МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФГБОУ ВО «Воронежский государственный
технический университет»

Е.Г. Глушко

ТЕОРИЯ ВЕРОЯТНОСТЕЙ.
ПРАКТИЧЕСКИЕ ЗАНЯТИЯ

Утверждено учебно-методическим советом
университета в качестве учебного пособия

Воронеж 2017

УДК 075.8
ББК 22.1я7
Г 555

Глушко Е.Г. Теория вероятностей. Практические занятия : учеб. пособие [Электронный ресурс]. – Электрон. текстовые, граф. данные (4,14 Мб) / Е.Г. Глушко. – Воронеж : ФГБОУ ВО «Воронежский государственный технический университет», 2017. – 1 электрон. опт. диск (CD-ROM) – Систем. требования: ПК 500 и выше; 256 Мб ОЗУ; Windows XP; SVGA с разрешением 1024x768; Adobe Acrobat; CD-ROM дисковод; мышь. – Загл. с экрана.

Учебное пособие содержит необходимые теоретические сведения и формулы, приведены решения типовых задач и задачи для самостоятельного решения по курсу «Спецглавы математики». Предназначается для проведения практических занятий, составления домашних заданий и самостоятельной подготовки студентов.

Издание соответствует требованиям Федерального государственного образовательного стандарта высшего образования по направлениям подготовки бакалавров 09.03.02 «Информационные системы и технологии», дисциплины «Спецглавы математики».

Табл. Ил. Библиогр.: 7 назв.

Рецензенты: кафедра уравнений в частных производных и теории вероятностей Воронежского государственного университета (зав. кафедрой д-р физ.-мат. наук, проф. А.В. Глушко); канд. физ.-мат. наук, доц. В.В. Горбунов

ВВЕДЕНИЕ

Пособие предназначено для студентов технических вузов, изучающих теорию вероятностей и математическую статистику в рамках курса спецглавы математики, и соответствует требованиям государственного образовательного стандарта.

В пособие приведены материалы для практических занятий по теории вероятностей и математической статистике, охватывающие темы: комбинаторика, непосредственное вычисление вероятностей, теоремы сложения и умножения вероятностей, формула полной вероятности и формула Байеса, схема независимых испытаний, законы распределения и числовые характеристики случайных величин, многомерные случайные величины, функции случайных величин, центральная предельная теорема и следствия из нее, закон больших чисел, неравенство Чебышева, элементы математической статистики, статистическая оценка неизвестных параметров распределения, точечные оценки, интервальные оценки параметров распределения, проверка статистических гипотез.

Каждое практическое занятие начинается с подробного перечня необходимых понятий и формул, что позволяет оперативно находить нужную информацию. После сводки формул и теоретических фактов приводятся задачи с решениями и необходимые методические рекомендации. Это позволяет быстро и эффективно формировать навыки решения задач. В конце каждой темы предложены задачи и упражнения для самостоятельного решения. Эти задачи подобраны так, что они по методам решения перекликаются с задачами, решения которых приведены в тексте.

ЗАНЯТИЕ №1. ЭЛЕМЕНТЫ КОМБИНАТОРИКИ

Пусть имеется k групп элементов, причем i-тая группа состоит из n_i – элементов. Сколькими способами можно составить новое множество, взяв из каждого исходного множества по одному элементу? Ответ на этот вопрос дает *основной комбинаторный принцип*:

Если некоторый первый выбор можно сделать n_1 способами, для каждого первого выбора некоторый второй можно сделать n_2 способами, для каждой пары первых двух - третий выбор можно сделать n_3 способами и так далее, то число способов для последовательности таких выборов равно

$$N = n_1 \cdot n_2 \cdot \ldots \cdot n_k .$$

Комбинаторные формулы в прикладных задачах теории вероятностей обычно связывают с выбором m элементов («выборкой объема m») из совокупности, состоящей из n элементов (элементов «генеральной совокупности»). Различают два способа выбора:

а) *повторный*, при котором выбранный элемент возвращается в генеральную совокупность и может быть выбран вновь;

б) *бесповторный*, при котором выбранный элемент в генеральную совокупность не возвращается и выборка не содержит повторяющихся элементов

При повторном выборе выборку объема m можно сделать n^m способами. Например, повторную выборку объема два из трех элементов $\{a, b, c\}$ можно сделать $3^2 = 9$ способами: $aa, ab, ba, bb, bc, cb, ac, ca, cc.$

При бесповторном выборе выборку объема m можно сделать $A_n^m = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-m+1)$ способами. Число A_n^m

называют *числом размещений* из n элементов по m. Размещения отличаются либо *составом* элементов, либо *порядком* их расположения. Например, размещений из трех элементов $\{a,b,c\}$ по два можно составить $A_3^2 = 3 \cdot 2 = 6$: $ab, ba, ac, ca, bc, cb.$ Число размещений вычисляется по формуле $A_n^m = \dfrac{n!}{(n-m)!}$.

Выборки объема m, которые отличаются друг от друга только составом, можно сделать $C_n^m = \dfrac{n!}{m!(n-m)!}$ способами.

Число C_n^m называют *числом сочетаний* из n элементов по m. Например, сочетаний из трех элементов $\{a,b,c\}$ по два существует $C_3^2 = 3$: $ab, ac, bc.$

При повторном выборе из n элементов число выборок объема m, которые отличаются только составом равно C_{n+m-1}^m.

Перестановкой из n элементов называется любой упорядоченный набор этих элементов.

Число различных перестановок из n элементов P_n вычисляется по формуле $P_n = n! = 1 \cdot 2 \cdot \ldots \cdot n$. Совокупность из n элементов разделить на m групп по $k_1, k_2, \ldots, k_m$ элементов соответственно ($k_1 + k_2 + \ldots + k_m = n$) можно $\dfrac{n!}{k_1! \cdot k_2! \cdot \ldots \cdot k_m!}$ способами.

Порядок элементов внутри каждой из этих m групп не имеет значения.

Пусть $A_1, A_2, \ldots, A_k$ - множества, число элементов в каждом из которых равно соответственно $n_1, n_2, \ldots, n_k$. Составить множество B из m_1 элементов множества A_1, m_2 элементов множества $A_2, \ldots, m_k$ элементов множества A_k, можно согласно основному комбинаторному принципу,

$$C_{n_1}^{m_1} \cdot C_{n_2}^{m_2} \cdot \ldots \cdot C_{n_k}^{m_k}$$

способами.

Для безошибочного выбора комбинаторной формулы достаточно последовательно ответить на вопросы в следующей схеме:

Откуда выбор? (n-?) Сколько выбираем? (m-?)	Состав	Бесповторный	C_n^m
		Повторный	C_{n+m-1}^m
	Состав и порядок	Бесповторный	A_n^m
		Повторный	n^m
	Порядок	Бесповторный	$n!$
		Повторный	$\dfrac{n!}{k_1!k_2!\ldots k_m!}$

Пример 1. Сколько сообщений можно послать посредством семи знаков точек и тире?

Решение. Выбор знака производится из множества двух элементов: точка и тире (n=2). Повторным способом выбирается семь элементов (m=7). Поэтому число различных сообщений равно $2^7 = 128$.

Пример 2. Сколько комбинаций из четырех букв можно составить? Сколько из них содержат только разные буквы?

Решение. Из совокупности 33 букв (n=33) необходимо выбрать четыре буквы (m=4). Если запрета на повторение букв нет, то выбор повторный и общее число комбинаций равно $(33)^4 = 1185921$. Если необходимо иметь только разные буквы, то выбор бесповторный и общее число комбинаций равно

$$A_{33}^4 = 33 \cdot 32 \cdot 31 \cdot 30 = 982080.$$

Пример 3. Сколькими способами можно разложить восемь книг на две пачки по четыре книги в каждой? Сколькими способами можно разложить эти книги на четыре пачки по две книги в каждой? Сколькими способами можно разослать эти книги восьми различным адресатам?

Решение. 1. Для распределения книг на две равные пачки достаточно из восьми книг (n=8) выбрать бесповторным образом любые четыре (m=4) для первой пачки, причем нас интересует только состав выбора, а остальные книги оставить для второй пачки. Поэтому общее число способов равно числу сочетаний из восьми элементов по четыре, то есть $C_8^4 = \dfrac{8!}{4! \cdot 4!} = 70.$

2. Для комплектации четырех пачек по две книги в каждой необходимо сначала бесповторным способом выбрать состав первой пачки ($C_8^2 = 28$ способов), затем из оставшихся шести книг выбрать две книги для второй пачки ($C_6^2 = 15$ способов), после этого из оставшихся четырех книг выбрать две книги для третьей пачки ($C_4^2 = 6$ способов), а оставшиеся две книги составят четвертую пачку (формально, $C_2^2 = 1$ способ). Заме-

тим, что при каждом выборе мы интересовались только составом. По комбинаторному принципу для описанной последовательности выборов существует $28 \cdot 15 \cdot 6 = 2520$ способов.

3. Рассылка восьми адресатам по одной книге каждому означает перестановку из восьми элементов, то есть имеется $8! = 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 40320$ способов.

Замечание. Для способов разделить n элементов на m групп по $k_1, k_2, ...k_m$ соответственно в каждой ($k_1 + k_2 + ... + k_m = n$) известна формула $\dfrac{n!}{k_1!k_2!...k_m!}$.

По этой формуле решение примера 3 (2) можно получить сразу $\dfrac{8!}{2!2!2!2!} = 2520.$

Пример 4. Сколькими способами можно составить трехцветный флаг, если имеется материал пяти цветов?

Решение. Для составления трехцветного флага нужно из пяти цветов (n=5) выбрать три различных цвета (m=3) , иначе флаг не будет трехцветным. При выборе нас интересует состав выбора и последовательность цветов. Поэтому число трехцветных флагов в нашем случае равно числу размещений из пяти по три, то есть равно $A_5^3 = 5 \cdot 4 \cdot 3 = 60.$

Пример 5. Каких чисел от 1 до 10000000 больше - тех, в записи которых встречается единица, или тех, в записи которых нет ни одной единицы?

Решение. Для того чтобы записать семизначное число, в записи которого нет ни одной единицы, необходимо повторным способом из девяти цифр (0,2,3,4,5,6,7,8,9) выбрать семь цифр. Это можно сделать $9^7 = 4769847 < 5 \cdot 10^6$ способами.

Следовательно, среди первых 10 миллионов чисел больше тех, в записи которых единица есть.

Пример 6. Каждый из 10 студентов может явиться на зачет в любой из двух назначенных дней. Сколькими способами могут студенты распределиться по дням явки на зачет? Сколькими способами могут распределиться студенты по дням явки на зачет, если каждый день должны сдавать зачет по пять студентов?

Решение. При распределении по дням явки каждый из 10 студентов производит выбор между двумя возможностями. По комбинаторному принципе всего способов выбора имеется $2^{10} = 1024$. Если же в каждый из дней должно явиться на зачет по пять студентов, то достаточно выбрать студентов для первого дня зачета, а остальные будут сдавать зачет во второй день. Из 10 студентов (n=10) следует выбрать бесповторным способом пять студентов (m=5) , причем нас интересует только состав выбора. Поэтому возможных комбинаций будет $C_{10}^{5} = 252$.

Пример 7. В некотором государстве не было двух жителей с одинаковым набором зубов. Какой может быть наибольшая численность населения в этом государстве?

Решение. В отношении каждого из 32 зубов может быть две возможности: есть этот зуб или его нет. Выбор из этих двух возможностей нужно произвести 32 раза. Поэтому число всех мыслимых комбинаций равно $2^{32} = 4294967296$.

Пример 8. Сколькими различными способами можно переставить между собой буквы:

а) A_1, A_2, B_1, B_2, B_3; б) A, A, B_1, B_2, B_3; в) A, A, B, B, B?

Решение. а) Так как все буквы различны, то число перестановок из них равно 5!=5·4·3·2·1=120.

б) В этом случае каждая пара перестановок отличающаяся только порядком расположения букв, например, A_1, B_1, A_2, B_2, B_3 и A_2, B_1, A_1, B_2, B_3, сливается в одну перестановку A, B_1, A, B_2, B_3. Поэтому различных перестановок будет $120 / 2 = 60$.

в) Итак, имеется 60 перестановок, которые отличаются друг от друга либо местами расположения букв B_1, B_2, B_3, либо порядком расположения этих букв на данных местах. Всего перестановок букв B_1, B_2, B_3 на заданных трех местах существует 3!=6. Поэтому при потере индексов у букв B_1, B_2, B_3 различимых комбинаций будет 60/6=10.

Заметим, что число различимых перестановок равно числу способов выбора из пяти мест любых двух и постановки на них буквы *A*. На остальные места ставим буквы *B*. Это число равно $C_5^2 = 10$.

Задачи для самостоятельного решения

1. Перечислите все перестановки из элементов А, В, С, Д. Перечислите все размещения по два элемента и все сочетания по два элемента из элементов А, В, С, Д.

2. Сколькими способами можно из колоды карт (36 штук) выбрать пять карт так, чтобы среди них было два туза?

3. Сколькими способами можно из колоды карт (36 штук) выбрать четыре карты разных мастей?

4. Сколько различимых перестановок можно сделать из букв слова: а) наука; б) математика?

5. Сколько автомобильных номеров можно составить из трех букв и трех цифр?

6. Сколькими способами можно переставить между собой цифры 1,2,3,4,5,6 так, чтобы на четных по порядку местах стояли четные цифры, а на нечетных - нечетные?

7. Сколькими способами можно расставить шесть книг по трем полкам? Сколько способов расставить книги так, чтобы ни одна полка не пустовала?

8. Сколько четырехзначных чисел состоят только из разных цифр?

9. Сколькими способами можно поставить на полку шесть книг так, чтобы три заданные книги оказались рядом (в произвольном порядке)?

10. Сколькими способами можно разделить 15 команд на три подгруппы в каждой по пять команд?

11.Сколько различных частных производных третьего порядка имеет функция трех переменных?

12. Сколькими способами можно расселить девять студентов в трех комнатах, рассчитанных на трех человек каждая? Сколькими способами это можно сделать, если какие-либо два из этих студентов отказываются поселиться в одной комнате?

13. Студенту нужно выбрать два факультативных курса из шести возможных. Сколькими способами он может это сделать?

14. Два города A и B соединены четырьмя различными дорогами. Сколькими способами можно проехать из A в B и обратно? Сколько существует таких способов, если на обратном пути непременно выбирать новую дорогу?

15. В соревнованиях принимают участие 16 равносильных

команд. Сколькими способами могут быть распределены золотая и серебряная медали?

16. Сколькими способами можно выбрать путь из начала координат в точку с координатами (6,4), если каждый шаг равен 1, и его можно совершать только вправо и вверх?

Ответы: 2. 29760; 3. 6561; 4. а) 60; б) 151200;
5. 35937000, 23569920; 6. 20; 7. 729, 540; 8. 4536; 9. 144;
10. 756756; 11. 10; 12. 1680, 1260; 13. 15; 14. 16, 12; 15. 240;
16. 210.

ЗАНЯТИЕ № 2. НЕПОСРЕДСТВЕННОЕ ВЫЧИСЛЕНИЕ ВЕРОЯТНОСТЕЙ

2.1. Классическое определение вероятности

Классическая схема служит моделью тех случайных явлений, для которых представляется естественным предположение о равновозможности (равновероятности) конечного числа элементарных событий. Общепринята следующая формулировка классического определения вероятности: *вероятностью* события *A* называется отношение числа m исходов опыта, благоприятствующих появления события А, к общему числу n равновозможных исходов опыта.

То есть вероятность события *A* определяется как $P(A)=\dfrac{m}{n}$.

Пример 1. Какова вероятность появления герба по крайней мере один раз при двукратном бросании монеты?

Решение. Пространство равновозможных элементарных событий данного опыта состоит из следующих событий: $\Omega=\{(г,г),(г,ц),(ц,г),(ц,ц)\}$, $n=4$. Событие *A*, состоящее в том, что при двукратном бросании монеты герб появится по крайней мере один раз, происходит при появлении одного из несовместных элементарных событий $(г,г)$, $(г,ц)$, $(ц,г)$. Следовательно, $A = \{(г,г), (г,ц), (ц,г)\}$, $m = 3$. Таким образом, $P(A) = \dfrac{m}{n} = \dfrac{3}{4}$.

Пример 2. Технический контроль проверяет из партии в 500 деталей 20 деталей, взятых наудачу. Партия содержит 15 нестандартных деталей. Какова вероятность того, что среди проверяемых деталей будет ровно две нестандартные?

Решение. Так как по условию задачи 20 деталей из 500 извлекаются наудачу, то все возможные варианты извлечения 20 деталей из 500 естественно считать равновозможными и для нахождения требуемой вероятности воспользоваться классической схемой (классическим определением вероятности).

Так как порядок следования стандартных и нестандартных деталей в извлекаемых 20 не играет роли, а играет роль только количество стандартных и нестандартных деталей, то количество всех возможных способов, которыми это можно сделать, равно C_{500}^{20}, то есть $n = C_{500}^{20}$.

Событию A, состоящему в том, что будут извлечены две нестандартные детали при извлечении 20 (следовательно, остальные 18 должны быть стандартными), будет соответствовать $C_{15}^{2} \cdot C_{485}^{18}$ исходов, то есть $m = C_{15}^{2} \cdot C_{485}^{18}$. Таким образом,

$$P(A) = \frac{C_{15}^{2} \cdot C_{485}^{18}}{C_{500}^{20}}.$$

Пример 3. Трехзначное число составляется следующим образом: бросаются три игральные кости белая, синяя и красная; число выпавших очков на белой кости – это число сотен, число выпавших очков на синей кости – это число десятков, а число выпавших очков на красной кости – это число единиц трехзначного числа. Какова вероятность того, что полученное таким образом число будет больше 456?

Решение. Количество всех чисел, которые можно получить указанным способом – это число размещений с повторениями из 6 по 3. Следовательно, $n = A_6^3 = 6^3$.

Числа большие 456 будут получаться, если число сотен будет больше 4, то есть 5 или 6 или число сотен будет равно 4, а число десятков будет больше чем 5, то есть 6. Количество таких чисел будет $2 \cdot A_6^2 + 1 \cdot A_6^1$. Следовательно, $m = 2 \cdot A_6^2 + 1 \cdot A_6^1$ $= 2 \cdot 6^2 + 6 = 78$ и $P(A) = \dfrac{78}{216} = \dfrac{13}{36}$.

2.2. Геометрическое определение вероятности

Геометрическое определение обобщает классическое определение вероятности на случай, когда пространство элементарных событий Ω представляет собой подмножество пространства R^n.

При этом на прямой будем рассматривать лишь промежутки или их объединения, т.е. подмножества, которые имеют длину, на плоскости – те подмножества, которые имеют площадь и т.д.

Под мерой $mes(A)$ множества $A \in \Omega$ будем понимать его длину, площадь или объем, в зависимости от того к какому пространству принадлежит Ω R^1, R^2 или R^n. Будем считать, что $0 < mes\,\Omega < \infty$, и вероятность попадания случайно брошенной точки в любое подмножество Ω пропорционально мере этого подмножества и не зависит от его расположения и формы. В этом случае вероятность считается по формуле:

$$P(A)=\frac{mes(A)}{mes(\Omega)} .$$

Пример 1. Телефонная линия длиной 2 км. соединяющая пункты А и В порвалась в неизвестном месте. Считая обрыв равновозможным в любой точке линии, найти вероятность того, что обрыв находится не далее чем 450 м. от пункта А.

Решение. Точка $x-$ место обрыва линии может с одинаковой вероятностью занимать любое положение на отрезке длиной 2000м. Следовательно, множество Ω непрерывно и его мера равна 2000. Событие C, состоящее в том, что обрыв произошел на расстоянии не более 450 от пункта А состоит из точек отрезка длиной 450 м. Следовательно, $mes(C)=450$ и

$$P\ \ C\ \ =\frac{450}{2000}=0,225 .$$

Пример 2. В эллипс с полуосями 2 и 3 наудачу ставится точка. Какова вероятность того, что она попадет во вписанную в эллипс окружность, центр которой совпадает с центром эллипса?

Решение. Точка $\ x,y\ $ может с одинаковой вероятностью занимать любое положение в области ограниченной эллипсом. Следовательно, множество Ω непрерывно и может быть записано в виде $\frac{x^2}{2^2}+\frac{y^2}{3^2}\le 1$. $mes(\Omega)=\pi\cdot 2\cdot 3=6\pi$. Событие А, состоящее в том, что точка попадет в круг, вписанный в эллипс, состоит из точек множества Ω, для которых выполняется условие $x^2 + y^2 \le 2^2$. $mes(A) = \pi \cdot 2^2 = 4\pi$. Следовательно, $P\ \ A\ = \dfrac{4\pi}{6\pi} = \dfrac{2}{3} \approx 0,67.$

Пример 3. На отрезок $\ 0,l\ $ длины l наудачу поставлены две точки x и y. Наблюдаемый результат – три отрезка. Какова вероятность того, что из полученных отрезков можно построить треугольник?

Решение. Точки x и y могут с одинаковой вероятностью занимать любое положение на отрезке $\ 0,l\ $. Следовательно, множество Ω непрерывно и может быть записано в виде $\Omega = \ \ x, y\ \ 0 < x, y < l\ $. Геометрически Ω – квадрат со стороной l.

При постановке точек x и y на отрезок $\ 0,l\ $ возможны два принципиально различных случая: $x < y$ и $y < x$.

Если $x < y$, то для того, чтобы из отрезков $\ 0,x\ ,\ \ x,y\ ,\ \ y,l\ $ можно было построить треугольник, необходимо, чтобы $x < l/2, y - x < l/2, l - y < l/2$ (каждый из отрезков должен быть меньше $l/2$).

Если $y < x$, то должны выполняться аналогичные условия $y < l/2,\ x - y < l/2,\ l - x < l/2$.

Если А – событие состоящее в том, что из полученных отрезков можно построить треугольник, то это событие будет состоять из точек области Ω, удовлетворяющим записанным выше условиям.

Области, соответствующие пространству элементарных событий Ω и событию А: из полученных отрезков можно построить треугольник, изображены на рис. 1:

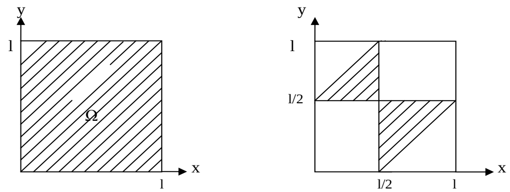


Рис. 1

Естественно, что $P\ \ A\ = \dfrac{S_A}{S_\Omega} = \dfrac{l^2/4}{l^2} = \dfrac{1}{4}$.

Пример 4. Для поражения точечной воздушной цели достаточно разрыва снаряда на расстоянии 10 м. от неё. Из-за ошибок прицеливания разрыв снаряда равновозможен в любой точке эллипсоида с центром в точки цели и полуосями 20, 20 и 60 м. Какова вероятность того, что цель будет поражена?

Решение. Точка x, y, z может с одинаковой вероятностью занимать любое положение в области ограниченной эллипсоидом. Следовательно, множество Ω непрерывно и может быть записано в виде $\dfrac{x^2}{20^2} + \dfrac{y^2}{20^2} + \dfrac{z^2}{60^2} \le 1$.

$$mes|\Omega| = \frac{4}{3}\pi \cdot 20 \cdot 20 \cdot 60 = 32000\pi.$$

Событие А, состоящее в том, что точка попадет в сферу радиуса 10 состоит из точек множества Ω, для которых выполняется условие $x^2 + y^2 + z^2 \le 10^2$. $mes|A| = \dfrac{4}{3}\pi \cdot 10^3 = \dfrac{4000}{3} \cdot \pi$.

Следовательно, $P\ A\ = \dfrac{4000\pi}{3\cdot 32000\pi} \approx 0{,}04$.

Задачи для самостоятельного решения

1. Десять команд случайным образом (по жребию) разбиваются на две равные подгруппы. Какова вероятность того, что две сильнейшие команды попадут в разные подгруппы? ….в одну подгруппу? …в первую подгруппу?

2. Из полной колоды карт (52 штуки) наугад выбраны три карты. Какова вероятность того, что это «тройка», «семерка», «туз»? Какова вероятность того, что эти карты выбраны в указанной последовательности?

3. На отрезок ОА длины L брошены «наугад» две точки B и C, причем точка C расположена правее точки B. Найти вероятность того, что длина отрезка BC меньше длины отрезка OB.

4. Десять билетов с номерами от 1 до 10 перемешаны на столе экзаменатора. Какова вероятность того, что эти билеты будут вытянуты студентами в порядке их номеров?

5. Четыре человека вошли в лифт на первом этаже девятиэтажного дома. Считая, что равновозможен выход каждого пассажира на любом из этажей со второго по девятый, найти вероятность того, что: а) все пассажиры выйдут на разных этажах; б) все пассажиры выйдут выше пятого этажа; в) на третьем этаже не выйдет ни одного пассажира.

6. Наугад выбираются четыре цифры и расставляются в случайном порядке. Какова вероятность того, что поручится четырехзначное число? Какова вероятность того, что это четырехзначное число делится на 5?

7. Имеется 20 экзаменационных билетов, расположенных на столе в случайном порядке. Десять студентов один за дру-

гим выбирают наугад по одному билету. Какова вероятность того, что билеты с номерами 1 и 2 не будут выбраны?

8. Из урны, в которой лежат, шесть белых, четыре черных и два красных шара, наудачу выбирают четыре шара. Какова вероятность того, что среди них только черные и красные шары?

9. Десять книг, из них три красные, в случайном порядке поставлены на полку. Какова вероятность того, что три красные книги в любом порядке стоят рядом?

10. Бросаются три игральные кости. Найдите вероятности следующих событий: A={кости выпадут разными гранями}; B={на всех костях выпадет одинаковой число очков}.

11. В шкафу лежат вперемешку пять пар ботинок. Наугад выбирается два ботинка. Какова вероятность того, что они образуют пару?

12. Каждый из шести призов в результате жеребьевки разыгрывается между десятью участниками. Какова вероятность того, что данные шесть участников получат по одному призу каждый?

13. Группа из четырех юношей и четырех девушек по жребию делится на две подгруппы по четыре человека. Какова вероятность того, что в каждую подгруппу попадет поровну юношей и девушек?

14. Внутрь круга радиусом наугад брошена точка. Какова вероятность того, что она попадет внутрь вписанного в круг : а) квадрата, б) правильного треугольника?

15. В течение суток к причалу независимо друг от друга должны подойти и разгрузиться два сухогруза. Одному из них требуется для разгрузки шесть часов, другому - восемь. Како-

ва вероятность того, что ни одному из сухогрузов не придется ждать очереди для разгрузки?

16. Коэффициенты b и c квадратного уравнения $x^2+bx+c=0$ наудачу выбираются из множества целых чисел $1,2,3,4$. Какова вероятность того, что полученное квадратное уравнение имеет действительные корни?

17. Параметры a и b уравнения кривой второго порядка $ax^2+by^2=1$ наудачу выбираются из множества целых чисел $-4,\ldots,8$. Какова вероятность того, что полученное уравнение является уравнением гиперболы?

18. Элементы a_{21} и a_{12} матрицы $\begin{pmatrix} 1 & a_{12} \\ a_{21} & 4 \end{pmatrix}$ наудачу выбираются из множества целых чисел $1,2,3,4$. Какова вероятность того, что определитель полученной матрицы будет неположительным?

19. Даны два вектора $\vec{a}\ 4;3$ и $\vec{b}\ x;y$. Координаты вектора $\vec{b}$ наудачу выбираются из множества целых чисел $4,5,6,7$. Какова вероятность того, что модуль вектора $\vec{a}$ будет меньше модуля вектора $\vec{b}$?

Ответы: 1. 5/9, 4/9, 2/9; 2. 16/5525≈0,003, 8/16575≈0,0005; 3. ½, 4. 1/10!; 5. а) 105/256, б) 1/16, 2401/4096≈0.59
6. 0,9, 0,18; 7. 9/38; 8. 1/33; 9. 1/15; 10 5/9, 1/36; 11/ 1/9;
12. 0,0072; 13. 6/35; 14. а) 2/π≈2/3, б) $(3\sqrt{3})\,/\,(4\pi)\approx 0,41$;
15. 25/72≈1/3.

ЗАНЯТИЕ №3. ТЕОРЕМЫ СЛОЖЕНИЯ И УМНОЖЕНИЯ ВЕРОЯТНОСТЕЙ

Условной вероятностью события *A* при условии, что произошло событие *B*, называется число $P\left(A/B\right)=\dfrac{P(AB)}{P(B)}$; $P(B)\neq 0$.

Из этого определения вытекает формула умножения вероятностей

$$P(AB)=P(A)P(B/A)=P(B)P(A/B)$$

для двух событий, которая допускает обобщение для *n* событий:

$$P(A_1\cdot A_2\cdot\ldots\cdot A_n)=$$
$$P(A_1)P(A_2/A_1)\cdot P(A_3/A_1A_2)\cdot\ldots\cdot P(A_n/A_1\ldots A_{n-1}).$$

События *A* и *B* называются *независимыми*, если $P(A/B)=P(A)$ или $P(AB)=P(A)P(B)$.

Для любых событий *A* и *B* имеет место формула

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

Если события А и В несовместны, $AB=\varnothing$ (где $\varnothing$ - невозможное событие), то $P(A+B)=P(A)+P(B)$.

При решении большинства задач теоремы сложения и умножения вероятностей используются вместе. При этом методика решения задач состоит в следующем: событие, вероятность которого требуется найти, выражается через элементарные или более простые события, вероятности которых известны или могут быть найдены; затем от равенства событий переходят к равенству вероятностей этих событий и, используя теоремы сложения и умножения вероятностей, переходят к умножению, сложению вероятностей элементарных или более простых событий, то есть чисел.

В тех случаях, когда событие, вероятность которого требуется найти, громоздко выражается через простые события, удобнее переходить к противоположному событию, находить его вероятность, а затем, используя формулу $P(A)=1-P(\bar{A})$ получать требуемый результат.

Пример 1. Для сигнализации об аварии установлены два независимо работающих сигнализатора. Вероятность того, что при аварии сработает первый сигнализатор, равна 0,95; для второго сигнализатора эта вероятность равна 0,9. Какова вероятность того, что при аварии сработает хотя бы один сигнализатор?

Решение. Введем в рассмотрение события:

A_1 – событие, состоящее в том, что при аварии сработает первый сигнализатор;

A_2 – событие, состоящее в том, что при аварии сработает второй сигнализатор;

B – событие, состоящее в том, что при аварии сработает хотя бы один сигнализатор.

Тогда $B=A_1\cdot\bar{A}_2+\bar{A}_1\cdot A_2+A_1\cdot A_2$. Перейдем от равенства событий к равенству вероятностей этих событий $P(B)=P(A_1\cdot\bar{A}_2+\bar{A}_1\cdot A_2+A_1\cdot A_2)$. Так как события $A_1\cdot\bar{A}_2$, $\bar{A}_1\cdot A_2$ и $A_1\cdot A_2$ несовместны, то

$$P(B)=P(A_1\cdot\bar{A}_2)+P(\bar{A}_1\cdot A_2)+P(A_1\cdot A_2).$$

События A_1 и $\overline{A}_2$, $\overline{A}_1$ и A_2, A_1 и A_2 независимые (по условию задачи сигнализаторы работают независимо). Следовательно,

$$P(B) = P(A_1)\cdot P(\overline{A}_2) + P(\overline{A}_1)\cdot P(A_2) + P(A_1)\cdot P(A_2).$$

Подставляя в последнее равенство известные из условия вероятности $P(A_1) = 0,95$, $P(A_2) = 0,9$ и учитывая то, что

$$P(\overline{A}_1) = 1 - P(A_1) = 1 - 0,95 = 0,05$$ и

$$P(\overline{A}_2) = 1 - P(A_2) = 1 - 0,9 = 0,1.$$

Получим $P(B) = 0,95\cdot 0,1 + 0,05\cdot 0,9 + 0,95\cdot 0,9 = 0,995$.

Пример 2. По каналу связи передаётся сообщение из трех символов, вероятности искажения которых при передаче сообщения соответственно равны 0,01, 0,005, 0,003. Сообщение считается принятым, если не искажено ни одного символа. Какова вероятность того, что сообщение не будет принято?

Решение. Введем в рассмотрение события A_1, A_2, A_3, состоящие в том , что при передаче будут искажены первый, второй и третий символы соответственно и событие B, состоящее в том, что сообщение не будет принято.

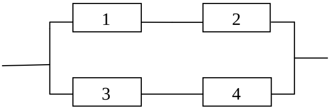
Тогда $B = A_1\cdot A_2\cdot A_3 + \overline{A}_1\cdot A_2\cdot A_3 + A_1\cdot \overline{A}_2\cdot A_3 + A_1\cdot A_2\cdot \overline{A}_3 +$
$+\overline{A}_1\cdot \overline{A}_2\cdot A_3 + \overline{A}_1\cdot A_2\cdot \overline{A}_3 + +A_1\cdot \overline{A}_2\cdot \overline{A}_3$, $\overline{B} = \overline{A}_1\cdot \overline{A}_2\cdot \overline{A}_3$.

Очевидно, что находить вероятность события $\overline{B}$ проще, чем вероятность события B, поэтому

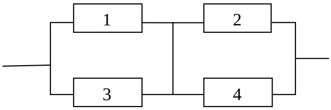
$$P(B) = 1 - P(\overline{B}) = 1 - P(\overline{A}_1\cdot \overline{A}_2\cdot \overline{A}_3) = 1 - 0,99\cdot 0,995\cdot 0,997 \approx$$
$$\approx 1 - 0,97 = 0,03.$$

Пример 3. Вероятность безотказной работы в течение заданного времени (надежность) каждого элемента равна 0,8. Из этих элементов составлены две системы:

а)



б)



Какая система надежнее? Иначе говоря, что выгоднее в системе дублировать каждый элемент отдельно или всю систему в целом?

Решение. Пусть событие A_i состоит в том, что *i*-тый элемент работает безотказно. Безотказная работа первой системы (событие B_1) равносильна безотказной работе первого элемента и второго или третьего элемента и четвертого. Символически это можно записать в виде $B_1 = A_1 \cdot A_2 + A_3 \cdot A_4$. События $A_1 \cdot A_2$ и $A_3 \cdot A_4$ совместны, а события A_i независимы. Поэтому

$$P(B_1) = P(A_1 \cdot A_2) + P(A_3 \cdot A_4) - P(A_1 \cdot A_2 \cdot A_3 \cdot A_4) =$$

$$= P(A_1) \cdot P(A_2) + P(A_3) \cdot P(A_4) - P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot P(A_4) =$$

$$= (0,8)^2 + (0,8)^2 - (0,8)^4 = 0,8704.$$

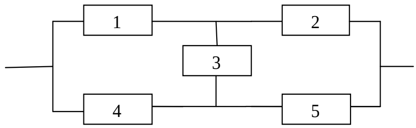
Безотказная работа второй системы (событие B_2) равносильна безотказной работе первого элемента или третьего и второго или четвертого, то есть $B_2 = (A_1 + A_3) \cdot (A_2 + A_4)$. Тогда

$$P(B_2) = P(A_1 + A_3) \cdot P(A_2 + A_4) = (P(A_1) + P(A_3) - P(A_1) \cdot P(A_3)) \cdot$$

$$\cdot (P(A_2) + P(A_4) - P(A_2) \cdot P(A_4)) = (0,8 + 0,8 - 0,64)^2 = 0,9216.$$

Результаты вычислений свидетельствуют о том, что выгоднее дублировать каждый элемент отдельно. Система *б*) надежнее.

Пример 4. Вероятность безотказной работы в течение заданного времени (надежность) каждого элемента равна 0,8. Из этих элементов составлена система.



Какова надежность системы?

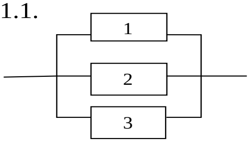
Решение. Сохраним обозначения примера 3. Если элемент №3 не работает, то система совпадает с системой *а*) из этого примера. Если же элемент №3 работает, то система совпадает с системой *б*) из предыдущего примера. Поэтому безотказная работа системы (событие *D*) равносильна событию $D = \overline{A}_3 \cdot B_1 + A_3 \cdot B_2$. События независимы, а события $\overline{A}_3 \cdot B_1$ и $A_3 \cdot B_2$ несовместны. Поэтому с учетом результатов предыдущей задачи имеем $P(D) = P(\overline{A}_3) \cdot P(B_1) + P(A_3) \cdot P(B_2) =$

$= 0,2 \cdot 0,8704 + 0,8 \cdot 0,91136.$

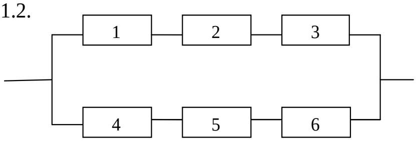
Задачи для самостоятельного решения

1. В следующих задачах приведены схемы соединения элементов, образующих цепь с одним входом и одним выходом. Предполагается, что отказы элементов являются независимыми в совокупности событиями. Считается известной надежность p_k *k*-го элемента (соответственно $q_k = 1 - p_k$ - вероятность его отказа). Отказ любого из элементов приводит к прерыванию сигнала в той ветви цепи, где находится данный элемент. Вычислить надежность *p* каждой из схем.

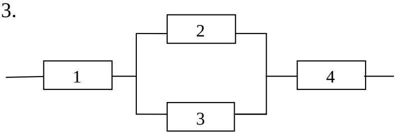
1.1.



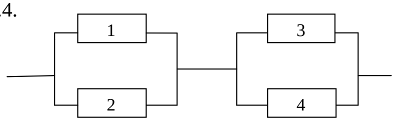
1.2.



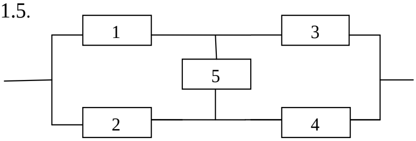
1.3.



1.4.

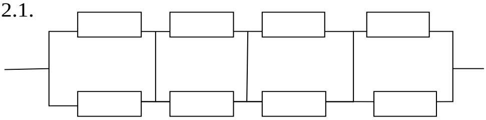


1.5.

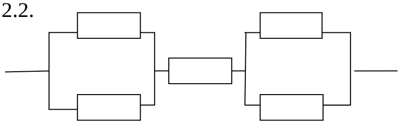


2. Вероятность безотказной работы в течение заданного времени (надежность) каждого элемента равна 0,8. Из этих элементов составлены три системы. Какова надежность каждой из систем?

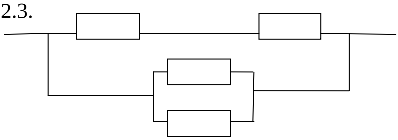
2.1.



2.2.



2.3.



3. Из урны, содержащей семь белых и три черных шара, наугад последовательно извлекают по одному шару до появления черного шара. Найти вероятность того, что придется извлечь четыре шара в предположении, что выбор производится: а) с возвращением; б) без возвращения.

4. Из полного набора костей домино (28 штук) наугад выбирают семь костей. Какова вероятность того, что среди них окажется меньшей мере одна кость с шестью очками?

5. Подбрасываются три игральных кости. Какова вероятность того, что на них выпадут разные грани?

6. В урне пять белых, семь красных и восемь синих шаров. Наугад выбрано два шара. Какова вероятность того, что шары одного цвета?

7. Два дуэлянта одновременно стреляют друг в друга. Для каждого вероятность убить противника равна 0,2. Какова вероятность того, что дуэль закончится гибелью одного из дуэлянтов?

8. Двадцать футбольных команд, среди которых четыре призера предыдущего первенства, по жеребьевке разбиваются на четыре занумерованных подгруппы по пять команд в каждой. Какова вероятность того, что в каждую подгруппу попадет по одному призеру? Какова вероятность того, что в первую подгруппу не попадет ни одного призера?

9. В первой урне два белых и три черных шара, во второй один белый и два синих шара, в третьей три белых и один красный шар. Из каждой урны наугад вынули по одному шару. Найдите вероятность следующих событий: A={вынут только один белый шар}; B={вынут хотя бы один белый шар}; C={вынуты шары разных цветов}.

10. Из колоды в 36 карт выбрали наугад две. Какова вероятность того, что обе карты красной масти?

11. В ящике находятся три неисправных лампочки и семь исправных. Лампочки извлекают наугад по одной и проверяют, пока не будут выбраны две исправные. Какова вероятность того, что придется проверить половину лампочек?

12. Из колоды карт одну за другой выбирают четыре карты. Какова вероятность того, что все они разных мастей? Какова вероятность того, что они все разного достоинства?

13. В ящике 10 теннисных мячей, в том числе семь новых и три побывавших в игре. Для игры наугад выбирают два мяча и после игры возвращают их обратно. Затем для второй игры также наугад извлекаются еще два мяча. Какова вероятность того, что вторая игра будет производиться новыми мячами?

14. На каждом крыле самолета по два мотора. Вероятность отказа в течение полета для каждого мотора равна p. Полет завершается благополучно, если на каждом крыле сохраняет работоспособность хотя бы один мотор. Какова вероятность благополучного полета?

15. Студент знает 20 вопросов из 30. Для сдачи экзамена достаточно правильно ответить на два предложенных вопроса или на один предложенный вопрос и один дополнительный. Какова вероятность того, что студент сдаст экзамен?

16. Два белых и два черных шара распределяют по двум урнам так, чтобы в каждой урне был хотя бы один шар. Затем наугад выбирают урну и из нее- один шар. Как следует распределить шары по урнам, чтобы вероятность вынуть белый шар (событие А) была наибольшей и какова эта вероятность?

17. Монету подбрасывают до тех пор, пока она два раза подряд не выпадет одной стороной. Найдите вероятность событий: А={опыт закончится до шестого броска}; В={понадобится более четырех бросков}.

18. Из колоды карт (36 штук) одну за другой наугад выбирают карты. Найдите вероятности следующих событий А={первый туз появится при третьем извлечении карты}; В={первый туз появится не ранее третьего извлечения карты}.

19. Из колоды карт (36 штук) выбирают карты по одной, пока не будет выбрана карта красной масти. Какова вероятность того, что придется выбрать более трех карт?

20. В урне восемь белых, шесть черных и два синих шара. Один за другим извлекаются наугад три шара. Какова вероятность того, что первый шар будет белым, второй- черным, третий – синим? Ответить на вопрос задачи в предположении: а) повторного выбора шаров; б) бесповторного выбора.

21. В первой урне четыре белых и два черных шара, а во второй урне три белых и пять черных шаров. Сначала подбрасывают монету и в зависимости от результата (герб или цифра) выбирают одну из урн, из которой затем наугад извлекают шар. Какова вероятность того, что шар окажется белым?

Ответы: 1.1. $1-q_1q_2q_3$; 1.2. $1-(1-p_1p_2p_3)(1-p_4p_5p_6)$; 1.3. $p_1p_4(1-q_2q_3)$; 1.4. $(1-q_1q_2)(1-q_3q_4)$; 1.5. $p_5(1-q_1q_2)(1-q_3q_4)+q_5(p_1p_3+p_2p_4-p_1p_2p_3p_4)$; 2.1. ≈0,85; 2.2. ≈0,74; 2.3. ≈0,98; 3. а) 1029/10000≈0,1, б) 1 /8; 4. 7/9; 5. 5/9; 6. 63/90≈1/3; 7. 0,32; 8. 125/969, 91/323; 9. $P(A)=5/12$, $P(B)=0,9$, $P(C)=31/60$; 10. 17/70; 11.3/20; 12. 3/32, 128/255; 13. ≈0,29; 14. $(1-p^2)^2$; 15. 152/203≈3/4; 16. Следует в одну урну положить белый шар, а остальные - во вторую урну; тогда $P(A)=2/3$. 17. $P(A)=15/16$, $P(B)=1/8$. 18. $P(A)=67/315\approx 0,21$, $P(B)=496/5355\approx 0,09$; 19. 4/35; 20. а) 3/128; б) 1/35. 21. 25/48.

ЗАНЯТИЕ № 4. ФОРМУЛА ПОЛНОЙ ВЕРОЯТНОСТИ И ФОРМУЛЫ БАЙЕСА

Формула полной вероятности является формулой, объединяющей теоремы сложения и умножения вероятностей:

$$P\ A = \sum_{i=1}^{n} P\ H_i \ \cdot P\left(A/H_i \right), \qquad (4.1)$$

где H_i $i = 1, 2, 3, ..., n$ такие, что $\sum_{i=1}^{n} H_i = \Omega$ и $H_i \cdot H_j = \varnothing$, $i, j = 1, 2, 3, ..., n$, $i \neq j$, $P(H_i) > 0$.

События H_i часто называют *гипотезами (предположениями).*

Естественно, что практическое использование формулы полной вероятности необходимо тогда, когда описание опыта содержит неопределенность, о которой можно сделать предположения $H_1, H_2, H_3, ..., H_n$ (ввести в рассмотрение гипотезы).

Пример 1. На складе имеется 1000 изделий: 700 из одной партии и 300 из другой. Вероятность брака в первой партии равна 0,06, во второй – 0,04. Какова вероятность того, что одно наудачу взятое изделие окажется бракованным?

Решение. В условии задачи не указано: из какой партии наудачу берется изделие. Поэтому возможны следующие гипотезы (предположения):

H_1– изделие принадлежит первой партии, H_2– изделие принадлежит второй партии.

Вероятности этих гипотез легко найти, используя классическое определение вероятности:

$$P(H_1)=\frac{700}{1000}=0,7,\ P(H_2)=\frac{300}{1000}=0,3.$$

Обозначим через А событие, состоящее в том, что наудачу извлеченное изделие окажется бракованным. Тогда из условия задачи ясно, что $P\left(A/H_1\right)=0,06$, а $P\left(A/H_2\right)=0,04$.

Следовательно, по формуле полной вероятности

$$P(A)=\sum_{i=1}^{2}P(H_i)\cdot P\left(A/H_i\right)=0,7\cdot0,06+0,3\cdot0,04=0,054.$$

Пример 2. По каналу связи, состоящему из передатчика, ретранслятора и приемника передается сигнал, состоящий из символов 0 и 1. Из-за помех символы независимо друг от друга могут искажаться: 0 переходит в 0 с вероятностью 0,96 и в 1 с вероятностью 0,04 на участке передатчик – ретранслятор и 0 переходит в 0 с вероятностью 0,95 и в 1 с вероятностью 0,05 на участке ретранслятор – приемник; 1 переходит в 1 с вероятностью 0,95 и в 0 с вероятностью 0,05 на участке передатчик – ретранслятор и 1 переходит в 1 с вероятностью 0,96 и в 0 с вероятностью 0,04 на участке ретранслятор – приемник. Какова вероятность того, что отправленный передатчиком сигнал 01 будет принят как 01?

Решение. В условии задачи не указано, искажаются ли символы или нет. Возможно 16 вариантов передачи сигнала 01, значит, введем в рассмотрение 16 гипотез:

$H_1:\ (01\to01\to01)$, $H_2:\ (01\to01\to00)$,

$H_3:\ (01\to01\to11)$, $H_4:\ (01\to01\to10)$,

$H_5:\ (01\to00\to01)$, $H_6:\ (01\to00\to10)$,

H_7 : $01 \to 00 \to 00$, H_8 : $01 \to 00 \to 11$,

H_9 : $01 \to 11 \to 10$, H_{10} : $01 \to 11 \to 01$,

H_{11} : $01 \to 11 \to 00$, H_{12} : $01 \to 11 \to 11$,

H_{13} : $01 \to 10 \to 11$, H_{14} : $01 \to 10 \to 10$,

H_{15} : $01 \to 10 \to 00$, H_{16} : $01 \to 10 \to 01$.

Вероятности этих гипотез находятся как вероятности произведения независимых событий:

$$P\ H_1\ = 0,96 \cdot 0,95 \cdot 0,95 \cdot 0,96 = 0,831744,$$

$$P\ H_2\ = 0,96 \cdot 0,95 \cdot 0,95 \cdot 0,04 = 0,034656.$$

Аналогично:

$P\ H_3\ = 0,043776$, $P\ H_4\ = 0,001824$, $P\ H_5\ = 0,00228$,

$P\ H_6\ = 0,090228$, $P\ H_7\ = 0,04332$, $P\ H_9\ = 0,0014592$,

$P\ H_{10}\ = 0,0014592$, $P\ H_{11}\ = 0,0006608$, $P\ H_{12}\ = 0,0350208$,

$P\ H_{13}\ = 0,000096$, $P\ H_{14}\ = 0,001824$, $P\ H_{15}\ = 0,000076$,

$P\ H_{16}\ = 0,000004$.

Правильность введения гипотез и вычисления их вероятностей можно проконтролировать.

Должно выполняться условие $\sum_{i=1}^{n} P\ H_i\ = 1$.

Если А – событие, состоящее в том, что приемником будет принят сигнал 01, то

$P\ A/H_1\ = P\ A/H_5\ = P\ A/H_{10}\ = P\ A/H_{16}\ = 1$, остальные $P\ A/H_k\ = 0$.

Таким образом, по формуле полной вероятности
$P(A) = 0,831744 + 0,0228 + 0,0014592 + 0,000004 = 0,8560072 \approx$
$\approx 0,86.$

Пример 3. Представим себе путника, который выходит из пункта O и на разветвлении дорог выбирает наугад один из возможных путей. Схема дорог изображена на рис. 2. Какова вероятность того, что путник, двигаясь описанным образом, попадет в пункт A?

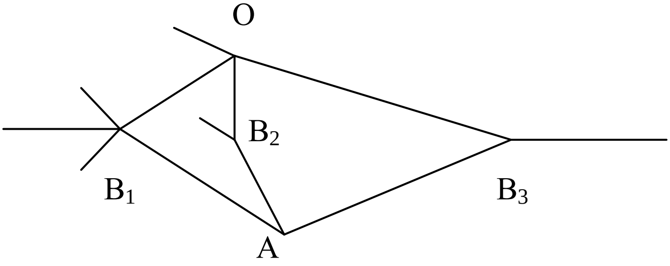


Рис. 2

Решение. Обозначим прибытие путника в пункт той же буквой, что и сам пункт. Легко видеть, что $P(B_1) = P(B_2) = P(B_3) = 1/4.$ Так как равновозможен выбор любого из путей, ведущих из пункта O. Путник попадет в пункт A, если он выберет дорогу в пункт B_1 и оттуда дорогу в пункт A, или он выберет дорогу в пункт B_2 и оттуда дорогу в пункт A, или он попадет сначала в пункт B_3 и оттуда – в пункт A. Символически это рассуждение можно записать в виде $A = B_1 \cdot A + B_2 \cdot A + B_3 \cdot A.$ Из чего следует, что

$$P(A) = P(B_1) \cdot P(A/B_1) + P(B_2) \cdot P(A/B_2) + P(B_3) \cdot P(A/B_3) =$$
$$= (1/4) \cdot (1/4) + (1/4) \cdot (1/2) + (1/4) \cdot (2/3) = 17/48 \approx 1/3,$$

Где вероятности $P(A/B_1), P(A/B_2), P(A/B_3)$ определены с учетом числа равновозможных путей, ведущих из соответствующего пункта.

Введенные, если этого требует условие задачи, гипотезы $H_i, i = 1, 2, 3, ..., n$ до опыта имеют определенные вероятности (априорные вероятности). Наблюдение результата опыта (появление события А) дает дополнительную информацию о гипотезах. Это приводит к перераспределению вероятностей гипотез.

Послеопытные (апостериорные) вероятности гипотез $P\left(H_i / A\right)$ могут быть найдены по формуле Байеса

$$P(H_i / A) = \frac{P(H_i)P(A/H_i)}{\sum_{i=1}^{n} P(H_i)P(A/H_i)} = \frac{P(H_i)P(A/H_i)}{P(A)}, \quad (4.2)$$

$P(A) > 0$,где $P\left(H_i\right)$ и $P\left(A/H_i\right)$ – априорные вероятности.

Пример 4. В магазин поступают однотипные изделия с трех заводов, причем первый завод поставляет 30% всех изделий, второй – 20%, а третий – 50%. Среди изделий первого завода 80% первосортных, второго – 90%, третьего – 70%. Куплено одно изделие. Оно оказалось первосортным. Какова вероятность того, что купленное изделие выпущено вторым заводом?

Решение. Так как неизвестно, изделие какого завода куплено, то введем в рассмотрение гипотезы:

H_1 – куплено изделие первого завода,

H_2 – куплено изделие второго завода,

H_3 – куплено изделие третьего завода.

Обозначим через А событие, состоящее в том, что куплено первосортное изделие.

Априорные вероятности гипотез равны $P\ H_1\ = 0,3$,

$P\ H_2\ = 0,2$, $P\ H_3\ = 0,5$.

Априорные условные вероятности появления события А при выполнении той или иной гипотезы равны $P\ A/H_1\ = 0,8$,

$P\ A/H_2\ = 0,9$, $P\ A/H_3\ = 0,7$.

Апостериорная вероятность того, что купленное первосортное изделие выпущено вторым заводом, найдем по формуле Байеса:

$$P(H_2/A) = \frac{P(H_2)\cdot P(A/H_2)}{\sum_{i=1}^{3} P(H_i)\cdot P(A/H_i)} = \frac{0,2\cdot 0,9}{0,3\cdot 0,8 + 0,2\cdot 0,9 + 0,5\cdot 0,7} =$$

$= 18/77 \approx 0,23.$

Пример 5. По каналу связи «0» и «1» с вероятностями 0,3 и 0,7 соответственно. Под действием помех возникают ошибки приема сигнала. Вероятность принять «1», если передавался «0», равна 0,05. Вероятность принять «0», если передавалась «1», равна 0,1. Какова вероятность правильного приема сигнала? Если принята «1», то какова вероятность того, что действительно передавалась «1»?

Решение. Обозначим передачу сигнала «1» через B_1, а сигнала «0» через B_2. Сигнал будет принят правильно (событие A), если будет передана «1» и она будет принята правильно или будет передан «0» и он будет принят правильно. Тогда $A = B_1\cdot A + B_2\cdot A$. В силу несовместности событий B_1 и B_2

имеем
$$P(A)=P(B_1\cdot A)+P(B_2\cdot A)=P(B_1)\cdot P(A/B_1)+P(B_2)\cdot P(A/B_2)=$$
$$=0,7\cdot 0,9+0,3\cdot 0,95=0,915.$$

Пусть событие C означает, что принят сигнал «1». При этом условии вычислим вероятность того, что действительно была передана «1», то есть вероятность $P(B_1\ /\ C)$. По формуле Байеса
$$P(B_1\ /\ C)=\frac{P(B_1)\cdot P(C\ /\ B_1)}{P(B_1)\cdot P(C\ /\ B_1)+P(B_2)\cdot P(C\ /\ B_2)}=$$
$$=\frac{0,7\cdot 0,9}{0,7\cdot 0,9+0,3\cdot 0,05}=0,9767441\approx 0,977.$$

Пример 6. Девяносто шесть процентов изделий некоторого производства удовлетворяют стандарту. Используется упрощенная схема проверки качества, дающая положительный результат с вероятностью 0,98, для стандартных изделий, для нестандартных изделий- с вероятностью 0,05. Изделие выдержало проверку. Какова вероятность того, что оно действительно стандартно?

Решение. Факт прохождения контроля обозначим через А. В отношении проверенного изделия можно выдвинуть два предположения: а) оно стандартно (событие B_1); б) оно нестандартно (событие B_2). По условию задачи необходимо найти вероятность $P(B_1\ /\ A)$. По формуле Байеса
$$P(B_1\ /\ A)=\frac{P(B_1)\cdot P(A/B_1)}{P(B_1)\cdot P(A/B_1)+P(B_2)\cdot P(A/B_2)}=$$
$$=\frac{0,96\cdot 0,98}{0,96\cdot 0,98+0,04\cdot 0,05}=\frac{2352}{2355}\approx 0,9988.$$

Пример 7. По каналу связи передается одна из последовательностей букв AAAA, BBBB, CCCC с вероятностями соответственно 0,5, 0,4, 0,1. Каждая передаваемая буква принимается правильно с вероятностью 0,8 и с вероятностями 0,1 и 0,1 за две другие буквы. Предполагается, что искажаются буквы при передаче независимо друг от друга. Найти вероятность того, что передано AAAA, если принято ABCA.

Решение. Для краткости записи формулы обозначим AAAA через T_1, BBBB через T_2, CCCCC через T_3. Тогда по формуле Байеса

$$P(T_1 / ABCA) =$$

$$\frac{P(T_1)\cdot P(ABCA/T_1)}{P(T_1)\cdot P(ABCA/T_1)+P(T_2)\cdot P(ABCA/T_2)+P(T_3)\cdot P(ABCA/T_3)} =$$

$$= \frac{0,5\cdot 0,8\cdot 0,1\cdot 0,1\cdot 0,8}{0,5\cdot 0,8\cdot 0,1\cdot 0,8+0,4\cdot 0,1\cdot 0,8\cdot 0,1+0,1\cdot 0,1\cdot 0,8\cdot 0,1} = \frac{8}{9}.$$

Задачи для самостоятельного решения

1. Из условий примера 3 стало известно, что путник пришел в пункт А. Какова вероятность того, что он попал в пункт А через пункт B_1?

2. Известно, что в партии из 10 изделий с равными вероятностями может оказаться от 0 до 2 изделий со скрытым дефектом. Проверили пять изделий, взятых наугад из этой партии. Среди проверенных не оказалось изделий с дефектом. Какова вероятность того, что в оставшейся половине партии нет изделий со скрытыми дефектами?

3. Три стрелка производят по одному выстрелу в одну и ту же мишень. Вероятности попадания в мишень при одном выстреле для этих стрелков соответственно равны 0,8,0,7, и 0,6.

Какова вероятность того, что третий стрелок промахнулся, если в мишени оказалось две пробоины?

4. Партия транзисторов, среди которых 10% дефектных, поступила на проверку. Схема проверки такова, что дефект обнаруживается с вероятностью 0,95. Вероятность того, что исправный транзистор будет признан дефектным, равна 0,03. Какова вероятность того, что выбранный наугад для проверки транзистор, будет признан дефектным?

5. В цехе болты изготавливают три автоматических станка A, B, и C. Станок A производит 25% болтов, станок B – 35% и станок C – 40% всех болтов. В продукции станков брак составляет соответственно 5%, 4% и 2%. Наугад взятый болт оказался дефектным. Какова вероятность того, что он был произведен станком A?

6. Из урны с четырьмя белыми и двумя черными шарами два шара, взятые наугад перенесены в урну с двумя белыми и тремя черными шарами. Какова после этого вероятность вынуть белый шар из второй урны?

7. Из урны, содержащей пять белых и три черных шара, наугад без возвращения выбирают три шара. Третий шар оказался белым. Какова вероятность того, что первые два шара были разного цвета?

8. Первая урна содержит один черный и один красный шар, а во второй урне два черных и три красных шара. Из первой урны вынули наугад один шар и переместили его во вторую урну. Затем вынули из второй урны наугад один шар. Какова вероятность того, что были вынуты шары одного цвета? Какова вероятность того, что из первой урны был вынут красный шар, если из второй урны был вынут черный шар?

9. Стрелок попадает в цель с вероятностью 1/при каждом выстреле. Монету подбрасывают три раза, и стрелку предоставляется столько выстрелов, сколько раз выпал герб. Найдите вероятность поражения цели.

10. Из урны, в которой было четыре белых и три черных шара, убрали один из шаров неизвестного цвета. Для определения состава урны наугад выбрали два шара. Они оказались черными. Какова вероятность того, что первый вынутый шар был белым?

11. По каналу связи передается одна из двух команд управления в виде кодовых комбинаций 11111 или 00000, причем вероятности этих комбинаций равны соответственно 0,7 и 0,3. Из-за наличия помех вероятность правильного приема каждого символа 0 или 1 равна 0,6 независимо от правильности приема остальных символов. На приемном устройстве получена комбинация 10110. Какова вероятность того, что была передана комбинация 11111?

Ответы:

1. $P(B_1 / A) = 3/17,\ P(B_2 / A) = 6/17,\ P(B_3 / A) = 8/17$;

2. 18/31≈3/5; 3. 56/113≈1/2; 4. 0,122; 5. 25/69≈0,36; 6. 10/21; 7. 4/7; 8. 7/12, 2/5; 9. 0,261≈1/4; 10. 0,8; 11. 0,78.

ЗАНЯТИЕ № 5. ПОСЛЕДОВАТЕЛЬНОСТЬ НЕЗАВИСИМЫХ ИСПЫТАНИЙ. СХЕМА БЕРНУЛЛИ

Самой простой моделью повторения испытаний являются независимые испытания, в каждом из которых вероятность появления некоторого события A постоянна и равна p. Тогда вероятность того, что в n опытах событие A появится ровно

m раз, определяется формулой Бернулли

$$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m} = \frac{n!}{m! \cdot (n-m)!} \cdot p^m \cdot q^{n-m}, \text{ где } q = 1 - p.$$

При больших n вычисления по формуле Бернулли становятся громоздкими и в этих случаях для вычисления вероятности появления события A ровно m раз в n независимых испытаниях используется локальная теорема Лапласа

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \phi(x), \text{ где } x = \frac{m - np}{\sqrt{npq}}, \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

или формула Пуассона $P_n(m) \approx \dfrac{a^m}{m!} \cdot e^{-a}$, где $a = np$.

Можно рекомендовать пользоваться теоремой Лапласа, если $npq > 9$, и формулой Пуассона, если $npq < 9$.

Если требуется найти вероятность того, что в n независимых испытаниях событие A появится не менее m_1, но не более чем m_2 раз, то используют формулы

$$P_n(m_1 \le m \le m_2) = \sum_{m=m_1}^{m_2} C_n^m \cdot p^m \cdot q^{n-m},$$

$$P_n(m_1 \le m \le m_2) = \Phi\left(\frac{m_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{m_1 - np}{\sqrt{npq}}\right),$$

где $\Phi(x) = \dfrac{1}{2\pi} \displaystyle\int_0^x e^{-\frac{x^2}{2}} dx$,

$$P_n(m_1 \le m \le m_2) = \sum_{m=m_1}^{m_2} \frac{a^m}{m!} \cdot e^{-a}.$$

Для функций $\phi\ \ x\ \ = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2}$ и $\varPhi\ \ x\ \ = \dfrac{1}{2\pi}\int\limits_0^x e^{-\frac{x^2}{2}}dx$ составлены таблицы.

Значение $m = m_0$, при котором вероятность $P_n(m)$ принимает наибольшее значение, называется *наивероятнейшим числом успехов*.

$m_0 = \ (n+1)p$, где $\ .\ $ – символ целой части числа.

Если $(n+1)p$ – целое число, то m_0 принимает два значения $m_0^{(1)} = (n+1)p - 1$ и $m_0^{(2)} = (n+1)p$.

Пример 1. Вероятность попадания в мишень при каждом выстреле равна 0,75. Какова вероятность того, что при пяти выстрелах будет ровно три попадания

Решение. Воспользуемся формулой Бернулли.

$n = 5, m = 3, p = 0{,}75, q = 1 - 0{,}75 = 0{,}25.$

$$P_5\ \ 3\ \ = C_5^3 \cdot 0{,}75^3 \cdot 0{,}25^2 = \frac{135}{512} \approx 0{,}26.$$

Пример 2. Вероятность попадания в объект равна 0,75. Для разрушения объекта необходимо не менее трех попаданий. Произведено пять выстрелов. Какова вероятность того, что объект будет разрушен?

Решение. Вероятность события А, состоящего в том, что объект будет разрушен, равна

$$P(A) = P_5(3 \le m \le 5) = \sum_{m=3}^{5} C_5^m \cdot 0{,}75^m \cdot 0{,}25^{5-m} = C_5^3 \cdot 0{,}75^3 \cdot 0{,}25^2 +$$

$$+C_5^4 \cdot 0{,}75^4 \cdot 0{,}25^1 + C_5^5 \cdot 0{,}75^5 \cdot 0{,}25^0 = \frac{459}{512} \approx 0{,}9.$$

Пример 3. Вероятность сбоя в работе телефонной станции при каждом вызове равна 0,012. Поступило 1000 вызовов. Какова вероятность 9 сбоев?

Решение. Так как число опытов $n = 1000$ велико, то воспользуемся локальной теоремой Лапласа

$$P_{1000}(9) \approx \frac{1}{\sqrt{1000 \cdot 0,012 \cdot 0,988}} \cdot \phi\left(\frac{9 - 1000 \cdot 0,012}{\sqrt{1000 \cdot 0,012 \cdot 0,988}}\right) \approx$$

$$\approx \frac{1}{3,43} \cdot \phi(-0,875).$$

По таблице находим $\phi\ \ -0,875\ \ = \phi\ \ 0,875\ \ = 0,2732.$

Окончательно $P_{1000}\ \ 9\ \ \approx \dfrac{0,2732}{3,43} \approx 0,07965.$

Пример 4. Вероятность выхода из строя за время T одного конденсатора равна 0,2. Определить вероятность того, что за время T из 100 конденсаторов выйдут из строя от 14 до 26 конденсаторов.

Решение. Для этой задачи математической моделью является схема Бернулли. Здесь $n = 100$, $p = 0,2$, $q = 1 - p = 1 - 0,2 = 0,8$, $npq = 100 \cdot 0,2 \cdot 0,8 = 16 > 9$.

Согласно теореме Муавра-Лапласа

$$P\left(x_1 \le \frac{m - np}{\sqrt{npq}} \le x_2\right) = F^*(x_2) - F^*(x_1), \text{ где}$$

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

Тогда

$$P\{14 \le m \le 26\} = \left| \begin{matrix} np = 100 \cdot 0,2 = 20 \\ \sqrt{npq} = \sqrt{100 \cdot 0,2 \cdot 0,8} = 4 \end{matrix} \right| =$$

$$= P\{14 - 20 \le m - np \le 26 - 20\} =$$

$$= P\{-6 \le m - np \le +6\} = P\left\{-\frac{6}{4} \le \frac{m - np}{\sqrt{npq}} \le \frac{6}{4}\right\} =$$

$$= F^*(1,5) - F^*(-1,5) = \left|F^*(-x) = 1 - F^*(x)\right| = 2F^*(1,5) - 1 =$$

$$= 2 \cdot 0,933 - 1 = 0,866.$$

Пример 5. Предположим, что 30% студентов нашего университета занимаются спортом. Какова вероятность того, что среди первых пяти встреченных студентов окажется только один спортсмен? Какова вероятность того, что среди них есть хотя бы один спортсмен? Каково наиболее вероятное число спортсменов среди них?

Решение. Так как студентов в университете много (несколько тысяч), то по мере опроса нескольких из них пропорции в оставшейся части практически не изменяются. Поэтому можно считать опрос каждого студента независимым опытом. Всего опытов производится $n = 5$, а вероятность положительного ответа $p = 0,3$. По формуле Бернулли имеем $P_5(1) = C_5^1 \cdot 0,3 \cdot (0,7)^4 = 0,36015$. Вероятность хотя бы одного правильного ответа проще вычислять, если перейти к противоположному событию:

$$P_5(k \ge 1) = 1 - P_5(0) = 1 - (0,7)^5 = 1 - 0,16807 = 0,83193.$$

Так как $(n+1)p = (5+1)0,3 = 1,8$ (целая часть числа равна 1), то наиболее вероятное число спортсменов среди пяти опрошенных $k_0 = 1$.

Пример 6. На каждый вопрос предлагается три ответа, среди которых следует выбрать один правильный. Задано пять вопросов. Какова вероятность того, что путем простого угадывания удастся правильно ответить на четыре вопроса? Какова вероятность угадать правильный ответ хотя бы на один вопрос?

Решение. Выбор ответа на вопрос можно рассматривать как независимый опыт. Всего таких опытов производится $n=5$, а вероятность успеха в каждом опыте равна $p=1/3$. Тогда вероятность путем простого угадывания правильно ответить на четыре вопроса равна

$$P_5(4)=C_5^4\cdot(1/3)^4\cdot(2/3)^1=10/243\approx 1/24.$$

Вероятность угадать хотя бы один правильный ответ равна

$$P_5(k\geq 1)=1-P_5(0)=1-(2/3)^5=1-32/243\approx 7/8.$$

Пример 7. Вероятность попадания в цель при выстреле равна 0,3. Сколько нужно сделать выстрелов, чтобы вероятность поражения цели была больше 0,9?

Решение. Каждый выстрел можно рассматривать как независимое испытание, и в каждом из них вероятность появления события (попадания в цель) равна $p=0,3$ Цель будет поражена, если в n выстрелах будет хотя бы одно попадание, вероятность чего равна:

$$P_n(k\geq 1)=1-P_n(0)>0,9\Rightarrow 1-(0,7)^n>0,9\Rightarrow(0,7)^n<0,1\Rightarrow n\geq 7.$$

Пример 8. Известно, что наборщик в среднем допускает одну ошибку на две страницы текста. В набранной книге взяли наугад страницу. Какова вероятность того, что на этой странице содержится более одной опечатки?

Решение. Опечатки появляются по одной и независимо друг от друга. Условия простейшего потока приблизительно выполняются, и формула Пуассона приблизительно верна. На одну страницу приходится в среднем $\lambda = 1/2$ опечатки. Поэтому вероятность того, что на данной странице содержится более одной опечатки, равна

$$P(k>1) = 1 - P(k=0) - P(k=1) = 1 - \frac{(1/2)^0}{0!} e^{-1/2} - \frac{(1/2)^1}{1!} e^{-1/2} \approx 0{,}1.$$

Задачи для самостоятельного решения

1. Игральный кубик подброшен пять раз. Какова вероятность того, что два раза выпадет шесть очков?

2. Вероятность рождения мальчика равна ½. Какова вероятность того, что в семье с четырьмя детьми два мальчика и две девочки?

3. Первый стрелок попадает в цель с вероятностью 1/3 и производит четыре выстрела. Второй стрелок попадает в цель с вероятностью ½ и производит три выстрела. Для какого стрелка в этих условиях вероятнее попасть в цель дважды?

4. Прибор состоит из пяти блоков. Надежность (вероятность безотказной работы в течение заданного времени t) для каждого блока равна 0,9. Узлы выходят из строя независимо друг от друга. Найти вероятность того, что за время t: а) откажет только один узел; б) откажет хотя бы один узел; в) откажут не менее двух узлов.

5. Вероятность выпуска детали со скрытым дефектом равна 0,02. Детали упаковываются в ящик по 100 штук. Какова вероятность того, что в данном ящике нет деталей со скрытым дефектом? Какова вероятность того, что в данном ящике больше двух деталей со скрытым дефектом? Каково наиболее вероятное число деталей со скрытым дефектом в дано ящике?

6. В крупной партии изделий 10% имеют низкое качество. Сколько нужно проверить изделий, чтобы с вероятностью не менее 0,95 обнаружить хотя бы одно изделие низкого качества?

7. В течение часа на коммутатор поступает в среднем 120 телефонных вызовов. Какова вероятность того, что в течение заданной минуты поступит четыре вызова?

8. Вероятность того, что при перевозке изделие повредится, равна $p = 0,005$. С завода отправлено четыреста изделий. Найдите вероятность того, что в пути повредится более двух изделий.

9. По каналу связи передается цифровой текст. Из-за помех каждая цифра может быть принята неправильно с вероятностью $p = 0,0025$. Какова вероятность того, что в тексте, состоящем из 800 цифр, все цифры будут приняты правильно?

10. При стрельбе из зенитного орудия в среднем только один выстрел из 200 достигает цели. Какова вероятность того, что при 100 выстрелах будет хотя бы одно попадание?

11. Дальтоники составляют 1% населения. Какова вероятность того, что среди пятидесяти студентов окажется по меньшей мере один дальтоник?

12. Двое бросают монету по пять раз каждый. Какова вероятность того, что у них выпадет одинаковое число гербов?

13. Частица в начальный момент времени находится в начале координат. Каждую последующую секунду она сдвигается вправо с вероятностью 1/3 или влево с вероятностью 2/3 независимо от того, как она двигалась в предыдущие секунды. Какова вероятность того, что через пять секунд частица окажется на отрезке [-2;0]?

14. В биатлоне на каждом из трех огневых рубежей спортсмен должен поразить пять мишеней в пяти выстрелах. За каждую непораженную мишень спортсмен обязан пробежать штрафной круг. Пусть вероятность поражения мишени при одном выстреле равна 0,95. Какова вероятность того, что спортсмен все три огневых рубежа пройдет без штрафных кругов? Какова вероятность того, что после каждого огневого рубежа спортсмен будет пробегать один штрафной круг?

15. Вероятность попадания в цель при одном выстреле равна 1/3. Стрелок стреляет по цели до тех пор, пока не накопится три попадания. Какова вероятность того, что к этому моменту у стрелка будет два промаха?

Ответы: 1. 1250/7776≈0,16; 2. 3/8; 3. Для второго;
4. а) 0,328, б) 1-$(0,9)^5$≈0,4, в)≈0,08;
5. $e^{-2} \approx 0,14$, $1-5e^{-2} \approx 0,3$, $k_0 = 2$; 6. $n \geq 29$;
7. $(2e^{-2})/3 \approx 0,09$; 8. $1-5e^{-2} \approx 0,31$; 9. $e^{-2} \approx 0,136$;
10. $1-e^{-1/2} \approx 0,4$; 11. $1-e^{-1/2} \approx 0,4$; 12. 63/256≈1/4;
13. 80/243≈1/3; 14. ≈0,46, ≈0,008; 15. 8/81.

ЗАНЯТИЕ № 6. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ И ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

Случайной величиной ξ называется функция $\xi = \xi(\omega)$, определенная на пространстве элементарных событий Ω, удовлетворяющая условию: для любых чисел a, b $(a < b)$ можно посчитать вероятность попадания случайной величины ξ в интервал (a, b).

Функцией распределения вероятности случайной величины ξ называется функция $F(x) = P(\xi < x)$, $|x| < \infty$.

Функция распределения случайной величины ξ $F(x)$ есть неубывающая функция; $F(-\infty) = 0$, $F(+\infty) = 1$, $0 \le F(x) \le 1$, $P(x_1 \le \xi < x_2) = F(x_2) - F(x_1)$.

Случайные величины, принимающие дискретное множество значений, называются *дискретными случайными величинами*.

Примеры дискретных распределений

1. Биномиальное распределение

$$P(\xi = k) = C_n^k p^k (1-p)^{n-k} \qquad 0 < p < 1, \qquad k = 0, \ldots, n.$$

2. Распределение Пуассона

$$P(\xi = k) = \frac{a^k}{k!} e^{-a}, \qquad a > 0, \qquad k = 0, 1, 2, \ldots .$$

3. Геометрическое распределение

$$P(\xi = k) = p(1-p)^{k-1}, \qquad 0 < p < 1, \qquad k = 1, 2, \ldots .$$

Вероятностной характеристикой дискретной СВ является закон распределения.

Законом распределения ДСВ называется перечень значений, которые принимает эта СВ, и соответствующих им вероятностей.

На практике закон распределения ДСВ задается в виде табл. 1, которую называют еще *рядом распределения*.

Таблица 1

X	x_1	x_2	…	x_k	…
P(X=x_k)	p_1	p_2	…	p_k	…

Зная закон распределения ДСВ, можно найти функцию распределения с помощью равенства

$$\mathrm{F}(x) = \sum_{k:\ x_k < x} \mathrm{P}(\mathrm{X} = x_k).$$

Суммирование в этом равенстве распространяется на все те индексы k, для которых $x_k<x$. Так как $P(X=x_k)=F(x_k+0)-F(x_k)$, $x_k \in \{x_1, x_2, ...\}$, то очевидно, что функция распределения имеет скачки при тех значениях x, которые являются ее возможными значениями. Величина скачка в точке $x=x_k$ как раз равна вероятности СВ принятия данного значения.

Для дискретных случайных величин функция распределения кусочно-постоянна, непрерывная слева, имеет разрывы 1 рода в точках $x = x_k$, и величина скачка равна p_k.

Непрерывной называется случайная величина ξ, функцию распределения которой $F(x)$ можно представить в виде

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

Функция $f(x)$ называется *плотностью распределения* вероятностей случайной величины ξ. Всюду в дальнейшем будем считать $f(x)$ непрерывной функцией.

Свойства функции распределения и плотности распределения вероятности непрерывной случайной величины:

1) $f(x) \geq 0$.

2) $\int_{-\infty}^{\infty} f(x)dx = 1$ *(условие нормировки).*

3) $f(x) = F'(x)$.

4) $P(x_1 < \xi < x_2) = \int_a^b f(x)dx = F(x_2) = F(x_1)$.

Примеры непрерывных распределений

1) Равномерное распределение

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \le x < b \\ 0, & x < a,\, x \ge b \end{cases}$$

2) Нормальное распределение (с параметрами (m, σ^2))

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \qquad |m| < \infty,\ \sigma > 0.$$

Запись $\xi \in N(m, \sigma^2)$ означает, что случайная величина ξ распределена нормально с параметрами m и σ^2.

3) Показательное распределение

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

4) Распределение Коши

$$f(x) = \frac{1}{\pi(1+x^2)}.$$

5) Распределение Релея

$$f(x) = \begin{cases} \dfrac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, & x > 0 \\ 0, & x \le 0. \end{cases}$$

6) Гамма-распределение с параметрами $\lambda > 0$, $\gamma > 0$

$\xi \in \Gamma(\lambda, \gamma)$

$$f(x)=\begin{cases} 0, & x\le 0, \\ \dfrac{\lambda^{\gamma}x^{\gamma-1}}{\Gamma(\gamma)}e^{-\lambda x}, & x>0. \end{cases}$$

Здесь $\Gamma(\gamma)$ – гамма-функция.

Математическим ожиданием случайной величины ξ называется число

$$M\xi=\begin{cases} \sum\limits_{k=1}^{n}x_k p_k & \text{для дискретной случайной величины,} \\ \int\limits_{-\infty}^{\infty}xf(x)dx & \text{для непрерывной случайной величины.} \end{cases} \qquad (6.1)$$

Говорят, что математическое ожидание у случайной величины существует, если ряд (интеграл) (1) сходится абсолютно.

Дисперсией случайной величины $D\xi$ называется число

$$D\xi=M\left(\xi-M\xi\right)^2. \qquad (6.2)$$

Дисперсия вычисляется по формулам:

$$D\xi=\sum_{i=1}^{n}(x_i-m)^2 p_i=\sum_{i=1}^{n}x_i^2 p_i-m^2 \qquad (6.3)$$

для дискретной случайной величины.

$$D\xi=\int_{-\infty}^{\infty}(x-m)^2 f(x)dx=\int_{-\infty}^{\infty}x^2 f(x)dx-m^2 \qquad (6.4)$$

для непрерывной случайной величины, где $m=M\xi$.

Пример 1. Изделия испытываются при перегрузочных режимах. Вероятности для каждого изделия пройти испытания равны 0,8 и независимы. Испытания заканчиваются после первого же изделия, не выдержавшего испытания. Вывести формулу для ряда распределения числа испытаний.

Решение. Испытания заканчиваются на *k*-ом изделии, если первые (*k*-1) изделия пройдут испытания, а *k* – изделие не пройдет. Если X – случайное число испытаний, то

$P(X = k) = (1 - 0,8) \cdot (0,8)^{k-1} = 0,8^{k-1} \cdot 0,2$, $k = 1, 2, \ldots$.

ЗРВ будет иметь вид

X	1	2	3	…	k	…
p	$0,2$	$0,2 \cdot 0,8$	$0,2 \cdot 0,8^2$		$0,2 \cdot 0,8^{k-1}$	

$$\sum_{k=1}^{6} p_k = 1.$$

Пример 2. В урне имеется четыре шара с номерами от 1 до 4. Извлекли два шара. Найти закон распределения и функцию распределения СВ X суммы номеров извлеченных шаров.

Решение. Ω={(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)}.

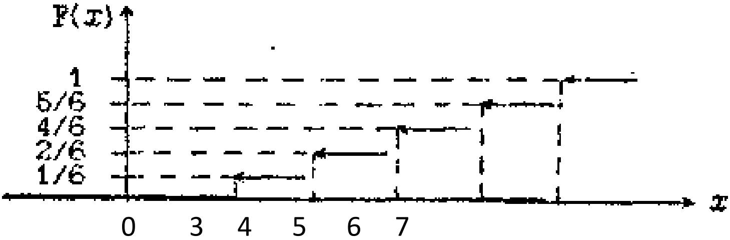
Закон распределения имеет вид:

X	3	4	5	6	7
P(X=x_k)	1/6	1/6	2/6	1/6	1/6

$$\sum_{k=1}^{6} p_k = 1.$$

$$\text{Функция распределения } F(x) = \begin{cases} 0, & x \le 3, \\ 1/6, & 3 < x \le 4, \\ 2/6, & 4 < x \le 5, \\ 4/6, & 5 < x \le 6, \\ 5/6, & 6 < x \le 7, \\ 1, & x > 7. \end{cases}$$

График F(x) имеет вид



Пример 3. Вероятность попадания стрелка при одном выстреле p=0,3. Найти закон распределения для СВ X - числа попаданий при 3 выстрелах.

Решение. Обозначим попадание через Y, непопадание - Н. Тогда
Ω={(ННН),(YНН),(НYН),(ННY),(YYН),(YНY),(НYY),
(YYY)},

X	0	1	2	3
$P_n(X=k)$	q^3=0,343	C_3^1 pq^2=0,441	C_3^2 p^2q=0,189	p^3=0,227

$$\sum_{k=1}^{3} P(X=k) = 1.$$

Пример 4. Случайная величина X распределена по закону, определяемому плотностью распределения вероятностей вида

$$f(x)=\begin{cases} C\cos x, \text{если } -\dfrac{\pi}{2}\le x\le \dfrac{\pi}{2}, \\ 0, \text{ если } |x|>\dfrac{\pi}{2}. \end{cases}$$

Найти константу C, вычислить $P\left\{|x|<\dfrac{\pi}{4}\right\}$, m_x, D_x.

Решение. 1. Для нахождения C воспользуемся свойством нормировки $\int\limits_{-\infty}^{\infty} f(x)dx=1$. Или $\int\limits_{-\frac{\pi}{2}}^{\frac{\pi}{2}} C\cos x dx=1,\ C\sin x\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}}=1,$

$2C=1,\ C=\dfrac{1}{2}.$

2. $P\left\{|x|<\dfrac{\pi}{4}\right\}=P\left\{-\dfrac{\pi}{4}<x<\dfrac{\pi}{4}\right\}=\dfrac{1}{2}\int\limits_{-\frac{\pi}{4}}^{\frac{\pi}{4}}\cos x dx=\dfrac{1}{2}\sin x\Big|_{-\frac{\pi}{4}}^{\frac{\pi}{4}}=$

$=\dfrac{1}{2}\left(\dfrac{\sqrt{2}}{2}+\dfrac{\sqrt{2}}{2}\right)=\dfrac{\sqrt{2}}{2}.$

3. $m_x=\dfrac{1}{2}\int\limits_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x\cos x dx=0.$

4. $D_x=\int\limits_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x^2 f(x)dx-(m_x)^2=\dfrac{1}{2}\int\limits_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x^2\cos x dx=\int\limits_{0}^{\frac{\pi}{2}} x^2\cos x dx=$

$$= x^2 \sin x \Big|_0^{\frac{\pi}{2}} - 2\int_0^{\frac{\pi}{2}} x \sin x dx = \frac{\pi^2}{4} - 2(-x\cos x\Big|_0^{\frac{\pi}{2}} + \int_0^{\frac{\pi}{2}} \cos x dx) =$$

$$= \frac{\pi^2}{4} - 2\sin x\Big|_0^{\frac{\pi}{2}} = \frac{\pi^2}{4} - 2.$$

Пример 5. Время безотказной работы некоторого узла сложного агрегата – экспоненциальная случайная величина со средним $M = 2$. Для увеличения надежности агрегата узел дублируется – ставят параллельно несколько одинаковых, но функционирующих независимо узлов. Сколько узлов следует запараллелить, чтобы с вероятностью, не меньшей чем 0,9, по крайней мере один из них не вышел из строя за 10 часов работы?

Решение. T – случайное время безотказной работы узла – имеет экспоненциальное распределение. Это означает, что

$$P\ T < t\ = F_T(t) = \begin{cases} 0, & t < 0, \\ 1 - \exp\ -\mu \cdot t\ , & t \geq 0. \end{cases}$$

Известно, что математическое ожидание экспоненциальной случайной величины есть величина, обратная параметру: $M = 1/\mu \Rightarrow \mu = 0,5$. Следовательно, вероятность отказа узла в течение 10 часов будет равна

$$P\ T < 10\ = F_T(10) = 1 - \exp-\ \ 0,5 \cdot 10\ \approx 0,9933.$$

Если запараллелено N идентичных узлов, то событие $A=$ {по крайней мере один из узлов не вышел из строя за 10 часов} является противоположным событию $\bar{A}=$ {все узлы вышли из строя за 10 часов}. Поэтому $P\ A\ = 1 - P\ \bar{A}\ $. Для последней вероятности (в силу независимости отказов запараллеленных узлов) получаем

$$P\left(\bar{A}\right)=P\left(T<10\right)^{N}.$$

Искомое количество N может теперь быть найдено как наименьшее целое решение неравенства

$P\left(A\right)=1-P\left(\bar{A}\right)=1-P\left(\bar{A}\right)=1-\left(P\left(T<10\right)\right)^{N}=1-(0,9933)^{N}\geq 0,9\Rightarrow N\geq 343$.

Задачи для самостоятельного решения

1. Некто имеет в связке пять ключей. При отмыкании замка он последовательно испытывает ключи, пока не подберет нужный. Полагая выбор ключей бесповторным, написать закон распределения числа испытанных ключей. Подсчитать математическое ожидание этой случайной величины и построить ее функцию распределения.

2. На электронное реле воздействует случайное напряжение, имеющее плотность вероятности $f\left(x\right)=\dfrac{x}{a^{2}}\cdot\exp\left\{-\dfrac{x^{2}}{2\sigma^{2}}\right\}, x\geq 0.$ Реле срабатывает всякий раз, когда напряжение на его входе превышает 3В. Какова вероятность срабатывания реле?

3. В кошельке было 5 монет по 10 копеек и три монеты по 50 копеек. Из кошелька вынули наугад четыре монеты. Найдите закон распределения случайной величины X, которая равна сумме вынутых копеек.

4. Время безотказной работы предохранителя имеет показательный закон распределения с функцией плотности вероятности $f\left(x\right)=0,002e^{-0,002x}$ при $x\geq 0$, $f\left(x\right)=0$ при $x<0$. Найдите функцию распределения времени безотказной работы. Найди-

те вероятность того, что предохранитель безотказно проработает 1000 часов.

5. Случайная величина X равна сумме выпавших очков на двух игральных кубиках. Напишите ее закон распределения и найдите ее математическое ожидание.

6. В урне лежат два черных три белых шара. Из этой урны вынимаются один за другим без возвращения шары до тех пор, пока не будет вынут черный шар. Найдите среднее число вынутых при этом шаров.

7. Случайная величина X имеет функцию плотности вероятности $f(x)=0,5\sin x$ при $x\in(0,\pi)$ и $f(x)=0$ при остальных x. Найдите: а) функцию распределения величины X; б) математическое ожидание этой величины; в) вероятность попадания в интервал $(0,\pi/3)$.

8. Монету бросают до первого выпадения герба, либо до тех пор, пока цифра не выпадет четыре раза подряд. Найдите среднее число бросков монеты.

9. Случайная величина X имеет функцию плотности вероятности $f(x)=\frac{1}{\pi}\cdot\frac{1}{1+x^2}$ (закон распределения Коши). Какова вероятность того, что при трех независимых наблюдениях этой случайной величины будут наблюдаться значения только из интервала $(-1;1)$?

10. Дан ряд распределения случайной величины *X*. Найти:
а) *M(X), D(X), σ(X), M(3X-2);*
б) функцию распределения *F(x)* и построить ее график.

X	-5	-3	1	4
p	0,2	0,3	0,4	0,1

11. В билете три задачи. Вероятность правильного решения первой задачи равна 0,9; второй – 0,8; третьей – 0,7. Составить закон распределения числа правильно решенных задач в билете и вычислить математическое ожидание и дисперсию этой случайной величины.

12. Дана функция распределения случайной величины *X*. Найти: а) плотность распределения вероятностей (дифференциальную функцию) *f(x);* б) построить графики *F(x)* и *f(x);* в) вычислить математическое ожидание *M(X),* моду $M_0(X)$ и медиану *Me(X);* г) найти $P(1<X<3)$.

$$F(x)=\begin{cases} 0, & \text{если } x\le 0;\\ \dfrac{x^2}{16}, & \text{если } 0<x\le 4;\\ 1, & \text{если } x>4. \end{cases}$$

13. Случайная величина X имеет плотность вероятности $f(x)$.

$$f(x)=\begin{cases} A\sin x, & \text{если } x\in(0;\pi];\\ 0, & \text{если } x\notin(0;\pi]. \end{cases}$$

Найти: а) параметр А; б) функцию распределения F(x).

14. Текущая цена акции может быть смоделирована с помощью нормального закона распределения с математическим ожиданием 15 ден. ед. и средним квадратическим отклонением 0,2 ден. ед. Найти вероятность того, что цена акции не выше 15,3 ден. ед.

Ответы

1. $M(X)=3;$; 2. $\exp\left(-9/2\sigma^2\right)$;

3. $P\left(X=40\right)=\dfrac{1}{14}, P\left(X=80\right)=\dfrac{3}{7}, P\left(X=120\right)=\dfrac{3}{7},$
$P\left(X=160\right)=\dfrac{1}{14};$

4. $F\left(x\right)=1-e^{-0{,}002x}$ при $x\ge 0,\ F(x)=0$, при $x<0, P=e^{-2}\approx 0{,}17;$ 5. $M\left(\xi\right)=7;$ 6. 2; 7. а) $F\left(x\right)=0$ при $x<0, F\left(x\right)=0{,}5\left(1-\cos x\right)$ при $x\in\left(0,\pi\right), F\left(x\right)=1$ при $\pi<x,$ б) $M\left(\xi\right)=\pi/2,$ в) $P\left(0<\xi<\pi/3\right)=1/4;$ 8. 15/8; 9. 1/8.

ЗАНЯТИЕ № 7. НАХОЖДЕНИЕ ЧИСЛОВЫХ ХАРАКТЕРИСТИК ДИСКРЕТНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН, ПРИНИМАЮЩИХ ЦЕЛОЧИСЛЕННЫЕ ЗНАЧЕНИЯ МЕТОДОМ ПРОИЗВОДЯЩЕЙ ФУНКЦИИ. РАСПРЕДЕЛЕНИЕ ПУАССОНА. НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Пусть СВ X принимает конечное или счетное число целочисленных значений. В соответствие заданному распределению ставится функцию

$$\Phi\left(z\right)=\sum_{n=0}^{\infty}p_n z^n ,$$

которая называется *производящей функцией* для заданного распределения вероятностей.

Тогда математическое ожидание и дисперсия могут быть найдены по формулам

$$\mathrm{M}(\mathrm{X}) = \Phi'(1),\ \mathrm{D}(\mathrm{X}) = \Phi''(1) + \mathrm{M}(\mathrm{X}) - \left(\mathrm{M}(\mathrm{X})\right)^2.$$

Случайная величина X называется распределенной по закону Пуассона с параметром $\lambda>0$, если ее возможные значения равны 0, 1, 2,..., а соответствующие вероятности определяются формулой

$$P(\mathrm{X}=k) = \frac{\lambda^k}{k!}e^{-\lambda}.$$

Закон распределения имеет вид

X	0	1	2	…	k	…
$P(\mathrm{X}=k)$	$\mathrm{e}^{-\lambda}$	$\frac{\lambda}{1!}\mathrm{e}^{-\lambda}$	$\frac{\lambda^2}{2!}\mathrm{e}^{-\lambda}$	…	$\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$	…

Для того, чтобы убедиться, что это действительно закон распределения, достаточно проверить, что сумма вероятностей всех возможных значений равна единице:

$$\sum_{k=0}^{\infty}\mathrm{P}(\mathrm{X}=k) = \sum_{k=0}^{\infty}\frac{\lambda^{\mathrm{k}}}{k!}\mathrm{e}^{-\lambda} = \mathrm{e}^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^{\mathrm{k}}}{k!} = \mathrm{e}^{-\lambda}\mathrm{e}^{\lambda} = 1.$$

Распределение Пуассона может быть получено из биномиального распределения путем предельного перехода, когда число испытаний стремится к бесконечности, а вероятность успеха стремится к нулю при условии, что $np=\lambda=\mathrm{const}$, то есть $n\to\infty$, $p\to 0$, $np=\lambda=\mathrm{const}$.

При больших n и малых p формулу Пуассона можно использовать в качестве приближения вместо формулы Бернулли для вероятностей k успехов в n испытаниях $\mathrm{P}_n(k) \cong \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda n}$.

Распределение Пуассона является приемлемой моделью описания случайного числа появления определенных событий в фиксированном промежутке времени в фиксированной области пространства.

Найдем M(X), D(X) пуассоновского распределения.

Построим производящую функцию этого распределения

$$\Phi(z)=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}z^k=\mathrm{e}^{\lambda(z-1)}.$$

Отсюда

$$\Phi'(z)=\lambda\mathrm{e}^{\lambda(z-1)},\Phi'(1)=\lambda;\Phi''(z)=\lambda^2\mathrm{e}^{\lambda(z-1)},\Phi''(1)=\lambda^2.$$

Таким образом, $M(X)=\lambda$; $D(X)=\lambda^2+\lambda-\lambda^2=\lambda$.

Вывод. В распределении Пуассона математическое ожидание и дисперсия равняются параметру λ.

Непрерывная СВ X называется *распределенной по нормальному (гауссовскому) закону с параметрами* $m\in R$ и $\sigma>0$, если плотность распределения вероятностей имеет вид

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

Для краткости говорят, что СВ распределена по закону $\mathrm{N}(m,\sigma)$. Здесь *m*- математическое ожидание, σ^2- дисперсия случайной величины.

График плотности распределения N(*m*,σ) называют *кривой Гаусса*.

Интеграл вида $\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_0^z\mathrm{e}^{-\frac{\alpha^2}{2}}d\alpha$ называется *интегралом Лапласа, интегралом вероятностей, или стандартной* *функцией ошибок*. Закон распределения в этом случае обозначают N(0,1).

Функция распределения вероятностей нормально распределенной СВ выражается через Φ(z) следующим образом:

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x-m}{\sigma}\right).$$

В этом случае непрерывная СВ распределена по нормальному закону N(m,σ).

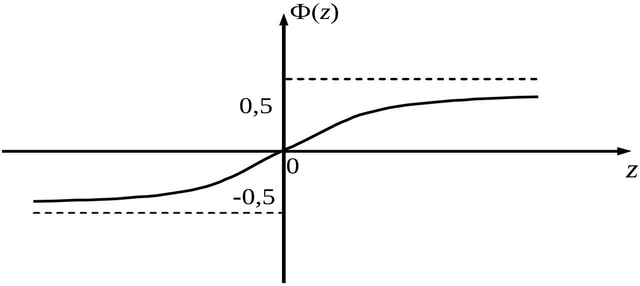
Простейшие свойства интеграла вероятностей

Свойство 1. Φ(0)=0.

Свойство 2. Φ(-z)=-Φ(z), то есть функция Φ(z) нечетная.

Свойство 3. Φ(+∞)=0,5.

График функции имеет вид



Свойство 4. Функция Φ(z) очень энергично стремится к своему пределу 0,5 при $z\to\infty$. Это стремление определяется следующей асимптотической формулой:

$$\Phi(z) = 0{,}5 - \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \left(\frac{1}{z} - \frac{1}{z^3} + \frac{1\cdot 3}{z^5} - \frac{1\cdot 3\cdot 5}{z^7} ...\right),$$

справедливой при достаточно больших z.

Вероятность отклонения случайной величины X от математического ожидания меньше чем на δ равна

$$P\left(\left|X\text{-}M(X)\right|<\delta\right)=2\Phi\left[\frac{\delta}{\sigma}\right].$$

Рассмотрим частный случай, когда δ=3σ. Тогда

$$P\left(\left|X-M(X)\right|<3\sigma\right)=2\Phi(3\sigma/\sigma)=2\Phi(3)=2\cdot 0,49865=0,9973$$

Это означает, что с вероятностью 0,997, близкой к единице (практически достоверно), все значения нормально распределенной СВ располагаются на интервале длиной 6σ, симметричном относительно математического ожидания.

Замечание. Часто рассматривают другую функцию:

$$\Phi(z)=\frac{1}{2}\left[1+\mathrm{erf}\left(\frac{z}{\sqrt{2}}\right)\right],$$

где $\mathrm{erf}\,z=\frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt=2\Phi(z\sqrt{2})-1$.

Для erf z справедливо erf(-z)=-erf z.

Иногда табулируется функция

$$\Phi_0(z)=\frac{1}{\sqrt{2\pi}}\int_0^z e^{-\frac{t^2}{2}}dt=\Phi(z)-\frac{1}{2}=\frac{1}{2}\mathrm{erf}\,(\frac{z}{\sqrt{2}}).$$

Очевидно, что $\Phi(z)=\frac{1}{2}+\Phi_0(z)$.

Пример 1. Найдем числовые характеристики биномиального распределения методом производящей функции:

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

Таким образом, математическое ожидание и дисперсия в биномиальном (распределении равны соответственно M(X)=*np*, D(X)=*npq*.

Пример 2. Нахождение числовых характеристик геометрического распределения.

Производящая функция имеет вид

$$\Phi(z)=\sum_{n=1}^{\infty} pq^{n-1}z^n = pz\sum_{n=1}^{\infty} (qz)^{n-1}.$$

Так как q<1 и |z|<1, то $\sum\limits_{n=1}^{\infty} (qz)^{n-1}$ есть сумма членов бесконечно убывающей геометрической прогрессии, поэтому

$$\Phi(z)=\frac{pz}{1-qz}=\frac{-\frac{p}{q}(1-qz)+\frac{p}{q}}{1-qz}=-\frac{p}{q}+\frac{p}{q}\frac{1}{1-qz}.$$

Вычисляем производные

$$\Phi'(z)=\frac{p}{q}\frac{q}{(1-qz)^2}=\frac{p}{(1-qz)^2},\Phi''(z)=\frac{2pq}{(1-qz)^3}.$$

Отсюда $\Phi'(1)=\frac{1}{p}$, $\Phi''(1)=\frac{2q}{p^2}$, следовательно,

$$\mathrm{M}(\mathrm{X})=\frac{1}{p},\quad \mathrm{D}(\mathrm{X})=\frac{2q}{p^2}+\frac{1}{p}-\frac{1}{p^2}=\frac{q}{p^2}.$$

Таким образом, $\mathrm{M}(\mathrm{X})=\frac{1}{p}$, $\mathrm{D}(\mathrm{X})=\frac{q}{p^2}$.

Пример 3. Радиоаппаратура состоит из 1000 электроэлементов. Вероятность отказа одного элемента в течение одного года работы равна 0,001 и не зависит от состояния других элементов. Какова вероятность отказа двух и не менее двух электроэлементов за год?

Решение.

1) $P_{1000}(2)=\frac{\lambda^2}{2!}e^{-\lambda}=\frac{1}{2e}$, так как $\lambda=np=1000\cdot 0{,}001=1$

2) $P_{1000}(m \geq 2) = 1 - P(0) - P(1) = 1 - \frac{1}{0!}e^{-1} - \frac{1}{1!}e^{-1} = 1 - \frac{2}{e} \approx 0,264.$

Пример 4. Завод отправил на базу 5000 изделий. Вероятность того, что в пути изделие испортится, равно p=0,0002. Найти вероятность того, что на базу прибудет а) ровно три негодных изделия; б) не более трех негодных изделий.

Решение. Используем биномиальное распределение

$$\mathrm{P}_{5000}(3) = \mathrm{C}_{5000}^{3}(0,0002)^{3} \cdot (0,9998)^{4997}.$$

Очевидно, что с помощью такой формулы вычисление вероятности затруднительно. Для упрощения вычислений естественно заменить (приближенно) биномиальное распределение распределением Пуассона:

$$\lambda_n = np = 5000 \cdot 0,0002 = 1.$$

Тогда а) $\mathrm{P}_{5000}(3) \cong \left.\frac{\lambda^3}{3!}\mathrm{e}^{-\lambda}\right|_{\lambda=1} = \frac{1}{6\mathrm{e}} \cong 0,06,$

б) $\mathrm{P}_{5000}(m \leq 3) = \mathrm{P}(0) + \mathrm{P}(1) + \mathrm{P}(2) + \mathrm{P}(3) = \mathrm{e}^{-1}\left(1 + 1 + \frac{1}{2} + \frac{1}{6}\right) \cong 0,95.$

Пример 5. Поток заявок, поступающих на телефонную станцию, представляет собой простейший поток. Математическое ожидание числа вызовов за час равно 30. Найти вероятность того, что за минуту поступит не менее двух вызовов.

Решение. Так как поток заявок представляет собой простейший поток, то число событий потока m, попадающий на любой участок времени τ, распределено по закону Пуассона

$$P_m = \frac{a^m}{m!}e^{-a}, \qquad a = \lambda\tau.$$

В нашем случае $\lambda = 30$, $\tau = \frac{1}{60}$, $a = \lambda\tau = \frac{30}{60} = 0,5$.

Обозначим через A – событие, состоящее в том, что за минуту поступит не менее двух вызовов. Тогда

$$P(A)=1-P_0-P_1=1-\frac{a^0}{0!}e^{-0,5}-\frac{a^1}{1!}e^{-0,5}=1-e^{-0,5}-0,5e^{-0,5}=$$

$0,0902$.

Пример 6. Случайная величина ξ имеет пуассоновское распределение и известно, что ее математическое ожидание m и дисперсия D связаны соотношением
$6-m^2=5\cdot D$. Найти вероятность $P(\xi<3)$, $M\xi^2$.

Решение. Известно, что математическое ожидание и дисперсия пуассоновского распределения совпадают и равны значению его параметра λ. Условие задачи приводит к уравнению относительно λ:

$$\lambda^2+5\lambda-6=0,$$

решениями которого являются числа $\lambda_1=1$, $\lambda_2=-6$. Последнее значение не может быть параметром пуассоновского распределения в силу положительности параметра. Таким образом, случайная величина ξ имеет ряд распределения

$$P\left(\xi=n\right)=e^{-1}\frac{1}{n!},\ n=0,1,2,\dots\ .$$

Для искомой вероятности получаем

$$P(\xi<3)=P(\xi=0)+P(\xi=1)+P(\xi=2)=e^{-1}+e^{-1}\frac{1}{1!}+e^{-1}\frac{1}{2!}\approx$$

$0,9197$.

$M\xi^2=D\xi+(M\xi)^2=1+1=2$.

Пример 7. Размер диаметра втулок подчиняется нормальному распределению с параметром M(X)= 2,5 см, σ=10^{-2} см.

В каких границах можно практически достоверно гарантировать размеры диаметра втулок?

Решение. P(|X-M(X)|<δ)=2Φ(δ/σ)=2Φ(δ/0,01)=0,997. По таблице находим δ/0,01=2,98 и δ=0,0298. α=2,5±0,028 (см), 2,4702≤α≤2,5298.

Задачи для самостоятельного решения

1. Случайная величина X –погрешность измерительного прибора распределена по нормальному закону распределения с дисперсией $\sigma^2 = 25 мB^2$. Систематическая погрешность прибора отсутствует. Найдите вероятность того, что при пяти независимых измерениях ошибка измерения хотя бы один раз превзойдет по модулю 10 мВ.

2. Взвешивание производится без систематической ошибки, а случайные ошибки подчинены нормальному закону распределения со средним квадратическим отклонением $\sigma = 20$ мг. Найдите вероятность того, что взвешивание будет произведено с ошибкой, не превосходящей по абсолютной величине 50 мг.

3**.** Плотность распределения вероятностей случайной величины ξ имеет вид $f(x) = \gamma e^{-ax^2+bx-\frac{b^2}{4a}}$. Найти: γ, математическое ожидание $M\xi$, дисперсию $D\xi$, вероятность выполнения неравенства $\left|\xi - \frac{b}{2a}\right| < 1$.

4. Случайные ошибки измерения ξ дальности до неподвижной цели подчинены гауссовскому закону с математическим ожиданием $m = a$ и $\sigma = b$.

Определить вероятности того, что:

1) $P\left\{|\xi-m|\le\frac{a+b}{3}\right\}$.

2) при трех независимых измерениях ошибка хотя бы одного измерения не превзойдет по абсолютной величине $\frac{(a+b)}{3}$.

Ответы: 1. $\approx 0,21$; 2. $\approx 0,98$.

ЗАНЯТИЯ № 8-10. МНОГОМЕРНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ (СЛУЧАЙНЫЕ ВЕКТОРЫ)

Вектор $\bar{\xi}=(\xi_1\xi_2)$, координаты которого есть случайные величины, заданные на одном и том же вероятностном пространстве, называется *случайным вектором*, а функция $F(x,y)=P\ \ \xi_1<x,\ \xi_2<y$ называется функцией распределения случайного вектора $\bar{\xi}$ или двумерной случайной величины $\bar{\xi}$.

Если координаты вектора $\bar{\xi}$ дискретные случайные величины, то $\bar{\xi}$ называют *дискретным случайным вектором.*

Законом распределения дискретного случайного вектора называется перечень всех возможных значений пар компонент $\{(x_i,y_j)|(x_i,y_j)\in G(x,y)\}$ и соответствующих каждой паре вероятностей p_{ij}=P(X=x_i,Y=y_j), удовлетворяющих условию $\sum_i\sum_j p_{ij}=1$, где суммирование распространяется на все возможные значения индексов i и j.

Закон распределения двумерного случайного вектора часто задается в виде табл. 2:

Таблица 2

X \ Y	x_1	x_2	…	x_i	…
y_1	p_{11}	p_{21}	…	p_{i1}	…
y_2	p_{12}	p_{22}	…	p_{i2}	…
…	…	…	…	…	…
y_j	p_{1j}	p_{2j}	…	p_{ij}	
…	…	…	…	…	…

$$\sum_{i,j} p_{ij} = 1.$$

Знаная закон распределения двумерного случайного вектора, можно получить закон распределения его компонент

$P\{X = x_i\} = \sum_j p_{ij}$, $P\{Y = y_j\} = \sum_i p_{ij}$ и ФР $F(x, y) = \sum_{i,j \in U} p_{ij}$,

где множество индексов U определяется следующим образом:

$$U=\{(i,j)|(X<x_i,\ Y<y_j)\}.$$

Если функцию распределения вероятности вектора $\bar{\xi}$ можно представить в виде $F(x, y) = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} f(t, s)dtds$, то случайную величину (ξ_1, ξ_2) называют *непрерывной двумерной случайной величиной*, а $f(x, y)$ – ее *плотностью распределения вероятности.*

Всюду в дальнейшем будем считать, что $f(x, y)$ – непрерывная функция по обоим аргументам.

Свойства функции и плотности распределения вероятности

1) $P(x_1 < \xi_1 < x_2, \quad y_1 < \xi_2 < y_2) = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1)$.

2) $0 \le F(x, y) \le 1$.

3) $F(-\infty, -\infty) = F(x, -\infty) = F(-\infty, y) = 0.$

4) $F(+\infty, +\infty) = 1$.

5) $F(x, +\infty) = F_1(x)$, $F(+\infty, y) = F_2(y)$, где $F_1(x)$ и $F_2(y)$ – функции распределения случайных величин ξ_1 и ξ_2.

6) $f(x, y) = F''_{xy}(x, y)$.

7) $f(x, y) \ge 0$.

8) $\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x, y)dxdy = 1$.

9) $P(\overline{\xi} \in D) = \iint\limits_D f(x, y)dxdy$.

10) $f_1(x) = \int\limits_{-\infty}^{\infty} f(x, y)dy$, $f_2(y) = \int\limits_{-\infty}^{\infty} f(x, y)dx$,

где $f_1(x)$ и $f_2(y)$ – плотности распределения случайных величин ξ_1 и ξ_2.

Условной плотностью распределения случайной величины ξ_1 при условии $\xi_2 = y$ называют отношение плотности совместного распределения $f(x, y)$ системы (ξ_1 , ξ_2) к плотности распределения составляющей ξ_2: $f(x / y) = \dfrac{f(x, y)}{f_2(y)}$.

Аналогично определяют $f(y/x)=\dfrac{f(x,y)}{f_1(x)}$.

Теорема умножения плотностей

$$f(x,y)=f_1(x)\cdot f\left(y/x\right)=f_2(y)\cdot f\left(x/y\right).$$

Случайные величины ξ_1 и ξ_2 называются *независимыми*, если для любых чисел x,y случайные события $\left(\xi_1<x\right)$ и $\left(\xi_2<y\right)$ независимы.

Случайные события независимы, если выполняется любое из условий:

1) $F(x,y)=F_1(x)\cdot F_2(y)$.

2) $f(x,y)=f_1(x)\cdot f_2(y)$.

3) $f(y/x)=f_2(y)$ или $f\left(x/y\right)=f_1(x)$.

Условным математическим ожиданием называют выражение

$M\left(\xi_2/\xi_1=x_i\right)=\sum\limits_{j=1}^{m} y_j P\left(\xi_2=y_i/\xi_1=x_i\right)$ для дискретного случайного вектора

$M\left(\xi_2/x\right)=\int\limits_{-\infty}^{\infty} yf\left(y/x\right)dy$ для непрерывного случайного вектора.

Величина $K_{\xi_1,\xi_2}=\operatorname{cov}\left(\xi_1,\xi_2\right)=M\left[\left(\xi-M\xi\right)\left(\eta-M\eta\right)\right]$ называется *корреляционным моментом (ковариацией)* двух случайных величин ξ_1 и ξ_2.

Если ξ_1, ξ_2 – непрерывная двумерная случайная величина с плотностью распределения $f(x, y)$, то

$$\operatorname{cov}(\xi_1, \xi_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - m_1)(y - m_2) f(x, y)dxdy =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x, y)dxdy - m_1 m_2,$$

где $m_1 = M\xi_1$, $m_2 = M\xi_2$.

Для дискретного случайного вектора

$$\operatorname{cov}(\xi_1, \xi_2) = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i - m_1)(y_i - m_2) p_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j p_{ij} - m_1 m_2 .$$

Величина $r_{\xi_1\xi_2} = \dfrac{K_{\xi_1\xi_2}}{\sqrt{D_{\xi_1} D_{\xi_2}}}$ называется коэффициентом корреляции случайных величин ξ_1 и ξ_2.

Если $r_{\xi_1\xi_2} = 0$, то случайные величины ξ_1 и ξ_2 называются некоррелированными.

Свойства корреляционного момента и коэффициента корреляции

1) $\left| r_{\xi_1, \xi_2} \right| \le 1$.

2) Если ξ_1 и ξ_2 независимы, то $r_{\xi,\xi_i} = 0$. Обратное неверно: из некоррелируемости случайных величин не следует их независимость.

3) Если $\xi_2 = a\xi_1 + b$, то $\left| r_{\xi_1\xi_2} \right| = 1$.

4) $\operatorname{cov}(\xi_1, \xi_2) = \operatorname{cov}(\xi_2, \xi_1)$.

5) $\operatorname{cov}(\xi, \xi) = D\xi$.

6) $\operatorname{cov}(a\xi_1, b\xi_2) = ab\operatorname{cov}(\xi_1, \xi_2)$.

7) $\operatorname{cov}(a\xi_1 + b\xi_2, \xi_3) = a\operatorname{cov}(\xi_1\xi_3) + b\operatorname{cov}(\xi_2, \xi_3)$.

Свойства математического ожидания и дисперсии случайного вектора

1) $MC = C$, где C – постоянная.

2) $MC\xi = CM\xi$.

3) $M(\xi_1 + \xi_2) = M\xi_1 + M\xi_2$.

4) $M(\xi_1 \cdot \xi_2) = M\xi_1 \cdot M\xi_2 + \operatorname{cov}(\xi_1\xi_2)$.

Если $\operatorname{cov}(\xi_1\xi_2) = 0$, то $M(\xi_1 \cdot \xi_2) = M\xi_1 \cdot M\xi_2$.

Случайная величина ξ называется *неотрицательной* $(\xi \geq 0)$, если она принимает только неотрицательные значения.

5) Если $\xi \geq 0$, $M\xi \geq 0$.

6) $DC = 0$, где C – постоянная.

7) $DC\xi = C^2 D\xi$.

8) $D(\xi_1 + \xi_2) = D\xi_1 + D\xi_2 + 2\operatorname{cov}(\xi_1, \xi_2)$.

Если $\operatorname{cov}(\xi_1\xi_2) = 0$, то $D(\xi_1 + \xi_2) = D\xi_1 + D\xi_2$.

9) $D(\xi + C) = D\xi$. C – постоянная.

10) $D\xi = M(\xi - M\xi)^2 \geq 0$.

11) $D\xi = M\xi^2 - M\xi^{\ 2} \geq 0 \Rightarrow M\xi^{\ 2} \leq M\xi^2$.

Двумерная случайная величина $\xi_1\xi_2$ называется *распределенной по нормальному закону*, если ее плотность распределения

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\cdot\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_1)^2}{\sigma_1^2}-\right.\right.$$

$$\left.\left.-\frac{2r(x-m_1)(y-m_2)}{\sigma_1\sigma_2}+\frac{(y-m_2)^2}{\sigma_2^2}\right]\right\}.$$

Здесь $m_1 = M\xi_1$, $m_2 = M\xi_2$, $\sigma_1^2 = D\xi_1$, $\sigma_2^2 = D\xi_2$, r – коэффициент корреляции случайных величин ξ_1 и ξ_2. Для нормальной случайной величины понятия независимости и некоррелируемости эквивалентны.

Двумерная случайная величина распределена равномерно в области D, если ее плотность распределения

$$f(x,y)=\begin{cases}\dfrac{1}{S}, & (x,y)\in D\\ 0, & (x,y)\in D\end{cases}$$

Здесь S – площадь области D.

Пример 1. Дискретная двумерная случайная величина ξ,η распределена по закону, приведенному в таблице

Таблица 3

$\xi \backslash \eta$	–1	0	2
–1	0,2	0,1	0,3
1	0,1	0,1	0,2

Определить:

1) Законы распределения составляющих ξ и η;

2) Условный закон распределения случайной величины ξ при условии, что $\eta=-1$;

3) $M\;\;\xi/\eta=-1\;$;

4) Коэффициент корреляции $r_{\xi,\eta}$.

Решение. Имеем

ξ	–1	1
	0,6	0,4

(8.1)

η	–1	0	2
	0,3	0,2	0,5

$M\xi=-1\cdot 0,6+1\cdot 0,4=-0,2$, $\qquad M\eta=-1\cdot 0,3+2\cdot 0,5=0,7$,

$D\xi=(-1)^2 0,6+1^2\cdot 0,4-(-0,2)^2=0,96$,

$D\eta=(-1)^2\cdot 0,3+2^2\cdot 0,5-(0,7)^2=1,81$.

$\xi/\eta=-1$	–1	1
	2/3	1/3

(8.2)

$$P\;\;\xi=-1/\eta=-1\;\;=\frac{P\;\;\xi=-1,\eta=-1}{P\;\;\eta=-1}=\frac{0,2}{0,3}=2/3.$$

$$P\;\;\xi=1/\eta=-1\;\;=\frac{P\;\;\xi=1,\eta=-1}{P\;\;\eta=-1}=\frac{0,1}{0,3}=1/3.$$

Сравнивая (8.1) и (8.2), видим, что ξ,η зависимые случайные величины:

$$M\;\;\xi/\eta=-1\;\;=-1\cdot\frac{2}{3}+1\cdot\frac{1}{3}=-\frac{1}{3};$$

$$K(\xi,\eta) = \sum_{i,j} P_{ij} x_i y_i - m_1 \cdot m_2 = (-1)\cdot(-1)\cdot 0,2 + (-1)\cdot 2\cdot 0,3 +$$

$$+1\cdot(-1)\cdot 0,1 + 1\cdot 2\cdot 0,2 + 0,2\cdot 0,7 = 0,2 - 0,6 - 0,1 + 0,4 + 0,14 =$$

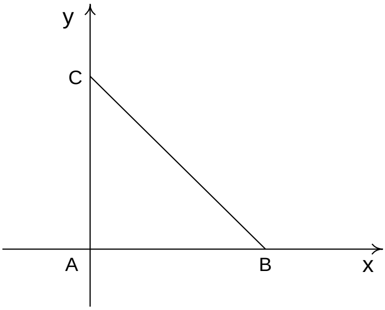
$$= 0,04 .$$

$$r_{\xi,\eta} = \frac{K_{\xi,\eta}}{\sigma_\xi \sigma_\eta} = \frac{0,04}{\sqrt{0,96}\cdot\sqrt{1,81}} = \frac{0,04}{0,98\cdot 1,345} = \frac{0,04}{1,32} = 0,03 .$$

Пример 2. Двумерная случайная величина (ξ,η) имеет равномерное распределение вероятностей в треугольной области ABC, то есть

$$f(x,y) = \begin{cases} c, & (x,y) \in ABC, \\ 0, & (x,y) \notin ABC. \end{cases}$$

Найти постоянную c, одномерные плотности $f_1(x)$, $f_2(y)$ случайных величин ξ и η, коэффициент корреляции r, условную плотность $f(y/x)$ и условное математическое ожидание $M(\eta/x)$.



т. $A(0,0)$, т. $B(1,0)$, т. $C(0,1)$.

1) Постоянную c найдем из условия нормировки

$$1=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x,y)dxdy=\iint\limits_{ABC} cdxdy=c\cdot S_{\Delta}=c\cdot 1/2,\quad c=2,$$

где S – площадь треугольника ABC. Обозначим область, ограниченную треугольником ABC через D. Тогда

$$f(x,y)=\begin{cases}2, & (x,y)\in D,\\ 0, & (x,y)\notin D.\end{cases}$$

2) Уравнение прямой BC : $y=1-x$. Тогда область D можно задать аналитически следующим образом:

$$D=\begin{Bmatrix}0\le x\le 1\\ 0\le y\le 1-x\end{Bmatrix}\quad \text{или}\quad D=\begin{Bmatrix}0\le y\le 1\\ 0\le x\le 1-y\end{Bmatrix}.$$

3) $$f_1(x)=\int\limits_{-\infty}^{\infty} f(x,y)dy=\begin{cases}2\int\limits_{0}^{1-x} dy & 0\le x\le 1\\ 0 & x\le 0,\ x>1\end{cases}=\begin{cases}2(1-x), 0<x<1,\\ 0, \quad x<0,\ x>1,\end{cases}$$

$$f_2(y)=\int\limits_{-\infty}^{\infty} f(x,y)dx=\begin{cases}2(1-y), & 0<y<1,\\ 0, & y<0,\ y>1.\end{cases}$$

$$M\xi=\int\limits_{-\infty}^{\infty} xf_1(x)dx=\int\limits_{0}^{1} x\cdot 2(1-x)dx=1/3.$$

$$M\eta=\int\limits_{-\infty}^{\infty} yf_2(y)dy=\int\limits_{0}^{1} y\cdot 2(1-y)dy=1/3.$$

$$D\xi=\int\limits_{-\infty}^{\infty} x^2 f_1(x)dx-m_x^2=\int\limits_{0}^{1} x^2\cdot 2(1-x)dx-\frac{1}{9}=0,055.$$

$$D\eta=\int\limits_{-\infty}^{\infty} y^2 f_2(y)dy-m_y^2=0,055.$$

4) $K_{xy} = \iint\limits_D xyf(x,y)dxdy - m_x m_y = 2\int\limits_0^1 dx \int\limits_0^{1-x} dy - \dfrac{1}{9} \cong -0,0278$.

$$r = \frac{K_{xy}}{\sigma_x \sigma_y} \cong \frac{-0,0278}{\sqrt{(0,055)^2}} \cong -0,5\,.$$

5) $f \;\; y/x \;\; = \dfrac{f(x,y)}{f_1(x)} = \begin{cases} \dfrac{1}{1-x}, & 0 < x < 1, \quad 0 < y < 1-x, \\ 0, & (x,y) \notin D. \end{cases}$

$$M(\eta/x) = \int\limits_{-\infty}^{\infty} yf(y/x)\,dy = \begin{cases} \int\limits_0^{1-x} y\dfrac{1}{1-x}dy, & 0 < x < 1, \\ 0, & x < 0, \;\; x \geq 1. \end{cases}$$

$$\int\limits_0^{1-x} y\frac{1}{1-x}\,dy = \frac{1}{1-x}\cdot\frac{y^2}{2}\bigg|_0^{1-x} = \frac{1-x}{2}\,.$$

$$M(\eta/x) = \begin{cases} \dfrac{1-x}{2}, & 0 < x < 1, \\ 0, & x \leq 0, \;\; x \geq 1. \end{cases}$$

Пример 3. Случайная точка ξ_1, ξ_2 распределена равномерно внутри круга радиуса R $D = \;\; x^2 + y^2 \leq R^2$. Найти математическое ожидание случайной величины $\eta = \xi_1 \cdot \xi_2$.

Решение. Плотность распределения вероятности

$$f(x,y) = \begin{cases} \dfrac{1}{\pi R^2}, & x^2 + y^2 \leq R^2 \\ 0, & x^2 + y^2 \geq R^2 \end{cases}$$

$$M\eta = \frac{1}{\pi R^2}\iint\limits_D xydxdy = \begin{vmatrix} x=\rho\cos\phi \quad 0\le\rho\le R \\ y=\rho\sin\phi \quad 0\le\phi\le 2\pi \\ dxdy=\rho d\rho d\phi \end{vmatrix} =$$

$$\frac{1}{\pi R}\int\limits_0^R\int\limits_0^{2\pi}\rho^3\cos\phi\,\sin\phi\, d\rho d\phi\; =$$

$$=\frac{1}{\pi R^2}\int\limits_0^R \rho^3 d\rho\cdot\int\limits_0^{2\pi}\cos\phi\sin\phi d\phi = \frac{1}{\pi R^2}\cdot\frac{\rho_4}{4}\bigg|_0^R\cdot\frac{\sin^2\phi}{2}\bigg|_0^{2\pi} =$$

$$=\frac{1}{\pi R^2}\cdot\frac{1}{4}R^4\cdot\ \sin^2 2\pi-\sin^2 0\ \ \frac{1}{2}=0.$$

Пример 4. Пара случайных величин ξ и η имеет совместное нормальное распределение с вектором математических ожиданий $-2,-1$ и ковариационной матрицей K

$$K\ \xi,\eta\ =\sigma^2\begin{bmatrix} 2 & 3 \\ 3 & 7 \end{bmatrix}.$$

Известно, что $P\ \ \xi-2\eta<3\ =0,65$. Найти $D\xi,\ D\eta$.

Решение. Совместная нормальность пары случайных величин ξ и η обеспечивает нормальность каждой из них и любой их линейной комбинации, в частности величина $\zeta=\xi-2\eta$ нормальна с параметрами

$$M\zeta = M\xi - 2M\eta = -2-2(-1)=0,\ D\zeta = D\xi+4D\eta-4\mathrm{cov}(\xi,\eta)$$

.

Подставляя в последнее соотношение элементы ковариационной матрицы:

получим $D\xi=2\sigma^2$, $D\eta=7\sigma^2$, $\mathrm{cov}(\xi,\eta)=3\sigma^2$,

$$D\zeta = 2\sigma^2+4\cdot7\sigma^2-4\cdot3\sigma^2\ =18\sigma^2.$$

По условию $P\left(\zeta < 3\right) = 0,65$, откуда, используя нормальность ζ,

$$F\left(\frac{3}{\sigma\sqrt{18}}\right) = 0,65 \Rightarrow \frac{3}{\sigma\sqrt{18}} = 0,385 \Rightarrow \sigma \approx 1,837.$$

Искомые дисперсии равны, соответственно,

$$D\xi = 2\sigma^2 \approx 6,747, \quad D\eta = 7\sigma^2 \approx 23,622.$$

Пример 5. Случайный вектор $\overline{\xi} = \left(\xi_1, \xi_2\right)$ имеет вектор математических ожиданий $\overline{m}_1 = (2,1)$ и корреляционную матрицу

$$K = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}.\ \eta_1 = 2\xi_1 - \xi_2,\ \ \eta_2 = \xi_1 + 3\xi_2.$$

Вычислить вектор математических ожиданий $\overline{m}_2 = \left(M\eta_1, M\eta_2\right)$ случайного вектора $\overline{\eta} = (\eta_1, \eta_2)$ и корреляционную матрицу вектора $\overline{\eta}$.

Решение. $M\eta_1 = M\left(2\xi_1 - \xi_2\right) = 2M\xi_1 - M\xi_2 = 2\cdot 2 - 1 = 3.$

$M\eta_2 = M\left(\xi_1 + 3\xi_2\right) = M\xi_1 + 3M\xi_2 = 2 + 3 = 5$. $\overline{m}_2 = (3,5)$.

$D\eta_1 = D\left(2\xi_1 - \xi_2\right) = 4D\xi_1 - 4K\left(\xi_1\xi_2\right) + D\xi_2 = 4\cdot 2 - 4\cdot 1 + 4 = 8.$

$D\eta_2 = D\left(\xi_1 + 3\xi_2\right) = D\xi_1 + 6K\left(\xi_1\xi_2\right) + 9D\xi_2 = 2 + 6\cdot 1 + 9\cdot 4 = 44$
.

$K\left(\eta_1\eta_2\right) = \operatorname{cov}\left(\eta_1, \eta_2\right) = \operatorname{cov}\left(2\xi_1 - \xi_2, \xi_1 + 3\xi_2\right) =$

$= 2\operatorname{cov}\left(\xi_1\xi_1\right) + 6\operatorname{cov}\left(\xi_1\xi_2\right) - \operatorname{cov}\left(\xi_2\xi_1\right) - 3\operatorname{cov}\left(\xi_2\xi_1\right) =$

$= 2D\xi_1 + 5K\left(\xi_1\xi_2\right) - 3D\xi_2 = 4 + 5 - 12 = -3.$

Ответ: $K(\eta_1,\eta_2)=\begin{pmatrix} 8 & -3 \\ -3 & 44 \end{pmatrix}$, $\bar{m}_2=(3,5)$.

Примеры для самостоятельного решения

1. Дискретные случайные величины X и Y независимы и имеют распределения:

X	2	3
P	0,3	0,7

$У$	4	5
P	0,4	0,6

Найдите закон распределения случайной величины $Z = X + Y$ и ее математическое ожидание.

2. Случайные величины X и Y независимы и каждая имеет показательный закон распределения с плотностью распределения $f(x)=\lambda e^{-\lambda x}$ при $x \geq 0$ и $f(x)=0$ при $x<0$. Найдите плотность вероятности суммы этих величин.

3. Найдите математическое ожидание и среднее квадратическое отклонение случайной величины $Z=3X-Y+5$, если $M(X)=3$, $M(Y)=5$, $D(X)=2$, $D(Y)=1$, а случайные величины X и Y независимы.

4. Случайные величины X и Y независимы и обе равномерно распределены на отрезке [0, 2]. Найдите функцию плотности вероятности случайной величины $Z = X + Y$.

5. Пусть X_1 и X_2 - независимые одинаково распределенные случайные величины, имеющие показательное распределение с параметром λ. Найти распределение случайной величины $Y = X_1 - X_2$.

6. Случайные величины X и Y независимы и каждая равномерно распределена на (0, 1). Найдите плотность вероятности случайной величины $Z = Y / X^2$.

7. Каждая из случайных величин X и Y равномерно распределена в интервале $(0,1)$. Полагая величины X и Y независимыми, найдите функцию распределения, математическое ожидание и дисперсию для каждой из величин $U = \min(X, Y)$ и $V = \max(X, Y)$.

8. Закон распределения двумерной случайной величины задан табл. 4.

Таблица 4

X	1	2	4
$Y=1$ $Y=3$	0,2 0,05	0,3 0,15	0,1 0,2

Найдите: а) безусловные законы распределения величин X и Y; б) закон распределения X при условии, что $Y=1$.

9. Равновозможны все положения случайной точки (X, Y) в треугольнике с вершинами $A(0,0), B(2,0)$ и $C(2,1)$. Найти коэффициент корреляции случайных величин X и Y. Найти линию регрессии Y на X.

10. В примере №8 найдите корреляции между X и Y.

11. По известной функции плотности вероятности $f(x)$ случайной величины X найдите функцию плотности вероятности $g(y)$ случайной величины.

12. Система случайных величин (X_1, X_2) имеет функцию плотности вероятности $f(x_1, x_2) = \dfrac{2}{\pi^2 (x_1^2 + 1)(x_2^2 + 4)}$. Найдите плотность распределения $g(y_1, y_2)$ двумерной случайной ве-

личины (Y_1, Y_2), если $X_1 = tg(Y_1)$, $X_2 = 2tg(Y_2)$, $|Y_1| < \pi / 2$, $|Y_2| < \pi / 2$.

13. Задана табл. 5 распределения дискретной двумерной случайной величины (X, Y). Определить:

а) безусловные законы распределения СВ (X, Y);

б) функцию распределения $F(x, y)$ системы СВ (X, Y);

в) условный закон распределения СВ Y при $X = x_i$ и $M[Y / X = x_i]$;

г) зависимость или независимость компонент X, Y;

д) центр рассеивания: точку $M(m_X, m_Y)$;

е) коэффициент корреляции r_{XY} .

Таблица 5

X / Y	-1	0	1
1	0.15	0.3	0.35
2	0.05	0.05	?

14. Двумерная СВ распределена равномерно в области D . D – треугольник с вершинами О(0,0), А(1,0), В(0,1).

Определить: 1) двумерную плотность вероятности $f(x, y)$;

2) одномерные плотности $f_1(x)$, $f_2(x)$;

3) зависимы или нет СВ (X, Y);

4) центр рассеивания: точку $M(m_X, m_Y)$;

5) дисперсии D_X , D_y ;

6) коэффициент корреляции r_{XY} .

15. Задана табл. 6 распределения дискретной двумерной случайной величины (X, Y) .Определить:

а) безусловные законы распределения СВ (X, Y);

б) функцию распределения $F(x,y)$ системы СВ (X,Y);

в) условный закон распределения СВ Y при $X = x_i$ и $M[Y/X = x_i]$;

г) зависимость или независимость компонент X,Y;

д) центр рассеивания: точку $M(m_X, m_Y)$;

е) коэффициент корреляции r_{XY} .

Таблица 6

X/Y	2	3	5
0	0.05	0.2	?
1	0405	015	0.2

16. Дана плотность вероятности $f(x,y)$ СВ (X,Y)

$$f(x,y) = \begin{cases} C(x^3 + y^3), \ 0 \le x \le 1, \ 0 \le y \le 1; \\ 0 \quad \text{в остальных случаях.} \end{cases}$$

Определить:

1) параметр С;

2) одномерные плотности $f_1(x), \ f_2(x)$;

3) зависимы или нет СВ (X,Y);

4) центр рассеивания: точку $M(m_X, m_Y)$;

5) дисперсии D_X , D_y ;

6) коэффициент корреляции r_{XY} .

Ответы

1.

Z	6	7	8
P	0,12	0,46	0,42

, $M(Z) = 7{,}3$;

2.

$f(z) = \lambda^2 z e^{-\lambda z}$ при $z \ge 0$ и $f(x) = 0$ при $z < 0$;

3. $M(Z) = 6$, $\sigma(Z) = 3$;

4. $f(z)=0,25z$ при $z\in[0,2]$, $f(z)=1-0,25z$ при $z\in[2,4]$, $f(z)=0$ при остальных z; 5. $f(y)=\dfrac{\lambda^2}{2}e^{-\lambda|x|}$;

6. $f(z)=0$ при $z\le 0$, $f(z)=1/3$ при $z\in(0,1]$, $f(z)=\dfrac{1}{3z^{1,5}}$ при $1<z$;

7. $F(v)=P(V<v)=0$ при $v\le 0$, $F(v)=2v-v^2$ при $0<v\le 1$, $F(v)=1$ при $1<v$, $M(V)=1/3$, $D(V)=1/18$,

$F(u)=P(U<u)=0$ при $u\le 0$, $F(u)=u^2$ при $0<u\le 1$,

$F(u)=1$ при $1<u$, $M(U)=2/3$, $D(U)=1/18$;

8.

а

X	1	2	4
P	0,25	0,45	0,3

,

Y	1	3
P	0,6	0,4

; б

X	1	2	4
P	1/3	1/2	1/6

9. $r_{xy}=1/2$, $y=0,5x$; 10. $r_{xy}=0,08$;

11. $g(y)=\dfrac{f(y)}{2\sqrt{y}}$ при $0<y\le 1$, $g(y)=f(x)$ при $y\le 0$ и $y>1$;

12. $g(y_1,y_2)=4/\pi^2$ при $|y_1|<\pi/2$, $|y_2|<\pi/2$ и $g(y_1,y_2)=0$ при остальных y_1 и y_2.

ЗАНЯТИЕ № 11-12. ФУНКЦИИ СЛУЧАЙНЫХ ВЕЛИЧИН

Если X - дискретная СВ, имеющая закон распределения (x_k, p_k) k=1, 2,..., который можно записать в виде табл. 7, и Y=φ

φ(X), где φ -неслучайная функция, то Y также является дискретной СВ, причем ее возможные значения $y_k=\varphi(x_k)$.

Таблица 7

X	x_1	x_2	x_3	…	x_k	…
P(X)	p_1	p_2	p_3	…	p_k	…

Если $\varphi(x)$ – монотонная функция, то все значения y_k различны и $P(Y=y_k)=P(X=x_k)$, то есть СВ Y имеет следующий закон распределения

Таблица 8

Y	$\varphi(x_1)$	$\varphi(x_2)$	…	$\varphi(x_k)$	…
P(Y)	p_1	p_2	…	p_k	…

Если при этом $\varphi(x)$ – немонотонная функция, то среди ее значений y_1, y_2, y_3,…,y_k,… могут быть одинаковые. В этом случае столбцы с равными значениями $\varphi(x_i)$ объединяются в один столбец, а соответствующие вероятности складываются, т. е.

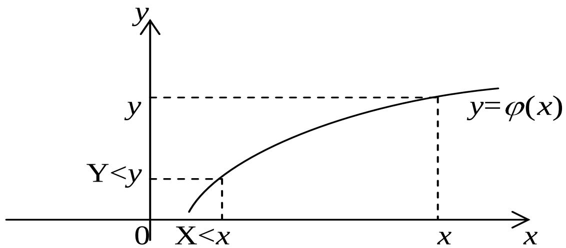
$$P\left(Y=y_k\right)=\sum_{i:\varphi\left(x_i\right)=y_k} P\left(X=x_i\right).$$

Если X – непрерывная СВ с ФР F(x) и плотностью вероятности $f(x)$ и $Y=\varphi(x)$, причем $\varphi(x)$ – монотонно возрастающая непрерывно дифференцируемая функция, а $x=\varphi^{-1}(y)$ – обратная функция, то

$F(y)=P(Y<y)=P(\varphi(x)<y)=P(X<\varphi^{-1}(x))=F[\varphi^{-1}(x)]$.

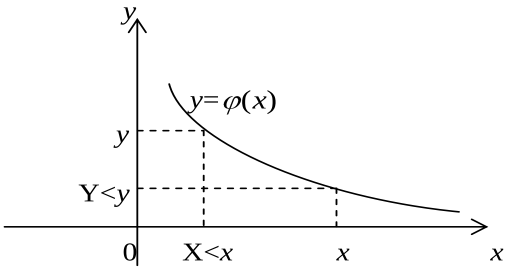
Дифференцируя последнее равенство по y, получаем

$$F'(y)=F'\left[\varphi^{-1}(y)\right]\frac{d}{dy}\left[\varphi^{-1}(y)\right] \quad \text{или} \quad f(y)=f\left[\varphi^{-1}(y)\right]\frac{d}{dy}\left[\varphi^{-1}(y)\right].$$



Если $y=\varphi(x)$ - монотонно убывающая функция, то аналогично получаются следующие соотношения:

$F(y)=1-F[\varphi^{-1}(y)]$, $f(y)=-f[\varphi^{-1}(y)]\dfrac{d}{dy}[\varphi^{-1}(y)]$.



Выражения для плотности вероятности СВ Y и для монотонно возрастающей и для монотонно убывающей функции $\varphi(x)$ можно объединить:

$$f(y)=f\left[\varphi^{-1}(y)\right]\left|\frac{d}{dy}\left[\varphi^{-1}(y)\right]\right|.$$

Рассмотрим непрерывную случайную величину ξ с плотностью распределения $f(x)$ и случайную величину $\eta=\phi(\xi)$ с плотностью распределения $g(y)$. По определению функция распределения $F_{\eta}(y)$ случайной величины η равна

$$F_\eta(y) = P(\eta < y) = \int_D f(x)dx = \left|D = \{x : \phi(x) < y\}\right| = \int_{\phi(x)<y} f(x)dx .$$

Числовые характеристики находятся по формулам:

$$m_y = \begin{cases} \sum_k \varphi(x_k) P(X = x_k), \text{ если } X - \text{дискретная } CB; \\ \int_{-\infty}^{\infty} \varphi(x) f(x) dx, \text{ если } X - \text{непрерывная } CB. \end{cases}$$

$$D_y = \begin{cases} \sum_k (\varphi(x_k) - m_y)^2 P(X = x_k), \text{ если } X - \text{дискретная } CB; \\ \int_{-\infty}^{\infty} (\varphi(x) - m_y)^2 f(x) dx, \text{ если } X - \text{непрерывная } CB. \end{cases}$$

Для случайного вектора $\bar{\xi} = (\xi_1, \xi_2)$ с плотностью распределения $f(x, y)$ если $\eta = \phi(\xi_1, \xi_2)$, то

$$M\eta = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(x, y) f(x, y) dxdy.$$

Пример 1. Найти закон распределения СВ. Y=(X-*m*)/σ, если СВ X подчиняется нормальному закону N(*m*,σ).

Решение. В нашем случае Y=(1/σ)X-*m*/σ, т.е. *a*=1/σ, *b*=-*m*/σ. Вспоминая формулу для плотности распределения нормального закона N(*m*,σ) и учитывая формулу (3), получаем

$$f(y) = \frac{1}{m/\sigma}\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{\left[\sigma(y + m/\sigma) - m\right]^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-y^2/2} ,$$

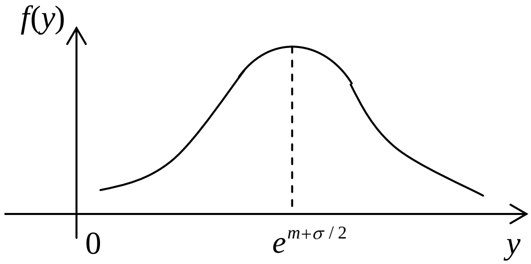
т.е. СВ Y распределена по закону N(0,1).

Отсюда $\mathrm{N}(m,\sigma) \xrightarrow{Y=(X-m)/\sigma} \mathrm{N}(0,1)$

Пример 2. *Логарифмически нормальное распределение*.

СВ X распределена нормально по закону N(m,σ). Найти плотность распределения СВ Y==φ(X); $y = \varphi(x) = e^x$ отсюда $x = \ln y = \varphi^{-1}(y)$ и следовательно,

$$\mathrm{N}\ \ m,\sigma\ \ \to f\ \ x\ \ =\ 1/\sigma\sqrt{2\pi}\ \ \mathrm{e}^{-\ \ x-m\ \ /2\sigma}.$$



Функция e^x – монотонно возрастающая всюду, поэтому находим $f\ \ y\ \ = \dfrac{1}{\sigma\sqrt{2\pi}}\dfrac{1}{y}\mathrm{e}^{\ \ln y-m\ \ /2\sigma}$.

$$\mathrm{M}\ \ y\ \ = \mathrm{e}^{m+\sigma/2}, \mathrm{D}\ \ y\ \ = \mathrm{e}^{2m+\sigma}\ \ \mathrm{e}^{\sigma}-1\ \ .$$

Пример 3. Распределение Пирсона χ_n^2 с n степенями свободы. Пусть $\xi_i \in N(0,1)$ и независимы. Тогда $\chi_n^2 = \xi_1^2 + \ldots + \xi_n^2$ имеет плотность распределения

$$f_n(x) = \begin{cases} \dfrac{1}{2^{n/2}\Gamma(n/2)} x^{\frac{n}{2}-1} \cdot e^{-x/2}, x > 0 \\ \qquad\qquad 0, \qquad x < 0, \end{cases}$$

где $\Gamma(\alpha) = \int\limits_0^\infty x^{\alpha-1}e^{-x}dx$ – Гамма-функция. $M\chi_n^2 = n$,

$D\chi_n^2 = 2n_0$.

Пример 4. Распределение Стьюдента с n степенями свободы $T(n)=u\sqrt{\dfrac{n}{\text{v}}}$, где $u\in N(0,1)$, $\text{v}=\chi_n^2$, u,v – независимые случайные величины, имеет плотность распределения

$$f_n(t)=\frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\sqrt{\pi n}}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}.$$

$$MT_n=0,\qquad DT_n=\frac{n}{n-2},\quad n>2.$$

Пример 5. Плотность распределения случайной величины ξ равна $f(x)=\dfrac{1}{\pi(1+x^2)}$. Найти плотность распределения $g(y)$ случайной величины $\eta=1/\xi$.

Решение. Решение задачи располагаем в виде двух столбцов; слева будем писать обозначения функций, принятые в общем случае; справа – конкретные функции, соответствующие данному примеру:

$f \quad x$	$f(x)=\dfrac{1}{\pi \quad 1+x^2}$
$y=\phi \quad x$	$y=1/x$
$x=\psi(y)$	$x=1/y$
$x'=\psi' \quad y$	$x'=-\dfrac{1}{y^2}$
$x\in \quad -\infty,+\infty$	$y\in \quad -\infty,+\infty$
$g \quad y$	$g \quad y \quad =\dfrac{1}{\pi \quad 1+1/y^2}\cdot\dfrac{1}{y^2}=\dfrac{1}{\pi \quad 1+y^2}.$

Пример 6. Бросаются 3 монеты. Пусть $\xi_i = 1$, если $i-$ая монета выпала орлом вверх и $\xi_i = 0$ в противном случае, $i = 1, 2, 3$. Построить ряд распределения случайной величины $\eta = \xi_1 + \xi_2 - \xi_3$.

Решение.

1. Определяем пространство элементарных исходов.

Элементарными исходами рассматриваемого случайного эксперимента являются упорядоченные наборы чисел n_1, n_2, n_3, где n_i либо нуль, либо единица $i = 1, 2, 3$.

2. Определяем множество возможных значений η.

Случайная величина η на элементарном исходе (n_1, n_2, n_3) принимает значение $\eta = n_1 + n_2 - n_3$.

3. Составляем табл. 9 элементарных исходов и соответствующих им значений η.

Таблица 9

n_1	0	1	0	0	1	1	0	1
n_2	0	0	1	0	1	0	1	1
n_3	0	0	0	1	0	1	1	1
η	0	1	1	–1	2	0	0	1

4. Определяем вероятности значений η и строим ее ряд распределения. Всего элементарных исходов $2^3 = 8$. Следовательно, вероятность элементарного исхода равна $1/8$. Имеем

η	–1	0	1	2
p	1/8	3/8	3/8	1/8

Задачи для самостоятельного решения

1. Случайная величина X имеет закон распределения:

X	-1	0	2	3
P	0,1	0,4	0,3	0,2

Найдите закон распределения случайной величины $Y = (X-1)^2$.

2. Случайная величина X распределена равномерно на отрезке $[a,b]: f(x) = \frac{1}{b-a}, 0 < a \le x \le b.$ Найдите плотность распределения случайной величины $Y = X^2$.

3. Случайная величина X имеет плотность вероятности $f(x) = \frac{1}{\sigma^2}\exp\left(\frac{x^2}{2\sigma^2}\right), x \ge 0$ (закон распределения Релея). Найдите плотность вероятности случайной величины $Y = \delta/X$.

4. На окружности единичного радиуса с центром в начале координат наугад выбирают точку. В выбранной точке проводят касательную к окружности. Найдите функцию распределения длины этой касательной от точки касания до точки ее пересечения с осью Оx.

5. Пусть случайная величина X равномерно распределена на отрезке $[0;1]$. Найдите функцию распределения случайной величины $Y = \ln(1/X)$.

6. Дискретная случайная величина задана распределением:

X	1	2	4
P	0,3	0,5	0,2

Найдите математическое ожидание случайной величины $Y = X^2 - X + 1$

7. Пусть X – число выпавших гербов при трех подбрасываниях монеты. Найдите математическое ожидание случайной величины $Y = X^2$.

8. Случайная величина X равномерно распределена на отрезке $(0;1)$. Найдите математическое ожидание случайной величины $Y = \sin(\pi X)$.

9. Непрерывная случайная величина X имеет функцию плотности вероятности: $f(x) = 0,5\sin x$ при $x \in (0,\pi)$ и $f(x) = 0$ при остальных x. Найдите плотность вероятности и математическое ожидание случайной величины $Y = 2X$.

10. Пусть случайная величина X имеет стандартное нормальное распределение с функцией плотности вероятности $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Найдите функцию плотности вероятности случайной величины $Y = \dfrac{1}{2} X^2$.

11. Случайная величина X имеет функцию плотности вероятности $f(x) = \lambda e^{-\lambda x}, \lambda > 0, x \ge 0$. Найдите закон распределения случайной величины $Y = \varphi(x)$, где $\varphi(x) = 0,5x$ при $x \in \left|0,4\right|$ и $\varphi(x) \equiv 2$ при $x \in (4,+\infty)$.

12. Две вершины треугольника совпадают с концами диаметра единичного круга, а третья вершина треугольника располагается в случайной точке внутри круга. Полагая равновозможными все положения случайной точки внутри круга, найдите функцию плотности вероятности площади треугольника S и математическое ожидание этой площади.

Ответы.

1.

X	1	4
P	0,7	0,3

2. $g(h)=\dfrac{1}{2(b-a)\sqrt{y}}, a^2\le x\le b^2$;

3. $g(y)=y\cdot\exp\left(\dfrac{-y^2}{2}\right), y\ge 0$; 4. $F(x)=\dfrac{2}{\pi}\operatorname{arc}tgx$ при $x>0$,

$F(x)=0$ при $x\ge 0$; 5. $F(x)=1-e^{-x}$; 6. 4,4; 7. 3; 8.1/2;

9. $f(y)=0.25\sin(y/2)$ при $y\in 0,2\pi$, $f(y)=0$ при осталь-ных y; $M(Y)=\pi$; 10. $g(x)=\dfrac{1}{\sqrt{\pi y}}e^{-y}$;

11. $g(y)=2\lambda e^{-2\lambda x}+e^{-4\lambda}\delta(y-2)$ при $y\in|0,2|$ и $g(y)=0$ при $y>2$; 12. $f(x)=\dfrac{4}{\pi}\sqrt{1-x^2}, M(S)=\dfrac{3}{4\pi}$.

ЗАНЯТИЕ № 13. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА И СЛЕДСТВИЯ ИЗ НЕЕ. ЗАКОН БОЛЬШИХ ЧИСЕЛ. НЕРАВЕНСТВО ЧЕБЫШЕВА

Пусть случайная величина ξ имеет конечную дисперсию $D\xi$. Тогда для любого $\varepsilon>0$ справедливо неравенство Чебышева

$$P(|\xi-M\xi|\ge\varepsilon)\le\frac{D\xi}{\varepsilon^2} \quad \text{или} \quad P(|\xi-M\xi|<\varepsilon)\ge 1-\frac{D\xi}{\varepsilon^2}.$$

Теорема Чебышева

Пусть случайные величины $\xi_1, \xi_2, \ldots \xi_n, \ldots$ независимы, существуют $M_{\xi_i} = m_i$, $D\xi_i = \sigma_i^2$ и $D\xi_i \le c$, $i = 1, \ldots, n, \ldots$, c – некоторая постоянная.

Тогда для любого $\varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum \xi_i - \frac{1}{n}\sum M\xi_i\right| < \varepsilon\right\} = 1.$$

В частности, если все ξ_i имеют одно и то же математическое ожидание и дисперсию $M\xi_i = m,\ D\xi_i = \sigma^2$, то

$$P\left\{\left|\frac{1}{n}\sum \xi_i - m\right| < \varepsilon\right\} \ge 1 - \frac{\sigma^2}{n\varepsilon^2}.$$

Для биномиального распределения

$$P\left\{\left|\frac{m}{n} - p\right| < \varepsilon\right\} \ge 1 - \frac{pq}{nq^2}.$$

Здесь p – вероятность появления события A в одном испытании, $q = 1 - p$, n – общее число испытаний, m – число испытаний, в которых событие А произошло.

Центральная предельная теорема.

Пусть $X_1, X_2, X_3, ..., X_n$ - последовательность независимых одинаково распределенных случайных величин. Если дисперсии случайных величин конечны и отличны от нуля, то при допустимых n закон распределения суммы $X_1 + X_2 + X_3 + ... + X_n$ сколь угодно близок к нормальному закону распределения. Это означает, что

$$P\left(\frac{\sum_{i=1}^{n} X_i - nm}{\sigma\sqrt{n}} < x\right) \xrightarrow{n\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dx,$$

где $m = M(X_i)$, $D(X_i) = \sigma^2$.

Следствие 1. Пусть k число появлений события в n независимых опытах, в каждом из которых вероятность появления события равна p $(0 < p < 1)$. Тогда при достаточно больших n (порядка десятков, сотен и т.д.) имеют место следующие формулы:

$$P_n(k) \approx \frac{1}{\sqrt{2\pi npq}} \exp\left\{-\frac{(k-np)^2}{2npq}\right\}, \qquad (13.1)$$

$$Pn(k_1 \le k \le k_2) \approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right). \qquad (13.2)$$

Первая формула дает приближенное значение вероятности того, что событие появится k раз в n опытах, и составляет содержание локальной теоремы Муавра-Лапласа. Вторая формула позволяет вычислить вероятность того, что в n независимых опытах событие появится от k_1 до k_2 раз. Эта формула основана на интегральной теореме Муавра-Лапласа. Обе формулы вытекают из того факта, что при большом числе независимых опытов число появлений события имеет (на основании центральной предельной теоремы) закон распределения близок к нормальному: $N(np; npq)$.

Следствие 2. Пусть k/n частота появлений события в n независимых опытах, в каждом из которых вероятность появления события равна $p(0 < p < 1)$. Тогда при достаточно больших n (порядка десятков, сотен и т.д.) имеет место следующая формула:

$$P\ |k/n-p|<\alpha\ \approx 2\Phi\left(\frac{\alpha}{\sqrt{pq/n}}\right). \qquad (13.3)$$

Пример 1. При изготовлении некоторой детали брак равен 5%. Оценить вероятность того, что при просмотре партии в 2000 штук выявляется отклонение доли бракованных деталей от установленного процента брака меньше чем на 1%.

Решение. Воспользуемся формулой

$$P\left\{\left|\frac{m}{n}-p\right|<\varepsilon\right\}\geq 1-\frac{pq}{n\varepsilon^2}.$$

Здесь $p=0,05$, $q=1-p=0,95$, $\varepsilon=0,01$, $n=2000$.

Тогда $P\left\{\left|\frac{m}{n}-0,05\right|<0,01\right\}\geq 1-\frac{0,05\cdot 0,95}{2000\cdot(0,01)^2}=0,7625$.

Пример 2. Сколько нужно произвести измерений, чтобы с вероятностью равной 0,95 утверждать, что погрешность средней арифметической результатов этих измерений не превысит 0,1, если $\sigma=2$.

Решение. Воспользуемся формулой

$$P\left\{\left|\frac{1}{n}\sum\xi_k-m\right|<\varepsilon\right\}\geq 1-\frac{D\xi}{n\varepsilon^2}.$$

Здесь $\sigma=2$, $\varepsilon=0,1$.Имеем $0,95=1-\frac{2^2}{n\cdot(0,1)^2}=1-\frac{400}{n}$,

$\frac{400}{n}=0,05$, $n=\frac{400}{0,05}=8000$.

Пример 3. Поезда метро идут с интервалами 2 минуты. Каждый из пассажиров независимо от других приходит на платформу в случайный момент времени и ожидает ближайшего поезда. В данный поезд село 75 пассажиров. Какова вероятность того, что их суммарное время ожидания превысило один час?

Решение. Обозначим время ожидания i-го пассажира через X_i. Естественно предположить, что равновозможен приход каждого пассажира в любой момент времени. Это означает, что случайная величина X_i имеет равномерный закон распределения с функцией плотности вероятности $f(x)=1/2$ при $x \in [0,2]$ и $f(x)=0$ при $x \notin [0,2]$. Тогда

$$M(X_i)=\int_0^2 x\frac{1}{2}dx=1 \text{ и } D(X)=\int_0^2 (x-1)^2\frac{1}{2}dx=\frac{1}{3}.$$

Суммарное время ожидания $Y=\sum_{i=1}^{75} X_i$ представляет собой сумму большего числа независимых одинаково распределенных случайных величин с ограниченными дисперсиями. В силу центральной предельной теоремы можно утверждать, что Y имеет закон распределения близкий к нормальному. Нормальный закон распределения определяется математическим ожиданием и дисперсией. Вычислим их:

$$M(Y)=M(\sum_1^{75} X_i)=\sum_1^{75} M(X_i)=75\cdot 1=75.$$

$$D(X)=D(\sum_1^{75} X_i)=\sum_1^{75} D(X_i)=75\cdot\frac{1}{3}=25.$$

В итоге можно утверждать, что случайная величина Y имеет закон распределения близкий к $N(75;25)$. Искомая вероятность может быть вычислена по формуле (13.2):

$$P(\eta > 60) = P(60 < \eta < 150) = \Phi\left(\frac{150-75}{5}\right) - \Phi\left(\frac{60-75}{5}\right) =$$

$$= \Phi(15) + \Phi(3) = 0,5 + 0,4986 = 0,9986 \approx 1.$$

Пример 4. Регулировка прибора занимает время от 4 до 10 минут. Регулировщику предстоит отрегулировать 50 приборов. Считая для каждого прибора равновозможными все значения времени регулировки в указанных пределах, оценить вероятность того, что регулировщик справиться с работой за шесть часов.

Решение. Пусть ξ_i - время регулировки i-го прибора, а $Y = X_1 + X_2 + X_3 + ... + X_{50}$ - время выполнения работы рабочим. требуется найти $P(Y < 360)$. Величина Y является суммой большего числа одинаково распределенных независимых случайных величин, каждая из которых ограничена. По центральной предельной теореме Y имеет закон распределения близкий к нормальному закону распределения. Найдем параметры этого закона, т.е. математическое ожидание и дисперсию величины Y. Так как случайные величины X_i независимы, то

$$M(Y) = M(X_1) + M(X_2) + ... + M(X_{50});$$

$$D(Y) = D(X_1) + D(X_2) + ... + D(X_{50}).$$

Вычислим $M(X_i)$ и $D(X_i)$.

По условию все значения случайной величины X_i равновозможны в отрезке $[4,10]$. Поэтому функция плотности вероятности этой случайной величины в указанном отрезке постоянна. Чтобы площадь, заключенная между графиком функции плотности вероятности и осью абсцисс, равнялась единицы, следует положить $f(x)=1/6$ при $x\in[4,10]$ и $f(x)=0$ при остальных значениях x. С учетом этого имеем:

$$M(X_i)=\int_4^{10} x\frac{1}{6}dx=7, \qquad D(X_i)=\int_4^{10}(x-7)^2\frac{1}{6}dx=3.$$

Поэтому $M(Y)=50\cdot 7=250, D(Y)=50\cdot 3=150, \sigma(Y)=12,25$.

Итак, $Y \sim N(350;150)$. Для вычисления искомой вероятности воспользуемся формулой и таблицей функции Лапласа:

$$P(Y<360)=P(200<\eta<360)=\Phi\left(\frac{360-350}{12,25}\right)-\Phi\left(\frac{200-350}{12,25}\right)=$$

$$=\Phi(0,82)+\Phi(12,24)==0,294+0,5\approx 0,8.$$

Пример 5. Стрелок в десятку попадает с вероятностью 0,5, в девятку – с вероятностью 0,2, в восьмерку – с вероятностью 0,15, в семерку – с вероятностью 0,1 и в шестерку – с вероятностью 0,05. Какова вероятность того, что при 24 выстрелах стрелок наберет не менее 210 очков?

Решение. Пусть при i-м выстреле стрелок набивает X_i очков. Величины X_i независимы и каждая имеет распределение:

X_i	6	7	8	9	10
P	0,05	0,1	0,15	0,2	0,5

Заметим, что

$$M(X_i)=6\cdot 0,05+7\cdot 0,1+8\cdot 0,15+9\cdot 0,2+10\cdot 0,5=9;$$

$$D(X_i)=(6-9)^2\cdot 0,05+(7-9)^2\cdot 0,1+(8-9)^2\cdot 0,15+$$

$$+(9-9)^2\cdot 0,2+(10-9)^2\cdot 0,5=1,5.$$

Сумма очков $Y=\sum_{i=1}^{24} X_i$, будучи суммой большого числа независимых одинаково распределенных слагаемых с ограниченными дисперсиями, имеет закон распределения близкий к нормальному с параметрами:

$$M(Y)=M\left(\sum_{i=1}^{24} X_i\right)=\sum_{i=1}^{24} M(X_i)=24\cdot 9=216;$$

$$D(Y)=D\left(\sum_{i=1}^{24} X_i\right)=\sum_{i=1}^{24} D(X_i)=24\cdot 1,5=36=\sigma^2.$$

В итоге $Y \sim N(216;36)$. Поэтому

$$P(210\le Y)=P(210\le Y\le 240)=\Phi\left(\frac{240-216}{6}\right)-$$

$$-\Phi\left(\frac{210-216}{6}\right)==\Phi(4)-\Phi(-1)=\Phi(4)+\Phi(1)=$$

$$=0,49997+0,34134\approx 0,84.$$

Пример 6. Восемьдесят процентов приборов после сборки нуждаются в регулировке. Какова вероятность того, что среди 400 собранных за смену приборов в регулировке нуждаются: а) не менее 310; б) не более 350; в) от 304 до 336?

Решение. Сборку каждого прибора можно считать независимым испытанием с вероятностью появления события равной $p=0,8$. Так как число опытов велико, то можно воспользоваться интегральной теоремой Муавра-Лапласа :

а) $P_{400}(310;400)=\Phi\left(\frac{400-400\cdot 0,8}{\sqrt{400\cdot 0,8\cdot 0,2}}\right)-\Phi\left(\frac{310-400\cdot 0,8\cdot 0,2}{\sqrt{400\cdot 0,8\cdot 0,2}}\right)=$

$$=\Phi(10)+\Phi(1,25)==0,5+0,3944=0,8944\approx 0,9;$$

б) $P_{400}(0;350)=\Phi\left(\dfrac{350-400\cdot 0,8}{\sqrt{400\cdot 0,8\cdot 0,2}}\right)-\Phi\left(\dfrac{0-400\cdot 0,8\cdot 0,2}{\sqrt{400\cdot 0,8\cdot 0,2}}\right)=$

$=\Phi(3,75)+\Phi(40)==0,4999+0,5=0,9999\approx 1;$

в) $P_{400}(304;336)=\Phi\left(\dfrac{336-400\cdot 0,8}{\sqrt{400\cdot 0,8\cdot 0,2}}\right)-\Phi\left(\dfrac{304-400\cdot 0,8\cdot 0,2}{\sqrt{400\cdot 0,8\cdot 0,2}}\right)=$

$=\Phi(2)+\Phi(2)==2\Phi(2)=0,9545.$

Правило *«Трех сигм»*: для случайной величины X, распределенной по нормальному закону распределения $N(m;\sigma^2)$,

$$P(|X-m|<3\sigma)=2\Phi\left(\frac{3\sigma}{\sigma}\right)=2\Phi(3)=0,997\approx 1.$$

Поэтому интервалом практически возможных значений такой случайной величины считают интервал $(m-3\sigma;m+3\sigma)$..

Напомним, что если вероятность события близка к единице, то событие называют *практически достоверным*. Можно быть практически уверенным, что в единичном опыте оно произойдет.

Пример 7. В страховой компании застраховано 10 000 автомобилей. Вероятность поломки любого автомобиля в результате дорожно-транспортного происшествия равна 0,02. Каждый владелец застрахованного автомобиля платит в год 24 у.е. страховых и в случае поломки автомобиля в результате аварии получает от компании 1000 у.е. Найдите вероятность того, что по истечении года работы компания потерпит убытки от этого вида страховой деятельности.

Решение. Страховой сбор с 10 000 автовладельцев составляет $24\cdot 10000=240000$ у.е. Компания потерпит убытки, если будет предъявлено более 240 исков по 1000 у.е. каждый. Вероятность поступления страхового иска от каждого автовладельца равна 0,02. Эксплуатация каждого автомобиля в течение страхового срока можно считать независимым испытанием. Так как число испытаний велико $(n = 10000)$, то можно воспользоваться интегральной теоремой Муавра-Лапласа.

$$P_{10000}\quad 240 \le k \le 10000 \quad =$$

$$= \Phi\left(\frac{10000 - 10000 \cdot 0,02}{\sqrt{10000 \cdot 0,02 \cdot 0,98}}\right) - \Phi\left(\frac{240 - 10000 \cdot 0,02}{\sqrt{10000 \cdot 0,02 \cdot 0,98}}\right) =$$

$$= \Phi(700) - \Phi(2,86) = 0,5 - 0,4979 = 0,0021.$$

Задачи для самостоятельного решения

1. Монету подбросили 900 раз. Герб выпал 403 раза. Можно ли считать, что подбрасывали симметричную монету?

2. В условиях примера 7 найдите вероятность того, что страховая фирма получит доход менее 60 000 у.е.

3. Сколько раз нужно подбросить монету, чтобы с вероятностью, не меньшей 0,9, утверждать, что частота выпадения гербы попадает в интервал (0,4;0,6)? Получить оценку числа бросков монеты: а) по неравенству Чебышева; б) с использованием следствия 10.2 из центральной предельной теоремы.

4. Время безотказной работы предохранителя X имеет показательный закон распределения (функция распределения $F(x) = 1 - e^{-\lambda x}, M(X) = 1/\lambda, D(X) = 1/\lambda^2$) с параметром $\lambda = 0,01$ отказов в час. Перегоревший предохранитель практически мгновенно заменяется новым. Какова вероятность того, что 20 предохранителей хватит на 2500 часов работы?

5. При дальней радиосвязи из-за помех 10% сигналов искажаются и принимаются неверно. Найдите вероятность того, что

при передаче 50 сигналов ошибок в приеме будет не более трех.

6. Вероятность поражения цели стрелком при одном выстреле равна 0,8. Найдите вероятность того, что при 100 выстрелах мишень будет поражена: а) не менее 75 раз; б) от 70 до 90 раз; в) не более 82 раз.

7. Сорок процентов жителей нашего города поддерживают некоторое мероприятие. Для изучения общественного мнения было опрошено 400 взятых наугад жителей. Какова вероятность того, что больше половины из опрошенных выскажутся в поддержку мероприятия?

8.Наблюдается простейший поток событий интенсивности μ (интервалы X_i между событиями независимы и имеют показательное распределение с функцией плотности вероятности $f(x) = \mu e^{-\mu x}$). Оцените вероятность того, что первые 100 событий потока произойдут в интервале времени от 90/μ до 100/μ .(Указание: $M(X_i) = 1\mu$, $D(X_i) = 1\mu^2$.)

9. Длительность телефонного разговора случайна. Известно, что у данного абонента средняя длительность разговора равна 4 мин, а среднее квадратическое отклонение длительности разговора равна 2 мин. Оцените вероятность того, что длительность 50 разговоров превысит 3 часа.

10. В крупной партии изделий 1% изделий обладает скрытыми дефектами. Оценить вероятно того, что среди взятых наугад 400 окажется не более k изделий со скрытыми дефектами. Ответ для $k = 4, k = 6, k = 8$.

11. Известно, что 5% студентов носят очки. На первый курс данного факультета принято 250 студентов. Какова вероятность того, что среди них не менее 15 носят очки?
12. Игральный кубик подбрасывают 15 раз. Оцените вероятность того, что суммарное число выпавших очков превысит 50.

Ответы: 1.Нет; 2. 0,92; 3. а) больше 250, б) больше 68; 4. 0,13; 5. 0,16; 6. а) $\approx 0{,}89$, б) $\approx 0{,}99$, в) $\approx 0{,}69$; 7. $\approx 0{,}02$; 8. $\approx 0{,}68$; 9. 0,92; 10. 0,5, 0,86, 0,998; 11. 0,24; 12. 0,96.

ЗАНЯТИЕ №14. ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

14.1. Основные определения

Математической статистикой называется наука, которая основывается на методах теории вероятностей, занимается систематизацией, обработкой и использованием экспериментальных данных для получения научных и практических выводов.

Одним из основных методов исследования случайных явлений в математической статистике является *выборочный метод*. Рассмотрим основные понятия этого метода.

Пусть рассматривается некоторый случайный эксперимент, связанный с СВ X, имеющей ФР F(x). Полный набор всех возможных результатов измерений СВ X в эксперименте называют *генеральной совокупностью* (ГС) с ФР F(x).

Число членов N, образующих генеральную совокупность, называют *объемом генеральной совокупности*.

Отметим, что объем генеральной совокупности может быть как конечным, так и бесконечным.

Выборкой (*выборочной совокупностью* (ВС)) объемом n из N генеральной совокупности (ГС) называется последовательность $x_1,x_2,...,x_n$ наблюдаемых значений СВ X, соответствующих n независимым повторениям эксперимента .

Метод, состоящий в том, что на основании изучения характеристик и свойств выборки $x_1,x_2,...,x_n$ даются заключения о числовых характеристиках и законе распределения СВ X, называется *выборочным методом*. Выборка может быть записана в виде вариационного ряда или в виде статистического ряда.

Вариационным рядом выборки $x_1,x_2,...,x_n$ называется способ ее записи, при котором элементы x_i упорядочиваются по величине, то есть записываются в виде последовательности $x^1, x^2, ..., x^n$, причем $x^1 \le x^2 \le ... \le x^n$.

Разность между максимальным и минимальным элементами выборки $x^n - x^1 = \omega$ называется *размахом выборки*.

Пусть в выборке объемом n элемент x_i встречается n_i раз. Число n_i называется *частотой элемента* x_i. Очевидно, что

$$n = \sum_{i=1}^{k} n_i \ .$$

Статистическим рядом называется последовательность пар (x_i,n_i), которая записывается обычно в виде табл. 10

Таблица 10

x_i	x_1	x_2	…	x_k	
i	n_1	n_2	…	n_k	$\sum_{i=1}^{k} n_i = n$

Отношение $\omega_i=n_i/n$ называется *относительной частотой*, или *частостью элемента* x_i выборки.

Статистическим распределением СВ X называется последовательность пар $(x_i, n_i / n)$, которая также записывается в виде табл. 11

Таблица 11

x_i	x_1	x_2	…	x_k	
$\omega_i=\frac{n_i}{n}$	$\frac{n_1}{n}$	$\frac{n_2}{n}$	…	$\frac{n_k}{n}$	$\sum_{i=1}^{k}\frac{n_i}{n}=1=\sum_{i=1}^{k}\omega_i$

При большом объеме выборки ее элементы объединяют в группы (разряды), представляя результаты опытов в виде группированного статистического ряда. Для этого интервал, содержащий все элементы выборки, разбивается на k частичных непересекающихся интервалов.

Обычно выбирают частичные интервалы одинаковой длины b=ω/k. После того, как частичные интервалы выбраны, определяют частоты - количество n_i^* элементов выборки, попавших в i-тый интервал (элемент, совпадающий с верхней границей интервала, относится к следующему интервалу). В группированный статистический ряд в верхнюю строку записываются середины x_i^* интервалов группировки, а в нижний - частоты n_i^*.

В зависимости от объема выборки число k интервалов группировки берется от 6 до 20. Наряду с частотами n_i^* удобно одновременно подсчитывать также *накопленные частоты* $\sum_{j=1}^{i} n_j{}^*$, *относительные частоты* и *накопленные относительные частоты* $\sum_{j=1}^{i}\omega_j^*$, $\quad i=1,2,...,k$.

Полученные результаты сводятся в таблицу, называемую таблицей частот группированной выборки.

Следует отметить, что группировка выборки вносит погрешность в последующие вычисления, которая становится тем больше, чем меньше выбирается число интервалов.

Пример 14.1. Дана выборка из некоторой ГС
11,15,12,9,13,
12,6,11,12,13,15,8,9,14,9,11,6. Определить объем, размах выборки, а также построить вариационный и статистический ряды.

Решение. Объем n=17.Вариационный ряд 6,6,8,9,9,11,11,11, 12,12,12,13,13,14,15,15. Размах ω=15-6=9.Статистический ряд и статистическое распределение имеют вид:

x_i	6	8	9	11	12	13	14	15
n	2	1	3	3	3	2	1	2
ω_i	2/17	1/17	3/17	3/17	3/17	2/17	1/17	2/17

$$\sum n_i = n; \qquad \sum \omega_i = 1$$

Пример 14.2. Время решений контрольной работы студентами 2-го курса дается выборкой

38	60	41	51	33	42	45	21	53	60
68	52	47	46	49	49	10	57	54	59
79	47	28	48	58	32	42	58	61	30
61	35	47	72	41	45	44	55	30	40
67	65	39	48	43	60	54	42	59	50

Построить выборку в виде табл. 12 частот группированной выборки, используя 7 интервалов группировки.

Решение. Размах выборки 79-10=69. Длина интервала

b=69/7$\cong$10.

Таблица 12

Инт.	[10-20)	[20-30)	[30-40)	[40-50)	[50-60)	[60-70)	[70-80)	
x_i^*	15	25	35	45	55	65	75	
n_i^*	1	2	7	18	12	8	2	$\sum n_i = 50$
$\sum n_i^*$	1	3	10	28	40	48	50	
ω_i	0,02	0,04	0,14	0,36	0,24	0,16	0,04	$\sum \omega_i = 1$
$\sum \omega_i^*$	0,02	0,06	0,20	0,56	0,80	0,96	1,00	

14.2. Графическое представление выборки

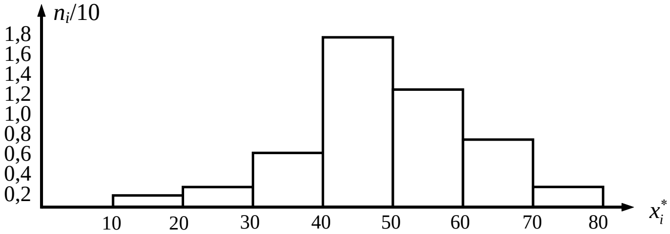
Для наглядности сгруппированные статистические ряды представляются графиками и диаграммами.

Полигоном частот группированной выборки называется ломаная с вершинами в точках (x_i^*, n_i^*), i=1, 2,...,k.

Полигоном относительных частот группированной выборки называется ломаная с вершинами в точках (x_i^*, ω_i^*).

Гистограммой частот группированной выборки называется ступенчатая фигура, составленная из прямоугольников, построенных на интервалах так, что площадь каждого прямоугольника равна частоте n_i^*, i=1, 2,...,k. Отсюда следует, что площадь гистограммы частот равна объему выборки n. В том случае, когда длины всех интервалов одинаковы и равны b, высоты прямоугольников равны $h_i = n_i^* / b$, i=1, 2,...,k. Аналогично строится гистограмма *относительных частот*. Площадь гистограммы относительных частот равна единице.

Гисто-
грамма



Полигоном накопленных частот группированной выборки называется ломаная с вершинами в точках ($x_i^* + b/2; \sum_{j=1}^{i} n_j^*$).

Полигоном относительных накопленных частот (кумулятивной кривой, кумулятой) называется ломаная с вершинами в точках ($x_i^* + b/2;\ \ (\sum_{j=1}^{i} n_j^*)/n$).

Замечание. Перечисленные графические представления аналогичным образом определяются и в случае негруппированной выборки.

Пример 14.3. Для выборки примера 14.2 построить гистограмму, полигон частот и кумулятивную кривую.

14.3. Эмпирическая функция распределения (ЭФР)

Эмпирической функцией распределения СВ Х называется функция F*(x), определяющая для каждого значения x относительную частоту события (X<x) $F^*(x)=n_x/n$, где n_x - число выборочных значений, меньших x, а n - объем выборки.

По значениям накопленных относительных частот ЭФР определяется следующим образом:

$$F^*\ x\ = \sum_{x_i<x} n_i\,/\,n\,.$$

В отличие от ЭФР, функция распределения генеральной совокупности F(x)=P(X<x) называется *теоретической функцией распределения* (ТФР).

Отметим, что разница между ТФР и ЭФР состоит в том, что ТФР определяет вероятность события (X<x), а ЭФР определяет относительную частоту этого же события.

ЭФР обладает всеми свойствами ТФР, то есть:

1) значения F*(x) принадлежат отрезку [0,1];

2) F*(x)- неубывающая функция аргумента x;

3) F*(x)=0, если $x < x_1$, и F*(x)=1, если $x > x_n$, где x_1 - наименьшее, а x_n - наибольшее наблюдаемые значения СВ X .

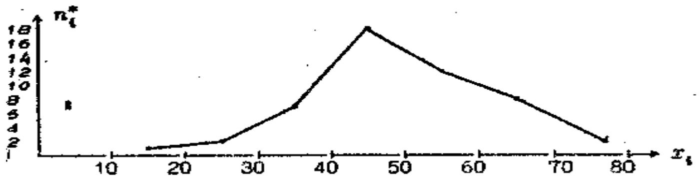
Из закона больших чисел, а именно из теоремы Бернулли, следует, что при объеме выборки $n \to \infty$ ЭФР сходится по вероятности к ТФР. Это означает, что при достаточно большом объеме выборки ЭФР F*(x) и ТФР мало отличаются друг от друга.

Основное значение ЭФР состоит в том, что она используется в качестве оценки ТФР.

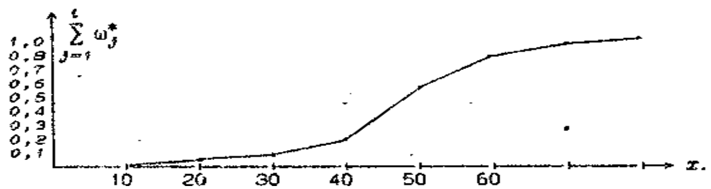
Пример 14.4. Построить график ЭФР по выборке примера 14.2.

Решение. ЭФР имеет вид

Полигон

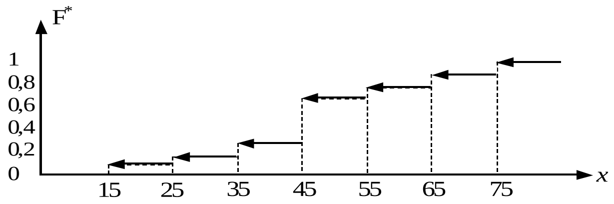


Кумулята



$$F^*(x)=\begin{cases}0 & x\le 15;\\ 0{,}02 & 15<x\le 25;\\ 0{,}06 & 25<x\le 35;\\ 0{,}20 & 35<x\le 45;\\ 0{,}56 & 45<x\le 55;\\ 0{,}80 & 55<x\le 65;\\ 0{,}96 & 65<x\le 75;\\ 1 & 75<x.\end{cases}$$

График $F^*(x)$ имеет вид



Задачи для самостоятельного решения

14.1. В течение суток измеряют напряжение X тока в электросети в вольтах. В результате опыта получена выборка объема $n = 30$:

107	108	110	109	110	111	109	110	111	107
108	109	110	108	107	110	109	111	111	110
109	112	113	110	106	110	109	110	108	112

Построить статистический ряд этой выборки. Построить полигон относительных частот. Найти эмпирическую функцию и построить ее график.

14.2. По данному распределению выборки

x_i	1	3	6
n_i	10	25	15

найти эмпирическую функцию и построить ее график.

14.3. Дана выборка:

x_i	2	4	5	7	10
n_i	15	20	10	10	45

Найти эмпирическую функцию распределения, построить ее график. Построить полигон относительных частот выборки.

14.5. Измерен рост n=500 студентов. Результаты измерений представлены в виде интервального статистического ряда:

[145;150)	[150;155)	[155;160)	[160;165)	[165;170)
1	2	28	90	169
[170;175)	[175;180)	[180;185)	[185;190)	[190;195]
132	55	16	6	1

Построить гистограмму относительных частот.

14.6. По данным выборки построить гистограмму относительных частот:

1.

Номер интервала	Интервал	Число вариант в интервале
1	(1;5)	10
2	(5;9)	20
3	(9;13)	50
4	(13;17)	12
5	(17;21)	8

2.

Номер интервала	Интервал	Число вариант в интервале
1	(2;5)	6
2	(5;8)	10
3	(8;11)	5
4	(11;14)	4

14.7. Дана выборка:

38	60	41	51	33	42	45	21	53	60
68	52	47	46	42	43	57	44	54	59
77	47	28	27	49	49	14	28	61	30
61	35	47	46	58	45	42	21	30	40
67	65	39	35	41	60	54	42	59	60

Построить гистограмму относительных частот.

14.8. Построить полигон относительных частот следующей выборки:

x_i	4	6	10	12
n_i	10	15	5	20

ЗАНЯТИЕ №15. СТАТИСТИЧЕСКАЯ ОЦЕНКА НЕИЗВЕСТНЫХ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ. ТОЧЕЧНЫЕ ОЦЕНКИ

15.1. Постановка задачи

При обработке опытных данных часто бывает так, что вид закона распределения генеральной совокупности (ЗРГС) известен, а требуется найти только некоторые параметры, от которых он зависит. Например, если известно, что СВ X распределена по нормальному закону, то на основании опытных данных (по выборке) необходимо "оценить", то есть найти приближенное значение двух параметров - МО и СКО.

Одна из задач математической статистики и состоит в нахождении оценок неизвестных параметров по выборке наблюдений.

Пусть из ГС с ФР F(x,θ), где θ - неизвестный параметр, произведена выборка x_1, x_2,…,x_n объемом n. В качестве оценки параметра θ рассматривают *функции элементов выборки* $\tilde{\theta} = \mathrm{U}(x_1, x_2, ..., x_n)$, которые называются *статистиками*.

Задача оценки неизвестного параметра θ сводится к нахождению таких статистик (выборочных функций) $\tilde{\theta} = \mathrm{U}(x_1, x_2, ..., x_n)$, которые могут быть использованы для приближенного определения значения неизвестного параметра θ.

Оценки параметров подразделяются на точечные и интервальные.

15.2. Основные свойства точечных статистических оценок распределения

Точечная оценка параметра θ определяется одним числом $\tilde{\theta} = \mathrm{U}(x_1, x_2, ..., x_n)$.

Качество оценок характеризуется некоторыми *свойствами*. Сформулируем основные из них.

Свойство 1. Оценка $\tilde{\theta}$ называется *несмещенной*, если ее МО равно оцениваемому параметру, то есть если М($\tilde{\theta}$)=θ. Разность М($\tilde{\theta}$)-θ называется *смещением*.

Свойство 2. Оценка $\tilde{\theta}(n) = \mathrm{U}(x_1, x_2, ..., x_n)$ называется *состоятельной*, если при увеличении объема выборки n оценка $\tilde{\theta}(n)$ сходится по вероятности к θ, то есть $\forall \varepsilon > 0$ $\lim \mathrm{P}\left(\left|\tilde{\theta}(n) - \theta\right| < \varepsilon\right) = 1$.

Свойство 3. Пусть $\tilde{\theta}_1$ и $\tilde{\theta}_2$ - две различные несмещенные
оценки параметра θ. Если D($\tilde{\theta}_1$)<D($\tilde{\theta}_2$), то говорят, что оценка $\tilde{\theta}_1$ *более эффективна*, чем оценка $\tilde{\theta}_2$.

Требование несмещенности устраняет систематические ошибки в определении оценок, обусловленных ограниченным объемом выборки.

Требование состоятельности гарантирует от совершения грубых ошибок ε в определении θ при достаточно большом объема n выборки.

Свойство эффективности используется для выбора оценки, обладающей наименьшим разбросом.

15.3. Статистическая оценка МО

В качестве статистической оценки МО выбирается выборочное среднее.

Выборочным средним $\overline{x}$ называется среднее арифметическое элементов выборки

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i . \qquad (15.1)$$

Если x_i - варианты выборки, n_i -частоты вариант x_i, $i = \overline{1,k}$, $n = \sum_{i=1}^{k} n_i$ - объем выборки, то

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i \qquad (15.2)$$

Для группированной выборки это соотношение принимает вид

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} x_i^* n_i^* \,. \qquad (15.3)$$

Выборочное среднее $\overline{x}$ является несмещенной и состоятельной оценкой.

По поводу эффективности $\overline{x}$ заметим, что если СВ X распределена по нормальному закону, то выборочное среднее является эффективной оценкой МО.

15.4. Статистическая оценка дисперсии

В качестве статистической оценки дисперсии D(X) СВ X примем *выборочную дисперсию*

$$s^2 = \frac{1}{n}\sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2 , \qquad (15.4)$$

или

$$s^2 = \frac{1}{n}\sum_{i=1}^{k} n_i \left(x_i - \overline{x}\right)^2 \qquad (15.5)$$

для группированной выборки

$$s^2 = \frac{1}{n}\sum_{i=1}^{k} \left(x_i^* - \overline{x}\right)^2 \cdot n_i^* \,. \qquad (15.6)$$

Но оценка s^2 является смещенной. Для получения несмещенной оценки дисперсии выборочную дисперсию s^2 исправляют, умножая ее на множитель $n/(n-1)$.

Исправленная дисперсия

$$s_0^2 = \frac{n}{n-1}\cdot s^2 = \frac{1}{(n-1)}\cdot\sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2 \qquad (15.7)$$

Или

$$s_0^2 = \frac{n}{n-1}\cdot s^2 = \frac{1}{(n-1)}\cdot\sum_{i=1}^{k} n_i \left(x_i - \overline{x}\right)^2 \qquad (15.8)$$

является уже несмещенной оценкой дисперсии.

Непосредственно из определений следует, что

$$s_0^2 = \frac{1}{(n-1)} \cdot \sum_{i=1}^{k} n_i x_i^{\,2} - \frac{n}{n-1}\bar{x}^2. \qquad (15.9)$$

Величина s_o называется *исправленным средним квадратическим отклонением*.

Оценка s^2 (а вместе с ней и s_0^2) состоятельна.

Пример 15.1. Оценить математическое ожидание и дисперсию случайной величины X по результатам ее независимых наблюдений: 7, 3, 4, 8, 4, 6, 3.

Решение. По формулам (15.1) и (15.4) имеем

$$M\ X\ \approx \tilde{X} = \frac{7+3+4+8+4+6+3}{7} = 5,$$

$$D\ X\ \approx s^2 = \frac{\ 7-5\ ^2 + \ 3-5\ ^2 + \ 4-5\ ^2 + \ldots + \ 3-5\ ^2}{6} = \frac{25}{6} \approx 4{,}17.$$

Пример 15.2. Данные 25 независимых наблюдений случайной величины X представлены в сгруппированном виде:

Границы интервалов	5-7	7-9	9-11	11-13	13-15
Число наблюдений	2	4	9	7	3

Оценить математическое ожидание и дисперсию этой случайной величины.

Решение. Представителем каждого интервала можно считать его середину. С учетом этого формулы (15.3) и (15.6) дают следующие оценки:

$$M(X) \approx \overline{x} = \frac{6\cdot 2+8\cdot 4+10\cdot 9+12\cdot 7+14\cdot 3}{25} = \frac{260}{25} = 10,4;$$

$$D(X) \approx s^2 = \frac{(6-10,4)^2\cdot 2+(8-10,4)^2\cdot 4+\ldots+(14-10,4)^2\cdot 3}{24} = 5.$$

Пример 15.3. По выборке признака X, заданной:

x_i	45	50	55	60	65	70	75
n_i	4	6	10	40	20	12	8

найти выборочное среднее, дисперсию и среднее квадратическое отклонение.

Решение. Объем выборки $n = \sum_{i=1}^{7} n_i = 100.$ По формуле (15.2) выборочное среднее

$$\overline{x} = \frac{45\cdot 4+50\cdot 6+55\cdot 10+60\cdot 40+65\cdot 20+70\cdot 12+75\cdot 8}{100} = 61,7.$$

Для вычисления дисперсии составляем таблицу квадратов значений СВ *X*:

x_i^2	2025	2500	3025	3600	4225	4900	5625
n_i	4	6	10	40	20	12	8

По формуле (15.9) имеем

$$s_0^2 = \frac{1}{99}\cdot\sum_{i=1}^{7} n_i x_i^2 - \frac{100}{99}(61,7)^2 = 50,11,$$

откуда $s_0 = \sqrt{50,11} \approx 7,08.$

Получили несмещенные оценки для дисперсии и среднего квадратического отклонения. Соответствующие смещенные оценки $s^2 = 49,61$ и $s = 7,04.$

Пример 15.4. Проведено несколько измерений расстояния. Результаты измерений в метрах представлены в виде ряда:

i	x_i	i	x_i	i	x_i	i	x_i
1	1235,6	5	1238,5	9	1234,5	13	1234,3
2	1237,5	6	1234,2	10	1236,8	14	1237,5
3	1232,9	7	1235,9	11	1237,6	15	1235,4
4	1236,2	8	1233,3	12	1233,1	16	1234,7

Вычислить оценки математического ожидания, дисперсии и среднего квадратического отклонения измеренного расстояния.

Решение. Так как значение вариант x_i большие, то удобно ввести условные варианты $u_i = x_i - a$, где в качестве a возьмем среднее число 1235, т.е. a=1235. В итоге получим результат для условных вариант:

i	u_i	i	u_i	i	u_i	i	u_i
1	0,6	5	3,5	9	-0,5	13	-0,7
2	2,5	6	-0,8	10	1,8	14	2,5
3	-2,1	7	0,9	11	2,6	15	0,4
4	1,2	8	-1,7	12	-1,9	16	-0,3

Выборочное среднее в данном случае вычисляется по формуле

$$\overline{x} = a + \frac{1}{n}\sum_{i=1}^{k} n_i u_i = 1235 + \frac{1}{16}(16-8) = 1235,5.$$

Статистическая дисперсия

$$s_0^2 = \frac{1}{(n-1)} \cdot \sum_{i=1}^{7} n_i u_i^2 - \frac{n}{n-1}(\overline{x} - a)^2 = \frac{1}{n}\sum_{i=1}^{k} n_i (u_i - \overline{u})^2,$$

где $\overline{u}$ - среднее значение условных вариант. Отсюда

$$s_0^2 = \frac{1}{15}(0,01 + 4 + 6,76 + 0,49 + 9 + 1,69 + 0,16 + 4,48 + 1 +$$

$$+1,69 + 4,41 + 5.76 + 1,44 + 4 + 0,01 + 0,64) = 3,06,$$

$s_0 = \sqrt{s_0^2} = \sqrt{3,06} \approx 1,75$ м.

Задачи для самостоятельного решения:

15.1. Оценить математическое ожидание и дисперсию случайной величины X по результатам ее независимых наблюдений: 9, 3, 7, 4, 3, 8, 7, 3.

15.2. Найти выборочное среднее по заданному распределению выборки объемом $n = 20$:

x_i	2560	2600	2620	2650	2700
n_i	2	3	10	4	1

15.3. Найти выборочную дисперсию по данному распределению выборки объемом $n = 10$:

x_i	186	192	194
n_i	2	5	3

По данным выборкам найти выборочные средние и средние квадратические отклонения.

15.4.

x_i	1250	1275	1280	1300
n_i	20	25	50	5

15.5.

x_i	12,5	17,5	22,5	27,5	32,5
n_i / n	0,1	0,2	0,4	0,2	0,1

15.6.

Интервал	(28;30)	(30;32)	(32;34)	(34;36)
n_i	8	15	15	12
Интервал	(36;38)	(38;40)	(40;42)	(42;44)
n_i	15	20	10	5

15.7. Данные 30 независимых наблюдений случайной величины X представлены в сгруппированном виде:

Границы интервалов	4-8	8-12	12-16	16-20
Число наблюдений	5	8	14	3

Оценить математическое ожидание и дисперсию этой случайной величины.

15.8. При сверлении отверстий одним и тем же сверлом и последующем измерении диаметров отверстий получены данные, представленные в виде интервального статистического ряда

[40,25;40,28)	[40,28;40,31)	[40,31;40,34)	[40,34;40,37)
2	10	18	25
[40,37;40,40)	[40,40;40,43)	[40,43;40,46)	
12	8	5	$\sum_{i=1}^{n} n_i = 80.$

По данным выборкам найти выборочное среднее и среднее квадратическое отклонение.

Ответы:

15.1. $M(X) \approx 5,5, D(X) \approx 7,43;$ 15.2. 2621. 15.3. 8,93.

15.4. $\bar{x} = 1273,75, s = 13,054.$ 15.5. $\bar{x} = 22,5, s = 5,45.$

15.6. $\bar{x} = 35,72, s = 4,012.$ 15.7. $M(X) \approx 12, D(X) \approx 12,97;$

15.8. $\bar{x} = 40,355, s = 0,04.$

Рассмотрены оценки неизвестных параметров распределения по выборке для частного случая, когда оцениваемый параметр θ является математическим ожиданием или дисперсией распределения, то есть первым начальным и вторым центральным моментами распределения. Приведем теперь общие методы нахождения точечных оценок.

15.5. Метод моментов

Пусть известен вид распределения ГС F(x, θ_1, θ_2,..., θ_k), зависящего от k параметров. Идея метода моментов заключается в следующем: по выборке вычисляют k выборочных моментов и приравнивают их к соответствующим моментам распределения ГС. Искомые оценки каждого из k параметров находятся как решение полученной системы k уравнений. При этом оценки начальных и центральных моментов k-го порядка вычисляются по формулам

$$\tilde{\alpha}_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k, \qquad \tilde{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^k \tag{15.10}$$

Замечание. Оценки математического ожидания и дисперсии, рассмотренные выше, получены по методу моментов.

Пример 15.5. СВ X распределена по показательному закону с плотностью вероятностей $f(x,\lambda) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0. \end{cases}$

Требуется по результатам наблюдений x_1, x_2,…, x_n оценить параметр λ.

Решение. Ранее мы находили M(X)=α_1=1/λ. Оценка математического ожидания есть выборочное среднее $\tilde{\alpha}_1 = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. Приравнивая теоретический момент α_1 эмпирическому $\tilde{\alpha}_1$, получаем $1/\lambda = \bar{x}_1$. Отсюда $\lambda = 1/\bar{x} = n/\sum_{i=1}^{n} x_i$.

Пример 15.6. Случайная величина X имеет пуассоновский закон распределения: $P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$, $k = 0, 1, 2, \ldots$. Найти оценку параметра λ по методу моментов.

Решение. Так как $M(X)=\sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}=\lambda$, то оценкой параметра по методу моментов будет величина $\tilde{\lambda}=\overline{x}$.

Метод моментов обладает тем недостатком, что оценки, полученные по нему, вообще говоря, не являются асимптотически эффективными и могут быть смещенными.

15.6. Метод максимального правдоподобия

Метод максимального правдоподобия наиболее распространен при нахождении точечных оценок параметров. Будем рассматривать результаты выборки как реализацию n-мерной случайной величины $(X_1, X_2, ... X_n)$ с независимыми компонентами. Для получения оценки неизвестного параметра θ естественно попытаться найти такое значение $\tilde{\theta}$, при котором вероятность реализации этой выборки x_1, x_2,..., x_n была бы максимальной.

Если СВ X дискретна, то закон распределения ее имеет вид $\mathrm{P}(\mathrm{X}=x_i)=p_i(\theta)$, $i=1,2,...,n$. Тогда вероятность при n независимых наблюдениях СВ X получить выборку (x_1, x_2,..., x_n) равна

$$\mathrm{L}(\theta)=\mathrm{L}(x_1,x_2,...,x_n,\theta)=\mathrm{P}(\mathrm{X}_1=x_1,\mathrm{X}_2=x_2,...,\mathrm{X}_n=x_n)=$$

$$=\mathrm{P}(\mathrm{X}_1=x_1)\mathrm{P}(\mathrm{X}_2=x_2),\mathrm{P}(\mathrm{X}_n=x_n)=p_1(\theta)\cdot p_2(\theta)\cdot...\cdot p_n(\theta).$$

Функция L(θ) называется функцией *правдоподобия*, а величина $\tilde{\theta}$, являющаяся точкой максимума этой функции, есть оценка параметра θ, полученная по методу максимального правдоподобия (сокращенно МП-оценкой).

Если определяется оценка непрерывной СВ X с плотностью распределения $f(x,\theta)$, то функция правдоподобия определяется так:

$$\mathrm{L}(\theta) = \mathrm{L}(x_1, x_2, ..., x_n, \theta) = f(x_1, \theta) f(x_2, \theta) f(x_n, \theta). \quad (15.11)$$

Если функция правдоподобия дифференцируема по θ и при любых возможных значениях x_i достигает максимума по θ внутри интервала возможных значений параметра θ, то $\tilde{\theta}$ находят, решая уравнение $\dfrac{\partial \mathrm{L}(x_1, x_2, ..., x_n, \theta)}{\partial \theta} = 0.$

Поскольку точки максимума функции $\mathrm{L}(x_1, x_2, ..., x_n, \theta)$ и $\ln \mathrm{L}(x_1, x_2, ..., x_n, \theta)$ при фиксированных $x_1,..., x_n$ совпадают, то в

некоторых случаях удобно решать уравнение

$$\frac{\partial \ln \mathrm{L}(x_1, x_2, ..., x_n, \theta)}{\partial \theta} = 0.$$

Если требуется оценить не один, а k неизвестных параметров $\theta_1, \theta_2, ..., \theta_k$, то оценки максимального правдоподобия для этих параметров находят, решая систему уравнений

$$\frac{\partial \mathrm{L}(x_1, x_2, ..., x_n, \theta_1, ..., \theta_k)}{\partial \theta_j} = 0, \qquad j = 1, 2, ..., k.$$

Пример 15.7. Пусть СВ X распределена по нормальному закону $\mathrm{N}(m,\sigma^2)$ с неизвестными параметрами m и σ^2. Найти МП-оценку параметров нормального распределения.

Решение. Рассмотрим выборку $x_1, x_2, \ldots, x_n$ как реализацию n-мерного СВ $(\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n)$. Тогда составляющие X_i также распределены по закону $\mathrm{N}(m,\sigma^2)$. Запишем функцию правдоподобия

$$L(x_1, x_2, ..., x_n, m, \sigma^2) = \frac{1}{(\sigma\sqrt{2\pi})^n} \exp\left\{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m)^2\right\}.$$

Здесь удобнее перейти к $\ln L$:

$$\ln L(x_1, x_2, ..., x_n, m, \sigma^2) = -(n/2)\cdot\ln(2\pi) - (n/2)\ln\sigma^2 -$$
$$-(1/2\sigma^2)\sum_{i=1}^{n}(x_i - m)^2. \tag{15.12}$$

Дифференцируя (15.12) по m и σ^2, получаем систему

$$\begin{cases} \dfrac{\partial \ln L}{\partial m} = \dfrac{-1}{\tilde{\sigma}^2}\sum\limits_{i=1}^{n}(x_i - \tilde{m}) = 0; \\ \dfrac{\partial \ln L}{\partial(\sigma^2)} = -\dfrac{n}{2\tilde{\sigma}^2} + \dfrac{1}{2\tilde{\sigma}^4}\sum\limits_{i=1}^{n}(x_i - \tilde{m})^2 = 0. \end{cases}$$

Из первого уравнения находим $\tilde{m} = (1/n)\cdot\sum_{i=1}^{n} x_i = \bar{x}$. Подставив это значение во второе уравнение, получим

$\tilde{\sigma}^2 = 1/n\cdot\sum_{i=1}(x_i - \bar{x})^2$. Заметим, что оценка $\tilde{m}$ совпадает с оценкой, полученной по методу моментов, а оценка $\tilde{\sigma}^2$ не совпадает.

Пример 15.8. Пусть имеется простейший поток событий неизвестной интенсивности λ. Для оценки параметра λ проведено наблюдение потока и зарегистрированы $x_1, x_2, \ldots x_n$ - длительности n последовательных интервалов времени между моментами наступления событий. Найти оценку для λ.

Решение. В простейшем потоке интервалы времени между последовательными моментами наступления событий потока имеют показательный закон распределения

$$F(x)=1-\exp(-\lambda x),\, x\geq 0.$$

Так как плотность вероятности показательного закона распределения равна $f(x,\lambda)=F'(x,\lambda)=\lambda\cdot\exp(-\lambda x)$, то функция правдоподобия (15.7) имеет вид:

$$L=f(x_1,\lambda)\cdot f(x_2,\lambda)\cdot f(x_3,\lambda)\cdot\ldots\cdot f(x_n,\lambda)=$$

$$=\lambda\cdot\exp(-\lambda x_1)\cdot\lambda\cdot\exp(-\lambda x_2)\cdot\ldots\cdot\lambda\cdot\exp(-\lambda x_n)=\lambda^n\exp\left\{-\lambda\sum_{i=1}^{n}x_i\right\}.$$

Тогда $\ln L=n\ln\lambda-\lambda\sum_{i=1}^{n}x_i$ и уравнение правдоподобия

$\dfrac{\partial\ln L}{\partial\lambda}=\dfrac{n}{\lambda}-\sum_{i=1}^{n}x_i=0$ имеет решение $\tilde{\lambda}=\dfrac{n}{\sum_{i=1}^{n}x_i}=\dfrac{1}{\overline{x}}$.

При таком значении $\lambda=\tilde{\lambda}$ функция правдоподобия действительно достигает наибольшего значения, так как

$$\frac{\partial^2\ln L}{\partial\lambda^2}=-\frac{n}{\lambda^2}<0.$$

Пример 15.9. Случайная величина X имеет равномерное распределение на отрезке $[\theta-b;\theta+b]$, где θ и b неизвестны. Пусть $x_1,x_2,\ldots,x_n$ - результаты n независимых наблюдений. Найти оценку параметра θ.

Решение. Функция плотности вероятности величины X имеет вид

$$f(x,b,\theta)=\begin{cases}1/(2b), & x\in(-b,b),\\ 0, & x\notin(-b,b).\end{cases}$$

В этом случае функция правдоподобия $L = 1/(2b)^n$ от θ явно не зависит. Дифференцировать по θ такую функцию нельзя и нет возможности записать уравнение правдоподобия. Однако, легко видеть, что L возрастает при уменьшении b. Все результаты наблюдений лежат в $\theta - b; \theta + b$, поэтому можно записать, что

$$\theta - b \le x^1, x^n \le \theta + b,$$

где x^1 - наименьший, а x^n - наибольший из результатов наблюдений. При минимально возможном b

$$\theta - b = x^1, x^n = \theta + b,$$

откуда $\theta - b + \theta + b = x^1 + x^n$ или $2\theta = x^1 + x^n$. Оценкой наибольшего правдоподобия для параметра θ будет величина

$$\tilde{\theta} = \frac{x^1 + x^n}{2}.$$

Задачи для самостоятельного решения

15.1. Методом моментов найти оценку параметра $\theta = p$, где p- есть вероятность «успеха» в любом из n независимых повторных наблюдений, а случайная величина k- число «успехов».

15.2. Случайная величина X (число появлений события A в n независимых испытаниях) подчинена биномиальному распределению с неизвестным параметром распределения p. Проведено 10 опытов по 5 испытаний в каждом. В результате получено эмпирическое распределение

x_i	0	1	2	3	4	5
n_i	4	2	1	1	1	1

где x_i - число появлений события А в одном опыте; n_i - количество опытов, в которых А появилось x_i раз. Методом моментов найти точечную оценку параметра p биномиального распределения.

15.3. Пусть дана случайная выборка $(x_1, x_2, ..., x_n)$ объема n из ГС Х, имеющей равномерный закон распределения

$$f\ x\ = \begin{cases} 1/(b-a), & x \in (a,b); \\ 0, & x \notin (a,b), \end{cases}$$

с неизвестными параметрами a и b. Найти методом моментов точечные оценки этих параметров.

15.4. Случайная величина Х- ошибка измерения дальности радиодальнометра - подчинена равномерному распределению с неизвестными параметрами a и b. Статистическое распределение СВ Х имеет вид

x_i	3	5	7	9	11	13	15	17	19	21
n_i	21	16	15	26	22	14	21	22	18	25

где x_i - средняя ошибка измерений; n_i - количество измерений, имеющих среднюю ошибку x_i. Методом моментов найти точечные оценки неизвестных параметров a и b равномерного распределения.

15.5. Методом наибольшего правдоподобия найти по выборке $x_1, x_2, ..., x_n$ точечную оценку параметра p геометрического распределения $P(X = x_i) = (1-p)^{x_i - 1} p$, где p- вероятность появления события в отдельном испытании.

15.6. Случайная величина имеет плотность вероятности $f(x) = \dfrac{x^{m-1}\lambda^m}{(m-1)!}e^{-\lambda x}$, где λ и m - параметры. Найдите оценку наибольшего правдоподобия для параметра λ.

15.7. При испытании 10 однотипных датчиков импульсного питания зафиксирована их наработка в часах до первого отказа: $x_1 = 52$, $x_2 = 354$, $x_3 = 600$, $x_4 = 418$, $x_5 = 97$, $x_6 = 452$, $x_7 = 553$, $x_8 = 127$, $x_9 = 211$, $x_{10} = 136$. Из теоретических соображений известно, что время безотказной работы датчика имеет функцию распределения $F(x) = 1 - \exp(-\lambda x)$, $0 \le x, \lambda > 0$ (показательный закон распределения). Найдите на основе опытных данных наиболее правдоподобное значение λ.

15.8. Случайная величина имеет закон распределения Рэлея. Функция плотности вероятности этого закона распределения имеет вид

$$f(x) = \frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right),\ x \ge 0.$$

Найдите оценку наибольшего правдоподобия для σ, если результаты наблюдений дали следующие значения случайной величины: $x_1 = 1{,}5$, $x_2 = 1{,}1$, $x_3 = 1{,}2$.

15.9. Отказ устройства произошел при k-м по счету испытании. Найдите оценку наибольшего правдоподобия для вероятности отказа устройства при одном испытании.

Ответы: 15.1. $\tilde{p} = k/n.$ 15.2. $p = 0{,}32.$

15.3. $\tilde{b} = \bar{x} + \sqrt{3}\tilde{\sigma},\ \tilde{a} = \bar{x} - \sqrt{3}\tilde{\sigma}.$ 15.4. $a = 2{,}24; b = 2{,}38.$

15.5. $p = 1/\bar{x}$. 15.6. $\lambda = mn / \sum_{i=1}^{n} X_i$; 15.7. $\tilde{\lambda} = 1/300$;

15.8. $\tilde{\sigma} = 1{,}4$; 15.9. $1/k$.

ЗАНЯТИЕ 16. ИНТЕРВАЛЬНЫЕ ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ

16.1. Доверительный интервал и доверительная Вероятность

Интервальной называют оценку, которая определяется двумя числами $\tilde{\theta}_1$ и $\tilde{\theta}_2$ - концами интервала, накрывающего оцениваемый параметр θ.

На основании выборки $x_1, x_2,\ldots, x_n$ указывают два значения $\theta_1(x_1, x_2,\ldots, x_n)$ и $\theta_2(x_1, x_2,\ldots, x_n)$, с помощью которых можно сделать статистический вывод о том, что истинное значение параметра θ лежит в интервале (θ_1, θ_2).

Доверительным интервалом для параметра θ называется интервал (θ_1, θ_2), содержащий истинное значение параметра с заданной вероятностью p=1-α. Таким образом,

$$P(\theta_1<\theta<\theta_2)=1-\alpha. \qquad (16.1)$$

Число 1-α называется *доверительной вероятностью* (надежностью), а значение α - *уровнем значимости*.

Часто применяют односторонние доверительные интервалы, границы которых определяются из условия $P(\theta<\theta_2)=1-\alpha$ или $P(\theta_1<\theta)=1-\alpha$.

Эти интервалы называются соответственно *левосторонними* и *правосторонними* доверительными интервалами. Выбор доверительной вероятности определяется в каждом случае конкретными условиями. Обычно используемые значения 1-α

равны 0,90; 0,95; 0,99. На практике часто рассматривают *симметричные доверительные интервалы* длиной 2δ. Соотношение (16.1) в этом случае записывается в виде

$$P\left(\left|\tilde{\theta}-\theta\right|<\alpha\right)=1-\alpha \text{ или } P\left(\tilde{\theta}-\delta<\theta<\tilde{\theta}+\delta\right)=1-\alpha. \qquad (16.2)$$

Длина доверительного интервала играет важную роль: чем меньше длина доверительного интервала, тем точнее оценка. Если же длина доверительного интервала велика, то оценка малопригодна для практики.

Из соотношения (16.1) или (16.2) и из того, что θ_1 и θ_2 являются функциями выборки, следует, что длина доверительного интервала определяется двумя величинами: доверительной вероятностью $1-\alpha$ и объемом n выборки.

Таким образом, величины δ, $(1-\alpha)$, n взаимосвязаны и, задавая определенные значения двум из них, можно определить значение третьей. Процедура нахождения границ доверительного интервала для параметра θ по заданной доверительной вероятности в простейшем случае состоит в следующем.

1. Из ГС с ФР $F(x,\theta)$ извлекается выборка объемом n. По этой выборке методом моментов или методом максимального правдоподобия находится точечная оценка $\tilde{\theta}$ неизвестного параметра θ.

2. Составляется некоторая функция элементов выборки - статистика $Y(\tilde{\theta},\theta)$, связанная с параметром θ, такая, что ее распределение не зависит от θ и других неизвестных параметров.

3. Задается доверительная вероятность $(1-\alpha)$.

4. Зная распределение статистики Y, определяют два числа y_1 и y_2, удовлетворяющих условию $P(y_1<Y<y_2)=1-\alpha$.

5. Границы доверительного интервала для параметра θ определяются из решения относительно θ неравенства

$$y_1 < \mathrm{Y}(\tilde{\theta}, \theta) < y_2.$$

Используя указанную схему, можно получить доверительные интервалы параметров m и σ нормального закона распределения.

16.2. Доверительный интервал для математического ожидания СВ X, распределенной по закону N(m, σ) при известном σ

Пусть СВ X имеет нормальное распределение N(m,σ). Тогда доверительный интервал для параметра m по результатам выборки $x_1, x_2, \ldots, x_n$ объемом n при условии, что дисперсия σ^2 известна, а доверительная вероятность равна 1-α, имеет вид

$$\left(\bar{\mathrm{X}} - \sigma/\sqrt{n}\cdot u_{1-\alpha/2};\ \bar{\mathrm{X}} + \sigma/\sqrt{n}\cdot u_{1-\alpha/2}\right). \qquad (16.3)$$

Здесь $u_{1-\alpha/2,}$ квантиль стандартизованного нормального распределения, определяется как решение уравнения

$$\Phi\left(u_{1-\alpha/2}\right) = 1-\alpha/2 \Rightarrow 2/\sqrt{2\pi}\int_0^{u_{1-\alpha/2}} \mathrm{e}^{-t^2/2}\mathrm{d}t = 1-\alpha.$$

Для $u_{1-\alpha/2}$:

1-α	0,90	0,95	0,99	0,997	0,999
$u_{1-\alpha/2}$	1,64	1,96	2,58	3,00	3,37

Из анализа полученных соотношений можно сделать следующие *выводы*.

1. *Увеличение объема n выборки приводит к уменьшению длины доверительного интервала*.

2. *Увеличение доверительной вероятности (1-α) приводит к увеличению длины доверительного интервала, то есть к уменьшению точности δ.*

3. *Если задать точность δ, то есть предельную погрешность интервальной оценки, по формуле (2) и доверительную вероятность 1-α, то из соотношения* $\delta = \sigma.\sqrt{n} \cdot u_{1-\alpha/2}$ *можно найти минимальный объем выборки, который обеспечивает заданную точность:*

$$n = \sigma^2 / \delta \cdot u_{1-\alpha/2}^2 .$$

16.3. Доверительный интервал для МО СВ X, распределенной по нормальному закону при неизвестном σ

Если генеральная совокупность $\xi \in N(m, \sigma^2)$ и σ неизвестно, то с вероятностью $P = 1 - \alpha$

$$m \in \left(\overline{x} - \frac{s_0}{\sqrt{n}} t_{1-\alpha/2, n-1}, \overline{x} + \frac{s_0}{\sqrt{n}} t_{1-\alpha/2, n-1} \right), \qquad (16.4)$$

$$\frac{(n-1)\tilde{s}^2}{\chi^2_{1-\alpha}(n-1)} \le \sigma^2 \le \frac{(n-1)\tilde{s}^2}{\chi^2_{\alpha}(n-1)},$$

где $s_0 = \dfrac{1}{n-1} \sum (x_i - \overline{x})^2$, $t_{1-\alpha/2, n-1}$ $\chi^2_\alpha(n-1)$ – квантиль распределения Стьюдента (Пирсона) с $(n-1)$ степенью свободы уровня α, n – объем выборки .

Для нахождения квантилей распределения Стьюдента t_p имеется таблица. Приведем ее для двух значений доверительной вероятности.

1-α \ n	5	10	20	30	∞
0,95	2,571	2,228	2,086	2,042	1,960
0,99	4,032	3,169	2,845	2,750	2,576

16.4. Доверительный интервал для σ^2 СВ X, распределенной по нормальному закону

Пусть СВ X имеет нормальное распределение N(m,σ), причем m и σ неизвестны. Тогда доверительный интервал для параметра σ^2 по выборке $x_1x_2,..,x_n$ объемом n с доверительной вероятностью 1-α имеет вид

$$\left(\frac{n-1\ s_0^2}{\chi^2_{1-\alpha/2,n-1}};\quad \frac{n-1\ s_0^2}{\chi^2_{\alpha/2,n-1}}\right). \qquad (16.5)$$

Для квантилей $\chi^2_{p,n-1}$ имеются таблицы.

Замечание. Так как при $n\to\infty$ распределение χ^2_n приближается к нормальному, то при достаточно большом объеме выборки ($n\geq 50$) доверительный интервал можно найти по формуле

$$\frac{s_0}{1+u_{1-\alpha/2}/\sqrt{2n}}<\sigma<\frac{s_0}{1-u_{1-\alpha/2}/\sqrt{2n}},$$

где $u_{1-\alpha/2}$ квантиль стандартизованного нормального распределения, соответствующая доверительной вероятности 1-α.

Пример 16.1. Измерения сопротивления резистора дали следующие результаты (в омах): $x_1=592$, $x_2=595$, $x_3=594$, $x_4=592$, $x_5=593$, $x_6=597$, $x_8=589$, $x_9=590$. Известно, что ошибки измерения имеют нормальный закон распределения. Систематическая ошибка отсутствует. Построить доверительный интервал для истинного сопротивления резистора с надежностью 0,99 в предположении: а) дисперсия ошибки измерения известна и равна 4; б) дисперсия ошибки измерения неизвестна.

Решение. В данной серии из девяти наблюдений

$$\bar{X} = \frac{592+595+\ldots+590}{9} = 593.$$

Если дисперсия ошибки измерения известна, то можно воспользоваться формулой (16.3). Для этого из таблицы функции Лапласа (см. приложение) находим, что $2\Phi(2,58) = 0,99$, т.е. уровню надежности 0,99 соответствует значение $t_\gamma = 2,58$. Тогда по формуле (16.3)

$$593 - 2,58\frac{2}{\sqrt{9}} < M(X) < 593 + 2,58\frac{2}{\sqrt{9}}$$

или $591,28 < M(X) < 594,72$ с вероятностью 0,99.

В случае неизвестной дисперсии ее можно оценить на основе тех же опытных данных:

$$\sigma^2 \approx s^2 = \frac{(592-593)^2 + (595-593)^2 + \ldots + (590-593)^2}{8} = 6,5,$$

$s = \sqrt{6,5} \approx 2,55$.По таблице распределения Стьюдента (см. приложение, табл. 3) для $n-1 = 9-1 = 8$ степеней свободы и заданной вероятности $\gamma = 0,99$ находим $t_\gamma = 3,355$. Тогда по формуле (16.4)

$$593-3,355\frac{2,55}{\sqrt{9}}<M(X)<593+3,355\frac{2,55}{\sqrt{9}}$$

или $590,15<M(X)<595,85$ с вероятностью $0,99$.

Пример 16.2. Приведены сгруппированные данные измерений роста у 50 наугад выбранных студентов:

Рост	166 – 170	170 – 174	174 – 178	178 – 182	182 – 186	186 – 190
Число Студентов	3	7	15	13	11	1

Оценить средний рост и дисперсию роста студентов. Построить доверительный интервал для среднего роста студентов с надежностью 0,9.

Решение. Так как данные сгруппированы, то в качестве представителя каждого интервала можно взять середину этого интервала. Тогда

$$M(X)\approx\bar{X}=\frac{\sum_k \hat{X}_k n_k}{n}=$$

$$=\frac{168\cdot 3+172\cdot 7+176\cdot 15+180\cdot 13+184\cdot 11+188\cdot 1}{50}=178,$$

$$D(X)\approx s^2=\frac{\sum_k\left(\hat{X}_k-\bar{X}\right)^2\cdot n_k}{n-1}=$$

$$=\frac{(168-178)^2\cdot 3+(172-178)^2\cdot 7+\ldots+(188-178)^2\cdot 1}{49}\approx 23,76,$$

Так как $2\Phi(1,65)=0,9$, то по формуле (16.4) имеем

$$178-1,65\sqrt{\frac{23,67}{50}}<M(X)<178+1,65\sqrt{\frac{23,76}{50}}$$

или $176,86<M(X)<179,14$ с вероятностью 0,9.►

Пример 16.3. По результатам девяти измерений емкости конденсатора получена оценка $\bar{X}=20$ мкФ. Среднеквадратическая ошибка измерения известна и равна 0,04 мкФ. Построить доверительный интервал для емкости конденсатора с надежностью 0,95.

Решение. В предположении, что ошибки измерения имеют нормальный закон распределения можно воспользоваться формулой (16.3). Так как $2\Phi(1,96)=0,95$, то

$$20-1,96\frac{0,04}{3}<M(X)<20+1,96\frac{0,04}{3}$$

или $19,974<M(X)<20,026$ с вероятностью 0,95.

вероятности 1-α.

Задачи для самостоятельного решения

16.1. Стрелок 20 раз попал в цель при 100 выстрелах. Построить доверительный интервал для вероятности попадания в цель при одном выстреле для уровня надежности $\gamma=0,90$.

16.2. Для проверки всхожести посеяли 900 семян. Из них проросло 810. Постройте доверительный интервал для доли всхожих семян с надежностью 0,95.

16.3. Для изучения общественного мнения было опрошено наугад 1600 жителей нашего города. Деятельность мэра города одобрили 1200 из них. Постройте с надежностью 0,95 доверительный интервал для доли жителей нашего города, одобряющих деятельность мэра.

16.4. По данным 16 наблюдений нормально распределенной случайной величины найдены ее среднее арифметическое $\overline{x} = 15,6$ и оценка среднего квадратического отклонения $s = 0,06$. Построить доверительный интервал для математического ожидания этой случайной величины при уровне надежности 0,95.

16.5. По результатам 10 измерений некоторой физической величины найдены среднее арифметическое результатов измерений $\overline{x} = 8,4$ и оценка среднего квадратического отклонения $s = 0,06$. Считая, что ошибки измерений имеют нормальный закон распределения, найдите интервальную оценку для измеряемой величины с вероятностью 0,95.

16.6. Из большой партии однотипных транзисторов наугад отобрали и проверили 100 штук. У 36 из них оказался коэффициент усиления меньше стандартного. Постройте 95%-й доверительный интервал для доли транзисторов с недостаточным коэффициентом усиления во всей партии.

Ответы:

16.1.(0,134;0,266); 16.2.(0,88;0,92); 16.3.От 73% до 77%;
16.4.(15,385;15,815);16.5.(8,383;8,437); 16.6.(0,26;0,45).

ЗАНЯТИЕ № 17. ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Статистическими гипотезами называют любые предложения относительно параметров (такие гипотезы называют параметрическими) или вида функции распределения случайной величины.

Наряду с выдвинутой гипотезой обычно рассматривают одну или несколько альтернативных (конкурирующих) гипотез. Если выдвинутая гипотеза будет отвергнута, то ее место занимает альтернативная.

Основную (выдвинутую) гипотезу называют *нулевой* и обозначают через H_0.

Альтернативную гипотезу обозначают через H_α.

Выбор альтернативной гипотезы определяется конкретной формулировкой задачи.

Пример 17.1. В теории надежности для многих классов изделий установлен экспоненциальный закон отказов: если X – время безотказной работы изделия, то

$$P\{X > x\} = e^{-x/T}, \quad x \geq 0,$$

где T – среднее время безотказной работы изделия. Допустим, что техническими условиями предусмотрено обеспечение безотказной работы с надежностью 1-α в течение времени T_0.

$$P\{X > T_0\} \geq 1-\alpha \Leftrightarrow e^{-T_0/T} \geq 1-\alpha \quad \text{или} \quad T \geq \frac{T_0}{\ln\frac{1}{1-\alpha}} = T_1.$$

Тогда

$$H_0 = \{F_\chi(x) \in \{1-e^{-x/T},\ T \geq T_1\}\}.$$

Альтернативная гипотеза

$$H_\alpha = \{F_\chi(x) \in \{1-e^{-x/T},\ T < T_1\}\}.$$

Задача проверки статистической гипотезы H_0 относительно генеральной совокупности ξ ставится так:

найти правило, позволяющее по выборке обоснованно решить вопрос о принятии или отклонении гипотезы H_0.

Это правило называется статистическим критерием или просто критерием К проверки гипотезы.

Для решения этой задачи выбирают критерий проверки К, т.е. некоторую функцию (статистику) от выборки

$$z = z(x_1, \ldots x_n) \qquad (17.1),$$

которая является случайной величиной, так как все x_i есть случайные величины. Предполагается, что для этой функции известны плотности распределения вероятностей $f_1(z/H_0)$ и $f_2(z/H_a)$, где H_a – альтернативная гипотеза. Зададим уровень значимости $\alpha = (0,1;\, 0,05;\, 0,01)$. Эта вероятность такова, что событиями, происходящими с такой вероятностью в данной ситуации можно пренебречь.

Критической областью G называют совокупность значений критерия z, при которой гипотезу H_0 отвергают.

Область G находят из условия

$$P\left(z \in G/H_0\right) = \int_G f_1\left(z/H_0\right) dz = \alpha \qquad (17.2)$$

Отметим, что условием (17.2) область G определяется неоднозначно.

Основной принцип проверки статистической гипотезы состоит в следующем: по выборке и формуле (17.1) считают величину $z = z_{\text{набл.}}$

Если $z_{\text{набл.}} \in G$, то H_0 отвергают в пользу альтернативной гипотезы H_a. Если же $z_{\text{набл.}} \notin G$, то оснований отвергнуть H_0 нет, так как выборочные данные не противоречат гипотезе H_0.

Число $r = P\ \ z \in G/H_a\ \ = \int\limits_G f_2\ \ z/H_a\ \ dz$ называют *мощностью критерия*.

$$\beta = 1 - r = P\ \ z \in G/H_a \tag{17.3}$$

При принятии или отклонении гипотезы H_0 возможны ошибки двоякого рода: 1) ошибка первого рода – H_0 отвергают, а она верна.

Вероятность ошибки 1 рода $P\ \ z \in G/H_0\ \ = \alpha$;

2) ошибка второго рода – H_0 принимают, а она не верна.

Вероятность ошибки второго рода $P\ \ z \in G/H_a\ \ = \beta$.

Из формулы (17.3) видно, что чем больше мощность r, тем меньше ошибка 2 рода. Обычно поступают следующим образом: фиксируют уровень значимости α, т.е. фиксируют приемлемую вероятность ошибки 1 рода, а затем ищут критерий z_α с наибольшей мощностью, то есть с наименьшей ошибкой 2 рода.

Таким образом, проверка параметрической статистической гипотезы может быть разбита на следующие этапы:

1) формулируем гипотезы H_0 и H_a;

2) назначаем уровень значимости α;

3) выбираем статистику z (17.1) для проверки гипотезы H_0;

4) находим плотности распределения $f_1\ \ z/H_0$ и $f_2\ \ z/H_a$;

5) в зависимости от гипотезы H_a находим критическую область G;

6) по выборке вычисляем $z_{\text{набл.}} = z(x_1, x_2, \ldots, x_n)$;

7) принимаем решение: если $z_{\text{набл.}} \notin G$, гипотезу H_0 оставляем. Если $z_{\text{набл.}} \in G$, гипотезу H_0 отклоняем в пользу альтернативной H_a.

Модель 1. Пусть известно, что генеральная совокупность $\xi \in N(m, \sigma^2)$, σ – известно. Требуется по выборке и уровню значимости α проверить нулевую гипотезу $H_0 : m = m_0$.

Решение. Для этой модели статистика $z = \dfrac{\overline{x} - m_0}{\sigma}$, где $\overline{x} = \dfrac{1}{n} \sum x_i$.

Тогда для альтернативной гипотезы

$H_a : \quad m > m_0$ критическая область $G = \quad z > u_{1-\alpha} \quad$,

$H_a : \quad m < m_0 \;\; G = \quad z < u_\alpha \quad$,

$H_a : \quad m \neq m_0 \;\; G = \quad |z| > |u_{\alpha/2}| \quad$.

Для простой альтернативной гипотезы

$H_a : \quad m = m_1, m_1 < m_0$ мощность $r = F\left(u_\alpha + \dfrac{m_0 - m_1}{\sigma} \sqrt{n} \right)$;

Для гипотезы $H_a : \quad m = m_1 \;\; m_1 > m_0$,

$$r = 1 - F\left(u_{1-\alpha} - \frac{a_1 - a_0}{\sigma_x} \sqrt{n} \right).$$

Здесь $F(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} e^{t^2/2} dt$, u_α – квантиль уровня α случайной величины $\eta \in N(0,1)$.

Модель 2. Пусть генеральная совокупность $\xi \in N(m,\sigma^2)$, но оба параметра неизвестны. По выборке найдем точечные оценки

$$\bar{x} = \frac{1}{n}\sum x_i \quad \text{и} \quad s_0^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

неизвестных параметров.

По уровню значимости α проверить нулевую гипотезу H_0 : $m = m_0$:

Решение. Для альтернативных гипотез H_a : $m > m_0$ критическая область $G = \{z > t_{1-\alpha}(n-1)\}$.

H_a : $m < m_0$ $G = \{z < t_\alpha(n-1)\}$.

H_a : $m \neq m_0$ $G = \{|z| > |t_{\alpha/2}(n-1)|\}$,

где $z = \sqrt{n}\,\dfrac{\bar{x} - m_0}{s_0}$, $t_\alpha(n-1)$ – квантиль уровня α распределения Стьюдента с $(n-1)$ степенью свободы.

Для H_a : $m = m_1$, $m_1 < m_0$ мощность

$$r = T_{n-1}\left(t_\alpha + \frac{m_0 - m_1}{\bar{s}}\sqrt{n}\right),$$

где $T_{n-1}(t)$ – функция распределения Стьюдента с $(n-1)$ степенью свободы.

Модель 3. Пусть имеем две независимые выборки $x_1, x_2, \ldots x_n$ и $y_1 \ldots y_m$ объемом n и m из нормальных генеральных совокупностей $\xi \in N(m_1, \sigma_1^2)$ и $\eta \in N(m_2, \sigma_2^2)$. Предположим, что σ_1 и σ_2 известны. Требуется на уровне значимости α проверить гипотезу $H_0 : m_x = m_y$:

Решение. Для этой модели статистика $z = \dfrac{\overline{x} - \overline{y}}{\sqrt{\dfrac{\sigma_1^2}{n} + \dfrac{\sigma_2^2}{m}}}$,

где $\overline{x} = \dfrac{1}{n}\sum x_i$, $\overline{y} = \dfrac{1}{m}\sum y_i$.

Для альтернативных гипотез $H_a: \; m_1 > m_2$ область $G = \; z > u_{1-\alpha}$, $H_a: \; m_1 \neq m_2$ область $G = \; |z| > |u_{\alpha/2}|$, $H_a: \; m_1 < m_2$ $G = \; z < u_\alpha$, где u_α квантиль уровня α случайной величины $\eta \in N(0,1)$.

Пример 17.2. По паспортным данным автомобильного двигателя расход топлива на 100 км пробега составляет 10 л. Ожидается, что после модернизации двигателя расход топлива уменьшится. Для проверки производятся испытания 25 случайно отобранных автомобилей с модернизированным двигателем. По результатам испытаний выборочная средняя расходов топлива на 100 км пробега составила $\overline{x} = 9{,}3$ л. Предполагая, что расход топлива есть нормальная случайная величина с $\sigma = 2,$ проверить гипотезу H_0 утверждающую, что изменение конструкции двигателя не повлияет на расход топлива при уровне значимости $\alpha = 0{,}05$.

Решение. Дано: $\overline{x} = 9{,}3$, $\sigma = 2$, $n = 25$, $\alpha = 0{,}05$, $H_0 : a_0 = 10$, $H_0 : a < 10$.

Вычислим статистику

$$z_{\text{набл.}} = \sqrt{n}\,\frac{\overline{x} - a_0}{\sigma} = \sqrt{25} \cdot \frac{9{,}3 - 10}{2} = -1{,}75.$$

Область $G=\{z<u_{\alpha}\}$, где u_{α} – квантиль уровня α случайной величины $\xi \in N(0,1)$. Имеем $u_{0,05}=-u_{0,95}=-1,65$, т.е. $G= \quad z<-1,65$. Так как $z_{\text{набл.}} \in G$

$-1,75<-1,65$, то гипотезу H_0 отвергаем в пользу альтернативной. То есть из опытных данных следует, что модернизация двигателя привела к уменьшению расхода топлива.

Замечание. Пусть в условиях задачи $H_a: a_1=9$. Вычислим мощность критерия r, вероятность ошибки второго рода $\beta=1-r$, и ответим на вопрос, какой минимальный объем выборки нужно взять, чтобы $\beta \le 0,05$.

Имеем $$r=F\left(u_{\alpha}+\frac{a_0-a_1}{\sigma}\sqrt{n}\right)=F\left(-1,65+\frac{10-9}{2}\sqrt{25}\right)=$$

$F(0,85)=0,802$.

$\beta=1-r=0,198$, где $F(x)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{x} e^{-t^2/2}dt$.

$\beta \le 0,05 \Rightarrow r=1-\beta \ge 0,95$.

Решим уравнение $0,95=F\left(u_{\alpha}+\dfrac{a_0-a_1}{\sigma}\sqrt{n}\right)$,

$$u_{\alpha}+\frac{a_0-a_1}{\sigma}\sqrt{n}=u_{0,95}=1,65, \quad -1,65+\frac{10-9}{2}\sqrt{n}=1,65$$

$\sqrt{n} \ge 6,6, \qquad n \ge 44$.

Пример 17.3. Из продукции двух станков-автоматов, выпускающих однотипные изделия, взяты выборки объемов $n_1=15$ и $n_2=18$. По результатам выборок найдены $\bar{x}_1=32$

мм, $\overline{x}_2 = 35$ мм. Дисперсии генеральных совокупностей известны $\sigma_1^2 = 1,5$, $\sigma_2^2 = 2,1$. В предположении о нормальном законе распределения погрешностей изготовления требуется на уровне значимости $\alpha = 0,05$ проверить гипотезу $H_0 : m_1 = m_2$ при альтернативной гипотезе $H_a : m_1 \neq m_2$.

Решение. У нас $n_1 = 15$, $n_2 = 18$, $\overline{x}_1 = 32$, $\overline{x}_2 = 35$, $\sigma_1^2 = 1,5$, $\sigma_2^2 = 2,1$, $\alpha = 0,05$.

Статистика $u = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$. $u_{\text{набл.}} = \dfrac{32 - 35}{\sqrt{\dfrac{1,5}{15} + \dfrac{2,1}{18}}} = \dfrac{-3}{\sqrt{0,1 + 0,117}} =$

$$= \frac{-3}{0,465} = -6,45.$$

Критическая область G для альтернативной гипотезы $H_a : m_1 \neq m_2$ имеет вид

$$G = \quad |u| > |u_{\alpha/2}| \quad .$$

$u_{\alpha/2} = u_{0,025} = -u_{0,975} = -1,96$, $\qquad G = \quad |u| > 1,96 \quad .$

Т.к. $u_{\text{набл.}} \in G$ $(6,45 > 1,96$, то отклоняем гипотезу H_0 в пользу альтернативной H_a.

Задачи для самостоятельного решения.

17.1. По результатам 230 замеров, установлено, что среднее время изготовления детали $\overline{x} = 48$ сек. Предполагая, что время изготовления есть нормальная случайная величина с $\sigma = 1$ сек., необходимо:

1) проверить на уровне значимости $\alpha = 0,05$ гипотезу H_0: $m = 50$ сек. против альтернативной гипотезы H_a: $m = 45$ сек.

2) вычислить мощность критерия и вероятность ошибки второго рода;

3) проверить на уровне значимости $\alpha = 0{,}05$ гипотезу H_0: $m = 50$ сек. против альтернативной гипотезы H_a: $m \neq 50$ сек.

17.2. По данным 170 рейсов установлено, что в среднем машина затрачивает на поездку до хлебоприемного пункта $\bar{x} = 73$ мин. Допустив, что время поездки есть нормальная случайная величина на уровнях значимости $\alpha = 0{,}05$ и $\alpha_1 = 0{,}1$ проверить гипотезу H_0: $a = 75$ мин.; при альтернативной гипотезе H_a: $a = 72$ мин.

1) если известно, что $\sigma = 1$ мин.;

2) если выборочное среднее квадратичное отклонение $s = 1$ мин.;

3) для условий 1) и 2) вычислить мощность критерия.

17.3. Из продукции двух автоматических линий, обрабатывающих корпуса вентилей одного типоразмера, взяты выборки объемов $n_1 = 150$ и $n_2 = 90$. По результатам выборочных наблюдений найдено $\bar{x}_1 = 182$ мм, $\bar{x}_2 = 185$ мм. Предварительно установлено, что погрешности изготовления есть нормальные случайные величины с дисперсиями $\sigma_1^2 = 9$ мм^2, $\sigma_2^2 = 15$ мм^2. Требуется на уровне значимости $\alpha = 0{,}05$ проверить гипотезу $H_0 : m_1 = m_2$:

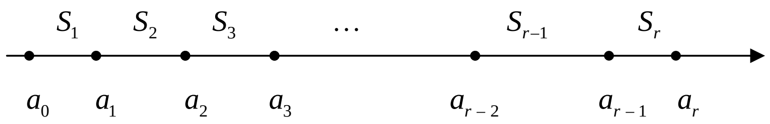
1) при альтернативной гипотезе H_a: $m_1 < m_2$;

2) при альтернативной гипотезе H_a: $m_1 \neq m_2$.

ЗАНЯТИЕ № 18. ПРОВЕРКА ГИПОТЕЗЫ О ВИДЕ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ. КРИТЕРИЙ χ^2

Пусть основная гипотеза H_0 состоит в том, что функция распределения случайной величины ξ есть функция $F(x)$, зависящая от ℓ неизвестных параметров. Наиболее часто применимым критерием проверки этой гипотезы является критерий, введенный К.Пирсоном. Его можно использовать для любых распределений, в том числе и многомерных.

Чтобы воспользоваться этим критерием, выборочные данные предварительно группируют следующим образом. Разбивают множество значений СВ X на *r* непересекающихся множеств S_i с помощью (*r*-1) чисел $a_0 < a_1 < a_2 < \ldots < a_r$:



Обозначим: $p_i = P(a_{i-1} < X < a_i) = F(a_i) - F(a_{i-1})$ - вероятность попадания X в интервал a_{i-1}, a_i в случае, когда предложенная гипотеза справедлива. Очевидно, что $\sum_{i=1}^{r} p_i = 1$. Пусть n_i, $i=1,\ldots,r$ - количество элементов выборки, попавших в интервал a_{i-1}, a_i, $\sum_{i=1}^{r} n_i = 1$. Тогда n_i/n есть относительная частота попадания величины X в интервал $S_i = \ a_{i-1}, a_i$ при n наблюдениях. Очевидно, что $\sum_{i=1}^{r} n_i / n = 1$.

Для приведенного на рисунке разбиения p_i есть приращение гипотетической ФР F(x)на интервале S_i, а n_i/n - приращение эмпирической ФР F*(x) на том же интервале S_i. В качестве статистики принимают следующую величину:

$$\chi^2 = \sum_{i=1}^{r} \frac{n}{p_i} \left(\frac{n_i}{n} - p_i \right)^2 = \sum_{i=1}^{r} \frac{n_i - np \quad ^2}{np_i},$$

являющуюся мерой отклонения эмпирической ФР от теоретической, а критическую область задают в виде $V_k = \ Z \geq Z_\alpha \ $.

Таким образом, процедура применения критерия χ^2 для проверки гипотезы H_0 состоит из следующих этапов:

1. По выборке (2) найдем точечные оценки неизвестных параметров предполагаемого закона распределения $F(x)$.

2. Разобьем числовую ось на r промежутков $\ a_0, a_1 \ : \ a_1, a_2 \ ,$ $\ldots \ a_{r-1}, a_r \ ,$

$a_0 = -\infty$, $a_r = +\infty$, $a_1 < a_2 < \ldots < a_{r-1}$.

Если гипотеза H_0 справедлива, то $i-$му промежутку $\ a_{i-1}, a_i$ соответствует вероятность $p_i = F(a_i) - F(a_{i-1})$, $i = 1 \ldots r$.

3. Пусть из выборки (17.1) n_i значений попадает в $i-$ый промежуток $\ a_{i-1}, a_i$

$\sum n_i = n$. Тогда отношение n_i/n представляет собой частоту попадания выборочных значений в $i-$ый интервал. Близость частот $\frac{n_i}{n}$ к p_i свидетельствует в пользу гипотезы H_0.

4. Вычисляем выборочное значение статистики $\chi^2_{\text{набл.}} = \sum_{i=1}^{r}\frac{(n_i - np_i)^2}{np_i}$, которая характеризует согласованность гипотезы H_0 с опытными данными.

5. Принимаем статистическое решение: гипотеза H_0 не противоречит опытным данным на заданном уровне значимости α, если $\chi^2_{\text{набл.}} < \chi^2_{1\text{-}\alpha}(r-\ell-1)$;

если же $\chi^2_{\text{набл.}} \geq \chi^2_{1\text{-}\alpha}(r-\ell-1)$, то гипотеза H_0 отклоняется. Здесь $\chi^2_{1\text{-}\alpha}$ – квантиль уровня $1-\alpha$ распределения Пирсона с $r-\ell-1$ степеней свободы, ℓ – число параметров распределения $F(x)$, которые оцениваются по выборке .

Замечание. Критерий χ^2 использует тот факт, что случайная величина $\frac{n_i - np_i}{\sqrt{np_i}}$

$i = 1,\ldots,r$, имеет распределение, близкое к нормальному $N(0,1)$. Чтобы это утверждение было достаточно точным, необходимо чтобы для всех интервалов выполнялось условие $np_i \geq 5$. Если в некоторых интервалах это условие не выполняется, то их следует объединить с соседними.

Задачи для самостоятельного решения

18.1. Используя критерий Пирсона, при уровне значимости $\alpha = 0,05$ установить, случайно или значимо расхождение между эмпирическими частотами n_i и теоретическими частотами n_i', которые вычислены исходя из гипотезы о нормальном распределении генеральной совокупности X, если

1.

n_i	5	10	20	8	7
n_i'	6	14	18	7	5

2.

n_i	6	8	13	15	20	16	10	7	5
n_i'	5	9	14	16	18	10	9	6	7

3.

n_i	14	18	32	70	20	36	10
n_i'	10	24	34	80	18	22	12

18.2. При уровне значимости $\alpha = 0,05$ проверить гипотезу о нормальном распределении генеральной совокупности, если известны эмпирические и теоретические частоты

1.

n_i	6	12	16	40	13	8	5
n_i'	4	11	15	43	15	6	6

2.

n_i	5	6	14	32	43	39	30	20	6	5
n_i'	4	7	12	29	48	35	34	18	7	6

3.

n_i	5	13	12	44	8	12	6
n_i'	2	20	12	35	15	10	6

18.3. Используя критерий Пирсона, при уровне значимости $\alpha = 0,05$ проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X с заданным эмпирическим распределением, если

Номер интервала	1	2	3	4	5	6	7
Интервал	(-20;-10)	(-10;0)	(0;10)	(10;20)	(20;30)	(30;40)	(40;50)
Частот n_i	20	47	80	89	40	16	8

Ответы:

18.1.1. $\chi^2_{набл} = 2,47,\ \chi^2_{кр} = 6,0.$ 18.1.2 $\chi^2_{набл} = 1,52,\ \chi^2_{кр} = 12,6.$

18.1.3. $\chi^2_{набл} = 13,93,\ \chi^2_{кр} = 9,5.$ 18.2.1. $\chi^2_{набл} = 2,65,\ \chi^2_{кр} = 9,5.$

18.2.2. $\chi^2_{набл} = 3,\ \chi^2_{кр} = 14{,}1.$ 18.2.3. $\chi^2_{набл} = 13{,}0,\ \chi^2_{кр} = 9{,}5.$

18.3. $\bar{x} = 10{,}4,\ \sigma = 13{,}67,\ k = 4,\ \chi^2_{набл} = 1{,}52,\ \chi^2_{кр} = 9{,}5.$

Дополнение. Распределение χ^2

(Хи-квадрат с *n* степенями свободы)

Пусть $X_1, X_2, \ldots, X_n$- независимые СВ, каждая из которых имеет стандартное нормальное распределение $X_i \in N(0,1)$. χ^2_n-*распределением* называется распределение СВ

$$\chi^2_n = X_1^2 + X_2^2 + \ldots + X_n^2.$$

Плотность распределения χ^2 имеет вид

$$f_n(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}, \qquad x > 0.$$

Математическое ожидание и дисперсия определяются формулами: $M(\chi^2_n) = n, \qquad D(\chi^2_n) = 2n$.

ИНДИВИДУАЛЬНОЕ ДОМАШНЕЕ ЗАДАНИЕ

Хронометраж затрат времени на сборку узла машин у n слесарей дал следующее распределение (мин.)

а) Записать значения результатов эксперимента в виде вариационного ряда. Найти размах варьирования и разбить его на 5 интервалов. Построить гистограмму относительных частот.

б) Найти числовые характеристики выборки $\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i'$ и

$$D_в = \frac{1}{n}\sum_{i=1}^{k} n_i (x_i' - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{k} (x_i')^2 n_i - (\overline{x})^2 ,$$

где x_i' - середины интервалов ($x_i' = \frac{x_i + x_{i+1}}{2}$);

в) Определить доверительные интервалы для неизвестных математического ожидания m_x и среднего квадратического отклонения σ, отвечающие заданной доверительной вероятности $\gamma = 0{,}95$, в предположении, что выборка взята из нормальной генеральной совокупности;

1. Доверительный интервал для математического ожидания в случае нормального распределения

$\overline{x} - t_\gamma \frac{s}{\sqrt{n}} < m_x < \overline{x} + t_\gamma \frac{s}{\sqrt{n}}$, где n- объем выборки, $\overline{x}$ - выборочное среднее, $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{k} n_i (x_i' - \overline{x})^2}$ - исправленное среднее квадратическое отклонение выборки, γ - доверительная вероятность, значение параметра t_γ определяется из таблицы приложений по заданному уровню значимости $\alpha = 1-\gamma$ при числе степеней свободы $k = n-1$.

2. Доверительный интервал для среднего квадратического отклонения σ с заданной надежностью γ $s(1-q) < \sigma < s(1+q)$ при $q<1$ и $0<\sigma<s(1+q)$ при $q>1$, где s - исправленное среднее квадратическое отклонение, параметр q находим из таблицы приложений.

г) Проверить гипотезу о нормальном законе распределения генеральной совокупности по критерию χ^2 при уровне значимости $\alpha = 0,05$.

Вычислить наблюдаемое значение критерия Пирсона

$$\chi^2_{набл} = \sum_{i=1}^{k} \frac{(n_i - n_i')^2}{n_i'},$$

где $n_i' = nP_i$, $P_i = \Phi(z_{i+1}) - \Phi(z_i)$, $z_i = \dfrac{x_i - \overline{x_в}}{\sigma_в}, z_{i+1} = \dfrac{x_{i+1} - \overline{x_в}}{\sigma_в}$.

$\Phi(x)$ - функция Лапласа, значения в таблице приложений . Для первого интервала левый конец устремляем в $-\infty$, для последнего интервала правый конец стремится к ∞. По таблице (приложений) критических точек распределения $\chi^2_{кр}$, уровню значимости $\alpha = 0,05$ и числу степеней свободы $k = l-3$, (l - число интервалов) находим $\chi^2_{кр}$. Если $\chi^2_{набл} < \chi^2_{кр}$, то гипотеза H_0 о нормальном распределении генеральной совокупности принимается, если $\chi^2_{набл} > \chi^2_{кр}$, то гипотеза отвергается.

Таблица 13

№ вар.	Выборка																
1	17	15	23	16	19	19	13	14	11	18	16	18	15	19	16	15	12
	18	19	10	8	19	16	12	12	17	17	11	16	18	22	14	18	10
	17	15	16	20	16	18	14	12	11	3	6	20	14	13	17	9	11
2	16	15	13	12	16	12	16	15	13	15	8	11	10	16	13	20	12
	14	7	5	16	9	13	9	8	14	13	14	15	19	11	14	15	13
	12	13	17	15	11	8	9	3	5	17	11	10	14	6	9	7	8
3	15	7	13	12	7	16	21	17	12	14	10	14	17	10	9	10	1
	1	12	2	15	4	9	15	18	16	9	12	14	11	14	16	16	15
	17	20	7	13	12	10	14	15	8	14	15	13	16	13	17	18	
4	14	13	10	11	4	9	12	9	11	1	5	11	14	5	12	8	10
	7	12	11	12	13	6	5	2	1	8	1	8	10	7	10	12	12
	11	13	16	3	13	8	6	11	5	14	12	8	10	10	12	11	12
5	17	19	25	12	10	21	18	21	15	16	20	13	20	18	21	17	20
	18	17	20	21	18	14	19	14	19	13	18	20	24	16	20	19	17
	22	18	20	18	16	14	13	5	22	18	15	19	14	11	12	10	15
6	14	16	9	22	15	13	7	12	13	10	17	15	17	18	14	15	18
	17	18	15	11	16	11	5	16	10	15	17	21	13	17	16	14	14
	15	17	13	11	10	2	19	18	13	12	16	8	11	13	18	15	19
7	18	11	17	16	24	17	20	14	15	15	12	19	17	19	20	16	20
	16	19	20	11	9	17	13	20	2	18	3	18	12	17	19	23	15
	16	17	17	21	16	19	15	13	12	4	7	15	21	11	13	9	17
8	20	18	16	17	20	15	19	18	17	19	12	15	14	19	16	22	14
	18	17	9	8	17	12	13	18	17	11	16	18	23	15	19	17	15
	16	17	21	11	12	10	18	14	15	20	4	16	18	15	13	12	7
	14	13	4														
9	15	17	24	17	19	14	11	9	20	19	15	16	20	15	19	17	15
	17	11	10	8	19	16	13	12	17	17	10	15	17	21	14	17	18
	9	16	21	16	19	12	15	12	11	4	5	19	14	13	18	10	12
	16	16	11														
10	15	16	11	9	11	10	14	10	11	17	5	17	13	16	11	18	12
	13	19	23	15	19	18	14	17	16	18	14	16	11	9	8	15	12
	15	20	14	18	15	17	8	14	16	7	20	13	16	9	16	9	13

Продолжение табл. 13

11	14	17	6	16	20	14	17	11	16	21	9	16	12	8	1	8	14
	16	10	15	17	16	8	24	17	20	2	18	4	4	7	12	13	8
	7	11	19	13	12	17	19	20	7	17	2	10	12	15	19	28	20
12	23	19	16	9	14	13	8	13	15	9	8	11	10	11	13	6	18
	13	21	9	5	22	11	7	16	19	13	17	20	8	11	19	14	16
	11	14	18	7	8	18	21	8	16	15	19	15	18	10	15	13	17
	15	19	25	5	12	13											
13	20	21	17	19	23	15	12	21	11	16	11	12	21	17	9	15	13
	15	23	7	13	15	12	7	10	16	17	12	14	10	14	17	10	10
	12	2	1	15	4	15	9	16	18	9	12	14	16	16	10	9	17
14	16	13	18	13	12	15	9	14	11	18	16	6	7	5	16	13	10
	14	14	21	13	14	16	17	8	12	2	13	6	16	17	8	11	13
	16	20	9	8	15	9	12	10	1	11	12	16	10	11	15	7	13
	11	14															
15	17	26	18	37	12	34	30	27	12	28	14	17	25	32	23	28	30
	17	21	19	28	22	27	29	30	32	21	26	29	30	29	21	26	29
	23	27	30	18	21	29	24	26	21	15	24	22	18	28	31	18	26
16	9	12	26	24	27	16	20	17	15	23	19	20	17	22	18	18	23
	23	24	27	22	28	20	24	23	21	21	22	26	22	24	20	18	17
	21	16	23	14	29	22	25	19	17	20	21	24	22	24	25	21	21
	25																
17	25	14	14	17	16	24	23	20	19	20	17	18	23	21	9	11	13
	26	25	14	14	17	16	24	23	20	19	20	20	13	23	21	9	11
	23	26	24	22	17	18	22	23	21	16	23	28	18	24	21	19	20
18	7	10	5	1	6	14	13	11	10	11	5	12	10	14	8	13	6
	11	8	12	3	10	11	9	6	4	5	5	6	0	14	13	17	11
	11	9	17	12	11	13	6	4	3	10	2	7	6	18	9	7	12
	13																
19	10	8	2	8	7	10	1	13	9	7	6	4	16	2	1	6	11
	9	17	10	11	13	11	5	4	0	8	12	10	11	13	4	12	5
	10	13	12	9	14	9	10	4	10	0	10	5	4	11	1	9	6
	5	12															

Продолжение табл. 13

20	12	9	15	4	8	6	6	9	13	2	3	6	8	4	1	5	4
	8	1	13	8	11	4	9	7	9	1	7	4	13	7	9	2	7
	10	1	4	8	6	4	2	1	2	11	7	4	7	7	7	1	10
	8																
21	7	13	5	5	9	4	12	6	8	5	0	6	8	0	2	4	2
	4	7	1	7	6	8	3	4	8	9	5	9	2	4	2	9	2
	10	3	7	6	7	3	6	1	10	6	7	9	6	5	9	1	7
	9	2	9														
22	14	17	6	16	20	14	17	11	16	21	9	16	12	8	1	8	14
	16	10	15	17	16	8	24	17	20	2	18	4	4	7	12	13	8
	7	11	19	13	12	17	19	20	7	17	2	10	12	15	19	28	20
23	1	0	7	5	2	0	1	8	5	0	1	0	8	7	8	0	2
	9	0	6	0	8	3	9	1	–5	6	3	5	2	0	7	27	4
	9	4	1	8	6	4	3	2	4	9	2	–2	8	0	7	–1	
24	9	5	7	6	7	6	4	1	4	0	5	6	3	6	–3	8	6
	1	2	1	4	0	5	12	6	1	2	1	11	7	–2	–4	7	5
	0	5	4	7	9	6	7	8	6	3	6	–1	3	4	10	8	0
	6																
25	5	8	4	7	9	2	13	–2	9	6	3	8	1	6	9	8	5
	2	7	2	7	1	6	8	0	8	7	9	4	1	8	7	5	6
	6	-1	2	0	3	6	1	12	4	7	5	10	8	12	1	10	–4
	–4	7															
26	1	8	1	–1	6	7	3	10	6	10	8	7	5	2	9	7	3
	1	9	10	5	9	7	6	4	7	10	7	4	8	8	5	6	2
	7	9	1	3	7	6	10	7	6	8	5	4	–1	3	-3	2	7
	5																
27	5	4	8	6	-1	1	5	3	12	5	5	5	2	2	11	0	7
	4	5	8	8	1	6	1	3	1	8	0	9	3	6	1	3	5
	3	4	6	4	9	5	4	1	0	6	9	3	2	3	–1	7	3
	6																
28	15	16	11	9	11	10	14	10	11	17	5	17	13	16	11	18	12
	13	19	23	15	19	18	14	17	16	18	14	16	11	9	8	15	12
	15	20	14	18	15	17	8	14	16	7	20	13	16	9	16	9	13

Продолжение табл. 13

29	4	9	5	6	8	6	4	5	7	–3	5	6	3	8	4	6	0
	6	1	1	7	–1	0	1	–2	–1	6	1	6	10	5	1	10	2
	4	3	5	13	1	3	5	4	-1	8	7	6	3	2	6	0	8
	4	6															
30	9	1	5	0	5	3	7	2	4	8	2	2	4	1	9	2	5
	7	2	7	1	2	7	5	6	0	9	1	8	1	9	8	3	9
	7	5	6	4	7	2	6	8	7	8	3	1	7	4	3	9	2
	8																

ЗАКЛЮЧЕНИЕ

При составлении настоящего пособия автор стремился, во-первых, привести достаточное число упражнений для выработки навыков решения типовых задач, во-вторых, дать задачи, способствующие уяснению основных понятий и их взаимной связи, в-третьих, дать задачи, дополняющие лекционный курс и содействующие расширению математического кругозора. Ряд приведенных задач и примеров ориентирован на дальнейшее изучение вероятностных дисциплин, таких, как теория случайных процессов, математическая теория надежности и т.д. Такой разнообразный набор задач позволяет использовать пособие на практических занятиях, при составлении домашних заданий и типовых расчетов, а также для организации самостоятельного изучения отдельных разделов теории вероятностей и математической статистики.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Теория вероятностей: Учеб. Для вузов / А.В. Печенкин, О.И. Тескин, Г.М. Цветкова и др.; Под ред. В.С. Зарубина, А.П. Крищенко.-М.: Изд-во МГТУ им. Н.Э. Баумана, 1999.-456 с.

2. Математическая статистика: Учеб. Для вузов / В.Б. Горяинов, И.В. Павлов, Г.М. Цветкова и др.; Под ред. В.С. Зарубина, А.П. Крищенко.-М.: Изд-во МГТУ им. Н.Э. Баумана, 2001.-424 с.

3. Петрушко И.М. Курс высшей математики. Теория вероятностей: Лекции и практические занятия/ И.М. Петрушко, В.И. Афанасьев, А.А, Бободжанов, В.Г. Крупин.-М.: Издательство МЭИ, 2004.-304 с.

4. Дубровская А.П. Курс теории вероятностей и математическая статистика: Учеб. пособие/ А.П. Дубровская, Е.Г. Глушко.- Воронеж: ВГТУ, 2004. 161 с.

5. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике/ В.Е. Гмурман. -М: Высш. шк., 1997.-400 с.

6. Сборник задач по математике для втузов. Ч.3. Теория вероятностей и математическая статистика: Учеб. Пособ./ Под ред. А.В. Ефимова. М.: Наука, 1990. 428 с.

7. Рябушко А.П. Индивидуальные задания по высшей математике: Операционное исчисление. Элементы теории устойчивости. Теория вероятностей. Математическая статистика: учеб. пособие/ А.П. Рябушко.-Мн.: Высш. шк., 2006.-336 с.

ОГЛАВЛЕНИЕ

Учебное издание

Глушко Елена Георгиевна

ТЕОРИЯ ВЕРОЯТНОСТЕЙ.
ПРАКТИЧЕСКИЕ ЗАНЯТИЯ

В авторской редакции

Подписано к изданию 13.12.2017.
Объем данных 4,14 Мб.

ФГБОУ ВО «Воронежский государственный
технический университет»
394026 Воронеж, Московский просп., 14