Белорусский национальный технический университет

Приборостроительный факультет

Кафедра «Инженерная математика»

СОГЛАСОВАНО	СОГЛАСОВАНО
Заведующий кафедрой	Декан факультета
____________ М.А. Князев	___________А.И. Свистун
___________2019 г.	_________2019 г.

ЭЛЕКТРОННЫЙ УЧЕБНО-МЕТОДИЧЕСКИЙ КОМПЛЕКС

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

для студентов технических специальностей
приборостроительного факультета

Составители: М.А. Князев, М.А. Гундина, Н.А. Кондратьева,
Н.К. Прихач, И.В. Прусова, Л.В. Бокуть

Рассмотрено и утверждено
на заседании совета приборостроительного факультета 25. 02. 2019 г.,
протокол № 7

ПЕРЕЧЕНЬ МАТЕРИАЛОВ

Электронный учебно-методический комплекс (ЭУМК) по учебной дисциплине «Математика» под названием «Теория вероятностей и элементы математической статистики» состоит из разделов:

I. Теоретическо-практический:

- учебный материал «Теория вероятностей и элементы математической статистики» представлен параграфами, каждый из которых имеет следующую структуру: теоретические сведения, практическая часть (примеры решения типовых задач), задания для самостоятельной работы и ответы к ним;

II. Контроль знаний:

- проверочные тесты по темам «Теория вероятностей» и «Элементы математической статистики» с таблицами верных ответов;
- контрольные работы по темам «Теория вероятностей» и «Элементы математической статистики» с решениями типовых вариантов;
- перечень вопросов, выносимых на экзамен (для самоконтроля знаний);
- список использованной литературы;

III. Вспомогательный раздел:

- «Приложения» в виде таблиц;
- учебная программа для учреждения высшего образования.

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Цели создания ЭУМК

Целью ЭУМК «Теория вероятностей и элементы математической статистики»» является формирование у студентов комплекса знаний по изучаемой учебной дисциплине «Математика», соответствующих академическим, социально-личностным и профессиональным компетенциям специалиста в рамках образовательного стандарта для технических специальностей приборостроительного факультета.

Особенности структурирования и подачи учебного материала соответствуют следующим теоретическим материалам:

- базовые сведения из разделов «Теория вероятностей» и «Элементы математической статистики», которые дополнены наглядными таблицами. Практическая часть состоит из примеров с решениями, задач для самостоятельного решения с ответами. Раздел контроля знаний содержит проверочные тесты, контрольные работы, вопросы к экзамену. Вспомогательный раздел содержит учебную программу по дисциплине «Математика» раздел «Теория вероятностей и элементы математической статистики».

Рекомендации по организации работы с ЭУМК: Материалы данного ЭУМК могут быть использованы студентами технических специальностей при выполнении практических и расчетно-графических работ, при подготовке к промежуточному контролю знаний.

СОДЕРЖАНИЕ

1. ОБЩИЕ РЕКОМЕНДАЦИИ СТУДЕНТУ ПО ИЗУЧЕНИЮ «ТЕОРИИ ВЕРОЯТНОСТЕЙ И ЭЛЕМЕНТОВ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ»

Данный ЭУМК позволяет организовать учебный процесс таким образом, чтобы сформировать глубокие знания, прочные умения и навыки у студентов.

Неотъемлемой часть учебного процесса является самостоятельная работа.

Использование ЭУМК предусматривает различные виды деятельности: изучение теоретического материала, закрепление полученных умений при решении индивидуальных заданий; проведение самоконтроля и проверки правильности выполненных задач; выполнение контрольных работ.

Цели данного ЭУМК: совершенствование учебно-методического обеспечения учреждений высшего образования, организация и совершенствование самостоятельной работы студентов, внедрение в образовательный процесс современных технологий, обеспечивающих повышение качества образования, формирование информационно-коммуникационной среды взаимодействия между участниками образовательного процесса.

Изучение теоретического материала предусматривает последовательный переход от темы к теме. Для большей наглядности в ЭУМК присутствует графический и табличный материал.

Особое внимание следует обращать на определение основных понятий курса, которые отражают количественную сторону или пространственные свойства реальных объектов и процессов и возникают в результате абстракции из этих свойств и процессов. В ЭУМК приведен набор примеров с разобранным последовательным решением.

В первой части ЭУМК приведен теоретический материал, в котором указаны основные определения, формулировки теорем, формулы, уравнения, которые в дальнейшем необходимы для решения прикладных задач теории вероятностей и математической статистики.

В процессе изучения дисциплины «Математика» студенту необходимо выполнить набор контрольных работ, целью которых является закрепление полученных знаний при решении конкретных типовых задач.

Данный ЭУМК может использоваться при проведении всех видов занятий по учебной дисциплине, для текущего и промежуточного контроля знаний студентов. Теоретический раздел ЭУМК может быть использован для подготовки к лекционным и практическим занятиям. Представленные в ЭУМК задачи и примеры могут применяться для организации индивидуальных расчетно-графических работ, при подготовке к зачету и экзамену по соответствующей дисциплине.

Работа студентов с данным ЭУМК способствует продуктивной учебной деятельности, позволяет сформировать профессиональные компетенции будущих специалистов, обеспечивает развитие познавательных способностей личности, а также обеспечивает рациональное распределение учебного времени по темам учебной дисциплины.

1.1 Методические указания по выполнению контрольных работ

Формирование навыков самостоятельной работы у современного специалиста является одним из важнейших направлений воспитательного процесса. В большинстве учебных планов общих и специальных дисциплин отводятся часы на самостоятельную работу студентов. Опираясь на потребность учащихся в знаниях, возникает возможность развития навыков умственной самостоятельности, формирования самодисциплины, самоорганизации учебного времени.

При подготовке сопутствующего методического руководства акцент был сделан на формирование навыков планирования, анализа и оценки деятельности обучающегося.

Важной формой обучения студента на дневном и заочном отделении является самостоятельная работа над учебным материалом. Она заключается в изучении предмета по учебникам и учебно-методическим пособиям, в ответах на вопросы для самопроверки, в выполнении тестов и контрольных работ. При изучении теоретического материала следует переходить к новому вопросу только после правильного понимания предыдущего, выполняя на бумаге все вычисления и строя все графики и чертежи. Особое внимание следует уделять изучению основных понятий и определений курса.

Рекомендуется вести конспект, в который необходимо вписывать определения, формулировки теорем, формулы, уравнения и т.п. Чтение учебника обязательно должно сопровождаться решением задач. Решение следует излагать достаточно подробно, чтобы его можно было легко восстановить при необходимости, и доводить до конечного ответа. Условия задач и их решения необходимо записывать в отдельную тетрадь.

При изучении курса студент должен выполнить ряд контрольных работ и тестов с целью закрепления материала и проверки его усвоения.

1.2 Правила оформления контрольной работы

При выполнении работ необходимо:

1) указывать на титульном листе номер работы, название дисциплины, номер курса и название факультета, номер зачетной книжки, фамилию, имя и отчество, обратный адрес;

2) решения задач приводить в порядке, указанном в задании;

3) перед каждым решением указывать полный номер задачи (например, 4.2.17 – четвертая работа, задание 2, вариант 17) и ее условие согласно заданию;

4) решения приводятся с необходимыми краткими пояснениями, крупным и разборчивым почерком;

5) после каждого решения оставлять место для возможных замечаний рецензента;

6) не зачтенные работы не оформлять заново (если на необходимость этого не указано рецензентом). Исправленные решения задач приводятся в конце работы.

При несоблюдении указанных требований работа не рецензируется.

Прорецензированные и зачтенные контрольные работы вместе со всеми исправлениями и дополнениями, сделанными по требованию рецензента, следует сохранять. Без предъявления зачтенных контрольных работ студент не допускается к сдаче зачета и экзамена.

2. ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ РАЗДЕЛА «ТЕОРИЯ ВЕРОЯТНОСТЕЙ»

2.1 Введение

Теория вероятностей изучает закономерности, возникающие в случайных экспериментах, раскрывает объективные закономерности, присущие массовым явлениям.

Развитие как науки теории вероятностей берет свое начало с переписки Паскаля и Ферма (1654 г.). Но и до этого многих ученых интересовали задачи, относящиеся к азартным играм, теоретико-вероятностные задачи, имеющие прикладное значение (Кардано, Галилей).

Кроме задач азартных игр появлялся интерес к построению таблиц смертности и вопросам страхования (Граунт, Ван Худде, Ван де Витт).

Факты устойчивости частот случайных событий в задачах обработки демографических данных были известны еще в Древнем Китае и Древнем Риме.

С течением времени объект изучения теории вероятностей менялся. Если вначале основной интерес вызывало исследование вероятностей случайных событий, то уже в XIX в. интерес вызывало исследование случайных величин.

Теория вероятностей тесно связана с прикладными исследованиями различной природы. Она применима как в задачах экономики, производства, так и задачах лингвистики и истории. Сейчас без применения понятия доверительного интервала, корреляции, уровня значимости, нормального закона распределения случайной величины сложно представить обширное исследование в педагогике, физике, механике и других науках.

В основе квантовой механики лежат принципы теории вероятностей. В случае радиоактивного распада нет закона природы, позволяющего определить точное время деления ядра. Существуют только законы, согласно которым можно говорить о вероятности рассада ядра за определенный промежуток времени.

2.2 Элементарная теория вероятностей

Во многих областях человеческой деятельности существуют ситуации, когда определенные явления могут повторяться неограниченное число раз в одинаковых условиях. Подбрасывание монеты, кости, выброс из колоды карт и т.д.

Заметим, что представляется возможным предсказать исход последующего эксперимента по результатам предыдущих, как бы ни было велико число проведенных испытаний.

Во-вторых, относительная частота определенных исходов по мере роста числа испытаний стабилизируется, приближаясь к определенному числу.

Рассмотрим эксперимент по подбрасыванию монеты. Его результат представлен в таблице 1.

Таблица 1. – Результаты эксперимента по подбрасыванию монеты

N / n	10^2	10^4	10^6
1	41	4985	499558
2	48	5004	499995
3	44	5085	500144

N – номер испытания, *n* – количество подбрасываний, в таблице указывается количество выпадений герба.

Наблюдалась стабилизация частот

$$p = \frac{m}{n} = \frac{1}{2}.$$

Обнаруженные закономерности, распространенные на испытания с произвольным числом исходов, позволяют построить простейшую математическую модель случайного эксперимента.

Под *опытом*, или *экспериментом*, или испытанием понимают осуществление конкретного комплекса условий. Опыт называется случайным, если его результат нельзя точно предсказать до его осуществления.

Например, если опыт заключается в подбрасывании монеты, то результат его – выпадение герба (Г) или решетки (Р) – нельзя предсказать заранее. Точно также при стрельбе по мишени нельзя заранее предсказать, будет ли точное попадание в цель или промах.

Построение математической модели эксперимента начинается с описания множества Ω всевозможных исходов, которые могут произойти в результате каждого испытания.

Пространство Ω называют *пространством элементарных исходов*, элемент этого пространства $\omega \in \Omega$ – элементарный исход (элементарное событие).

Событием является любое подмножество $A \subset \Omega$.

Событие называется *достоверным*, если оно обязательно произойдет в условиях данного опыта. Например, выбор одной годной детали из партии *n* годных деталей есть событие достоверное. Так как достоверное событие является совокупностью всех элементарных событий из Ω, то оно совпадает с пространством Ω и также обозначается Ω.

Невозможным называется событие, которое в условиях данного опыта не может произойти. Невозможное событие в пространстве не имеет точек в Ω и обозначается $\varnothing$. Например, невозможно поразить одну и ту же мишень три раза при двух выстрелах.

Если ограничиться рассмотрением пространства элементарных исходов, состоящих из не более, чем счетного числа элементов, то построение вероятностной модели по существу состоит в задании распределения вероятностей на пространстве Ω, в соответствие с которым каждому элементарному исходу $\omega \in \Omega$ ставится в соответствие число $P(\omega)$, называемое *вероятностью* элементарного события ω.

$$0 \le P(\omega) \le 1,$$

$$\sum_{\omega \in \Omega} P(\omega) = 1.$$

Различают *элементарные* и *составные* события. События, которые невозможно разложить на более простые, называются элементарными. Все остальные события называются составными. Например, пусть событие состоит в том, что сумма очков, выпавших при бросании двух игральных костей, равна шести. Это событие состоит из пяти возможных элементарных событий – выпадение на гранях костей следующих пар цифр: (1,5), (2,4), (3,3), (4,2), (5,1) соответственно.

Вероятность любого составного события A:

$$P(A) = \sum_{\omega \in A} P(\omega).$$

Число $P(A)$ интерпретируется как относительная частота появления события A в статистическом эксперименте.

События называются *несовместными*, если появление одного из них исключает появление других событий в условиях одного и того же опыта.

Два или несколько событий называются *равновозможными*, если нет оснований утверждать, что одно из них имеет больше данных появиться в итоге опыта по сравнению с другими. Например, извлечение туза, валета, короля или дамы из колоды карт.

Событие $\overline{A}$, которое обязательно произойдет, если не произойдет событие A, называется противоположным событию A. Например, выигрыш и проигрыш в лотерее – противоположные события.

Если в задаче дана вероятность $P(A)$, тогда чтобы найти вероятность противоположного события, необходимо воспользоваться следующей формулой:

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

где $P(\overline{A})$ – вероятность противоположного события.

Говорят, что несколько событий в условиях данного опыта образуют полную группу событий, если в результате опыта обязательно произойдет хотя бы одно из них. Например, события “извлечение белого шара”, “извлечение красного шара”, “извлечение голубого шара” образуют полную группу событий в опыте извлечения шара из урны, в которой находятся белые, красные и голубые шары.

Примеры

1. Подбрасывается монета и регистрируется сторона монеты, которая обращена к наблюдателю после падения. Найти пространство элементарных исходов.

Решение. Пусть событие Г = {выпал герб}, Р = {выпала решка}.

Тогда $\Omega = \{Г, Р\}$.

2. Бросается игральная кость и регистрируется число выпавших очков. Найти пространство элементарных исходов. Найти событие, состоящее в выпадении четного числа очков.

Решение. $\Omega = \{1,2,3,4,5,6\}$, $A = \{2,4,6\}$.

3. Бросаются две игральные кости. Описать событие, состоящее в том, что сумма очков больше 10.

Решение. $\Omega = \{(1,1),(1,2),...,(6,6)\}$, $A = \{(5,6),(6,5),(6,6)\}$.

Задачи для самостоятельного решения

1. Среди студентов, собравшихся на лекцию по теории вероятностей, наудачу выбирают одного. Пусть событие *A* заключается в том, что выбранный студент окажется юношей; событие *B* – в том, что он не курит, а событие *C* – в том, что он живет в общежитии. Описать событие ABC.

2. Монета подбрасывается три раза подряд. Построить пространство Ω элементарных исходов. Описать событие *A*, состоящее в том, что выпало не менее двух гербов.

3. Пусть *A*, *B* и *C* — произвольные события. Найти выражения для событий, состоящих в том, что из *A*, *B* и *C*:
а) произошло только *A*; б) произошли *A* и *B*, но *C* не произошло; в) все три события произошли.

4. Брошены три монеты. Найти вероятности событий:
а) *A* = {выпало ровно два герба};
б) *B* = {выпало не больше двух гербов}.

5. В лифт 9-этажного дома на первом этаже вошли 5 человек. Известно, что каждый из них с равной вероятностью может выйти на любом из этажей, начиная со второго. Найти вероятность того, что: а) все пятеро выйдут на пятом этаже;
б) все пятеро выйдут одновременно (на одном и том же этаже).

6. Бросаются 4 игральные кости. Найти вероятность того, что на них выпадет по одинаковому числу очков.

Ответы

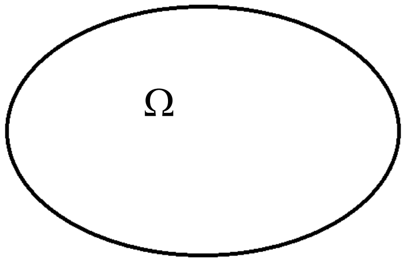
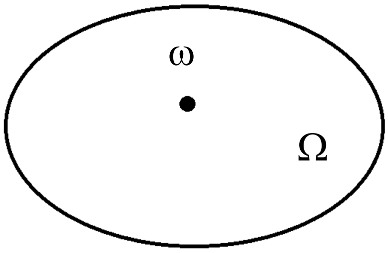
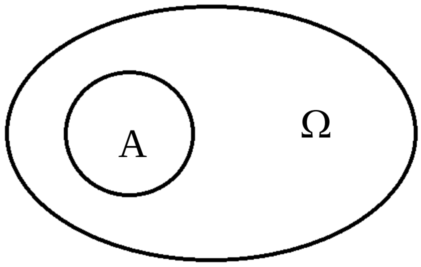
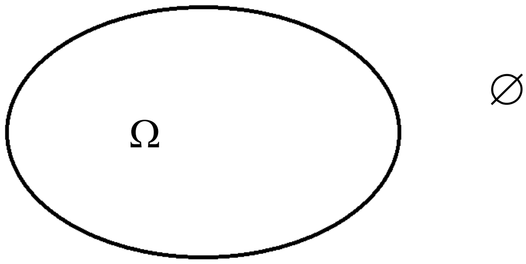
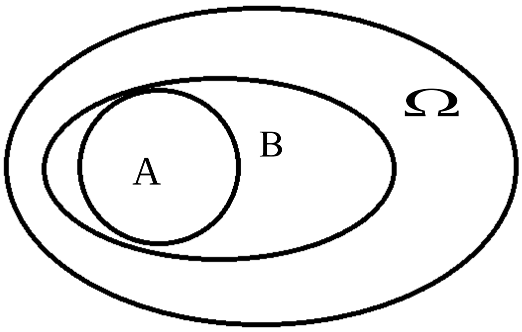
1	2	3	4	5	6
{юноша не курит и проживает в общежитии}	*A*={{ГГР}, {РГГ}, {ГРГ}, {ГГГ}}	$A\overline{B}\overline{C}$, $AB\overline{C}$, ABC	*P*(*A*) = 0,375. *P*(*B*) = 0,875	1/32768, 1/4096	1/216

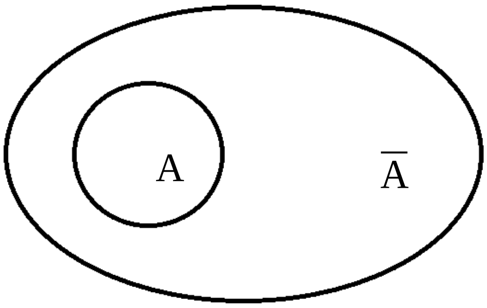
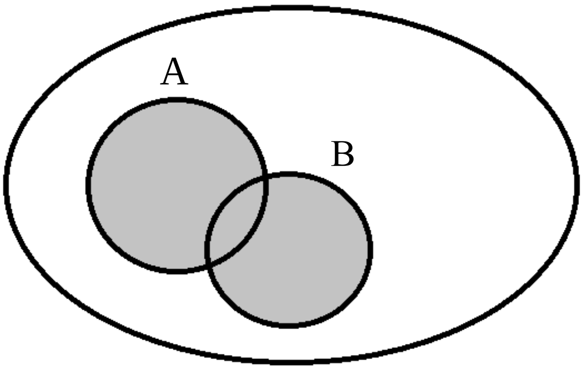
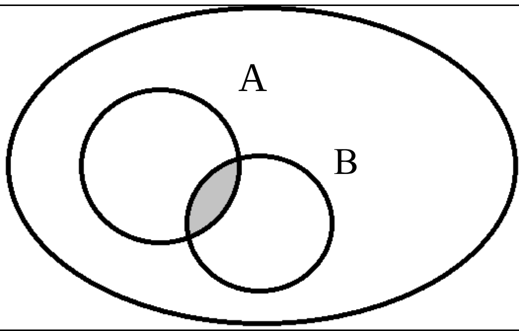
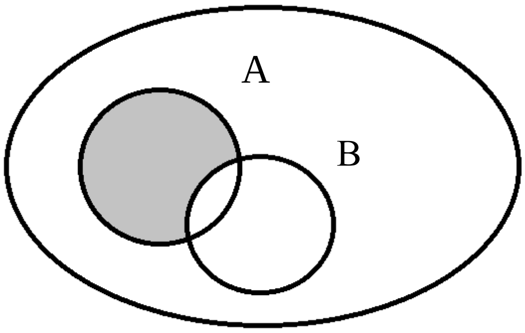
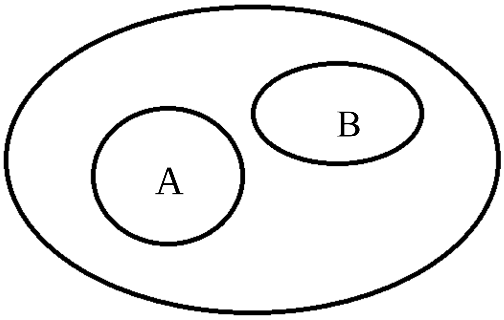
2.3 Вероятностное пространство

Пусть Ω – множество элементарных исходов.

Подмножество пространства Ω называется *событием* $A \subset \Omega$, если статистический эксперимент закончился элементарным исходом $\omega \in A$.

Рассмотрим теоретико-множественные операции в данном пространстве, которые представлены в следующей таблице.

Теоретико-множественные объекты и операции	Вероятностная трактовка	Геометрическая интерпретация
множество Ω	пространство элементарных исходов, достоверное событие	
ω – элемент Ω	элементарный исход эксперимента, элементарное событие	
A – подмножество множества Ω	Событие	
$\varnothing$ – пустое множество	невозможное событие	
$A \subset B$ (подмножество A есть часть B)	из события B следует событие A	

$\overline{A}$ – дополнение к подмножеству *A*	событие *A* не произошло	
$A \cup B$ – объединение подмножеств *A* и *B*	произошло хотя бы одно из событий *A* и *B*	
$A \cap B$ – пересечение подмножеств *A* и *B*	произошли одновременно оба события *A* и *B*	
$A \setminus B$ – разность подмножеств *A* и *B* (за *A* вычитается *B*)	произошло событие *A*, в то время как не произошло *B*	
$A \cap B = \varnothing$ множества *A* и *B* не пересекаются	события *A* и *B* несовместны	

Пусть A и B – обозначают события выпадения при бросании игральной кости соответственно нечетного числа очков и числа очков, кратного трем. Тогда $A = \{1, 3, 5\}$, $B = \{3, 6\}$, и, значит, $A \cup B = \{1, 3, 5, 6\}$.

2.4 Булева алгебра и понятие вероятности

Булевой алгеброй называют такой класс $\tilde{A}$ подмножеств Ω, что:

1) $\Omega \in \tilde{A}$,

2) $A \in \tilde{A} \Rightarrow \bar{A} \in \tilde{A}$,

3) $A, B \in \tilde{A} \Rightarrow A \cup B \in \tilde{A}$.

Вероятностью P на булевой алгебре $\tilde{A}$ подмножеств Ω называется отображение $\tilde{A}$ в отрезок [0, 1], обладающее следующими свойствами:

1) $P(\Omega) = 1$.

2) Если события $A_1, A_2, ..., A_n$ несовместны, то $P(\sum_1^n A_i) = \sum_1^n P(A_i)$.

3) Если $\{A_n, n \geq 1\}$ -- монотонно убывающая последовательность элементов из $\tilde{A}$ и $\bigcap_1^\infty A_n = \varnothing$, то $\lim_{n\to\infty} P(A_n) = 0$. Это может быть записано, как $A_n \downarrow \varnothing$.

Замечание. Вероятность P на $\tilde{A}$ обладает свойствами:

1) $P(\varnothing) = 0$.

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

3) Если $A \subset B$, то $P(A) \leq P(B)$, $P(B \setminus A) = P(B) - P(A)$.

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

5) $P(\bigcup_1^n A_i) \leq \sum_1^n P(A_i)$.

6) Если $A_n \downarrow A$ или $A_n \uparrow A$, то $\lim_{n\to\infty} P(A_n) = P(A)$.

7) $P(\bigcup_1^\infty A_i) \leq \sum_1^\infty P(A_i)$.

Пара $(\Omega, \tilde{A})$, состоящая из пространства элементарных исходов Ω и булевой σ-алгебры $\tilde{A}$ его подмножеств, называется *измеримым пространством*. Только элементы $\tilde{A}$ называются событиями.

Тройка $(\Omega, \tilde{A}, P)$, где P – вероятность на σ- алгебре $\tilde{A}$, называется *вероятностным пространством*.

2.5 Элементы комбинаторики

Комбинаторика – раздел математики, изучающий комбинации конечных множеств элементов различной природы.

Пусть все элементы рассматриваемых множеств различны. Будем изучать комбинации этих элементов, различающихся количеством и/или порядком.

Дано конечное число n объектов произвольной природы, которые назовем элементами.

Из них по определенному правилу можно образовать некоторые группы. Подсчетом числа таких возможных групп и занимается комбинаторика.

Будем рассматривать такие множества, в которых каждый элемент входит не более одного раза (соединения *без повторений*).

Перестановкой из n элементов называется конечное множество элементов, в котором установлен порядок. Так, например, из букв a, b, c можно составить следующие перестановки: (a, b, c), (b, a, c), (a, c, b), (c, a, b), (c, b, a), (b, c, a).

Число возможных перестановок из n элементов равно:

$$P_n = 1 \cdot 2 ... \cdot n = n!.$$

Множество, для которого указан порядок расположения элементов, называется *упорядоченным*. Упорядоченные конечные подмножества некоторого множества называются *размещениями*.

Число всех возможных *размещений*, содержащих по m элементов из множества, содержащего n элементов $(m \le n)$, определяется по формуле:

$$A_n^m = \frac{n!}{(n-m)!}.$$

Всякое конечное подмножество, состоящее из m элементов данного множества из n элементов, называется *сочетанием* m элементов из n, если каждое подмножества из m элементов отличается одно от другого хотя бы одним элементом.

Число всех возможных *сочетаний* обозначается:

$$C_n^m = \frac{n!}{(n-m)! \cdot m!}.$$

Примеры

1. В группе 10 юношей и 7 девушек. Из группы случайным образом отбирается 5 студентов. Найти вероятность того, что среди них окажется 4 девушки?

Решение. Пусть событие A состоит в том, что из 5 случайно отобранных студентов окажутся 4 девушки. Общее число исходов будет равно количеству способов, сколькими из 17 студентов можно отобрать по 5 студентов $n = C_{17}^5$. Благоприятствовать событию A будут те исходы, в которых будет 4 девушки и 1 юноша $m = C_7^4 \cdot C_{10}^1$. Тогда $P(A) = \dfrac{m}{n} = \dfrac{C_7^4 \cdot C_{10}^1}{C_{17}^5} = \dfrac{175}{3094} = 0,057$.

2. Сколько способов существует для выбора команды участников субботника, если известно, что в команде должно быть 5 человек, а в студенческой группе 25 человек?

Решение. Поскольку порядок следования элементов в подгруппе не имеет значения, значит речь идет о количестве сочетаний

$$C_{25}^{5}=\frac{25!}{20!\cdot 5!}=\frac{21\cdot 22\cdot 23\cdot 24\cdot 25}{1\cdot 2\cdot 3\cdot 4\cdot 5}=53130\,.$$

Задачи для самостоятельного решения

1. В соревнованиях участвуют 10 равных по силе шахматистов. Сколько существует вариантов распределения мест между ними?

2. Сколько вариантов бригад по 6 человек в каждой можно составить из 12 человек?

3. Сколько способов существует расставить 5 поездов на 8 запасных путях, если на один путь можно поставить только один поезд?

4. Сколько способов существует для распределения первой, второй и третьей премии на конкурсе, в котором принимает участие 20 человек?

5. Девочка Женя на день Рождения позвала 5 подружек. Сколько существует способов рассадить всех девочек за праздничный стол?

Ответы

1	2	3	4	5
3628800	924	6840	6840	720

2.6 Гипергеометрическое распределение

Большой класс задач, которые интерпретируются в рамках урновой схемы.

Типовая задача: Пусть в эксперименте рассматриваются:

M – черных шаров,

N-M – белых шаров.

Отбирается n шаров из урны. Какова вероятность, что выборка содержит k черных шаров?

Нахождение вероятности в рамках данной схемы осуществляется по следующей формуле: $P(A)=\dfrac{C_M^k C_{N-M}^{n-k}}{C_N^n}$.

Примеры

1. Автомат с 30 мягкими игрушками, содержит фигурки зверей и супергероев в пропорции 2:1 соответственно. В случае победы автомат выдает случайным образом две игрушки. Какова вероятность, что это окажутся супергерои?

Решение. Поскольку в эксперименте есть два ярко выделенных признака, по которым объект можно отнести либо к первому типу (мягкая игрушка), либо ко второму типу (супергерой), речь идет о гипергеометрическом распределении. $M = 1k$ (количество супергероев), N-$M = 2k$ (количество зверей). Тогда общее количество $N = 30 = 3k$, выбирают $n = 2$ игрушек, $k = 2$ (среди тех, которые выбрали, оба оказались супергероями). Тогда по формуле гипергеометрического распределения:

$$P(A) = \frac{C_{10}^{2} \cdot C_{20}^{0}}{C_{30}^{2}} = \frac{\frac{10!}{2!\cdot 8!} \cdot \frac{20!}{0!\cdot 20!}}{\frac{30!}{2!\cdot 28!}} = \frac{45 \cdot 1}{435} = 0{,}103.$$

2. На складе обоев 10 трубок первой партии и 7 трубок второй партии. Продавец случайным образом выбирает 3 трубки, какова вероятность, что все трубки окажутся одной партии?

Решение. По вопросу задачи можно сделать вывод, что исходами, благоприятствующими наступлению события A = { все три трубки окажутся одной партии}, являются следующие: {три трубки первой партии}, {три трубки второй партии}. Тогда вероятность может быть найдена по следующей формуле:

$$P(A) = \frac{C_{10}^{3} \cdot C_{7}^{0}}{C_{17}^{3}} + \frac{C_{10}^{0} \cdot C_{7}^{3}}{C_{17}^{3}} = \frac{\frac{10!}{3!\cdot 7!} \cdot \frac{7!}{0!\cdot 7!}}{\frac{17!}{3!\cdot 14!}} + \frac{\frac{10!}{0!\cdot 10!} \cdot \frac{7!}{3!\cdot 4!}}{\frac{17!}{3!\cdot 14!}} = 0{,}228.$$

Задачи для самостоятельного решения

1. Из сокращенной колоды в 36 карт достают три карты. Какова вероятность, что среди них два туза?

2. В ящике 10 красных, 4 синих и 2 зеленых пуговицы. Какова вероятность того, что среди трех наудачу выбранных пуговиц будет одна синяя, одна зеленая и одна красная?

3. Выигрыш в лотерее «Спорт-лото. 6 из 49». В лотерее рассматриваются 49 видов спорта. Участник называет 6 видов. Выигрыш определяется тем, сколько наименований он угадал из 6 других наименований, которые были заранее выделены комиссией. Какова вероятность того, что участник угадает все 6 наименований?

4. Девочка Даша в класс принесла мешок конфет, в котором 10 конфет «Красная шапочка», 15 «Мишки на севере». Какова вероятность, что 3 случайно выбранные конфеты окажутся с названием «Красная шапочка»?

Ответы

1	**2**	**3**	**4**
0,0269	0,0086	0,00000007	0,052

2.7 Примеры вероятностных пространств

Рассмотрим в таблице примеры вероятностных пространств.

Название пространства	Содержание	Формула для нахождения вероятности	Пример использования
Классическое вероятностное пространство	Число исходов конечно, они равно возможны $\Omega=\{\omega_1,...\omega_n\}$, Ã – множество всех подмножеств Ω, $P(\omega_i)=\frac{1}{n}$.	$P(A)=\frac{m}{n}$, n – число элементарных исходов, m – число элементарных исходов, принадлежащих A.	Однократное подбрасывание монеты. Нахождение вероятности выпадение герба.
Конечное вероятностное пространство	Конечное число исходов, исходы не равно возможны $\Omega=\{\omega_1,...\omega_n\}$, Ã – множество всех подмножеств Ω, $P(\omega_i)=p_i$, $p_i\geq 0,\ \sum p_i=1$.	$P(A)=\sum\limits_{i:\omega_i\in A} p_i$.	Подбрасывается кубик. Вероятность выпадения числа, которое делится на 5.
Дискретное вероятностное пространство	Число элементарных исходов не более чем счетно: $\Omega=\{\omega_1,...\omega_n,...\}$, Ã – множество всех подмножеств Ω, $P(\omega_i)=p_i$, $p_i\geq 0,\ \sum p_i=1$.	$P(A)=\sum\limits_{i:\omega_i\in A} p_i$.	Двое по очереди бросают одну монету. Выигрывает тот, у кого раньше выпадет герб. Какова вероятность, что первый выиграет?
Геометрическое вероятностное пространство	$\Omega\subset R^n$, Ã – некоторая σ алгебра подмножеств, λ – мера, заданная на этом пространстве.	$P(A)=\frac{\lambda(A)}{\lambda(\Omega)}$	Два человека договорились встретиться в течение определенного часа. Пришедший первым человек ждет другого че-

			ловека только 10 минут, после чего уходит. Какова вероятность их встречи?

Задачи для самостоятельного решения

1. Студент пришел на зачет, зная из 30 вопросов только 24. Если студент не знает ответа на поставленный вопрос, преподаватель задает ему еще один, дополнительный. Зачет ставится, если студент правильно отвечает хотя бы на один вопрос. Какова вероятность получения зачета?

2. Два лица А и В условились встретиться в определенном месте между двумя и тремя часами дня. Человек, пришедший первым, ждет другого в течение 15 минут, после чего уходит. Чему равна вероятность встречи этих лиц, если каждый из них может прийти в любое время в течение указанного часа?

3. Найти $P(x+y\le 2)$, где *x, y* – любые числа [0, 4].

4. Найти $P(y\le \frac{1}{x})$, где x,y – любые числа [0.5, 2].

5. Четырехтомное сочинение расположено на полке в произвольном порядке. Какова вероятность, что номера томов идут в порядке возрастания?

Ответы

1	2	3	4	5
0,96	0,4375	0,125	(ln2-ln0.5) / 2.25	1/24

2.8 Разбиение на группы: перестановки, сочетания и размещения с повторениями

Пусть $k_1,k_2,...,k_n$ – целые неотрицательные числа, причем $k_1+k_2+...+k_m=n$. Число способов, которыми можно представить множество Ω из *n* элементов в виде суммы *m* множеств $B_1,B_2,...,B_m$, число элементов которых составляет соответственно $k_1,k_2,...,k_m$ равно:

$$P_n(k_1,k_2,...,k_m)=\frac{n!}{k_1!\cdot k_2!\cdot ...\cdot k_m!}.$$

Сочетаниями из *m* элементов по *n* элементов с повторениями называются группы, содержащие *n* элементов, причем каждый элемент принадлежит одному из *m* типов.

Число различных сочетаний из *m* типов по *n* объектов с повторениями равно:

$$f_m^n=C_{m+n-1}^n.$$

Отображение множества k первых натуральных чисел 1, 2, 3, …, k в данное множество $\{a_1,...,a_n\}$ называется *размещением с повторением*, составленным из данных n элементов (количество типов) по k. Количество размещений с повторениями находится по следующей формуле:

$$m = n^k.$$

Примеры

1. Найдем число различных слов, которые можно получить, переставляя буквы в слове «Математика».

Решение: $P_n(3,2,2,1,1,1) = \dfrac{10!}{3!\cdot 2!\cdot 2!} = 151200$.

2. Найти число способов, которыми можно выбрать три буквы из АААТТТГГГЦЦЦ.

Решение: $f_4^3 = C_{3+4-1}^3 = 20$.

3. Найти количество всевозможных размещений с повторениями из букв *a, b, c* по две буквы.

Решение: $n^k = 3^2 = 9$.

Задачи для самостоятельного решения

1. В урне 6 белых шаров, 11 черных. Одновременно наугад вынимают 2 шара. Найти вероятность, что оба а) белые, б) разных цветов.

2. В цехе работают 12 человек: 5 женщин, 7 мужчин. Какова вероятность, что в бригаде из 7 человек будет 3 женщины?

3. В турнире участвуют семь команд. Сколько вариантов распределения мест между ними возможно?

4. На третьем курсе изучается 12 предметов. Сколько существует способов составления расписания на один день, если в этот день предусмотрено четыре пары по разным предметам?

5. Из цифр 1,2,3,4,5 без повторений составлены всевозможные пятизначные числа. Сколько среди этих чисел таких, которые начинаются цифрой 3?

6. Сколько способов существует выбрать по 3 дежурных из группы в 20 человек?

7. Из 60 вопросов, входящих в экзаменационные билеты, студент подготовил 50. Какова вероятность того, что наудачу взятый студентом билет, содержащий 2 вопроса, будет состоять из подготовленных им вопросов?

8. Один из мальчиков родился в марте, а другой в апреле. Найти вероятность того, что оба они родились в первой неделе месяца (учитывается, что в каждом месяце 5 неполных недель)?

9. Из колоды в 52 карты наугад вынимают 3 карты. Найти вероятность, что среди них окажутся только 2 дамы.

Ответы

1	2	3	4	5
0,1103; 0,4853	0,4419	5040	11880	24
6	**7**	**8**	**9**	
1140	0,692	1 / 25	0,678	

2.9 Независимость. Условные вероятности

Зная распределения вероятностей, мы в состоянии оптимизировать свое поведение при игре, производя ставки на те события из $\tilde{\mathrm{A}}$, которые обладают наибольшей вероятностью.

Дальнейшая оптимизация такой игры обычно осуществляется за счет дополнительной информации, которой может располагать игрок, и учет такой информации осуществляется в терминах так называемой условной вероятности.

Рассмотрим два случайных события *A* и *B*. Пусть известно, что событие *B* наступило, но неизвестно, какое конкретно из элементарных событий ω, составляющих событие *B*, наступило. Что можно сказать в этом случае о вероятности наступления события *A*?

Пусть вероятность события *B* – положительная величина. *Условной вероятностью* события *A* при условии, что произошло событие *B*, называют число:

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

Теорема умножения. Пусть $P(A)>0, P(B)>0$. Тогда $P(A\cap B)=P(A)\cdot P(B|A)=P(B)\cdot P(A|B)$.

Теорема. $P(A_1\cap A_2\cap ...\cap A_{n-1})>0$. Тогда

$$P(A_1\cap A_2\cap ...\cap A_n)=P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1\cap A_2)...P(A_n|A_1\cap ...\cap A_{n-1}).$$

Задача. Студент знает 20 вопросов из 30. Экзаменатор задает три вопроса. Какова вероятность того, что студент ответит на все вопросы?

Два события называются *независимыми,* если вероятность появления одного из них не влияет на вероятность наступления другого. Говорят, что событие *A не зависит* от события *B*, если $P(A|B)=P(A)$, т.к. его вероятность не зависит от того, произошло ли событие В или нет. Независимость двух событий – свойство симметричное.

События *A* и *B* называются *независимыми*, если $P(A\cap B)=P(A)\cdot P(B)$.

Случайные события $A_1, A_2, ..., A_n$ называются *попарно независимыми*, если для любых $i \neq j$, $i, j = 1, 2, ..., n$ $P(A_i \cap A_j) = P(A_i) \cdot P(A_j)$.

Случайные события $A_1, A_2, ..., A_n$ называются *независимыми в совокупности*, если для любого подмножества индексов:

$\{i_1, i_2, ..., i_k\} \subseteq \{1, 2, ..., n\}$ выполняется: $P(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_k}) = \prod_{j=1}^{k} P(A_{i_j})$.

Задача (Пример Бернштейна). На плоскость бросают тетраэдр, три грани которого окрашены соответственно в красный, зеленый и синий цвета, а на четвертой грани есть все цвета. Рассмотреть вероятности событий «выпала грань, которая содержит красный цвет», «выпала грань, которая содержит синий цвет», «выпала грань, которая содержит зеленый цвет». Будут ли эти события попарно независимыми и независимыми в совокупности?

Примеры

1. В тире девушке и юноше выдали по одному патрону для попадания в цель и получения плюшевого медведя. Вероятность того, что попадет в цель девушка, равна 0,01. Вероятность того, что попадет юноша, равна 0,95. Каждый сделал по одному выстрелу. Какова вероятность, что мишка будет выигран?

Решение. Исходы, благоприятствующие наступлению этого события:

{юноша попал и девушка попала},{юноша не попал и девушка попала},{юноша попал и девушка не попала}. $P(A) = 0,9505$.

2. В вазе стоит 5 роз и 4 гвоздики. Случайным образом выбирается один цветок. После этого выбирается еще один. Какова вероятность того, что второй цветок – роза?

Решение. Первым выбранным цветком могла оказаться роза, тогда после ее изъятия в вазе останется только 4 розы. Первой могла оказаться гвоздика, тогда после первого изъятия цветка останется 5 роз. Вероятность того, что второй выбранный цветок роза, вычисляется следующим образом:

$$P(A) = \frac{5}{9} \cdot \frac{4}{8} + \frac{4}{9} \cdot \frac{5}{8}.$$

Задачи для самостоятельного решения

1. Подбрасывается кубик до первого появления шестерки. Какова вероятность, что для этого необходимо подбросить кубик 3 раза?

2. Проверено 25 контрольных работ, среди которых 5 оценены на отлично. Какова вероятность, что три выбранные работы окажутся выполненными на отлично?

3. Вы останавливаете на улице трех человек и спрашиваете наугад у этих трех человек, в какой день недели они родились. Какова вероятность того, что хотя бы двое из них родились в четверг?

4. Из букв резной азбуки составлено слово ВЕРОЯТНОСТЬ. Какова вероятность того, что перемешав буквы и укладывая в ряд по одной (наудачу) мы получили слово ВЕРОЯТНОСТЬ?

5. Узел автомашины состоит из четырех деталей. Вероятности выхода этих деталей из строя соответственно равны: 0,04; 0,07; 0,05; 0,02. Узел выходит из строя, если выходит из строя хотя бы одна деталь. Найти вероятность того, что узел не выйдет из строя, если детали выходят из строя независимо друг от друга.

Ответы

1	**2**	**3**	**4**	**5**
0,116	0,0043	0,055	$1{,}002\cdot 10^{-7}$	0,83

2.10 Формула полной вероятности. Формулы Байеса

Конечное или счетное число случайных событий $A_1, A_2, ..., A_n, ...$ образует полную группу событий (разбиение) если:

1) $P(A_i) > 0, i = 1,2,...;$

2) $A_i \cap A_j = \varnothing, i \neq j$;

3) $\sum_k A_k = \Omega$.

Теорема (Формула полной вероятности). Пусть случайные события $A_1, A_2, ..., A_n, ...$ образует полную группу событий. Тогда для произвольного события *B*, рассматриваемого на том же вероятностном пространстве выполняется следующее:

$$P(B) = \sum_{k=1}^{n} P(B \mid A_k) \cdot P(A_k).$$

Пусть до опыта об исследуемом случайном явлении имеются гипотезы $A_1, A_2, ..., A_n$. После опыта становится известной информация о результатах этого явления, но не полная. Результаты наблюдений показывают, не какой конкретно элементарный исход $\omega \in \Omega$ произошел, а что наступило некоторое событие *B*. Считая, что до опыта были известны (априорные) вероятности $P(A_1), P(A_2), ..., P(A_n)$ и условные вероятности $P(B \mid A_1), P(B \mid A_2), ..., P(B \mid A_n)$, необходимо определить апостериорные вероятности $P(A_1 \mid B), P(A_2 \mid B), ..., P(A_n \mid B)$. Решение поставленной задачи дают формулы Байеса.

Теорема (Формулы Байеса). Пусть случайные события $A_1, A_2, ..., A_n$ образуют полную группу событий. Пусть для произвольного события *B*: $P(B) > 0$.

Тогда для любых значений $k = 1, 2, .., n$ имеют место формулы:

$$P(A_k \mid B) = \frac{P(A_k) \cdot P(B \mid A_k)}{\sum_{k=1}^{n} P(A_k) \cdot P(B \mid A_k)}.$$

Примеры

1. Студент выучил 20 билетов из 25 и идет отвечать вторым. Какова вероятность, что он вытянет «удачный билет»?

Решение. Рассмотрим следующие события:

$A_1 = \{\textit{первый студент выбрал удачный билет}\}$, $A_2 = \overline{A_1}$,

$B = \{\textit{второй студент выбрал удачный билет}\}$.

Тогда $P(B) = P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) = 4/5$.

2. Соотношение грузовых автомобилей, проезжающих по шоссе, на котором стоит бензоколонка, к числу легковых машин, проезжающих по тому же шоссе, равно 2 : 3. Вероятность того, что будет заправляться грузовая автомашина равна 0,1; для легковой машины эта вероятность равна 0,3. К бензоколонке подъехала для заправки автомашина. Найти вероятность того, что это грузовая автомашина.

Решение. Пусть событие *A* – к бензоколонке подъехала для заправки автомашина; H_1 – подъехала грузовая автомашина; H_2 – подъехала легковая автомашина. Тогда

$$P(H_1 / A) = \frac{P(H_1)P(A/H_1)}{P(H_1)P(A/H_1) + P(H_2)P(A/H_2)},$$

$$P(H_1) = \frac{2}{5} \quad P(A/H_1) = 0{,}1 \quad P(H_2) = \frac{3}{5} \quad P(A/H_2) = 0{,}3,$$

$$P(H_1 / A) = \frac{\frac{2}{5} \cdot 0{,}1}{\frac{2}{5} \cdot 0{,}1 + \frac{3}{5} \cdot 0{,}3} = \frac{0{,}04}{0{,}04 + 0{,}18} = \frac{0{,}04}{0{,}22} = 0{,}182.$$

3. При лечении больному необходимо принять лекарства двух видов одинаковой дозировки. Вероятность того, что больному станет легче от первого лекарства равна 0,9; от второго – 0,97. Больному стало легче. Какова вероятность того, что на его состояние повлияло первое лекарство?

Решение. Рассмотрим равновероятные гипотезы A_1={больной принимает первое лекарство}, A_2={больной принимает второе лекарство}.

$$P(A_1) = P(A_2)=0{,}5.$$

Также рассмотрим событие B = {больному стало легче}. Условные вероятности: $P(B|A_1)=0{,}9$. $P(B|A_2)=0{,}97$. Поскольку известно событие, которое наступило, необходимо использовать формулы Байеса. Вероятность того, что на состояние больного повлияло первое лекарство, будет найдена по формуле:

$$P(A_1 | B) = \frac{P(B | A_1) \cdot P(A_1)}{P(B | A_1) \cdot P(A_1) + P(B | A_2) \cdot P(A_2)} = 0{,}48.$$

4. На огороде посажены семена гороха и перца в одинаковых пропорциях. Всхожесть гороха равна 0,06. Всхожесть перца составляет 0,15. Растение проросло, какова вероятность, что это взошел перец?

Решение. Рассмотрим взаимоисключающие гипотезы A_1={посажено семя гороха}, A_2={посажено семя перца}.

$$P(A_1) = P(A_2)=0{,}5.$$

Также рассмотрим событие B = {всхожесть семени}. $P(B|A_1) = 0{,}06$. $P(B|A_2)=0{,}15$. Поскольку известно событие, которое наступило (растение проросло), необходимо использовать формулы Байеса. Вероятность того, что взошел перец, будет найдена по формуле:

$$P(A_1 | B) = \frac{P(B | A_2) \cdot P(A_2)}{P(B | A_1) \cdot P(A_1) + P(B | A_2) \cdot P(A_2)} = 0{,}714.$$

Задачи для самостоятельного решения

1. Изделия некоторого производства удовлетворяют стандарту с вероятностью 0,96. Предлагается упрощенная система испытаний, дающая положительный результат с вероятностью 0,98 для изделий, удовлетворяющих стандарту, а для изделий, которые не удовлетворяют стандарту, с вероятностью 0,05. Какова вероятность того, что система даст положительный результат? Какова вероятность, что изделие удовлетворяет стандарту, если известно, что система дала положительный результат?

2. Предположим, что 5 мужчин из 100 и 25 женщин из 10000 являются дальтониками. Наугад выбранное лицо страдает дальтонизмом. Какова вероятность того, что это мужчина?

3. Лето может оказаться засушливым в 20% случаев, чрезмерно влажным в 30% случаев и нормальным в остальных случаях. Вероятности вызревания урожая составляют 0,7; 0,6; 0,9 соответственно. А) Найти вероятность вызревания урожая в случайно выбранный год; Б) Найти вероятность того, что лето было засушливым, если урожай вызрел.

4. В телевизионном ателье имеется четыре кинескопа. Вероятности того, что кинескоп выдержит гарантийный срок службы, соответственно равны 0,8; 0,85; 0,9; 0,95. Найти вероятность того, что взятый наудачу кинескоп выдержит гарантийный срок.

5. На базе посажены семена укропа и фасоли в одинаковых пропорциях. Всхожесть укропа равна 0,14. Всхожесть фасоли составляет 0,07. Растение проросло, какова вероятность, что это взошел укроп?

Ответы

1	**2**	**3**	**4**	**5**
0,942;0,496	0,952	0,1818	0,875	0,667

2.11 Схема Бернулли

Под *испытанием* следует понимать эксперимент со случайным исходом.

Пусть производятся n независимых испытаний. Известно, что в каждом испытании возможны два исхода: либо происходит событие A (успех), либо событие A не происходит (неудача). Данная схема называется схемой Бернулли. При том предполагается, что вероятность p успеха и q = 1– p неудачи не изменяются при переходе от испытания к испытанию.

$\Omega=\{(\omega_1,...\omega_n),\omega_i=0,1\}$, $\tilde{A}$ – множество всех подмножеств Ω.

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

Задача. Известно, что левши составляют 1% от жителей Земли. Найти вероятность того, что среди 200 человек найдется хотя бы 3 левши.

Наивероятнейшее число появления события A в n независимых испытаниях m – число испытаний, при котором достигается максимальная вероятность в n независимых испытаниях:

$$np-q\le m\le np+p.$$

Примеры

1. Прибор состоит из четырех узлов. Вероятность безотказной работы в течение смены для каждого узла равна 0,85. Узлы выходят из строя независимо друг от друга. Найти вероятность того, что в течение смены откажут ровно два узла.

Решение. Из условия задачи $n=4,\ m=2,\ p=0{,}15,\ q=0{,}85$. Используя формулу Бернулли, получим:

$$P_4(2)=C_4^2p^2q^2=\frac{4\cdot 3}{1\cdot 2}\cdot 0{,}15^2\cdot 0{,}85^2=0{,}098.$$

2. Определить вероятность того, что в семье, имеющей пять детей, будет три девочки и два мальчика. Вероятности рождения мальчика и девочки предполагаются одинаковыми.

Решение. Из условия задачи $n=5,\ m=3,\ p=0{,}5,\ q=0{,}5$. Используя формулу Бернулли, получим:

$$P_5(3)=C_5^3 p^3 q^2=\frac{5\cdot 4\cdot 3}{1\cdot 2\cdot 3}\cdot 0{,}5^3\cdot 0{,}5^2=0{,}3125.$$

3. В условиях предыдущей задачи найти вероятность того, что среди детей будет не больше трех девочек.

Решение.

$$P_5(0)=C_5^0 p^0 q^5=0{,}5^5=0{,}03125;$$
$$P_5(1)=C_5^1 p^1 q^5=5\cdot 0{,}5^1\cdot 0{,}5^4=0{,}15625;$$
$$P_5(2)=C_5^2 p^2 q^3=10\cdot 0{,}5^5=0{,}3125;$$
$$P_5(3)=C_5^3 p^3 q^2=10\cdot 0{,}5^5=0{,}3125;$$
$$P=P_5(0)+P_5(1)+P_5(2)+P_5(3)=0{,}8125.$$

4. Вероятность попадания в цель стрелком равна 0,75. Сделано 20 выстрелов. Определить наивероятнейшее число попаданий в цель.

Решение. Здесь $n=20,\ p=0{,}75,\ q=0{,}25$. Следовательно, применим формулу $np-q\le m\le np+p$. Получим: $20\cdot 0{,}75-0{,}25\le m\le 20\cdot 0{,}75+0{,}75$, т. е. $14{,}75\le m\le 15{,}75$. Наивероятнейшее число попаданий в цель равно 15.

Задачи для самостоятельного решения

1. Пусть вероятность попадания в цель равна 1/5. Производится 10 независимых выстрелов. Какова вероятность по меньшей мере дважды попадания в цель?

2. В цехе шесть моторов. Для каждого мотора вероятность того, что он в данный момент включен, равна 0,8. Найти вероятность того, что в данный момент включено четыре мотора.

3. Вероятность того, что болезнь перейдет в тяжелую форму, равна 0,6. Проводится наблюдение за тремя больными. Какова вероятность того, что у двух из них болезнь перейдет в тяжелую форму.

4. Имеется 20 ящиков одинаковых деталей. Вероятность того, что в одном взятом наудачу ящике детали окажутся стандартными, равна 0,75. Найти наивероятнейшее число ящиков, в которых все детали стандартные.

5. В результате многолетних наблюдений установлено, что вероятность выпадения дождя 1 августа в данном городе равна 0,143. Определить наивероятнейшее число дождливых дней 1 августа в данном городе за 40 лет.

Ответы

1	2	3	4	5
0,624	0,246	0,432	15	5

2.12 Предельные теоремы в схеме Бернулли

Схема независимых испытаний служит вероятностной моделью многих реальных явлений, поэтому представляет значительный интерес задача подсчета вероятности $P_n(m)$. При больших значениях m и n есть трудности в получении численного значения этих вероятностей.

Естественным образом возникает задача нахождения асимптотических форм, позволяющих приближенно вычислять вероятности $P_n(m)$ для достаточно больших n и малых p.

Теорема (Локальная предельная теорема Пуассона). Если $p_n \to 0$, $n \to \infty$, так что $p_n \cdot n \to a$, то

$$P(\mu_n = m) = C_n^m \cdot p_n^m \cdot (1-p_n)^{n-m} \to \frac{a^m}{m!}e^{-a}.$$

Теорема (Интегральная предельная теорема Пуассона). В схеме Бернулли для любого натурального числа n, любого $p \in (0;1)$ и для любого числового множества B справедливо неравенство:

$$\left| P(\mu_n \in B) - \sum_{m \in B} \frac{a^m}{m!} e^{-a} \right| \le \frac{a^2}{n}, \; a = n\,p.$$

Теперь рассмотрим асимптотическую формулу для вероятности не близкой к нулю.

Теорема (Локальная предельная теорема Муавра-Лапласа). Если в схеме Бернулли $\sigma = \sqrt{npq} \to \infty$, то для любого положительного с равномерно по всем x таких:

$$x \in \{x \in R : |x| \le c, x = \frac{m-np}{\sigma}, m \in N \cup \{0\}\}$$

справедливо соотношение:

$$P(\{\frac{\mu_n - np}{\sigma} = x\}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2}}(1+o(1)),$$

где $o(1)$ – бесконечно малая величина при $\sigma \to \infty$.

Теорема (Интегральная предельная теорема Муавра-Лапласа). При выполнении условий предыдущей теоремы равномерно $-\infty \le a < b \le \infty$ выполнено предельное соотношение:

$$P\left(a \le \frac{\mu_n - np}{\sigma} \le b\right) \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx.$$

Заметим, что при использовании интегральной формулы Муавра-Лапласа формула обеспечивает достаточную точность уже при $npq \ge 10$.

По полученным теоремам составим таблицу.

Формула Бернулли	$P_n(m) = C_n^k p^k q^{n-k}$	В каждом из 5 опытов событие А может появится с вероятностью p=0,4. Найти вероятность того, что событие А появится 3 раза.
Локальная теорема Муавра-Лапласа	$P_n(m) \approx \frac{1}{\sqrt{npq}} \varphi(x)$, $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $x = \frac{m - np}{\sqrt{npq}}$	Найти вероятность того, что в 243 испытаниях событие А наступит ровно 70 раз, если вероятность появления этого события p=0,25 в каждом испытании.
Интегральная теорема Муавра-Лапласа	$P_n(m_1 \le m \le m_2) \approx \Phi(x_2) - \Phi(x_1)$ $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$, $x_1 = \frac{m_1 - np}{\sqrt{npq}}$, $x_2 = \frac{m_2 - np}{\sqrt{npq}}$	Фабрика выпускает 70% продукции I сорта. Чему равна вероятность того, что в партии из 1000 изделий число изделий I сорта будет в диапазоне [652, 760].

Примеры

1. В каждом из 5 опытов событие *А* может появится с вероятностью p = 0,4. Найти вероятность того, что событие *А* появится 3 раза.

Решение. Применим формулу Бернулли:

$$P(3) = C_5^3 0{,}4^3 (1 - 0{,}4)^2 = 0{,}23.$$

2. Найти вероятность того, что в 243 испытаниях событие *А* наступит ровно 70 раз, если вероятность появления этого события p = 0,25 в каждом испытании.

Решение. Применим локальную теорему Муавра-Лапласа:

$$x=\frac{m-np}{\sqrt{npq}}=\frac{70-243\cdot\frac{1}{4}}{\sqrt{243\cdot\frac{1}{4}\cdot\frac{3}{4}}}=1{,}37,$$

$$\varphi(1{,}37)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1{,}37^2}{2}}=0{,}156,$$

$$P_{243}(70)=\frac{1}{\sqrt{243\cdot\frac{1}{4}\cdot\frac{3}{4}}}\cdot 0{,}156=0{,}231.$$

3. Фабрика выпускает 70% продукции I сорта. Чему равна вероятность того, что в партии из 1000 изделий число изделий I сорта будет в диапазоне [652, 760]?

Решение. Применим интегральную теорему Муавра-Лапласа.

$$x_1=\frac{m_1-np}{\sqrt{npq}}=\frac{652-1000\cdot 0{,}7}{\sqrt{1000\cdot 0{,}7\cdot 0{,}3}}=-3{,}312,$$

$$x_2=\frac{m_2-np}{\sqrt{npq}}=\frac{760-1000\cdot 0{,}7}{\sqrt{1000\cdot 0{,}7\cdot 0{,}3}}=4{,}14,$$

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt,\ \Phi(x_1)=-0{,}49954,\ \Phi(x_2)=0{,}49998.$$

$$P(652\le m\le 760)=0{,}49998+0{,}49954=0{,}99952.$$

Задачи для самостоятельного решения

1. Цех выпускает 60% шоколадной продукции. Чему равна вероятность того, что в партии из 1000 изделий число изделий, содержащих шоколад, будет не превышать 650 упаковок?

2. Найти вероятность того. Что при 300 подбрасываниях игрального кубика выпадет ровно 29 шестерки.

3. Вероятность выпуска бракованного микроскопа равна 0,17. Найти вероятность того, что в партии из 1000 изделий 160 бракованных.

4. В сводках опубликовано, что утром 31 декабря вероятность осадков составляет 35%. Собраны сведения за 150 лет. Какова вероятность того, что осадки 31 декабря наблюдались 50 раз?

5. По статистике врачи разводятся в 29% случаев. Выбрано для исследования 1000 женатых врачей, какова вероятность того, что разведутся не более 300 исследуемых врачей?

Ответы

1	2	3	4	5
0,998	0,0003	0,0236	0,0623	0,757

2.13 Отклонение относительной частоты от постоянной вероятности в независимых испытаниях

Вероятность того, что в n независимых испытаниях, в каждом из которых вероятность появления события равна p $(0<p<1)$, абсолютная величина отклонения относительной частоты появления события не превысит положительного числа ε, приближенно равна удвоенной функции Лапласа при $x=\varepsilon\sqrt{npq}$.

$$P\left(\left|\frac{m}{n}-p\right|\le\varepsilon\right)\to 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right).$$

Относительная частота события A определяется равенством $P(A)=\frac{m}{n}$, где m – число испытаний, в которых A наступило, n – общее число произвольных испытаний.

Примеры

1. Вероятность появления события в каждом из независимых испытаний равна 0,5. Найти число испытаний n, при котором с вероятностью 0,7698 можно ожидать, что относительная частота появления события отклонится от его вероятности по абсолютной величине не более чем на 0,02.

Решение. Их рассмотренной формулы: $P\left(\left|\frac{m}{n}-p\right|\le\varepsilon\right)\to 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)$ получим, что $p=0{,}5; q=0{,}5; \varepsilon=0{,}02; P=0{,}7698$.

$$2\Phi\left(\varepsilon\sqrt{\frac{n}{0{,}25}}\right)=0{,}7698,\ \varepsilon\sqrt{\frac{n}{0{,}25}})=1.2,\ n=900.$$

2. Вероятность выигрыша на турнире по баскетболу равна 0,58. Найти количество турниров n, при котором с вероятностью приблизительно равной 0,9 можно ожидать, что относительная частота побед отклонится от вероятности по абсолютной величине не более чем на 0,1.

Решение. $p=0{,}58; q=0{,}42; \varepsilon=0{,}1; P=0{,}9$.

$$2\Phi\left(0{,}1\cdot\sqrt{\frac{n}{0{,}58\cdot 0{,}42}}\right)=0{,}9,\ 0{,}1\cdot\sqrt{\frac{n}{0{,}58\cdot 0.42}})=0{,}9,\ n=4.$$

Задачи для самостоятельного решения

1. Найти количество подбрасываний игрального кубика n, при котором с вероятностью приблизительно равной 0,95 можно ожидать, что относительная частота выпадений шестерки отклонится от вероятности по абсолютной величине не более чем на 0,05.

2. Найти количество подбрасываний монеты n, при котором с вероятностью приблизительно равной 0,95 можно ожидать, что относительная частота выпадений герба отклонится от вероятности по абсолютной величине не более чем на 0,1.

3. Найти количество опрошенных мужчин n, при котором с вероятностью приблизительно равной 0,95 можно ожидать, что относительная частота выбранных женатых мужчин отклонится от вероятности по абсолютной величине не более чем на 0,05, если считать, что случайно выбранный мужчина с равной вероятностью может оказаться женатым или не женатым.

4. Найти количество подбрасываний кубика n, при котором с вероятностью приблизительно равной 0,76 можно ожидать, что относительная частота выпадений четного числа очков отклонится от вероятности по абсолютной величине не более чем на 0,05.

5. Танцор Вася обычно выигрывает 1 из 10 турниров по бальным танцам. Найти количество турниров n, в которых участвовал Вася. Если известно, что с вероятностью приблизительно равной 0,9 можно ожидать, что относительная частота его побед отклонится от вероятности по абсолютной величине не более чем на 0,05.

Ответы

1	**2**	**3**	**4**	**5**
26	12	48	138	97

2.14 Случайные величины и их распределения

В азартных играх интерес играющих вызывает не наступление случайного исхода, а связанный с ним выигрыш или проигрыш, т.е. определенная числовая величина, которая соответствует исходу.

Примером случайной величины может быть число очков, выпавших при подбрасывании кубика, число бракованных изделий среди общего числа изделий.

Случайная величина ξ есть число, которое ставится в соответствие каждому возможному исходу эксперимента, т.е. ее можно рассматривать как функцию $\xi(\omega)$ на пространстве элементарных событий Ω.

Пусть $(\Omega, \tilde{A}, P)$ – произвольное вероятностное пространство. Случайной величиной называется функция $\xi : \Omega \to R$, такая что для любого $C \in R$ выполняется следующее: $\{\omega \in \Omega : \xi(\omega) \le c\} \in \tilde{A}$.

Определим функцию распределения случайной величины, которая несет всю информацию, заложенную в случайной величине.

Функцией распределения случайной величины ξ называется функция $F_\xi(x): R \to [0,1]$, такая, что для любого действительного *x* выполняется:

$$\mathrm{F}_\xi(\mathrm{x}) = \mathrm{P}\{\omega \in \Omega : \xi(\omega) < \mathrm{x}\}.$$

Любая функция распределения обладает следующими свойствами:

1) $x_1 < x_2 \Rightarrow F_\xi(x_1) \le F_\xi(x_2)$,

2) существуют пределы $\underset{x \to -\infty}{Lim} F_\xi(x) = 0, \underset{x \to \infty}{Lim} F_\xi(x) = 1$.

3) функция непрерывна слева, т.е. $F_\xi(x_0 - 0) = F_\xi(x_0)$.

4) $\forall x_0 : F_\xi(x_0 + 0) - F_\xi(x_0) = P(\xi = x_0)$.

5) $\forall \xi : P(a \le \xi < b) = F_\xi(b) - F_\xi(a)$.

2.15 Классификация дискретных случайных величин

Дискретная случайная величина – это случайная величина, которая принимает не более чем счетное число значений.

Пусть ее значения $\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_k, ...$ такие, что $P(\xi = x_k) = p_k > 0, k = 1, 2, ...$.

Тогда $\sum_{k=1}^{n(\infty)} p_k = 1$.

Совокупность значений x_k и соответствующих вероятностей p_k *называется распределением дискретной случайной величины*.

Закон распределения такой величины может быть таблично следующим образом:

ξ	x_1	x_2	…	x_n	…
P	p_1	p_2		p_n	

Примеры дискретных случайных величин:

<table>
<tr><td>Случайная величина Бернулли</td><td><table><tr><td>ξ</td><td>0</td><td>1</td></tr><tr><td>P</td><td>p</td><td>1-p</td></tr></table>$0 < p < 1$</td><td>Результат эксперимента по подбрасыванию монеты</td></tr>
<tr><td>Биномиальная случайная величина</td><td><table><tr><td>ξ</td><td>0</td><td>1</td><td>…</td><td>n</td></tr><tr><td>p</td><td>$P_n(0)$</td><td>$P_n(1)$</td><td></td><td>$P_n(\mathrm{n})$</td></tr></table>$P_n(m) = C_n^k p^k q^{n-k}$</td><td>Число успехов в n испытаниях схемы Бернулли</td></tr>
<tr><td>Случайна величина Пуассона</td><td>$P(\xi = k) = \frac{a^k}{k!} e^{-a}$,
$a > 0$,
$a = n \cdot p$,
$n \to \infty$,
$p \to 0$.</td><td>Закон редких событий: число вызовов, поступивших на телефонную станцию; число распавшихся нестабильных частиц</td></tr>
</table>

Геометрическая случайная величина	$P(\xi=k)=p\cdot(1-p)^{k-1}$ $0<p<1$	Производятся независимые испытания, причем в каждом есть только два варианта исхода. Тогда эта случайная величина соответствует числу испытаний до появления успешного исхода
Гипергеометрическая случайная величина	$P(\xi=k)=\dfrac{C_M^k C_{N-M}^{n-k}}{C_N^n}$	Число бракованных изделий
Равномерная случайная величина	<table><tr><td>ξ</td><td>1</td><td>2</td><td>…</td><td>N</td></tr><tr><td>p</td><td>1/n</td><td>1/n</td><td></td><td>1/n</td></tr></table>	Число очков на кубике
Логарифмическая случайная величина	$P(\xi=k)=-\dfrac{p^k}{k\cdot\ln(1-p)}$ $0<p<1$	Распределение по размерам астероидов в Солнечной системе

Закон распределения дискретной случайной величины можно изобразить графически, для чего в прямоугольной системе координат строят точки $M_1(x_1,p_1)$, $M_2(x_2,p_2),\ldots,\ M_n(x_n,p_n)$, где x_i – возможные значения *X*, p_i – соответствующие вероятности; и соединяют их отрезками прямых. Полученную фигуру называют многоугольником распределения (полигоном).

Примеры

1. Найти функцию распределения случайной величины, которая представлена таблицей:

ξ	0	1
P	0,5	0,5

Решение. Запишем функцию распределения в виде сложной функции:

$$F_\xi(x)=\begin{cases}0, & x\in(-\infty,0]\\ 0{,}5 & x\in(0,1]\\ 1 & x\in(1,\infty)\end{cases}.$$

2. Два шахматиста Миша и Коля делают по одному ходу. Вероятность удачного хода Мишей равна 0,7, а для Коли эта вероятность равна 0,76. Найти ряд распределения суммарного числа удачных ходов шахматистами.

Решение.

ξ	**0**	**1**	**2**
P	$0{,}3\cdot 0{,}24$	$0{,}3\cdot 0{,}76+0{,}7\cdot 0{,}24$	$0{,}7\cdot 0{,}76$

3. Партия изделий содержит 10 % нестандартных. Пусть случайная величина X – число стандартных изделий в партии из пяти изделий. Требуется составить закон распределения случайной величины и записать функцию распределения.

Решение. Случайная величина X может принимать значения x_k = 0, 1, 2, 3, 4, 5. Вероятность $P(X = x_k)$ найдем по формуле Бернулли: $P_n(k) = C_n^k p^k q^{n-k}$. По условию задачи $n = 5$; $p = 0{,}9$; $q = 0{,}1$.

$$p_0 = P(X = 0) = C_5^0 p^0 q^5 = 0{,}00001,$$

$$p_1 = P(X = 1) = C_5^1 \cdot pq^4 = 0{,}00045,$$

$$p_2 = P(X = 2) = C_5^2 \cdot p^2 q^3 = 0{,}0081$$

$$p_3 = P(X = 3) = C_5^3 \cdot p^3 q^2 = 0{,}0729,$$

$$p_4 = P(X = 4) = C_5^4 p^4 q = 0{,}32805 \cdot,$$

$$p_5 = P(X = 5) = C_5^5 \cdot p^5 q^0 = 0{,}59049.$$

Запишем закон распределения случайной величины:

x_i	0	1	2	3	4	5
p_i	0,00001	0,00045	0,0081	0,0729	0,32805	0,59049

Найдем функцию распределения. По определению:

$$F(x) = P(X < x) = \sum_{i: x_i < x} p_i .$$

При $x \le 0 \quad F(x) = 0$,

при $0 < x \le 1 \quad F(x) = p_0 = 0{,}00001$,

при $1 < x \le 2 \quad F(x) = p_0 + p_1 = 0{,}00046$,

при $2 < x \le 3 \quad F(x) = p_0 + p_1 + p_2 = 0{,}00856$,

при $3 < x \le 4 \quad F(x) = p_0 + p_1 + p_2 + p_3 = 0{,}081146$,

при $4 < x \le 5 \quad F(x) = p_0 + p_1 + p_2 + p_3 + p_4 = 0{,}40951$,

при $x > 5 \quad F(x) = 1$.

Окончательно получим:

$$F(x) = \begin{cases} 0, & \text{если} \quad x \le 0, \\ 0{,}00001, & \text{если} \quad 0 < x \le 1, \\ 0{,}00046, & \text{если} \quad 1 < x \le 2, \\ 0{,}00856, & \text{если} \quad 2 < x \le 3, \\ 0{,}08146, & \text{если} \quad 3 < x \le 4, \\ 0{,}40951, & \text{если} \quad 4 < x \le 5, \\ 1, & \text{если} \quad x > 5. \end{cases}$$

Задачи для самостоятельного решения

1. Орудие выпускает по цели три снаряда. Событие A – попадание снаряда в цель. $P(A) = 0{,}8$. Найти для случайной величины X (количество попаданий) ряд распределения.

2. Найти функцию распределения для случайной величины X:

X	-1	1
P	½	½

3.Вероятность сдачи экзамена первым студентом равна 0,6, а вторым – 0,9. Составить ряд распределения случайной величины X—числа студентов, успешно сдавших экзамен, если известно, что экзамены пересдавать нельзя.

4. Найти функцию распределения для случайной величины X:

X	-1	0	1
P	3/16	5/16	1/2

5. Два стрелка делают по одному выстрелу в одну мишень. Вероятность попадания для первого стрелка равна 0,5, а для второго 0,8. Найти и построить функцию распределения случайной величины X (число попаданий в мишень).

Ответы

<table>
<tr><th>1</th><th colspan="2">2</th></tr>
<tr><td>$\begin{array}{|c|c|c|c|c|}\hline X & 0 & 1 & 2 & 3 \\ \hline P & 0{,}008 & 0{,}096 & 0{,}384 & 0{,}512 \\ \hline\end{array}$</td><td colspan="2">$F_\xi(x)=\begin{cases}0, & x\in(-\infty,-1] \\ 0{,}5 & x\in(-1,1] \\ 1 & x\in(1,\infty)\end{cases}$</td></tr>
<tr><th>3</th><th>4</th><th>5</th></tr>
<tr><td>$\begin{array}{|c|c|c|c|}\hline X & 0 & 1 & 2 \\ \hline P & 0{,}04 & 0{,}42 & 0{,}54 \\ \hline\end{array}$</td><td>$F_\xi(x)=\begin{cases}0, x\le -1 \\ 3/16, x\in(-1,0] \\ 1/2, x\in(0,1] \\ 1, x>1\end{cases}$</td><td>$F_\xi(x)=\begin{cases}0, x\le 0 \\ 0{,}1; x\in(0,1] \\ 0{,}6; x\in(1,2] \\ 1, x>2\end{cases}$</td></tr>
</table>

2.16 Классификация абсолютно непрерывных случайных величин

Если случайная величина X принимает любые значения из некоторых интервалов или отрезков числовой оси, то она называется непрерывной случайной величиной. Примерами такой величины являются дальность полета снаряда, время безотказной работ прибора.

Плотностью распределения вероятностей случайной величины X в точке x $p(x)$ называется предел:

$$p(x) = \underset{\Delta x\to 0}{Lim}\frac{P(x < X < x+\Delta x)}{\Delta x}.$$

Теорема. Для того, чтобы случайная величина ξ была абсолютно непрерывной, необходимо и достаточно, чтобы:

$$\forall x \in R : F_\xi(x) = \int_{-\infty}^{x} p_\xi(t)dt .$$

Распределение случайной величины ξ называется непрерывным, а сама случайная величина – абсолютно непрерывной случайной величиной, если

$$\forall B \in \tilde{B}(R) : P_\xi(B) = \int_B p_\xi(x)dx ,$$

где $\tilde{B}$ – минимальная σ - алгебра.

Свойства плотности распределения:

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

2) $p_\xi(x) \geq 0$.

3) $F'_\xi(x) = p_\xi(x)$.

Эти три свойства выполняются для любой точки непрерывности функции.

4) $P(a < x < b) = \int_a^b p_\xi(x)dx$.

Примеры абсолютно непрерывных случайных величин.

Нормальная случайная величина (Случайна величина Гаусса)	$p_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{\sigma^2}}$, $P(\alpha \leq X \leq \beta) = \Phi(\frac{\beta - a}{\sigma}) - \Phi(\frac{\alpha - a}{\sigma})$	Ошибка выборки, связь между признаками, скорость роста растений, колебания курса акций
Экспоненциальная случайная величина	$p_\xi(x) = \begin{cases} 0, x < 0 \\ \lambda e^{-\lambda x}, x \geq 0 \end{cases}$, $\lambda > 0$	Распад атомов. Длительность работы оборудования
Равномерная случайная величина	$p_\xi(x) = \begin{cases} 0, x \notin [a,b] \\ \frac{1}{b-a}, x \in [a,b] \end{cases}$	Время ожидания пассажирского транспорта
Случайная величина Коши	$p_\xi(x) = \frac{1}{\pi} \cdot \frac{c}{x^2 + c^2}$, $c > 0, x \in R$	Амплитудно-частотные характеристики линейных колебательных систем
Гамма-	$\alpha, \lambda > 0$,	Время, необхо-

распределенная величина	$p_\xi(x)=\begin{cases}0, x\le 0\\ \dfrac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\cdot e^{-\lambda x}, x>0\end{cases}$ $\Gamma(\alpha)=\int\limits_0^\infty y^{\alpha-1}e^{-y}dy$	димое для появления отказов, при условии, что они независимы и появляются с определенной интенсивностью. Время восстановления сигнала
Бета-распределенная величина	$p_\xi(x)=\begin{cases}0, x\notin(0,1)\\ \dfrac{1}{B(\alpha,\beta)}x^{\alpha-1}\cdot(1-x)^{\beta-1}, x\in(0,1)\end{cases}$ $\alpha,\beta>0$ $B(\alpha)=\dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$	Применяется для описания процессов, обладающих естественными нижними и верхними пределами
Логарифмическая нормальная случайная величина	$\sigma>0,\mu\in R,$ $p_\xi(x)=\begin{cases}0, x\le 0\\ \dfrac{1}{\sigma x\sqrt{2\pi}}e^{-\frac{\ln(x/\mu)^2}{2\sigma^2}}, x>0\end{cases}$	Распределение частот частиц по их размерам при случайном дроблении (градин при граде)
Отраженная нормальная величина	$p_\xi(x)=\begin{cases}0, x<0\\ \dfrac{1}{\sigma}\sqrt{\dfrac{2}{\pi}}e^{-\frac{x^2}{2\sigma^2}}, x\ge 0\end{cases}$	*Упражнение. Найти самостоятельно применение данной случайной величины*
Распределение экстремальных значений	$p_\xi(x)=e^{\frac{\alpha-x}{\beta}\cdot e^{\frac{\alpha-x}{\beta}}}/\beta$	*Упражнение. Найти самостоятельно применение данной случайной величины*
Случайная величина Лапласа	$p_\xi(x)=\dfrac{1}{2b}e^{-\frac{\lvert x-a\rvert}{b}}, b>0, a\in R$	*Упражнение. Найти самостоятельно применение данной случайной величины*

Логистическая случайная величина	$p_\xi(x)=\dfrac{e^{-\frac{x-a}{b}}}{b(1+e^{-\frac{x-a}{b}})^2},b>0,a\in R$	Исследование медико-биологических объектов
Случайная величина Парето	$p_\xi(x)=\begin{cases}0,x<a\\ a^b x^{-b-1}b,x\ge a\end{cases},a,b>0$	Описывает величину дохода, а – минимальный доход. Зависимость абсолютной частоты слов в длинном тексте. Кривая для популярности имен.
Случайна величина Гнеденко	$p_\xi(x)=\begin{cases}0,x\le 0\\ \alpha\lambda x^{\alpha-1}e^{-\lambda x^\alpha},x>0\end{cases}$	Время безотказной работы
Случайная величина Рэлея	$p_\xi(x)=\begin{cases}0,x\le 0\\ \dfrac{1}{\sigma^2}xe^{-\frac{x^2}{2\sigma^2}},x>0\end{cases}$	В теории стрельбы, теории связи. Отклонение от цели изучается. В радиотехнике при исследовании радиосигнала. Плотность распределения излучения абсолютно черного тела по частотам

Примеры

1. Плотность вероятности случайных амплитуд боковой качки корабля подчиняется закону Рэлея с параметрами $\sigma=0{,}1$. Найти вероятность того, что значение случайной амплитуды будет находиться в диапазоне 0,1 до 0,6.

Решение. $P(A)=\int\limits_{0,1}^{0,6}p(x)dx=\int\limits_{0,1}^{0,6}\frac{1}{0{,}1^2}xe^{-\frac{x^2}{2\cdot 0{,}1^2}}dx=0{,}607$**.**

2. Непрерывная случайная величина задана плотностью распределения:

$$p(x)=\begin{cases}0, & при \quad x\le 0,\\ cx^2, & при \quad 0<x\le 2,\\ 0, & при \quad x>2.\end{cases}$$

Требуется найти значение параметра c и записать функцию распределения.

Решение. Значение параметра c определим, используя свойство плотности распределения: $\int\limits_{-\infty}^{\infty} p(x)\,dx=1.$

$$\int\limits_{-\infty}^{\infty} p(x)\,dx=\int\limits_{-\infty}^{0} 0\,dx+\int\limits_{0}^{2} cx^2\,dx+\int\limits_{2}^{\infty} 0\,dx=\frac{cx^3}{3}\Bigg|_0^2=\frac{8}{3}c=1\Rightarrow c=\frac{3}{8}.$$

Функцию распределения определим из условия:

$$F(x)=\int\limits_{-\infty}^{x} p(x)\,dx..$$

Для $x\le 0$ $F(x)=\int\limits_{-\infty}^{x} 0\,dx=0$,

для $0<x\le 2$ $F(x)=\int\limits_{-\infty}^{0} 0\,dx+\int\limits_{0}^{x}\frac{3}{8}x^2\,dx=\frac{x^3}{8}$,

для $x>2$ $F(x)=\int\limits_{-\infty}^{0} 0\,dx+\int\limits_{0}^{2}\frac{3}{8}x^2\,dx+\int\limits_{2}^{x}0\,dx=\frac{x^3}{8}\Bigg|_0^2=1.$

Значит,

$$F(x)=\begin{cases}0, & x\le 0,\\ \dfrac{x^3}{3}, & 0<x\le 2,\\ 1, & x>2.\end{cases}$$

3. Дана функция распределения случайной величины. Найти ее плотность распределения.

$$F(x)=\begin{cases}0, & x\le 0,\\ 1-e^{-4x}, & x>0.\end{cases}$$

Решение. Плотность распределения определим из свойства плотности распределения: $p(x) = F'(x)$.

$$p(x) = \begin{cases} 0, & x \le 0, \\ 4e^{-4x}, & x > 0. \end{cases}$$

2.17 Некоторые законы распределения случайных величин

1. Биномиальный закон $p_m = P(X = m) = C_n^m p^m (1-p)^{n-m}$ Число успехов с вероятностью p в схеме Бернулли проведения n независимых опытов	Наличие синей глины указывает на возможность алмазного месторождения $P(A) = 0{,}4$. Синяя глина обнаружена в трех районах. Построить ряд распределения количества алмазных месторождений.
2. Распределение Пуассона $p_m = P(X = m) = \dfrac{a^m e^{-a}}{m!}$ Число успехов с вероятностью p в схеме Бернулли проведения n>>1, независимых опытов	В ящике находится n = 100 деталей. Вероятность достать бракованное изделие p = 0,01. Мы вынимаем изделие, определяем, бракованное оно или нет, и кладем его обратно. Получилось, что из 100 изделий, которые мы перебрали, два оказались бракованными. Какова вероятность этого?
3. Геометрическое распределение $p_m = P(X = m) = (1-p)^{m-1} p$ Число опытов в схеме Бернулли, проведенных до первого успеха, с вероятностью успеха p в единичном случае	1.Пусть игральная кость вбрасывается до выпадения первой шестерки. Найти вероятность того, что нам понадобится не больше 3-х вбросов. 2.Вероятность поражения цели равна 0,6. Производится стрельба по мишени до первого попадания. Определить вероятность того, что для поражения цели потребуется не более 2 патронов.
4. Гипергеометрическое распределение $p_m = P(X = m) = \dfrac{C_M^m C_{N-M}^{n-m}}{C_N^n}$ В урне N шаров, M – белых, N–M – черных. Вынимается n шаров. Найти вероятность того,	В группе из 21 студента 5 девочек. Наудачу из этой группы отбирается 2 студента. Составить закон распределения случайной величины X – число девушек из отобранных студентов.

что среди извлеченных шаров ровно m белых	
5. Равномерный закон распределения $p(x)=\begin{cases}\frac{1}{b-a}, x\in[a,b]\\ 0, x\langle a, x\rangle b\end{cases}$ Время ожидания пассажирского транспорта, курсирующего с определенным интервалов	Автобус ходит регулярно с интервалом 10 минут. Пассажир приходит на остановку в случайный момент времени. Какова вероятность того, что ждать пассажиру придётся не больше двух минут?
6. Показательный закон распределения $p(x)=\begin{cases}\lambda e^{-\lambda x}, x\geq 0\\ 0, x<0\end{cases}$ $F(x)=\begin{cases}1-e^{-\lambda x}, x\geq 0\\ 0, x<0\end{cases}$ Длительность работы прибора до первого отказа	Время работы радиолампы имеет показательное распределение. Найти вероятность того, что лампа проработает не менее 800 часов, если среднее время работы радиолампы 400 часов.

Примеры

1. Автобусы некоторого маршрута ходят строго по расписанию. Интервал движения 5 мин. Найти вероятность того, что пассажир, подошедший к остановке, будет ожидать очередной автобус менее 3 минут.

Решение. Случайная величина X – время прихода пассажира на остановку, распределена равномерно на [0; 5]. Плотность распределения вероятностей имеет вид:

$$p(x)=\begin{cases}0, & x\notin[0,5],\\ 1/5, & x\in[0,5].\end{cases}$$

Пассажир будет ожидать автобус менее 3 минут, если он подойдет к остановке в интервале времени от 2 до 5 минут после отправления автобуса.

$$P(2\leq x\leq 5)=\int_2^5\frac{1}{5}dx=\frac{x}{5}\bigg|_2^5=1-\frac{2}{5}=\frac{3}{5}.$$

2. Телефонная станция обслуживает 400 абонентов. Вероятность того, что в течение часа абонент позвонит на станцию, равна 0,01 и постоянна для всех абонентов. Найти вероятность того, что на станцию в течение часа позвонят не более двух абонентов.

Решение. Дискретная случайная величина распределена *по закону Пуассона*. Воспользуемся формулой Пуассона:

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

По условию задачи n = 400, p = 0,01, $m \le 2$, λ = 4.

$$P_{400}(m\le 2)=P_{400}(0)+P_{400}(1)+P_{400}(2)=\frac{4^0}{0!}e^{-4}+\frac{4^1}{1!}e^{-4}+\frac{4^2}{2!}e^{-4}=$$

$$=e^{-4}(1+4+8)=\frac{13}{e^4}=\frac{13}{54{,}576}=0{,}238.$$

3. Время T безотказной работы двигателя автомобиля распределено по показательному закону. Известно, что среднее время наработки двигателя на отказ между техническим обслуживанием – 100 ч. Определить вероятность безотказной работы двигателя в течение 80 ч.

Решение. По условию задачи математическое ожидание случайной величины T равно 100 ч. Следовательно, $\frac{1}{\lambda}=100,\ \lambda=10^{-2}$. Тогда плотность распределения времени безотказной работы двигателя имеет вид:

$$p(t)=\begin{cases}0, & t<0,\\ 0{,}01e^{-0{,}01t}, & t\ge 0.\end{cases}$$

Функция распределения случайной величины T принимает вид:

$$F(t)=P(T<t)=\begin{cases}0, & t\le 0,\\ 1-e^{-0{,}01t}, & t>0\end{cases}$$

и определяет вероятность отказа двигателя за время продолжительностью t. Тогда вероятность безотказной работы двигателя за это время будет равна:

$$R(t)=1-P(T<t)=e^{-0{,}01t}.$$

Функцию $R(t)$ называют *функцией надежности*. Для нашего случая

$$P=R(80)=e^{-0{,}01\cdot 80}=e^{-0{,}8}=0{,}45.$$

Задачи для самостоятельного решения

1. Наличие синей глины указывает на возможность алмазного месторождения $P(A)$ = 0.4. Синяя глина обнаружена в трех районах. Построить ряд распределения количества алмазных месторождений.

2. В ящике находится n = 100 деталей. Вероятность достать бракованное изделие p = 0.01. Мы вынимаем изделие, определяем, бракованное оно или нет, и кладем его обратно. Получилось, что из 100 изделий, которые мы перебрали, два оказались бракованными. Какова вероятность этого?

3. а) Пусть игральная кость вбрасывается до выпадения первой шестерки. Найти вероятность того, что нам понадобится не больше 3-х вбросов.

б) Вероятность поражения цели равна 0,6. Производится стрельба по мишени до первого попадания. Определить вероятность того, что для поражения цели потребуется не более 2 патронов.

4. В группе из 21 студента 5 девочек. Наудачу из этой группы отбирается 2 студента. Составить закон распределения случайной величины *X* – число девушек из отобранных студентов.

5. Автобус ходит регулярно с интервалом 10 минут. Пассажир приходит на остановку в случайный момент времени. Какова вероятность того, что ждать пассажиру придётся не больше двух минут?

6. Время работы радиолампы имеет показательное распределение. Найти вероятность того, что лампа проработает не менее 800 часов, если среднее время работы радиолампы 400 часов.

Ответы

<table>
<tr><th colspan="5">1</th><th colspan="2">2</th><th>3</th></tr>
<tr><td>X</td><td>0</td><td>1</td><td>2</td><td>3</td><td colspan="2" rowspan="2">0,185</td><td rowspan="2">0,42;0,84</td></tr>
<tr><td>P</td><td>0,216</td><td>0,432</td><td>0,288</td><td>0,064</td></tr>
<tr><th colspan="4">4</th><th colspan="2">5</th><th colspan="2">6</th></tr>
<tr><td>X</td><td>0</td><td>1</td><td>2</td><td colspan="2" rowspan="2">0,2</td><td colspan="2" rowspan="2">0,137</td></tr>
<tr><td>P</td><td>0,57</td><td>0,38</td><td>0,05</td></tr>
</table>

2.18 Основные числовые характеристики случайных величин

	Дискретная случайная величина	Абсолютно непрерывная случайна величина
Математическое ожидание	$M(x)=\sum_{i=1}^{n} x_i p_i$	$M(x)=\int_{-\infty}^{\infty} xp(x)dx$

Свойства математического ожидания:

1) Математическое ожидание числа появлений события *A* в одном испытании равно вероятности *p* наступления этого испытания.

2) $M(C) = C$, $C = const$.

3) $M(CX) = C\,M(X)$.

4) $M(X + Y) = M(X) + M(Y)$.

5) $M(X\,Y) = M(X)M(Y)$.

6) $M(X\text{-}MX) = 0$.

Заметим, что математическое ожидание случайной величины ξ можно трактовать как вероятностное среднее этой величины.

Для любой случайной величины X случайная величина $X^0 = X - MX$ называется центрированной случайной величиной или отклонением.

Пусть случайная величина ξ определена на вероятностном пространстве $(\Omega, \tilde{A}, P)$. Для $m \in R$ величина $M\xi^m$, если она определена, называется моментом m-го порядка случайной величины ξ.

Величина $M\left|\xi^m\right|$ называется абсолютным моментом *m*-го порядка случайной величины ξ.

Моменты случайной величины $\xi - M\xi$ называются центральными моментами случайной величины ξ.

Центральные моменты четного порядка случайной величины ξ характеризуют степень разброса значений относительно ее среднего значения.

Дисперсией случайной величины ξ называется число $D\xi = M(\xi - M\xi)^2$, число $\sigma_\xi = \sqrt{D\xi}$ называется среднеквадратическим отклонением случайной величины ξ.

	Дискретная случайная величина	Абсолютно непрерывная случайна величина
Дисперсия	$D(X) = \sum_{i=1}^{n}(x_i - MX)^2 p_i$	$D(X) = \int_{-\infty}^{\infty}(x - MX)^2 p(x)dx$

Свойства дисперсии случайной величины:

1) $DX = M(X^2) - (MX)^2$.

2) $D(C) = 0$, $C = const$.

3) $D(CX) = C^2 D(X)$.

4) $D(X + Y) = D(X) + D(Y)$, X,Y – независимые случайные величины.

5) $D(X+C) = D(X)$.

6) $D(XY) = MX^2MY^2\text{-}(MX)^2(MY)^2$.

Формулы вычисления математического ожидания и дисперсии для некоторых случайных величин:

<table>
<tr><td>Случайная величина Бернулли
<table><tr><td>ξ</td><td>0</td><td>1</td></tr><tr><td>P</td><td>p</td><td>1-p</td></tr></table></td><td>$M\xi = p$</td><td>$D\xi = p(1-p)$</td></tr>
<tr><td>Биномиальная случайная величина
$P_n(m) = C_n^k p^k q^{n-k}$</td><td>$M\xi = np$</td><td>$D\xi = np(1-p)$</td></tr>
<tr><td>Геометрическая случайная величина
$P(\xi = k) = p \cdot (1-p)^{k-1}$</td><td>$M\xi = \frac{1-p}{p}$</td><td>$D\xi = \frac{1-p}{p^2}$</td></tr>
<tr><td>Случайная величина Пуассона</td><td>$M\xi = a$</td><td>$D\xi = a$</td></tr>
</table>

$P(\xi=k)=\dfrac{a^k}{k!}e^{-a}$		
Равномерная случайная величина $p_\xi(x)=\begin{cases}0, x\notin[a,b]\\ \dfrac{1}{b-a}, x\in[a,b]\end{cases}$	$M\xi=\dfrac{a+b}{2}$	$D\xi=\dfrac{(a-b)^2}{12}$
Показательная случайная величина $p_\xi(x)=\begin{cases}0, x<0\\ \lambda e^{-\lambda x}, x\ge 0\end{cases}$	$M\xi=1/\lambda$	$D\xi=\dfrac{2}{\lambda^2}$
Нормальная случайная величина $p_\xi(x)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{\sigma^2}}$	$M\xi=a$	$D\xi=\sigma^2$

Ковариацией случайной величины ξ и η называется число:

$$\operatorname{cov}(\xi,\eta)=M((\xi-M\xi)\cdot(\eta-M\eta)).$$

Если математическое ожидание случайной величины X является характеристикой ее положения, средним значением, около которого группируются значения случайной величины, то дисперсия и среднеквадратического отклонение являются характеристиками рассеяния случайной величины около математического ожидания.

Примеры

1. Найти математическое ожидание, дисперсию и среднеквадратическое отклонение случайной величины X

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

Решение.

$$MX=0\cdot 0{,}1+1\cdot 0{,}1+2\cdot 0{,}3+3\cdot 0{,}5=0{,}1+0{,}6+1{,}5=2{,}2.$$

$$DX=(0-2{,}2)^2\cdot 0{,}1+(1-2{,}2)^2\cdot 0{,}1+(2-2{,}2)^2\cdot 0{,}3+(3-2{,}2)^2\cdot 0{,}5=0{,}96.$$

$$\sigma=\sqrt{DX}=0{,}98.$$

2. Найти математическое ожидание, дисперсию и среднеквадратическое отклонение непрерывной случайной величины *X*, заданной плотностью распределения:

$$p(x)=\begin{cases}0, & x\le 0,\quad x>1,\\ 3x^2, & 0<x\le 1.\end{cases}$$

Решение.

$$M(X)=\int\limits_{-\infty}^{\infty} xp(x)\,dx=\int\limits_0^1 3x^3\,dx=\left.\frac{3x^4}{4}\right|_0^1=\frac{3}{4},$$

$$M(X^2)=\int\limits_{-\infty}^{\infty} x^2 p(x)\,dx=\int\limits_0^1 3x^4\,dx=\left.\frac{3x^5}{5}\right|_0^1=\frac{3}{5},$$

$$D(X)=M(X^2)-(M(X))^2=\frac{3}{5}-\frac{9}{16}=\frac{48-45}{80}=\frac{3}{80},$$

$$\sigma(X)=\sqrt{D(X)}=\sqrt{\frac{3}{80}}=0{,}194.$$

Задачи для самостоятельного решения

1. Найти функции распределения случайных величин X и $У$ $p_X(x)=\begin{cases}c_1 e^{-x}, x\ge 0\\ 0, x<0\end{cases}$, $p_Y(x)=\begin{cases}c_2/x^2, x\ge 1\\ 0, x<1\end{cases}$.

2. Найти математическое ожидание, дисперсию и среднеквадратическое отклонение случайной величины X:

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

3. Найти математическое ожидание, дисперсию и среднеквадратическое отклонение случайной величины X:

X	-2	-1	0	1	2
P	1/8	¼	3/16	1/16	3/8

4. Найти математическое ожидание, дисперсию и среднеквадратическое отклонение случайной величины X:

$$p_X(x)=\begin{cases}e^{-x}, x\ge 0\\ 0, x<0.\end{cases}$$

5. Найти математическое ожидание, дисперсию и среднеквадратическое отклонение случайной величины X

$$p_X(x)=\begin{cases}0, x\le 0, x>\pi\\ \frac{1}{2}\sin x, x\in(0,\pi].\end{cases}$$

Ответы

1	2	3	4	5
$F(x)=\begin{cases}0, & x<0,\\ -c_1(e^{-x}-1), & x\ge 0.\end{cases}$ $F(x)=\begin{cases}0, & x<0,\\ -c_2\left(\frac{1}{x}-1\right), & x\ge 1.\end{cases}$	$M(X)$=0,9 $D(X)$=1,29 $\sigma(X)$=1,136	$M(X)=\frac{5}{16}$ $D(X)=\frac{567}{256}$ $\sigma(X)$= 1,488	$M(X)$= $=D(X)$= $=\sigma(X)$= =1	$M(X)=\pi/2$ $D(X)$= $\pi^2/4-2$ $\sigma(X)$= $=\frac{1}{2}\sqrt{\pi^2-8}$

2.19 Другие характеристики случайных величин

Мода M_0	$P(M_0)=\max P(X)$	Наиболее вероятное значение по сравнению с соседними значениями. Если мода единственная, то случайная величина называется унимодальная, в противном случае – полимодальная.
Медиана M_e	$P(X<M_e)=P(X>M_e)$	В случае симметричного распределения случайной величины мода и медиана совпадают с математическим ожиданием.
Коэффициент асимметрии *A*	$\dfrac{M(X-MX)^3}{\sigma^3}$	Характеризует асимметрию распределения. Если *A*>0 то кривая более полога справа от *M(X)*, при *A* < 0 полога слева от *M(X)*.
Эксцесс	$\dfrac{M(X-MX)^4}{\sigma^4}-3$	Характеризует островершинность графика плотности распределения

Кроме рассмотренных выше числовых характеристик случайной величины, в приложениях используются так называемые квантили.

Квантилью уровня p случайной величины X называется решение уравнения: $F_\xi(x_p) = p$, где $p \in (0,1)$.

Квантили $X_{0.25}, X_{0.5}$, $X_{0.75}$ имеют названия нижняя квартиль, медиана, верхняя квартиль. Они делят числовую прямую на четыре части, вероятности попадания в которые равны 0,25.

Примеры

1. Найти моду случайной величины *X*, заданной распределением:

X	-2	-1	0	1	2
P	1/8	¼	3/16	1/16	3/8

Решение. Поскольку для моды выполняется равенство: $P(M_0) = \max P(X)$.

Наибольшая вероятность достигается при $M_0 = 2$.

2. Найти эксцесс случайной величины *X*:

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

Решение.

Найдем начальные моменты случайной величины *X* первых четырех порядков:

$\nu_1 = M(X) = 0 \cdot 0,1 + 1 \cdot 0,5 + 2 \cdot 0,4 = 1,3.$

$\nu_2 = M(X^2) = 0^2 \cdot 0,1 + 1^2 \cdot 0,5 + 2^2 \cdot 0,4 = 2,1.$

$\nu_3 = M(X^3) = 0^3 \cdot 0,1 + 1^3 \cdot 0,5 + 2^3 \cdot 0,4 = 3,7.$

$\nu_4 = M(X^4) = 0^4 \cdot 0,1 + 1^4 \cdot 0,5 + 2^4 \cdot 0,4 = 6,9.$

Найдем центральные моменты случайной величины *X* первых четырех порядков:

$\mu_1 = 0.$

$\mu_2 = \nu_2 - \nu_1^2 = 2,1 - 1,3^2 = 0,41.$

$\mu_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3 = 3,7 - 3 \cdot 1,3 \cdot 2,1 + 2 \cdot 1,3^3 = -0,096.$

$\mu_4 = \nu_4 - 4\nu_1\nu_3 + 6\nu_1^2\nu_2 - 3\nu_1^4 = 6,9 - 4 \cdot 1,3 \cdot 3,7 + 6 \cdot 1,3^2 \cdot 2,1 - 3 \cdot 1,3^4 = 0,3857$

.

Тогда среднеквадратическое отклонение случайной величины *X*:

$\sigma = \sqrt{D(X)} = \sqrt{\mu_2} = \sqrt{0,41} = 0,6403.$

Эксцесс случайной величины *X* найдем по формуле:

$E = \dfrac{\mu_4}{\sigma^4} - 3.$

$E = \dfrac{0,3857}{0,41^2} - 3 = -0,7055.$

Задачи для самостоятельного решения

1. Найти моду случайной величины *X*:

X	0	10	20	30	40
P	1/8	3/8	3/16	1/16	1/4

2. Найти коэффициент асимметрии случайной величины *X*:

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

3. Найти $X_{0.25}$ случайной величины *X*:

X	2	4	6
P	0,1	0,5	0,4

4. Найти эксцесс случайной величины *X*:

X	-1	0	1
P	0,2	0,2	0,6

5. Найти моду случайной величины *X*:

X	8	9	10
P	1/10	5/10	4/10

Ответы

1	**2**	**3**	**4**	**5**
10	-0,3657	2	-0,9219	9

2.20 Нормальный закон распределения

Непрерывная случайная величина *X* распределена по нормальному закону с параметрами $a, \sigma > 0$, если ее плотность распределения имеет вид:

$$p_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{\sigma^2}}, \; x \in R.$$

То, что случайная величина *X* имеет нормальное распределение, обозначается следующим образом:

$$X \sim N(a,\sigma).$$

Функция распределения: $F(x) = \int_{-\infty}^{x} p(t)dt$ примет вид:

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

Если $a = 0, \sigma = 1$, то нормальной распределения с такими параметрами называется стандартным.

Если $X \sim N(a,\sigma)$, то $M_0 = M_e = a, A = 0, E = 0$.

Свойства нормальной случайной величины:

1) $p(x) > 0,\ \forall x \in (-\infty, \infty)$, график функции расположен выше оси *Ox*.

2) Ось *Ox* служит асимптотой графика функции $p(x)$, т.к. $\underset{x\to\pm\infty}{Lim}\, p(x) = 0$.

3) Функция $p(x)$ имеет один максимум при $x = a$, равный $p(a) = \frac{1}{\sigma\sqrt{2\pi}}$.

4) График функции $p(x)$ симметричен относительно прямой $x = a$.

5) Точки $M_1(a-\sigma, \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2}), M_2(a+\sigma, \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2})$ являются точками перегиба графика функции $p(x)$.

Вероятность попадания случайной величины $X \sim N(a, \sigma)$ на заданный участок (*a*, *b*):

$$P(a < X < b) = \int_a^b p(x)dx.$$

Случайная величина $X \sim N(a, \sigma)$ принимает свое значения в промежутке $(a-3\sigma, a+3\sigma)$ с вероятностью 0,9973.

Вероятность отклонения нормально распределенной случайной величины от ее математического ожидания по абсолютной величине определяется по формуле:

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

Вероятность отклонения относительной частоты $\varpi = \frac{m}{n}$ от вероятности наступления события *p* в серии из *n* независимых испытаний выражается формулой:

$$P(|\varpi - p| < \varepsilon) = 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right).$$

Примеры

1. Суточное потребление электроэнергии исправной печью является случайной величиной, распределенной по нормальному закону со средним 1000 кВт и среднеквадратичным отклонением 50. Если суточное потребление превысит 1100 кВт, то по инструкции печь отключают и ремонтируют. Найти вероятность ремонта печи.

Решение. Случайная величина *X* есть суточное потребление электроэнергии печью. M(X) = $1000, \sigma(X) = 50$. Найдем вероятность попадания случайной величины *X* в интервал (0; 1100). Для этого воспользуемся формулой:

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right).$$

$$P\left(\left|0 < X < 1100\right|\right) = \Phi\left(\frac{1100-0}{50}\right) - \Phi\left(\frac{0-1000}{50}\right) = 2\Phi(2) = 2 \cdot 0,4772 = 0,9544.$$

Тогда вероятность ремонта печи равна 1-0,9544 = 0,0456.

2. Рост мальчиков возрастной группы 15 лет есть нормально распределённая случайная величина X с параметрами $\alpha = 161$ см и $\sigma = 4$ см.

Какую долю костюмов для мальчиков, имеющих рост от 152 до 158 см, нужно предусмотреть в объёме производства для данной возрастной группы.

Решение.

$M(X) = 161$, $\sigma(X) = 4.$ Найдем вероятность попадания случайной величины X в интервал (152; 158). Для этого воспользуемся формулой:

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right).$$

$$P(152 < X < 158) = \Phi\left(\frac{158-161}{4}\right) - \Phi\left(\frac{152-161}{4}\right) = 0,214.$$

3. Текущая оценка ценной бумаги представляет собой нормально распределенную случайную величину со средним значением 100 у. е. и дисперсией 9. Найти вероятность того, что цена актива (ценной бумаги) будет находиться в пределах от 91 до 109 у. е.

Решение. Так как $a = 100,\quad \sigma = \sqrt{D} = 3$, тогда

$$P(91 < X < 109) = P(|X - 100| < 9) = 2\Phi\left(\frac{9}{3}\right) = 2\Phi(3) = 0,9973$$

Задачи для самостоятельного решения

1.Пусть случайная величина X подчиняется нормальному закону распределения вероятностей с параметрами $a = 0, \sigma = 2$. Определить 1) $P(-2 < X < 3)$, 2) $P(|X| < 0,1)$.

2.Известно, что случайная величина $X \sim N(3,2)$. Найти $P(-3 < X < 5)$.

3.Нормальное распределение случайной величины X задана плотностью распределения вероятностей $p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2}$. Найти вероятность попадания в интервал (1,3).

4. Ошибка взвешивания – случайная величина, распределенная по нормальному закону с математическим ожиданием, равным 0, и среднеквадратическим отклонением, равным 2 грамма. Найти вероятность того, что взвешивание проведено с ошибкой, не превышающей по модулю 4 грамма.

5. Уровень воды в реке – случайная величина со средним значением 2,5 м и стандартным отклонением 0,2 м. Оценить вероятность того, что в наудачу выбранный день уровень воды окажется в пределах от 2м 20см до 2м 80см.

Ответы

1	**2**	**3**	**4**	**5**
0,775; 0,04	0,84	0,157	0,954	0,866

2.21 Неравенство Чебышева

Необходимо рассмотреть условия, при выполнении которых совокупное действие очень многих случайных причин приводит к результату, почти не зависящему от случая, так как позволяет предвидеть ход явлений. Эти условия и указываются в теоремах, носящих общее название *закона больших чисел*. К ним относятся теоремы Чебышева и Бернулли. Теорема Чебышева является наиболее общим законом больших чисел, теорема Бернулли – простейшим.

Теорема (Неравенство Чебышева). Вероятность того, что отклонение случайной величины X от ее математического ожидания по абсолютной величине меньше положительного числа ε, не меньше чем $1-\dfrac{D(X)}{\varepsilon^2}$:

$$P(|X-MX|<\varepsilon)\geq 1-\frac{D(X)}{\varepsilon^2}.$$

Теорема Чебышева. Если последовательность попарно независимых случайных величин $X_1, X_2, ..., X_n, ...$ имеет конечное математическое ожидание и дисперсии этих величин равномерно ограничены (не превышают постоянного числа С), то среднее арифметическое случайной величины сходится по вероятности к среднему арифметическому их математического ожидания, т.е. если ε – любое положительное число, то

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}M(X_i)\right|<\varepsilon\right)=1.$$

В этом случае среднее арифметическое достаточно большого числа независимых случайных величин (дисперсии которых равномерно ограничены) утрачивает характер случайной величины.

Если проводится серия измерений какой-либо физической величины, причем:

1) результат каждого измерения не зависит от результатов остальных (измерения попарно независимы);

2) измерения производятся без систематических ошибок, (имеют одно и то же математическое ожидание);

3) обеспечена определенная точность измерений, (дисперсии их ограничены)

то при достаточно большом числе измерений их среднее арифметическое окажется сколь угодно близким к истинному значению измеряемой величины.

Теорема Бернулли. Если в каждом из независимых опытов вероятность появления события A постоянна, то при достаточно большом числе испытаний вероятность того, что модуль отклонения относительной частоты появлений A в n опытах от p будет сколь угодно малым, как угодно близка к 1.

Закон больших чисел не исследует вид предельного закона распределения суммы случайных величин. Этот вопрос рассмотрен в группе теорем, называемых *центральной предельной теоремой*. Они утверждают, что закон распределения суммы случайных величин, каждая из которых может иметь различные распределения, приближается к нормальному при достаточно большом числе слагаемых. Этим объясняется важность нормального закона для практических приложений.

Центральная предельная теорема. Если $X_1, X_2, ..., X_n$ – независимые случайные величины с одинаковым законом распределения, математическим ожиданием a и дисперсией σ^2, то при неограниченном увеличении n закон распределения суммы $Y_n = \sum_{k=1}^{n} X_k$ неограниченно приближается к нормальному.

Примеры

1. Найти вероятность того, что $|X - MX| < 2$ для случайной величины X:

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

Решение. По неравенству Чебышева: $P(|X - MX| < 2) \geq 1 - \frac{D(X)}{4}$. Найдем математическое ожидание случайной величины X.

MX =0,2-0,2+0,2+0,2 = 0,

DX = 0,4+0,2+0,2+0,4 = 1,2,

$$P(|X - MX| < 2) \geq 1 - \frac{1,2}{4} = 0,7.$$

2. Найти вероятность того, что $1,5 < X - MX < 1,5$ для случайной величины X:

X	-2	-1	0	1	2
P	1/7	1/7	3/7	1/7	1/7

Решение.

Математическое ожидание случайной величины принимает значение:

MX =0,

Дисперсия равна

DX = 10/7,

$$P(|X - MX| < 1,5) \geq 1 - \frac{10/7}{2,25} = 0,63.$$

Задачи для самостоятельного решения

1. Найти вероятность того, что $|X - MX| < 1,7$ для случайной величины *X*:

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

2. Найти вероятность того, что $|X - MX| < 1,2$ для случайной величины *X*:

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

3. Найти вероятность того, что $-2 < X - MX < 2$ для случайной величины *X*:

X	-2	1	3	7	10
P	0,2	0,1	0,1	0,1	0,5

4. Найти вероятность того, что $-5 < X - MX < 5$ для случайной величины *X*:

X	0	1	2	7	17
P	0,1	0,1	0,05	0,05	0,7

5. Найти вероятность того, что $-1,9 < X - MX < 1,9$ для случайной величины *X*:

X	5	6	7	8	9
P	0,15	0,05	0,1	0,1	0,6

3. ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ РАЗДЕЛА «ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ»

3.1 Статистический материал и его обработка

Результаты наблюдений массовых явлений, случайных величин составляют статистические данные или *статистический материал*. *Выборкой* объёма n называется совокупность n случайно отобранных объектов. Множество всех объектов, из которых производится выборка, называется *генеральной совокупностью* (ГС).

Выборочный метод состоит в том, что на основании изучения некоторого количественного признака X у некоторой части статистической совокупности (выборки), полученной в результате статистического отбора, можно сделать вывод о характере распределения этого признака по всей статистической совокупности (генеральной совокупности).

Результаты наблюдений выборки n объёма записываются, в частности, в виде статистической совокупности:

$\boldsymbol{i}$	1	2	…	n	номера наблюдений, измерений
$\boldsymbol{x}$	x_1	x_2	…	x_n	значение наблюдаемой величины

При больших значениях n и различных значениях x_i статистическую совокупность подвергают специальным видам статистической обработки.

Расположим значения x_i, которые назовём *вариантами*, в порядке возрастания и обозначим $a = \min x_i$; $b = \max x_i$. Величина $R = b - a$ называется *размахом* статистической совокупности. Среди значений x_i могут быть одинаковые. Пусть значение x_1 наблюдалось n_1 раз, x_2 – n_2 раз, …, x_k наблюдалось n_k раз. Тогда общий объём выборки равен $n = \sum_{i=1}^{k} n_i$. Число n_i, показывающее, сколько раз встречается варианта (значение) x_i, называется *частотой* x_i, а число $p_i^* = \frac{n_i}{n}$ – *относительной частотой* варианты x_i.

Последовательность x_i, записанная в *порядке возрастания* с указанием частот и (или) относительных частот, называется *вариационным рядом*. Статистическим рядом называется последовательность пар (x_i, n_i). Обычно статистический ряд записывается в виде следующей таблицы:

Варианта x_i	x_1	x_2	…	x_k	**Сумма**
Частота n_i	n_1	n_2	…	n_k	$\boldsymbol{n}$
Относительная частота p_i^*	p_1^*	p_2^*	…	p_k^*	**1**

Геометрическим изображением вариационного ряда является *эмпирический полигон распределения*, являющийся аналогом плотности распределения случайной величины X_i– ломаная с вершинами (x_i, p_i^*) – см. рисунок 1.

Вариационный ряд обозрим при небольших значениях k ($k \leq 20$). В противном случае его (или первоначальную статистическую совокупность) подвергают интервальной обработке.

Все варианты x_i принадлежат отрезку $[a, b]$. Пусть k некоторое (не больше 20) натуральное число. Отрезок $[a, b]$ разобьём на k равных частей длины $h = \frac{b-a}{k}$. Обозначим эти промежутки следующим образом: $[a_0,\ a_1[;\ [a_1, a_2[;\ \dots, [a_{k-1}, a_k]$. Через m_i обозначим число вариант, попавших в интервал $[a_{i-1}, a_i[$, *при этом будем считать, что каждый промежуток содержит свой левый конец, но лишь последний промежуток содержит и свой правый конец*. Пусть $\omega_i^* = \frac{m_i}{n}$ (числа m_i и ω_i можно также отнести к середине a_i^* a_i^* интервала (a_{i-1}, a_i)). Полученные данные занесём в таблицу, называемую *интервальной обработкой ряда,* или статистической совокупности.

Интервал	$[a_0, a_1[$	$[a_1, a_2[$	…	$[a_{k-1}, a_k]$	Сумма
Середины	a_1^*	a_2^*	…	a_k^*	–
Частоты	m_1	m_2	…	m_k	$\boldsymbol{n}$
Относительные частоты	p_1^*	p_2^*	…	p_k^*	**1**

Количество интервалов k можно рассчитать по *формуле Стерджеса* $k = 1 + 3{,}332 \cdot lgn$ либо с помощью таблицы:

Объём выборки n	25-40	40-60	60-100	100-200	> 200
Число интервалов k	5-6	6-8	7-10	8-12	10-15

Геометрическим изображением интервальной обработки служит *гистограмма* (см. рисунок 1). *Гистограммой частот* называется множество прямоугольников с основаниями (a_{i-1}, a_i) и высотами $\frac{m_i}{h}$. Площадь гистограммы равна объёму выборки n.

Нормированная гистограмма (гистограмма относительных частот) представляет собой ступенчатую фигуру из прямоугольников с основаниями равными интервалам значений признака (a_{i-1}, a_i) и высотами, равными плотности частоты $\frac{m_i}{h \cdot n}$. Если соединить прямолинейными отрезками середины верхних оснований прямоугольников, получим полигон распределения. Суммарная площадь всех прямоугольников гистограммы равна 1:

$$\sum_{i=1}^{k} \frac{m_i}{h \cdot n} \cdot h = \sum_{i=1}^{k} \frac{m_i}{n} = \frac{1}{n} \sum_{i=1}^{k} m_i = \frac{n}{n} = 1$$

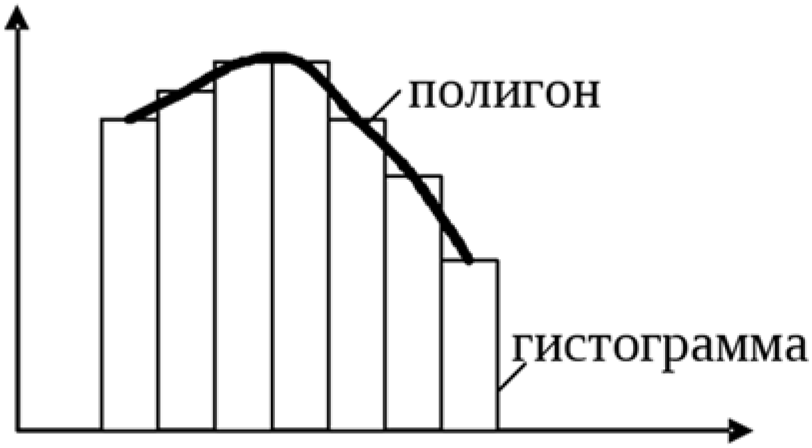


Рисунок 1

Эмпирической функцией распределения (функцией распределения выборки) называется функция $F^*(x)$, определяющая для каждого значения x относительную частоту события $X < x$ (см. рисунок 2):

$$F^*(x) = \frac{n_x}{n}, \qquad (1)$$

где n_x – число вариант x_i, меньших чем x, n – объем выборки.

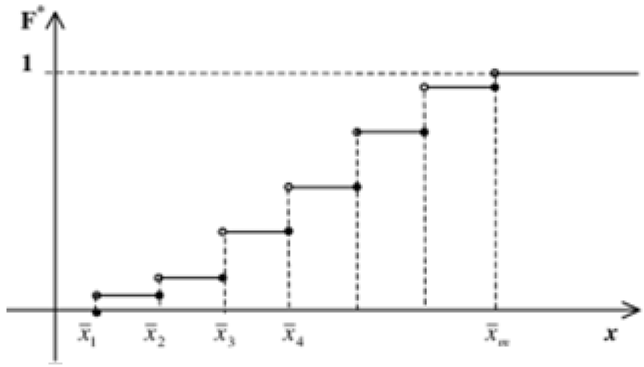


Рисунок 2

Функция $F^*(x)$ обладает следующими свойствами (здесь $a = \min x_i$; $b = \max x_i$):

1) $F^*(x) \in [0; 1]$; 2) $F^*(x) = 0$ при $x < a$, $F^*(x) = 1$ при $x > b$; 3) $F^*(x)$ – монотонно неубывающая, непрерывная слева функция.

Функция $F^*(x)$ является статистическим аналогом функции распределения $F(x)$ генеральной совокупности. Функцию распределения $F(x)$ в математической статистике называют *теоретической функцией распределения*. Различие между теоретической и эмпирической функциями распределения состоит в том, что $F(x)$ определяет вероятность события $X < x$, а $F^*(x)$ – относительную частоту этого события.

Т. е. эмпирическая функция распределения служит для оценки вида теоретической функции распределения случайного признака, полигон и гистограмма – для оценки вида теоретической кривой распределения.

3.2 Числовые характеристики законов распределения эмпирических величин

Одна из задач математической статистики состоит в установлении закона распределения случайной величины *X* (генеральной совокупности) и оценке параметров этого закона.

Вид закона выбирается из каких-либо теоретических или практических соображений, а параметры следует вычислять, исходя из параметров этого закона.

Важнейшим этапом обработки статистических данных является вычисление *оценок* числовых характеристик исследуемой случайной величины.

Полученные оценки позволяют в числовой форме описать характерные черты статистического распределения и являются базой для построения математической модели изучаемого случайного явления.

Любая величина $\tilde{\theta}$, определяемая как функция выборочных значений $\tilde{\theta} = \varphi(x_1, x_2, \ldots x_k)$, называется *выборочной статистикой* или просто *статистикой*. Статистика $\tilde{\theta}$, используемая в качестве приближённого значения неизвестного параметра θ, называется *статистической оценкой* параметра θ.

Существует два вида оценок параметров: точечные и интервальные.

Точечной называется статистическая оценка, которая определяется одним числом.

К точечным статистическим оценкам предъявляется ряд требований.

Если $\tilde{\theta}$ – статистическая оценка параметра θ, то она должна удовлетворять следующим условиям:

1) быть *несмещенной*, что означает, что $M(\tilde{\theta}) = \theta$.

2) быть *состоятельной*, т.е. предел по вероятности при $n \to +\infty$ последовательности таких оценок должен быть равен искомому параметру, т.е. вероятность того, что $|\theta - \tilde{\theta}| > \delta > 0$, стремится к нулю при $n \to +\infty$.

3) быть *эффективной*, т.е. дисперсия $D(\tilde{\theta})$ – наименьшая или быть *асимптотически эффективной*, что означает, что $\lim_{n\to\infty} D(\tilde{\theta}) = 0$.

Число $\delta > 0$ называется *точностью оценки*, если имеет место равенство $|\theta - \tilde{\theta}| < \delta$. Если это неравенство имеет место с некоторой вероятностью γ, то число γ называется *надёжностью оценки* или *уровнем надёжности*. Наиболее употребительными уровнями надёжности являются $\gamma = 0{,}95$; $\gamma = 0{,}99$; $\gamma = 0{,}999$.

Выборочной средней $\bar{x}$ называют среднее арифметическое значение случайной величины X по выборочной совокупности объёма n:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i n_i. \qquad (2)$$

Выборочная средняя служит *несмещённой оценкой математического ожидания* признака X или генеральной совокупности.

Кроме выборочной средней в статистическом анализе применяются структурные средние: *медиана* и *мода*.

Модой Mo называют варианту, которая имеет наибольшую частоту. Если распределение интервальное, то определяется модальный интервал $[x_i, x_{i+1}[$, которому соответствует наибольшая частота n_i, мода вычисляется по формуле:

$$Mo = x_i + h \cdot \frac{m_i - m_{i-1}}{(m_i - m_{i-1}) + (m_i - m_{i+1})}, \qquad (3)$$

где h – величина модального интервала; n_i и n_{i+1} – частоты предмодального и послемодального интервала.

Медианой Me называют варианту, которая делит вариационный ряд на две части, равные по числу вариант. Если $n = 2k + 1$, то $Me = x_{k+1}$, а если $n = 2k$, то $Me = \frac{x_k + x_{k+1}}{2}$. При вычислении медианы интервального ряда распределения используется формула:

$$Me = x_i + h \cdot \frac{0,5n - s_{i-1}}{m_i}, \qquad (4)$$

где s_{i-1} - накопленная частота интервала, предшествующего медианному, включая интервал, предшествующий медианному; x_i – начальное значение интервала, который содержит медиану. Номер медианного интервала определяется из неравенства $\sum_{i=1}^{L} m_i \le \frac{n}{2} < \sum_{i=1}^{L+1} m_i$. В случае выполнения равенства номер медианного интервала равен L, в противном случае – $L + 1$.

Средние величины не отражают изменчивости (вариации) значений признака. Чтобы охарактеризовать рассеяние наблюдаемых значений количественного признака выборки вокруг своего среднего значения $\bar{x}$ вводят свободную характеристику – *выборочную дисперсию*.

Выборочной дисперсией D_B называют среднее арифметическое квадратов отклонения наблюдаемых значений признака от их среднего значения $\bar{x}$:

$$D_B = \frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})^2 \cdot n_i = \frac{1}{n}\sum_{i=1}^{k} x_i^2 n_i - \bar{x}^2. \qquad (5)$$

Выборочным средним квадратическим отклонением (стандартом) называют квадратный корень из выборочной дисперсии:

$$\sigma_B = \sqrt{D_B}. \qquad (6)$$

Выборочная дисперсия является смещённой оценкой генеральной дисперсии, так как $M(D_B) = \frac{n}{n-1} D_Г$. В качестве н*есмещенной оценки генеральной дисперсии* служит «исправленная» выборочная дисперсия:

$$S^2 = \frac{n}{n-1} D_B = \frac{1}{n-1}\sum_{i=1}^{k}(x_i - \bar{x})^2 n_i. \qquad (7)$$

При достаточно больших n *выборочная и исправленная дисперсии мало отличаются, поэтому на практике исправленной дисперсией пользуются, если* $n < 50$.

Выборочная средняя и дисперсия вариационного ряда являются частными случаями более общего понятия – *выборочных (эмпирических) моментов*.

Начальный момент $\tilde{\nu}_k$ $k-$го порядка вариационного ряда определяется по формуле

$$\tilde{\nu}_k = \frac{1}{n}\sum_i x_i^k n_i. \qquad (8)$$

Центральный момент $\tilde{\mu}_k$ $k-$го порядка вариационного ряда определяется по формуле

$$\tilde{\mu}_k = \frac{1}{n}\sum_i (x_i - \bar{x})^2 n_i. \qquad (9)$$

В частности $\tilde{\nu}_1 = \bar{x}, \tilde{\mu}_1 = 0, \tilde{\mu}_2 = D_B$.

Центральные моменты первых четырёх порядков выборки $\{x_i\}$ выражаются через начальные моменты $\tilde{\nu}_k$ по формулам:

$$\tilde{\mu}_2 = \tilde{\nu}_2 - \tilde{\nu}_1^2;\ \tilde{\mu}_3 = \tilde{\nu}_3 - 3\tilde{\nu}_1\tilde{\nu}_2 + 2\tilde{\nu}_1^3;\ \ \tilde{\mu}_4 = \tilde{\nu}_4 - 4\tilde{\nu}_1\tilde{\nu}_3 + 6\tilde{\nu}_2\tilde{\nu}_1^2 - 3\tilde{\nu}_1^4. \qquad (10)$$

Коэффициентом асимметрии вариационного ряда называется число

$$\tilde{A}_s = \frac{\tilde{\mu}_3}{\sigma_B^3}. \qquad (11)$$

Эксцессом вариационного ряда называется число:

$$\tilde{E}_x = \frac{\tilde{\mu}_4}{\sigma_B^4} - 3. \qquad (12)$$

Асимметрия называется также нормированным третьим центральным моментом, а эксцесс – нормированным четвертым центральным моментом признака X. Знаки асимметрии и эксцесса указывают на отклонения графика закона распределения X от нормального распределения, для которого $\tilde{A}_s = 0$; $\tilde{E}_x = 0$. При $\tilde{A}_s > 0$ большая часть вариант будет расположена слева от $\bar{x}$ – имеет место левосторонняя асимметрия распределения, при $\tilde{A}_s < 0$– правосторонняя. Если $\tilde{A}_s = 0$, в этом случае распределение имеет симметричную форму (рисунок 3).

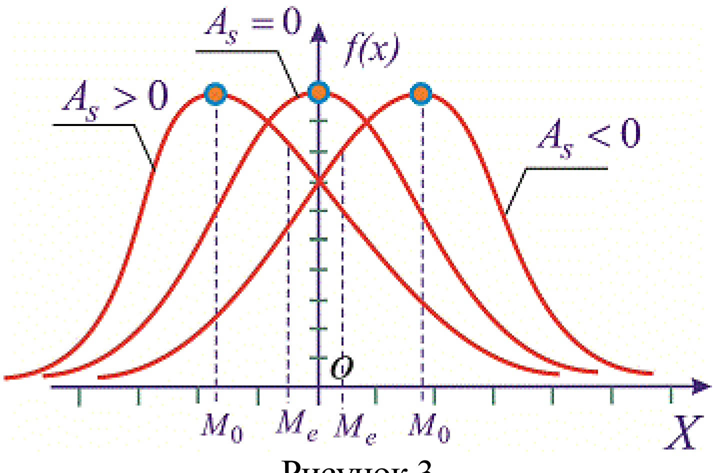


Рисунок 3

Положительное значение эксцесса указывает на то, что полигон распределения около моды имеет более высокую острую вершину, чем нормальная кривая, с тем же центром и той же дисперсией.

Отрицательное значение эксцесса имеет место для кривых с более низким и более плоским характером вершины по сравнению с нормальной кривой (см. рисунок 4).

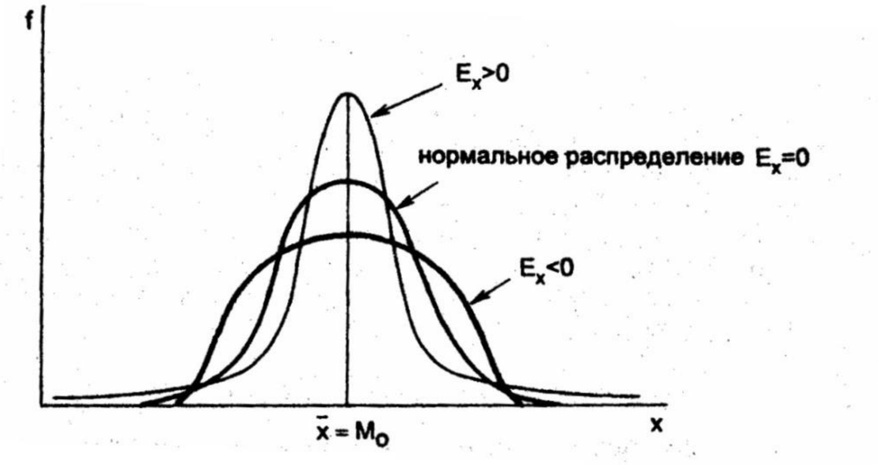


Рисунок 4

Точечные оценки не указывают величину ошибки, которая совершается при замене m_x и $\sigma(X)$ их приближёнными значениями (оценками). Поэтому иногда выгодно пользоваться *интервальной оценкой*, которая определяется двумя числами $\tilde{\theta}_1$ и $\tilde{\theta}_2$ – концами интервала, накрывающего оцениваемый параметр θ с заданной вероятностью (надёжностью).

Пусть $\tilde{\theta}$ – точечная оценка параметра θ. Она тем лучше, чем меньше разность $\left|\theta - \tilde{\theta}\right|$. Тогда в качестве характеристики *точности* оценки можно взять некоторое $\delta > 0$, такое, что $\left|\theta - \tilde{\theta}\right| < \delta$. Но в статистике можно говорить лишь о вероятности (надёжности) $\gamma = 1 - \alpha$, с которой выполняется это неравенство. Число α называется *уровнем значимости*.

Доверительной вероятностью оценки называется вероятность $\gamma = 1 - \alpha$ выполнения неравенства $\left|\theta - \tilde{\theta}\right| < \delta$. Обычно γ задаётся заранее и наиболее часто полагают $\gamma = 0{,}9; 0{,}95; 0{,}99$ и пр. Таким образом:

$$P\left(\left|\theta - \tilde{\theta} < \delta\right|\right) \Leftrightarrow P\left(\tilde{\theta} - \delta < \theta < \tilde{\theta} + \delta\right).$$

Доверительный интервал – это интервал $\left(\tilde{\theta} - \delta,\ \tilde{\theta} + \delta\right)$, который накрывает неизвестный параметр θ с заданной надёжностью $\gamma = 1 - \alpha$.

Границы интервала и его величина находятся по выборочным данным и поэтому являются случайными величинами в отличие от оцениваемого параметра θ, поэтому говорят, что $\left(\tilde{\theta} - \delta,\ \tilde{\theta} + \delta\right)$ накрывает, а не содержит истинное значение θ.

Величина доверительного интервала существенно зависит от объёма выборки n (уменьшается с ростом n) и значения доверительной вероятности γ (увеличивается с приближением γ к единице).

Интервальной оценкой с надёжностью γ математического ожидания a нормально распределённой случайной величины (признака) Х по выборочной средней $\bar{x}$ при известном СКО σ генеральной совокупности служит доверительный интервал

$$\bar{x} - t \cdot \frac{\sigma}{\sqrt{n}} < a < \bar{x} + t \cdot \frac{\sigma}{\sqrt{n}},$$

где $\delta = t \cdot \frac{\sigma}{\sqrt{n}}$ – точность оценки, n – объём выборки, t – значение аргумента функции Лапласа $\Phi(t)$, при котором $\Phi(t) = \frac{\gamma}{2}$.

При неизвестном σ (в условиях эксперимента σ обычно неизвестно) доверительный интервал для математического ожидания a нормально распределённой случайной величины Х имеет вид:

$$\bar{x} - t_\gamma \cdot \frac{S}{\sqrt{n}} < a < \bar{x} + t_\gamma \cdot \frac{S}{\sqrt{n}}, \qquad (13)$$

где t_γ – квантиль распределения Стьюдента, определяемый по таблицам, а параметры $\bar{x}, S$ находятся по данным выборки.

При больших выборках $(n > 30)$ распределение Стьюдента приближается к нормальному, и тогда можно пользоваться теоремами о нормальном распределении.

Доверительный интервал для σ задаётся неравенствами:

$$S \cdot \sqrt{\frac{n-1}{\chi^2_{\frac{\alpha}{2},n-1}}} < \sigma < S \cdot \sqrt{\frac{n-1}{\chi^2_{1-\frac{\alpha}{2},n-1}}},$$

где $\chi^2_{\frac{\alpha}{2},n-1}$, $\chi^2_{1-\frac{\alpha}{2},n-1}$ – квантили χ^2 распределения, определяемые по соответствующим таблицам по заданному уровню значимости α и числу степеней свободы $k = n - 1$, либо

$$S(1 - q) < \sigma < S(1 + q), \text{ если } q < 1;$$
$$0 < \sigma < S(1 + q), \text{ если } q > 1.$$

Величина q находится по таблице $q = q(\gamma, n)$ и зависит от надежности и объема выборки.

3.3 Статистическая проверка гипотез. Основные понятия

Статистической гипотезой называется предположение относительно параметров или вида распределения изучаемой случайной величины.

Статистические гипотезы можно разделить на следующие основные группы:

1)гипотезы о параметрах распределения;

2)гипотезы о виде распределения.

Выдвинутую гипотезу называют *нулевой* и обозначают ее через H_0. Наряду с H_0 рассматривают *конкурирующую* (или *альтернативную*) гипотезу H_1.

Таким образом, ставится задача проверки гипотезы H_0 относительно конкурирующей гипотезы H_1 на основе выборки X объема n. Правило, по которому принимается или отвергается гипотеза, называется *статистическим критерием*. Принципы проверки статистических гипотез впервые были сформулированы в работах известных математиков Е. Неймана и Э. Пирсона. Они исходили из того, что принимая или отвергая гипотезу H_0, можно допустить ошибки двух видов.

Ошибка первого рода: H_0 отвергается (принимается H_1) в то время как в действительности верна гипотеза H_0. Вероятность ошибки первого рода называют *уровнем значимости* и обозначают α:

$$P(H_1/H_0) = \alpha.$$

Величину $\gamma = 1 - \alpha$, то есть вероятность принять верную гипотезу, называют *уровнем доверия* (*доверительным уровнем*).

Ошибка второго рода: H_0 принимается, в то время как верна гипотеза H_1. Вероятность ошибки второго рода обозначается β.

Вероятность принять гипотезу H_1, если она верна, называют *мощностью критерия.*

Суть проверки статистической гипотезы заключается в том, что используется специально составленная выборочная характеристика (*статистика*) $K = K(x_1, x_2, \dots, x_n)$, полученная по выборке X, так, чтобы в случае, если гипотеза H_0 верна, точное или приближенное распределение K было бы известным. Построение критерия, в зависимости от вида гипотезы H_0, заключается в выборе таких значений K_{kp}^1 и K_{kp}^2, что если $K_{kp}^1 < K < K_{kp}^2$, то гипотеза H_0 принимается. Значения K_{kp}^1 и K_{kp}^2 называются *критическими*, а область $K_D = \left\{K : K_{kp}^1 < K < K_{kp}^2 \right\}$ называется *областью допустимых значений*.

Множество возможных значений статистики K разбивается на 2 непересекающихся подмножества: *критическую область* – множество значений K, при которых H_0 отвергается – $\overline{K}_D$, и *область допустимых значений* – множество значений K, при которых H_0 принимается – K_D. Если фактически наблюдаемое (полученное по выборке) значение статистики критерия K попадает в критическую область, то гипотезу H_0 отвергают, в противном случае принимают.

3.4 Проверка гипотезы о нормальном распределении генеральной совокупности. Критерий согласия Пирсона

Одной из задач математической статистики является установление истинного закона распределения случайной величины на основании экспериментальных данных. Критерии, устанавливающие закон распределения, называются *критериями согласия*.

Алгоритм применения критерия Пирсона.

1)Из генеральной совокупности образовывается случайная выборка, и на ее основе делается предположение о нормальном законе распределения. Выдвигается гипотеза H_0: «*генеральная совокупность распределена нормально*».

2)Вычисляются выборочные числовые характеристики $\bar{x}, \sigma_B$.

3)Вычисляются теоретические частоты:

а) Для дискретного ряда

$$n_i' = \frac{n \cdot h}{\sigma_B} \cdot \varphi(u_i),$$

где n – объем выборки, h – шаг (разность между двумя соседними вариантами),

$$u_i = \frac{x_i - \bar{x}}{\sigma_B}; \;\; \varphi(u) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{u^2}{2}}$$

Значения $\varphi(u)$ определяются из таблицы приложения 1.

б) Для интервального ряда $n_i' = n \cdot p_i$, где n – объем выборки, $p_i = \Phi(z_{i+1}) - \Phi(z_i)$ – теоретические вероятности попадания в интервалы $x_i - x_{i+1}$, $z_i = \frac{x_i - \bar{x}}{\sigma_B}$, $z_{i+1} = \frac{x_{i+1} - \bar{x}}{\sigma_B}$, $\Phi(z)$ – функция Лапласа, значения которой определяются по таблице приложения 2.

4)Находится наблюдаемое значение критерия Пирсона по формуле

$$\chi^2 = \sum_{i=1}^{l} \frac{(n_i - n_i')^2}{n_i'}.$$

5)По таблице критических точек распределения χ^2, по заданному уровню значимости α и числу степеней свободы $k = l - 3$ (l – число групп для дискретного ряда или число интервалов для интервального ряда) находят критическую точку $\chi^2_{kp}(\alpha; k)$ правосторонней критической области.

6)Если $\chi^2 < \chi^2_{kp}(\alpha; k)$ – нет оснований отвергнуть гипотезу о нормальном распределении генеральной совокупности. Другими словами, эмпирические и теоретические частоты различаются незначимо. Если $\chi^2 > \chi^2_{kp}(\alpha; k)$ – гипотезу отвергают.

Замечание. Малочисленные варианты и интервалы (содержащие малочисленные частоты $n_i < 5$) следует объединить, а соответствующие им частоты сложить. Если производилось объединение частот, то в формуле $k = l - 3$ следует в качестве l принять число групп или интервалов выборки, оставшихся после объединения частот.

3.5 Элементы теории регрессионного и корреляционного анализа

Методы теории корреляции позволяют определять зависимость между различными факторами или случайными величинами. Термин «корреляция» происходит от латинского «correlatio» - соотношение, взаимосвязь.

В естественных науках часто речь идёт о *функциональной зависимости*, когда каждому значению одной величины соответствует вполне определённое значение другой. Случайные величины обычно не связаны функциональной зависимостью. В большинстве случаев между переменными существуют зависимости, когда каждому значению одной переменной соответствует не какое-то определённое, а множество возможных значений другой переменной. Такая зависимость получила название *статистической* (или *стохастической, вероятностной*).

В силу неоднозначной статистической зависимости между случайными величинами X и Y для исследователя представляет интерес усреднённая схема зависимости – зависимость условного математического ожидания $M_x(Y) = M(Y/X = x)$ или его статистического аналога $\bar{y}_x$ от значений x случайной величины X, то есть $M_x(Y) = f(x)$ или $\bar{y}_x = f(x)$. Здесь $\bar{y}_x$ – условная средняя, которая определяется как среднее арифметическое значений Y, то есть y_i, соответствующих значению $X = x$. Такая зависимость получила название *корреляционной. Корреляционной зависимостью* Y *от* X называют функциональную зависимость условной средней $\bar{y}_x$ от x:

$$\bar{y}_x = f(x). \tag{14}$$

Уравнение (14) называют *уравнением регрессии* Y *на* X; функцию $f(x)$ называют *регрессией* Y *на* X X, а её график – *линией регрессии* Y *на* X.

Статистические связи между переменными можно изучать методами *корреляционного* и *регрессионного* анализа. Основной задачей корреляционного анализа является выявление связи между случайными величинами и оценка ее тесноты. Основной задачей регрессионного анализа – установление и изучение формы зависимости между переменными.

Данные о статистической зависимости удобно представлять в виде корреляционной таблицы:

X	Y						
	y_1	y_2		y_j		y_m	n_x
x_1	n_{11}	n_{12}		n_{1j}		n_{1m}	n_{x1}
x_2	n_{21}	n_{22}		n_{2j}		n_{2m}	n_{x2}
…							…
x_i	n_{i1}	n_{i2}		n_{ij}		n_{im}	n_{xi}
…							…
x_k	n_{k1}	n_{k2}		n_{kj}		n_{km}	n_{xk}
n_y	n_{y1}	n_{y2}		n_{yj}		n_{ym}	n

Здесь $x_1, x_2, \ldots, x_k$; $y_1, y_2, \ldots, y_m$ – значения случайных величин X и Y соответственно, а $n_{x1}, n_{x2}, \ldots, n_{xk}$; $n_{y1}, n_{y2}, \ldots, n_{ym}$ – соответствующие частоты, n_{ij} – частота, с которой встречается пара (x_i, y_j); $n = \sum_{i=1}^{k}\sum_{j=1}^{m} n_{ij}$.

По направлению корреляционная связь может быть положительной ("прямой") и отрицательной ("обратной"). При положительной прямолинейной корреляции более высоким значениям одного признака соответствуют более высокие значения другого, а более низким значениям одного признака - низкие значения другого. При отрицательной корреляции соотношения обратные.

Наличие корреляции приближенно может быть определено с помощью *корреляционного поля*. Его получим, если нанесем на график в определенном масштабе точки, соответствующие наблюдаемым одновременным значениям двух величин

(x_i, y_i) – если точки рассеяны хаотично, то связь между Y и X отсутствует; если точки группируются около какой-то линии, то связь есть, и она тем теснее, чем ближе они группируются (рисунок 5).

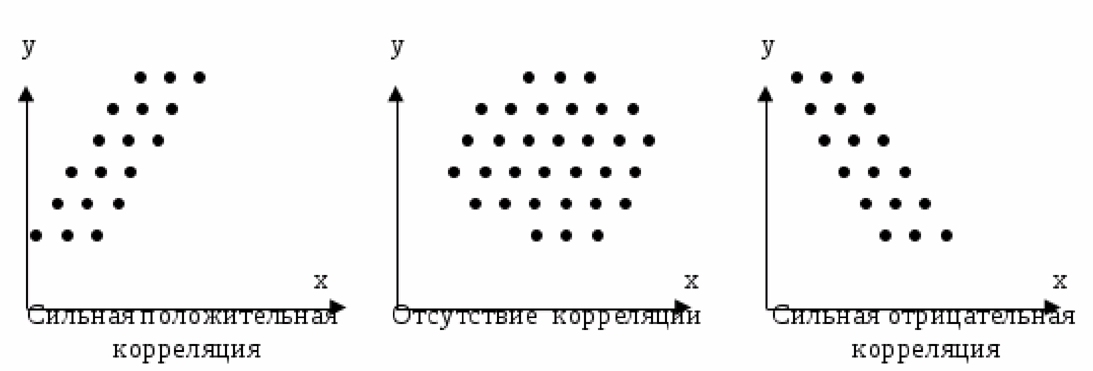


Рисунок 5

Рассмотрим наиболее важный для практики случай линейной зависимости между величинами. В теории вероятностей показателем тесноты линейной зависимости являлся коэффициент корреляции, в математической статистике таким показателем является выборочный коэффициент корреляции.

Выборочным коэффициентом корреляции называется величина, рассчитываемая по формуле:

$$r_B = \frac{s_{xy}}{s_x \cdot s_y}, \qquad (15)$$

где $s_{xy} = \frac{1}{n-1}\left(\sum_i \sum_j n_{ij} x_i y_j - \frac{1}{n}\left(\sum_i n_{xi} x_i\right) \sum_j n_{yj} y_j\right)$ – оценка корреляционного момента; s_x и s_y – исправленные средние квадратические отклонения.

Выборочный коэффициент корреляции обладает некоторыми свойствами:

1. $|r_B| \leq 1$;

2. Чем ближе значение $|r_B|$ к единице, тем более тесная *линейная* зависимость между изучаемыми величинами. В зависимости от того, насколько $|r_B|$ приближается к единице, различают слабую, умеренную, заметную, достаточно тесную и весьма тесную линейную связь.

3. Если $r_B > 0$, то говорят о прямой связи между изучаемыми величинами (т.е. с увеличением одной случайной величины увеличивается и другая), если же $r_B < 0$, говорят об обратной связи (с увеличением одной случайной величины вторая уменьшается).

4. Если все значения переменных увеличить (уменьшить) на одно и тоже число или в одно и то же число раз, то величина коэффициента корреляции не изменится. Коэффициент корреляции есть безразмерная характеристика тесноты линейной связи.

5. При $r_B = \pm 1$ корреляционная связь представляет линейную функциональную зависимость, при этом все точки поля корреляции лежат на одной прямой.

6. При $r_B = 0$ или $|r_B|$ близком к нулю линейная корреляционная связь отсутствует, но это не означает отсутствие другой зависимости, например, нелинейная связь может быть очень тесной.

Для ответа на вопрос о значимости коэффициента корреляции проверяют нулевую гипотезу $H_0\colon r_{Г} = 0$ о равенстве нулю генерального коэффициента корреляции. Если гипотеза принимается, то говорят, что между X и Y нет линейной корреляционной зависимости, иначе линейная зависимость признается значимой.

Для того чтобы при уровне значимости α проверить нулевую гипотезу о равенстве нулю генерального коэффициента корреляции при конкурирующей $H_1\colon r_{Г} \neq 0$, надо вычислить наблюдаемое значение критерия:

$$T_{\text{набл}} = r_B \cdot \sqrt{\frac{n-2}{1-r_B^2}}, \tag{16}$$

затем, пользуясь таблицей критических точек распределения Стьюдента, по заданному уровню значимости и числу степеней свободы $k = n - 2$ найти критическую точку $t_{kp}(\alpha; k)$ для двухсторонней критической области. Если сравнить данные величины, то можно сделать вывод о степени коррелированности исходных признаков:

•если $|T_{\text{набл}}| < t_{kp}$, то верна нулевая гипотеза и, следовательно, величины X, Y не коррелированны;

•если же $|T_{\text{набл}}| > t_{kp}$, то нулевая гипотеза отвергается.

Рассмотрим *уравнение парной линейной регрессии* $\bar{y}_x = a + bx$. Найдём формулы расчёта неизвестных параметров a и b по имеющимся статистическим данным $(x_i, y_i), i = \overline{1, n}$.

Согласно *методу наименьших квадратов* неизвестные параметры выбираются таким образом, чтобы сумма квадратов отклонений выборочных значений y_i от значений $\bar{y}_{x_i} = a + bx_i$, полученных по уравнению регрессии, была минимальна:

$$S = \sum_{i=1}^{n} (y_i - a - bx_i)^2 \to min.$$

На основании необходимого условия экстремума, приравнивая нулю частные производные, получим:

$$\begin{cases} 2\sum_{i=1}^{n} (y_i - a - bx_i) = 0 \\ 2\sum_{i=1}^{n} (y_i - a - bx_i)x_i = 0 \end{cases} \sim \begin{cases} n \cdot a + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \\ a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \end{cases}$$

После преобразования получаем систему нормальных уравнений для определения параметров линейной регрессии:

$$\begin{cases} a + b\bar{x} = \bar{y} \\ a\bar{x} + b\overline{x^2} = \overline{xy}. \end{cases} \tag{17}$$

Из последней системы следуют формулы для определения параметров уравнения парной линейной регрессии Y на X:

$$b = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} \sim b = r_B \frac{s_y}{s_x}; \quad a = \bar{y} - b\bar{x}. \tag{18}$$

Уравнение регрессии $\bar{y}_x = a + bx$ можно с учётом формулы вычисления параметра a записать в виде

$$\bar{y}_x - \bar{y} = b(x - \bar{x}).$$

Коэффициент b показывает, на сколько единиц в среднем изменится переменная Y при увеличении переменной X на одну единицу.

Уравнение регрессии может быть использовано для прогнозирования значений Y при значениях X, не указанных в корреляционной таблице.

Величину $D_{\text{ост}} = s_y^2(1 - r_B^2)$ называют *остаточной дисперсией* случайной величины Y относительно случайной величины X; она характеризует величину ошибки, которая возникает при замене Y линейной функцией. При $r_B = \pm 1$ остаточная дисперсия равна нулю, т.е. при представлении Y в виде линейной функции от X не возникает ошибки, а Y и X связаны *линейной функциональной* зависимостью.

Примеры

1. Из генеральной совокупности извлечена выборка объёма $n = 21$:

x_i	2	6	8	10	$\sum$
n_i	3	4	6	8	21

Требуется:

1)Найти и построить эмпирическую функцию распределения;

2)Найти выборочное среднее, «исправленное» СКО, выборочную моду и медиану.

Решение.

1) Согласно определению эмпирической функции распределения её значение при любом x равно $F^*(x) = \frac{n_x}{n}$, где n_x – количество элементов x_i выборки, меньших, чем x; $n = \sum n_i$ – объём выборки.

Например, при $x = 1$ имеем $n_x = 0,\ F^*(x) = 0$; при $x = 4$ $n_x = 3$ $F^*(x) = \frac{3}{21} = \frac{1}{7} \approx 0{,}1$; при $x = 7$ $n_x = 3 + 4 = 7,\ F^*(x) = \frac{7}{21} \approx 0{,}3$; при $x = 9$ $n_x = 3 + 4 + 6 = 13,\ F^*(x) = \frac{13}{21} \approx 0{,}6$**;** при $x = 11$ $n_x = 3 + 4 + 6 + 8 = 21;,$ $F^*(x) = \frac{21}{21} = 1$.

Тогда:

$$F^*(x) = \begin{cases} 0 & x \le 2 \\ 0{,}1 & 2 < x \le 6 \\ 0{,}3 & 6 < x \le 8 \\ 0{,}8 & 8 < x \le 10 \\ 1 & x > 10. \end{cases}$$

График эмпирической функции распределения изображён на рисунке 6.

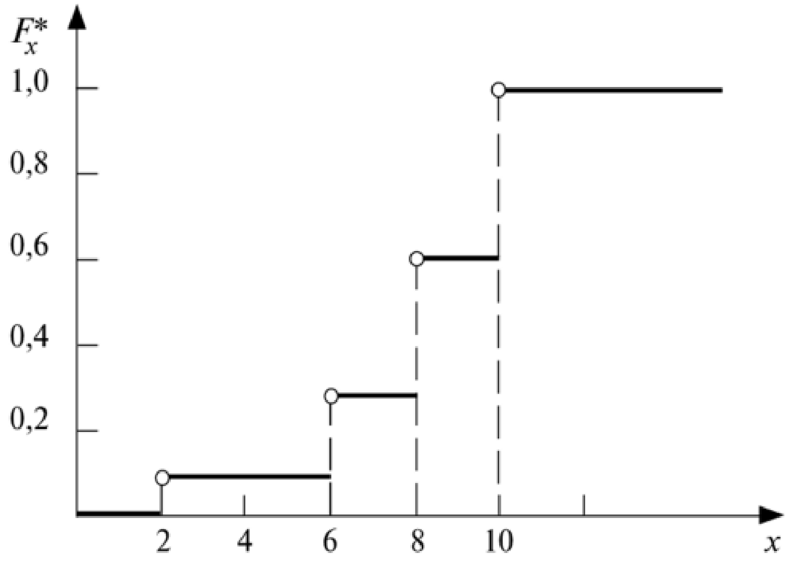


Рисунок 6

2) Определим выборочное среднее выборки по формуле (2):

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i n_i = \frac{1}{21}(2 \cdot 3 + 6 \cdot 4 + 8 \cdot 6 + 10 \cdot 8) = \frac{116}{21} \approx 7{,}52.$$

«Исправленную» дисперсию найдём, используя следующую формулу:

$S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot n_i.$

$$S^2 = \frac{1}{20} \cdot ((2 - 7{,}52)^2 \cdot 3 + (6 - 7{,}52)^2 \cdot 4 + (8 - 7{,}52)^2 \cdot 6 + (10 - 7{,}52)^2 \cdot 8)$$
$$= \frac{151{,}2384}{20} \approx 7{,}2 \Rightarrow S = \sqrt{7{,}2} \approx 2{,}683.$$

Так как мода – это варианта, которой соответствует наибольшая частота, то $Mo = 10$.

Не сгруппированные данные образуют дискретный вариационный ряд, содержащий нечётное число вариант ($n = 21 = 2 \cdot 10 + 1$):

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
x_i	2	2	2	6	6	6	6	8	8	8	8	8	8	10	10	10	10	10	10	10	10

Значит, медиана равна $Me = x_{10} = 8$.

2. Записать в виде вариационного ряда выборку 20, 19, 12, 13, 16, 17, 17, 14, 16, 14, 13, 19, 18, 16, 14. Представить статистическое распре деление выборки. Построить полигон относительных частот для статистического ряда. Вычислить числовые характеристики выборки: выборочное среднее, «исправленную» и выборочную дисперсии, «исправленное» среднеквадратическое отклонение (СКО).

Решение. Объём выборки $n = 15$. Упорядочив элементы выборки по возрастанию, получим вариационный ряд:

12, 13, 13, 14, 14, 14, 16, 16, 16, 17, 17, 18, 19, 19, 20.

Статистическое распределение исходной выборки можно записать в виде следующей таблицы:

x_i	12	13	14	16	17	18	19	20	Сумма
n_i	1	2	3	3	2	1	2	1	15
p_i^*	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{3}{15}$	$\frac{2}{15}$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{1}{15}$	1

Полигон относительных частот изображён на рисунке 7.

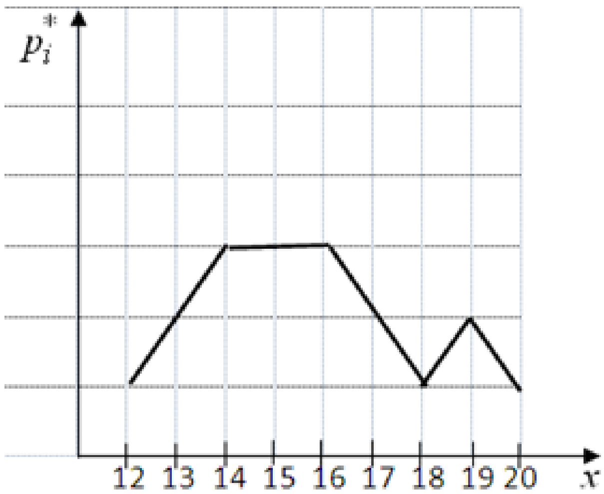


Рисунок 7

Находим выборочное среднее по формуле (2):

$\bar{x} = \frac{(12+13\cdot2+14\cdot3+16\cdot3+17\cdot2+18+19\cdot2+20)}{15} = \frac{238}{15} \approx 15{,}9.$

Для вычисления выборочной дисперсии используем формулу (5):

$$D_B = \frac{1}{15}\cdot\left((12-15{,}9)^2 + 2\cdot(13-15{,}9)^2 + 3\cdot(14-15{,}9)^2 + \cdots\right.$$
$$\left. + (20-15{,}9)^2\right) = \frac{85{,}75}{15} \approx 5{,}72.$$

«Исправленная» дисперсия и СКО:

$$S^2 = \frac{n}{n-1}\cdot D_B = \frac{15}{14}\cdot 5{,}72 \approx 6{,}13; \quad S = \sqrt{6{,}13} \approx 2{,}48.$$

3. Найти выборочное среднее, моду, медиану и выборочное СКО выборки объёмом $n = 70$, распределение которой задано следующей таблицей:

Интервалы	0 – 1,02	1,02 – 2,04	2,04 – 3,06	3,06 – 4,08	4,08 – 5,1	Сумма
Частота	4	28	19	12	7	140

Построить гистограмму и полигон частот.

Решение. Для построения гистограммы все частоты необходимо разделить на длину интервала, равную 1,02, и откладывать по оси ординат. По оси абсцисс отмечаются границы интервалов (рисунок 8).

Для построения полигона частот найдем середины интервалов и дополним исходную таблицу:

Интервалы $a_i - a_{i+1}$	0 – 1,02	1,02 – 2,04	2,04 – 3,06	3,06 – 4,08	4,08 – 5,1	Сумма
Середины интервалов a_i^*	0,51	1,53	2,55	3,57	4,59	
Частота m_i	4	28	19	12	7	140
Накопленная частота	4	32	51	63	70	

Ломаная линия (рисунок 8) будет соединять точки с координатами $\left(a_i^*; \frac{n_i}{h}\right)$.

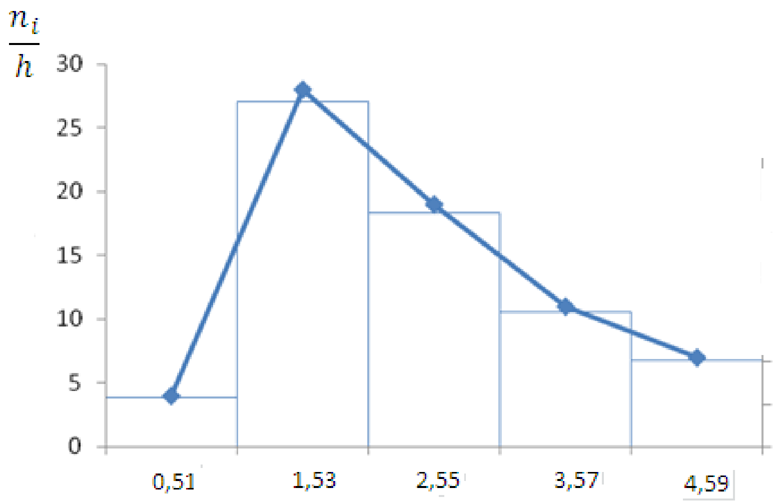


Рисунок 8

Для расчёта выборочного среднего и выборочного СКО составляем вариационный ряд, принимая в качестве вариант середины соответствующих интервалов:

a_i^*	0,51	1,53	2,55	3,57	4,59	**Σ**
m_i	4	28	19	12	7	**70**
$a_i^* \cdot m_i$	2,04	42,84	48,45	42,84	32,13	**168,3**
$(a_i^*)^2 \cdot m_i$	1,0404	65,5452	123,5475	152,9388	147,4767`	**490,5486**

Таким образом:

$$\bar{x} = \frac{168{,}3}{70} = 2{,}404; \quad \bar{x}^2 = \frac{490{,}5486}{70} = 7{,}008$$

$$D_B = \bar{x}^2 - (\bar{x})^2 = 7{,}008 - 2{,}404^2 = 1{,}229; \quad \sigma_B = \sqrt{D_B} = \sqrt{1{,}229} \approx 1{,}11.$$

Так как наибольшая частота $m_i = 28$ отвечает интервалу $1{,}02 - 2{,}04$, то $x_i = 1{,}02;\ m_{i-1} = 4;\ m_{i+1} = 19;\ h = 1{,}02$.

Мода (согласно формуле (3)) равна:

$$Mo = 1{,}02 + 1{,}02 \cdot \frac{28-4}{(28-4)+(28-19)} \approx 1{,}76.$$

Определим номер медианного интервала. Так как $32 < \frac{70}{2} < 51$, то номер медианного интервала равен 3, а сам интервал – $2{,}04 - 3{,}06$. Тогда, по формуле (4), получаем:

$$Me = 2{,}04 + 1{,}02 \cdot \frac{35-32}{19} = 2{,}202.$$

4. Дан статистический ряд признака X:

x_i	3,8	3,9	4	4,1	4,2	4,3	4,4	4,5	Σ
n_i	2	2	5	7	6	4	3	1	30

Найти начальные и центральные моменты первых четырёх порядков признака X, а также определить асимметрию и эксцесс.

Решение. Вычисления проводим по формулам (8) для $\tilde{\nu}_k$ и по формулам (10) для $\tilde{\mu}_k$.

Начальные моменты:

$$\tilde{\nu}_1 = \frac{1}{30}(2\cdot 3{,}9 + 2\cdot 4{,}1 + 5\cdot 4 + 7\cdot 4{,}1 + 6\cdot 4{,}2 + 4\cdot 4{,}3 + 3\cdot 4{,}4 + 1\cdot 4{,}5) = \frac{124{,}2}{30} = 4{,}14.$$

$$\tilde{\nu}_2 = \frac{1}{30}(2\cdot 3{,}9^2 + 2\cdot 4{,}1^2 + 5\cdot 4^2 + 7\cdot 4{,}1^2 + 6\cdot 4{,}2^2 + 4\cdot 4{,}3^2 + 3\cdot 4{,}4^2 + 1 \cdot 4{,}5^2) = \frac{515{,}1}{30} = 17{,}17.$$

$$\tilde{\nu}_3 = \frac{1}{30}(2\cdot 3{,}9^3 + 2\cdot 4{,}1^3 + 5\cdot 4^3 + 7\cdot 4{,}1^3 + 6\cdot 4{,}2^3 + 4\cdot 4{,}3^3 + 3\cdot 4{,}4^3 + 1 \cdot 4{,}5^3) = \frac{2140{,}1}{30} = 71{,}34.$$

$$\tilde{\nu}_4 = \frac{1}{30}(2\cdot 3{,}9^4 + 2\cdot 4{,}1^4 + 5\cdot 4^4 + 7\cdot 4{,}1^4 + 6\cdot 4{,}2^4 + 4\cdot 4{,}3^4 + 3\cdot 4{,}4^4 + 1 \cdot 4{,}5^4) = \frac{8906{,}8}{30} = 296{,}89.$$

Центральные моменты:

$\tilde{\mu}_1 = 0$

$\tilde{\mu}_2 = \tilde{\nu}_2 - \tilde{\nu}_1^2 = 17{,}17 - 4{,}14^2 = 0{,}0304;$

$\tilde{\mu}_3 = \tilde{\nu}_3 - 3\tilde{\nu}_1\tilde{\nu}_2 + 2\tilde{\nu}_1^3 = 71{,}34 - 3\cdot 4{,}14\cdot 17{,}17 + 2\cdot 4{,}14^3 = -70{,}95;$

$$\tilde{\mu}_4 = \nu_4 - 4\tilde{\nu}_1\tilde{\nu}_3 + 6\tilde{\nu}_2\tilde{\nu}_1^2 - 3\tilde{\nu}_1^4 = 296{,}89 - 4\cdot 4{,}14\cdot 71{,}34 + 6\cdot 17{,}17\cdot 4{,}14^2 - 3\cdot 4{,}14^4 = -11{,}34$$

Тогда, так как $\sigma_B = \sqrt{D_B} = \sqrt{\tilde{\mu}_2} = 0{,}174$, то

$$\tilde{A}_s = \frac{\tilde{\mu}_3}{\sigma_B^3} = \frac{-70{,}95}{0{,}174^3} = -13{,}5; \quad E_x = \frac{\tilde{\mu}_4}{\sigma_B^4} - 3 = \frac{-11{,}34}{0{,}174^4} - 3 = -2{,}15.$$

5. Предельная нагрузка для выборки из 50 стальных стержней характеризуется следующим рядом:

x_i	11	13	14	16	17
n_i	4	5	30	7	4

Считая распределение предельной нагрузки X нормальным, построить доверительные интервалы для оценки с надёжностью $\gamma = 0{,}99$ средней предельной нагрузки и СКО предельной нагрузки стальных стержней партии, из которой произведена выборка.

Решение. Вычислим выборочное среднее и исправленное СКО соответственно по формулам

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i n_i}{n} = \frac{11 \cdot 4 + 13 \cdot 5 + 14 \cdot 30 + 16 \cdot 7 + 17 \cdot 4}{50} = \frac{709}{50} = 14{,}18$$

$$S = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2 n_i}{n-1}} = \sqrt{\frac{(11 - 14{,}18)^2 \cdot 3 + \dots + (17 - 14{,}17)^2 \cdot 4}{49}} = \sqrt{\frac{103{,}38}{49}}$$
$$= 1{,}45$$

По таблице (см. приложение 3) найдём $t_\gamma = t(0{,}99; 50) = 2{,}679$. Точность оценки:

$$\delta = t_\gamma \cdot \frac{S}{\sqrt{n}} = 2{,}679 \cdot \frac{1{,}45}{7{,}071} = 0{,}55.$$

Доверительный интервал для средней предельной нагрузки найдём по формуле (13):

$$\bar{x} - t_\gamma \cdot \frac{S}{\sqrt{n}} = 14{,}18 - 0{,}55 = 13{,}63; \quad \bar{x} + t_\gamma \cdot \frac{S}{\sqrt{n}} = 14{,}18 + 0{,}55 = 14{,}73.$$

$$13{,}63 < a < 14{,}73..$$

Доверительный интервал для СКО предельной нагрузки будем искать по формуле $S(1-q) < \sigma < S(1+q)$, так как $q = q(0{,}99; 50) = 0{,}3$ (см. приложение 4):

$$1{,}45 \cdot (1 - 0{,}3) < \sigma < 1{,}45 \cdot (1 + 0{,}3) \Rightarrow 1{,}015 < \sigma < 1{,}885$$

6. В результате эксперимента получены данные, представленные в виде статистического ряда:

38 60 41 51 33 42 45 21 53 60 60 52 47 46 49 49 14 57 54 59 30 40 50 59 30 61 56 58 42 54 44 42 32 45 60 43 41 58 48 72 48 47 39 28 47 35 65 61 77 67.

Требуется:

1)Записать значения результатов эксперимента в виде вариационного ряда.

2)Представить данную выборку в виде интервального статистического ряда.

3)Найти числовые характеристики выборки: $\bar{x}$, D_B, S^2, S.

4)Определить доверительные интервалы неизвестного математического ожидания и неизвестного среднего квадратического отклонения. Предполагается, что генеральная совокупность имеет нормальное распределение. Доверительную вероятность принять равной 0,95.

Решение.

1) Расположим значения результатов эксперимента в порядке возрастания, т.е. записываем вариационный ряд:

14 21 28 30 30 32 33 35 38 39 40 41 41 42 42 42 43 44 45 45 46 47 47 47 48 48 49 49 50 51 52 53 54 54 56 57 58 58 59 59 60 60 60 60 61 61 65 67 72 77.

2) Объём выборки $n = 50$. Наибольшая варианта – 77, наименьшая – 14. Найдём длину интервала:

$$h = \frac{77 - 14}{1 + 3{,}322 \cdot lg50} = \frac{63}{6{,}644} \approx 9.48.$$

Выбираем длину интервала 9. Интервальный статистический ряд примет вид:

Границы интервалов	$[14; 23[$	$[23; 32[$	$[32; 41[$	$[41; 50[$	$[50; 59[$	$[59; 68[$	$[68; 77]$
m_i	2	3	6	17	10	10	2

3) Для вычисления числовых характеристик составляем вариационный ряд, принимая в качестве вариант середины соответствующих интервалов:

Границы интервалов	$[14; 23[$	$[23; 32[$	$[32; 41[$	$[41; 50[$	$[50; 59[$	$[59; 68[$	$[68; 77[$	$\sum$
Середины интервалов a_i^*	18,5	27,5	36,5	45,5	54,5	63,5	72,5	
m_i	2	3	6	17	10	10	2	**50**
$a_i^* \cdot m_i$	37	82,5	219	773,5	545	635	145	**2437**
$(a_i^*)^2 \cdot m_i$	684,5	2268,75	7993,5	35194,25	29702,5	40322,5	10512,5	**126678,5**

Таким образом:

$$\bar{x} = \frac{2437}{50} = 48{,}74; \;\; \bar{x^2} = \frac{126678{,}5}{50} = 2533{,}57.$$

$$D_B = \bar{x^2} - (\bar{x})^2 = 2533{,}57 - 48{,}74^2 = 158{,}9824;$$

$$S^2 = \frac{50}{49} \cdot 158{,}9824 = 161{,}2 \Rightarrow S = \sqrt{161{,}2} = 12{,}7.$$

4) Доверительный интервал для оценки математического ожидания нормально распределённой случайной величины найдём по формуле:

$$\bar{x} - t_\gamma \cdot \frac{S}{\sqrt{n}} < a < \bar{x} + t_\gamma \cdot \frac{S}{\sqrt{n}}.$$

Из приложения 3 для $\gamma = 0{,}95$ находим $t_\gamma = 2{,}009$. Далее, $t_\gamma \cdot \frac{S}{\sqrt{n}} = \frac{2{,}009 \cdot 12{,}7}{7{,}071} = 3{,}608$ и $45{,}132 < a < 52{,}348$, т.е. $a \in (45{,}132; 52{,}348)$.

Доверительный интервал для оценки σ нормального распределения по несмещённой оценке S определяется из неравенства $S(1-q) < \sigma < S(1+q)$, где величина $q = q(\gamma, n)$ определяется из таблицы (приложение 4).

Имеем $n = 50$; $S = 12{,}7$. При $\gamma = 0{,}95$ и $n = 50$ в таблице приложения находим $q(0{,}95; 50) = 0{,}21$.

Следовательно, $S(1-q) = 12{,}7 \cdot 0{,}79 = 10{,}033$; $S(1+q) = 12{,}7 \cdot 1{,}21 = 15{,}367$. Значит, $\sigma \in (10{,}033; 15{,}367)$.

7. Требуется при уровне значимости $\alpha = 0{,}05$ проверить по критерию согласия Пирсона гипотезу о нормальном распределении генеральной совокупности, если известны эмпирические частоты n_i и теоретические частоты n_i':

n_i	8	10	18	27	17	11	9
n_i'	5	15	16	25	20	12	7

Решение. Определим наблюдаемое значение критерия Пирсона по формуле

$$\chi^2 = \sum_{i=1}^{l} \frac{(n_i - n_i')^2}{n_i'} = \frac{(8-5)^2}{5} + \frac{(10-15)^2}{15} + \frac{(18-16)^2}{16} + \dots + \frac{(9-7)^2}{7} \approx 5$$

В таблице критических точек χ^2 (приложение 5) находим при уровне значимости $\alpha = 0{,}05$ значение $\chi^2_{kp}(0{,}05; 4) = 9{,}488$ (имеем $k = l - 3 = 7 - 3 = 4$ степени свободы). Значение $\chi^2 = 5 < \chi^2_{kp}$. Следовательно, выдвинутая гипотеза о нормальном распределении генеральной совокупности не отвергается.

8. Из генеральной совокупности извлечена выборка, представленная в виде ряда.

Требуется проверить, согласуются ли выборочные данные с гипотезой о нормальном распределении случайной величины Х с помощью критерия согласия Пирсона при уровне значимости $\alpha = 0{,}05$, разбив отрезок $[x_{min}; x_{max}]$ на l интервалов одинаковой длины. Величину l рассчитать по формуле Стерджеса $l = 1 + 3{,}322 \cdot \lg n$.

2,1	2,3	1,5	3,1	2,7	1,9	2,4	0,9	2,5	1,1
1,3	2,9	2,3	3,9	2,4	3,6	1,6	3,2	2,9	2
2,1	3,3	0,8	3,5	1,7	2,6	4,1	2,8	1,2	2,5
1,1	2,4	1,5	3,2	2,7	1,5	3,7	1,9	3,1	4
4,1	2,9	2	1,1	0,7	3,3	2,5	1,6	2,4	2,1
3,2	0,9	2,8	4,2	2,8	1,9	1,2	1,7	3,5	2
2,7	3,9	2,4	1,7	3,6	2,5	0,8	3,1	2,1	1,3

3,2	1,6	0,7	2,6	1,3	2	3,7	2,9	4	3,1
2,8	4,1	1,9	3,6	3,3	2,9	0,6	1,5	1,2	2,4
1,1	3,5	1,6	2,4	3,9	2,7	2,5	1,9	2,6	3,2

Решение.

Подсчитаем количество интервалов разбиения:

$l = 1 + 3{,}322 \cdot \lg n = 1 + 3{,}322 \cdot \lg 100 = 7{,}644 \approx 8.$

Из ряда видно, что $x_{min} = 0{,}6$; $x_{max} = 4{,}2$, поэтому $\frac{x_{max} - x_{min}}{8} = \frac{4{,}2 - 0{,}6}{8} \approx 0{,}5$.

Границы интервалов будут:

$x_0 = 0{,}6$; $x_1 = 0{,}6 + 0{,}5 = 1{,}1$; $x_2 = 1{,}1 + 0{,}5 = 1{,}6$; $x_3 = 1{,}6 + 0{,}5 = 2{,}1$; $x_4 = 2{,}1 + 0{,}5 = 2{,}6$; $x_5 = 2{,}6 + 0{,}5 = 3{,}1$; $x_6 = 3{,}1 + 0{,}5 = 3{,}6$; $x_7 = 3{,}1 + 0{,}5 = 4{,}1$; $x_8 = 4{,}1 + 0{,}5 = 4{,}6$.

Частота n_i интервала $[x_i; x_{i+1}[$ $(i = \overline{0{,}7})$ подсчитывается с помощью ряда как число наблюдений, попавших в интервал. Так в первый $(i = 0)$ интервал $[0{,}6;\ 1{,}1[$ попало 7 значений, во второй $[1{,}1;\ 1{,}6[$ – 14 значений. Сведём полученные данные в таблицу:

$x_i - x_{i+1}$	0,6 – 1,1	1,1 – 1,6	1,6 – 2,1	2,1 - 2,6	2,6 - 3,1	3,1 – 3,6	3,6 – 4,1	4,1 – 4,6
x_i^*	0,85	1,35	1,85	2,35	2,85	3,35	3,85	4,35
n_i	7	14	16	18	16	15	10	4

Объем выборки равен $n = \sum n_i = 100$. Выборочное среднее и дисперсия определяются по формулам:

$\bar{x} = \frac{\sum x_i^* n_i}{n} = \frac{1}{100}(0{,}85 \cdot 7 + 1{,}35 \cdot 14 + 1{,}85 \cdot 16 + \cdots + 4{,}35 \cdot 4) = \frac{248{,}5}{100} = 2{,}485$

$\sigma_B = \sqrt{\frac{\sum (x_i^* - \bar{x})^2 n_i}{n}} = \sqrt{\frac{1}{100}((0{,}85 - 2{,}485)^2 \cdot 7 + (1{,}35 - 2{,}485)^2 \cdot 14 + \cdots + (4{,}35 - 2{,}485)^2 \cdot 4)} = \sqrt{0{,}89} \approx 0{,}95.$

Найдём теоретические вероятности p_i по формуле

$$p_i = P(z_i < z < z_{i+1}) = \Phi\left(\frac{x_{i+1} - \bar{x}}{\sigma_B}\right) - \Phi\left(\frac{x_i - \bar{x}}{\sigma_B}\right),$$

где $\Phi(x)$ – функция Лапласа, значения которой даются в приложении 3. Результаты вычислений сведём в таблицу:

i	x_i	x_{i+1}	$x_i - \bar{x}$	$x_{i+1} - \bar{x}$	z_i	z_{i+1}	$\Phi(z_i)$	$\Phi(z_{i+1})$	p_i	$n_i' = n \cdot p_i$
1	0,6	1,1	–	-1,385	–	-	-0,5	-0,4279	0,0721	7,21

						1,46				
2	1,1	1,6	-1,385	-0,885	-1,46	-0,94	-0,4279	-0,3264	0,1015	10,15
3	1,6	2,1	-0,885	-0,385	-0,94	-0,41	-0,3264	-0,1591	0,1673	16,73
4	2,1	2,6	-0,385	0,115	-0,41	0,12	-0,1591	0,0478	0,2069	20,69
5	2,6	3,1	0,115	0,615	0,12	0,65	0,0478	0,2422	0,1944	19,44
6	3,1	3,6	0,615	1,115	0,65	1,18	0,2422	0,381	0,1388	13,88
7	3,6	4,1	1,115	1,615	1,18	1,71	0,381	0,4564	0,0754	7,54
8	4,1	4,6	1,615	–	1,71	–	0,4564	0,5	0,0436	4,36

Вычислим наблюдаемое значение критерия Пирсона. Для этого составим следующую расчетную таблицу:

i	n_i	n_i'	$n_i - n_i'$	$(n_i - n_i')^2$	$\dfrac{(n_i - n_i')^2}{n_i'}$
1	7	7,21	-0,210	0,044	0,006
2	14	10,15	3,850	14,823	1,460
3	16	16,73	-0,730	0,533	0,032
4	18	20,69	-2,690	7,236	0,350
5	16	19,44	-3,440	11,834	0,609
6	15	13,88	1,120	1,254	0,090
7	10	7,54	2,460	6,052	0,803
8	4	4,36	-0,360	0,130	0,030
$\sum$	**100**	**100**			**3,379**

По таблице критических точек распределения χ^2, уровню значимости $\alpha = 0{,}05$ и числу степеней свободы $k = 8 - 2 - 1 = 5$ находим $\chi^2_{kp} = 11{,}1$.. Так как $\chi^2 = 3{,}379 < \chi^2_{kp}(0{,}05; 5)$, то нет оснований отвергнуть гипотезу о нормальном распределении генеральной совокупности.

9. По заданной таблице зависимости признаков X и Y

X	0	1.7	4.7	7.5	8.5
Y	-3.2	-2.7	-1.0	0.2	1.8

вычислить выборочный коэффициент корреляции и остаточную дисперсию. Записать уравнения прямой регрессии Y на X. Построить корреляционное поле и линию регрессии на корреляционном поле.

Решение. Вычислим основные выборочные характеристики:

Выборочные средние:

$$\bar{x} = \frac{1}{5}(0 + 1{,}7 + 4{,}7 + 7{,}5 + 8{,}5) = \frac{22{,}4}{5} = 4{,}48;$$

$$\bar{y} = \frac{1}{5}(-3{,}2 - 2{,}7 - 1 + 0{,}2 + 1{,}8) = \frac{-4{,}9}{5} = -0{,}98.$$

Найдем оценки для средних квадратичных отклонений и корреляционного момента, для чего составим следующую вспомогательную таблицу:

X	0	1,7	4,7	7,5	8,5	$\sum$
Y	-3,2	-2,7	-1	0,2	1,8	
$(x_i - \bar{x})^2$	20,07	7,7284	0,0484	9,1204	16,16	**53,128**
$(y_i - \bar{y})^2$	4,9284	2,9584	0,0004	1,3924	7,7284	**17,008**
$(x_i - \bar{x})(y_i - \bar{y})$	9,9456	4,7816	-0,004	3,5636	11,176	**29,462**

$$s_x = \sqrt{\frac{1}{n-1}\sum_i (x_i - \bar{x})^2} = \sqrt{\frac{53{,}128}{4}} \approx 3{,}6444.$$

$$s_y = \sqrt{\frac{1}{n-1}\sum_i (y_i - \bar{y})^2} = \sqrt{\frac{17{,}008}{4}} \approx 2{,}062.$$

$$s_{xy} = \frac{1}{n-1}\sum_i (x_i - \bar{x})(y_i - \bar{y}) = \frac{29{,}462}{4} = 7{,}3655.$$

Согласно формуле (15):

$$r_B = \frac{s_{xy}}{s_x \cdot s_y} = \frac{7{,}3655}{3{,}6444 \cdot 2{,}062} \approx 0{,}98.$$

Найдем методом наименьших квадратов эмпирическую формулу вида $\bar{y}_x = a + bx$. Составим систему нормальных уравнений (17) для определения параметров линейной регрессии.

Так как

$\bar{x} = 4{,}48;\ \bar{y} = -0{,}98;\ \overline{x^2} = \frac{1}{5}(0 + 1{,}7^2 + 4{,}7^2 + 7{,}5^2 + 8{,}5^2) = \frac{153{,}48}{5} = 30{,}696;$

$$\overline{xy} = \frac{1}{5}(0 - 1{,}7 \cdot 2{,}7 - 4{,}7 + 7{,}5 \cdot 0{,}2 + 8{,}5 \cdot 1{,}8) = \frac{7{,}51}{5} = 1{,}502,$$

то:

$$\begin{cases} a + 4{,}48b = -0{,}98 \\ 4{,}48a + 30{,}696b = 1{,}502 \end{cases} \Rightarrow b = \frac{1{,}502 + 4{,}48 \cdot 0{,}98}{30{,}696 - 4{,}48^2} = 0{,}5545;$$

$$a = -0{,}98 - 0{,}5545 \cdot 4{,}48 = -3{,}464.$$

Уравнение регрессии Y на X имеет вид:

$$\bar{y}_x = -3{,}464 + 0{,}5545x.$$

Остаточная дисперсия:

$$D_{\text{ост}} = s_y^2(1 - r_B^2) = 2{,}062^2 \cdot (1 - 0{,}98^2) = 0{,}168.$$

Корреляционное поле и линия регрессии на корреляционном поле изображены на рисунке 9.

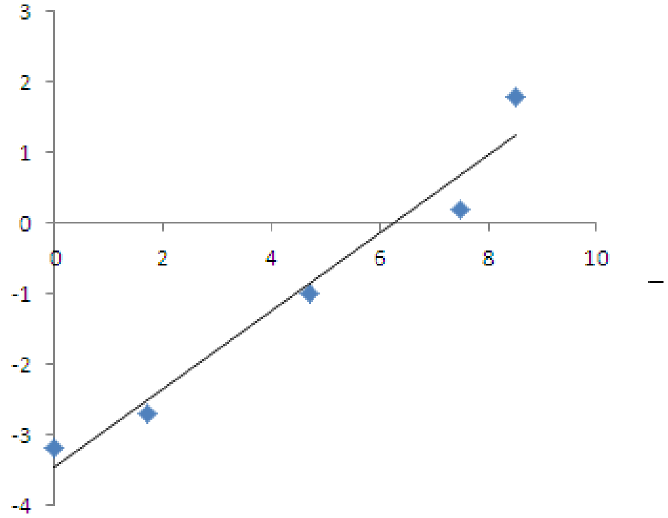


Рисунок 9

10. Таблица значений признака Y при данных значениях признака X имеет вид:

Y	X				
	5	10	15	20	n_y
10	2	-	-	-	2
20	5	4	1	-	10
30	3	8	6	3	20
40	-	3	6	6	15
50	-	-	2	1	3
n_x	10	15	15	10	$n = 50$

Построить корреляционное поле. Найти выборочный коэффициент корреляции, оценить его значимость. Записать уравнения прямой линии регрессии Y на X.

Решение. Корреляционное поле данной двумерной выборки приведено на рисунке 10.

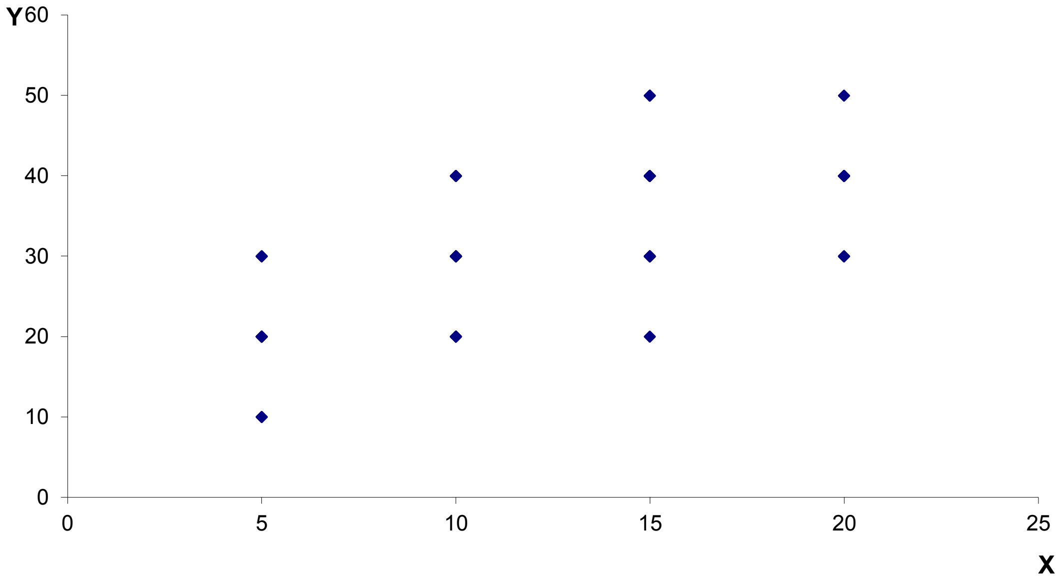


Рисунок 10

По виду поля корреляции можно судить о том, что между величинами существует зависимость.

Для вычисления выборочных числовых характеристик составляем следующую расчётную таблицу:

Y	**X**					$n_{yi}y_j$	$\sum_i n_{ij}x_i$	$y_j^2 n_{yj}$	$y_j \sum_i n_{ij}x_i$
	5	**10**	**15**	**20**	n_y				
10	2	-	-	-	2	20	10	200	100
20	5	4	1	-	10	200	80	4000	1600
30	3	8	6	3	20	600	245	18000	7350
40	-	3	6	6	15	600	240	24000	9600
50	-	-	2	1	3	150	50	7500	2500
n_x	10	15	15	10	$\boldsymbol{n=50}$	**1570**	**625**	**53700**	**21150**
$n_{xi}x_i$	50	150	225	200	**625**				
$\sum_j n_{ij}y_j$	210	440	540	380	**1570**				
$n_{xi}x_i^2$	250	1500	3375	4000	**9125**				
$x_i \sum_j n_{ij}y_j$	1050	4400	8100	7600	**21150**				

Замечание. Строка $\sum_j n_{ij} y_j$ получается следующим образом:
$210 = 10 \cdot 2 + 20 \cdot 5 + 30 \cdot 3$; $440 = 20 \cdot 4 + 30 \cdot 8 + 40 \cdot 3$; … и т. д;

Столбец $\sum_i n_{ij} x_i$: $10 = 5 \cdot 2; 30 = 5 \cdot 5 + 10 \cdot 4 + 15 \cdot 1$; …

Вычислим выборочные средние $\bar{x}$ и $\bar{y}$ $\left(i = \overline{1,4};\ j = \overline{1,5}\right)$:
$\bar{x} = \frac{1}{n} \sum_i n_{xi} x_i = \frac{625}{50} = 12,5$; $\bar{y} = \frac{1}{n} \sum_j n_{xj} x_j = \frac{1570}{50} = 31,4$.
«Исправленные» дисперсии находим по формулам:
$s_x^2 = \frac{1}{n-1}\left(\sum_i n_{xi} x_i^2 - \frac{1}{n}(\sum_i n_{xi}\, x_i)^2\right) = \frac{1}{49}\left(9125 - \frac{1}{50} \cdot 625^2\right) = 26,786$.
$s_y^2 = \frac{1}{n-1}\left(\sum_j n_{yj} y_j^2 - \frac{1}{n}\left(\sum_j n_{yj}\, y_j\right)^2\right) = \frac{1}{49}\left(53700 - \frac{1}{50} \cdot 1570^2\right) = 89,837$.
Оценку корреляционного момента вычисляем по формуле:
$s_{xy} = \frac{1}{n-1}\left(\sum_i \sum_j n_{ij} x_i y_j - \frac{1}{n}(\sum_i n_{xi}\, x_i) \cdot \left(\sum_j n_{yj}\, y_j\right)\right) = \frac{1}{49}\Big(21150 -$
$\frac{1}{50} 625 \cdot 1570\Big) = 31,1$.

Рассчитав все нужные величины, можно вычислить выборочный коэффициент корреляции:

$$r_B = \frac{31,1}{\sqrt{26,786} \cdot \sqrt{89,837}} = \frac{31,1}{5,176 \cdot 9,478} \approx 0,63.$$

Для оценки значимости выборочного коэффициента корреляции вычислим наблюдаемое значение критерия, воспользовавшись формулой (16):

$$T_{\text{набл}} = r_B \cdot \sqrt{\frac{n-2}{1-r_B^2}} = 0,63 \cdot \sqrt{\frac{48}{1-0,63^2}} = 5,62.$$

Затем по таблице критических точек распределения Стьюдента, по заданному уровню значимости $\alpha = 0,05$ и числу степеней свободы $k = n - 2 = 48$ найдем критическую точку t_{kp} для двухсторонней критической области:

$$t_{kp}(0,05; 48) = 2,01.$$

Сравнивая $T_{\text{набл}}$ и t_{kp}, получим, что $|T_{\text{набл}}| > t_{kp}$, следовательно, величины X, Y коррелированы.

Оценкой теоретической линии регрессии является эмпирическая линия регрессии, уравнение которой имеет вид $\bar{y}_x - \bar{y} = b(x - \bar{x})$, где $b = r_B \frac{s_y}{s_x}$.

Тогда:
$\bar{y}_x - 31,4 = 0,63 \cdot \frac{9,478}{5,176} \cdot (x - 12,5)$ или $\bar{y}_x = 1$.

Задачи для самостоятельного решения

1. Дана выборка:

x_i	1,5	2	2,5	3	3,5

n_i	10	15	30	20	25

Найти эмпирическую функцию распределения, построить её график. Построить полигон относительных частот выборки.

2. Записать в виде вариационного и статистического рядов выборку **7,** 18, 16, 16, 17, 18, 19, 17, 15, 17, 19, 18, 16, 16, 16, 18, 18. Определить размах выборки. Построить полигон частот и эмпирическую функцию распределения.

3. Построить гистограмму относительных частот по заданному распределению выборки

Номер интервала	Частичный интервал	Число наблюдений, **попавших** в интервал
1	2 – 5	6
2	5 – 8	10
3	8 – 11	4
4	11 – 14	5

4. Дана статистическая совокупность чисел. Написать вариационный ряд, построить интервальную обработку. Построить полигон распределения, гистограмму, эмпирическую функцию распределения:

а)

1,31	0,48	0,76	1,71	1,2	0,54	0,2	0,67
0,62	0,15	0,05	0,78	0,24	0,29	1,47	1,11
0,67	0,99	1,02	0,51	0,65	1,56	0,16	0,49

б)

31	32	29	31	38	32	15	41	46	17
23	24	45	43	41	32	31	38	23	33
45	43	31	32	12	26	28	39	29	26
12	31	25	24	41	40	51	16	19	27
27	23	22	43	42	40	32	31	29	21
19	17	19	21						

5. Найти выборочное среднее, выборочную и «исправленную» дисперсии, а также «исправленное» СКО по заданному распределению выборки. Построить полигон частот и эмпирическую функцию распределения.

x_i	2	3	5	6	8	9	10
n_i	4	10	21	30	20	10	5

6. Инженер по контролю качества продукции обнаружил в 10 партиях электроламп, произведенных заводом, следующее число бракованных изделий: **5, 3, 7, 1, 0, 6, 3, 4, 5, 2** . Найти среднее число (выборочное среднее) и стандартное отклонение («исправленое» среднее квадратическое отклонение) бракованных ламп. Вычислить медиану и моду.

7. Найти асимметрию и эксцесс эмпирического распределения:

x_i	10,6	10,8	11	11,2	11,4	11,6	11,8
n_i	5	10	17	30	20	12	6

8. Дан статистический ряд признака X:

x_i	2,6	3	3,4	3,8	4,2
n_i	8	20	45	15	12

Найти начальные и центральные моменты первых четырёх порядков признака, а также определить асимметрию и эксцесс.

9. Найти точечные оценки параметров a и σ нормального распределения, исходя из данной выборки:

x_i	0,3	0,5	0,7	0,9	1,1	1,3	1,5	1,7	1,9	2,1	2,3
n_i	6	9	26	25	30	26	21	24	20	8	7

10. Найти выборочное среднее и «исправленную» дисперсию для интервального статистического ряда. Построить гистограмму относительных частот.

Границы интервалов	$[0; 0{,}02[$	$[0{,}02; 0{,}04[$	$[0{,}04; 0{,}06[$	$[0{,}06; 0{,}08[$
$\boldsymbol{m_i}$	4	8	5	14

11. Задан интервальный вариационный ряд (в первой строке – границы интервалов, во второй строке – число точек, попавших в данный интервал):

$x_i - x_{i+1}$	0,4 - 0,8	0,8 – 1,2	1,2 – 1,6	1,6 – 2	2 – 2,4	2,4 – 2,8	2,8-3,2	3,2 – 3,6
m_i	2	5	8	14	16	10	8	7

Построить полигон и гистограмму распределения. Вычислить эмпирические характеристики: математическое ожидание, выборочную и исправленную дисперсии, моду и медиану. Построить доверительный интервал с доверительной вероятностью $\gamma = 0{,}95$ для неизвестного математического ожидания случайной величины X.

12. При исследовании эффективности работы системы массового обслуживания были зафиксированы интервалы времени обслуживания 60 заявок:

0,5 0,6 1,4 0,8 1,0 1,8 0,2 0,4 0,1 0,3 1,1 0,9 0,7 0,2 1,2
0,1 0,6 0,4 0,8 1,2 0,3 1,7 0,2 1,6 0,5 0,2 0,1 1,5 1,0 0,9
1,3 0,4 1,6 0,3 0,1 0,6 1,5 0,1 0,5 0,8 1,1 0,7 1,1 0,6 0,5
0,7 0,4 1,4 0,6 0,5 1,3 0,3 1,2 0,2 1,0 0,1 0,8 0,4 0,6 0,1

Построить интервальный статистический ряд. Вычислить эмпирические характеристики времени обслуживания заявки: математическое ожидание, дисперсию, асимметрию и эксцесс.

13. Ниже приведены результаты измерения роста (в см) случайно отобранных 100 студентов:

Рост	154-158	158-162	162-16	166-170	170-174	174-178	178-182
Число студентов	10	14	26	28	12	8	2

Найти выборочное среднее, «исправленное» СКО, выборочную моду и медиану.

14. Из генеральной совокупности извлечена выборка

x_i	0,2	0,4	0,8	0,9	1	1,1	1,2	1,3	1,5	1,8
n_i	5	7	12	15	20	18	13	5	3	2

Оценить с надежностью 0,95 математическое ожидание a нормально распределенного признака генеральной совокупности по выборочному среднему при помощи доверительного интервала. Построить доверительный интервал для дисперсии.

15. Среднее время сборки изделия составляло 90 минут. Инженер изобрел новый метод сборки этого изделия, и продолжительность сборки 10 изделий новым способом составила 79; 74; 112; 95; 83; 96; 77; 84; 70; 90 (мин.). Построить доверительный интервал для нового среднего времени сборки с надежностью 95%.

16. В ОТК были измерены диаметры 300 валиков из партии, изготовленной одним станком-автоматом. Отклонения измеренных диаметров от номинала (в нм) даны в таблице.

Границы отклонений	**Середина интервала**	**Число валиков**	**Границы отклонений**	**Середина интервала**	**Число валиков**
–30…–25	-27,5	3	0…5	2,5	55
–25…–	-22,5	8	5…10	7,5	30

20					
–20…–15	-17,5	15	10…15	12,5	25
–15…–10	-12,5	35	15…20	17,5	14
–10…–5	-7,5	40	20…25	22,5	8
–5…0	-2,5	60	25…30	27,5	7

Требуется:

1)Найти оценки математического ожидания, дисперсии и среднего квадратичного отклонения случайной величины X – отклонения диаметра валика от номинала.

2)Определить доверительные интервалы неизвестного математического ожидания и неизвестного среднего квадратического отклонения случайной величины X. Предполагается, что случайная величина X имеет нормальное распределение. Доверительную вероятность принять равной 0,95.

17. Используя критерий Пирсона при уровне значимости $\alpha = 0{,}01$ установить, согласуется ли гипотеза о выборе закона распределения случайной величины X со случайной выборкой, если её эмпирические и теоретические частоты заданы таблицей:

а)

n_i	8	16	40	72	36	18	10
n_i'	6	18	36	76	39	18	7

б)

n_i	6	8	13	15	20	16	10	7	5
n_i'	5	9	14	16	18	16	9	6	7

18. В итоге многократных измерений некоторой физической величины одним прибором получены результаты, представленные в таблице. Проверить, согласуются ли выборочные данные с гипотезой о нормальном распределении случайной величины X графически и с помощью критерия согласия Пирсона при уровне значимости $\alpha = 0{,}05$.

а)

19,2	20,0	18,2	21,6	20,3	18,9	16,5	17,2	18,7	20,5
21,3	12,9	19,5	21,3	16,7	21,2	17,5	17,5	15,8	16,7
17,6	19,6	15,8	18,5	15,4	12,2	17,1	12,6	23,6	17,8
15,5	16,9	16,1	15,5	16,6	19,6	19,9	16,9	16,0	18,2
19,8	15,0	19,0	19,2	17,0	16,3	14,9	16,8	18,8	16,5

б)

1,03	1,1	0,95	0,75	0,58	0,59	0,56	0,69	0,71	0,65
1,02	0,66	0,68	0,57	0,54	0,7	0,72	0,82	0,83	0,91
0,65	0,95	0,85	0,55	0,74	0,76	0,77	0,82	0,66	0,69
0,99	0,88	1,02	1,05	0,81	0,79	0,72	0,61	0,58	0,59
0,91	0,81	0,71	0,73	0,84	0,69	0,82	0,9	0,8	0,77

19. По заданной таблице зависимости признаков X и Y вычислить выборочный коэффициент корреляции и остаточную дисперсию. Записать уравнения прямой регрессии Y на X. Построить корреляционное поле и линию регрессии на корреляционном поле.

X	2.6	5.4	4.0	0.7	5.8
Y	-1.2	-0.1	-1.0	-3.2	0.4

б)**20.** Для исследования зависимости случайных величин X и Y получены статистические данные, представленные в корреляционной таблице Требуется:

а) построить корреляционное поле, сделать вывод;

б) определить выборочный коэффициент корреляции;

в) при уровне значимости $\alpha = 0{,}05$ проверить нулевую гипотезу о равенстве генерального коэффициента корреляции нулю при конкурирующей гипотезе $H_1\colon r_\Gamma \neq 0$;

г) найти уравнение прямой регрессии Y на X;

д) построить линию регрессии на корреляционном поле.

1)

X	Y					n_x
	3	6,5	10	13,5	17	
0	15	5				20
0,5	7	10	3			20
1		4	8	9	2	23
1,5				6	1	7
n_y	22	19	11	15	3	$n = 70$

Ответы

5.n=100**,**

$$\overline{x_в} = \frac{1}{n}\sum x_i n_i = \frac{1}{100}(2\cdot 4 + 3\cdot 10 + 5\cdot 21 + 6\cdot 30 + 8\cdot 20 + 9\cdot 10 + 10\cdot 5 = 6.23.$$

$$D_в = \frac{1}{n}\sum (x_i - \overline{x_в})^2 n_i = \frac{1}{n}\sum x_i^2 n_i - .\overline{x_в}^2 =$$

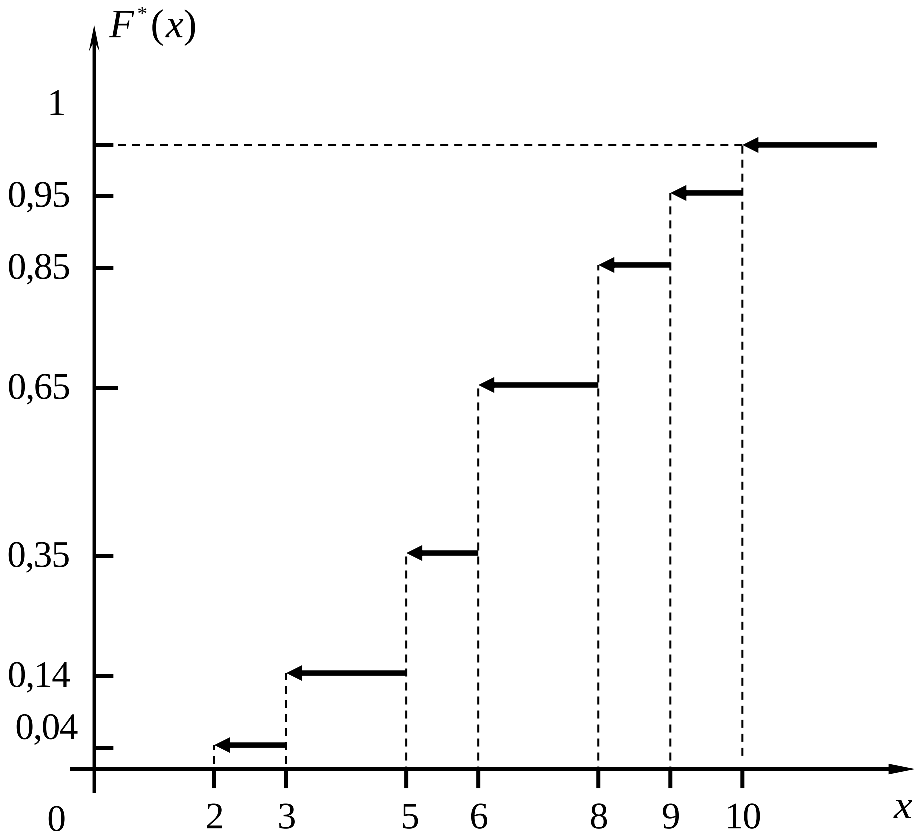
$$= \frac{1}{100}(2^2\cdot 4 + 3^2\cdot 10 + 5^2\cdot 21 + 6^2\cdot 30 + 8^2\cdot 20 + 9^2\cdot 10 + 10^2\cdot 5) - (6.23)^2 \approx 4.20.$$

$$S^2=\frac{1}{n-1}\sum(x_i-\overline{x}_в)^2n_i=\frac{n}{n-1}D_в=\frac{100}{99}\cdot D_в\approx 4.24.$$

$$S=\sqrt{S^2}=\sqrt{4.24}\approx 2.06.$$

$$F^*(x)=\begin{cases}0, & x\le 2\\ 0.04, & 2<x\le 3\\ 0.14, & 3<x\le 5\\ 0.35, & 5<x\le 6\\ 0.65, & 6<x\le 8\\ 0.85, & 8<x\le 9\\ 0.95, & 9<x\le 10\\ 1, & 10>x.\end{cases}$$

График функции $F^*(x)$:



7. $\widetilde{A}_s=-0{,}063$; $\widetilde{E}_x=-0{,}4358$.
8. 3,142; 11,8256; 41,61424; 148,6083. 0; 0,183856; 0,01066; 0,08999.
$\tilde{A}_s=0{,}13522$; $\widetilde{E}_x=-0{,}3378$.
9. $\overline{x}=1{,}2723$; $S^2=0{,}24952$. **10.** $\overline{x}=0{,}049$; $S^2=0{,}00048$.
11. $\overline{x}=2{,}1657$; $D_B=0{,}52003$; $S^2=0{,}52756$; $Mo=2{,}1$; $Me=2{,}15$.
(0,199243; 2,33897).

4. ПРОВЕРОЧНЫЕ ТЕСТЫ

4.1 Проверочный тест по разделу «Теория вероятностей»

Условие задачи	Варианты ответов
1.В лифт 7-этажного корпуса общежития на первом этаже вошли 2 студента. Каждый из них с одинаковой вероятностью выйдет на любом из этажей общежития, начиная со второго. Найти вероятность того, что все студенты выйдут на четвертом этаже.	1)1/49 2)1/36 3)1/64 4)1/7 5)2/7
2. Издание трудов студенческой конференции состоит из 4 томов. Книги расположены на полке в произвольном порядке. Какова вероятность, что номера томов идут подряд по возрастанию?	1)1/4 2)3/4 3)1/2 4)0 5)1/24
3. В группе выступающих в мероприятии «Весна БНТУ» 5 девушек и 6 юношей. Из группы случайным образом выделены 3 выступающих. Найти вероятность того, что все выбранные выступающие окажутся девушками.	1)2/33 2)5/6 3)3/11 4)5/11 5)3/5
4. Найти вероятность $P(y \geq x^2)$, где *x, y* – числа из отрезка [0, 1].	1)1 2)0 3)1/3 4)2/3 5)1/2
5. Студент знает 7 вопросов из 20 по дисциплине «Математика». Преподаватель задает три вопроса. Какова вероятность, что студент ответит хотя бы на 2 вопроса?	1)7/20 2)0,7 3)1/20 4)0,27 5)1/2
6. Во время производственной практики студенты первого курса изготавливают детали на трех автоматических станках. Известно, что 10% деталей производится первым станком, 25% -- вторым, остальные третьим. Вероятность изготовления детали, отвечающей стандарту, на первом станке равна 0,7; на втором – 0,988, на третьем 0,99. Изготовленные в течение дня на трех станках нерассортированные детали оцениваются преподавателем. Определить вероятность того, что взятая наугад деталь не соответствует стандарту.	1)0,682 2)0,683 3)0,6 4)0,654 5)0,632
7. В студенческой группе *10%* посещают тренажерный зал. Найти вероятность того, что среди наугад выбранных 8 студентов окажется 3 посещающих тренажерный зал.	1)0,33 2)0,3 3)0,1 4)0,134

	5)0,033
8. Пусть случайная величина X – равномерно распределена на отрезке [0,10]. Найти вероятность попадания СВ X в интервал (2,7).	1)1/10 2)2/7 3)2/3 4)1/7 5)0,5
9. Вероятность того, что купленный товар со скидкой равна 0,62. Составить ряд распределения для числа товаров со скидкой из общего числа 2 купленных товаров. Найти математическое ожидание рассмотренной случайной величины.	1)0,55 2)0,558 3)0,855 4)0,8 5)0,876
10. Ошибка взвешивания – случайная величина, распределенная по нормальному закону с математическим ожиданием, равным 0, и среднеквадратическим отклонением, равным 1 грамма. Найти вероятность того, что взвешивание проведено с ошибкой, не превышающей по модулю 0,2 грамма.	1)0,4578 2)0,1543 3)0,1566 4)0,1586 5)0,1598

4.1.1 Ответы к тесту по теории вероятностей

№	1	2	3	4	5	6	7	8	9	10
	2	5	1	4	4	2	5	5	3	4

4.2 Проверочный тест по разделу «Элементы математической статистики»

1	Предметом математической статистики является изучение …	1.Случайных величин по результатам наблюдений. 2. Случайных явлений. 3.Совокупностей. 4. Числовых характеристик.
2	Совокупность объектов, из которых производится выборка, называется…	1. Генеральной 2. Средней 3. Вероятной 4. Выборочной 5. Массовой
3	Выберите номер неправильного ответа. Генеральные совокупности могут быть…	1. Конечными 2. Бесконечными 3. Интервальными 4. Счётными
4	Часть отобранных объектов из гене-	1. Генеральной выборкой

	ральной совокупности называется…	2. Выборочной совокупностью 3. Репрезентативной совокупностью 4. Вариантами
5	Для того чтобы по выборке можно было судить о случайной величине, выборка должна быть…	1. Бесповторной 2. Повторной 3. Репрезентативной 4. Безвозвратной
6	Различные значения признака (случайной величины X) называются…	1. Частотами 2. Выборкой 3. Вариантами 4. Частостями
7	Числа, показывающие, сколько раз встречаются варианты из данного интервала, называются….	1. Частостями 2. Выборкой 3. Вариантами 4. Частотами
8	Отношение частоты данного варианта к общей сумме частот всех вариантов называется ….	1. Частостью 2. Группой 3. Вариантами 4. Частотой
9	3, 3, 1, 2, 5, 4, 2, 2, 4, 0, 2, 3, 2, 0, 2 – выборка. Частость варианты 2 составляет…	1. 5 2. $\frac{1}{3}$ 3. $\frac{1}{5}$ 4. 3
10	Полигон служит для изображения…	1. Дискретного ряда 2. Гистограммы 3. Интервального ряда 4. Выборочной функции
11	Ступенчатая фигура из прямоугольников с основаниями, равными интервалам значения признака $x_i - x_{i+1} (i = \overline{1, n})$ и высотами, равными частотам (частостям) интервалов, носит название…..	1. Полигона 2. Гистограммы 3. Кумулянты 4. Выборочной функции
12	Эмпирической функцией распределения $F^*(x)$ называется относительная частота того, что признак (случайная величина X) примет значение….	1. меньше заданного x 2. больше заданного x 3. равное заданному x
13	Статистическая оценка, математическое ожидание которой равно оцениваемому параметру при любом объеме выборки, называется…	1. Несмещенной 2. Состоятельной 3. Эффективной 4. Прямой

		5. Обратной
14	Статистическая оценка, которая (при заданном объеме выборки) имеет наименьшую возможную дисперсию, называется	1. Несмещенной 2. Состоятельной 3. Эффективной 4. Прямой 5. Обратной
15	Что является оценкой математического ожидания?	1. Выборочная средняя 2. Выборочная дисперсия 3. Исправленная дисперсия 4. Относительная частота
16	Проведено 5 измерений (без систематических ошибок) некоторой случайной величины: 5, 7, 8, 9, 10. Тогда несмещенная оценка математического ожидания равна…	1. 7 2. 7,2 3. 7,5 4. 7,8
17	Вариант, которому соответствует наибольшая частота вариационного ряда, называется…	1. медианой 2. модой 3. дисперсией 4. вариантом
18	Что является несмещённой оценкой генеральной дисперсии?	1. Выборочная средняя 2. Выборочная дисперсия 3. Исправленная дисперсия 4. Относительная частота
19	Мода вариационного ряда 21; 21; 23; 23; 23; 23; 24; 25; 27; 29; 29; 30 равна...	1. 27 2. 21 3. 29 4. 23
20	Медиана вариационного ряда 22; 23; 24; 25; 26; 28; 28; 28; 31; 32 равна	1. 22 2. 24 3. 27 4. 28
21	В результате измерений некоторой физической величины одним прибором (без систематических ошибок) получены следующие результаты (в мм): 14, 16, 17. Тогда несмещенная оценка дис-	1. 4 2. 3 3. 5 4. 2

<table>
<tr><td></td><td>персии измерений равна…</td><td></td></tr>
<tr><td>22</td><td>Из генеральной совокупности извлечена выборка объёма $n = 10$:
<table><tr><td>x_i</td><td>4</td><td>5</td><td>8</td></tr><tr><td>n_i</td><td>5</td><td>1</td><td>4</td></tr></table>Тогда выборочная дисперсия равна ...</td><td>1. 3,61
2. 3,49
3. 5,7
4. 4,1</td></tr>
<tr><td>23</td><td>Чему равна сумма доверительной вероятности и уровня значимости $\gamma+\alpha$?</td><td>1. 0
2. неотрицательному числу
3. какому-то числу от 0 до 1
4. 1</td></tr>
<tr><td>24</td><td>Симметричный ли интервал строится при оценивании генеральной средней?</td><td>1. нет
2. зависит от изучаемого явления
3. да</td></tr>
<tr><td>25</td><td>При построении доверительного интервала для генеральной дисперсии при малых объёмах выборки используют</td><td>1. распределение Пирсона
2. нормальный закон распределения
3. распределение Фишера-Снедекора
4. распределение Стьюдента</td></tr>
<tr><td>26</td><td>При построении доверительного интервала для генеральной дисперсии при больших объёмах выборки используют</td><td>1. распределение Пирсона
2. нормальный закон распределения
3. распределение Фишера-Снедекора
4. распределение Стьюдента</td></tr>
<tr><td>27</td><td>В результате измерений некоторой физической величины одним прибором (без систематических ошибок) получены следующие результаты (в Н): 12,5; 14,5; 16,5. Тогда несмещенная оценка среднего квадратического отклонения равна…</td><td>1. 2 2. 3
3. 5 4. 4</td></tr>
<tr><td>28</td><td>На основе значений некоторой вы-</td><td>1. 78 2. 84</td></tr>
</table>

	борки объема n=20,была вычислена выборочная дисперсия, значение которой 76. Тогда исправленная выборочная дисперсия данной выборки равна …	3. 80 4. 83
29	Точечная оценка параметра распределения равна 21. Тогда его интервальная оценка может иметь вид…	1. (20; 21) 2. (21; 22) 3. (0; 21) 4. (20; 22)
30	Дан доверительный интервал (18,5; 25,9) для оценки математического ожидания нормально распределённого количественного признака. Тогда точечная оценка математического ожидания равна ...	1. 22,2 2. 22,5 3. 22 4. 22,6
31	Ширина доверительного интервала зависит от:	1. уровня значимости и числа наблюдений 2. уровня значимости 3. числа наблюдений
32	Статистической гипотезой называют предположение:	1. о равенстве двух параметров 2. о неравенстве двух величин 3. о виде или параметрах неизвестного закона распределения случайной величины
33	Нулевая гипотеза — это:	1. выдвинутая гипотеза, которую нужно проверить 2. альтернативная гипотеза 3. гипотеза, определяющая закон распределения 4. гипотеза о равенстве нулю параметра распределения
34	Если нулевую гипотезу в результате проверки критерия отвергают, какова вероятность при этом совершить ошибку?	1. β 2. α 3. $1-\beta$ 4. γ

35	Что называют ошибкой первого рода?	1.Гипотеза H_0 верна и ее принимают согласно критерию 2. Гипотеза H_0 верна, но ее отвергают согласно критерию 3. Гипотеза H_0 неверна и ее отвергают согласно критерию 4. Гипотеза H_0 неверна, но ее принимают согласно критерию
36	Что называют ошибкой второго рода?	1.Гипотеза H_0 верна и ее принимают согласно критерию 2. Гипотеза H_0 верна, но ее отвергают согласно критерию 3. Гипотеза H_0 неверна и ее отвергают согласно критерию 4. Гипотеза H_0 неверна, но ее принимают согласно критерию
37	При проверке гипотезы о нормальном распределении генеральной совокупности при заданном уровне значимости определено критическое значение критерия $\chi^2_{kp} = 14{,}4$. Тогда эмпирические и теоретические частоты будут различаться значимо, если наблюдаемое значение статистического критерия $\chi^2_{\text{набл}}$будет равно ...	1. 14,0 2. 14,1 3. 14,8 4. 14,3
38	Связь считается сильной, если значение выборочного коэффициента корреляции:	1. Равно 0 2. В диапазоне от 0 до 0,3 3. В диапазоне от 0,7 до 1 4. В диапазоне от 1 до 2 5. Принимает положительные значения
39	Если одному значению первого	1. Функциональная

	признака соответствует несколько значений второго – это связь:	2. Положительная 3. Регрессионная 4. Прямолинейная 5. Корреляционная
40	Коэффициент линейной корреляции может принимать значения:	1. Близкое к нулю 2. От –1 до 1 3. Близкое к единице (+1 или -1) 4. Измеряется в процентах
41	Аналитическое выражение связи определяется с помощью метода анализа:	1. корреляционного; 2. регрессионного; 3. группировок
42	Для изображения корреляционной зависимости используется график:	1. линейный 2. радиальный 3. график рассеяния точек 4. динамический
43	Выборочное уравнение прямой линии регрессии Y на X имеет вид $y = 1{,}4 - 1{,}8x$, а средние квадратические отклонения равны $s_x = 0{,}12$; $s_y = 0{,}54$. Тогда коэффициент корреляции равен….	1. –0,4 2. -3,6 3. -0,02 4. 0,4
44	По результатам выборки, извлечённой из генеральной совокупности (X, Y) вычислены выборочный коэффициент регрессии Y на X $(b = 2{,}5)$ и выборочные средние $\bar{x} = 30$; $\bar{y} = 41$. Тогда выборочное уравнение прямой линии регрессии Y на X имеет вид …	1. $\bar{y}_x = 2{,}5x + 116$ 2. $\bar{y}_x = 2{,}5x - 75$ 3. $\bar{y}_x = 2{,}5x - 116$ 4. $\bar{y}_x = 2{,}5x + 132$
45	Зависимость средней выборки одного рабочего за смену y (штук) от квалификации x (разряды) приведена в таблице: x: 2, 3, 4, 5 y: 12, 19, 23, 30 Уравнение регрессии $y = a + bx$. Тогда коэффициент b равен…	1. 2,8 2. 5,3 3. 5,8 4. 7,2

4.2.1 Ответы к тесту по теме «Элементы математической статистики»

№ задачи	**№ правильного ответа**	**№ задачи**	**№ правильного ответа**	**№ задачи**	**№ правильного ответа**
1	1	**16**	4	**31**	1
2	1	**17**	2	**32**	3
3	3	**18**	3	**33**	1
4	2	**19**	4	**34**	2
5	3	**20**	3	**35**	2
6	3	**21**	3	**36**	4
7	4	**22**	1	**37**	3
8	1	**23**	4	**38**	3
9	3	**24**	3	**39**	5
10	1	**25**	4	**40**	2
11	2	**26**	2	**41**	2
12	1	**27**	1	**42**	3
13	1	**28**	3	**43**	1
14	3	**29**	4	**44**	1
15	1	**30**	1	**45**	3

5. КОНТРОЛЬНАЯ РАБОТА №1 ПО ТЕМЕ «ТЕОРИЯ ВЕРОЯТНОСТЕЙ»

Задание 1

1.Наугад взяты два числа x, y. Каждое число не превышает 2. Найти вероятность, что сумма $(x+y)$ не превышает 1,5.

2.Два биатлониста произвели по одному выстрелу по мишени с вероятностью $p_1 = 0{,}6$, $p_2 = 0{,}7$. Найти вероятность: а) только одного попадания; б) хотя бы одного попадания.

3.Вероятности попадания в баскетбольное кольцо для двух спортсменов равны $p_1 = 0{,}8$, $p_2 = 0{,}7$ соответственно. Найти вероятность того, что: а) только один из них попал в кольцо; б) хотя бы один из них попал в кольцо.

4.Три посылки отправлены в три почтовых отделения. Вероятность своевременной доставки в первое отделение 0,95, второе – 0,9, в третье – 0,95. Найти вероятность, что все три посылки доставлены вовремя.

5.К месту аварии направлены милиция и скорая, которые могут успеть к потерпевшим вовремя с вероятностями $p_1 = 0{,}9$, $p_2 = 0{,}8$. Какова вероятность, что обе машины успели вовремя; хотя бы одна машина успела вовремя.

6.Для сигнализации установлены два датчика. Вероятность того, что при пожаре сработает первый датчик 0,95. Второй сработает с вероятностью 0,99. Найти вероятность того, что при пожаре сработает только один датчик.

7.В прямоугольник 5 см x 4 см вписан круг с радиусом 2 см. Какова вероятность того, что точка, поставленная в прямоугольник, окажется в круге?

8.Вероятности выполнения месячного плана двумя цехами предприятия равны p_1=0,9, p_2=0,7. Полагая, что цеха работают независимо друг от друга, найти вероятности того, что: а) только один цех выполнит план; б) хотя бы один цех выполнит план.

9.Вероятность хотя бы одного попадания в цель при 4 выстрелах равна 0,9984. Найти вероятность попадания в цель при одном выстреле.

10.Два орудия выпускают в цель по одному снаряду с вероятностями попадания $p_1 = 0{,}6$, $p_2 = 0{,}7$. Найти вероятность того, что в цель попадет: а) только один снаряд; б) хотя бы один снаряд.

11.Ребенок имеет на руках 4 кубика с буквами А,А,М,М. Какова вероятность того, что ребенок соберет слово «мама»?

12.Каждая из двух команд по 5 гребцов проводит жеребьевку для присвоения номеров. Два брата входят в состав разных команд. Найти вероятность того, что братья получат: а) номер 4; б) одинаковый номер.

13.Каких чисел от 1 до 10000 больше: тех, в записи которых встречается 9, или тех, в которых она не встречается?

14.Заказ на покупку книг отправили в два магазина. Вероятность своевременной доставки в каждый из них равна 0,8. Найти вероятность того, что своевременно получит книги: а) только один магазин; б) хотя бы один магазин.

15.Рейсовый автобус может опоздать вследствие двух независимых причин: плохой погоды и неисправности оборудования. Вероятность плохой погоды равна 0,3, вероятность неисправности 0,4. Найти вероятность того, что автобус опоздает: а) только по причине плохой погоды; б) по любым причинам.

16.Условия дуэли предусматривают по 2 выстрела каждого из дуэлянтов по очереди до первого попадания. Вероятности их попадания при одном выстреле равны 0,2 и 0,3 соответственно. Найти вероятность того, что первый дуэлянт: а) поразит соперника вторым выстрелом; б) поразит соперника.

17.Вероятность забить гол нападающим при одном ударе по воротам равна 0,3. Найти вероятность того, что после двух ударов будет забит: а) только один гол; б) хотя бы один гол.

18.Вероятность выловить щуку утром 0,48, а днем 0,39, 0,39 вечером, 0 ночью. Найти вероятность того, что две щуки выловили днем; хотя бы одна из 2 щук выловлена утром.

19.Автомобильный номер содержит четыре цифры. Найти вероятность того, что у встречного автомобиля сумма цифр номера: а) равна двум; б) не более двух.

20.Группу из 20 студентов нужно разделить на три подгруппы. В первую должно входить 3 человека, во вторую –12, в третью – 5. Сколько способов существует это сделать?

21.В коробке пять белых зефира и два шоколадных. Найти вероятность того, что наугад извлеченные два зефира будут: а) одного цвета; б) шоколадные.

22.Двое независимо друг от друга садятся в состав в метро из восьми вагонов. Найти вероятность их встречи.

23.У мамы 2 яблока и 7 киви. Каждый день в течение 5 дней подряд она выдает ребенку в школу по одному фрукту. Сколько способов существует это сделать?

24.В коробке конфет пять «Красная шапочка» и три «Суфле». Найти вероятность того, что наугад извлеченные две конфеты будут: а) разные; б) «Суфле».

25.Найти вероятность того, что двое встречных прохожих родились: а) в один месяц; б) зимой.

26.Найти вероятность того, что сумма цифр наугад выбранного двузначного числа: а) равна пяти; б) меньше пяти.

27.Найти вероятность того, что произведение цифр наугад выбранного двузначного числа: а) равно трем; б) меньше трех.

28.Вероятности поймать ерша при поклевке у рыболовов равны 0,57 и 0,48 соответственно. У каждого произошла одна поклевка. Найти вероятность того, что их общий улов составит: а) одну рыбу; б) не менее одной рыбы.

29.Телефонный номер содержит 9 цифр. Вначале номера идут цифры +37529. Найти вероятность того, что сумма оставшихся цифр наугад выбранного номера: а) равна 10; б) меньше 7.

30. Найти вероятность того, что при 10 случайных нажатиях на клавиши пишущей машинки будет напечатано слово "математика". Клавиатура содержит 40 клавишей.

Задание 2

Вариант 1–5

В банке обслуживаются 100 клиентов. Из них K_1 клиентов взяли краткосрочные кредиты, K_2 – среднесрочные, K_3 – долгосрочные. Вероятность того, что клиент вернет кредит в зависимости от срока пользования кредитом составляет P_1, P_2, P_3 соответственно.

Найти вероятность того, что клиент вернет кредит.

Клиент вернул кредит. Найти вероятность, что это был i –й тип кредита.

1)	K_1=44 P_1=0,82	K_2=36 P_2=0,9	K_3=20 P_3=0,75	i =1
2)	K_1=30 P_1=0,87	K_2=38 P_2=0,85	K_3=32 P_3=0,92	i =2
3)	K_1=35 P_1=0,91	K_2=35 P_2=0,95	K_3=30 P_3=0,84	i =3
4)	K_1=28 P_1=0,8	K_2=37 P_2=0,75	K_3=35 P_3=0,92	i =1
5)	K_1=40 P_1=0,88	K_2=33 P_2=0,91	K_3=27 P_3=0,84	i =3

Вариант 6–10

Мониторинг для определения уровня знаний проходил в трех потоках, содержащих K_1, K_2, K_3 студентов. Вероятность успешно пройти мониторинг для каждого потока: $P_1\%, P_2\%, P_3\%$ (учитывалась доля отличников к общему количеству студентов в потоках). Какова вероятность, что студент, успешно прошедший мониторинг, учится в i –м потоке.

6)	K=30 P_1=0,21	K_1=12 P_2=0,1	K_2=10 P_3=0,3	K_3=8 i =2
7)	K=25 P_1=0,15	K_1=10 P_2=0,12	K_2=7 P_3=0,22	K_3=8 i =1

8)	$K=28$ $P_1=0{,}14$	$K_1=9$ $P_2=0{,}12$	$K_2=8$ $P_3=0{,}12$	$K_3=11$ $i=3$
9)	$K=24$ $P_1=0{,}09$	$K_1=8$ $P_2=0{,}12$	$K_2=6$ $P_3=0{,}14$	$K_3=10$ $i=3$
10)	$K=27$ $P_1=0{,}13$	$K_1=7$ $P_2=0{,}08$	$K_2=10$ $P_3=0{,}12$	$K_3=10$ $i=2$

Вариант 11–15

По данным журнала K_1% туристов посещают Израиль, K_2% – Турцию, K_3% – остальные направления. Вероятности того, что туристы едут к родственникам равны P_1, P_2 и P_3 соответственно.

Найти вероятность того, что случайным образом выбранный турист едет посетить родственников.

Известно, что турист посетил родственников, какова вероятность, что он ездил по i-ому направлению?

1)	$K_1=48\%$ $P_1=0{,}11$	$K_2=16\%$ $P_2=0{,}12$	$K_3=36\%$ $P_3=0{,}05$	$i=3$
2)	$K_1=53\%$ $P_1=0{,}08$	$K_2=21\%$ $P_2=0{,}10$	$K_3=26\%$ $P_3=0{,}13$	$i=1$
3)	$K_1=33\%$ $P_1=0{,}12$	$K_2=37\%$ $P_2=0{,}06$	$K_3=30\%$ $P_3=0{,}15$	$i=3$
4)	$K_1=41\%$ $P_1=0{,}09$	$K_2=22\%$ $P_2=0{,}14$	$K_3=37\%$ $P_3=0{,}19$	$i=2$
5)	$K_1=29\%$ $P_1=0{,}12$	$K_2=44\%$ $P_2=0{,}12$	$K_3=27\%$ $P_3=0{,}08$	$i=2$

Вариант 16–20

На витрине три полки: на первой a_1 отечественных товаров и b_1 зарубежных товаров; на второй a_2 отечественных и b_2 зарубежных товаров, на третьей a_3 отечественных и b_3 зарубежных товаров. Менчендайзер выбирает наугад одну полку и выбирает одну продукцию. Найти вероятность того, что эта продукция отечественного производства.

Выбранная продукция отечественного производства. Найти вероятность того, что она находится на k–ой полке.

16) $a_1 = 8; b_1 = 14; a_2 = 7, b_2 = 5; a_3 = 10; b_3 = 0; k = 1.$
17) $a_1 = 8; b_1 = 0; a_2 = 6; b_2 = 14; a_3 = 9; b_3 = 3; k = 3.$
18) $a_1 = 11; b_1 = 4; a_2 = 12; b_2 = 0; a_3 = 6; b_3 = 8; k = 3.$
19) $a_1 = 7; b_1 = 9; a_2 = 12; b_2 = 5; a_3 = 8; b_3 = 0; k = 2.$
20) $a_1 = 11; b_1 = 4; a_2 = 6; b_2 = 0; a_3 = 3; b_3 = 7; k = 1.$

Вариант 21–25

В исследовании участвуют k_1 мужчина, имеющий врожденные аномалии и k_2 мужчины, имеющие инфекционные заболевания. Врожденные аномалии могут вызывать бесплодие с вероятностью p_1, а инфекционные заболевания p_2. Найти вероятность того, что наугад выбранный мужчина, участвующий в исследовании, оказался бесплодным. Какова причина бесплодия вероятнее всего в этом случае?

21) $k_1 = 5; k_2 = 7;\ p_1 = 0{,}82;\ p_2 = 0{,}75.$
22) $k_1 = 8; k_2 = 6;\ p_1 = 0{,}90;\ p_2 = 0{,}84.$
23) $k_1 = 3; k_2 = 7;\ p_1 = 0{,}88;\ p_2 = 0{,}94.$
24) $k_1 = 6; k_2 = 9;\ p_1 = 0{,}78;\ p_2 = 0{,}86.$
25) $k_1 = 8; k_2 = 7;\ p_1 = 0{,}92;\ p_2 = 0{,}77.$

Вариант 26–30

Учебный годовой план учитывает изучение общеобразовательных k_1% дисциплин и k_2% специальных дисциплин. При изучении общеобразовательных дисциплин доля занятий с привлечением информационных технологий составляет p_1, а для специальных дисциплин эта доля равна p_2. Какова вероятность того, что случайно выбранное занятие проводится с привлечением информационных технологий? В рассмотренном занятии использовались информационные технологии, какова вероятность, что это было занятие по специальной дисциплине?

26) $k_1 = 42\%; k_2 = 58\%;\ p_1 = 0{,}24;\ p_2 = 0{,}33.$
27) $k_1 = 63\%; k_2 = 37\%;\ p_1 = 0{,}12;\ p_2 = 0{,}26.$
28) $k_1 = 57\%; k_2 = 43\%;\ p_1 = 0{,}18;\ p_2 = 0{,}11.$
29) $k_1 = 72\%; k_2 = 28\%;\ p_1 = 0{,}14;\ p_2 = 0{,}35.$

30) $k_1 = 38\%; k_2 = 62\%; p_1 = 0{,}15; p_2 = 0{,}27.$

Задание 3

а)

Проведены исследования среднесуточной температуры X первого марта. Результаты представлены законом распределения.

Найти математическое ожидание $M(x)$, дисперсию $D(x)$, среднеквадратическое отклонение $\sigma(x)$, начальные и центральные моменты первого, второго порядков.

1.	X_i	2	3	4	6	**16.**	X_i	–1	1	2	3
	P_i	¼	2/16	7/16	3/16		P_i	0,15	0,2	0,25	0,4
2.	X_i	1	2	3	5	**17.**	X_i	0	2	3	5
	P_i	0,3	0,2	0,2	0,3		P_i	1/7	1/2	2/14	3/14
3.	X_i	–5	2	3	4	**18.**	X_i	–2	2	3	4
	P_i	0,4	0,3	0,1	0,2		P_i	0,1	0,38	0,42	0,1
4.	X_i	2	4	6	8	**19.**	X_i	4	5	7	8
	P_i	0,1	0,2	0,3	0,4		P_i	0,23	0,27	0,13	0,37
5.	X_i	3	5	7	9	**20.**	X_i	–3	–1	2	3
	P_i	0,2	0,1	0,4	0,3		P_i	0,25	0,2	0,3	0,25
6.	X_i	1	3	4	6	**21.**	X_i	1	3	5	6
	P_i	0,1	0,3	0,2	0,4		P_i	0,7	0,12	0,13	0,05
7.	X_i	0	1	2	3	**22.**	X_i	–10	–5	0	5
	P_i	0,729	0,343	0,027	0,001		P_i	0,17	0,33	0,22	0,28
8.	X_i	3	4	7	10	**23.**	X_i	4	5	6	7
	P_i	0,2	0,4	0,1	0,3		P_i	0,2	0,22	0,31	0,27
9.	X_i	–2	0	2	4	**24.**	X_i	2	4	5	8
	P_i	0,21	0,15	0,36	0,28		P_i	0,05	0,15	0,45	0,35
10.	X_i	0	1	2	3	**25.**	X_i	1	2	5	7
	P_i	0,225	0,245	0,255	0,275		P_i	0,4	0,2	0,1	0,3
11.	X_i	2	3	5	7	**26.**	X_i	–3	1	4	6
	P_i	1/3	1/3	2/9	1/9		P_i	1/5	2/25	12/25	6/25
12.	X_i	1	3	4	7	**27.**	X_i	–8	–7	–4	2
	P_i	0,1	0,3	0,5	0,1		P_i	0,25	0,1	0,15	0,5
13.	X_i	–3	0	1	4	**28.**	X_i	10	12	14	15
	P_i	0,345	0,455	0,115	0,085		P_i	0,12	0,33	0,31	0,24

14.	X_i	4	6	7	10	**29.**	X_i	0	1	3	4
	P_i	0,4	0,1	0,3	0,2		P_i	0,34	0,22	0,26	0,18
15.	X_i	1	2	3	4	**30.**	X_i	3	5	7	9
	P_i	0,5	0,2	0,15	0,35		P_i	1/6	2/6	5/12	1/12

б)

По заданной функции распределения $F(x)$ случайной величины ξ найти плотность распределения и построить ее график. Вычислить вероятность попадания случайной величины в заданный интервал $[a,b]$, математическое ожидание и дисперсию.

$$F(x)=\begin{cases} 0, x\le 0 \\ ax+b, 0<x<a \\ 1, x>a \end{cases}$$ a = N, b=1/N, N—номер варианта.

Задание 4

а)

Варианты 1–15

Построить ряд распределения, функцию распределения и ее график, найти математическое ожидание и дисперсию случайной величины X – числа наступлений случайного события A в указанной ниже серии независимых испытаний.

1.В результате переполюсовки стартер выгорает с вероятностью 0,6. Рассматриваются три стартера.

2.Фотоумножитель конвертирует фотоны в фотоэлектроны с вероятностью 0,1. Счетчик создал четыре фотона.

3.В результате переполюсовки провод выгорает с вероятностью 0,1. Рассмотри три провода.

4.Поезд может прибыть по расписанию (событие A) с вероятностью $P(A)=0,9$. Рассматриваются три рейса.

5.В результате сдергивания плюсовой клеммы АКБ на работающем двигателе реле выгорает с вероятностью 0,6. Рассматриваются три реле.

6.В среднем при профессиональном наборе текста специалист совершает ошибку (событие A) с вероятностью 0,02. Рассматриваются 4 текста.

7.Процент мужчин среди заболевших СПИДом составляет 89,1%. Исследуются 4 человека.

8.Инфекционные заболевания становятся причиной смерти с вероятностью 0,25. Рассматривается три летальных случая.

9.В ядерной реакции может образоваться резонансная частица (событие А) вероятностью $P(A)=0,2$. Рассматриваются три реакции.

10.Травма стала причиной летального исхода с вероятностью 0,09. Рассматриваются три летальных исхода.

11.Вероятность выздоровления от дизентерии 0,97. Рассматриваются три больных дизентерией.

12.Грипп вызывает летальный исход с вероятностью 0,004. Рассматриваются причины трех летальных исходов.

13.При изготовлении детали она может оказаться бракованной (событие A) с вероятностью $P(A)$=0,02. Изготовлено три детали.

14.Полевая всхожесть семени хазара равна $P(A)$=0,358. Посажено три семени.

15.В результате переполюсовки диодный мост выгорает с вероятностью $P(A)$=0,99. Рассматриваются три диодных моста.

Варианты 16–30

Решить задачи, используя приближенные формулы: локальную формулу Муавра-Лапласа, интегральную формулу Муавра-Лапласа и формулу Пуассона.

16.Вероятность наступления беременности в 43 года равна 0,12. Какова вероятность забеременеть 75 женщинам из 100.

17.Найти вероятность того, что 30 каналов из 100 свободны, если вероятность того, что канал свободен 0,12.

18.Z-бозон с вероятность появления 0,6991 распадается на адроны. Найти вероятность того, что событие наступит 50 раз в 243 испытаниях.

19.Вероятность того, что деталь высшего сорта, равна 0,5. Найти вероятность того, что из 1000 деталей 501 высшего сорта.

20.Биатлонист на тренировке сделал 20 выстрелов с вероятностью попадания при отдельном выстреле 0,7. Найти вероятность того, то при этом будет 9 попаданий.

21.Вероятность нарушения стандарта при штамповке металлических клемм равна 0,1. Найти вероятность того, что для 700 клемм число бракованных колец заключено между 205 и 250.

22.Полевая всхожесть семян рапана равна 0,372. Найти вероятность того, что из 1000 посаженных семян число проросших будет заключено между 790 и 833.

23.Согласно полученным сотрудниками Маастрихского университета данным, вероятность рождения мальчика равна 0,58. Найти вероятность того, что из 1000 рождающихся детей мальчиков будет не менее 500, но не более 750.

24.При установившемся технологическом процессе цех по производству муки выпускает в среднем 90% продукции первого сорта. Найти вероятность того, что в партии из 1000 упаковок число первосортных заключено между 672 и 760.

25.Вероятность выхода из строя нагревательного элемента равна 0,02. Какова вероятность того, что в партии из 600 изделий, 20 изделий выйдут из строя по причине выхода из строя нагревательного элемента?

26.На лекции по теории вероятностей присутствуют 120 студентов второго курса. Найти вероятность того, что: 1) $k\ (k=0{,}1{,}2)$ студентов из присутствующих изучали эту дисциплину в колледже. 2) Хотя бы один студент курса изучал эту дисциплину в колледже.

27.В магазин со склада отправили 500 деревянных полок. Вероятность повреждения каждой полки при транспортировке равна 0,002. Найти вероятность того, что среди 500 полок будет повреждено при транспортировке: а) ровно 3 полки; б) ровно одна полка; в) не более 3 полок; г) более 3 полок.

28.На экспертизу передано 1000 телефонов. Вероятность того, что телефон не соответствует должному уровню качества, равна 0,003. Найти вероятность того, что а) хотя один телефон соответствует качеству; б) менее 2; в) ровно 2 телефона; г) более 2 телефонов.

29.Микросхема состоит из 100 элементов. Вероятность неисправлености одного элемента равна 0,001 и не зависит от состояния других элементов. Какова вероятность отказа: а) трех элементов; б) не менее 2 и не более 4 элементов; в) не менее 2 элементов.

30.Учебно-методическое пособие издано тиражом 100 экземпляров. Вероятность того, что экземпляр будет бракованный равна 0,0001. Найти вероятность того, что: а) тираж содержит 15 бракованных книг; б) менее 5; в) 99 экземпляров не содержит брак.

б)

Колебания среднесуточной температуры – случайная величина *X* с математическим ожиданием a и среднеквадратическим отклонением σ, которая распределена нормально. Записать плотность распределения и функцию распределения случайной величины *X*. Найти вероятность того, что среднесуточная температура находится в диапазоне $(\alpha;\beta)$.

№ варианта	a	σ	α	β
1	2,4	1,2	3,3	5,2
2	2,3	0,7	2,1	3,2
3	1,8	0,6	1,3	4,8
4	1,5	2,2	0,9	4,1
5	3,4	1,1	0,8	2,9
6	2,8	0,4	0,5	3,8
7	1,2	2,9	1	4
8	0	3	0	2,4

9	20	5	15	25
10	10	2	12	14
11	10	4	2	13
12	8	1	5	9
13	2	4	6	10
14	3	2	3	10
15	4	5	2	11
16	3,3	2,3	3,2	4,7
17	0,7	0,1	0,5	1,1
18	6,3	0,9	5,1	8,2
19	8,2	4,3	5,4	7,2
20	2,8	1,4	3,1	4,1
21	2	1	0	3
22	1	4	-5	0
23	9	5	5	14
24	6	3	2	11
25	3	1	0,5	3,5
26	4	1	3	7
27	6	2	3	10
28	10	2	9	12
29	4	3	0	9
30	-8	3	-9	0

5.1 Решение типового варианта контрольной работы №1 по теме «Теория вероятностей»

Задание 1.

1. В ящике 20 деталей, из которых 4 бракованные.

Найти вероятность того, что наугад взятая из ящика деталь окажется бракованной.

Решение.

Так как каждая из имеющихся деталей может быть из ящика, то число всех равновозможных элементарных исходов $n = 20$. Число исходов, благоприятствующих появлению бракованной детали, $m = 4$. Если событие A означает, что взятая деталь бракованная, то

$$P(A) = \frac{m}{n} = \frac{4}{20} = \frac{1}{5}.$$

2. Для сигнализации об аварии установлены два независимо работающих сигнализатора. Вероятность того, что при аварии сигнализатор сработает, равна 0,95 для первого сигнализатора и 0,9 для второго.

Найти, вероятность того, что при аварии сработает только один сигнализатор.

Решение.

Вероятности появления каждого из двух независимых событий A_1 и A_2 соответственно равны $p_1 = 0{,}95$ и $p_2 = 0{,}9$.

B_1 – появилось только событие A_1;

B_2 – появилось только событие A_2.

Появление события B_1 равносильно появлению события $A_1\overline{A_2}$ (появилось первое событие и не появилось второе), т.е. $B_1 = A_1\overline{A_2}$.

Появление события B_2 равносильно появлению события $\overline{A_1}A_2$ (появилось второе событие и не появилось первое). $B_2 = \overline{A_1}A_2$.

Таким образом, чтобы найти вероятность появления одного из событий A_1 и A_2, достаточно найти вероятность появления одного, безразлично какого, из событий B_1 и B_2. События A_1 и A_2 несовместны, поэтому применима теорема сложения:

$$P(B_1 + B_2) = P(B_1) + P(B_2);$$
$$P(B_1) = 0{,}95 \cdot 0{,}1 = 0{,}095;$$
$$P(B_2) = 0{,}9 \cdot 0{,}05 = 0{,}045.$$

События A_1 и A_2 – независимы, следовательно, независимы события A_1 и $\overline{A_2}$, а также $\overline{A_1}$ и A_2, поэтому применима теорема умножения:

$$P(B_1) = P(A_1 \cdot \overline{A_2}) = P(A_1) \cdot P(\overline{A_2}) = p_1 \cdot q_2;$$
$$P(B_2) = P(\overline{A_1} \cdot A_2) = P(\overline{A_1}) \cdot P(A_2) = q_1 \cdot p_2.$$

Подставив эти вероятности, найдем искомую вероятность появления только одного из событий A_1 и A_2:

$$P(B_1 + B_2) = p_1 \cdot q_2 + q_1 \cdot p_2;$$
$$p_1 = 0{,}95;\ p_2 = 0{,}9;$$
$$q_1 = 0{,}05;\ p_2 = 0{,}1;$$
$$P(B_1 + B_2) = 0{,}95_1 \cdot 0{,}1 + 0{,}05 \cdot 0{,}9 = 0{,}095 + 0{,}045 = 0{,}14.$$

Задание 2.

1. В вычислительной лаборатории имеется 6 компьютеров одного типа и 4 – другого. Вероятность того, что на время выполнения некоторого расчета компьютер I типа не выйдет из строя, равна 0,95; для компьютера другого типа – 0,8. Студент производит расчет на наудачу взятом компьютере.

Найти вероятность того, что до окончания расчета компьютер не выйдет из строя.

Решение.

Обозначим через A – компьютер не выйдет из строя до окончания расчета. Возможны следующие предположения (гипотезы) о первоначальном качестве компьютеров; всего компьютеров в лаборатории – 10, из них:

$$P(B_1) = \frac{6}{10} = 0{,}6;$$

$$P(B_2) = \frac{4}{10} = 0{,}4.$$

В сумме гипотезы всегда равны 1, проверим:

$$P(B_1) + P(B_2) = 0{,}6 + 0{,}4 = 1 \text{(верно)}.$$

Условная вероятность того, что студент воспользуется компьютером 1-го типа, равна $P_{B_1}(A) = 0{,}95$; 2-го типа – $P_{B_2}(A) = 0{,}8$.

Искомая вероятность того, что до окончания расчета компьютер любого типа не выйдет из строя, равна:

$$P(A) = P(B_1) \cdot P_{B_1}(A) + P(B_2) \cdot P_{B_2}(A) = 0{,}6 \cdot 0{,}95 + 0{,}4 \cdot 0{,}8 = 0{,}89$$

2. Два автомата производят одинаковые детали, которые сбрасываются на общий конвейер. Производительность первого автомата вдвое больше производительности второго. Первый автомат производит в среднем 78% деталей отличного качества, а второй – 86%.

Найти вероятность того, что эта деталь произведена первым автоматом.

Решение.

Обозначим событие A – деталь отличного качества. Можно сделать два предположения (гипотезы): B_1 – деталь произведена первым автоматом; $P(B_1) = \frac{2}{3}$, B_2 – деталь произведена вторым автоматом; $P(B_2) = \frac{1}{3}$.

Условная вероятность того, что деталь будет отличного качества, если она произведена первым автоматом, равна

$$P_{B_1}(A) = 0{,}78.$$

Условная вероятность того, что деталь будет отличного качества, если она произведена вторым автоматом, равна

$$P_{B_2}(A) = 0{,}86.$$

Вероятность того, что наудачу взятая деталь окажется отличного качества, по формуле полной вероятности равна

$$P(A)=P(B_1)\cdot P_{B_1}(A)+P(B_2)\cdot P_{B_2}(A)=\frac{2}{3}\cdot 0{,}78+\frac{1}{3}\cdot 0{,}86=0{,}52\cdot 0{,}29=0{,}81$$

Искомая вероятность того, что взятая отличная деталь произведена первым автоматом, по формуле Байеса равна

$$P_A(B_1)=\frac{P(B_1)\cdot P_{B_1}(A)}{P(A)}=\frac{\frac{2}{3}\cdot 0{,}78}{0{,}81}=\frac{0{,}52}{0{,}81}=0{,}64.$$

Задание 3.

При измерении роста у 18 студентов установлено, что у трех рост – 188 см; у четверых – 182 см; у пятерых – 180 см; у шестерых 178 см. СВ X – рост студента.

Записать закон распределения СВ X. Вычислить математическое ожидание $M(X)$, дисперсию $D(X)$ и среднеквадратическое отклонение σ.

Решение.

Вероятность обнаружения среди 18 студентов троих с ростом 188 см: $p_1=\frac{3}{18}=0{,}17.$

Аналогично вероятность обнаружения среди 18 студентов четверых с ростом 182 см: $p_2=\frac{4}{18}=0{,}22$ и т.д.

Получаем закон распределения в виде следующей таблицы:

X	188	182	180	178
p	0,17	0,22	0,28	0,33

$$p_3=\frac{5}{18}=0{,}28;\ p_4=\frac{6}{18}=0{,}33;$$

$$p=p_1+p_2+p_3+p_4=0{,}17+0{,}22+0{,}28+0{,}33=1.$$

Далее находим:

$$M(X)=188\cdot 0{,}17+182\cdot 0{,}22+180\cdot 0{,}28+178\cdot 0{,}33=31{,}96+40{,}04+50{,}4+58{,}74=$$
$$=191{,}14.$$

$$M(X^2)=188^2\cdot 0{,}17+182^2\cdot 0{,}22+180^2\cdot 0{,}28+178^2\cdot 0{,}33=6008{,}5+7287{,}3+$$
$$+9072+10455{,}7=32823{,}5.$$

$$D(X)=M(X^2)-(M(X))^2=32823{,}5-32811{,}7=11{,}8.$$

$$\sigma(X)=\sqrt{D(X)}=\sqrt{11{,}8}=3{,}44.$$

Задание 4.

Устройство состоит из четырех независимо работающих элементов. Вероятность отказа каждого элемента равна 0,2.

Записать закон распределения числа отказавших элементов устройства, найти математическое ожидание и дисперсию.

Решение.

Дискретная случайная величина X может принимать значения $x_1 = 0$; $x_2 = 1$; $x_3 = 2$; $x_4 = 3$; $x_5 = 4$. Так как $n = 4$; $p = 0{,}2$; $q = 0{,}8$, то по формуле Бернулли находим:

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k};$$

$$P_4(0) = q^4 = 0{,}8^4 \approx 0{,}41;$$

$$P_4(1) = C_4^1 \cdot p^1 \cdot q^3 = 4 \cdot 0{,}2 \cdot 0{,}8^3 \approx 0{,}41;$$

$$P_4(2) = C_4^2 \cdot p^2 \cdot q^2 = \frac{4 \cdot 3}{2} \cdot 0{,}2^2 \cdot 0{,}8^2 = 6 \cdot 0{,}04 \cdot 0{,}64 \approx 0{,}15;$$

$$P_4(3) = C_4^3 \cdot p^3 \cdot q^1 = \frac{4 \cdot 3 \cdot 2}{3 \cdot 2} \cdot 0{,}2^3 \cdot 0{,}8 \approx 0{,}03;$$

$$P_4(4) = (0{,}2)^4 \approx 0{,}002.$$

X	0	1	2	3	4
p	0,41	0,41	0,15	0,03	0,002

$$p = 0{,}41 + 0{,}41 + 0{,}15 + 0{,}03 + 0{,}001 \approx 1.$$

Математическое ожидание равно

$$M(X) = n \cdot p = 4 \cdot 0{,}2 = 0{,}8;$$

$$D(X) = n \cdot p \cdot q = 4 \cdot 0{,}2 \cdot 0{,}8 = 0{,}64.$$

6. КОНТРОЛЬНАЯ РАБОТА №2 ПО ТЕМЕ «ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ»

Задание 1. Из генеральной совокупности извлечена выборка, представленная в виде статистического ряда. Требуется:

1) вычислить выборочное среднее $\bar{x}$, выборочную дисперсию D_B, исправленную выборочную дисперсию S^2и среднее квадратичное отклонение s;

2) найти размах варьирования; моду и медиану;

3) построить полигон частот эмпирическую функцию распределения;

4) проверить, согласуются ли выборочные данные с гипотезой о нормальном распределении случайной величины X графически (с помощью гистограммы частот) и с помощью критерия согласия Пирсона при уровне значимости $\alpha = 0{,}05$ $(\alpha = 0{,}01)$, представив данную выборку в виде интервального ряда. Количество интервалов рассчитать по формуле Стерджеса $l = 1 + 3{,}322 \lg n$;

5) найти с доверительной вероятностью $\gamma = 0{,}95$ $(\gamma = 0{,}99)$ доверительный интервал для математического ожидания, а также доверительный интервал для $\sigma(x)$.

Варианты заданий

1	x_i	3	6	9	11	16	20	21	23	25
	n_i	4	5	6	8	10	7	6	5	4
2	x_i	0	6	9	12	15	20	22	26	28
	n_i	9	8	7	11	10	6	4	3	2
3	x_i	0	1	2	3	4	5	6	7	8
	n_i	2	3	5	7	6	10	6	4	2
4	x_i	2	3	4	5	6	7	8	9	10
	n_i	2	4	6	8	12	7	5	4	4
5	x_i	5	8	10	12	14	15	16	17	20
	n_i	3	5	10	12	15	12	4	3	1
6	x_i	4	5	6	7	9	10	12	13	15
	n_i	2	3	4	10	13	12	6	3	2
7	x_i	3	3,5	3,7	4,0	4,1	4,5	4,8	5,1	5,3
	n_i	3	5	6	8	9	10	7	4	3
8	x_i	5	5,1	5,4	5,5	5,8	6,1	6,4	6,9	7
	n_i	4	7	8	10	8	7	3	2	1
9	x_i	0	2	5	7	10	12	15	18	20
	n_i	2	3	4	10	11	10	8	4	3

10	x_i	0,6	0,7	0,8	0,9	1,2	1,3	1,4	1,5	1,6
	n_i	4	5	6	8	10	9	7	4	2

11	x_i	2,5	2,7	2,8	3	3,1	3,3	3,5	3,6	3,7
	n_i	1	4	7	8	10	9	5	4	2
12	x_i	1,2	1,4	1,6	1,8	2	2,2	2,4	2,5	3
	n_i	2	3	5	7	10	11	6	7	4
13	x_i	-5	-3	-1	1	3	5	7	9	11
	n_i	4	6	8	10	8	7	3	7	2
14	x_i	2,5	2,6	2,7	2,9	3	3,2	3,5	3,4	3,7
	n_i	4	5	8	9	10	6	5	2	1
15	x_i	4	8	9	11	13	14	15	18	20
	n_i	1	2	5	7	9	10	8	3	5
16	x_i	1	3	5	4	6	7	8	10	13
	n_i	5	7	10	12	8	5	7	4	2
17	x_i	3,5	3,7	4	4,2	4,4	4,5	5	5,1	5,2
	n_i	2	4	7	6	12	9	7	2	1
18	x_i	-2	0	2	4	6	8	10	12	14
	n_i	3	5	6	8	10	9	7	4	3
19	x_i	4,1	4,4	4,7	4,8	5	5,2	5,5	5,7	5,8
	n_i	3	5	7	9	10	9	8	7	2
20	x_i	6	8	10	12	15	18	20	22	24
	n_i	5	6	7	8	9	11	8	4	2
21	x_i	-6	-4	-2	0	2	4	6	8	10
	n_i	2	4	5	7	12	10	9	6	5
22	x_i	2,3	2,5	2,7	2,8	2,9	3	3,2	3,5	4
	n_i	1	2	4	5	6	8	11	10	3
23	x_i	2,1	2,4	2,5	3,2	3,3	3,4	3,7	3,8	4
	n_i	2	4	5	10	8	6	5	4	1
24	x_i	1	1,3	1,5	1,7	1,9	2	2,2	2,5	2,8
	n_i	4	6	7	10	11	8	6	5	3
25	x_i	10	12	14	15	16	20	22	24	30
	n_i	2	4	5	7	8	10	12	5	2
26	x_i	0	3	5	7	9	11	13	15	17
	n_i	2	3	5	7	10	5	3	2	2

27	x_i	1	1,2	1,8	2,2	2,4	2,6	3,2	3,5	4
	n_i	6	8	13	15	20	16	10	7	5

28	x_i	2	2,1	2,4	2,5	2,7	2,9	3	3,2	3,4
	n_i	5	7	14	15	21	16	9	7	6
29	x_i	4	6	8	9	11	13	15	16	18
	n_i	5	9	14	16	18	16	9	6	7
30	x_i	0,3	0,5	0,7	0,9	1	1,1	1,3	1,5	1,7
	n_i	5	6	14	15	22	15	9	8	6

Задание 2. По заданной таблице зависимости признаков X и Y, соответствующей номеру варианта:

1) Вычислить выборочный коэффициент корреляции; проверить его на значимость, приняв $\alpha = 0{,}05$.

2) Методом наименьших квадратов выровнять зависимость Y от X по прямой $y = a + bx$.

3) Вычислить остаточную дисперсию, сделать вывод.

4) Построить корреляционное поле и линию регрессии на корреляционном поле.

Варианты заданий

Вариант 1						**Вариант 2**					
X	0	0,4	1	1,5	2,1	X	9.5	5.8	4.0	1.3	3.4
Y	6,7	7,1	7,6	8,1	8,5	Y	1.5	0.1	-1.3	-2.1	-1.6
Вариант 3						**Вариант 4**					
X	2	4	6	8	9	X	-2	-1	0	1	2
Y	-0,8	-1,5	-2	-3	-3,7	Y	2,3	2,8	3,6	4	4,7
Вариант 5						**Вариант 6**					
X	1	1,5	2	2,5	3	X	6.1	0.8	0.3	1.2	0.4
Y	0,3	0,8	1,3	1,9	2,5	Y	0	-2.0	-3.3	-1.8	-2.9
Вариант 7						**Вариант 8**					
X	5.6	4.8	9.6	5.0	5.3	X	3.7	1.5	9.3	0.4	6.5
Y	-0.4	-1.6	1.3	-0.2	-0.1	Y	-1.7	-2.0	2.2	-3.2	0.9
Вариант 9						**Вариант 10**					
X	2	4	5	6	8	X	-2	0	1	2	4
Y	-1	5	8,5	12	18	Y	0,5	1	1,5	2	3
Вариант 11.						**Вариант 12.**					

X	4.2	7.4	10	6.6	1.6	X	0.5	7.7	5.4	3.7	3.6
Y	-1.0	1.4	1.3	1.5	-2.8	Y	-3.2	1.2	0.2	-0.6	-1.1
Вариант 13						**Вариант 14**					
X	8.6	4.3	5.2	9.2	4.8	X	4.7	3.3	0.7	2.9	6.7
Y	1.2	-0.4	-0.6	1.1	-0.9	Y	-0.6	-1.8	-2.5	-1.6	0.6
Вариант 15						**Вариант 16**					
X	3.0	1.2	0.3	1.7	8.0	X	9.1	0.2	5.9	5.4	9.9
Y	-1.2	-1.8	-2.9	-1.9	0.1	Y	2.0	-2.8	0.1	0.0	1.6
Вариант 17						**Вариант 18**					
X	7.2	6.9	5.6	0.9	1.1	X	3.5	7.1	1.0	7.9	3.1
Y	0.2	-0.2	0.6	-2.1	-2.4	Y	-1.2	0.4	-2.2	1.2	-1.1
Вариант 19						**Вариант 20**					
X	1.6	3.8	0.7	3.4	4.3	X	7.7	2.8	6.3	7.1	6.2
Y	-2.2	-1.6	-2.5	-0.9	-0.2	Y	1.4	-2.1	0.1	0.5	-0.3
Вариант 21						**Вариант 22**					
X	0	1	2	4	6	X	1	2	3	4	5
Y	6	7,2	9,4	11	15,2	Y	3,2	4,2	2,7	0,7	1,2
Вариант 23						**Вариант 24**					
X	0	1	3	6	8	X	4,1	5	8,1	10,4	12
Y	3,2	4,3	5,4	8,3	9	Y	4	8	10	14	16
Вариант 25						**Вариант 26**					
X	1,4	1,5	1,8	2	2,4	X	0	0,4	1	1,5	2,1
Y	2	1,9	2,3	2,6	3	Y	6,7	7,1	7,6	8,1	8,6
Вариант 27						**Вариант 28**					
X	0,3	0,9	1,5	2	2,2	X	1	4	9	16	25
Y	0,2	0,4	0,3	0,5	0,8	Y	0,1	3	8,1	14,9	23,9
Вариант 29						**Вариант 30**					
X	-2	0	1	2	3	X	-3	-2	-1	0	1
Y	4,7	1	-1,2	-3,1	-5	Y	-1,2	-0,8	-0,4	0	0,1

6.1 Решение типового варианта контрольной работы №2 по теме «Элементы математической статистики»

Задание 1.

Из генеральной совокупности извлечена выборка, представленная в виде статистического ряда:

x_i	4	6	7	8	10	11	12
n_i	3	7	10	11	9	6	4

Требуется:

1)вычислить выборочное среднее $\bar{x}$, выборочную дисперсию D_B, исправленную выборочную дисперсию S^2 и среднее квадратичное отклонение s;

2)найти размах варьирования; моду и медиану;

3)построить полигон частот и эмпирическую функцию распределения;

4)проверить, согласуются ли выборочные данные с гипотезой о нормальном распределении случайной величины X графически и с помощью критерия согласия Пирсона при уровне значимости $\alpha = 0{,}05$, представив данную выборку в виде интервального ряда. Количество интервалов рассчитать по формуле Стерджеса $l = 1 + 3{,}322 \lg n$;

5)найти с доверительной вероятностью $\gamma = 0{,}95$ $(\gamma = 0{,}99)$ доверительный интервал для математического ожидания, а также доверительный интервал для $\sigma(x)$.

Решение.

1) Объем выборки равен $n = \sum n_i = 3 + 7 + 10 + 11 + 9 + 6 + 4 = 50$.

Выборочное среднее определим по формуле:

$$\bar{x} = \frac{1}{n}\sum x_i = \frac{1}{50}(4 \cdot 3 + 6 \cdot 7 + 7 \cdot 10 + 8 \cdot 11 + 10 \cdot 9 + 11 \cdot 6 + 12 \cdot 4)$$
$$= \frac{416}{50} = 8{,}32$$

Для нахождения выборочной дисперсии составим следующую вспомогательную таблицу:

x_i	4	6	7	8	10	11	12	$\sum$
n_i	3	7	10	11	9	6	4	**50**
$(x_i - \bar{x})^2 \cdot n_i$	55,987	37,677	17,424	1,1264	25,402	43,094	54,17	**234,88**

Тогда $D_B = \frac{1}{n}\sum(x_i - \bar{x})^2 \cdot n_i = \frac{1}{50} \cdot 234{,}88 = 4{,}697$; $S^2 = \frac{n}{n-1} D_B = \frac{50}{49} \cdot 4{,}6976 \approx 4{,}8$

Исправленное среднее квадратичное отклонение будет $s = \sqrt{4{,}8} \approx 2{,}19$.

2) Размах варьирования находится по формуле $R = \max x_i - \min x_i = 12 - 4 = 8$. Так как мода – это варианта, которой соответствует наибольшая частота, то $Mo = 8$. Не сгруппированные данные образуют дискретный вариационный ряд, содержащий чётное число вариант $(n = 50 = 2 \cdot 25)$, поэтому $M_e = \frac{x_{25} + x_{26}}{2} = \frac{8+8}{2} = 8$.

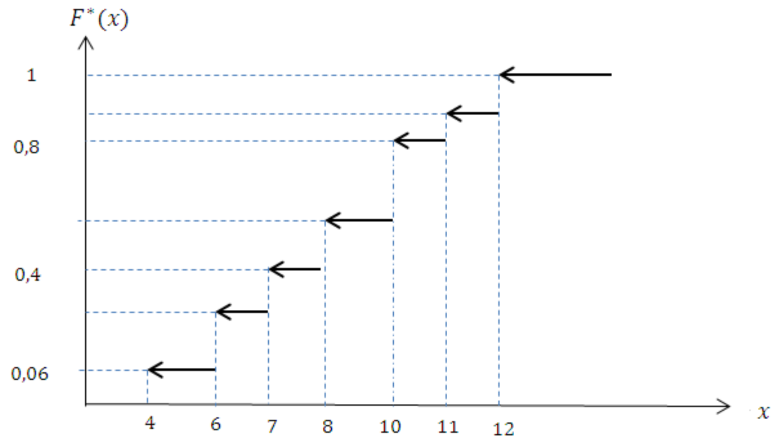
3) Согласно определению эмпирической функции распределения её значение при любом x равно $F^*(x) = \frac{n_x}{n}$, где n_x – количество элементов x_i выборки, меньших, чем x.

Например, при $x = 3$ имеем $n_x = 0$, $F^*(x) = 0$; при $x = 5$ $n_x = 3$ $F^*(x) = \frac{3}{50} = 0{,}06$; при $x = 6{,}2$ $n_x = 3 + 7 = 10$, $F^*(x) = \frac{10}{50} = 0{,}2$; при $x = 7{,}9$ $n_x = 3 + 7 + 10 = 20$, $F^*(x) = \frac{20}{50} = 0{,}4$; при $x = 9$ $F^*(x) = \frac{31}{50} = 0{,}6$; при $x = 10{,}6$ $F^*(x) = \frac{40}{50} = 0{,}8$; при $x = 11{,}2$ $F^*(x) = \frac{46}{50} = 0{,}92$.

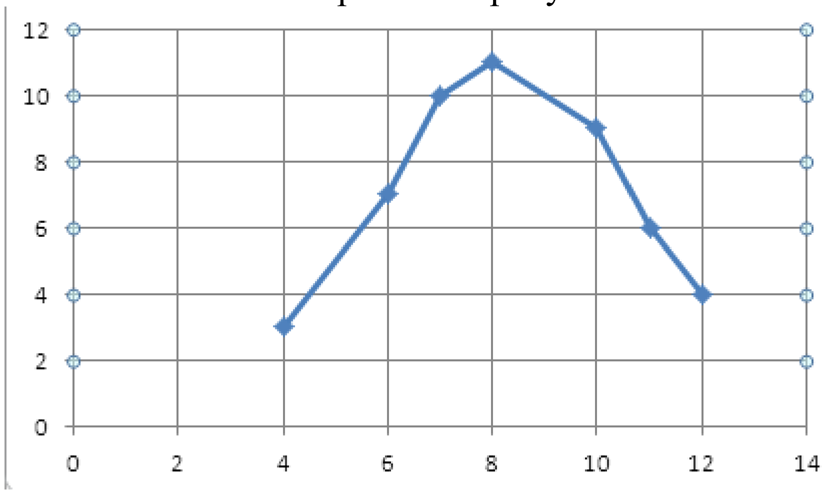
Тогда:

$$F^*(x) = \begin{cases} 0 & x \le 4 \\ 0{,}06 & 4 < x \le 6 \\ 0{,}2 & 6 < x \le 7 \\ 0{,}4 & 7 < x \le 8 \\ 0{,}62 & 8 < x \le 10 \\ 0{,}8 & 10 < x \le 11 \\ 0{,}92 & 11 < x \le 12 \\ 1 & x > 12 \end{cases}$$

График эмпирической функции распределения:



Полигон частот изображён на рисунке:



4) Так как полигон частот по форме напоминает кривую Гаусса, то можно сделать предположение о том, что случайная величина X распределена по нормальному закону. Проверим данное утверждение по критерию Пирсона.

Вычислим количество интервалов:

$$l = 1 + 3{,}322 \lg n = 1 + 3{,}322 \lg 50 = 6{,}644 \approx 7.$$

Длина интервала: $h = \frac{R}{h} = \frac{8}{7} \approx 1{,}2$. Границы интервалов будут:

$$a_0 = 4,\ a_1 = 4 + 1{,}2 = 5{,}2,\ \ a_2 = 6{,}4,\ \ a_3 = 7{,}6,\ \ a_4 = 8{,}8,$$
$$a_5 = 10,\qquad a_6 = 11{,}2,\qquad a_7 = 12{,}4$$

Посчитаем число выборочных значений, попавших в каждый интервал. Частота m_i интервала $[a_i;\ a_{i+1}]$ $(i = \overline{0{,}6})$ подсчитывается с помощью ряда, как число наблюдений, попавших в интервал. Так, в первый $(i = 0)$ интервал $[4;\ 5{,}2]$ попало 3 значения; во второй $(i = 1)$ – $[5{,}2;\ 6{,}4]$ попало 7 значений. Аналогично получаем частоты 3 – 7 интервалов.

Полученные данные сведём в следующую таблицу:

$[a_i;\ a_{i+1}]$	4 – 5,2	5,2 – 6,4	6,4 – 7,6	7,6 – 8,8	8,8 – 10	10 – 11,2	11,2 – 12,4
m_i	3	7	10	11	9	6	4

Найдем теоретические вероятности p_i по формуле:

$$p_i = \Phi\left(\frac{a_{i+1} - \bar{x}}{s}\right) - \Phi\left(\frac{a_i - \bar{x}}{s}\right)$$

Результаты вычислений сведем в таблицу:

i	a_i	a_{i+1}	$\frac{a_i - \bar{x}}{s}$	$\frac{a_{i+1} - \bar{x}}{s}$	$\Phi\left(\frac{a_i - \bar{x}}{s}\right)$	$\Phi\left(\frac{a_{i+1} - \bar{x}}{s}\right)$	p_i	np_i
1	4	5,2	–	-1,42	-0,5	-0,4222	0,0778	3,89
2	5,2	6,4	-1,42	-0,88	-0,4222	-0,3106	0,1116	5,58
3	6,4	7,6	-0,88	-0,33	-0,3106	-0,1293	0,1813	9,065
4	7,6	8,8	-0,33	0,22	-0,1293	0,0871	0,2164	10,82
5	8,8	10	0,22	0,77	0,0871	0,2794	0,1923	9,615
6	10	11,2	0,77	1,32	0,2794	0,4066	0,1272	6,36
7	11,2	12,4	1,32	–	0,4066	0,5	0,0934	4,67

Так как ожидаемые (эмпирические) частоты первого и седьмого интервалов группировки не удовлетворяют условию $min(n_i) \geq 5$, объединим эти интервалы (первый со вторым; а седьмой – с шестым).

Вычислим наблюдаемое значение критерия Пирсона. Для этого составим следующую расчетную таблицу:

i	n_i	np_i	$(n_i - np_i)^2$	$\frac{(n_i - np_i)^2}{np_i}$
1	10	9,47	0,2809	0,029662
2	10	9,065	0,874225	0,09644
3	11	10,82	0,0324	0,002994
4	9	9,615	0,378225	0,039337
5	10	11,03	1,0609	0,096183
Σ	**50**	**50**		**0,264616**

По таблице критических точек распределения χ^2, уровню значимости $\alpha = 0{,}05$ и числу степеней свободы $k = 5 - 3 = 2$ находим $\chi^2_{kp}(0{,}05; 2) = 6$. Так как $\chi^2_{\text{набл}} = 0{,}264616 < 6$, то гипотеза о нормальном распределении принимается.

5) Доверительный интервал для математического ожидания найдём по формуле $\bar{x} - t_\gamma \cdot \frac{s}{\sqrt{n}} < a < \bar{x} + t_\gamma \cdot \frac{s}{\sqrt{n}}$. Значение t_γ определим по таблице для доверительной вероятности $\gamma = 0{,}95$ $(\gamma = 0{,}99)$ и объёму выборки $n = 50$:

$$t_\gamma(0{,}95; 50) = 2{,}009; \quad t_\gamma(0{,}99; 50) = 2{,}679$$

Тогда доверительный интервал имеет вид:

для $\gamma = 0{,}95$: $8{,}32 - 2{,}009 \cdot \frac{2{,}19}{\sqrt{50}} < a < 8{,}32 + 2{,}009 \cdot \frac{2{,}19}{\sqrt{50}} \Rightarrow 7{,}7 < a < 8{,}9$

для $\gamma = 0{,}99$: $8{,}32 + 2{,}679 \cdot \frac{2{,}19}{\sqrt{50}} < a < 8{,}32 + 2{,}679 \cdot \frac{2{,}19}{\sqrt{50}} \Rightarrow$
$7{,}49 < a < 9{,}15$.

Задание 2.

По заданной таблице зависимости признаков *X* и *Y*:

1)Вычислить выборочный коэффициент корреляции; проверить его на значимость, приняв $\alpha = 0{,}05$.

2)Методом наименьших квадратов выровнять зависимость Y от X по прямой $y = a + bx$.

3)Вычислить остаточную дисперсию, сделать вывод.

4)Построить корреляционное поле и линию регрессии на корреляционном поле.

X	-2	-1	0	1	2	3
Y	5,6	5	4,3	4	3,6	3

Решение. Найдём выборочные средние $\bar{x}, \bar{y}$, а также оценки для средних квадратических отклонений и корреляционного момента, для чего составим следующую вспомогательную таблицу:

i	**1**	**2**	**3**	**4**	**5**	**6**	$\sum$
X	-2	-1	0	1	2	3	**3**
Y	5,6	5	4,3	4	3,6	3	**25,5**
$(x_i - \bar{x})^2$	6,25	2,25	0,25	0,25	2,25	6,25	**17,5**
$(y_i - \bar{y})^2$	1,8225	25	18,49	16	12,96	9	**83,2725**
$(x_i - \bar{x})(y_i - \bar{y})$	-3,375	-1,125	-0,025	-0,125	-0,975	-3,125	**-8,75**

Здесь $\bar{x} = \frac{3}{6} = 0{,}5$; $\bar{y} = \frac{25{,}5}{6} = 4{,}25$. Тогда:

$$s_x = \sqrt{\frac{1}{n-1}\sum_i (x_i - \bar{x})^2} = \sqrt{\frac{17{,}5}{5}} \approx 1{,}87$$

$$s_y = \sqrt{\frac{1}{n-1}\sum_i (y_i - \bar{y})^2} = \sqrt{\frac{4{,}435}{5}} \approx 0{,}94$$

$$s_{xy} = \frac{1}{n-1}\sum_i (x_i - \bar{x})(y_i - \bar{y}) = \frac{-8{,}75}{5} = -1{,}75$$

Выборочное значение коэффициента корреляции:

$$r_B = \frac{s_{xy}}{s_x \cdot s_y} = \frac{-1{,}75}{1{,}87 \cdot 0{,}94} \approx -0{,}996$$

Проверим значимость полученного выборочного коэффициента корреляции. Найдём наблюдаемое значение критерия:

$$T_{\text{набл}} = r_B \cdot \sqrt{\frac{n-2}{1-r_B^2}} = -0{,}996 \cdot \sqrt{\frac{4}{1-0{,}996^2}} = -22{,}4$$

По таблице критических точек распределения Стьюдента, по уровню значимости $\alpha = 0{,}05$ и числу степеней свободы $k = n - 2 = 4$ находим критическую точку двусторонней критической области $t_{kp}(0{,}05; 4) = 2{,}78$.

Так как $|T_{\text{набл}}| > t_{kp}$, то отвергаем гипотезу о равенстве нулю генерального коэффициента корреляции, значит *X* и *Y* – коррелированы.

2) Для вычисления коэффициентов a и b воспользуемся таблицей:

i	x_i	y_i	x_i^2	$x_i \cdot y_i$
1	-2	5,6	4	-11,2
2	-1	5	1	-5
3	0	4,3	0	0
4	1	4	1	4
5	2	3,6	4	7,2
6	3	3	9	9
$\sum$	**3**	**25,5**	**19**	**4**

Запишем нормальную систему уравнений. Так как $\bar{x} = 0{,}5$; $\bar{y} = 4{,}25$; $\overline{x^2} = \frac{19}{6} \approx 3{,}17$; $\overline{xy} = \frac{4}{6} \approx 0{,}67$, то:

$$\begin{cases} a + 0{,}5b = 4{,}25 \\ 0{,}5a + 3{,}17b = 0{,}67 \end{cases}$$

Решая систему по формулам Крамера, получим:

$$a = \frac{\begin{vmatrix} 4{,}25 & 0{,}5 \\ 0{,}67 & 3{,}17 \end{vmatrix}}{\begin{vmatrix} 1 & 0{,}5 \\ 0{,}5 & 3{,}17 \end{vmatrix}} = \frac{4{,}25 \cdot 3{,}17 - 0{,}5 \cdot 0{,}67}{3{,}17 - 0{,}5 \cdot 0{,}5} = \frac{13{,}1375}{2{,}92} \approx 4{,}5;$$

$$b = \frac{\begin{vmatrix} 1 & 4{,}25 \\ 0{,}5 & 0{,}67 \end{vmatrix}}{\begin{vmatrix} 1 & 0{,}5 \\ 0{,}5 & 3{,}17 \end{vmatrix}} = \frac{0{,}67 - 0{,}5 \cdot 4{,}25}{3{,}17 - 0{,}5 \cdot 0{,}5} = \frac{-1{,}455}{2{,}92} \approx -0{,}5$$

Следовательно, зависимость между величинами X и Y выражается приближённой формулой

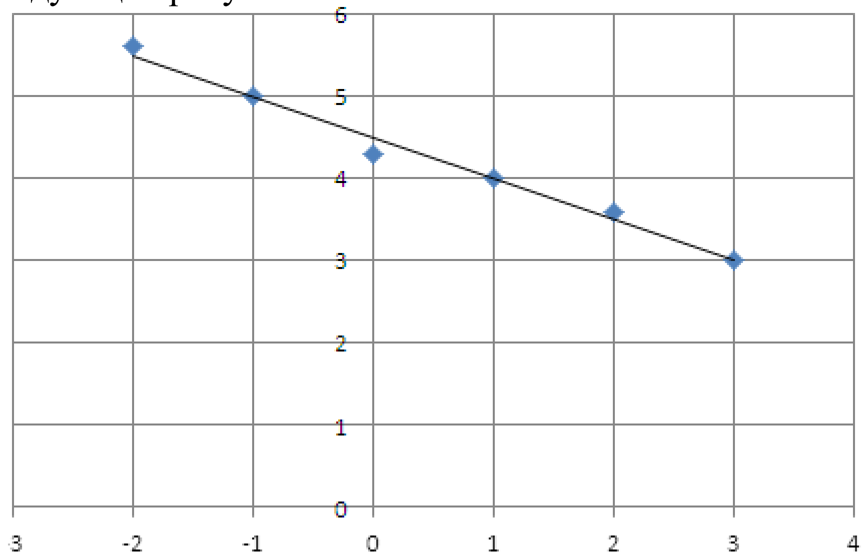
$\bar{y}_x = 4{,}5 - 0{,}5x.$

3) Остаточная дисперсия:

$$D_{\text{ост}} = s_y^2(1 - r_B^2) = 4{,}08^2 \cdot (1 - 0{,}996^2) = 0{,}133$$

То есть величина ошибки, которая возникает при замене Y линейной функцией, невелика – можно сделать вывод, что между величинами X и Y существует приближённая линейная зависимость.

4) Корреляционное поле и линия регрессии на корреляционном поле представлены на следующем рисунке:



7. ВОПРОСЫ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ И ЭЛЕМЕНТАМ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ ДЛЯ ПОДГОТОВКИ К ЭКЗАМЕНУ (ЗАЧЕТУ)

1.Пространство элементарных событий. Алгебра событий.

2.Классическое, геометрическое и статистическое определения вероятности.

3.Свойства вероятности.

4.Теорема сложения вероятностей.

5.Теорема умножения вероятностей.

6.Формула полной вероятности.

7.Формулы Байеса.

8.Схема повторных независимых испытаний. Формула Бернулли.

9.Локальная и интегральная теоремы Муавра–Лапласа.

10.Формула Пуассона.

11.Случайные величины (СВ). Закон распределения СВ. Непрерывные и дискретные СВ.

12.Математическое ожидание и его свойства.

13.Дисперсия и ее свойства.

14.Функция распределения и ее свойства.

15.Плотность распределения вероятностей и ее свойства.

16.Биномиальный закон распределения. Математическое ожидание и дисперсия СВ, распределенной по биномиальному закону.

17.Распределение Пуассона. Математическое ожидание и дисперсия.

18.Равномерный закон распределения. Математическое ожидание и дисперсия.

19.Показательный закон распределения. Математическое ожидание и дисперсия.

20.Нормальный закон распределения. Функция и плотность распределения. Математическое ожидание и дисперсия.

21.Вероятность попадания нормально распределенной СВ в интервал. Вероятность отклонения СВ от математического ожидания по модулю. Правило трех сигм.

22.Двумерные СВ. Закон распределения. Условный закон распределения.

23.Числовые характеристики двумерных СВ. Условное математическое ожидание и условная дисперсия.

24.Корреляционный момент и его свойства.

25.Коэффициент корреляции и его свойства.

26.Закон больших чисел. Неравенство Чебышева.

27.Теорема Чебышева.

28.Теорема Бернулли.

29.Центральная предельная теорема Ляпунова.

30.Генеральная и выборочная совокупности. Вариационный ряд.

31.Полигон и гистограмма.

32.Эмпирическая функция распределения и ее свойства.

33.Понятия выборки и выборочной функции (статистики). Выборочная средняя и выборочная дисперсия.

34.Оценки параметров распределения. Точечные оценки и требования, предъявляемые к ним.

35.Точечные оценки для математического ожидания и дисперсии генеральной совокупности.

36.Интервальные оценки. Доверительный интервал.

37.Построение доверительного интервала для математического ожидания нормально распределенной СВ при известном среднем квадратическом отклонении.

38.Распределение Стьюдента.

39.Построение доверительного интервала для математического ожидания нормально распределенной СВ при неизвестном среднем квадратическом отклонении.

40.Распределение Пирсона.

41.Построение доверительного интервала для среднего квадратического отклонения нормально распределенной СВ.

42.Понятие о статистических гипотезах и критериях согласия.

43.Критерий согласия Пирсона (χ^2).

44.Выборочный коэффициент корреляции и его свойства.

45.Уравнение регрессии. Линейная регрессия. Определение коэффициентов линейной регрессии методом наименьших квадратов.

46.Нелинейная регрессия. Определение параметров нелинейной регрессии.

8. СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1. Латышева, И.Г. Конспект лекций по математике для студентов инженерно-технических специальностей. В четырех частях. Часть 3 / И. Г. Латышева, И.В. Прусова, Л.А. Барминова, О.Г. Вишневская, Е.А. Глинская, Н.А. Кондратьева, А. Н. Мелешко. – Электрон. дан. – БНТУ, 2007. – Режим доступа: http://rep.bntu.by/handle/data/1214.

2. Кондратьева, Н.А. Математика. Специальные разделы. Элементы теории функций комплексной переменной и операционного исчисления. Теория вероятностей. Элементы математической статистики / Н.А. Кондратьева, Н.К. Прихач, Н.Н. Буснюк, А.Н. Мелешко. – Электрон. дан. – БНТУ, 2014. – Режим доступа: http://rep.bntu.by/handle/data/9383.

3. Кондратьева, Н.А Электронный учебно-методический комплекс по учебной дисциплине «Математика» для студентов экономических специальностей второго курса обучения заочного отделения ПСФ в двух частях. Часть1 [Электронный ресурс] / Н.А. Кондратьева, Л.В. Бокуть, А.Н.Мелешко, Т.Г. Крупенкова. – Минск: БНТУ, 2014. – БНТУ/ЭУМК-ПСФ85-168.– Режим доступа: http://rep.bntu.by/handle/data/26446

4. Семенов, В. Теория вероятностей и математическая статистика. Стандарт третьего поколения / В. Семенов // С.-Пб. : Питер, 2013. – 192 с.

5. Просветов, Г. Теория вероятностей и математическая статистика. Задачи и решения / Г. Просветов // М. : Альфа-Плюс, 2009. – 272 с.

6. Золотаревская, Д. Теория вероятностей. Задачи с решениями / Д. Золотаревская // М. : Либроком, 2016. – 170 с.

7. Спирин, П. Теория вероятностей и математическая статистика / П. Спирин, М. Спирина // С.-Пб. : Academia, 2013. – 352 с.

8. Рябушко, А. Индивидуальные задания по высшей математике. В 4 частях. Часть 4. Операционное исчисление. Элементы теории устойчивости. Теория вероятностей. Математическая статистика / А. Рябушко // Мн. : Вышэйшая школа, 2013. – 336 с.

9. Бричкова, Е. Теория вероятностей. Примеры и задачи / Е. Бричкова, А. Гусак // Мн. : ТетраСистем, 2013. – 288 с.

10. Микулик, Н.А. Теория вероятностей и математическая статистика / Н.А. Микулик, А.В. Метельский. – Минск: Пион, 2002.

11. Фигурин, В.В. Теория вероятностей и математическая статистика / В.В. Фигурин. – Минск: Новое знание, 2000.

12. Гайшун, Л.Н. Теория вероятностей / Л.Н. Гайшун, Г.К. Игнатьева, О.А. Велько. – Минск: МПУ, 2002.

13. Минюк, С.А. Математика для инженеров: учебник: в 2 т. / С.А. Минюк [и др.]. – Минск: Элайда, 2006. – Т. 2.

14.Матальыцкий, М.А. Теория вероятностей и математическая статистика: пособие / М.А. Маталыцкий, Т.В. Русилко. – 2-е изд., перераб. и доп. – Гродно: ГрГУ, 2009 – 219 с.

15.Савич, Л.К. Теория вероятностей и математическая статистика: учеб. пособие для студентов эконом. специальностей учреждений, обеспечивающих получение высш. образования/ Л.К. Савич, Н.А. Смольская; науч. ред. О.И. Лаврова. – Мн.: Адукацыя і выхаванне, 2006. – 208 с.

16.Гмурман, В.Е. Теория вероятностей и математическая статистика. Изд. 4-е, доп. Учебное пособие для вузов. – М., «Высшая школа», 1972. – 368 с.

17.Лисьев, В.П. Теория вероятностей и математическая статистика: Учебное пособие/ Московский государственный университет экономики, статистики и информатики. – М., 2006. – 199 с.

18.Лунгу, К.Н. Высшая математика. Руководство к решению задач. Ч. 2 / К.Н. Лунгу, Е.В. Макаров – М., ФИЗМАТЛИТ, 2007 – 384 с.

19.Математическая статистика в примерах и задачах: учебное пособие / Г.С. Евдокимова; Смол. гос. ун-т. – Смоленск: Изд-во СмолГУ, 2014. – 98 с. https://studfiles.net/preview/3544158/

20.Сборник задач по математике для втузов. В 4 частях. Ч. 4: Учебное пособие для втузов/Под общ. ред. А.В. Ефимова и А.С. Поспелова. – 3-е изд. перераб. и доп. – М., Издательство физико-математической литературы, 2003. –432 с.

21.Еськова, О.И. Основы статистической обработки информации: пособие / О.И. Еськова, Л.П. Авдашкова, М.А. Грибовская. – Минск, Беларусь, 2011. – 175с.

22.Булдык, Г.М. Руководство к решению задач по теории вероятностей и математической статистике: Для практической и самостоятельной работы студентов экономических специальностей. – Минск, ФУ Аинформ, 2009 –228 с.

23.Гмурман, В.Е. Руководство к решению задач по теории вероятностей и математической статистике: Учебное пособие для студентов вузов/ В.Е. Гмурман. – 7-е изд., доп. – М.: Высш. шк., 2003. – 405 с.

24. Рябушко, А.П. Индивидуальные задания по высшей математике: Операционное исчисление. Элементы теории устойчивости. Теория вероятностей. Математическая статистика: учеб. пособие/А. П. Рябушко. – Мн.: Выш. шк., 2006. – 336 с.

25. Тест по математической статистики [Электронный ресурс]. – Режим доступа: https://multiurok.ru/files/tiest-po-matiematichieskoi-statistikie.html. Дата доступа: 20.12.2018.

26. Тесты по математической статистики [Электронный ресурс]. – Режим доступа: https://studfiles.net/preview/4347353/. Дата доступа: 20.12.2018.

27. Основы математической статистики [Электронный ресурс]. – Режим доступа: http://megamozg.kz/index.php?page=view_mat&id=4724&partition=student&subpartition=mathematics. Дата доступа: 20.12.2018.

27. Тесты по курсу теория вероятностей [Электронный ресурс]. – Режим доступа: https://poisk-ru.ru/s17335t8.html. Дата доступа: 20.12.2018.

28. Дубровина, О.В. Прикладная математика. – Учебное пособие для практических и лабораторных работ для студентов заочного отделения специальности 1–54 01 01 «Метрология, стандартизация и сертификация» // О.В. Дубровина, Н.К. Прихач, В.М. Романчак. – Минск: БНТУ, 2009 – 70с.

9. ПРИЛОЖЕНИЯ

Приложение 1

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

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3084	0,3982	0,3980	0,3977	0,3973
0,1	0,3970	0,3965	0,3961	0,3956	0,3951	0,3945	0,3939	0,3932	0,3025	0,3918
0,2	0,3910	0,3902	0,3894	0,3885	0,3876	0,3867	0,3857	0,3847	0,3836	0,3825
0,3	0,3814	0,3802	0,3790	0,3778	0,3765	0,3752	0,3739	0,3726	0,3712	0,3697
0,4	0,3683	0,3668	0,3652	0,3637	0,3621	0,3605	0,3589	0,3572	0,3555	0,3538
0,5	0,3521	0,3503	0,3485	0,3467	0,3448	0,3429	0,3410	0,3391	0,3372	0,3352
0,6	0,3332	0,3312	0,3292	0,3271	0,3251	0,3230	0,3209	0,3187	0,3166	0,3144
0,7	0,3123	0,3101	0,3079	0,3056	0,3034	0,3011	0,2989	0,2966	0,2943	0,2920
0,8	0,2897	0,2874	0,2850	0,2827	0,2804	0,2780	0,2756	0,2732	0,2709	0,2685
0,9	0,2661	0,2637	0,2613	0,2589	0,2565	0,2541	0,2516	0,2492	0,2468	0,2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	0,2179	0,2155	0,2131	0,2107	0,2083	0,2059	0,2036	0,2012	0,1989	0,1965
1,2	0,1942	0,1919	0,1895	0,1872	0,1849	0,1826	0,1804	0,1781	0,1758	0,1736
1,3	0,1714	0,1691	0,1669	0,1647	0,1626	0,1604	0,1582	0,1561	0,1539	0,1518
1,4	0,1497	0,1476	0,1456	0,1435	0,1415	0,1394	0,1374	0,1354	0,1334	0,1315
1,5	0,1295	0,1276	0,1257	0,1238	0,1219	0,1200	0,1182	0,1163	0,1145	0,1127
1,6	0,1109	0,1092	0,1074	0,1057	0,1040	0,1023	1006	0989	0973	0957
1,7	0,0940	0,0925	0,0909	0,0893	0,0878	0863	0,0846	0,0833	0,0818	0,0804
1,8	0,0790	0,0775	0,0761	0,0748	0,0734	0,0721	0,0707	0,0694	0,0681	0,0669
1,9	0,0656	0,0644	0,0632	0,0620	0,0608	0,0596	0,0584	0,0573	0,0562	0,0551
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0,0440	0,0431	0,0422	0,0413	0,0404	0,0396	0,0387	0,0379	0,0371	0,0363
2,2	0,0355	0,0347	0,0339	0,0332	0,0325	0,0317	0,0310	0,0303	0,0297	0,0290
2,3	0,0283	0,0277	0,0270	0,0264	0,0258	0,0252	0,0246	0,0241	0,0235	0,0229
2,4	0,0224	0,0219	0,0213	0,0208	0,0203	0,0198	0,0194	0,0189	0,0184	0,0180

x	0	1	2	3	4	5	6	7	8	9
2,5	0,0175	0,0171	0,0167	0,0163	0,0158	0,0154	0,0151	0,0147	0,0143	0,0139
2,6	0,0136	0132	0,0129	0,0126	0,0122	0,0119	0,0116	0,0113	0,0110	0,0107
2,7	0,0104	0,0101	0,0099	0,0096	0,0093	0,0091	0,0088	0,0086	0,0084	0,0081
2,8	0,0079	0,0077	0,0075	0,0,0073	0,0071	0,0069	0,0067	0,0065	0,0063	0,0061
2,9	0,0060	0,0058	0,0056	0,0055	0,0053	0,0051	0,0050	0,0048	0,0047	0,0046
3,0	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0,0033	0,0032	0,0032	0,0030	0,0029	0,0028	0,0027	0,0026	0025	0025
3,2	0,0024	0,0023	0,0022	0,0022	0,0021	0,0020	0,0020	0,0019	0,0018	0,0018
3,3	0,0017	0,0017	0,0012	0,0016	0,0015	0,0015	0,0014	0,0014	0,0013	0,0013
3,4	0,0012	0,0012	0,0010	0,0011	0,0011	0,0010	0,0010	0,0010	0,0009	0,0009
3,5	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007	0,0007	0,0007	0,0006
3,6	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005	0,0004
3,7	0,0004	0,0004	0,0004	0004	0,0004	0,0004	0,0003	0,0003	0,0003	0,0003
3,8	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002	0,0002
3,9	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0001	0,0001

Приложение 2

Значения функции Лапласа $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,01	0,0040	0,33	0,1293	0,65	0,2422	0,97	0,3340
0,02	0,0080	0734	0,1331	0,66	0,2454	0,98	0,3365
0,03	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,04	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,05	0,0199	0,37	0,1443	0,69	0,2549	2,01	0,3438
0,06	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,07	0,0279	0,39	0,1517	0,71	0,2611	1,03	0,3485
0,08	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,09	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,10	0,0398	0,42	0,1628	0,74	0,2703	1,06	0,3554
0,11	0,0438	0,43	0,1664	0,75	0,2734	1,07	0,3577
0,12	0,0478	0,44	0,1700	0,76	0,2764	1,08	0,3599
0,13	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,14	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,15	0,0596	0,47	0,1808	0,79	0,2852	1,11	0,3665
0,16	0,0636	0,48	0,1844	0,80	0,2881	1,12	0,3686
0,17	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,18	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,19	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3749
0,20	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,21	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,22	0,0871	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,23	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,24	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,25	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869
0,26	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3883
0,27	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,28	0,1103	0,6,	0,2257	0,92	0,3212	1,24	0,3925
0,29	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,30	0,1179	0,62	0,2324	0,94	0,3264		

0,31	0,1217	0,63	0,2357	0,95	0,3289		

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,26	0,3962	1,59	0,4441	1,92	0,4726	2,50	0,4938
1,27	0,3980	1,60	0,4452	1,93	0,4732	2,52	0,4941
1,28	0,3997	1,61	0,4463	1,94	0,4738	1,54	0,4945
1,29	0,4015	1,62	0,4474	1,95	0,4744	2,56	0,4948
1,30	0,4032	1,63	0,4484	1,96	0,4750	2,58	0,4951
1,31	0,4049	1,64	0,4495	1,97	0,4756	2,60	0,4953
1,32	0,4066	1,65	0,4505	1,98	0,4761	2,62	0,4956
1,33	0,4082	1,66	0,4515	1,99	0,4767	2,64	0,4959
1,34	0,4099	1,67	0,4525	2,00	0,4772	2,66	0,4961
1,35	0,4115	1,68	0,4535	2,02	0,4783	2,68	0,4963
1,36	0,4131	1,69	0,4545	2,04	0,4793	2,70	0,4965
1,37	0,4147	1,70	0,4554	2,06	0,4803	2,72	0,4967
1,38	0,4162	1,71	0,4564	2,08	0,4812	2,74	0,4969
1,39	0,4177	1,72	0,4573	2,10	0,4821	2,76	0,4971
1,40	0,4192	1,73	0,4582	2,12	0,4830	2,78	0,4973
1,41	0,4207	1,74	0,4591	2,14	0,4838	2,80	0,4974
1,42	0,4222	1,75	0,4599	2,16	0,4846	2,82	0,4976
1,43	0,4236	1,76	0,4608	2,18	0,4854	2,84	0,4977
1,44	0,4251	1,77	0,4616	2,20	0,4861	2,86	0,4979
1,45	0,4265	1,78	0,4625	2,22	0,4868	2,88	0,4980
1,46	0,4279	1,79	0,4633	2,24	0,4875	2,90	0,4981
1,47	0,4292	1,80	0,4641	2,26	0,4881	2,92	0,4982
1,48	0,4306	1,81	0,4649	2,28	0,4887	2,94	0,4984
1,49	0,4319	1,82	0,4656	2,30	0,4893	2,96	0,4985
1,50	0,4332	1,83	0,4664	2,32	0,4898	2,98	0,4986
1,51	0,4345	1,84	0,4671	2,34	0,4904	3,00	0,49865
1,52	0,4357	1,85	0,4678	2,36	0,4909	3,20	0,49931
1,53	0,4370	1,86	0,4686	2,38	0,4913	3,40	0,49966
1,54	0,4382	1,87	0,4693	2,40	0,4918	3,60	0,499841
1,55	0,4394	1,88	0,4699	2,42	0,4922	3,80	0,499928
1,56	0,4406	1,89	0,4706	2,44	0,4927	4,00	0,499968
1,57	0,4418	1,90	0,4713	2,46	0,4931	4,50	0,499997

1,58	0,4429	1,91	0,4719	2,48	0,4934	5,00	0,499997

Приложение 3

Значения функции $\chi^2_{\alpha;\nu}$; $P(\chi^2 \geq \chi^2_{\alpha;\nu}) = \alpha$

$\nu \backslash \alpha$	0,20	0,10	0,05	0,02	0,01	0,001
1	1,642	2,706	3,841	5,412	6,635	10,827
2	3,219	4,605	5,991	7,824	9,210	13,815
3	4,642	6,251	7,815	9,837	11,345	16,266
4	5,989	7,779	9,488	11,668	13,277	18,467
5	7,289	9,236	11,070	13,388	15,086	20,515
6	8,558	10,645	12,592	15,033	16,812	22,457
7	9,803	12,017	14,067	16,622	18,475	24,322
8	11,030	13,362	15,507	18,168	20,090	26,125
9	12,242	14,684	16,919	19,679	21,666	27,877
10	13,442	15,987	18,307	21,161	23,209	29,588
11	14,631	17,275	19,675	22,618	24,725	31,264
12	15,812	18,549	21,026	24,054	26,217	32,909
13	16,985	19,812	22,362	25,472	27,688	34,528
14	18,151	21,064	23,685	26,683	29,141	36,123
15	19,311	22,307	24,996	28,259	30,578	37,697
16	20,465	23.542	26,296	29,633	32,000	39,252
17	21,615	24,769	27,587	30,995	33,409	40,790
18	22,760	25,989	28,869	32,346	34,805	42,312
19	23,900	27,204	30,144	33,687	36,191	43,820
20	25,038	28,412	31,410	35,020	37,566	45,315
21	26,171	29,615	32,671	36,343	38,932	46,797
22	27,301	30,813	33,924	37,659	40,289	48,268
23	28,429	32,007	35,172	38,968	41,638	49,728
24	29,553	33,196	36,415	40,270	42,980	51,179
25	30,675	34,382	37,652	41,566	44,312	52,620
26	31,795	35,563	38,885	42,856	45,642	54,052
27	32,912	36,741	40,113	44,140	46,963	55,476
28	34,027	37,916	41,337	45,419	48,278	56,893
29	35,139	39,087	42,557	46,693	49,588	58,302
30	36,250	40,256	43,773	47,962	50,892	59,703

Приложение 4

Распределение Стьюдента

Значения $t_{\alpha;\nu}$ удовлетворяют условию $P(t \geq t_{\alpha;\nu}) = \int\limits_{t_{\alpha;\nu}}^{\infty} S(t,\nu)dt = \alpha$

$\nu \backslash \alpha$	0,40	0,30	0,20	0,10	0,05	0,025	0,010	0,005	0,001	0,0005
1	0,325	0,727	1,376	3,078	6,314	12,71	31,82	63,66	318,3	636,6
2	0,289	0,617	1,061	1,886	2,920	4,303	6,965	9,925	22,33	32,60
3	0,277	0,584	0,978	1,638	2,353	3,182	4,541	5,841	10,22	12,94
4	0,271	0,569	0,941	1,533	2,132	2,776	3.747	4,604	7,173	8,610
5	0,267	0,559	0,920	1,476	2,015	2,571	3,365	5,032	5,893	6,859
6	0,265	0,553	0,906	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,263	0,549	0,896	1,415	1,895	2,365	2,998	3,499	4,785	5,405
8	0,262	0,546	0,889	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,261	0,543	0,883	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,260	0,542	0,879	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,260	0,540	0,876	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,259	0,539	0,873	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,259	0,538	0,870	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,258	0,537	0,868	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	0.258	0,536	0,866	1.341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,258	0,535	0,865	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,257	0,534	0,863	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,257	0,534	0,862	1,330	1,734	2,101	2,552	2,878	3,611	3.922
19	0,257	0,533	0,861	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,257	0,533	0,860	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	0,257	0,532	0,859	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	0,256	0,532	0,858	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	0,256	0,532	0,858	1,319	1,714	2,069	2,500	2,807	3,485	3,767
24	0,256	0,531	0,857	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	0,256	0,531	0,856	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	0,256	0,531	0,856	1,315	1,7j6	2,056	2,479	2,779	3,435	3,707
27	0,256	0,531	0,855	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	0,256	0,530	0,855	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	0,256	0,530	0,854	1,311	1,699	2,045	2,462	7,756	3,396	3,659
30	0,256	0,530	0,854	1,310	1,697	2,042	2,457	2,750	3,385	3,646
40	0,255	0,529	0,851	1,303	1,684	2,021	2,423	2,704	3,307	3,551
50	0,255	0,528	0,849	1,298	1,676	2,009	2,403	2,678	3,262	3,495
60	0,254	0,527	0,848	1,296	1,671	2,000	2,390	2,660	3,232	3,460
80	0,254	0,527	0,846	1,292	1,664	1,990	2,374	2,639	3,195	3,415
100	0,254	0,526	0,845	1,290	1,660	1,984	2,365	2,626	3,174	3,389
200	0,254	0,525	0,843	1,286	1,653	1,972	2,345	2,601	3,131	3,339
500	0,253	0,525	0,842	1,283	1,648	1.965	2,334	2,586	3,106	3,310
∞	0,253	0,524	0,842	1,282	1,645	1,960	2,326	2,576	3,090	3,291

Приложение 5

Значения функции $t_{\gamma;n}$: $\overline{x}_{\text{в}} - t_{\gamma;n}\dfrac{s}{\sqrt{n}} < a < \overline{x}_{\text{в}} + t_{\gamma;n}\dfrac{s}{\sqrt{n}}$

$n \backslash \gamma$	0,95	0,99	0,999	$n \backslash \gamma$	0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

Приложение 6

Значения коэффициентов q_1 и q_2; $q_1 S < \sigma < q_2 S$

г	0,99		0,98		0,95		0,00	
= n – 1	q_1	q_2	q_1	q_2	q_1	q_2	q_1	q_2
1	0,356	15,0	0,388	79,8	0,446	31,9	0,510	15,9
2	0,434	14,1	0,466	9,97	0,521	6,28	0,578	4,40
3	0,483	6,47	0,514	5,11	0,566	3,73	0,620	2,92
4	0,519	4,39	0,549	3,67	0,599	2,87	0,649	2,37
5	0,546	3,48	0,576	3,00	0,624	2,45	0,672	2,090
6	0,569	2,98	0,597	2,62	0,644	2,202	0,690	1,916
7	0,588	2,66	0,616	2,377	0,661	2,035	0,705	1,797
8	0,604	2,440	0,631	2,205	0,675	1,916	0,718	1,711
9	0,618	2,277	0,644	2,076	0,688	1,826	0,729	1,645
10	0,630	2,154	0,656	1,977	0,699	1,755	0,739	1,593
11	0,641	2,056	0,667	1,898	0,708	1,698	0,748	1,550
12	0,651	1,976	0,676	1,833	0,717	1,651	0,755	1,515
13	0,660	1,910	0,685	1,779	0,725	1,611	0,762	1,485
14	0,669	1,854	0,693	1,733	0,732	1,577	0,769	1,460
15	0,676	1,806	0,700	1,694	0,739	1,548	0,775	1,437
16	0,683	1,764	0,707	1,659	0,745	1,522	0,780	1,418
17	0,690	1,727	0,713	1,629	0,750	1,499	0,785	1,400
18	0,696	1,695	0,719	1,602	0,756	1,479	0,790	1,385
19	0,702	1,668	0,725	1,578	0,760	1,460	0,794	1,370
20	0,707	1,640	0,730	1,556	0,765	1,414	0,798	1,358
21	0,712	1,617	0,734	1,536	0,769	1,429	0,802	1,346
23	0,722	1,576	0,743	1,502	0,777	1,402	0,809	1,326
24	0,726	1,558	0,747	1,487	0,781	1,391	0,812	1,316
25	0,730	1,541	0,751	1,473	0,784	1,380	0,815	1,308
26	0,734	1,526	0,755	1,460	0,788	1,371	0,818	1,300
27	0,737	1,512	0,758	1,448	0,791	1,361	0,820	1,293
29	0,744	1,487	0,765	1,426	0,796	1,344	0,825	1,279
30	0,748	1,475	0,768	1,417	0,799	1,337	0,828	1,274
40	0,774	1,390	0,792	1,344	0,821	1,279	0,847	1,228
50	0,793	1,336	0,810	1,297	0,837	1,243	0,861	1,199
60	0,808	1,299	0,824	1,265	0,849	1,217	0,871	1,179
70	0,820	1,272	0,835	1,241	0,858	1,198	0,879	1,163
80	0,829	1,250	0,844	1,222	0,866	1,183	0,886	1,151
90	0,838	1,233	0,852	1,207	0,873	1,171	0,892	1,141
100	0,845	1,219	0,858	1,195	0,878	1,161	0,897	1,133
200	0,887	1,15	0,897	1,13	0,912	1,11	0,925	1,09

10. РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Раздел. Теория вероятностей

1. Элементы комбинаторики. Перестановки, размещения и сочетания.

2. Пространство элементарных событий, алгебра событий. Относительная частота и вероятность события. Аксиоматическое и классическое определения вероятности. Теоремы сложения и умножения.

3. Условная вероятность. Зависимые и независимые события. Формула полной вероятности, формулы Байеса.

4. Последовательность независимых испытаний. Схема Бернулли. Предельные теоремы Муавра-Лапласа и Пуассона. Случайные величины. Функция распределения случайной величины, ее свойства. Дискретные случайные величины. Непрерывные случайные величины, функция и плотность распределения.

5. Математическое ожидание, дисперсия и среднее квадратическое отклонение случайной величины. Моменты случайной величины.

6. Основные законы распределения. Биномиальный закон распределения, закон распределения Пуассона, равномерный закон распределения, показательный закон распределения, нормальный закон распределения. Функция Лапласа, правило трёх сигм.

7. Закон больших чисел и предельные теоремы. Неравенства Чебышева. Теоремы Чебышева и Бернулли. Центральная предельная теорема Ляпунова.

8. Системы случайных величин (случайные векторы). Функция и плотность распределения систем двух случайных величин, их свойства. Вероятность попадания случайной точки в заданную область. Зависимые и независимые случайные величины. Числовые характеристики систем случайных величин. Начальные и центральные моменты. Корреляционный момент. Коэффициент корреляции.

Раздел. Математическая статистика

1.Задачи математической статистики. Генеральная совокупность и выборка. Статистические ряды. Числовые характеристики выборки. Полигон и гистограмма.

2.Основные статистические распределения: χ^2 -распределение, распределение Фишера и Стьюдента.

3. Статистические оценки параметров. Точечные и интервальные оценки. Методы нахождения точечных оценок: метод моментов Пирсона, метод максимального правдоподобия, метод наименьших квадратов. Интервальные оценки: доверительный интервал, уровень значимости. Доверительный интервал для математического ожидания при известной и неизвестной дисперсии.

4. Статистическая проверка гипотезы. Ошибки первого и второго родов. Проверка гипотезы о равенстве математических ожиданий. Критерии

согласия Неймана-Пирсона, χ^2 -Пирсона, А. Н. Колмогорова.

5. Понятие о регрессионном и корреляционном анализе. Линейная регрессия. Определение параметров линейной регрессии методом наименьших квадратов.

6. Нелинейная регрессия. Корреляционное отношение.