}_n > \theta_{кр}) = \alpha$$

– для левосторонней критической области

$$P(\widetilde{\theta}_n < \theta_{кр}) = \alpha$$

– для двусторонней критической области

$$P(\widetilde{\theta}_n > \theta_{кр.2}) = \alpha / 2,$$

$$P(\widetilde{\theta}_n < \theta_{кр.1}) = \alpha / 2, \text{ где } \theta_{кр.1} < \theta_{кр.2}.$$

Гипотезы принято различать на *параметрические* (о параметрах распределения известного вида) и *непараметрические* (о виде закона неизвестного распределения). Рассмотрим более конкретные примеры .

4.2. Проверка гипотезы о числовом значении дисперсии генеральной совокупности.

Предположим, что имеется генеральная совокупность и значение некоторого ее признака X является случайной величиной имеющей нормальное распределение $N(a,\sigma)$ с неизвестной дисперсией σ^2. Предположим также, что из

этой совокупности взяли случайную выборку объемом n. При этом, дисперсия S^2 определенная по выборке (исправленная выборочная дисперсия $S^2 = \frac{\sum_{i=1}^{m}(X_i - \overline{X}_в)}{(n-1)}$) дает только приближенное представление о σ^2.

Требуется проверить нулевую гипотезу H_0: $\sigma^2 = \sigma_0^2$, где σ_0^2 – определенное заданное значение дисперсии.

В качестве критерия, при оценке нулевой гипотезы H_0, используем следующую выборочную характеристику: $\widetilde{\theta}_n = \frac{nS^2}{\sigma_2}$, которая, при выполнении гипотезы H_0, имеет распределение χ^2 с $(n-1)$ степенями свободы, т.е. $\widetilde{\theta}_n = \frac{nS^2}{\sigma_2} = \chi^2$.

В зависимости от конкурирующей гипотезы, следует выбирать правостороннюю, левостороннюю или двустороннюю критические области. Границы критической области ($\chi^2_{кр}$) определяют по таблице распределения χ^2 для заданного уровня значимости α и числа степеней свободы $\nu = n-1$.

Для дальнейших рассуждений, необходимо знать уровень значимости α (вероятность ошибки первого рода). Эту величину обычно задают заранее, используя стандартные значения: $\alpha = 0{,}05$; $\alpha = 0{,}01$; $\alpha = 0{,}005$; $\alpha = 0{,}0)1$. (Выбор конкретной величины зависит от специфики решаемой задачи !).

Итак, зададимся уровнем значимости α и перейдем к построению критических областей для проверки гипотезы H_0 при следующих трех альтернативных гипотезах H_1:

1). H_1: $\sigma^2 = \sigma_1^2 > \sigma_0^2$. В этом случае, выбирают правостороннюю критическую область и $\chi^2_{кр}$ находят из условия

$$P(\chi^2 > \chi^2_{кр}(\alpha, n-1)) = \alpha,$$

где $\chi^2_{кр}(\alpha, n-1)$ – табличное значение χ^2, найденное для уровня значимости α и числа степеней свободы $(n-1)$.

Правило проверки гипотезы следующее: если $\chi^2_{набл}$ (наблюдаемое значение) оказывается больше чем $\chi^2_{кр}$, т.е. $\chi^2_{набл} > \chi^2_{кр}$, то нулевую гипотезу $H_0: \sigma^2 = \sigma_0^2$ отвергают; если же $\chi^2_{набл} \le \chi^2_{кр}$, то считается, что нулевая гипотеза не противоречит опытным данным.

При этом, для вычисления мощности критерия можно воспользоваться формулой

$$1-\beta = P(\chi^2 > \frac{\sigma_0^2}{\sigma_1^2}\chi^2_{кр}(\alpha, n-1))$$

2). При конкурирующей гипотезе $H_1: \sigma^2 = \sigma_1^2 < \sigma_0^2$ строят левостороннюю критическую область. Границу критической области определяют по таблице распределения χ^2 из условия

$$1-\alpha = P(\chi^2 > \chi^2_{кр}(1-\alpha; n-1))$$

Если $\chi^2_{набл} < \chi^2_{кр}(1-\alpha, n-1)$, то гипотеза $H_0: \sigma^2 = \sigma_0^2$ отвергается, если же $\chi^2_{набл} \ge \chi^2_{кр}(1-\alpha, n-1)$, то гипотеза H_0 не отвергается.

Для вычисления мощности критерия можно воспользоваться формулой

$$1-\beta = P(\chi^2 > \frac{\sigma_0^2}{\sigma_1^2}\chi^2_{кр}(1-\alpha, n-1)).$$

3). Если конкурирующая гипотеза $H_1: \sigma^2 = \sigma_1^2 \ne \sigma_0^2$, то строят двустороннюю критическую область. При этом, левую ($\chi^2_{кр.лев.}$) и правую ($\chi^2_{кр.прав.}$) границы критической области находят из условий:

$$1-(\alpha/2)=P(\chi^2>\chi^2_{кр.лев.}(1-\frac{\alpha}{2};n-1))$$

$$\alpha/2=P(\chi^2>\chi^2_{кр.прав.}(\frac{\alpha}{2};n-1)).$$

В этом случае, правило проверки гипотезы сводится к следующему: если $\chi^2_{кр.прав.}\geq\chi^2_{набл.}\geq\chi^2_{кр.лев.}$, то нет оснований опровергать гипотезу. Если же $\chi^2_{набл.}<\chi^2_{кр.лев.}$ или $\chi^2_{набл.}>\chi^2_{кр.прав.}$, то гипотеза отвергается.

Пример 1: Точность работы станка-автомата проверяется по дисперсии контролируемого размера детали. По выборке из 25 деталей вычислена величина $S^2=0{,}25$. Требуется проверить гипотезу H_0: $\sigma^2=0{,}15$ при уровне значимости $\alpha=0{,}05$.

Решение: За альтернативную примем гипотезу H_1: $\sigma^2>0{,}15$, т.е. имеем случай 1. По соответствующим таблицам для критерия Пирсона, находим правую границу интервала $\chi^2(\alpha,m-1)=\chi^2(0{,}05;24)=36{,}4$. Следовательно, критическая область $(36{,}4;\infty)$. По формуле $\widetilde{\theta}_n=\frac{nS^2}{\sigma_2}=\chi^2$ определяем так называемое наблюдаемое значение критерия $\chi^2_{набл.}=\frac{25\cdot0{,}25}{0{,}15}\approx42$.

Поскольку $\chi^2_{набл.}$ попадает в критическую область, то гипотезу H_0 отвергаем.

4.3. Проверка гипотезы о законе распределения.

Предположим, что нужно проверить гипотезу H_0 о том, что случайная величина X подчиняется определенному закону распределения, заданному функцией распределения $F_0(x)$, т.е. $F_X(x)=F_0(x)$. Под альтернативной гипотезой будем понимать то, что просто не выполняется основная т.е. H_1: $F_X(x)\neq F_0(x)$. Для проверки гипотезы о распределении рассматриваемой случайной величины, проводят выборку объемом n, которую оформляют в виде статистического ряда. Для того, чтобы сделать заключение о том согласуются

ли результаты наблюдений с высказанным предположением, используют специальную величину получившую название *критерий согласия*.

Критерием согласия называют статистический критерий проверки гипотезы о предполагаемом законе неизвестного распределения. Существуют различные критерии согласия, среди которых наиболее часто употребляемым является критерий Пирсона.

Для проверки гипотезы H_0 всю область значений X разбивают на m интервалов: $\Delta_1, \Delta_2, \ldots, \Delta_m$ и подсчитывают вероятности p_i $(i = 1,2,....,m)$ попадания случайной величины X в интервал Δ_i. С этой целью используют формулу: $P(\alpha \le X \le \beta) = F_0(\beta) - F_0(\alpha)$. При этом, теоретическое число значений случайной величины X, попавших в интервал Δ_i, можно подсчитать по формуле $n \cdot p_i$. Таким образом, кроме статистического ряда распределения величины X, который был получен в результате проведенной выборки

x_i	x_1	x_2	…..	x_m
n_i	n_1	n_i	…..	n_m

где $\sum_{i=1}^{m} n_i = n$,

получаем еще и теоретический ряд распределения:

Δ_1	Δ_2	Δ_3	…..	Δ_m
$n_1' = np_1$	$n_2' = np_2$	$n_3' = np_3$	…..	$n_m' = np_m$

В том случае, если так называемые эмпирические частоты n_i сильно отличаются от теоретических $(n_i' = np_i)$, то проверяемую гипотезу H_0 отвергают, в противном случае – принимают. В качестве меры расхождения между этими частотами используют критерий (статистику) Пирсона

$$\chi^2 = \sum_{i=1}^{m} \frac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^{m} \frac{n_i^2}{np_i} - n\,.$$

Эта статистика, при $n \to \infty$, имеет χ^2 распределение с $k = m - r - 1$ степенями свободы, где n – число групп (интервалов) выборки, r – число параметров предполагаемого распределения. (В частном случае, если предполагаемое распределение нормально, то следует оценивать два параметра (α и σ). Поэтому: $k = m - 3$).

Замечание: Необходимым условием применения критерия Пирсона является наличие в каждом интервале не менее пяти значений (т.е. $n_i \geq 5$). Если в отдельных интервалах их оказывается меньше, то число интервалов следует уменьшать путем объединения соседних.

Правило применения χ^2:

1). Вычисляют выборочное (наблюдаемое) значение $\chi^2_{набл}$ статистики критерия;

2). Выбирают уровень значимости α критерия и по таблице χ^2 - распределения находят критическую точку $\chi^2_{\alpha,k}$.

3). Если $\chi^2_{набл} > \chi^2_{\alpha,k}$, то гипотеза H_0 отвергается. Если $\chi^2_{набл} \leq \chi^2_{\alpha,k}$, то считается что гипотеза H_0 не противоречит опытным данным.

Пример 2: Измерены 100 обработанных деталей. Отклонения от заданного размера приведены в таблице:

$[x_i, x_{i+1})$	[-3,-2)	[-2,-1)	[-1,0)	[0,1)	[1,2)	[2,3)	[3,4)	[4,5)
n_i	3	10	15	24	25	13	7	3

Требуется, при заданном уровне значимости $\alpha = 0{,}01$, гипотезу H_0 о том, что отклонения от проектного размера подчиняются нормальному закону распределения.

Решение: Поскольку число наблюдений в крайних интервалах меньше пяти, то их следует объединить с соседними. В результате, получается следующий ряд распределения:

$[x_i, x_{i+1}$	[-3,-1)	[-1,0)	[0,1)	[1,2)	[2,3)	[3,5)
n_i	13	15	24	25	13	10

Отклонение размера, которое представляет собой случайную величину, обозначим через X. Для определения вероятностей p_i, вычислим параметры определяющие нормальный закон распределения (α и σ).

Согласно ранее сделанной выборке:

$$\bar{x} = \frac{1}{100}(-2 \cdot 13 + (-0{,}5) \cdot 15 + ... + 4 \cdot 10) = 0{,}855 \approx 0{,}9,$$

$$D_в = \frac{1}{100}(4 \cdot 13 + 0б25 \cdot 15 + ... + 16 \cdot 10) - (0{,}855)^2 \approx 2{,}809,\ \sigma \approx 1{,}676 \approx 1{,}7.$$

Поскольку случайная величина X имеет нормальное распределение и определена на всей числовой оси , т.е. на интервале $(-\infty, \infty)$, то крайние интервалы в ряде распределения заменяем, соответственно, на $(-\infty, -1)$ и $(3, +\infty)$. В этом случае

$p_1 = P(-\infty < X < -1) = \Phi_0\left(\frac{-1-0{,}9}{1{,}7}\right) - \Phi_0(-\infty) = \frac{1}{2} - \Phi_0(1{,}12) = 0{,}1314$. Аналогично получаем: $p_2 = 0{,}1667$, $p_3 = 0{,}2258$, $p_4 = 0{,}2183$, $p_5 = 0{,}1503$,

$p_6 = P(3 \le X < \infty) = \Phi_0(\infty) - \Phi_0\left(\frac{3-0{,}9}{1{,}7}\right) = 0{,}5 - \Phi_0(1{,}24) = 0{,}1075$.

Полученные результаты представим в виде таблицы:

$[x_i, x_{i+1})$	$[-\infty,-1)$	$[-1,0)$	$[0,1)$	$[1,2)$	$[2,3)$	$[3, -\infty)$
n_i	13	15	24	25	13	10
$n' = np_i$	13,14	16,67	22,58	21,83	15,03	10,75

Далее, вычисляем $\chi^2_{набл}$:

$$\chi^2_{набл} = \sum_{i=1}^{6} \frac{n_i^2}{np_i} - n = \left(\frac{13^2}{13,4} + \frac{15^2}{16,67} + ... + \frac{10^2}{10,75} \right) - 100 = 101,045 - 100 \approx 1,045$$

Поскольку по выборке рассчитаны два параметра, то $r = 2$. Количество интервалов 6, т.е. $m = 6$. Следовательно $k = 6 - 2 - 1 = 3$. Зная что $\alpha = 0,01$ и $k = 3$, по таблице χ^2 - распределения определяем $\chi^2_{\alpha,k} = 11,3$.

Поскольку $\chi^2_{набл} < \chi^2_{\alpha,k}$, следовательно, отвергать проверяемую гипотезу нет оснований.

Section 4.

4.1. Statistical hypotheses and their testing. Basic concepts

Any assumption about the type or parameters of an unknown distribution law is called a *statistical hypothesis*.

If a hypothesis unambiguously determines the distribution law, it is called a *simple* hypothesis, otherwise it is called a *complex* hypothesis.

The null hypothesis (H_0) is a hypothesis that is to be tested. Any hypothesis (H_1) that is opposite to the null hypothesis is called a *competing* one (or an *alternative* hypothesis).

The statistical criterion is a uniquely defined rule that establishes the conditions under which the hypothesis H_0 to be tested should be either accepted or rejected.

A specially compiled characteristic of a sample (statistic) $\widetilde{\theta}_n = f(x_1, x_2, ..., x_n)$, whose exact or approximate distribution is known, is the basis of the criterion.

Let us assume that there is a sample with the size n: $X_1, X_2, ..., X_n$. Each criterion divides the entire set of possible statistic values into two non-intersecting subsets: the so-called critical area ("area of hypothesis rejection") and the area of hypothesis acceptance.

The basic principle of hypothesis testing assumes the hypothesis rejection if the observed values of the criterion statistics fall into the critical area. Otherwise, it is accepted. This principle does not provide a logical proof or disproof of the hypothesis. When using it, four cases are possible:

- the hypothesis H_0 is true and it is accepted according to the criterion;

- the hypothesis H_0 is not true and it is rejected according to the criterion;

- the hypothesis H_0 is true, but it is rejected according to the criterion (error of the 1st kind);

- the hypothesis H_0 is not true, but it is accepted according to the criterion (error of the 2nd kind).

The probability of making an error of the first kind (i.e., rejecting a null hypothesis H_0 when in fact it is true) is called a *significance level* α. With decrease of α, the probability of making the error β of the second kind (to accept H_0 when it is not true) increases.

The probability that the null hypothesis H_0 will be rejected if the competing hypothesis H_1 is true is called the *power of the criterion* $(1-\beta)$. In other words, it is necessary to prevent an error of the second kind.

Let us define $P(\tilde{\theta}_n \in W / H)$ as the probability of the criterion statistics $\tilde{\theta}_n$ falling into the critical area W, if hypothesis H is true. Then the requirements to the critical area can be written analytically in the following form:

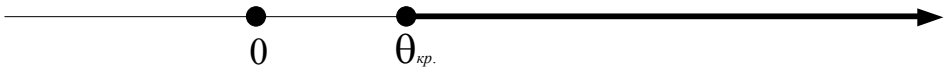
$$\begin{cases} P(\widetilde{\theta}_n \in W / H_0) = \alpha, \\ P(\widetilde{\theta}_n \in W / H_1) = 1 - \beta = \max, \end{cases}$$

in which the second condition expresses the requirement of the maximum criterion power.

According to the last condition, the critical area is chosen so that the probability of falling into it is minimal (equal to α), if the null hypothesis H_0 is true, and in the opposite case - the probability is maximal.

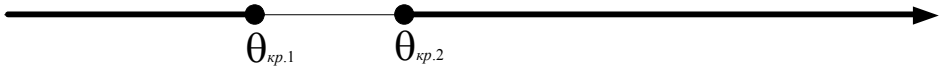
Depending on the type of competing hypothesis H_1 one chooses

right-handed ($\widetilde{\theta}_n > \theta_{кр}$):



left-handed ($\widetilde{\theta}_n < \theta_{кр}$):

and bidirectional ($\widetilde{\theta}_n < \theta_{кр.1}$ и $\widetilde{\theta}_n > \theta_{кр.2}$)



critical areas.

At a given significance level α, the boundaries of the critical areas are determined from the following relations

– for the right-hand critical area

$$P(\widetilde{\theta}_n > \theta_{кр}) = \alpha$$

– for the left-hand critical area

$$P(\widetilde{\theta}_n < \theta_{кр}) = \alpha$$

– for the bidirectional critical area

$$P(\widetilde{\theta}_n > \theta_{кр.2}) = \alpha/2,$$

$$P(\widetilde{\theta}_n < \theta_{кр.1}) = \alpha/2, \text{ in which } \theta_{кр.1} < \theta_{кр.2}.$$

Generally, hypotheses are classified into *parametric* (about the parameters of a known kind of distribution) and *nonparametric* (about the kind of law of an unknown distribution). Let us consider more specific examples.

4.2. Testing the hypothesis about the numerical value of the entire assembly variance.

Let us assume that there is an entire assembly and the value of some attribute X is a random variable that has a normal distribution $N(a,\sigma)$ with an unknown variance σ^2. Let us also assume that a random sample of volume n was taken from this assembly. In this case, the variance of S^2 determined from the sample (the corrected sample variance of $S^2 = \dfrac{\sum_{i=1}^{m}(X_i - \overline{X}_в)}{(n-1)}$) gives only an approximate representation of σ^2.

We need to test the null hypothesis H_0: $\sigma^2 = \sigma_0^2$, where σ_0^2 is a certain given value of the variance.

When evaluating the null hypothesis H_0, we use as a criterion the following sample characteristic: $\widetilde{\theta}_n = \dfrac{nS^2}{\sigma_2}$, which, if hypothesis H_0 is satisfied, has a distribution χ^2 with $(n-1)$ degrees of freedom, i.e. $\widetilde{\theta}_n = \dfrac{nS^2}{\sigma_2} = \chi^2$.

Depending on the competing hypothesis, the right-hand, left-hand, or bidirectional critical areas should be chosen. The boundaries of the critical area ($\chi^2_{кр}$) are deter-

mined by χ^2 *distribution table for a given level of significance x and number of degrees of freedom* $\nu = n - 1$.

In order to do further reasoning, it is necessary to know the significance level x (probability of error of the first kind). This value is usually set in advance, using standard values: $\alpha = 0.05$; $\alpha = 0.01$; $\alpha = 0.005$; $\alpha = 0.0)1$. (The choice of a particular value depends on the specifics of the problem to be solved!).

So we set the significance level H_0 and proceed to construct critical areas for testing hypothesis x under the three following alternative hypotheses H_1:

1). H_1: $\sigma^2 = \sigma_1^2 > \sigma_0^2$. In this case, the right-hand critical area $\chi^2_{кр}$ is chosen and found from the condition

$$P(\chi^2 > \chi^2_{кр}(\alpha, n-1)) = \alpha,$$

In which $\chi^2_{кр}(\alpha, n-1)$ – - is the table value, found for the significance level α and the number of degrees of freedom ($n - 1$).

The rule for hypothesis testing is as follows: if $\chi^2_{набл}$ (observed value) is greater than $\chi^2_{кр}$, i.e. $\chi^2_{набл} > \chi^2_{кр}$, then the null hypothesis H_0: $\sigma^2 = \sigma_0^2$ is rejected; if $\chi^2_{набл} \le \chi^2_{кр}$, then it is considered that the null hypothesis does not contradict the experimental data.

To calculate the power of the criterion, you can use the following formula

$$1 - \beta = P(\chi^2 > \frac{\sigma_0^2}{\sigma_1^2}\chi^2_{кр}(\alpha, n-1))$$

2). In the competing hypothesis H_1: $\sigma^2 = \sigma_1^2 < \sigma_0^2$ the left-hand critical area is plotted. The boundary of the critical area is determined by the χ^2 distribution table from the condition

$$1-\alpha = P(\chi^2 > \chi^2_{кр}(1-\alpha; n-1))$$

If $\chi^2_{набл} < \chi^2_{кр}(1-\alpha, n-1)$, then hypothesis H_0: $\sigma^2 = \sigma_0^2$ is rejected, but if $\chi^2_{набл} \geq \chi^2_{кр}(1-\alpha, n-1)$, then hypothesis H_0 is not rejected.

The following formula can be used to calculate the power of the criterion

$$1-\beta = P(\chi^2 > \frac{\sigma_0^2}{\sigma_1^2}\chi^2_{кр}(1-\alpha, n-1)).$$

3). If the competing hypothesis H_1: $\sigma^2 = \sigma_1^2 \neq \sigma_0^2$, then a bilateral critical area is plotted. In this case, the left ($\chi^2_{кр.лев.}$) and right ($\chi^2_{кр.прав.}$) boundaries of the critical area are found from the conditions:

$$1-(\alpha/2) = P(\chi^2 > \chi^2_{кр.лев.}(1-\frac{\alpha}{2}; n-1))$$

$$\alpha/2 = P(\chi^2 > \chi^2_{кр.прав.}(\frac{\alpha}{2}; n-1)).$$

In this case, the rule of hypothesis testing is as follows: if $\chi^2_{кр.прав.} \geq \chi^2_{набл.} \geq \chi^2_{кр.лев.}$, there is no reason to disprove the hypothesis. If $\chi^2_{набл.} < \chi^2_{кр.лев.}$ or $\chi^2_{набл.} > \chi^2_{кр.прав.}$, then the hypothesis is rejected

Example 1: The accuracy of the automatic machine is checked by the variance of the controlled part size. The value $S^2 = 0.25$ is calculated on a sample of 25 parts. We need to test the hypothesis H_0: $\sigma^2 = 0.15$ at the significance level of $\alpha = 0{,}05$.

Solution: We consider H_1: $\sigma^2 > 0.15$ as an alternative hypothesis, i.e. we have case 1. According to the corresponding tables for the Pearson's criterion, we find the right boundary of the interval $\chi^2(\alpha, m-1) = \chi^2(0.05; 24) = 36.4$. Hence, the critical area

$(36,4;\infty)$. Using the formula $\widetilde{\theta}_n = \frac{nS^2}{\sigma_2} = \chi^2$, we determine the so-called observed value of the criterion $\chi^2_{набл.} = \frac{25 \cdot 0.25}{0.15} \approx 42$.

Since $\chi^2_{набл}$ is within the critical area, we reject the hypothesis H_0.

4.3. Testing the hypothesis about the distribution law.

Let us assume that we need to test the hypothesis H_0 that a random variable X follows a certain distribution law defined by the distribution function $F_0(x)$, i.e. $F_X(x) = F_0(x)$. We will define the alternative hypothesis as something that is simply not satisfied by the main hypothesis, i.e. $H_1: F_X(x) \neq F_0(x)$. To test the hypothesis about the distribution of the random variable in question, we conduct a sample of n, which is drawn in the form of a statistical series. In order to conclude whether the results of observations are consistent with the stated hypothesis, a special value called the *goodness-of-fit criterion* is used.

A statistical criterion for testing a hypothesis about the assumed law of an unknown distribution is called a *goodness-of-fit criterion*. There are various goodness of fit criteria, among which the Pearson's criterion is the most commonly used.

For hypothesis H_0 testing, the entire range of X values is divided into m intervals: $\Delta_1, \Delta_2, \ldots, \Delta_m$, then we calculate the probabilities p_i $(i = 1,2,....,m)$ of a random variable X falling into the interval Δ_i. In order to do this, the following formula is used: $P(\alpha \le X \le \beta) = F_0(\beta) - F_0(\alpha)$. Thus, the theoretical number of values of a random variable X which fall into the interval Δ_i can be calculated using the formula $n \cdot p_i$. Consequently, in addition to the statistical series of the X value distribution, which was obtained as a result of sampling

x_i	x_1	x_2	…..	x_m
n_i	n_1	n_i	…..	n_m

In which $\sum_{i=1}^{m} n_i = n$,

we also get a theoretical distribution series:

Δ_1	Δ_2	Δ_3	…..	Δ_m
$n_1' = np_1$	$n_2' = np_2$	$n_3' = np_3$	…..	$n_m' = np_m$

Should the so-called empirical frequencies n_i differ significantly from the theoretical frequencies ($n_i' = np_i$), then the hypothesis H_0 is rejected, while otherwise it is accepted. As a measure of deviation between these frequencies, Pearson's criterion (statistic) is used

$$\chi^2 = \sum_{i=1}^{m} \frac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^{m} \frac{n_i^2}{np_i} - n.$$

Given $n \to \infty$, this statistic has an χ^2 distribution with $k = m - r - 1$ degrees of freedom, where n is the number of sample groups (intervals), r is the number of parameters of the assumed distribution. (In the special case, if the assumed distribution is normal, then two parameters (α and σ) shall be estimated. Therefore: $k = m - 3$).

Note: A prerequisite for applying Pearson's criterion is to have at least five values (i.e. $n_i \geq 5$) in each interval. If there are fewer values in individual intervals, the number of intervals should be reduced by combining neighboring intervals.

Application rule χ^2:

1). The sample (observed) value $\chi^2_{набл}$ of the criterion statistic is calculated;

2). The significance level α of the criterion is chosen and the critical point $\chi^2_{\alpha,k}$ is found according to the table of χ^2-distribution.

3). If $\chi^2_{набл} > \chi^2_{\alpha,k}$, then H_0 hypothesis is rejected. If $\chi^2_{набл} \le \chi^2_{\alpha,k}$, it is considered that H_0 hypothesis does not contradict the experimental data.

Example 2: 100 machined parts were measured. The deviations from the specified size are shown in the table:

$[x_i, x_{i+1})$	[-3,-2)	[-2,-1)	[-1,0)	[0,1)	[1,2)	[2,3)	[3,4)	[4,5)
n_i	3	10	15	24	25	13	7	3

At a given significance level $\alpha = 0.01$, it is required to hypothesize H_0 that the deviations from the designed size follow the normal law of distribution.

Solution: Since the number of observations in the extreme intervals is less than five, they should be combined with the neighboring ones. As a result, the following distribution series is obtained:

$[x_i, x_{i+1}$	[-3,-1)	[-1,0)	[0,1)	[1,2)	[2,3)	[3,5)
n_i	13	15	24	25	13	10

We denote the size deviation, which is a random variable, by X. In order to determine the probabilities p_i, let us calculate the parameters defining the normal law of distribution (α and σ).

According to the previously made sample:

$$\bar{x} = \frac{1}{100}(-2 \cdot 13 + (-0.5) \cdot 15 + ... + 4 \cdot 10) = 0.855 \approx 0.9,$$

$$D_в = \frac{1}{100}(4 \cdot 13 + 0.25 \cdot 15 + ... + 16 \cdot 10) - (0.855)^2 \approx 2.809,\ \sigma \approx 1.676 \approx 1.7.$$

Considering that the random variable X has a normal distribution and is defined on the entire numerical axis, i.e. on the interval $(-\infty,\infty)$, we replace the extreme intervals in the distribution series with $(-\infty,-1)$ and $(3,+\infty)$, respectively. In this case

$$p_1 = P(-\infty < X < -1) = \Phi_0\left(\frac{-1-0.9}{1.7}\right) - \Phi_0(-\infty) = \frac{1}{2} - \Phi_0(1.12) = 0.1314.$$

Similarly, we get: $p_2 = 0.1667$, $p_3 = 0.2258$, $p_4 = 0.2183$, $p_5 = 0.1503$,

$$p_6 = P(3 \le X < \infty) = \Phi_0(\infty) - \Phi_0\left(\frac{3-0.9}{1.7}\right) = 0.5 - \Phi_0(1.24) = 0.1075.$$

The results are presented in the form of a table:

$[x_i, x_{i+1})$	$[-\infty,-1)$	$[-1,0)$	$[0,1)$	$[1,2)$	$[2,3)$	$[3, -\infty)$
n_i	13	15	24	25	13	10
$n' = np_i$	13.14	16.67	22.58	21.83	15.03	10.75

Then, we calculate $\chi^2_{набл}$:

$$\chi^2_{набл} = \sum_{i=1}^{6} \frac{n_i^2}{np_i} - n = \left(\frac{13^2}{13.4} + \frac{15^2}{16.67} + \ldots\ldots + \frac{10^2}{10.75}\right) - 100 = 101.045 - 100 \approx 1.045$$

As there are two parameters calculated on the sample, then $r = 2$. The number of intervals is 6, i.e. $m = 6$. Consequently, $k = 6 - 2 - 1 = 3$. Knowing that $\alpha = 0.01$ and $k = 3$, we use the table of χ^2 - distributions to determine $\chi^2_{\alpha,k} = 11.3$.

Since $\chi^2_{набл} < \chi^2_{\alpha,k}$, therefore there is no reason to reject the tested hypothesis.

Информационное обеспечение обучения
(Informationsupportoftraining)

Основнаялитература:
(Main literature):

1. *Кремер, Н. Ш.* Теория вероятностей и математическая статистика: учебник и практикум для академического бакалавриата / Н. Ш. Кремер. — М. : Издательство Юрайт, 2019. — 539 с. https://www.biblio-online.ru/book/teoriya-veroyatnostey-i-matematicheskaya-statistika-431167
https://edu-lib.com/izbrannoe/kremer-n-sh-teoriya-veroyatnostey-i-matematicheskaya-statistika-onlayn
2. *Гмурман, В. Е.*Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие для прикладного бакалавриата /— 11-е изд., перераб. и доп. — М. : Издательство Юрайт, 2017. — 404 с. https://www.biblio-online.ru/book/F6DC17CF-66E8-400F-9CDA-8067F86D996A
http://padabum.com/d.php?id=10680
3. *Tobodga M.* Lecture on Probability Theory and Mathematical Statistics//—2-nd Edition — New York, USA, 2012 , — 657 Pages.
https://www.pdfdrive.com/lectures-on-probability-theory-and-mathematical-statistics-e188437236.html
4. *Suhov Y.,Kelbert M.* Probability and Statistics by Example: 1. Basic Probability and Statistics/—2-nd Edition — Cambridge University Press, 2014, — 476 Pages.
https://www.pdfdrive.com/probability-and-statistics-by-example-volume-1-basic-probability-and-statistics-e166589935.html

Дополнительнаялитература:
(Additional literature):

1. *Гмурман, В. Е.* Теориявероятностей и математическая статистика: учебник для прикладного бакалавриата / В. Е. Гмурман. — 12-е изд. — М. : Издательство Юрайт, 2017. — 480 с. https://www.biblio-online.ru/book/F6DC17CF-66E8-400F-9CDA-8067F86D996A
https://alleng.org/d/math/math321.htm
2. *Кремер, Н. Ш.* Теория вероятностей: учебник и практикум для академического бакалавриата / Н. Ш. Кремер — М. : Издательство Юрайт, 2018.

— 272 с. https://www.biblio-online.ru/book/F6DC17CF-66E8-400F-9CDA-8067F86D996A

3. *Soong T.T.* Fundamentals of Probability and Statistics for engineers/— New York, USA, 2004 , — 408 Pages. https://www.pdfdrive.com/fundamentals-of-probability-and-statistics-for-engineers-e6851455.html

4. *Письменный, Д. Т.* Конспект лекций по теории вероятностей и математической статистике / Д. Т. Письменный — М. : Издательство АЙРИС пресс, 2015 — 288 с. http://padabum.com/d.php?id=22233
5. *Семенов В. А.* Теория вероятностей и математическая статистика / В. А. Семенов — М., С.-Петербург, Н.-Новгород, Воронеж, Ростов-на-Дону, Екатеринбург, Самара, Новосибирск, Киев, Харьков, Минск, : Издательство ПИТЕР, 2013 — 192 с. https://www.razym.ru/naukaobraz/disciplini/matem/302650-semenov-va-teoriya-veroyatnostey-i-matematicheskaya-statistika.html
6. *Rohatgi V.K.,, A.K. Md. Ehsanes Saleh A.K An Introduction to Probability and Statistics*/—2-nd Edition — New York, USA, 1976 , — 747 Pages. https://www.pdfdrive.com/an-introduction-to-probability-and-statistics-wiley-series-in-probability-and-statistics-e168585572.html

Литература, из которой бралсяматериал используемый при подготовке лекционного курса:

(Literature from which the material used in the preparation of the lecture course was taken):

[1].*Гмурман, В. Е.* Теория вероятностей и математическая статистика: учебник для прикладного бакалавриата / В. Е. Гмурман. — 12-е изд. — М. : Издательство Юрайт, 2017. — 480 с. https://www.biblio-online.ru/book/F6DC17CF-66E8-400F-9CDA-8067F86D996A
https://alleng.org/d/math/math321.htm

[2].*Кремер, Н. Ш.* Теория вероятностей: учебник и практикум для академического бакалавриата / Н. Ш. Кремер — М. : Издательство Юрайт, 2018. — 272 с. https://www.biblio-online.ru/book/F6DC17CF-66E8-400F-9CDA-8067F86D996A

[3]. *Письменный, Д. Т.* Конспект лекций по теории вероятностей и математической статистике / Д. Т. Письменный — М. : Издательство АЙРИС пресс, 2015 — 288 с. http://padabum.com/d.php?id=22233

[4].*Халафян А. А., Боровиков В. П., Калайдина Г. В.* Теория вероятностей, математическая статистика и анализ данных. Основы теории и практика на компьютере. / А. А. Халафяни др. . — M. URSS, 2017 — 320 с. https://geminibook.xyz/books/teoriya-veroyatnostey-matemati

[5].*Семенов В. А.* Теория вероятностей и математическая статистика / В. А. Семенов — М., С.-Петербург, Н.-Новгород, Воронеж, Ростов-на-Дону, Екатеринбург, Самара, Новосибирск, Киев, Харьков, Минск, : Издательство ПИТЕР, 2013 — 192 с. https://www.razym.ru/naukaobraz/disciplini/matem/302650-semenov-va-teoriya-veroyatnostey-i-matematicheskaya-statistika.html

[6].*Suhov Y.,Kelbert M.* Probability and Statistics by Example: 1. Basic Probability and Statistics/—2-nd Edition — Cambridge University Press, 2014, — 476 Pages.

https://www.pdfdrive.com/probability-and-statistics-by-example-volume-1-basic-probability-and-statistics-e166589935.html

[7]. *Tobodga M.* Lecture on Probability Theory and Mathematical Statistics//—2-nd Edition — New York, USA, 2012 , — 657 Pages. https://www.pdfdrive.com/lectures-on-probability-theory-and-mathematical-statistics-e188437236.html

[8].*Suhov Y.,Kelbert M.* Probability and Statistics by Example: 1. Basic Probability and Statistics/—2-nd Edition — Cambridge University Press, 2014, — 476 Pages. https://www.pdfdrive.com/probability-and-statistics-by-example-volume-1-basic-probability-and-statistics-e166589935.html

[9]. *Виленкин Н. Я., Потапов В. Г.* Задачник-практикум по теории вероятностей с элементами комбинаторики и математической статистики / Виленкин Н. Я., Потапов В. Г. . — М. Издательство ПРОСВЕЩЕНИЕ, 1979 — 112 с. http://padabum.com/d.php?id=29021

[3]. *Емельянов Г. В., Скитович В. П.* Задачник по теории вероятностей и математической статистике / Емельянов Г. В., Скитович В. П. — Ленинград, Издательство ленинградского университета, 1967 — 333 с. http://mexalib.com/view/33788

[5]. *Лунгу К. Н. и др. Сборник задач по высшей математике с контроль-ными работами /* К. Н. Лунгу и др. — М. : Издательство АЙРИС пресс, 2017 — 592 с.

https://www.razym.ru/naukaobraz/obrazov/53551-sbornik-zadach-po-vysshej-matematike-2-kurs.html

[6]. Сборник задач по теории вероятностей, математической статистике и теории случайных функций: Учебное пособие / Под общей ред. А.А.Свешникова, 4-е изд., - СПб.; Изд. «Лань», 2008. — 448 с. http://en.bookfi.net/book/1501516

[8]*Прохоров А.В.,Ушаков В.Г., Ушаков Н.Г.,* Задачи по теории вероятностей: Основные понятия. Предельные теоремы. Случайные процессы: Учебное пособие. — М.: Наука. Гл. ред. физ.-мат лит., 1968. 1968. — 328 с.http://mexalib.com/download/10676

Андрей Борисович Колпаков

Анна Сергеевна Рукомина

Краткий курс лекций по дисциплине «*Теория вероятностей и Математическая статистика*» Часть 2. Математическаястатистика. **(***Short course of lectures on the discipline "Probability theory and Mathematical statistics"* Part 2. Mathematical statistics**)**

Учебно-методическоепособие

Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского».
603950, Нижний Новгород, пр. Гагарина, 23.