Российский государственный педагогический университет
им. А.И. Герцена

Чурилова М.Ю.

Теория вероятностей

Оглавление

Предисловие

Данное пособие по теории вероятностей предназначено для студентов различных естественно-научных и экономических специальностей РГПУ им. А. И. Герцена. Оно состоит из трех основных частей, посвященных элементам комбинаторики, случайным событиям и случайным величинам. В каждом разделе теория излагается на уровне, соответствующем государственному образовательному стандарту дисциплины «Математика (теория вероятностей)».

Большое место в настоящем издании отводится методическим указаниям к решению задач по теории вероятностей. Большинство задач — авторские, некоторые задачи заимствованы из различных сборников [1, 3, 5, 6, 8–10]. В каждом разделе пособия помещены задачи для самостоятельного решения. Они разбиты на два уровня сложности: первый и второй. Для решения задач первого уровня требуется знание основных определений и теорем, приведенных в теоретической части (конспекте лекций). Для решения задач второго уровня требуется знание приемов, показанных в пособии. Задачи первого уровня снабжены подсказками и ответами. Ответы к большинству задач второго уровня, как правило, отсутствуют, поскольку автор считает более важным навыком подбор подходящей вероятностной модели, чем получение некоего числа, совпадающего с указанным в книге. Предполагается, что задачи второго уровня будут обсуждаться с преподавателем, на семинарских занятиях.

Глава 1. Элементы комбинаторики

Введение

В различных областях человеческой деятельности приходится рассматривать наборы предметов разной природы и подсчитывать количество таких наборов, которые удовлетворяют требуемым условиям. Задачи этого типа возникают при изучении структуры вещества, при подборе лекарственных средств, при составлении карты посевов, при размещении промышленных объектов и так далее. Раздел математики, посвященный решению задач выбора и расположения элементов некоторого, обычно конечного, множества в соответствии с заданными правилами, называется *комбинаторикой*. Комбинаторные методы играют важную роль не только в практической, но и в чисто научной деятельности человека: они широко используются в различных разделах математики, в частности, при решении задач теории вероятностей. В комбинаторике по каждой конкретной задаче составляется своя математическая модель, описывающая данное множество и те наборы элементов, которые надо изучить. Вместо слова «наборы» обычно употребляется термин *выборки*. Две конкретные выборки могут отличаться друг от друга либо составом (то есть входящими в них элементами), либо порядком их расположения. Простейшими видами выборок из элементов конечного множества S являются так называемые «стандартные выборки». Так же, как в математическом анализе понятие «элементарная функция» раскрывается с помощью перечисления и описания конкретных функций, так и в комбинаторике стандартные выборки определяются описанием конкретных моделей. К стандартным выборкам относятся сочетания и кортежи. Им будут посвящены отдельные разделы.

1.1. Теорема сложения. Разбиение множества на классы

Здесь мы хотим особо выделить теоремы, на основе которых будут далее проводиться подсчеты количества стандартных выборок. Чтобы облегчить понимание этих теорем, сначала будем рассматривать задачи.

Задача 1.1. *В лаборатории обследовались 20 проб колодезной воды. Первый лаборант проверил эти пробы на вредные бактерии. Оказалось, что только пять из 20 проб не содержат вредных бактерий. Второй лаборант исследовал все 20 проб на вредные химические примеси. Он обнаружил такие примеси в 14 пробах из 20. Среди этих 14 проб ровно 12 содержат также и вредные бактерии. Сколько проб из 20 содержат либо вредные бактерии, либо вредные химические примеси?*

Вычтем из 20 проб те пять, в которых нет бактерий. Каждая из оставшихся 15 проб содержит бактерии. Известно, что химические примеси есть в 14 пробах. Если сложить 15 и 14, то получим 29 проб. В этой сумме 12 проб, содержащих бактерии и химию одновременно, учтены ровно два раза, следовательно, если их количество вычесть из 29, получится количество проб, содержащих либо бактерии, либо химию: $29 - 12 = 17$.

Задача 1.2. *Пусть в условиях первой задачи проведено дополнительное обследование на вирусы. Оказалось, что 10 проб из 20 содержат вирусы, семь проб — вирусы и бактерии, восемь — вирусы и химию, шесть — все три вредные примеси. Сколько проб не содержат ни одной примеси?*

Мы уже знаем, что 17 проб содержат бактерии или химию. Количество проб, содержащих не только вирусы, равно $7 + 8 - 6 = 9$. Следовательно, число проб, содержащих только вирусы, — одна. Число проб, содержащих вредные примеси, равно $17 + 1 = 18$. Проб без примесей — две $(20 - 18)$.

Сформулируем теперь в общем виде, как должны быть связаны между собой количества элементов, содержащихся в объединении и пересечении конечных множеств. Для этого обозначим через $M(S)$ число элементов произвольного конечного множества S.

Теорема 1.1 (теорема сложения). *Пусть S_1 и S_2 — конечные множества. Тогда их пересечение содержит следующее число элементов:*

$$M(S_1 \cup S_2) = M(S_1) + M(S_2) - M(S_1 \cap S_2).$$

Пусть S_3 — также конечное множество. Тогда

$$M(S_1 \cup S_2 \cup S_3) = M(S_1) + M(S_2) + M(S_3) - M(S_1 \cap S_2) - \\ - M(S_1 \cap S_3) - M(S_2 \cap S_3) + M(S_1 \cap S_2 \cap S_3).$$

Здесь символ $\cup$ означает объединение множеств, а символ $\cap$ означает их пересечение (общие элементы). Аналогичные соотношения можно доказать для любого конечного числа множеств. (Доказательство опускаем.)

Следствие из теоремы 1.1. Если множества $S_1, S_2, \ldots, S_k$ попарно не пересекаются, тогда

$$M(S_1 \cup S_2 \cup S_3 \cup \cdots \cup S_k) = M(S_1) + M(S_2) + M(S_3) + \cdots + M(S_k).$$

Определение классификации. Представление данного множества S в виде объединения некоторого (не более, чем счетного) количества попарно не пересекающихся подмножеств S_1, S_2, $\ldots$, S_k называется *классификацией* S (или разбиением его на классы). При этом подмножества S_1, S_2, $\ldots$, S_k называются классами, на которые разбито множество S.

Таким образом, следствие из теоремы 1.1 утверждает, что для конечного множества S количество его элементов можно найти, сложив количества элементов во всех классах, на которые разбито S произвольным образом. Обычно элементы данного множества классифицируются по каким-то признакам. Самая простая классификация проводится по одному признаку, причем в один класс помещаются элементы, имеющие указанный признак, а в другой — не имеющие такого признака.

Задача 1.3. *Дан список государств: Беларусь, Замбия, Франция, Германия, Перу, Гватемала. Требуется разбить список на классы по признаку: страна находится в северном полушарии.*

Очевидно, получаем два класса. В северном полушарии находятся государства: Беларусь, Франция, Германия, Гватемала. Остальные два государства находятся в южном полушарии. Заметим, что не всякий список государств можно разбить на два класса, учитывая расположение в северном или южном полушарии. Действительно, к какому классу отнести, например, Кению и другие страны, лежащие по обе стороны экватора? Если добавить в исходный список такие страны, то придется разбивать этот список на три класса, поместив в первый класс те страны, которые лежат целиком в северном полушарии, во второй — те, что лежат целиком в южном полушарии, а в третий класс — все остальные страны.

Если элементы множества классифицируются по двум признакам (обозначим их, условно, A и B), то получаем четыре класса. Это класс AB, где есть оба признака; два класса, где имеется только один признак: $A\bar{B}$ (A есть и нет B) или $\bar{A}B$ (нет A и есть B), а также класс $\bar{A}\bar{B}$, где нет ни одного из двух признаков.

Задача 1.4. *Список тех же государств разбить по двум признакам: страна лежит в северном полушарии (A), страна лежит в восточном полушарии (B).*

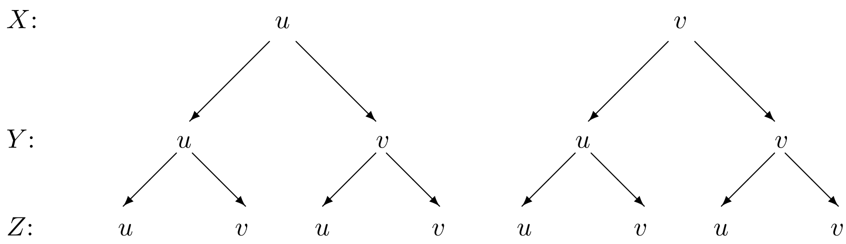
К классу AB относятся государства, лежащие в северном и в восточном полушариях одновременно: Беларусь, Франция, Германия. К классу $A\bar{B}$ относится Гватемала, находящаяся на юге Северной Америки. К классу $\bar{A}B$ относится Замбия (юг Африки). К классу $\bar{A}\bar{B}$ относится Перу (Южная Америка). Если в список добавить Россию, лежащую в западном и восточном полушариях одновременно, то разбиение списка на четыре класса не состоится. Придется поместить Россию в отдельный класс.

Классификации широко используются в комбинаторике для подсчета количества возможных исходов опыта.

Задача 1.5. *Рассмотрим такой опыт: каждую из трех монет, X, Y и Z, независимо от других укладывают в один из двух карманов, u или v. Разобьем все способы размещения монет на три класса: S_1 — пуст карман u, S_2 — пуст карман v, S_3 — оба кармана не пусты. Требуется определить, сколько способов в каждом классе. (Способы различаются только выбором кармана для каждой монеты.)*

Очевидно, что если пуст карман u, то все монеты лежат в v. Поэтому, $M(S_1) = 1$. Аналогично $M(S_2) = 1$. Чтобы найти $M(S_3)$, определим число всех способов и отнимем от него $M(S_1)$ и $M(S_2)$. Все способы будем классифицировать по трем признакам: карман для X, карман для Y, карман для Z. Получим сначала два класса по положению X. Каждый из этих двух классов разобьем еще на два более мелких класса по размещению Y. Затем каждый из получившихся четырех классов разобьем еще на два по положению Z. В результате имеем восемь классов, каждый из которых содержит ровно один способ расположения трех монет по карманам.

Составим схему, поясняющую наши подсчеты:



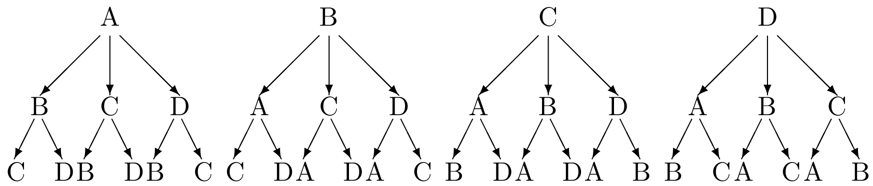
Таким образом, $M(S_3) = 8 - 2 = 6$.

1.2. Кортежи. Теорема умножения

Рассмотрим еще две задачи.

Задача 1.6. *Четыре вида молекул — A, B, C, D — могут соединяться по три молекулы в ряд, то есть образовывать упорядоченные тройки, причем в тройке все молекулы различны. Сколько разных троек может образоваться при этих условиях?*

Чтобы перебрать все варианты соединений и ничего не пропустить, составим схему:



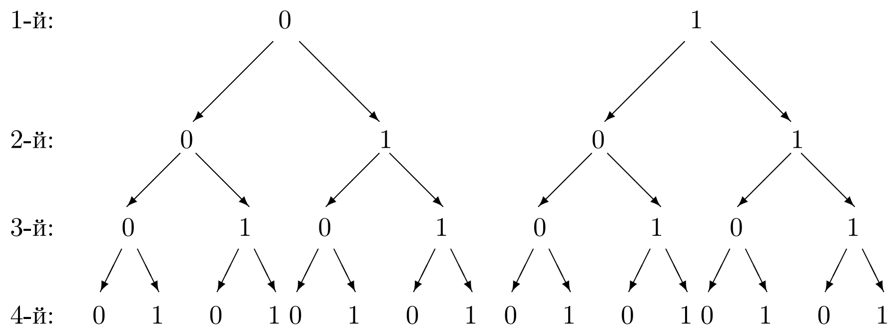
Рассмотрите внимательно эту схему. В первой строке — все варианты первой молекулы. Она определяет четыре класса по признаку, какова первая молекула. Во второй строке указано, какой может быть вторая молекула, следующая за первой. Каждый из четырех классов разбит еще на три. Получилось $4 \cdot 3 = 12$ классов. В третьей строке указаны все варианты для третьей молекулы, при условии, что первая и вторая заданы.

Каждый из 12 классов разбит еще на два. Всего $12 \cdot 2 = 24$ класса, по одному способу в каждом.

По схеме не трудно выписать все тройки: $ABC, ABD, ACB, \ldots, DCB$. Этих троек 24 штуки. Их количество подсчитывается умножением: $4 \cdot 3 \cdot 2$.

Задача 1.7. *В ЭВМ числа представлены в виде упорядоченных наборов из нулей и единиц. Рассмотрим наборы, состоящие из четырех символов (каждый — либо нуль, либо единица). Сколько существует таких наборов?*

Составим схему, похожую на предыдущую:



Для первого символа имеются две возможности: 0 или 1. Для второго символа — две такие же. Получается $2 \cdot 2 = 4$ пары: 00, 01, 10, 11. К каждой паре двумя способами можно приписать третий символ. Получим $4 \cdot 2 = 8$ троек. К каждой тройке двумя способами можно приписать четвертый символ. В результате этого получится $8 \cdot 2 = 16$ упорядоченных четверок.

Обобщением решений задач 1.5, 1.6 и 1.7 является

Теорема 1.2 (теорема умножения). *Пусть требуется задать элементы упорядоченного набора из k компонент. Будем последовательно определять первую, вторую, ..., k-ю компоненты. Если первая компонента может быть задана C_1 способами, вторая — C_2 способами и т.д., последняя — C_k способами, то общее число определенных таким образом наборов равно*

$$C_1 \cdot C_2 \cdot C_3 \ldots C_k.$$

Строгое доказательство теоремы опускаем.

Определение кортежа. В математике упорядоченные наборы из k элементов, среди которых могут быть и одинаковые, называются *кортежами длины* k. Число k называется *длиной кортежа*, а элементы — его *компонентами*. Номера компонент иногда называют *местами кортежа*.

Рассмотрим теперь несколько примеров применения теорем сложения и умножения.

Задача 1.8. *Сколько чисел от единицы до тысячи делятся на два или на три?*

Обозначим через S_1 множество чисел от 1 до 1000, делящихся на два. Обозначим через S_2 множество чисел от 1 до 1000, делящихся на три. В пересечении S_1 и S_2 содержатся числа, делящиеся на шесть. Поэтому по теореме сложения $M(S_1 \cup S_2) = M(S_1) + M(S_2) - M(S_1 \cap S_2) = 500 + 333 - 166 = 667$.

Задача 1.9. *Сколько существует трехзначных натуральных чисел, у которых первая цифра нечетная?*

Обозначим трехзначное число так: (v_1, v_2, v_3), v_1 — первая цифра, v_2 — вторая, v_3 — третья. Число v_1 принадлежит множеству $\{\,1,\ 3,\ 5,\ 7,\ 9\,\}$, а числа v_2 и v_3 — множеству $\{\,0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9\,\}$. Поэтому, по теореме умножения, количество всех трехзначных натуральных чисел с первой нечетной цифрой равно $5 \cdot 10 \cdot 10 = 500$.

Задача 1.10. *Сколько существует трехзначных натуральных чисел, у которых ровно одна нечетная цифра?*

Обозначим трехзначное число: $(v_1,\ v_2,\ v_3)$. Все числа, удовлетворяющие условию задачи, разобьем на три непересекающихся множества: S_1, S_2, S_3. В S_1 поместим все трехзначные числа, у которых нечетна цифра v_1, в S_2 — числа с нечетной цифрой v_2, в S_3 — с нечетной цифрой v_3. По теореме умножения $M(S_1) = 5 \cdot 5 \cdot 5$ (v_1 из $\{1, 3, 5, 7, 9\}$, v_2 и v_3 из $\{0,\ 2,\ 4,\ 6,\ 8\}$). Поскольку первая цифра не равна нулю, $M(S_2) = 4 \cdot 5 \cdot 5$ (v_1 из $\{2,\ 4,\ 6,\ 8\}$, v_2 из $\{1, 3, 5, 7, 9\}$, v_3 из $\{0,\ 2,\ 4,\ 6,\ 8\}$). Аналогично $M(S_3) = 4 \cdot 5 \cdot 5$. По теореме сложения $M(S_1 \cup S_2 \cup S_3) = 325$.

Задача 1.11. *Сколько существует трехзначных натуральных чисел, у которых есть хотя бы одна нечетная цифра?*

Определим сперва общее количество трехзначных чисел:

$$M(S) = 9 \cdot 10 \cdot 10$$

Все эти числа (множество S) разобьем на два непересекающихся множества: S_1 — числа, удовлетворяющие условию задачи, S_2 — не удовлетворяющие (то есть из четных цифр). По теореме умножения $M(S_2) = 4 \cdot 5 \cdot 5 = 100$. По теореме сложения $M(S) = M(S_1) + M(S_2)$. Поэтому $M(S_1) = 900 - 100 = 800$.

Задача 1.12. *Из 10 членов правления АО «Рога и копыта» надо выбрать трех на должности председателя, заместителя председателя, секретаря. Сколькими способами это можно сделать?*

Составим модель задачи. Пусть множество S из 10 элементов представляет наше правление АО. Будем из этих 10 элементов выбирать всевозможные тройки и упорядочивать их по должностям: первый — председатель, второй — заместитель, третий — секретарь. Способы различаются либо составом, либо распределением ролей. Полученные выборки являются кортежами длины 3, на места которых мы должны поместить элементы S без повторений. По теореме умножения получаем, что число кортежей равно $10 \cdot 9 \cdot 8 = 720$.

Задача 1.13. *Сколькими способами можно два разных кольца надеть на четыре (без большого) пальца левой руки так, чтобы ни на одном пальце не было двух колец сразу?*

Каждому кольцу мы должны сопоставить единственный палец, разным — разные. Зададим каждый способ двухместным кортежем: первое место — палец для первого кольца, второе — для второго. Значит, способов столько, сколько двухместных кортежей из четырех элементов без повторений. $4 \cdot 3 = 12$.

Задача 1.14. *Сколькими способами пять человек могут выстроиться в одну очередь?*

Каждому человеку мы должны сопоставить номер от 1 до 5, единственный и неповторимый. Все номера будут использованы. Следовательно, ответ: $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$.

Задача 1.15. *Сколькими способами можно раздать четыре разные конфеты четырем детям, по одной конфете каждому?*

Пронумеруем конфеты: 1, 2, 3, 4. Каждой сопоставим «своего» ребенка. Всего таких способов будет $4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24$.

Задача 1.16. *Сколькими способами три человека могут выйти из лифта в 9-этажном доме, если все они вошли в лифт на первом этаже?*

Пусть G — множество людей (например, $\{A, B, C\}$), S — множество этажей: $\{2, 3, 4, ..., 9\}$. Каждому человеку надо сопоставить номер этажа. Значит, всего будет $8 \cdot 8 \cdot 8 = 512$ способов.

1.3. Сочетания

В комбинаторике всевозможные k-элементные подмножества n-элементного множества S называются *сочетаниями* из n по k элементов.

Количество сочетаний из n по k $(n \geqslant k)$ обозначают

$$C_n^k.$$

В отличие от кортежей сочетания не являются упорядоченными выборками. Пусть $S = \{\,A, B, C, D\,\}$, $n = 4$. Выпишем все сочетания из n по k.

При $k = 0$ получаем одно пустое множество. $C_4^0 = 1$, $C_n^0 = 1$.

При $k = 1$ получаем четыре одноэлементных множества: $\{A\}$, $\{B\}$, $\{C\}$, $\{D\}$. $C_4^1 = 4$, $C_n^1 = n$.

При $k = 2$ получаем шесть пар: $\{A, B\}$, $\{A, C\}$, $\{A, D\}$, $\{B, C\}$, $\{B, D\}$, $\{C, D\}$. $C_4^2 = 6$.

При $k = 3$ имеем четыре тройки: $\{A, B, C\}$, $\{A, B, D\}$, $\{A, C, D\}$, $\{B, C, D\}$. $C_4^3 = 4$.

Троек столько же, сколько одноэлементных множеств, так как они являются их дополнениями до S.

При $k = 4$ получаем одно множество $S : \{A, B, C, D\}$. $C_4^4 = 1$, $C_n^n = 1$.

Теорема 1.3.

$$C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\ldots(n-k+1)}{1 \cdot 2 \ldots k}.$$

Здесь $0! = 1$.

Можно доказать эту теорему, основываясь на том, что каждое сочетание определяет $k!$ различных кортежей длины k, составленных из элементов S без повторений.

$$C_n^k\,k! = n(n-1)\ldots(n-k-1).$$

Задача 1.17. *Сколько существует способов составить подарочный набор из трех предметов, если всего имеется восемь различных сувениров?*

Все сувениры можно принять за множество S из восьми элементов. По условию надо выбрать из него трехэлементное подмножество (порядок не важен). Значит получится

$$\frac{8!}{3!\,5!} = \frac{8\cdot 7\cdot 6}{1\cdot 2\cdot 3} = 56 \quad \text{способов.}$$

Задача 1.18. *В языке племени «АББА» всего две буквы: А и Б. Сколько слов из трех А и двух Б может образовать это племя?*

Изобразим каждое слово кортежем из пяти компонент. Например: (А, А, А, Б, Б), (А, Б, А, Б, А), (Б, А, Б, А, А), ... Каждый кортеж определяется тем, какие из пяти мест занимают буквы Б (остальные — А). Значит, слов будет столько, сколькими способами можно выбрать два места из пяти, то есть

$$C_5^2 = \frac{5\cdot 4}{1\cdot 2} = 10 \quad \text{(слов).}$$

Задача 1.19. *Сколько различных семизначных чисел можно составить из двух единиц, четырех двоек и одной тройки?*

Каждое число представим как кортеж из цифр 1, 2, 3 с заданным числом повторений, то есть кортежем состава $\{1(2), 2(4), 3(1)\}$. Здесь в скобках указано число повторений для каждой цифры. Таких кортежей будет $C_7^2 \cdot C_5^1 = 105$. Действительно, разобьем все кортежи на классы по расположению единиц, затем каждый класс разобьем на классы по расположению тройки на одном из оставшихся пяти мест. Места под двойки определяются однозначно.

1.4. Задания для самостоятельной работы

1.4.1. Первый уровень сложности

1. Сколько существует пятизначных чисел, все цифры которых различны и принадлежат множеству $\{1, 2, 3, 4, 5\}$?
ОТВЕТ: 5!=120

2. Сколькими способами можно рассадить на восемь мест в кинозале компанию из восьми человек (по одному на место)?
ОТВЕТ: 8!

3. Сколькими способами можно из 20 человек выбрать 10 для игры в КВН?
ОТВЕТ: C_{20}^{10}

4. Реплики Ивана Ивановича Петушкова состоят из двух частей. Сначала он говорит либо «грубо говоря», либо «мягко выражаясь», а затем одну из фраз: «это маразм», «я балдею», «это просто экстаз». Сколько различных по форме реплик он может выдать?
ОТВЕТ: 6

5. Сколькими способами можно 12 разных монет разложить в три разные копилки?
ОТВЕТ: 3^{12}

6. Сколько четырехзначных чисел, кратных пяти, можно составить из цифр 1, 2, 3, 4, 5 с повторениями?
ОТВЕТ: $5 \cdot 5 \cdot 5 \cdot 1$

7. Сколько мелодий можно сделать из «Чижика-пыжика», включая его самого, если переставлять между собой его ноты: си, соль, си, соль, до, си, ля? (Длительности нот и паузы не учитываем.)
ОТВЕТ: $C_7^3 \cdot C_4^2 \cdot C_2^1$

8. Сколько 10-символьных цепочек можно составить из символов «точка», «тире»?
ОТВЕТ: $2^{10} = 1024$

9. Тридцать три богатыря поделили по жребию восемь одинаковых бутылок «Пепси». Каждому не более одной. Сколько способов?
ОТВЕТ: C_{33}^8

10. Сколькими способами семь мужчин могут пригласить на танец семь женщин (танцуют все)?

ОТВЕТ: 7!

11. Сколькими способами семь кавалеров и десять дам могут образовать семь пар (м — ж) для одновременного вальсирования?
ОТВЕТ: $10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4$

12. Дом раскрашивается в три разных цвета: первым — стены, вторым — крыша, третьим — рамы и двери. В наличии имеется пять цветов красок. Сколько вариантов окрасить дом?
ОТВЕТ: 60

13. В столовой предлагается три первых, четыре вторых и два третьих блюда. Сколько разных комплексных обедов можно составить?
ОТВЕТ: 24

14. Сколькими способами 10 пассажиров могут разместиться в двухвагонном трамвае?
ОТВЕТ: 1024

15. Из 20 студентов два не знают математику и не любят детей, 11 знают математику, 12 любят детей. Сколько студентов годятся в учителя математики (знают математику и любят детей)?
ОТВЕТ: $20 - 2 = 11 + 12 - x$

1.4.2. Второй уровень сложности

1. Из цифр 1, 2, 3, 4 составлены всевозможные трехзначные числа, все три цифры которых различны. Сколько таких чисел кратны трем?
ПОДСКАЗКА. Смотрите сумму цифр в числе.

2. У трех поросят есть четыре разные чашки, пять разных блюдечек и шесть разных ложечек. Сколькими способами они могут раздать себе посуду для чаепития?

3. Группа людей состоит из пяти мужчин и четырех женщин. Сколькими способами можно образовать подгруппу из двух мужчин и двух женщин?
ПОДСКАЗКА. Сочетания, произведение.

4. Сколько пятикомпонентных кортежей, состоящих из нулей и (или) единиц, содержат по крайней мере один нуль?
ПОДСКАЗКА. От общего числа кортежей отнять число кортежей без нулей.

5. Сколько пятикомпонентных кортежей, состоящих из нулей и (или) единиц, содержат по крайней мере две единицы?

6. Слово «гамма» разрезано на буквы. Сколько разных «слов», считая исходное, можно составить из всех этих букв?

ПОДСКАЗКА. Выберите из пяти мест в «слове» место под букву «г». Из оставшихся четырех мест выберите два места под буквы «а». Остальные места займут буквы «м».

7. Сколькими способами можно группу из десяти человек разбить на две (неупорядоченная пара) команды по пять человек?

8. Имеются шесть различных по цвету шаров и три коробки с номерами 1, 2, 3. Сколькими способами можно разложить шары так, чтобы в первой коробке был один шар, во второй — два, в третьей — три?

ПОДСКАЗКА. Сравните с заданием 1.6. (Выбираются шары для коробок.)

9. На окружности отметили сто точек. Их попарно соединили отрезками. Сколько получилось отрезков?

10. Сколько существует различных шестизначных чисел, состоящих из трех единиц и трех двоек, где никакие три одинаковые цифры не стоят подряд?

ПОДСКАЗКА. Из общего количества таких чисел отнимите количество чисел, где три единицы или три двойки стоят подряд.

11. Сколько натуральных чисел, меньших, чем миллион, состоят из цифр, принадлежащих множеству $\{1, 2, 3\}$?

ПОДСКАЗКА. Теорема умножения. Сумма.

12. В коробке пять разных хрямзиков и восемь разных жуж. Сколькими способами можно составить себе подарок из трех жуж и двух хрямзиков?

ПОДСКАЗКА. Сочетания, произведение.

13. В колоде 36 карт. Раздаются три. Сколько случаев получить ровно одного туза среди розданных карт?

14. Два разбойника делят пять разных монет. Сколько существует дележей, при которых каждый получит не меньше двух монет?

ПОДСКАЗКА. Сочетания, сумма.

15. Круг разбит на семь равных секторов. Они выкрашены в семь разных цветов. Сколько раскрасок не совместимы поворотом?

ПОДСКАЗКА. Надо фиксировать один цвет на каком-то месте. Для остальных мест применить теорему умножения.

16. Сколько натуральных чисел, меньших, чем миллион, состоят из цифр, принадлежащих множеству $\{8, 9\}$?

17. Куб, все грани которого окрашены, распилен на 1000 одинаковых кубиков плоскостями, параллельными граням. Сколько маленьких кубиков не имеют ни одной окрашенной грани?

ПОДСКАЗКА. Длина ребра меньше на два.

18. Куб, все грани которого окрашены, распилен на 1000 одинаковых кубиков плоскостями, параллельными граням. Сколько маленьких кубиков имеют ровно две окрашенные грани?

ПОДСКАЗКА. Вдоль ребер.

19. Куб, все грани которого окрашены, распилен на 1000 одинаковых кубиков плоскостями, параллельными граням. Сколько маленьких кубиков имеют ровно одну окрашенную грань?

ПОДСКАЗКА. На гранях.

20. Пять мальчиков поделили пять одинаковых конфет так, что ровно одному мальчику не досталось конфеты. Сколько способов?

ПОДСКАЗКА. Произведение; одна «лишняя».

21. Пять мальчиков поделили пять одинаковых конфет так, что только одному мальчику Васе не досталось конфеты. Сколько существует способов?

22. Четыре студента сдали экзамен не меньше чем на три. Сколько вариантов распределения оценок?

23. В продажу поступили открытки трех видов и марки двух видов. Сколькими способами можно составить набор из двух открыток с приклеенными марками?

24. Пять мальчиков поделили произвольным образом пять разных игрушек так, что только одному мальчику ничего не досталось. Сколько способов?

25. В группе людей каждый владеет по крайней мере одним из трех языков: английским, немецким, французским. Шесть человек владеют английским, шесть — немецким, семь — французским, четыре — английским и немецким, три — немецким и французским, два — французским и английским, один — всеми тремя языками. Сколько людей в группе?

ПОДСКАЗКА. Теорема сложения.

26. На первом этаже семиэтажного дома в лифт вошли три человека. Каждый может выйти на любом из шести этажей (от второго до седьмого). В скольких случаях все трое могут выйти на разных этажах?

27. На первом этаже семиэтажного дома в лифт вошли три человека. Каждый может выйти на любом из шести этажей (от второго до седьмого). В скольких случаях по крайней мере двое могут выйти на одном этаже?

28. На первом этаже семиэтажного дома в лифт вошли три человека. Каждый может выйти на любом из шести этажей (от второго до седьмого). В скольких случаях ровно двое могут выйти на одном этаже?

29. Сколькими способами можно четыре разных предмета разложить в две одинаковые коробки? В каждой от нуля до четырех предметов.
ПОДСКАЗКА. Классификация возможностей.

30. Про 92 коротышек известно, что из них взяли в поход печенье — 47, конфеты — 38, пироги — 42, печенье и конфеты — 28, печенье и пироги — 31, конфеты и пироги — 26, все три вида закуски — 25. А некоторые взяли только воду. Сколько их?
ПОДСКАЗКА. Теорема сложения.

Глава 2. Случайные события и их вероятности

Введение

Теория вероятностей — это наука о математических закономерностях случайных явлений. Речь идет о явлениях, наблюдения над которыми не всегда приводят к одним и тем же результатам, но для которых при массовом повторении наблюдений определенные результаты повторяются примерно в одной и той же доле случаев. Результат наблюдения (эксперимента) называют случайным событием, если он заранее не предопределен. В теории вероятностей случайные события описываются с помощью множеств, и каждому случайному событию приписывается числовая характеристика $\mathsf{P}(A)$, называемая вероятностью этого события и оценивающая «степень случайности». Таким образом осуществляется переход от реальных, «физических», явлений к их абстрактным математическим моделям («вероятностным» моделям). Построение вероятностных моделей и математический анализ закономерностей для $\mathsf{P}(A)$ составляют основное содержание теории вероятностей.

2.1. Математическое описание случайных событий

2.1.1. События и высказывания. Алгебра событий

Допустим, что проводится некоторый эксперимент, и о его предполагаемом результате делается высказывание A. Например, если бросаются одновременно три монеты, то A может иметь вид: «из трех монет хотя бы одна упала гербом вверх». Составление высказываний подобного рода является первым этапом математического описания случайных событий. Причем понятие «событие произошло» отождествляется с истинностью высказывания A. Известно из курса математической логики, что данные высказывания A, B можно соединять «логическими связками»: A и B; A или B; если A, то B. Проделывая такие операции, будем получать словесные описания «сложных» событий на основе описания «простых». В теории вероятностей на этот счет принята следующая терминология. Если A и B — случайные события, то их *суммой* $A + B$, *произведением* AB, *разностью* $A - B$ называются события, состоящие соответственно в том, что «происходит A или B», «A и B», «A и не B». Событие, которое обязательно происходит в рамках изучаемого эксперимента, называется *достоверным* и описывается тождественно истинным высказыванием. Его принято обозначать буквой U. Событие, которое никогда не произойдет в результате эксперимента, называется *невозможным*, описывается тождественно ложным высказыванием и обозначается буквой V. Если $AB = V$, то A и B называются *несовместными.* Если $A = B_1 + B_2 + \cdots + B_k$ и слагаемые попарно не совместны, то говорят, что *A распадается на частные случаи.* Если U распадается на частные случаи, то говорят, что они образуют *полную группу событий.* Если какое-то событие вместе с A образует полную группу, то оно называется *противоположным* A и обозначается $\bar{A}$ ($A + \bar{A} = U$, $A\bar{A} = V$). Запись $A \subset B$ обозначает, что «наступление события A *влечет за собой* B». Очевидно, что равенство $A = B$ равносильно тому, что $A \subset B$ и $B \subset A$. Операции, определенные для конечного числа событий, обладают всеми известными свойствами «алгебры высказываний».

Задача 2.1. *В аудитории случайным образом выбирается один слушатель. Через A обозначим тот факт, что выбрана девушка; B — что выбранный слушатель любит танцевать; C — что слушатель живет в общежитии. Выяснить следующие вопросы:*

1) что означает $A\bar{B}C$;

2) когда $ABC = A$;

3) когда $\bar{C} \subset B$;

4) когда $A = B$; справедливо ли это, если все девушки любят танцевать.

В первом вопросе речь идет о событии, словесное описание которого: «выбрана девушка, любящая танцевать и не живущая в общежитии». Во втором случае включение $ABC \subset A$ очевидно, поэтому равенство равносильно включению $A \subset ABC$, а значит, включению $A \subset BC$. Последнее можно описать высказыванием: «все девушки-слушательницы любят танцевать и живут в общежитии». Соотношение $\bar{C} \subset B$ означает, что все слушатели, не живущие в общежитии, любят танцевать. Равенство событий A и B означает совпадение множества девушек-слушательниц с множеством слушателей, любящих танцевать (юноши танцевать не любят), в то время как утверждение «все девушки любят танцевать» не исключает существования юношей, также любящих танцевать.

Простейшие свойства операций над событиями. (Алгебра событий.)

1) $A \subset A$;
2) $A \subset B$ и $B \subset C$, следовательно, $A \subset C$;
3) $(AB)C = A(BC)$;
4) $AB = BA$;
5) $(A+B)+C = A+(B+C)$;
6) $A+B = B+A$;
7) $A+A = A$;
8) $AA = A$;
9) $A(B+C) = AB+AC$;
10) $A+BC = (A+B)(A+C)$;
11) $A+V = A$;
12) $AU = A$;
13) $A \subset A+B$;
14) $AB \subset A$;
15) $A \subset C$ и $B \subset C$, следовательно, $A+B \subset C$.

Задача 2.2. *Доказать, что* $\overline{\overline{A}} = A$; $\overline{A+B} = \bar{A}\bar{B}$; $\overline{AB} = \bar{A}+\bar{B}$.

Действительно, $U \setminus \bar{A} = A$ и $U \setminus \bar{A} = \overline{\overline{A}}$, так как $\bar{A}+\overline{\overline{A}} = U$ и $\bar{A}\,\overline{\overline{A}} = V$. Поскольку $(A+B)+\bar{A}\bar{B} = U$ и $(A+B)\bar{A}\bar{B} = V$, то $\overline{A+B} = \bar{A}\,\bar{B}$. Наконец, $\bar{A}+\bar{B} = \overline{\overline{\bar{A}+\bar{B}}} = \overline{\overline{\overline{A}}\,\overline{\overline{B}}} = \overline{AB}$.

Для того чтобы вычислять вероятности событий, надо каким-то образом численно сравнивать их. Алгебра логики тут бессильна, поэтому привлекается теория множеств.

2.1.2. События и множества. Пространство элементарных событий

Пусть у нас есть высказывание A о результатах некоторого эксперимента, то есть сформулировано, в чем состоит событие A. Допустим также, что нам удалось так подобрать некоторое множество Ω, что все простейшие логически возможные результаты рассматриваемого эксперимента можно описать в виде элементов Ω, а событие A представить в виде высказывания с переменной, принадлежащей Ω. Тогда в Ω можно выделить множество истинности, соответствующее A. Подразумевается, что простейшие рассматриваемые логически возможные результаты образуют полную группу событий. Назовем такие результаты и соответствующие им элементы Ω *элементарными событиями*, а само множество Ω — *пространством элементарных событий.* На самом деле Ω —

это образ, теоретико-множественная модель пространства элементарных событий, но для краткости слово «модель» пропускается. Множество истинности события A является объединением нескольких элементарных событий, и про каждое из входящих в него элементарных событий говорят, что оно *благоприятствует* событию A (влечет A). При решении задачи на нахождение вероятности события A рассматривают не само высказывание, а его теоретико-множественную модель, то есть подмножество некоторого пространства элементарных событий Ω, поэтому подмножество также называют «событием» в Ω, соответствующим высказыванию A (описывающим A). Подмножество, описывающее A, обычно также обозначают A. Само пространство Ω считается, по определению, событием в Ω, описывающим U. Пустое подмножество Ω, также называется, по определению, событием. Оно описывает V. Далее, говоря об описании событий с помощью множеств, будем употреблять символы Ω и $\varnothing$ вместо U и V.

Задача 2.3. *Студент X проводит эксперимент по сдаче экзамена по теории вероятностей. Логически возможные результаты выражаются в стандартной системе оценок. Запишем их в виде множества:* $\Omega = \{2, 3, 4, 5\}$. *Определите, каким из нижеперечисленных высказываний (утверждений) соответствуют множества истинности (события) в рассматриваемом пространстве элементарных событий Ω:*

а) экзамен не провален;

б) студент X предпочитает сдавать на «отлично»;

в) преподаватель оценил знания студента как не блестящие;

г) подруга студента X может им гордиться;

д) преподаватель любит ставить двойки.

Ответим на поставленные вопросы. В случае (а) студент получил оценку больше двух. Это, конечно, событие! Ему благоприятствуют элементарные исходы: 3, 4, 5. В случае (б) высказывание касается предпочтений студента и не может быть описано подмножеством пространства Ω. В случае (в) оценка преподавателем знаний должна адекватно выражаться поставленным в ведомость баллом, следовательно, высказыванию (в) соответствует событие $A = \{2, 3, 4\}$ (то есть оценка не равна пяти). В случае (г) не известно, однозначно ли определяются отношения между X и его подругой результатом экзамена по теории вероятностей, поэтому высказыванию не соответствует событие в Ω. Чтобы описать множество истинности высказывания (г) подмножеством, нужно ввести другое Ω. Аналогичная ситуация имеется и для высказывания (д).

Задача 2.4. *В качестве эксперимента студент Y сел в трамвай, не имея ни денег, ни проездной карточки. Примем за результат эксперимента время, за которое кондуктор подойдет к студенту, при условии, что продолжительность поездки студента Y равна пяти минутам. Назовем это время «временем ожидания». Необходимо описать пространство элементарных событий в виде множества и выяснить, какие подмножества (события) могут соответствовать высказываниям: а) время ожидания не более пяти минут; б) ждать не пришлось; в) долго ждать не пришлось; г) через некоторое время кондуктор подошел к студенту Y.*

Для этой задачи естественно предположить, что время ожидания — это действительное число, принадлежащее отрезку $[\,0,\,5\,]$, то есть $\Omega = [\,0,\,5\,]$. В этом пространстве несчетное число элементарных событий вида: «время ожидания равно числу t из отрезка $[\,0,\,5\,]$». Высказыванию (а) соответствует достоверное событие. Высказыванию (б) соответствует элементарное событие $\{0\}$. Высказыванию (г) соответствует событие $(0,\,5]$. Для высказывания (в) нельзя подобрать подходящее подмножество Ω, так как понятие «долго» не записано в виде численной оценки. Мы не утверждаем, что приведенное здесь определение Ω является единственно возможным.

2.1.3. Операции над событиями в Ω. Поле событий

Над подмножествами Ω можно проделывать операции объединения, пересечения, дополнения. Можно определить отношения включения и равенства. Это будет алгебра подмножеств в точности дублирующая алгебру высказываний. Так, например, высказыванию типа «A или B» будет соответствовать объединение подмножеств, описывающих A, B. Это объединение также принято обозначать $A + B$. Высказыванию «A и B» будет соответствовать пересечение подмножеств, обозначаемое AB. Высказыванию $\bar{A}$ соответствует дополнение A до Ω. Его принято обозначать $\bar{A}$. Для нахождения вероятности событий типа $A + B$, AB с помощью Ω важно, чтобы A и B оба были описаны подмножествами, то есть были событиями, *в одном и том же* пространстве элементарных событий.

При решении задач на вычисление вероятности, как правило, нет необходимости рассматривать все подмножества Ω. В некоторых случаях они могут соответствовать столь экзотическим событиям, связанным с исходным экспериментом, рассматривать которые не имеет смысла. При решении задач среди всех подмножеств пространства элементарных событий Ω выделяют некую совокупность F событий, обладающую свойствами:

1) все элементарные события принадлежат F;

2) Ω и $\varnothing$ принадлежат F;

3) если события A и B принадлежат F, то AB, $A + B$, $\bar{A}$, $\bar{B}$ также принадлежат F.

Совокупность с такими свойствами называется *полем событий, определенных на* Ω.

Поле событий называется *борелевским*, если в пункте (3) предыдущего определения допускается сложение и умножение не только конечного, но и *счетного* количества событий.

Алгебра подмножеств данного множества, замкнутая относительно операций конечного или счетного объединения, конечного или счетного пересечения, дополнения до исходного множества, а также непременно содержащая пустое множество, называется в теории множеств σ-алгеброй. Поэтому определенное нами борелевское поле событий можно также назвать σ-алгеброй событий.

В рамках поля перечисленные операции над событиями дают в результате события из того же поля. Это важно при определении вероятности и ее свойств. Подробности можно найти в разделе 2.4.

2.2. Классическое определение вероятности (классическая дискретная модель)

Первые попытки определить вероятность случайного события математическими методами были сделаны в переписке двух известных ученых: Б. Паскаля и П. Ферма, начатой в 1654 году. Первая книга по теории вероятностей была издана Х. Гюйгенсом в 1657 году. Приведенное в этом разделе определение вероятности было дано Я. Бернулли в 1713 году.

Пусть некоторый эксперимент имеет конечное, не равное нулю, число попарно несовместных исходов. Назовем их элементарными событиями. Объединение всех элементарных событий назовем пространством элементарных событий Ω. Совокупность всех подмножеств Ω, включающая само Ω и пустое множество, очевидно, является полем событий (см. раздел 2.1.3). Обозначим это поле через F. Допустим далее, что все элементарные события равновозможны в силу каких-то объективных не математических причин. Тогда, по определению, вероятностью события A из F назовем число

$$\mathsf{P}(A) = \frac{M(A)}{M(\Omega)}, \tag{2.1}$$

где $M(A)$ обозначает число элементов в множестве A, $M(\Omega)$ — число элементов в Ω.

Совокупность (Ω, F, P) называется *классической вероятностной моделью.* Общее понятие о вероятностной модели будет дано в разделе 2.4.

Рассмотрим несколько примеров вычисления вероятности с помощью классической модели (как еще говорят, «по классической схеме»).

Задача 2.5. *Монета бросается три раза. Событие A состоит в том, что ровно два раза выпал герб. Чему равна его вероятность?*

Определим пространство элементарных событий так, чтобы A в нем было описано подмножеством. Сохраним за этим подмножеством обозначение A. Итак, пусть

$$\begin{aligned}
\Omega &= \{(\text{ггг}), (\text{ггц}), (\text{гцг}), (\text{цгг}), (\text{ццг}), (\text{цгц}), (\text{гцц}), (\text{ццц})\};\\
A &= \{(\text{ггц}), (\text{цгг}), (\text{гцг})\}.
\end{aligned}$$

Элементами Ω являются, по определению, упорядоченные тройки символов г(герб), ц(цифра). Первый символ показывает, каков исход первого бросания монеты, второй — второго, третий — третьего. Если предположить, что восемь элементарных событий, составляющих Ω, равновозможны, то вероятность A, вычисленная по классической схеме, равна $3/8$.

Даже в таком простом примере не ясно, почему Ω выбрано именно так, а не иначе. Можно предложить другой вариант. Пусть Ω содержит только четыре элементарных события: число гербов из трех равно 0, 1, 2 или 3. Если предположить, что все четыре элементарных события равновозможны, то вероятность A будет равна $1/4$. Таким образом, для разных моделей получаем разные оценки вероятности. Какая модель лучше описывает эксперимент? В общем случае это довольно сложный вопрос, ответить на который помогает статистика (то есть обработка реальных экспериментальных данных). Если модель достоверно отражает закономерности реального эксперимента (адекватна ему), то относительные частоты, с которыми наблюдаются результаты этого эксперимента, при большом числе повторений эксперимента должны быть «близки» к вероятностям, рассчитанным с помощью модели. Под относительной частотой появления события A в статистике понимается число k/n, где n — число проведенных экспериментов, а k — число экспериментов из n, в которых произошло событие A. Понятию «близости» вероятности и относительной частоты можно дать строгое определение [2, 11]. Здесь достаточно интуитивного понимания: числа приближенно равны.

В рассмотренном примере можно предложить логическое «тестирование» двух моделей. В качестве теста предлагается сравнить вероятности событий X и Y:

X — «все три монеты упали одинаковой стороной вверх»;

Y — «не все три монеты упали одинаковой стороной вверх».

С точки зрения «физики эксперимента» интуитивно ясно (и можно подтвердить на опытах), что событие X менее вероятно, чем Y. Первая модель адекватно отражает эту закономерность, тогда как во второй вероятности X и Y равны.

Заключая разбор первого примера, отметим, что в теории вероятностей разработаны приемы построения различных «стандартных» моделей, адекватных определенным классам практических задач. С некоторыми из этих моделей мы познакомимся в следующих разделах.

Задача 2.6. *Ребенок случайным образом ставит в ряд кубики с буквами: к, л, м, о, о, о. С какой вероятностью может получиться слово «молоко»?*

В качестве элементарных событий возьмем все кортежи из шести кубиков. Тогда Ω будет состоять из $6!$ элементов. Вычислим по теореме умножения число кортежей, дающих в результате слово «молоко»: $1 \cdot 3 \cdot 1 \cdot 2 \cdot 1 \cdot 1 = 3!$. Поэтому вероятность рассматриваемого события равна $3!/6! = 1/120$.

Очевидны свойства *вероятностной функции* $\mathsf{P}(A)$, вытекающие из определения (2.1):

1) $\mathsf{P}(A) \geqslant 0$, $\mathsf{P}(A) \leqslant 1$;

2) $\mathsf{P}(\Omega) = 1$, $\mathsf{P}(\varnothing) = 0$;

3) если $AB = \varnothing$, то $\mathsf{P}(A + B) = \mathsf{P}(A) + \mathsf{P}(B)$;

4) в общем случае $\mathsf{P}(A + B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB)$, что следует из теоремы сложения ($M(A \cup B) = M(A) + M(B) - M(A \cap B)$);

5) $\mathsf{P}(\bar{A}) = 1 - \mathsf{P}(A)$, так как $\mathsf{P}(A) + \mathsf{P}(\bar{A}) = \mathsf{P}(\Omega) = 1$;

6) если $A \subset B$, то $\mathsf{P}(A) \leqslant \mathsf{P}(B)$, так как $\mathsf{P}(A) \leqslant \mathsf{P}(A)+\mathsf{P}(B-A) = \mathsf{P}(B)$, где $B - A$ состоит из всех элементов B, не принадлежащих множеству A.

Пользуясь теоремой сложения, можно обобщить свойства (3) и (4) на любое конечное число событий.

В качестве иллюстрации к применению классической вероятностной модели приведем решение *задачи о выборке данного состава.*

Задача 2.7. *Пусть среди n предметов ровно m обладают некоторым свойством. Берутся наугад n_1 предметов из n. Какова вероятность, что среди них ровно m_1 предметов обладают данным свойством?*

Элементарными событиями назовем всевозможные подмножества данного n-элементного множества, состоящие из n_1 элементов (то есть сочетания из n по n_1). Их количество равно $C_n^{n_1}$. Выбрать m_1 предметов из m предметов, обладающих заданным свойством, можно $C_m^{m_1}$ способами. Выбрать оставшиеся $n_1 - m_1$ предметов из $n - m$ предметов, не обладающих этим свойством, можно $C_{n-m}^{n_1-m_1}$ способами. Следовательно, по теореме умножения получаем, что изучаемому событию благоприятствуют $C_m^{m_1}\,C_{n-m}^{n_1-m_1}$ элементарных исходов. Поэтому искомая вероятность равна

$$\mathsf{P}(A) = \frac{C_m^{m_1}\,C_{n-m}^{n_1-m_1}}{C_n^{n_1}}. \tag{2.2}$$

Задача 2.8. *В киоске шесть лотерейных билетов. Из них два выигрышных. Куплено наугад три билета. Какова вероятность, что ровно один из них выигрышный?*

Здесь $n = 6$, $m = 2$, $n_1 = 3$, $m_1 = 1$. Поэтому $\mathsf{P}(A) = C_2^1\,C_4^1/C_6^3 = 0.6$.

Замечание. Вычисление $n!$ при больших n затруднительно. Если воспользоваться *формулой Стирлинга*

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\theta_n}, \quad \text{где} \quad |\theta_n| < \frac{1}{12n},$$

и пренебречь в ней последним множителем, то можно получить приближенное значение

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n. \tag{2.3}$$

Практическая ценность этой формулы очевидна.

2.3. Геометрическое определение вероятности (геометрическая модель)

Классическое определение вероятности неприменимо, если логически возможных исходов эксперимента бесконечно много. В качестве примера рассмотрим следующую геометрическую задачу. Пусть Ω — квадрируемое (то есть имеющее площадь) множество, A — его квадрируемое подмножество. Какова вероятность, что случайно выбранная точка M из Ω

принадлежит также и A (как говорят, «попадет в A»)? Если предположить, что вероятность попадания в произвольную квадрируемую часть Ω зависит только от площади этой части (причем прямо пропорционально) и не зависит от ее расположения в Ω, то естественно за эту вероятность принять, по определению, отношение площадей:

$$\mathsf{P}(A) = \frac{\text{площадь}(A)}{\text{площадь}(\Omega)}. \tag{2.4}$$

Это определение хорошо согласуется с классическим. Действительно, если множество Ω разбито на n частей равной площади, то вероятность попадания случайной точки в каждую такую часть по обоим определениям равна $1/n$.

Хорошо известно, что квадрируемые подмножества Ω образуют σ-алгебру. В силу этого можно рассматривать поле F событий на Ω, состоящее из всех его квадрируемых подмножеств. На этом поле будут иметь смысл сумма, произведение, разность событий и дополнение их до Ω. Вероятностная функция, определенная по формуле (2.4), будет обладать всеми теми же свойствами, перечисленными в разделе 2.2 для классического определения.

Аналогичная модель $(\Omega,\, F,\, \mathsf{P})$ может быть построена при условии, что Ω — спрямляемое множество (имеющее длину) или кубируемое множество (имеющее объем). Все модели такого типа носят общее название *«геометрические вероятностные модели»*.

Задача 2.9. *Отрезок случайным образом делится на три части. Какова вероятность, что эти части могут быть сторонами одного треугольника?*

Пусть длина отрезка равна L, длина одной части равна x, другой — y, третьей — $L-x-y$. В качестве Ω рассмотрим множество всех таких точек $(x,\, y)$ на координатной плоскости, координаты которых удовлетворяют трем неравенствам одновременно:

$$0 < x < L, \quad 0 < y < L, \quad 0 < L - x - y < L.$$

Множество Ω (пространство элементарных событий) представляет собой треугольник OAB (см. рис. 2.1). Известно, что для того, чтобы три числа были длинами сторон одного треугольника, необходимо и достаточно, чтобы каждое из них было меньше суммы двух других. Следовательно, высказыванию «три части могут быть сторонами одного треугольника» соответствует событие A — множество всех таких пар $(x,\, y)$,

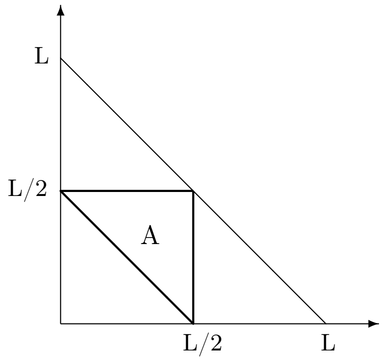


Рис. 2.1

принадлежащих Ω, для которых одновременно истинны неравенства:

$$x + y > L - x - y, \quad x + (L - x - y) > y, \quad y + (L - x - y) > x.$$

Множество A обведено жирными отрезками на рис. 2.1. Его площадь равна $0.5(L/2)^2$, в то время как площадь Ω равна $0.5L^2$. Вероятность события A в рассмотренной геометрической модели равна 0.25.

Задача 2.10. *В квадратном трехчлене $x^2 + px + q$ коэффициенты случайны и по модулю не превосходят девяти. Какие корни трехчлена более вероятны: действительные или мнимые?*

Пару (p, q) рассмотрим как точку на плоскости. Так как $|p|, |q| \leqslant 9$, то все такие точки заполняют квадрат. Корни трехчлена являются мнимыми тогда и только тогда, когда $p^2 - 4q < 0$, следовательно, событию «корни мнимые» соответствует фигура A на рис. 2.2. Площадь фигуры A можно вычислить с помощью определенного интеграла:

$$S(A) = \int\limits_{-6}^{6} (9 - 0.25p^2)\, dp = \left(9p - \frac{p^3}{12}\right)\Bigg|_{-6}^{6} = 72.$$

Следовательно, вероятность мнимых корней равна $72/18^2 = 2/9$. Вероятность действительных корней больше. Она равна $1 - 2/9 = 7/9$.

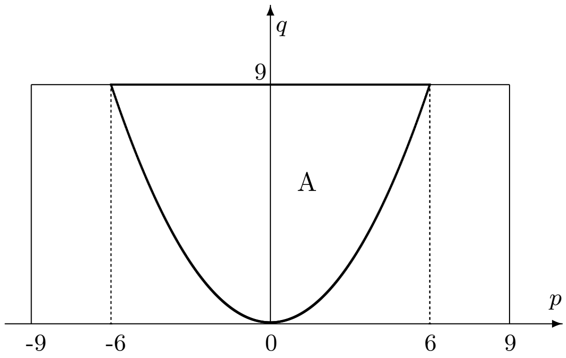


Рис. 2.2

2.4. Аксиоматическое определение вероятности

Современная теория вероятностей основана на аксиомах, предложенных А.Н. Колмогоровым [7]. Аксиоматика Колмогорова, по сути, формулирует общие требования к вероятностным моделям.

Определение. Пусть задано множество Ω (пространство элементарных событий) и борелевская σ-алгебра F его подмножеств (поле событий). *Вероятностной функцией* $\mathsf{P}(A)$, определенной на событиях, принадлежащих F, называется отображение F в множество вещественных чисел, удовлетворяющее условиям:

1) $\mathsf{P}(A) \geqslant 0$;

2) $\mathsf{P}(\Omega) = 1$;

3) если $A = \sum_{i=1}^{\infty} A_i$ (все A_i принадлежат F) и $A_i\, A_j = \varnothing$ при $i \neq j$, то $\mathsf{P}(A) = \sum_{i=1}^{\infty} \mathsf{P}(A_i)$.

Тройка объектов (Ω, F, P) называется *вероятностной моделью* или *вероятностным пространством.*

Замечания. Свойства (1)–(3) являются аксиомами вероятности. Из них легко выводятся другие свойства, перечисленные в разделе 2.2 для классической модели. Ряд, указанный в свойстве (3) вероятностной функции, сходится, так как слагаемые неотрицательны, а частные суммы ограничены единицей. Легко проверить, что вероятностные модели, описанные во втором и третьем разделах, удовлетворяют требованиям аксиоматики Колмогорова и, следовательно, не зря названы вероятностными моделями.

Если пространство элементарных событий не более чем счетно, то, очевидно, поле F должно состоять из всех подмножеств Ω. Докажите это самостоятельно.

По конкретной «физической» задаче можно построить различные вероятностные модели: по-разному вводить Ω, на том же Ω по-разному определять F. Но если эти модели описывают одну и ту же исходную задачу, объективно отражают «физические» закономерности, то одни и те же явления должны иметь одну и ту же вероятность.

2.5. Условная вероятность. Последовательные испытания (эксперименты)

Пусть ставятся два эксперимента. Рассмотрим высказывание A, описывающее некий результат первого эксперимента, и высказывание B, описывающее некий результат второго. Пусть до события B произошло событие A. Изменит ли оценку шанса появления B информация о том, что A уже произошло? Рассмотрим пример такой ситуации.

В ящике два шарика: черный и белый. Наугад вытаскивается один шар. Событие A: «первый шар — белый». Событие B: «второй шар — черный». Если мы не знаем, какого цвета первый шар (например, шар вынут в абсолютной темноте), то естественно считать, что вероятность B равна 0.5. Если же известно, что A произошло, то вероятность B равна единице.

Для математического описания зависимостей между результатами последовательно проводимых экспериментов используется понятие *условная вероятность.*

Определение условной вероятности. Пусть дана вероятностная модель $(\Omega,\ F,\ \mathsf{P})$. Выделим из F событие A, имеющее ненулевую вероятность. Для произвольного события B из F *вероятностью B при условии A* называется отношение

$$\mathsf{P}(B/A) = \frac{\mathsf{P}(AB)}{\mathsf{P}(A)}. \tag{2.5}$$

С помощью формулы (2.5) определяется вероятность события B в новом пространстве A. Действительно, примем множество A за новое пространство элементарных событий Ω_1. Пересечем все события поля F с множеством A. Получится новое борелевское поле F_1 в пространстве Ω_1 (докажите это). Вероятностную функцию P_1 на F_1 определим по формуле (2.5), полагая для любого события B из F_1, что $\mathsf{P}_1(B) = \mathsf{P}(AB)/\mathsf{P}(A)$.

Легко доказать, что вероятностная функция P_1 на F_1 имеет те же свойства, что и P на F. Будем в дальнейшем называть A *подпространством* Ω.

Формула (2.5) является теоретическим обобщением свойства относительных частот: если из N экспериментов событие A наблюдалось N_A раз, а событие AB — N_{AB} раз, то событие B наблюдалось в той части случаев появления A, которая равна N_{AB}/N_A (относительная частота B при условии A). С другой стороны, то же число равно $(N_{AB}/N)/(N_A/N)$.

Задача 2.11. *В группе 30 студентов, из них 10 — юноши, 20 — девушки. Отличников в группе — восемь, из них трое — юноши. Какова вероятность, что случайно выбранный из списка этой группы человек — отличник (событие B)? Какова вероятность, что выбран отличник, если известно, что выбран юноша (событие B при условии A)?*

По классической модели вероятность B равна $8/30 = 4/15$. Вероятность, что выбран юноша-отличник, равна $3/30$, это $\mathsf{P}(A \text{ и } B)$. Вероятность, что выбран юноша, равна $\mathsf{P}(A) = 10/30 = 1/3$. По определению вероятности B при условии A имеем $\mathsf{P}(B/A) = (3/30) : (1/3) = 3/10$. Тот же результат мы бы получили, введя классическую модель на A.

Задача 2.12. *Вероятность, что любит и не поцелует, равна* 0.1. *Вероятность, что не любит и поцелует, равна* 0.05. *Вероятность, что поцелует при условии, что любит, равна* 0.8. *Какова вероятность, что любит? Какова вероятность, что поцелует?*

В качестве элементарных событий рассмотрим четыре события: A (любит и поцелует), B (любит и не поцелует), C (не любит и поцелует), D (не любит и не поцелует). Они образуют полную группу. Даны вероятности: $\mathsf{P}(B) = 0.1$, $\mathsf{P}(C) = 0.05$, $\mathsf{P}(A+C$ при условии $A+B) = 0.8$. Надо найти вероятность $A+B$. Обозначим ее через x. Тогда по определению условной вероятности $(x-0.1)/x = 0.8$, откуда получаем $x = 0.5$. Так как $\mathsf{P}(A) = x - 0.1 = 0.4$, то $\mathsf{P}(A+C) = 0.45$.

Из определения условной вероятности непосредственно вытекает

Теорема 2.1 (теорема умножения вероятностей). *Если события A и B из F имеют положительные вероятности, то*

$$\mathsf{P}(AB) = \mathsf{P}(A)\,\mathsf{P}(B/A) = \mathsf{P}(B)\,\mathsf{P}(A/B). \tag{2.6}$$

Добавляя по одному событию, можно получить обобщение

$$\begin{aligned}\mathsf{P}(A_1A_2\ldots A_k) = \mathsf{P}(A_1)\,\mathsf{P}(A_2/A_1)\,\mathsf{P}(A_3/A_1A_2)\ldots \\ \mathsf{P}(A_k/A_1A_2\ldots A_{k-1}).\end{aligned} \tag{2.7}$$

Заметим, что в формулах (2.6), (2.7) все вероятности рассчитываются в рамках одной и той же модели (Ω, F, P). На практике при решении вероятностных задач часто возникает ситуация, когда вероятность A просчитана в одном пространстве, а условные вероятности — в другом. Использование формул (2.6), (2.7) в таких случаях приобретает несколько иной смысл.

Задача 2.13. *Студент на экзамене знает 20 вопросов из 25. Преподаватель задает два вопроса. Какова вероятность, что студент знает оба эти вопроса?*

Рассмотрим последовательность из двух испытаний: 1) входит ли первый вопрос в 20 известных; 2) входит ли второй в 20 известных. Обозначим через A высказывание: «первый вопрос входит в 20 известных», а через B — высказывание: «второй вопрос входит в 20 известных». Из постановки задачи следует, что надо найти вероятность события, описываемого высказыванием «A и B», то есть AB. Если в качестве Ω взять множество из 25 вопросов, то вероятность события A: «первый вопрос входит в 20 известных», вычисленная по классической схеме, равна $20/25 = 0.8$. Если событие A произошло, то остается 19 известных вопросов из 24, иначе — 20. Если в качестве Ω взять подмножество, соответствующее A, то можно на нем по классической схеме рассчитать вероятность события X, что и второй вопрос известен студенту. Эта вероятность будет равна $19/24$. Некорректно сразу утверждать, что полученная вероятность и есть вероятность B при условии A, поскольку в определении условной вероятности предполагается, что значения вероятностей A и AB получены в рамках одной модели. Для решения задачи построим модель, описывающую *совокупный результат двух последовательных экспериментов*, в которой определены события A, B, AB, а X является событием в подпространстве A и имеет в нем вероятность $19/24$.

В качестве Ω возьмем множество из четырех элементов, соответствующих высказываниям: AB, $\bar{A}B$, $A\bar{B}$, $\bar{A}\bar{B}$. События в Ω — это по сути дела *упорядоченные пары множеств*, которые мы приняли за элементы. Сумме событий $AB + A\bar{B}$ припишем вероятность 0.8, так как такое событие происходит тогда и только тогда, когда A. Событие B отождествим с $\bar{A}B + AB$, поскольку оно происходит тогда и только тогда, когда B. Сумму $AB + A\bar{B}$ можно рассматривать как подпространство A в Ω, определив в нем вероятностную функцию по формуле $\mathsf{P}(Y/A) = \mathsf{P}(AY)/\mathsf{P}(A)$. С другой стороны, известно, что A может быть само рассмотрено как пространство из двух элементарных событий: X и $\bar{X}$, на котором введена вероятностная функция P_A. Событие X в A получается пересечением

A с B. Его вероятность в A как в подпространстве Ω должна равняться $\mathsf{P}(B/A)$.

Следовательно, *если конкретному явлению приписывается единственное значение вероятности*, то должны выполняться соотношения:

$$\mathsf{P}_A(X) = \mathsf{P}(B/A) = 19/24, \quad \mathsf{P}(AB) = \mathsf{P}(A) \cdot \mathsf{P}(B/A) = (20/25) \cdot (19/24).$$

Итак, при решении задачи мы используем формулу (2.6) по следующей схеме:

1. Считаем, что некоторое событие A имеет ненулевую вероятность, рассчитанную каким-то образом в Ω_1. Устройство Ω_1 перестает нас интересовать, как только получена вероятность A. Обозначим ее $\mathsf{P}(A/\Omega_1)$.

2. Рассматриваем A как новое пространство. В нем каким-то способом для некоторого события X вычисляем вероятность $\mathsf{P}_A(X)$.

3. Вводим на A новое F_A, состоящее лишь из X, $\bar{X}$, A и пустого множества. Достраиваем A до Ω так, чтобы A было подпространством: достаточно добавить $\bar{A}$ («не A»). Элементарными событиями в Ω будут X, $\bar{X}$ и $\bar{A}$. F определяем как всевозможные подмножества Ω.

4. Полагаем $\mathsf{P}(A) = \mathsf{P}(A/\Omega_1)$. Рассматривая (A, F_A, P_A) как подпространство (Ω, F, P), получаем, что $\mathsf{P}_A(X) = \mathsf{P}(X/A)$. Следовательно, по формуле (2.6) $\mathsf{P}(X) = \mathsf{P}(A)\mathsf{P}_A(X)$.

Вывод: формула (2.6), естественным образом обобщая свойство относительных частот, позволяет от вероятности события X в подпространстве A перейти к вероятности X в пространстве Ω, куда «вложено» A. В дальнейшем мы не будем без особой надобности подробно описывать «вложение» A в Ω.

Задача 2.14. *В коробке десять красных, пять синих и два желтых шара. Один за другим (без возвращения) вынимаются три шара. Какова вероятность, что последовательно вынуты шары: желтый, синий, красный?*

По формуле (2.7) получаем: P(ж-с-к)$= (2/17)(5/16)(10/15)$. Точно такой же ответ дает классическая модель на пространстве всех упорядоченных троек из 17 шаров.

Задача 2.15. *В условиях предыдущей задачи найти вероятность, что вынуты три разноцветных шара.*

Представим изучаемое событие A как произведение исходов трех экспериментов: вытаскивание первого, второго, третьего шаров. Найдем сумму вероятностей всех произведений, соответствующих последовательностям шаров: ж-с-к, ж-к-с, к-ж-с, к-с-ж, с-ж-к, с-к-ж.

$$\begin{aligned}\mathsf{P}(A) &= \frac{2}{17}\frac{5}{16}\frac{10}{15}+\frac{2}{17}\frac{10}{16}\frac{5}{15}+\frac{10}{17}\frac{2}{16}\frac{5}{15}+\frac{10}{17}\frac{5}{16}\frac{2}{15}+\\ &+\frac{5}{17}\frac{2}{16}\frac{10}{15}+\frac{5}{17}\frac{10}{16}\frac{2}{15}=6\cdot\frac{2\cdot 5\cdot 10}{17\cdot 16\cdot 15}.\end{aligned}$$

Такой же ответ получаем по классической схеме двумя способами: считая выборку трех из 17 шаров упорядоченной или не упорядоченной.

2.6. Независимость событий

В этом разделе мы рассмотрим, как рассчитать вероятность совокупного результата двух или более экспериментов, которые проводятся в независимых друг от друга условиях с точки зрения «физики». Начнем опять со строгого определения, а потом применим его следствия к построению модели, описывающей совокупный результат опытов.

Определение независимости двух событий в рамках вероятностной модели. Пусть задана вероятностная модель (Ω, F, P). Два события A и B из F будем называть независимыми, если

$$\mathsf{P}(AB)=\mathsf{P}(A)\,\mathsf{P}(B). \tag{2.8}$$

Смысл такого определения хорошо раскрывается с помощью следующей теоремы, связывающей понятия условной вероятности и независимости.

Теорема 2.2 (связь условной вероятности и независимости). *Два события A и B из F, имеющие не нулевую вероятность, независимы тогда и только тогда, когда $\mathsf{P}(A/B)=\mathsf{P}(A)$. Последнее равенство равносильно каждому из следующих равенств: $\mathsf{P}(B/A)=\mathsf{P}(B)$, $\mathsf{P}(A/\bar{B})=\mathsf{P}(A)$, $\mathsf{P}(B/\bar{A})=\mathsf{P}(B)$.*

Доказательство.

1) Пусть A и B независимы, то есть $\mathsf{P}(AB)=\mathsf{P}(A)\,\mathsf{P}(B)$, причем $\mathsf{P}(A)$ и $\mathsf{P}(B)$ положительны. Тогда по определению условной вероятности получаем, что $\mathsf{P}(A/B)=\mathsf{P}(AB)/\mathsf{P}(B)=\mathsf{P}(A)$.

2) Пусть $\mathsf{P}(A/B)=\mathsf{P}(A)$, тогда $\mathsf{P}(AB)/\mathsf{P}(B)=\mathsf{P}(A)$, следовательно, $\mathsf{P}(AB)=\mathsf{P}(A)\,\mathsf{P}(B)$.

3) Покажем, что равенство $\mathsf{P}(A/B) = \mathsf{P}(A)$ равносильно равенству $\mathsf{P}(A/\bar{B}) = \mathsf{P}(A)$ (остальное аналогично). Это станет очевидно, если принять во внимание, что оно равносильно равенству $\mathsf{P}(AB) = \mathsf{P}(A)\,\mathsf{P}(B)$, а $AB + A\bar{B} = A$, и $\mathsf{P}(AB) + \mathsf{P}(A\bar{B}) = \mathsf{P}(A)$, следовательно, $\mathsf{P}(A/\bar{B}) = \mathsf{P}(A\bar{B})/\mathsf{P}(\bar{B}) = (\mathsf{P}(A) - \mathsf{P}(AB))/(1 - \mathsf{P}(B))$.

Доказанная теорема придает абстрактному понятию независимости двух событий «физический» смысл: вероятность A остается постоянной как при условии, что B произошло, так и при условии, что B не произошло. Именно это мы принимаем за независимость в обыденной жизни.

Задача 2.16. *Пусть выполнены условия задачи 2.12 предыдущего раздела. Зависимы ли события: «любит» и «поцелует»?*

Для ответа на этот вопрос сравним вероятность события «любит и поцелует» с произведением вероятностей событий «любит», «поцелует». Первая вероятность равна $\mathsf{P}(A) = 0.4$, а две другие — 0.5 и 0.45. Легко видеть, что условие (2.8) не выполнено и события зависимы.

Определение совокупной независимости нескольких событий. События $A_1, A_2, \ldots, A_k$ из F будем называть независимыми в совокупности, если вероятность совместного наступления любых r ($r \geqslant 2$) из этих событий равна произведению их вероятностей.

В частности, вероятность совместного наступления всех данных событий равна

$$\mathsf{P}(A_1 A_2 \ldots A_k) = \mathsf{P}(A_1)\,\mathsf{P}(A_2)\ldots\mathsf{P}(A_k). \tag{2.9}$$

Заметим, что *из попарной независимости не следует независимость в совокупности.*

Задача 2.17. *Пусть события A, B, C состоят в попадании точки в фигуры, указанные на рис. 2.3, расположенные внутри общего правильного треугольника (проведены средние линии). Независимы ли события A, B, C в совокупности?*

Здесь $\mathsf{P}(A) = \mathsf{P}(B) = \mathsf{P}(C) = 0.5$; $\mathsf{P}(AB) = \mathsf{P}(BC) = \mathsf{P}(AC) = 0.25$, следовательно, есть попарная независимость. Однако $\mathsf{P}(ABC) = 0.25$, в то время как $\mathsf{P}(A)\,\mathsf{P}(B)\,\mathsf{P}(C) = 1/8$.

Перейдем теперь к расчету вероятностей совокупных результатов последовательных или совместных испытаний, независимых с точки зрения «физики».

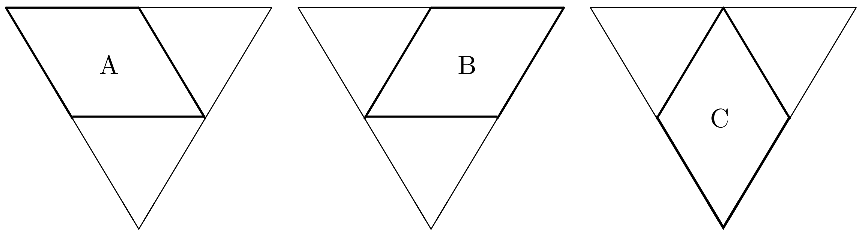


Рис. 2.3

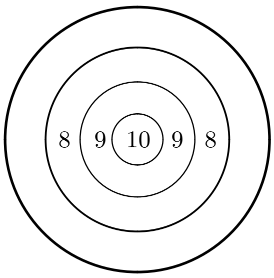


Рис. 2.4

Задача 2.18. *Мишень состоит из концентрических кругов. За попадание в определенную часть мишени начисляются очки, схема изображена на рис. 2.4. Первый стрелок попадает в «десятку» мишени с вероятностью* 0.2*, в «девятку» — с вероятностью* 0.3*, в «восьмерку» — с вероятностью* 0.4*, а в остальную часть мишени или мимо — с вероятностью* 0.1*. (Подумайте, почему нет смысла вычислять вероятность с помощью отношения площадей.)*

Вероятности поражения указанных в примере частей мишени вторым стрелком таковы: «в десятку» — 0.3, «в девятку» — 0.6, «в восьмерку» — 0.1. Эксперимент заключается в том, что два стрелка дают залп по мишени (делают по одному выстрелу одновременно). Будем считать, что результат выстрела второго стрелка не зависит от результата первого с точки зрения «физики» эксперимента, то есть мы имеем дело с *совместными независимыми испытаниями.* Каковы вероятности событий: «у первого стрелка — 10 очков и у второго — 9», «сумма очков за выстрел у команды из двух стрелков не менее 18»?

«Физическая» независимость явлений A и B, как уже обсуждалось, основана на том, что шанс появления A при условии B, что B произошло, равен шансу появления A при условии, что B не произошло. Поэтому во всякой модели для данной задачи, где определены события A и B, вероятность A при условии B должна равняться вероятности A при условии $\bar{B}$ (см. теорему 2.2).

Пусть A — «у первого стрелка — 10 очков»; B — «у второго стрелка — 9 очков». По формуле (2.6) получаем, что

$$\mathsf{P}(AB) = \mathsf{P}(A)\,\mathsf{P}(B/A) = \mathsf{P}(A)\,\mathsf{P}(B) = 0.2 \cdot 0.3 = 0.06.$$

Рассуждая аналогично, получаем вероятность, что команда набрала не менее 18 очков, равную сумме вероятностей событий, описываемых ниже.

Событие	Число очков у первого	Число очков у второго
A_1	10	10
A_2	10	9
A_3	10	8
A_4	9	10
A_5	9	9
A_6	8	10

$$\begin{aligned}\mathsf{P} = \mathsf{P}(A_1 + A_2 + A_3 + A_4 + A_5 + A_6) &= 0.2 \cdot 0.3 + 0.2 \cdot 0.6 + \\ &+ 0.2 \cdot 0.1 + 0.3 \cdot 0.3 + 0.3 \cdot 0.6 + 0.4 \cdot 0.3 = \\ &= 0.2 + 0.3(0.3 + 0.6) + 0.12 = 0.2 + 0.27 + 0.12 = 0.59.\end{aligned}$$

Более подробно похожие примеры рассматриваются в разделе 2.8.

Выделим два очевидных следствия из теоремы 2.2, которые относятся к вычислению вероятности суммы независимых событий.

Теорема 2.3. *Пусть события A и B независимы. Тогда*

$$\mathsf{P}(A+B) = 1 - \mathsf{P}(\bar{A}) \cdot \mathsf{P}(\bar{B}), \tag{2.10}$$

$$\mathsf{P}(A+B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A) \cdot \mathsf{P}(B). \tag{2.11}$$

Задача 2.19. *Два охотника одновременно выстрелили в зайца. Первый попадает в 70% случаев, а второй — в 20%. Какова вероятность, что заяц будет подстрелен?*

Если A, B — события, состоящие в попадании соответственно, первого, второго охотников, то $A + B$ — попадание хотя бы одного из них. Поэтому $\mathsf{P}(A+B) = 0.7 + 0.2 - 0.7 \cdot 0.2 = 1 - 0.3 \cdot 0.8 = 0.76$.

Здесь считалось, что A и B — независимые события. Заметим, что одновременный залп повышает вероятность попадания и с точки зрения теории.

2.7. Полная вероятность. Формула Байеса

Теорема умножения вероятностей (формула (2.6) раздела 2.6), как уже говорилось, позволяет определить вероятность события X в пространстве Ω, если известна вероятность X в подпространстве A и вероятность A в Ω. Но на этом пути нельзя определить вероятность любого события B в Ω. Чтобы пояснить это, рассмотрим следующую задачу.

Задача 2.20. *В появлении двух студентов A и B на лекциях по теории вероятностей обнаружилась следующая закономерность: A посещает 60% лекций, если A пришел на лекцию, то в 99% случаев придет и B. Какой процент лекций посещает B?*

Рассмотрим пространство Ω из четырех элементарных событий: AB, $A\bar{B}$, $\bar{A}B$, $\bar{A}\bar{B}$. Событию A соответствует высказывание: «студент A пришел на лекцию», событию B — «студент B пришел на лекцию». Как описывалось в предыдущем разделе, можно определить вероятность AB по теореме умножения: $\mathsf{P}(AB) = \mathsf{P}(B/A)\,\mathsf{P}(A) = 0.6 \cdot 0.99 = 0.594$. Но событие B в этой модели является суммой событий AB и $\bar{A}B$. Вероятность же $\bar{A}B$ не определить из условия: не известно, как часто ходит студент B на лекции в отсутствие студента A.

Если в отсутствие A студент B приходит также в 99% случаев, то посещаемость B не зависит от A, и $\mathsf{P}(B) = 0.99$. Если же, например, $\mathsf{P}(B/\bar{A}) = 0.01$, то $\mathsf{P}(B\bar{A}) = \mathsf{P}(\bar{A})\,\mathsf{P}(B/\bar{A}) = (1-0.6)0.01 = 0.004$, следовательно, $\mathsf{P}(B) = 0.594 + 0.004 = 0.598$. То есть B посещает 59.8% лекций.

Обобщая решение задачи, делаем следующие выводы: вероятность любого события B в пространстве Ω может быть определена через известные вероятности событий A, $\bar{A}$, B/A и $B/\bar{A}$, вычисленные до того, как было введено Ω (см. схему рассуждений при решении задачи 2.13 в разделе 2.6). Аналогичный результат получается, когда в Ω рассматриваются не два подпространства A и $\bar{A}$, а любое, не более чем счетное, количество подпространств $H_1, H_2, \ldots, H_k, \ldots$, образующих в Ω полную группу событий:

$$\mathsf{P}(B) = \mathsf{P}(B/H_1)\mathsf{P}(H_1) + \mathsf{P}(B/H_2)\mathsf{P}(H_2) + \cdots + \mathsf{P}(B/H_k)\mathsf{P}(H_k) + \ldots \quad (2.12)$$

Ряд в формуле (2.12) сходится, так как все частные суммы ограничены сверху числом 1, и последовательность частных сумм не убывает.

$H_1, H_2, \ldots, H_k, \ldots$ называют *гипотезами*, а их вероятности называют *априорными* (доопытными), поскольку вероятность B вычисляется «по частям» в предположении, что реализовалась либо гипотеза H_1, либо H_2, и так далее.

Формула (2.12) называется «формулой полной вероятности события B».

Замечание. Так же как и теорема умножения вероятностей, формула полной вероятности становится тривиальной, если предположить, что все входящие в нее события и их вероятности заведомо определены одной вероятностной моделью (Ω, F, P). Основная же ценность формулы заключается в том, что она позволяет использовать информацию о подпространствах для получения «исчерпывающей» информации о пространстве Ω.

Задача 2.21. *Город получает тетради от трех фабрик. Первая поставляет 30% общего числа тетрадей, вторая — 50%, третья — 20%. Среди тетрадей, сделанных на первой фабрике, — 60% имеют розовую обложку, на второй — 20%, на третьей — 80%. Какова вероятность, что купленная в этом городе наугад тетрадь будет в розовой обложке?*

Здесь H_i ($i = 1, 2, 3$) — событие, соответствующее высказыванию «купить тетрадь i-й фабрики». B — «купить тетрадь в розовой обложке». По условию задачи: $\mathsf{P}(H_1) = 0.3$, $\mathsf{P}(H_2) = 0.5$, $\mathsf{P}(H_3) = 0.2$, $\mathsf{P}(B/H_1) = 0.6$, $\mathsf{P}(B/H_2) = 0.2$, $\mathsf{P}(B/H_3) = 0.8$. По формуле полной вероятности получаем, что $\mathsf{P}(B) = 0.3 \cdot 0.6 + 0.5 \cdot 0.2 + 0.2 \cdot 0.8 = 0.44$.

Задача 2.22. *Студент из n билетов знает ответы лишь на m билетов. Что вероятнее: сдать экзамен, взяв билет первым или взяв билет вторым?*

Если студент берет билет первым, то вероятность равна m/n (классическая схема). Пусть он берет билет вторым, тогда введем две гипотезы: H_1 — первый отвечающий забрал хороший билет, H_2 — первый отвечающий забрал плохой билет. Очевидно, что $\mathsf{P}(H_1) = m/n$, $\mathsf{P}(H_2) = (n-m)/n$. По формуле (2.12) получаем:

$$\mathsf{P}(B) = \frac{m}{n}\,\frac{m-1}{n-1} + \frac{n-m}{n}\,\frac{m}{n-1} = \frac{m}{n},$$

где $\mathsf{P}(B)$ — вероятность сдать экзамен, идя вторым. Она такая же, как в первом случае! Докажите, что она не зависит от того, каким по счету идти на экзамен.

Следствием из формулы (2.12) является «формула Байеса»:

$$\mathsf{P}(H_i/B) = \mathsf{P}(BH_i)/\mathsf{P}(B) = \\ = \frac{\mathsf{P}(H_i)\,\mathsf{P}(B/H_i)}{\mathsf{P}(B/H_1)\,\mathsf{P}(H_1) + \dots + \mathsf{P}(B/H_k)\,\mathsf{P}(H_k) + \dots}. \tag{2.13}$$

Смысл применения этой формулы заключается в следующем. Некоторые предварительные соображения приводят к вычислению вероятностей гипотез H_1, H_2, ..., H_k, ... Проведен опыт, в результате которого произошло событие B. Какова вероятность, что реализовалась гипотеза H_i?

Вероятности $\mathsf{P}(H_i/B)$ называют *апостериорными* (послеопытными). Они позволяют оценить неизвестный результат первого опыта после проведения второго, если известны *априорные* вероятностные связи.

Задача 2.23. *У больного возможна ровно одна из трех болезней. Априорные вероятности заболевания этими болезнями равны* 0.3, 0.2, 0.5. *Положительная реакция при анализе крови для этих болезней имеет вероятность, соответственно,* 0.1, 0.2, 0.8. *Сделаны три последовательных анализа крови. В первых двух — результаты положительные, в третьем — отрицательный. Каковы апостериорные вероятности болезней?*

Здесь $\mathsf{P}(H_1) = 0.3$, $\mathsf{P}(H_2) = 0.2$, $\mathsf{P}(H_3) = 0.5$. Событие A — три последовательных результата анализов крови.

По условию имеем:

$$\begin{aligned} \mathsf{P}(A/H_1) &= 0.1 \cdot 0.1 \cdot 00.9 = 0.009, \\ \mathsf{P}(A/H_2) &= 0.2 \cdot 0.2 \cdot 0.8 = 0.032, \\ \mathsf{P}(A/H_3) &= 0.8 \cdot 0.8 \cdot 0.2 = 0.128. \end{aligned}$$

По формуле полной вероятности

$$\mathsf{P}(A) = 0.3 \cdot 0.009 + 0.2 \cdot 0.032 + 0.5 \cdot 0.128 = 0.0731.$$

По формуле Байеса

$$\begin{aligned} \mathsf{P}(H/A_1) &= 0.3 \cdot 0.009/0.0731 = 0.04, \\ \mathsf{P}(H/A_2) &= 0.2 \cdot 0.032/0.0731 = 0.09, \\ \mathsf{P}(H/A_3) &= 0.5 \cdot 0.128/0.0731 = 0.87. \end{aligned}$$

2.8. Независимые испытания. Схема Бернулли

В разделе 2.6 рассматривались примеры последовательных испытаний (опытов) и подробно обсуждалось, как, пользуясь теоремой умножения, можно вычислять вероятности событий типа: «в первом опыте — результат A_1, во втором — A_2, ...».

В этом разделе мы более подробно рассмотрим один важный частный случай, когда проводятся ровно n одинаковых последовательных или совместных независимых испытаний, в каждом из которых проверяется, наступит ли некое событие A, вероятность которого постоянна. Покажем, как должна быть построена вероятностная модель, описывающая *все* совокупные результаты этих испытаний. Но для начала нам потребуется поточнее определить, что такое «независимые испытания».

Определение независимости испытаний. Пусть имеется модель, в которой событие A описывает некоторый результат первого испытания (эксперимента), а событие B описывает некоторый результат второго испытания. Допустима совместность указанных событий A и B. Испытания будем называть *независимыми*, если все пары вида A, B независимы.

Аналогично определяются N независимых испытаний через совокупную независимость всевозможных наборов по N событий (первое — результат первого испытания, второе — второго, и так далее).

Определение схемы Бернулли. Будем говорить, что опыты (испытания) проводятся по схеме Бернулли, если выполнены следующие условия:

1. Число опытов известно. Оно равно натуральному числу n.
2. Опыты независимы.
3. В каждом опыте одно и то же (с точки зрения «физики») событие A может произойти с постоянной вероятностью $p = \mathsf{P}(A)$.

Наша цель — построить вероятностную модель, описывающую совокупный результат n опытов в схеме Бернулли. В качестве элементарного события в этой модели будем рассматривать кортеж из n компонент, каждая из которых является одним из двух исходов: A или $\bar{A}$. Всего таких кортежей будет, очевидно, 2^n, но они не равновозможны, если A и $\bar{A}$ не являются равновозможными. По формуле (2.9) мы должны перемножить вероятности исходов, образующих кортеж, если хотим, чтобы в новой модели результаты n опытов были независимы в совокупности, а следовательно, сами опыты были независимы с точки зрения математики (то есть по определению).

Рассмотрим событие, соответствующее высказыванию: «из n испытаний ровно в m случаях произошло событие A». Обозначим это событие $E_{n,m}$. Очевидно, что оно распадается на частные случаи, имеющие вид произведений m множителей A и $n-m$ множителей $\bar{A}$. Например, при $n=3$, $m=2$

$$E_{3,2} = AA\bar{A} + A\bar{A}A + \bar{A}AA.$$

Обозначим вероятность $\bar{A}$ через q. Очевидно, что $q = 1-p$. Вычислим с помощью умножения вероятности всех элементарных исходов, благоприятствующих $E_{n,m}$. Они все равны $\mathsf{P}(A_1 A_2 \dots A_n) = p^m\, q^{n-m}$, где m — число A_i, равных A.

Теперь вероятность события $E_{n,m}$ можно найти по формуле

$$\mathsf{P}(E_{n,m}) = C_n^m\, p^m\, q^{n-m}. \tag{2.14}$$

Формула (2.14) называется «биномиальным законом расчета вероятности», поскольку описывает слагаемые бинома Ньютона:

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

Эта же формула в литературе называется формулой Бернулли по имени создателя.

Задача 2.24. *В контрольной работе пять задач. Вероятность не решить любую из них равна* 0.2. *Какова вероятность, что будет решено не менее трех задач?*

Здесь $q = 0.2$, поэтому $p = 0.8$. Нужно найти вероятность суммы $E_{5,3} + E_{5,4} + E_{5,5}$. По теореме сложения и формуле (2.14) получаем:

$$\begin{aligned}\mathsf{P}(E_{5,3} + E_{5,4} + E_{5,5}) &= \mathsf{P}(E_{5,3}) + \mathsf{P}(E_{5,4}) + \mathsf{P}(E_{5,5}) = \\ &= C_5^3 p^3 q^2 + C_5^4 p^4 q + C_5^5 p^5 = \\ &= 10 \cdot (0.8)^3 \cdot (0.2)^2 + 5 \cdot (0.8)^4 \cdot 0.2 + (0.8)^5 \approx 0.942.\end{aligned}$$

Выясним, как меняется $\mathsf{P}(E_{n,m})$ при изменении m от нуля до n. Для этого рассмотрим отношение

$$\frac{\mathsf{P}(E_{n,m+1})}{\mathsf{P}(E_{n,m})} = \frac{C_n^{m+1} p^{m+1} q^{n-m+1}}{C_n^m p^m q^{n-m}} = \frac{(n-m)p}{(m+1)q} = 1 + \frac{np-q-m}{mq+q}.$$

Если $m \leqslant np - q$, то отношение больше единицы, и при таких m вероятность $\mathsf{P}(E_{n,m})$ растет. Если же $m \geqslant np - q + 1 = np + p$, то $\mathsf{P}(E_{n,m})$ с

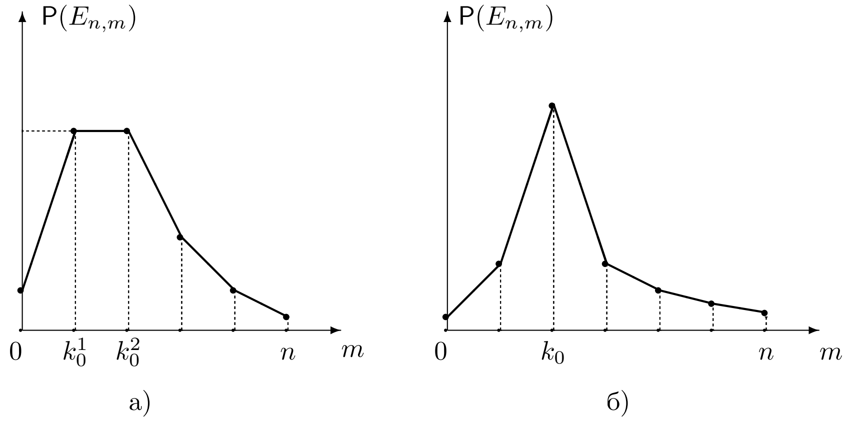


Рис. 2.5

ростом m убывает. Поэтому целые числа k_0, удовлетворяющие неравенству

$$np - q \leqslant k_0 \leqslant np + p, \tag{2.15}$$

являются *наивероятнейшими числами* появления A при n испытаниях (рис. 2.5). На рис. 2.5,а число $np - q$ — целое. На рис. 2.5,б число $np - q$ — дробное.

Задача 2.25. *Вероятность выиграть по лотерейному билету равна* 0.08*. Каково наиболее вероятное число выигрышных билетов из 90? А из 99? Сколько билетов надо купить, чтобы наиболее вероятное число выигрышных было равно единице?*

При $n = 90$, $p = 0.08$, $q = 0.92$ имеем $np - q = 6.28$, $np + p = 7.28$, поэтому $k_0 = 7$.

При $n = 99$, $p = 0.08$, $q = 0.92$ имеем $np - q = 7$, $np + p = 8$, следовательно, k_0 может принимать любое из двух значений: 7 или 8.

При $k_0 = 1$ имеем $0.08\,n - 0.92 \leqslant 1$, то есть $n \leqslant 24$, и $0.08\,n + 0.08 \geqslant 1$, то есть $n \geqslant 12$. Следовательно, купить нужно не менее 12 и не более 24 билетов.

Заметим, что возможны значения $k_0 = 0$ или $k_0 = n$. Подтвердите это примерами самостоятельно.

2.9. Задачи для самостоятельного решения

2.9.1. Первый уровень сложности

1. Из цифр 0, 1, 8 составлено четырехзначное число A. Какова вероятность, что $A = 8888$? Какова вероятность, что A содержит хоть одну восьмерку?

ПОДСКАЗКА. Числа не начинаются на нуль. Классическая схема.

ОТВЕТ: $1/54$, $1 - 8/54$.

2. Монета брошена четыре раза. Какова вероятность, что выпало не менее двух гербов?

ПОДСКАЗКА. Классическая схема.

ОТВЕТ: $1 - 1/16 - 4/16$.

3. В колоде — 36 карт (от шестерки до туза). Выбираются три. Какова вероятность, что среди них — ровно два туза?

ПОДСКАЗКА. Классическая схема.

ОТВЕТ: $C_4^2 C_{32}^1 / C_{36}^3$.

4. В круге единичного радиуса отметили случайно точку. Какова вероятность, что она попала в заранее вписанный правильный треугольник?

ПОДСКАЗКА. Геометрическая модель. Площади.

ОТВЕТ: $3\sqrt{3}/(4\pi)$.

5. На отрезке $[5,\ 18]$ случайно выбрано число A. Какова вероятность, что A принадлежит также отрезку $[8,\ 10]$?

ПОДСКАЗКА. Геометрическая модель. Длины.

ОТВЕТ: $2/13$.

6. Кубический кусок сыра, имеющий объем A^3, заплесневел по поверхности на глубину $A/8$, считая от каждой грани. Кусок разрезали на равные кубики параллельно граням. Объем каждого кубика равен $1/64$ от A^3. Какова вероятность, что случайно выбранный кусок не имеет плесени?

ПОДСКАЗКА. Классическая схема.

ОТВЕТ: $8/64$.

7. Из полного набора домино выбирают одну кость и кладут на стол. Затем извлекают вторую кость. Найти вероятность, что вторую кость можно приложить к первой.

ПОДСКАЗКА. Первая кость — дубль или не дубль. Формула полной вероятности.

ОТВЕТ: $\frac{7}{28} \cdot \frac{6}{27} + \frac{21}{28} \cdot \frac{12}{27}$.

8. Вероятность, что двойняшки одного пола, в два раза больше, чем разного. Равновозможны последовательности: мж и жм. Вероятность, что первый — мальчик, равна 0.51. Достоверно, что будет двойня. Какова вероятность, что второй — мальчик, при условии, что уже известно, что первый — мальчик.

ПОДСКАЗКА. Рассмотреть полную группу событий: мм, мж, жм, жж. Воспользоваться определением условной вероятности.

ОТВЕТ: $\frac{0.51-1/6}{0.51}$.

9. Вероятность попадания в цель первым стрелком равна p_1, вторым — p_2. Результаты стрельбы независимы. Стрелки выстрелили одновременно. Какова вероятность, что ровно один попал в цель?

ПОДСКАЗКА. Умножение вероятностей.

ОТВЕТ: $p_1 \cdot (1-p_2) + p_2 \cdot (1-p_1)$.

10. В круг радиуса R вписан квадрат. Чему равна вероятность, что поставленные наудачу внутри круга две точки окажутся внутри квадрата?

ПОДСКАЗКА. Теорема умножения.

ОТВЕТ: $4/\pi^2$.

11. Три лампочки: A, B, C соединены параллельно. Вероятность исправной работы A равна 0.6, B — 0.7, C — 0.8. Вычислить вероятность событий: 1) горят три лампы; 2) горит хотя бы одна.

ПОДСКАЗКА. Произведение вероятностей.

ОТВЕТ: $0.6 \cdot 0.7 \cdot 0.8$; $1 - 0.4 \cdot 0.3 \cdot 0.2$

12. Имеются два ящика: в первом — три белых и два черных шара; во втором — четыре белых и четыре черных. Из первого ящика во второй перекладывают два шара случайным образом. После этого из второго ящика извлекают один шар. Определить вероятность, что он белый.

ПОДСКАЗКА. Формула полной вероятности.

ОТВЕТ: $\frac{3}{10} \cdot \frac{6}{10} + \frac{1}{10} \cdot \frac{4}{10} + \frac{6}{10} \cdot \frac{5}{10}$.

13. На гранях правильного тетраэдра проставлены номера: 1, 2, 3, 4. Тетраэдр бросают 99 раз. Найти вероятнейшее число его приземлений на грань номер 3.

ПОДСКАЗКА. Схема Бернулли.

ОТВЕТЫ: 24 или 25.

14. Сколько раз надо бросить тетраэдр, чтобы наивероятнейшее число приземлений на грань номер 3 равнялось 30?

ПОДСКАЗКА. Схема Бернулли.

ОТВЕТЫ: от 119 до 123, включительно.

15. Из 13 упавших бутербродов наивероятнейшее число упавших вниз маслом равно девяти. Оценить вероятность падения бутерброда маслом вниз.

ПОДСКАЗКА. Схема Бернулли.

ОТВЕТ: от 9/14 до 10/14.

2.9.2. Второй уровень сложности

1. Десять разных шаров разложены в четыре разных ящика. Каждому шару равновозможно попасть в любой ящик. Найти вероятность событий: а) все шары в одном ящике; б) все шары в двух ящиках; в) нет пустых ящиков; г) в первом ящике — один шар, во втором — два, в третьем — три, в четвертом — четыре.

ПОДСКАЗКА. Произведение моделей. Размещения и перестановки с повторениями.

2. Коэффициенты трехчлена $x^2 + bx + c$ — случайные вещественные числа, по модулю не превосходящие единицы. Какова вероятность, что трехчлен не имеет вещественных корней? Что он имеет вещественные положительные корни?

ПОДСКАЗКА. На плоскости рассмотреть точки с координатами $(b,\ c)$.

3. Из начала координат в первый квадрант выпущен луч. Все углы с осью OX равновозможны. Найти вероятность, что луч пересечет круг с центром в точке $(3,\ 5)$ и радиусом 1.

ПОДСКАЗКА. Построить геометрическую модель на отрезке $[0,\ \pi/2]$.

4. На полке 50 книг. В 10 книгах сын нарисовал чертиков. Отец наугад берет три книги. Если хоть в одной он найдет чертика, то сын будет наказан. Вероятность найти чертика в книге, если он там есть, равна 0.3. С какой вероятностью сын будет наказан?

ПОДСКАЗКА. Полная вероятность+схема Бернулли.

5. В вазе было шесть конфет. Вася съел две, фантики набил бумагой и положил в вазу. Его братец Петя взял три конфеты, возможно, не все полноценные. Полноценные съел, все три набил бумагой и положил в вазу. Снова Вася берет одну конфету. Найти вероятность, что она полноценная.

ПОДСКАЗКА. Формула полной вероятности.

6. В условия задачи 5 внесено изменение. Снова Вася берет одну конфету. Она оказалась неполноценной. Найти вероятность, что Петя съел конфет больше, чем Вася.

ПОДСКАЗКА. Формула Байеса.

7. В правом кармане — три жетона метро и четыре монеты по 50 рублей, в левом — шесть жетонов и три монеты по 50 рублей. Из правого кармана в левый переложили пять предметов. Какова теперь вероятность вытащить наугад из левого кармана жетон?

ПОДСКАЗКА. Формула полной вероятности. Перебор составов переложенных предметов.

8. В тесто положили сто изюминок. Тщательно перемешали, разделили на десять равных частей и испекли десять булочек. Найти вероятность, что в случайно выбранной булочке будет а) не менее одной изюмины; б) ровно десять.

ПОДСКАЗКА. Совместные сто независимых испытаний для изюминок с вероятностью $1/10$.

9. Каждый из трех шаров попадает в один из трех ящиков равновозможно, независимо от двух других шаров. Событие A: в первом ящике нет шаров. Событие B: первый шар попал в третий ящик. Определите, являются ли эти события независимыми.

10. На первом этаже семиэтажного дома в лифт вошли три человека. Предположим, что каждому равновозможно выйти на любом этаже от второго до седьмого. Найдите вероятность, что все выйдут а) на одном этаже; б) на разных; в) ровно на двух.

ПОДСКАЗКА. Размещения трех разных шаров в семь разных ящиков.

11. Про группу коротышек известно, что из них взяли в поход печенье $\geqslant 47\%$, конфеты $\geqslant 38\%$, пироги $\geqslant 42\%$. Оценить вероятность того, что случайно выбранный коротышка взял в поход все три вида закуски.

ПОДСКАЗКА. Теорема сложения.

12. Получена партия из восьми изделий одного образца. По данным проверки половины партии три изделия оказались исправными, а одно — бракованным. Какова вероятность, что при проверке трех последующих изделий одно окажется исправным, а два — бракованными, если любое число бракованных в партии равновозможно?

ПОДСКАЗКА. Формулы Байеса и полной вероятности.

13. В ящике три шара: черный, красный и белый. Пять раз извлекалось по одному шару с последующим возвращением. Определить вероятность, что черный и белый шары извлекались не менее, чем по два раза.

14. Внутри отрезка [0, 10] выделяется случайный отрезок единичной длины. С какой вероятностью он будет содержать точку 5?

15. Двенадцать разбойников делят 30 одинаковых монет так, что любая монета может попасть к любому разбойнику с одинаковой вероятностью. Найти вероятность, что ровно два разбойника ничего не получат.

16. (Шевалье де Мере). Сколько раз нужно бросить пару игральных костей, чтобы с вероятностью, большей 0.5, ожидать сумму очков, равную 12, хотя бы один раз?

17. По клиническому анализу крови врач выявляет пневмонию с вероятностью 0.8, если она есть, и ошибочно указывает на нее при ее отсутствии с вероятностью 0.2. У направленных на анализ по подозрению в пневмонии эта болезнь встречается с вероятностью 0.7. Найти вероятность того, что пациенту, направленному на анализ, будет поставлен диагноз «пневмония». Вычислить вероятность того, что у пациента пневмонии нет при условии, что по анализу поставлен такой диагноз.

18. Из ста монет одна является фальшивой: у нее с двух сторон изображен герб. Берем случайную монету из ста и подбрасываем четыре раза. Все четыре раза выпал герб. Найти вероятность, что монета не фальшивая.

19. На карточках написаны числа от 1 до 1000. Наудачу извлекается одна карточка. Найти вероятность, что число, написанное на ней имеет в записи не менее двух цифр, равных 3.

20. Охотник стреляет в бегущую дичь до первого попадания. Вероятность попасть при одном выстреле равна 0.6. Найти вероятность, что он попадет, сделав не более трех выстрелов.

21. В партии из 1000 ламп каждая из ламп может быть бракованной с вероятностью 0.05. Выбираются любые четыре лампы из этой партии. Найдите вероятность, что среди них есть бракованные.

22. В группе из десяти человек равновозможно любое количество любителей футбола. Найдите вероятность того, что в случайной выборке из трех представителей этой группы будет ровно два любителя футбола.

23. На олимпиаде предложено две задачи ценой в одно очко (вероятность решить такую равна 0.9); три задачи по два очка (вероятность решить такую равна 0.7); две задачи по пять очков (вероятность решить такую равна 0.5). При условии, что все задачи решаются участником независимо, найти вероятность того, что он наберет не менее 16 очков.

24. Каждое из 20 изделий общества с ограниченной ответственностью может равновозможно оказаться изделием первого, второго или третьего сорта (при госпроверке). Найти вероятность того, что из 20 изделий не более трех относятся к третьему сорту.

25. В первой коробке лежат два черных шара и три белых. Во второй — три черных и два белых. Из каждой коробки выбирают случайным образом по одному шару и кладут в третью коробку, где уже лежат один черный шар и четыре белых. Затем, из третьей коробки достают какой-то случайный шар. Найти вероятность того, что он черный.

26. По прогнозу погоды на октябрь вероятность дождя в любой день равна 0.6, вероятность снега — 0.2, снега и дождя вместе — 0.1. Найти вероятность того, что в течение дня не будет ни дождя, ни снега. Найти вероятность того, что три дня подряд будет идти дождь, а снег идти не будет.

27. В условиях предыдущей задачи вероятность прогулки равна 0.8, если пойдет только снег; она равна 0.2, если пойдет снег с дождем; если пойдет только дождь, эта вероятность равна 0.3; в день без осадков — 0.9. Найти вероятность прогулки в случайный день октября.

28. Любительница теории вероятностей имеет три пары разных серег. В каждой паре серьги одинаковые. Все шесть штук серег перемешаны в коробке. Каждое утро берутся две случайные серьги и надеваются. Найдите наивероятнейшее число дней из шести, когда будут надеты парные серьги.

29. ЭВМ составляет вариант самостоятельной работы, выбирая случайно три задачи из 20, среди которых имеются 10 трудных и 10 легких. Найдите вероятность того, что среди выбранных трех задач будут хотя бы две трудные.

30. В условиях предыдущей задачи вероятность получить зачет за самостоятельную работу равна 0.2, если все три задачи — трудные; если трудных задач ровно две, эта вероятность равна 0.5; если одна — 0.7; если все задачи легкие — 0.9. Найти полную вероятность получить зачет за самостоятельную работу.

Глава 3. Случайные величины

Введение

При изучении случайных событий рассматриваются элементарные исходы экспериментов, имеющие любую природу. Например, при бросании монеты элементарными исходами считаются выпадения герба или цифры, при распределении частиц по энергетическим уровням — различные комбинации пустых и занятых уровней, при стрельбе по мишени — попадание или промах. В некоторых случаях результат эксперимента выражается числом, подсчитывается или измеряется с помощью эталона. Так, например, при многократном бросании монеты подсчитывается число выпавших гербов, при случайном выборе двух точек на отрезке определяется расстояние между ними. Многие биологические, химические, медицинские эксперименты так же связаны с измерениями изучаемых объектов, причем результаты таких измерений являются случайными числами, закономерность получения которых требуется установить.

Определение случайной величины. Пусть все простейшие логически возможные исходы эксперимента описаны как элементы пространства элементарных событий Ω, и каждому элементарному событию сопоставлено одно определенное вещественное число. Тогда говорят, что задана *случайная величина*.

Таким образом, случайная величина является отображением пространства элементарных событий Ω в множество действительных чисел.

Примеры случайных величин.

1. Число гербов, выпавшее при 10 бросаниях монеты. Ω — все 1024 упорядоченных набора по 10 символов «г» или «ц».
2. Длина листа, сорванного с конкретного дерева. Ω — все листья дерева.

3. Время ожидания автобуса на остановке. Пусть известно, что автобус прибывает на остановку регулярно через каждые T минут. Если элементарный исход опыта — одно, определенное значение времени ожидания автобуса, то $\Omega = [0, T]$. (Если автобус ходит не регулярно, то приходится рассматривать все неотрицательные числа.) Множество значений рассматриваемой случайной величины совпадает с ее областью определения.

4. Число произведенных выстрелов при стрельбе до первого попадания. Здесь также Ω совпадает с множеством значений случайной величины и представляет собой множество всех целых неотрицательных чисел.

При изучении случайных величин перед исследователем возникают следующие естественные задачи: найти множество значений случайной величины; выяснить, с какой вероятностью принимается каждое значение; определить, с какой вероятностью значения случайной величины попадают в заранее заданное множество; предсказать среднее значение случайной величины, которое можно ожидать при проведении n опытов; оценить вероятность отклонения наблюдаемых значений случайной величины от ожидаемого среднего. Этим и другим проблемам, связанным со случайными величинами, посвящается данная часть нашего учебного пособия.

В заключение отметим, что введение случайных величин помогает обобщать свойства вероятностных пространств с элементами различной природы. Отображение исходов эксперимента на числовую ось позволяет абстрагироваться от его физического содержания и изучить результаты не одного, а всех идентичных в некотором смысле экспериментов.

3.1. Общие принципы описания и исследования случайных величин

Мы определили случайную величину как отображение, следовательно, для того, чтобы задать конкретную случайную величину, связанную с данным опытом, необходимо указать ее область определения и правило, по которому вычисляются значения. Областью определения случайной величины является пространство элементарных событий Ω. Это множество, элементы которого описывают все логически возможные, попарно несовместные исходы рассматриваемого опыта. Примеры пространств элементарных событий, на которых заданы случайные величины, приводились в предыдущем разделе. Правило, по которому вычисляются значения случайной величины, обычно связано с описанием опыта, проводимых в его рамках измерений или подсчетов.

Первый вопрос, с которым мы сталкиваемся при исследовании случайной величины, — это вопрос о нахождении множества ее значений. Иногда все логически возможные значения можно перечислить. Например, при 10 бросаниях монеты возможно любое целое число выпадений герба от 0 до 10. К сожалению, во многих случаях точно указать множество значений случайной величины нельзя. Измеряя, к примеру, вес случайно выбранного человека, мы не можем определить точные границы, в которых будет содержаться полученное число. Аналогичная ситуация наблюдается в рассмотренных выше втором и третьем примерах случайных величин. В таких случаях приходиться рассматривать какое-нибудь множество, заведомо содержащее все возможные значения случайной величины, в частности, все множество $\mathbb{R}$ действительных чисел.

Второй вопрос, связанный с изучением случайной величины, — это вопрос о том, с какой вероятностью она принимает свои значения, а также с какой вероятностью значения принадлежат заданному множеству. Для ответа на этот вопрос необходимо построить вероятностную модель на множестве значений случайной величины или на каком-то более широком множестве. Универсальным ответом является построение модели на $\mathbb{R}$.

Определение закона распределения. Вероятностная модель на множестве V_ξ значений случайной величины ξ или на более широком множестве $\mathbb{R}$ называется *законом распределения вероятности* (более кратко — *распределением*) на этом множестве.

Задача 3.1. *Случайная величина ξ — это число очков, выпадающее при одном бросании игральной кости. Найти закон распределения вероятности на множестве V_ξ всех значений ξ.*

Область определения случайной величины: различные способы падения кубика при бросании. Этим способам однозначно соответствуют числа на гранях. Предположим, что равновозможно выпадение любой из граней, следовательно, равновозможно выпадение любого целого числа очков от 1 до 6. Построим на множестве $V_\xi = \{1, 2, 3, 4, 5, 6\}$ классическую вероятностную модель: каждому элементу V_ξ сопоставим вероятность $1/6$, событиями назовем все подмножества V_ξ, их вероятности найдем суммированием.

Задача 3.2. *Случайная величина ξ — это число гербов, которое может выпасть при четырех бросаниях монеты. Найти распределение вероятности на V_ξ.*

Область определения ξ — множество последовательностей из четырех символов (г или ц). Каждой последовательности сопоставлено число гербов. Очевидно, что $V_\xi = \{0, 1, 2, 3, 4\}$. Предполагая, что опыты по подбрасыванию монеты проходят по схеме Бернулли, где $n = 4$, $p = 0.5$, $q = 0.5$, определим вероятности элементарных событий в V_ξ:

$$\mathsf{P}_\xi(\{0\}) = 0.5^4, \quad \mathsf{P}_\xi(\{1\}) = C_4^1\, 0.5^4, \quad \mathsf{P}_\xi(\{2\}) = C_4^2\, 0.5^4,$$
$$\mathsf{P}_\xi(\{3\}) = C_4^3\, 0.5^4, \quad \mathsf{P}_\xi(\{4\}) = 0.5^4.$$

Вероятности остальных событий в V_ξ находим суммированием.

Задача 3.3. *Случайная величина ξ — вес случайно выбранного яблока. Построить для нее вероятностную модель (найти закон распределения).*

Поскольку множество значений ξ не определено (мы не можем взвесить все существующие в мире яблоки), то попробуем построить модель на $\mathbb{R}$. Необходимо задать поле событий F_ξ, удовлетворяющее аксиомам Колмогорова. Элементарными событиями будут одноэлементные подмножества $\mathbb{R}$. Яблок в мире очень много, поэтому вероятность получить случайным образом очень точный конкретный вес a крайне мала. Для неположительных a эта вероятность, очевидно, равна нулю. Для положительных — «приблизительно» равна нулю. В своей модели положим ее равной нулю.

Практический интерес представляет вероятность попадания веса в заранее указанный интервал (a, b) на действительной оси. Для определения этой вероятности у нас нет никаких теоретических предпосылок. В таких случаях проводятся массовые статистические исследования: взвешивания разнообразных партий яблок различных сортов. Вычисляется средняя относительная частота попадания веса в заданный интервал. Интервалов в $\mathbb{R}$ бесконечно много, нельзя для каждого проделать статистическое исследование. Приходится разбить $\mathbb{R}$ на конечное число интервалов и попытаться вывести общую закономерность «распределения частот по интервалам».

На этом пути в теории вероятностей получены типовые модели с параметрами, которые подбираются по конкретной задаче с помощью статистического материала. В таких моделях борелевское поле событий F_ξ представляет собой все интервалы вида $(-\infty, x)$, а также результаты их конечного или счетного объединения, пересечения, дополнения до $\mathbb{R}$.

Следует различать два способа построения вероятностной модели на V_ξ или на $\mathbb{R}$. В первом случае она строится на основе обобщения статистического материала. Во втором случае — на основе детального анализа того, какая вероятностная модель может быть построена на пространстве элементарных событий Ω. Конструируя вероятностную модель на $\mathbb{R}$ или на каком-то его подмножестве, содержащем все значения случайной величины, приходится учитывать уже известные вероятности случайных событий в Ω. Проиллюстрируем сказанное еще одним простым примером.

Задача 3.4. *Стрелок попадает в мишень с постоянной вероятностью* 0.6. *У него имеется всего три патрона, и он ведет стрельбу до первого попадания. Случайная величина ξ — число истраченных патронов. Множество V_ξ, очевидно, равно* $\{1, 2, 3\}$. *Построить на нем вероятностную модель* $(V_\xi, F_\xi, \mathsf{P}_\xi)$.

Введем на V_ξ борелевское поле событий F_ξ, состоящее из всех его подмножеств. Каждому подмножеству надо приписать вероятность. Это можно сделать, рассмотрев пространство элементарных событий Ω, состоящее из трех элементов: ω_1 — «попал с первого выстрела»; ω_2 — «первый выстрел — промах, второй — попадание»; ω_3 — «первые два выстрела – промахи, третий — любой». По теореме умножения вероятностей можно вычислить вероятности элементарных исходов в Ω:

$$\mathsf{P}_\Omega(\omega_1) = 0.6, \quad \mathsf{P}_\Omega(\omega_2) = 0.4 \cdot 0.6, \quad \mathsf{P}_\Omega(\omega_3) = 0.4 \cdot 0.4 \cdot 1.$$

Случайная величина ξ отображает Ω на V_ξ так, что ω_1 соответствует число 1, ω_2 соответствует число 2, ω_3 соответствует число 3. При этом случайная величина принимает значение 1 тогда и только тогда, когда наступает событие ω_1; значение 2 — тогда и только тогда, когда наступает ω_2; значение 3 — тогда и только тогда, когда наступает ω_3. Поэтому положим $\mathsf{P}_\xi(1) = 0.6$, $\mathsf{P}_\xi(2) = 0.24$, $\mathsf{P}_\xi(3) = 0.16$. Поскольку всякое событие A из F_ξ представляет собой объединение элементов V_ξ, то $\mathsf{P}_\xi(A)$ находится суммированием вероятностей составляющих его элементарных событий из V_ξ. Эта вероятность, с другой стороны, равна $\mathsf{P}_\Omega(B)$, где B — это объединение таких элементов Ω, которые ξ отображает в A. Например, если $A = \{1; 2\}$, то

$$\mathsf{P}_\xi(A) = \mathsf{P}_\xi(\,\{1\}\,) + \mathsf{P}_\xi(\,\{2\}\,) = \mathsf{P}_\Omega(\omega_1) + \mathsf{P}_\Omega(\omega_2) = \mathsf{P}_\Omega(B), \quad B = \omega_1 \cup \omega_2.$$

Таким образом, для вычисления вероятности события A из F_ξ находим такое событие B из F_Ω, что A имеет место тогда и только тогда, когда B.

Подведем итоги сказанному выше и выделим основные принципы построения вероятностных моделей на множестве V_ξ значений случайной величины ξ (аналогично, на включающем его множестве $\mathbb{R}$).

1. Если на области определения Ω случайной величины ξ не удается построить подходящую стандартную вероятностную модель, с помощью которой можно было бы найти вероятности попадания значений ξ в интересующие исследователя подмножества V_ξ, то применяется статистическое исследование с последующим подбором одной из стандартных моделей на V_ξ или более широком множестве.

2. Для данного подмножества A множества V_ξ назовем *прообразом* в пространстве Ω множество B, состоящее из всех таких элементарных событий, для которых значения ξ принадлежат A. Если A объявлено событием в V_ξ, и существует модель $(\,\Omega, F_\Omega, \mathsf{P}_\Omega\,)$ на Ω, в которой B является событием, то вероятности B и A должны быть связаны соотношением

$$\mathsf{P}_\xi(A) = \mathsf{P}_\Omega(B). \tag{3.1}$$

Действительно, ведь событие A происходит тогда и только тогда, когда B. Если на V_ξ введено борелевское поле событий F_ξ , и удалось так подобрать модель $(\,\Omega, F_\Omega, \mathsf{P}_\Omega\,)$, что всякое событие из F_ξ имеет прообразом событие в F_Ω, то соотношение (3.1) задает вероятностную функцию на V_ξ. Чтобы доказать это утверждение, надо проверить, выполняются ли для F_ξ аксиомы вероятности. Последнее следует из того, что прообразы множеств, не имеющих общих точек, также не имеют общих точек; а прообраз объединения множеств является объединением их прообразов. (Докажите самостоятельно.)

Задача 3.5. *Известно, что некий пакет весит целое число килограмм, равновозможное от 1 до 7. Его взвешивают на весах с помощью гирь из набора: 1, 2, 4 (кг). Гири кладут на одну чашку весов, пакет — на другую. Число использованных гирь — случайная величина ξ. Построить вероятностную модель на V_ξ.*

Рассмотрим подробно функцию ξ. Для этого составим таблицу:

Вес пакета (в кг)	1	2	3	4	5	6	7
Число гирь (ξ)	1	1	2	1	2	2	3

Заметим, что число использованных гирь всегда определено однозначно (в противном случае мы не имели бы функции). Перестроим теперь эту таблицу так, чтобы каждому значению ξ соответствовал его прообраз:

Число гирь (ξ)	1	2	3
Вес пакета	{1,2,4}	{3,5,6}	{7}

В верхней строке — все элементы V_ξ, в нижней — их прообразы, являющиеся событиями в классической модели. По этой модели вычислим вероятности указанных прообразов и положим их равными вероятностям соответствующих значений ξ:

Число гирь (ξ)	1	2	3
Вероятность	3/7	3/7	1/7

Вероятности остальных событий в дискретном пространстве V_ξ находятся суммированием.

Задача 3.6. *На отрезке AB единичной длины случайным образом выбирается точка так, что вероятность ее попадания в любую часть отрезка, имеющую длину, вычисляется по геометрической модели (то есть равна этой длине, поскольку длина AB равна единице). Найти закон распределения ξ — расстояния от выбранной точки до середины отрезка C.*

Область определения ξ — отрезок AB. Множество ее значений — числовой отрезок $[0,\, 1/2] = V_\xi$. Объявим в V_ξ событиями все точки и все промежутки вида $(\,a,\, b\,)$, входящие в $[0,\, 1/2]$, а также все множества, которые могут получаться из указанных операцией дополнения до $[0,\, 1/2]$. Объявим событиями в V_ξ и все множества, которые получаются в результате не более, чем счетного объединения или пересечения ранее определенных в V_ξ событий. В результате получится борелевское поле событий F_ξ, удовлетворяющее аксиомам. Прообразом элементарного события $\{a\}$ является пара точек $\{A_1; A_2\}$ на AB, равноудаленных от C (рис. 3.1). Следовательно, вероятность получить конкретное расстояние равна нулю. Прообразом числового интервала $(\,a, b\,)$ является объединение открытых отрезков B_1A_1 и A_2B_2, симметричных относительно точки C (эти отрезки выделены жирными линиями на рис. 3.1).

Сумма длин этих отрезков равна $2(b-a)$, следовательно, вероятность события $a < \xi < b$, равна $2(b-a)$. Вероятности остальных событий в F_ξ находятся по аддитивности.

В случае, когда какой-то из перечисленных этапов поиска закона распределения случайной величины не удается реализовать, пользуются помощью статистики. Но и она не всегда помогает полностью выявить вероятностные свойства этой величины.

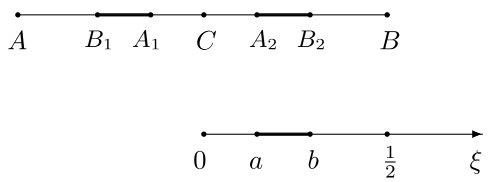


Рис. 3.1

В большинстве книг по теории вероятностей отображения, заданные на Ω, для которых не удается построить вероятностную модель на $\mathbb{R}$, где событиями объявлена борелевская σ-алгебра интервалов, по определению не рассматриваются как случайные величины. Наше определение более широко, из него вытекает, что случайная величина может не иметь вероятностной модели или иметь разные модели на V_ξ (или $\mathbb{R}$). Такое определение мы считаем более удобным для постановки и решения исследовательских задач.

3.2. Дискретные случайные величины

В приведенных ранее примерах некоторые случайные величины имели не более, чем счетное число различных значений.

Определение дискретной случайной величины. Случайная величина, множество значений которой не более, чем счетно, называется *дискретной случайной величиной*.

Распределение вероятности на множестве V_ξ значений дискретной случайной величины ξ представляет собой модель, в которой событиями являются все подмножества V_ξ. Для построения такой модели достаточно определить вероятности всех значений случайной величины ξ. Вероятности остальных событий в V_ξ находятся суммированием.

Определение закона распределения дискретной случайной величины. *Законом (рядом) распределения* дискретной случайной величины называется следующая таблица:

Значения ξ	x_1	x_2	$\dots$	x_n	$\dots$
Вероятности	p_1	p_2	$\dots$	p_n	$\dots$

В верхней строке этой таблицы помещены все значения данной дискретной случайной величины, расположенные в порядке нестрогого возрастания, в нижней — соответствующие им вероятности.

Если закон распределения на V_ξ определяется при помощи вероятностной модели (Ω, F_Ω, P_Ω) на пространстве элементарных исходов опыта, то каждому значению x_i величины ξ приписывается вероятность события B из F_Ω, состоящего из всех элементарных исходов, которым соответствует одно и то же значение x_i. В случае, когда само пространство Ω не более, чем счетно, введенная на нем модель обязательно дискретна, и всякое подмножество Ω принадлежит F_Ω. Следовательно, указанный способ определения вероятности значений ξ корректен. В случае несчетного Ω вопрос о корректности такого определения остается открытым.

Очевидно, что всякая таблица рассмотренного выше вида (где в верхней строке все числа различны, а в нижней строке все числа принадлежат отрезку [0, 1], и их сумма равна единице) определяет некоторую случайную величину, например, такую у которой каждое значение соответствует ровно одной точке Ω.

Рассмотрим теперь некоторые широко применяемыемые законы распределения дискретных случайных величин.

Равномерное дискретное распределение. Пусть ряд распределения дискретной случайной величины имеет вид:

Значения ξ	x_1	x_2	$\dots$	x_n
Вероятности	$1/n$	$1/n$	$\dots$	$1/n$

Тогда соответствующее ему распределение называется *равномерным дискретным*.

Если пространство Ω состоит из n равновозможных элементарных событий, а случайная величина ξ сопоставляет каждому элементарному событию ω_i свое, отличное от других, число x_i (например, число очков, выпадающее на грани игрального кубика при одном бросании), то, очевидно, ξ имеет равномерное дискретное распределение.

Геометрическое распределение. Пусть p и q — неотрицательные числа, и $q = 1 - p$. Распределение дискретной случайной величины называется *геометрическим*, если ее ряд распределения имеет вид:

Значения ξ	1	2	$\dots$	k	$\dots$
Вероятности	p	$q \cdot p$	$\dots$	$q^{k-1} \cdot p$	$\dots$

Геометрическое распределение имеет *число опытов* , проводимых в одних и тех же условиях *до первого наступления события* A, если вероятность события A в каждом опыте равна p. В таких случаях говорят, что опыты проводятся до первого «успеха». Вероятность того, что

A наступит в первом же опыте, равна p. Этому событию соответствует $\xi = 1$. Если в первом опыте A не произошло, а во втором произошло, то $\xi = 2$, и вероятность этого события равна $q \cdot p$ (по теореме умножения). Аналогично вычисляется вероятность, с которой $\xi = k$ (в первых $k-1$ опытах событие A не произошло, а в k-м опыте произошло). Разобраться в устройстве Ω предоставляем читателю в качестве упражнения.

Биномиальное распределение. Пусть n опытов проводятся по «схеме Бернулли», в которой вероятность наступления «успеха» (события A) равна p, а «неуспеха» (события $\bar{A}$), соответственно, $q = 1 - p$. Число «успехов» из n — это дискретная случайная величина. Обозначим ее ξ. Ряд распределения ξ, очевидно, имеет вид:

Значения ξ	0	1	$\dots$	k	$\dots$	n
Вероятности	q^n	$C_n^1 \cdot p \cdot q^{n-1}$	$\dots$	$C_n^k \cdot p^k \cdot q^{n-k}$	$\dots$	p^n

Соответствующее такому ряду распределение дискретной случайной величины называется *биномиальным*.

Биномиальное распределение имеют, например, такие случайные величины: число гербов, выпадающее при n бросаниях монеты; число попаданий в цель при n выстрелах (если вероятность попадания при одном выстреле не меняется).

Индикатор события. Пусть случайная величина ξ принимает значение 1, если результатом опыта является событие A, имеющее вероятность p. Если результатом опыта является событие $\bar{A}$, то ξ принимает значение 0. Областью определения такой случайной величины является пространство Ω, состоящее из двух элементарных событий: A и $\bar{A}$. Множество значений ξ равно $V_\xi = \{\, 0,\, 1 \,\}$.

Описанная случайная величина называется *индикатором события* A. Ее закон распределения полностью определен следующей таблицей.

Значения ξ	0	1
Вероятности	$1-p$	p

Очевидно, что индикатор события A имеет биномиальное распределение, в котором $n = 1$.

Распределение Пуассона Рассмотрим случайную величину ξ, которая может принимать только любые целые неотрицательные значения: $0, 1, \dots, k, \dots$ Говорят, что случайная величина ξ *имеет распределение*

Пуассона, если вероятность того, что она примет определенное значение k, выражается формулой

$$\mathsf{P}_\xi(k) = \frac{\lambda^k}{k!}\, e^{-\lambda}, \tag{3.2}$$

где λ — постоянный положительный параметр.

Ряд распределения Пуассона (синоним — *закон Пуассона*) имеет вид:

Значения ξ	0	1	2	$\dots$	k	$\dots$
Вероятности	$e^{-\lambda}$	$\lambda \cdot e^{-\lambda}$	$\dfrac{\lambda^2}{2!}e^{-\lambda}$	$\dots$	$\dfrac{\lambda^k}{k!}e^{-\lambda}$	$\dots$

Легко убедиться, что сумма вероятностей во второй строке таблицы равна единице. Это следует из того, что сумма ряда $\sum_{k=0}^{\infty} \lambda^k/k!$ равна e^{λ}.

Как известно из предельной теоремы Пуассона, формула (3.2) получается из формулы Бернулли (2.14) при условии, что n стремится к бесконечности, а произведение $n{\cdot}p = \lambda$. Эта формула может быть использована для расчета вероятности события “при n испытаниях «успех» A наступил k раз если испытания проводятся по схеме Бернулли при больших (более 300) значениях n и настолько малых p, что $n \cdot p < 10$. Произведение $n \cdot p$ и обозначается через λ в формуле (3.2). Поскольку здесь вероятность «успеха» A в одном испытании очень мала, закон Пуассона называют также *законом редких явлений*.

В качестве примера редкого явления рассмотрим телефонный звонок в Вашу квартиру в определенный промежуток времени. Число n в этом случае — это число всех людей, которые «теоретически» могут вам позвонить, их, как правило, очень много (ведь звонки могут быть и ошибочными). Вероятность p того, что позвонит конкретный человек, для разных людей различна, но ее усредненное значение мало. Если, по вашим наблюдениям, среднее число звонков в рассматриваемый промежуток времени равно λ, то для расчета вероятности k звонков за этот промежуток можно использовать формулу (3.2). Это подтверждается статистически.

Методами математической статистики установлено, что распределение Пуассона имеют такие случайные величины, как число опечаток на странице машинописного текста, поступление вызовов на АТС в единицу времени, число покупателей в магазине в определенный час и тому подобные.

Гипергеометрическое распределение. Рассмотрим задачу о выборке данного состава. Пусть даны n предметов, из которых ровно m

отмечены. Все n предметов тщательно перемешаны. Наугад без возвращения выбираются n_1 предметов из n. Рассмотрим случай, когда $n_1 \leqslant m$ и $n_1 \leqslant n - m$. Обозначим через ξ число отмеченных предметов среди n_1 выбранных. Очевидно, что ξ — это дискретная случайная величина, имеющая множество значений $\{0, 1, \ldots, n_1\}$. Ряд распределения для ξ имеет вид

Значения ξ	0	$\ldots$	k	$\ldots$	n_1
Вероятности	$\dfrac{C_{n-m}^{n_1}}{C_n^{n_1}}$	$\ldots$	$\dfrac{C_m^k C_{n-m}^{n_1-k}}{C_n^{n_1}}$	$\ldots$	$\dfrac{C_m^{n_1}}{C_n^{n_1}}$

Распределение дискретной случайной величины, соответствующее такому ряду, называется *гипергеометрическим*. Аналогичные формулы могут быть записаны для выборок, где содержатся элементы трех и более сортов, а также для случая, когда $n_1 > m$ или $n_1 > n - m$. Важно, что мы «следим» только за одним из сортов.

Рассмотрим теперь несколько типовых задач, в которых требуется выдвинуть гипотезу о законе распределения данной случайной величины и произвести расчет вероятностей по соответствующим гипотезе формулам. Для разнообразия будем обозначать случайные величины различными буквами, а не только ξ.

Задача 3.7. *Каждая из десяти проб грунта может равновозможно быть кислой, нейтральной или щелочной. Случайная величина G — число щелочных проб из данных десяти. Найти закон распределения G и вычислить вероятность того, что из десяти проб щелочными окажутся хотя бы две.*

Для выдвижения гипотезы о законе распределения представим себе десять независимых испытаний, в каждом из которых определяется кислотность одной, определенной, пробы грунта. Очевидно, что в каждом испытании вероятность того, что проба является щелочной, должна быть равна $1/3$ (так как имеем три равновозможных исхода). Таким образом, испытания проводятся по схеме Бернулли, следовательно, закон распределения случайной величины G, предположительно, биномиальный. Для расчета нужной нам вероятности вычтем из единицы вероятность того, что щелочных проб меньше двух (нуль или одна):

$$\mathsf{P}_G(\{2, \ldots, 10\}) = 1 - \mathsf{P}_G(\{0, 1\}) = 1 - \left(\frac{2}{3}\right)^{10} - 10 \cdot \left(\frac{1}{3}\right) \cdot \left(\frac{2}{3}\right)^9$$

Задача 3.8. *Машинистка делает в среднем две опечатки на трех страницах текста. Случайная величина F – число опечаток на одной странице. Выдвинуть гипотезу о законе распределения F и вычислить вероятность того, что на странице не будет ни одной опечатки.*

Представим себе работу машинистки над страницей как повторные независимые испытания, проводимые по схеме Бернулли: каждый печатный знак (из n штук на странице) может быть напечатан верно или нет. Предположим, что вероятность опечатки в одном знаке постоянна и равна p. Она настолько мала, что произведение $n \cdot p = 2/3$. Исходя из таких допущений, предполагаем, что случайная величина F имеет распределение Пуассона, следовательно, искомая вероятность равна $\mathsf{P}_F(0) = e^{-2/3}$.

Задача 3.9. *Изыскатель бурит скважины в садоводстве «Ладога» до первого обнаружения чистой питьевой воды. Вероятность получить воду при каждом бурении постоянна и равна* 0.1. *Случайная величина T – число сделанных скважин. Найти закон распределения T и вероятность того, что будет сделано менее 40 попыток.*

Если предположить, что попытки делаются независимо друг от друга, то получаем стандартное геометрическое распределение с параметром $p = 0.1$. Искомая вероятность равна $\mathsf{P}_T(\{1, \ldots, 39\}) = 0.1 + 0.9 \cdot 0.1 + \cdots + (0.9)^{38} \cdot 0.1$. Используя формулу суммы для геометрической прогрессии, получим, что эта вероятность равна $1 - (0.9)^{39}$, что, примерно, равно 0.984.

Задача 3.10. *В пакете перемешаны семена декоративных тыкв разных сортов: двенадцать семян грушевидной, десять семян пупырчатой и пять семян круглой. Выбираем случайным образом три семечка из этого пакета. Случайная величина S – число семечек пупырчатой тыквы среди выбранных. Выдвинуть гипотезу о законе распределения S и найти вероятность того, что выбрано ровно одно семя пупырчатой тыквы.*

Если предположить, что все сочетания из 27 по три равновозможны, то получим гипергеометрический закон распределения. Искомая вероятность равна $\mathsf{P}_S(\{1\}) = C_{10}^1 \cdot C_{17}^2 / C_{27}^3$.

Задача 3.11. *Среди купленных на рынке десяти маек равновозможно от одной до трех имеют брак. Другие варианты отсутствуют. Случайная величина W – число маек из десяти без брака. Найти закон распределения W и вероятность того, что куплено не менее половины маек без брака.*

По условию не имеют брака девять, восемь или семь маек с одинаковой вероятностью. Других вариантов нет, следовательно, эта вероятность равна $1/3$. Величина W имеет равномерный дискретный закон распределения. Искомая вероятность равна 1.

3.3. Числовые характеристики дискретных случайных величин

Как отмечалось во введении, при изучении одномерной случайной величины возникает проблема предсказания среднего значения M, которое она может принимать при n измерениях. Кроме того, для случайных величин, имеющих большое количество возможных значений, актуальна проблема выделения «наиболее вероятной» части из всего множества значений. Последняя проблема может быть более четко сформулирована следующим образом: определить окрестность M, в которую значения случайной величины попадут с определенной вероятностью (например, 0.99). Для ответа на поставленные вопросы используются так называемые *числовые характеристики* случайных величин: математическое ожидание, дисперсия и среднее квадратическое отклонение. Здесь мы рассмотрим их определения для случайных величин дискретного типа.

Определение математического ожидания дискретной случайной величины. Пусть дискретная случайная величина ξ имеет известный закон распределения:

Значения ξ	x_1	x_2	$\dots$	x_i	$\dots$
Вероятности	p_1	p_2	$\dots$	p_i	$\dots$

Математическим ожиданием M_ξ случайной величины ξ называется сумма всех произведений вида $x_i \cdot p_i$:

$$M_\xi = x_1 \cdot p_1 + x_2 \cdot p_2 + \dots \tag{3.3}$$

Если множество значений ξ конечно, то математическое ожидание ξ представляет собой сумму нескольких чисел, следовательно, всегда существует. Если же множество значений ξ счетно, то M_ξ представляет собой сумму числового ряда (бесконечно много слагаемых). Такая сумма может быть не определена (ряд расходится). В таком случае говорят, что математическое ожидание не существует.

Покажем теперь, почему математическое ожидание является «предсказанием» среднего значения случайной величины ξ, которое она может принимать в результате n измерений. Пусть, действительно, эти n измерений сделаны. Их результатами являются числа: $X_1, X_2, \dots, X_n$. Найдем среднее арифметическое этих чисел:

$$M = \frac{X_1 + X_2 + \dots + X_n}{n}.$$

Если в числителе этой дроби привести подобные слагаемые, то он будет равен $x_1 \cdot a_1 + x_2 \cdot a_2 + \dots$, где $x_1, x_2, \dots$ — различные значения случайной величины, $a_1, a_2, \dots$ — их абсолютные частоты (то есть количества значений $x_1, x_2, \dots$, наблюдавшихся среди данных n результатов измерений). Если число измерений n велико (стремится к бесконечности), то все возможные значения ξ будут получены на опыте. Перепишем среднее значение в виде

$$M = x_1 \cdot \frac{a_1}{n} + x_2 \cdot \frac{a_2}{n} + \dots$$

Отношения абсолютных частот a_i к n называются относительными частотами событий вида $\xi = x_i$. При большом числе измерений эти относительные частоты должны мало отличаться от вероятностей p_i, иначе закон распределения неправильно подобран для данной случайной величины. Таким образом, при большом количестве измерений величина среднего значения M должна мало отличаться от M_ξ, если оно существует.

Приведем далее без доказательства формулы для вычисления математического ожидания случайных величин, имеющих стандартные дискретные распределения:

1) биномиальный закон: $M_\xi = n \cdot p$,
2) геометрический закон: $M_\xi = 1/p$,
3) индикатор: $M_\xi = p$,
4) закон Пуассона: $M_\xi = \lambda$,
5) гипергеометрический закон: $M_\xi = n_1 \cdot (m/n)$.

В случае, когда закон распределения не является стандартным, можно найти математическое ожидание по определению.

Задача 3.12. *В задачах 3.7–3.11 вычислить математические ожидания случайных величин.*

В задаче 3.7 имеем биномиальный закон распределения с параметрами $n = 10$, $p = 1/3$, следовательно, $M_G = 10/3$. В задаче 3.8 идет речь о распределении Пуассона, следовательно, $M_F = 2/3$. Случайная величина T в задаче 3.9 имеет геометрический закон распределения, поэтому $M_T =$

10. В задаче 3.10 идет речь о гипергеометрическом распределении, для которого $n = 27$, $m = 10$, $n_1 = 3$, следовательно, $M_S = 30/27$. В последней задаче, 3.11, вычислим математическое ожидание W по определению: $9 \cdot \frac{1}{3} + 8 \cdot \frac{1}{3} + 7 \cdot \frac{1}{3} = 8$.

Определение дисперсии дискретной случайной величины. Пусть дискретная случайная величина ξ имеет известный закон распределения:

Значения ξ	x_1	x_2	$\ldots$	x_i	$\ldots$
Вероятности	p_1	p_2	$\ldots$	p_i	$\ldots$

Пусть существует математическое ожидание M_ξ. Дисперсией случайной величины ξ называется сумма всех произведений вида $(x_i - M_\xi)^2 \cdot p_i$:

$$D_\xi = (x_1 - M_\xi)^2 \cdot p_1 + (x_2 - M_\xi)^2 \cdot p_2 + \ldots \tag{3.4}$$

Если множество значений ξ конечно, то дисперсия ξ, очевидно, существует. Если же множество значений ξ счетно, то D_ξ представляет собой сумму числового ряда, который может расходиться. В таком случае говорят, что дисперсия не существует. Для вычисления дисперсии можно применить более удобную формулу

$$D_\xi = {x_1}^2 \cdot p_1 + {x_2}^2 \cdot p_2 + \cdots - (M_\xi)^2 = M_{\xi^2} - (M_\xi)^2. \tag{3.5}$$

Равносильность формул (3.4) и (3.5) доказывается с помощью простых алгебраических преобразований. Приведем, далее, без доказательства формулы для вычисления дисперсии случайных величин, имеющих стандартные дискретные распределения:

1) биномиальный закон: $D_\xi = n \cdot p \cdot (1 - p)$,
2) геометрический закон: $D_\xi = (1 - p)/p^2$,
3) индикатор: $D_\xi = p \cdot (1 - p)$,
4) закон Пуассона: $D_\xi = \lambda$,
5) гипергеометрический закон: $D_\xi = n_1 \cdot (n - n_1) \cdot \dfrac{m \cdot (n - m)}{n^2 \cdot (n - 1)}$.

Задача 3.13. *В задачах 3.7–3.11 вычислить дисперсии случайных величин.*

В задаче 3.7 имеем биномиальный закон распределения с параметрами $n = 10$, $p = 1/3$, следовательно, $D_G = 20/9$. В задаче 3.8 идет речь о распределении Пуассона, следовательно, $D_F = 2/3$. Случайная величина T в задаче 3.9 имеет геометрический закон распределения, поэтому

$D_T = 90$. В задаче 3.10 идет речь о гипергеометрическом распределении, для которого $n = 27$, $m = 10$, $n_1 = 3$, следовательно,

$$D_S = \frac{30 \cdot 17 \cdot 24}{27^2 \cdot 26}\,.$$

В последней задаче, 3.11, вычислим дисперсию W по формуле (3.5): $81 \cdot \frac{1}{3} + 64 \cdot \frac{1}{3} + 49 \cdot \frac{1}{3} - 64 = \frac{2}{3}$.

Как нетрудно заметить, дисперсия измеряется не в таких единицах, как математическое ожидание: единицы измерения возводятся в квадрат. Это не всегда удобно. Для единообразия единиц измерения из дисперсии извлекают квадратный корень.

Определение среднего квадратического отклонения. Средним квадратическим отклонением случайной величины ξ (не обязательно дискретной) называется квадратный корень из ее дисперсии: $\sigma_\xi = \sqrt{D_\xi}$.

Рассмотрим теперь проблему определения минимального интервала вида $(M_\xi - r, M_\xi + r)$, в который значения ξ попадают с заданной вероятностью γ. Если закон распределения дискретной случайной величины известен, то такой интервал можно определить, добавляя одно за другим значения ξ слева и справа от M_ξ в множество $\{M_\xi\}$. При этом вероятности добавленных значений надо складывать, пока не получим число γ или число больше него. Для большинства случайных величин решение такой задачи требует больших вычислений. Существует простой, но приближенный метод получения интервала вида $(M_\xi - r,\, M_\xi + r)$, в который значения ξ попадают с вероятностью заведомо большей, чем γ. Этот метод не требует знания закона распределения, а только знание M_ξ, D_ξ и γ.

Теорема 3.1 (неравенство Чебышева). *Пусть случайная величина ξ имеет математическое ожидание M_ξ и дисперсию D_ξ. Тогда для любого положительного числа r имеет место неравенство:*

$$\mathsf{P}_\xi((M_\xi - r,\, M_\xi + r)) \geqslant 1 - \frac{D_\xi}{r^2}. \tag{3.6}$$

Это неравенство означает, что значения случайной величины ξ попадают в интервал с центром в точке M_ξ и любым заданным радиусом r с вероятностью, которая больше или равна $1 - D_\xi / r^2$.

Доказательство теоремы опускаем. Заметим, что она верна не только для дискретных, но и для абсолютно непрерывных случайных величин, описанных в следующих параграфах.

Для приближенного определения радиуса окрестности математического ожидания, в которую значения случайной величины попадают с вероятностью больше, чем γ, надо правую часть неравенства (3.6) приравнять числу γ.

Следствие. Если в неравенстве (3.6) вместо r подставить $3 \cdot \sigma_\xi$, то соответствующая вероятность будет не меньше, чем $8/9$. Этот факт называют *«правилом трех сигм»*.

Задача 3.14. *Для случайной величины T из задачи 3.9 найти с помощью неравенства Чебышева интервал $(M_T - r,\ M_T + r)$, в который значения T попадают с вероятностью более, чем 0.9. (Таким образом отбрасываются маловероятные значения из бесконечного множества значений.)*

Ранее мы уже подсчитали, что $M_T = 10$, $D_T = 90$. Для определения r из неравенства Чебышева решим уравнение $1 - 90/r^2 = 0.9$. Оно равносильно $r^2 = 900$, откуда $r = 30$. Для сравнения, $3 \cdot \sigma_T = 9 \cdot \sqrt{10}$, что, примерно, равно 30, так как 0.9 приближенно равно $8/9$. Таким образом, приближенно, с вероятностью больше 0.9 случайная величина T принимает значения от 1 до 39. В задаче 3.9 была определена вероятность для этого множества: она равна 0.984.

К неравенству Чебышева мы вернемся в разделе 3.8.

3.4. Задачи для самостоятельного решения

3.4.1. Первый уровень сложности

1. В елочной гирлянде старого образца все сто лампочек соединены последовательно: если одна лампочка «не в порядке» (неисправна или выкрутилась), то и вся гирлянда не горит. Вероятность того, что одна лампочка «не в порядке», равна 0.2 для каждой лампочки, независимо от других. Пусть случайная величина ξ — это число лампочек в гирлянде, которые «не в порядке». Найдите закон распределения ξ, математическое ожидание, дисперсию и $r(0.9)$ с помощью неравенства Чебышева.

ПОДСКАЗКА. Биномиальный закон.

2. Все лампочки в старой гирлянде, кроме одной, «в порядке». Эта одна лампочка может равновозможно находиться на любом из ста мест. Начинаем проверять всю цепочку, начиная с первой лампочки, пока не найдем ту, что «не в порядке». Теперь ξ — число проверенных лампочек.

Найдите закон распределения и математическое ожидание ξ. Определите вероятность события: $\xi \leqslant 20$.

ПОДСКАЗКА. Равномерное дискретное распределение.

3. Среднее количество бракованных изделий в большой партии равно четырем. Случайная величина X — количество бракованных изделий в такой партии. Определите вероятность события: $X \geqslant 3$.

ПОДСКАЗКА. Закон Пуассона.

4. Каждую из десяти сложных задач цикла для самостоятельной работы студентов сам преподаватель может решить правильно с вероятностью 0.8, независимо от других задач того же цикла. Случайная величина f — число задач из десяти, которые преподаватель может решить правильно. Найдите закон распределения f, математическое ожидание, дисперсию и вероятность события: $f \geqslant 8$.

ПОДСКАЗКА. Схема Бернулли.

5. Студент ходит переписывать самостоятельную работу до первого успеха. Вероятность успеха постоянна и равна 0.6, независимо от номера попытки. Пусть Y — число попыток до первого успеха, включительно. Найдите закон распределения Y, математическое ожидание, дисперсию и $r(0.9)$ с помощью неравенства Чебышева.

ПОДСКАЗКА. Геометрический закон.

6. Из десяти садоводов каждый выбирает сорта моркови для посева. С вероятностью 0.6 каждый садовод, независимо от других, посеет новый сорт «Королева осени». Пусть Z — число садоводов из десяти, посеявших этот сорт. Найдите закон распределения Z, математическое ожидание, дисперсию и вероятность события: $Z \leqslant 5$.

ПОДСКАЗКА. Схема Бернулли.

7. В старом пакете семян сельдерея — более 1000 семян, но среднее количество всхожих семян равно пяти. Пусть K — случайное количество всхожих семян в пакете. Определите закон распределения K, математическое ожидание, дисперсию и вероятность того, что в пакете есть хотя бы одно семя, способное взойти.

ПОДСКАЗКА. Закон Пуассона.

8. На окружности отметили четыре разные точки. Каждые две из них соединили отрезком. Каждый из шести полученных отрезков закрасили целиком в синий или в красный цвет (независимо от других). Вероятность того, что отрезок красный, равна 0.8 (синий — 0.2). Пусть W —

число синих отрезков из шести. Определите закон распределения W, математическое ожидание, дисперсию и вероятность того, что среди отрезков будет хотя бы один красный.

ПОДСКАЗКА. Биномиальный закон.

9. Мальчик бросает мяч в баскетбольную корзину до первого попадания. Вероятность попадания при одном броске постоянна и равна 0.3, независимо от номера броска. Пусть R — число попыток, считая удачную. Найдите закон распределения R, математическое ожидание, дисперсию и вероятность того, что мальчик попадет, сделав меньше трех попыток.

ПОДСКАЗКА. Геометрический закон.

10. В партии равновозможно может содержаться от нуля до пяти бракованных электроламп. Других вариантов нет. Случайная величина ξ — число бракованных ламп в этой партии. Найдите закон распределения, математическое ожидание, дисперсию и вероятность того, что в партии будет больше одной бракованной лампы.

ПОДСКАЗКА. Равномерный дискретный закон распределения.

11. В футбольном клубе «Закат» — 25 футболистов, среди которых пять иностранцев. По жребию выбираются 11 случайных футболистов из 25 для участия в турнире. Число иностранцев, попавших в выбранную группу из 11 человек, — случайная величина F. Найдите ее закон распределения и математическое ожидание. Подсчитайте вероятность того, что будут выбраны не более двух иностранцев.

ПОДСКАЗКА. Сочетания. Гипергеометрический закон.

12. Команда «Закат» играет серию матчей с командой «Рассвет». Всего в серии пять матчей. Вероятность того, что матч выиграет «Закат» равна 0.6, вероятность свести матч в ничью равна 0.2 (для всех матчей независимо). Пусть D — число матчей из пяти, выигранных командой «Рассвет». Найдите закон распределения D, математическое ожидание, дисперсию и вероятность того, что «Рассвет» победит хотя бы один раз.

ПОДСКАЗКА. Биномиальный закон.

13. Команда «Закат» играет с командой «Зенит» до первой победы «Заката». Вероятность такой победы в одном матче постоянна и равна 0.2. Пусть W — число сыгранных матчей. Найдите закон распределения, математическое ожидание, дисперсию W. Определите $r(0.8)$ с помощью неравенства Чебышева.

ПОДСКАЗКА. Геометрический закон.

14. В коробке десять красных шаров и пять синих. Наугад выберем четыре шара из пятнадцати. Пусть T — число синих шаров в выборке.

Найдите закон распределения T, математическое ожидание T и вероятность выбрать красных и синих шаров поровну.

ПОДСКАЗКА. Сочетания. Гипергеометрическое распределение.

15. Среди 15 сорванных яблок ровно три — червивые. Выберем ровно два яблока из 15 случайным образом. Q – число червивых среди них. Найдите закон распределения Q, математическое ожидание, дисперсию Q и вероятность того, что среди выбранных двух яблок червивых нет.

ПОДСКАЗКА. Гипергеометрическое распределение.

3.4.2. Второй уровень сложности

1. В тексте учебника содержатся опечатки: в среднем, одна опечатка на десять страниц. Случайная величина f — число опечаток на одной странице. Найти закон распределения, математическое ожидание, дисперсию f и вероятность того, что на странице есть хотя бы одна опечатка.

ПОДСКАЗКА. Закон Пуассона.

2. Каждый из четырех мальчиков (по очереди) пытается разломить черствый пряник. Если кто-то пряник разломит, то пряник будет съеден. Если никто не разломит — пряник выбрасывают. Каждый мальчик делает по одной попытке разломить пряник. Вероятность успеха для первого мальчика равна 0.5, для второго — 0.6, для третьего — 0.7, для четвертого — 0.8. Пусть W — число неудачных попыток (от 0 до 4). Найдите закон распределения W и вероятность того, что пряник не выбросят.

ПОДСКАЗКА. Умножение вероятностей.

3. В большой мешок сахарного песка заползли муравьи. Если брать песок для чая, то, в среднем, попадаются два муравья на пять порций. Пусть W — число муравьев в одной порции. Найдите закон распределения, числовые характеристики и вероятность того, что взяв десять порций песка, мы не подцепим ни одного муравья.

ПОДСКАЗКА. Пуассон и Бернулли.

4. На окружности отметили четыре точки. Каждые две соединили отрезком. Каждый отрезок покрасили целиком в красный или синий цвет: в красный — с вероятностью 0.8, в синий — 0.2. Случайная величина R — число получившихся красных треугольников с вершинами в отмеченных точках. Найдите закон распределения.

ПОДСКАЗКА. Проанализируйте, сколько может быть красных треугольников.

5. Ровно один из десяти ключей может открыть сундук с сокровищами. Пробуем ключи поочередно. Непригодные откладываем в сторону. T — число попыток открыть сундук. Найдите закон распределения T и вероятность того, что будет сделано не больше трех попыток.

ПОДСКАЗКА. Умножение вероятностей.

6. В сундуке восемь драгоценных камней. Каждый может достаться разбойнику A с вероятностью 0.4 или B (с вероятностью 0.6), но не обоим одновременно. Пусть S — число камней из восьми, которые достанутся A. Найти закон распределения S, числовые характеристики и вероятность того, что разбойники получат поровну камней.

7. Игрок A бросает две монеты. Игрок B бросает три монеты. Если у A выпало больше гербов, чем у B, то B платит ему 300 рублей; если у B выпало больше гербов, чем у A, то A платит B 200 рублей; если гербов поровну, никто не платит. Пусть D — приращение денег A за одну партию игры. Найдите закон D и математическое ожидание.

ПОДСКАЗКА. Схема Бернулли.

8. На вызов хозяйки приходит один из двух дежурных электриков (равновозможно). Если придет электрик A, то он потребует 50 рублей с вероятностью 0.7 или 100 рублей с вероятностью 0.3. Если придет B, то он потребует 50 рублей с вероятностью 0.5, либо 100 рублей с вероятностью 0.4, либо 200 рублей с вероятностью 0.1. Случайная величина G — количество денег, которые придется заплатить электрику. Найдите закон G и математическое ожидание.

ПОДСКАЗКА. Формула полной вероятности.

9. Преподаватель сочиняет вариант из трех задач. Вероятность того, что задача сформулирована корректно, равна 0.8 для любой из трех задач, независимо от других. Некорректную задачу решить нельзя. Если задача корректна, то сам преподаватель может ее правильно решить с вероятностью 0.5, а студент — с вероятностью 0.1. X — число задач из трех, которые может решить преподаватель, Y — число задач из тех же трех, которые может решить студент. $F = X - Y$. Если $F \leqslant 2$, то ставится «зачет». Найдите закон распределения F и вероятность остаться без «зачета».

ПОДСКАЗКА. Произведение вероятностей.

10. Вероятность решить на контрольной первую задачу (из двух предложенных) равна 0.2, списать ее — 0.1. Вероятность решить вторую задачу равна 0.4, списать ее — 0.3. Пусть E — число оформленных на листке

задач (решенных или списанных). Найти закон E и вероятность того, что оформлена хотя бы одна задача.

11. В коробке лежали двадцать красных, десять зеленых и восемь желтых шариков. Случайным образом из смеси выбрали три шарика. Пусть R — число красных среди выбранных. Найдите закон R, математическое ожидание R и вероятность того, что будет выбран хотя бы один красный шарик.

12. У Вовочки пять пулек. Он стреляет по воздушным шарикам на детском празднике. Вероятность попадания постоянна и равна 0.2. Стрельба идет до первого попадания или до окончания боезапаса. V — число сделанных выстрелов. (При выстреле пулька пропадает.) Найдите закон распределения V.

13. На три свободные места в вагоне трамвая претендовали четыре юноши и пять девушек. Их силы были равны, поэтому победители определились случайным образом (все варианты по три из девяти равновозможны). Пусть G — число девушек, получивших место. Найти закон распределения, числовые характеристики G и вероятность того, что все места достались девушкам.

14. Вероятность того, что на остановке не откроется первая дверь трамвая равна 0.2, вторая — 0.3, третья — 0.4. Двери портятся независимо. D — число дверей, которые откроются-таки на остановке. Найти закон распределения D, математическое ожидание D и вероятность того, что откроется хотя бы одна дверь.

ПОДСКАЗКА Произведение вероятностей.

15. Шесть монет разложены по двум карманам так, что для каждой из монет вероятность попасть в первый карман равна 0.7, а во второй — 0.3. Каждая монета попадает ровно в один из двух карманов. Пусть C — число монет, попавших в первый карман. Найдите закон распределения, математическое ожидание, дисперсию C, а также вероятность того, что оба кармана не пусты.

ПОДСКАЗКА Бернулли.

16. Три монеты раскладываются в три копилки так, что каждая попадает в определенную копилку. Для каждой монеты равновозможно попасть в любую из трех копилок. Случайная величина Q — число копилок, в которые не попало ни одной монеты из этих трех. Найдите закон распределения случайной величины Q.

ПОДСКАЗКА Используйте классическую схему подсчета вероятностей.

17. В ящике шесть красных носков и четыре синих носка. Все одного размера. Наугад берем четыре носка из десяти. P — число пар одноцветных носков среди выбранных. Найдите закон распределения P. (Пары не имеют общих элементов.)

ПОДСКАЗКА Используйте классическую схему.

18. В компьютере возникают неполадки из-за комплектующих в среднем два раза за три года. Пусть A — число таких неполадок в год. Подберите закон распределения и подсчитайте вероятность того, что за год не будет ни одной неполадки такого вида.

19. Из тринадцати дворников каждый способен приступить к работе во вторник с вероятностью 0.5, а в понедельник — с вероятностью 0.2. Пусть S — число дней из двух (понедельник, вторник), когда к работе приступали хотя бы по одному дворнику из тринадцати. Найдите закон распределения S.

20. По пятницам дворники пытаются убрать весь мусор в контейнеры. Это можно сделать с вероятностью 0.7. Если в пятницу весь мусор не убран, то до следующей пятницы весь мусор точно не удастся убрать (так, как по пятницам — два контейнера, а с понедельника по четверг — по одному). Пусть W — число недель, которые протекут после удачной пятницы до следующей удачной пятницы. Найдите закон распределения W и вероятность того, что пройдет не более трех недель.

21. В наблюдениях Гейгера радиоактивное вещество за промежуток времени 7.5 секунд испускало в среднем 3.87 α-частицы. Найти вероятность того, что за одну секунду это вещество испустит хотя бы одну α-частицу.

ПОДСКАЗКА. Закон Пуассона.

22. Опытный врач ставит сто диагнозов в месяц, ошибаясь в среднем два раза в месяц. Верно ли, что за 200 месяцев работы месяцев, когда он ошибся ровно два раза, будет больше, чем месяцев, когда он ошибся по одному разу?

ПОДСКАЗКА (известна Симеону Дени Пуассону).

23. У девочки три подруги: A, B, C. Сегодня она может встретить A с вероятностью 0.6, B — с вероятностью 0.7, C — с вероятностью 0.8. Если девочка встретит A, то угостит ее конфетой с вероятностью 0.5. Если она встретит B, то угостит ее с вероятностью 0.9. Если встретит

C, то угостит ее с вероятностью 0.6. Каждой подруге девочка дает не более одной конфеты. Пусть K — число конфет, которые девочка отдаст подругам. Найдите закон распределения.

ПОДСКАЗКА Умножение вероятностей. Сложение вероятностей.

24. В коробке десять ламп фирмы «OSRAM» и пять ламп фирмы «GENERAL ELECTRIC». Вероятность брака для первой фирмы равна 0.002, для второй — 0.01. Выбираем случайно две лампы из пятнадцати. Пусть B — число бракованных среди них. Найдите закон распределения.

ПОДСКАЗКА Формулы полной вероятности.

25. Три мудреца плывут в одном тазу из Шлиссельбурга в Санкт-Петербург. Вероятность того, что таз перевернется равна 0.5 (тогда никто не доплывет в тазу). Кроме того, каждый мудрец может выпасть из таза, не перевернув его, с вероятностью 0.2 (независимо от других). Если таз не перевернулся и мудрец не выпал, то он доберется до цели. M — число доплывших до цели мудрецов. Найти закон распределения.

ПОДСКАЗКА Формула полной вероятности.

26. Мудрец достает из коробки шар, изучает его цвет и кладет шар обратно в коробку. В коробке десять черных шаров и два белых. Опыты продолжаются до первого белого шара. S — число «изученных» шаров. Найти закон распределения и числовые характеристики S.

27. Ромашка имеет равновозможно от 15 до 25 лепестков. (Других вариантов нет.) Студентка гадает на случайно выбранной ромашке: «зачет — незачет». Опыт считается удачным, если она нагадает «зачет». Пусть проведено 4 таких независимых опыта, M — число удач. Найдите закон распределения M, математическое ожидание и вероятность нагадать «зачет» хотя бы один раз из четырех.

28. На каждой из данных двух ромашек водятся равновозможно от одной до трех букашек (других вариантов нет). Случайная величина J — число букашек на двух ромашках. Найти закон распределения.

29. Яблоко может быть червивым с вероятностью 0.7. Берем 5 яблок и начинаем резать по очереди до тех пор, пока не разрежем все или до двух червивых яблок подряд. (Тогда все бросаем.) Случайная величина R — число разрезанных яблок. Найдите закон распределения.

30. В партии равновозможно любое количество бракованных изделий. Выбираем случайным образом четыре изделия. Q — число бракованных изделий среди выбранных. Найдите закон распределения и вероятность

того, что среди выбранных изделий будет меньше половины бракованных.

ПОДСКАЗКА. Используйте следствие из формулы полной вероятности: если среди данных N элементов имеется равновозможно любое количество отмеченных, то и в любом, случайно выбранном сочетании из этих N элементов по k ($k \leqslant N$) будет равновозможно любое количество отмеченных.

31. Рубль меняем на три монеты так, что каждая равновозможно имеет достоинство 10 или 50 копеек (других вариантов нет). Случайная величина R — приращение денег на рубль (-70, -30, $+10$, $+50$). Найдите закон распределения и вероятность положительного приращения.

32. В условие задачи 31 внесем изменение: каждая из трех монет имеет достоинство 10 копеек с вероятностью 0.4 или достоинство 50 копеек с вероятностью 0.6. Остальное аналогично.

3.5. Закон распределения на $\mathbb{R}$. Функция распределения

В этом разделе мы более подробно обсудим принципы построения вероятностной модели на множестве $\mathbb{R}$ всех действительных чисел. Такое построение необходимо при исследовании случайных величин, которые имеют несчетное множество значений. Кроме этого, построение вероятностной модели на $\mathbb{R}$ позволяет «стандартизировать» законы распределения случайных величин всех видов, поскольку вероятностные функции будут заданы на одних и тех же событиях, определенных в множестве $\mathbb{R}$.

Допустим, что случайная величина ξ может принимать какие-то действительные значения (все или нет — не важно). Для построения вероятностной модели на множестве $\mathbb{R}$, необходимо ввести на $\mathbb{R}$ борелевское поле событий F_R. Это делается следующим стандартным способом:

1) все точки $\mathbb{R}$ — элементарные события;

2) само $\mathbb{R}$ — достоверное событие, пустое множество — невозможное событие;

3) все промежутки вида $(-\infty, x)$ — события;

4) если A — событие, то дополнение A до $\mathbb{R}$ — событие;

5) если A и B — события, то $A \cap B$ — событие;

6) объединение конечного или счетного числа событий — событие.

Правила 4)–6) позволяют строить F_R «индуктивно»: из точек $\mathbb{R}$ и интервалов вида $(-\infty, x)$ получаются некие множества «первой ступени», из них – множества «второй ступени» и так далее. В результате получается совокупность подмножеств $\mathbb{R}$, каждое из которых получается из точек и интервалов $(-\infty, x)$ с помощью не более, чем счетной цепочки операций объединения, пересечения или дополнения. По построению, полученная совокупность подмножеств является *борелевским полем событий (σ-алгеброй)* и обозначается символом $\mathfrak{B}$ (B готическое).

Упражнения

1. Докажите, что $\mathfrak{B}$, действительно, является борелевским полем событий.

2. Докажите также, что событиями в нем будут, в частности, множества вида:

$$[x, +\infty), \quad [x, y), \quad (x, y], \quad [x, y], \quad (-\infty, x], \quad (x, +\infty), \quad (x, y).$$

3. Докажите, что пересечение счетного числа событий — событие (в любом борелевском поле).

После введения на $\mathbb{R}$ борелевского поля событий надо задать на этом поле вероятностную функцию P_ξ. Как это сделать? Для того, чтобы понять принцип определения P_ξ, принятого в теории вероятностей, рассмотрим следующую теорему, доказательство которой предоставим читателю в качестве упражнения.

Теорема 3.2. *Пусть на борелевском поле событий $\mathfrak{B}$ задана произвольная вероятностная функция P (выполнены аксиомы вероятности). Для вычисления значений этой функции от любых событий из $\mathfrak{B}$ достаточно знать только ее значения P от событий вида $(-\infty, x)$ и тот факт, что $\mathsf{P}(\mathbb{R}) = 1$. Вероятности же всех остальных событий поля вычисляются на основании свойств вероятностной функции, заложенных в аксиомы Колмогорова. В частности,*

$$\mathsf{P}(\,[a, +\infty)\,) = 1 - \mathsf{P}(\,(-\infty, a)\,), \tag{3.7}$$

$$\mathsf{P}(\,[a, b)\,) = \mathsf{P}(\,(-\infty, b)\,) - \mathsf{P}(\,(-\infty, a)\,), \tag{3.8}$$

$$\mathsf{P}(\{a\}) = \lim_{n\to\infty} \mathsf{P}\Big(\,\Big[a, a + \frac{1}{n}\Big)\,\Big). \tag{3.9}$$

Вернемся теперь к проблеме построения вероятностной модели на множестве $\mathbb{R}$ для случайной величины ξ. Как уже говорилось выше, следует различать два случая. В первом случае на области определения Ω

случайной величины ξ существует вероятностная модель $(\,\Omega, F_\Omega, \mathsf{P}_\Omega\,)$ такая, что любое событие A из $\mathfrak{B}$ имеет прообразом событие B из F_Ω. В этом случае вероятностная функция P_ξ на $\mathfrak{B}$ задается формулой (3.1). Во втором случае приходится прибегать к помощи статистики.

Рассмотрим первый случай более подробно. Поскольку вероятностная функция на $\mathfrak{B}$ определена по формуле (3.1), то в силу теоремы 3.2 достаточно вычислить вероятности событий вида $(\,-\infty,\, x\,)$ для всех действительных x, через которые выражать, по мере необходимости, вероятности других событий.

Определение функции распределения. Если на множестве $\mathbb{R}$ всех действительных чисел построена вероятностная модель $(\,\mathbb{R},\ \mathfrak{B},\ \mathsf{P}_\xi\,)$, являющаяся законом распределения случайной величины ξ, то функция

$$\mathcal{F}_\xi(x) = \mathsf{P}_\xi(\,(-\infty, x)\,) \tag{3.10}$$

называется *функцией распределения случайной величины* ξ.

Функция распределения однозначно определяется законом распределения. Она используется для вычисления вероятностей событий из $\mathfrak{B}$. Так, например, формула (3.8) может быть теперь переписана в виде равенства:

$$\mathsf{P}(\,[a,\, b)\,) = \mathcal{F}(b) - \mathcal{F}(a). \tag{3.11}$$

Формула (3.9) приобретает вид

$$\mathsf{P}(\{a\}) = \lim_{n\to\infty} \mathcal{F}\Big(\,a + \frac{1}{n}\,\Big) - \mathcal{F}(a). \tag{3.12}$$

Задача 3.15. *На отрезке AB единичной длины случайным образом выбирается точка так, что вероятность ее попадания в любую часть отрезка, имеющую длину, вычисляется по геометрической модели (то есть равна этой длине, поскольку длина AB равна единице). (Задача 3.6.) Найти на $\mathbb{R}$ закон распределения ξ — расстояния от выбранной точки до середины отрезка C.*

Как уже говорилось, область определения ξ — отрезок AB. Множество ее значений — числовой отрезок $[0,\, 1/2] = D_\xi$.

Введем на $\mathbb{R}$ борелевское поле событий $\mathfrak{B}$ так, как это делалось выше. Определим вероятностную функцию P_ξ по формуле (3.1), используя геометрическую модель на AB. Значения же вероятностной функции вычислим только для событий вида $(-\infty, x)$. Это равносильно определению

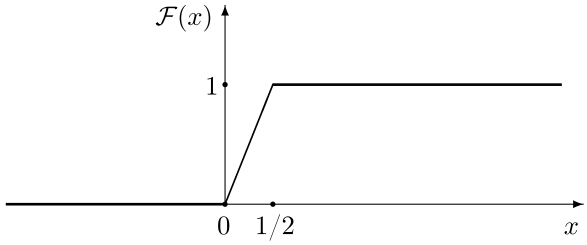


Рис. 3.2

функции $\mathcal{F}$:

$$\mathcal{F}(x) = \begin{cases} 0 & \text{при } x \leqslant 0, \\ \mathsf{P}_\xi(\,[0, x)\,) = 2x & \text{при } 0 \leqslant x \leqslant 1/2, \\ 1 & \text{при } x > 1/2. \end{cases}$$

Последнее равенство вытекает из того, что расстояние до точки C равно числу x из промежутка $[0,\, 1/2)$ ровно для двух точек из AB, симметричных относительно середины отрезка (см. рис. 3.1). График функции $\mathcal{F}(x)$ имеет вид, изображенный на рис. 3.2. Из геометрической модели вытекает, что

$$\mathsf{P}_\xi(\,\{x\}\,) = 0 \text{ для любого } x \text{ из } \mathbb{R}.$$

Тот же результат получается из формулы (3.12). Вероятность любого события $[a,\, b)$ вычисляется по формуле (3.11). Для $b \leqslant 0$ или $a \geqslant 1/2$ вероятность попадания значений случайной величины в промежуток вида $[a,\, b)$, очевидно, равна нулю.

В общем случае, значение вероятностной функции на произвольном событии A из $\mathfrak{B}$ можно определить по следующей схеме (аналогично тому, как это делалось в задаче 3.6):

1) каждому интервалу $(-\infty,\, x)$ сопоставить его прообраз — событие в геометрической модели, уже имеющее там определенную вероятность; эту вероятность приписать рассматриваемому интервалу;

2) представить A как результат конечной или счетной последовательности операций объединения, пересечения или дополнения, производимых над интервалами вида $(-\infty,\, x)$;

3) восстановить аналогичную цепочку операций с прообразами и найти прообраз A;

4) приписать событию A вероятность его прообраза.

Задача 3.16. *Рассмотрим еще раз задачу (3.5). Известно, что некий пакет весит целое число килограмм, равновозможное от одного до семи. Его взвешивают на весах с помощью гирь из набора: 1, 2, 4 (кг). Гири кладут на одну чашку весов, пакет — на другую. Число использованных гирь — случайная величина ξ. Построить для ξ вероятностную модель на $\mathbb{R}$.*

Будем использовать таблицу, построенную при решении задачи 3.5:

Число гирь (ξ)	1	2	3
Вероятность	3/7	3/7	1/7

Эта таблица задает вероятностную функцию на множестве значений D_ξ рассматриваемой случайной величины (то есть закон распределения на множестве значений). Расширим область определения вероятностной функции. Для этого введем на $\mathbb{R}$ борелевское поле событий $\mathfrak{B}$. Событиям $\{1\}$, $\{2\}$, $\{3\}$ припишем вероятности из таблицы. Назовем эти элементарные события «возможными». Другие элементарные события произойти не могут. Им припишем вероятность, равную нулю. Произвольно выбранному событию из $\mathfrak{B}$ припишем вероятность, равную сумме вероятностей благоприятствующих ему «возможных» элементарных событий из D_ξ. В частности, для интервалов вида $(-\infty, x)$ получатся вероятности:

$$\mathcal{F}(x) = \begin{cases} 0 & \text{при } x \leqslant 1, \\ 3/7 & \text{при } 1 < x \leqslant 2, \\ 6/7 & \text{при } 2 < x \leqslant 3, \\ 1 & \text{при } x > 3. \end{cases}$$

График этой функции распределения изображен на рис. 3.3. Остается проверить, что «расширенная» указанным образом вероятностная функция удовлетворяет аксиомам Колмогорова. Первые два свойства очевидны. Третье имеет место, поскольку и сумма вероятностей несовместных событий A_i, и вероятность их объединения вычисляются с помощью сложения вероятностей одних и тех же «возможных» элементарных исходов. В нашем случае число слагаемых, отличных от нуля, всегда не больше трех.

Аналогичным образом можно построить на $\mathbb{R}$ вероятностную модель для любой дискретной случайной величины, если известна модель на

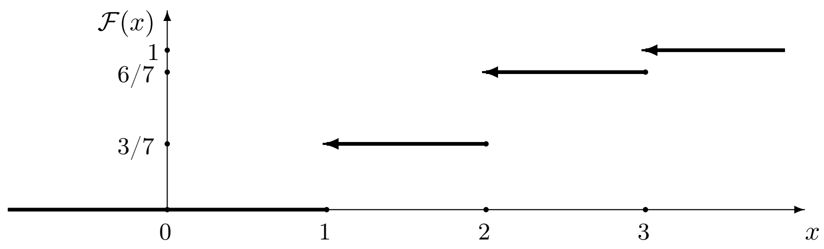


Рис. 3.3

множестве ее значений (то есть ряд распределения). Сформулируем это в виде следующей простой теоремы, доказательство которой предоставляем читателю в качестве упражнения.

Теорема 3.3. *Пусть случайная величина ξ имеет не более, чем счетное множество значений D_ξ, и пусть известен ряд распределения этой величины.*

Тогда на множестве $\mathbb{R}$ всех действительных чисел можно построить вероятностную модель для ξ, доопределив вероятностную функцию P_ξ на борелевском поле $\mathfrak{B}$ по формуле

$$\mathsf{P}_\xi(A) = \sum_{x_i \in A} \mathsf{P}_\xi(x_i), \tag{3.13}$$

где через x_i обозначены «возможные» элементарные события, благоприятствующие событию A, принадлежащему $\mathfrak{B}$, а через $\mathsf{P}_\xi(x_i)$ обозначены вероятности, соответствующие x_i в ряду распределения.

Задача 3.17. *Гражданка X обещала позвонить гражданину Y по телефону в течение дня, с 8 до 24 часов. Если она не забудет выполнить свое обещание, то равновозможен ее звонок в любой момент из указанного промежутка времени. Но, к сожалению, с вероятностью 0.2 она может вообще забыть о своем обещании. Случайная величина ξ — время ожидания телефонного звонка гражданином Y. Найти на $\mathbb{R}$ закон распределения ξ.*

Рассмотрим две гипотезы: «X не забыла позвонить», «X забыла позвонить». Обозначим их, соответственно, через H_1 и H_2. Вероятности гипотез известны из условия задачи: 0.8 и 0.2.

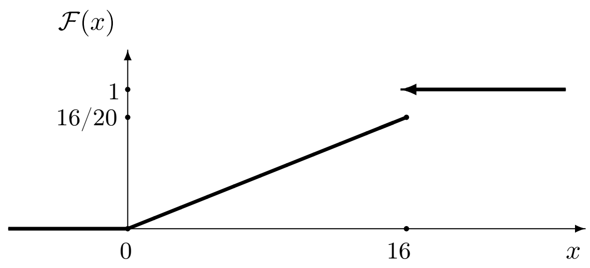


Рис. 3.4

Рассмотрим теперь событие A: время ожидания телефонного звонка принадлежит интервалу длины a, содержащемуся в отрезке $[0,\ 16]$. Если гражданка X не забудет позвонить Y, то вероятность события A при условии H_1 можно вычислить по геометрической схеме через отношение длин: $a/16$. Это число надо умножить на вероятность гипотезы H_1. Получится вероятность произведения AH_1. Вероятность события A при условии H_2, очевидно, равна нулю (бедный Y должен ждать все 16 часов). Следовательно, по формуле полной вероятности событие A должно иметь вероятность $a/20$.

Все события типа «время ожидания равно числу T из промежутка $[0,\ 16)$» имеют нулевую вероятность (это следует из геометрической схемы или из гипотезы H_2). Только одно элементарное событие $T = 16$ имеет вероятность 0.2 по условию задачи.

На основе сказанного можно определить вероятности событий вида $(-\infty, x)$, принадлежащих $\mathfrak{B}$:

$$\mathcal{F}(x) = \begin{cases} 0 & \text{при } x \leqslant 0, \\ x/20 & \text{при } 0 < x \leqslant 16, \\ 1 & \text{при } x > 16. \end{cases}$$

На рис. 3.4 приведен график функции $\mathcal{F}(x)$. Для остальных событий из $\mathfrak{B}$ вероятность определяется аналогично, по формуле полной вероятности. Она может быть вычислена через вероятности событий вида $(-\infty,\ x)$ с помощью конечного или счетного числа операций сложения или вычитания.

Рассмотренные примеры иллюстрируют стандартный подход к построению вероятностной модели на множестве $\mathbb{R}$ (то есть к выяснению закона распределения случайной величины на $\mathbb{R}$). Сперва определяются вероятности событий типа $(-\infty,\, x)$, после чего вычисляются вероятности остальных событий «по аддитивности». Применение аддитивности является обоснованным, если вероятностная функция определена соотношением (3.1).

В случае, когда не удается найти модель на Ω и приходится прибегать к помощи статистики, можно попытаться подобрать формулу для функции распределения, приближенно вычисляя вероятности событий вида $(-\infty,\, x)$, а затем восстановить по этой функции всю модель на $\mathbb{R}$. Успех реализации такого подхода, оказывается, полностью зависит от свойств полученной «функции распределения». (Кавычки подчеркивают, что распределения-то пока нет!) Для решения задачи восстановления модели на $\mathbb{R}$ отметим необходимые свойства, которыми должна обладать функция распределения любой случайной величины.

Теорема 3.4. *Если функция $\mathcal{F}(x)$ определяется равенством (3.10) как функция распределения случайной величины ξ, то она должна обладать следующими свойствами:*

1) $0 \leqslant \mathcal{F}(x) \leqslant 1$*;*

2) $\mathcal{F}(x)$ *не убывает;*

3) $\mathcal{F}(x)$ *непрерывна слева;*

4) $\lim\limits_{x\to+\infty} \mathcal{F}(x) = 1$*;*

5) $\lim\limits_{x\to-\infty} \mathcal{F}(x) = 0$.

Доказательство. Первое свойство функции распределения вытекает из того, что ее значения являются вероятностями, а вероятности любых событий содержатся в промежутке $[0,\, 1]$. Второе свойство следует из равенства (3.11) для любых действительных чисел a и b таких, что $a < b$.

Для доказательства третьего свойства зафиксируем произвольное число a и рассмотрим любую последовательность чисел x_n, сходящуюся слева к a. Покажем, что

$$\lim_{x_n\to a-0} \mathcal{F}(x_n) = \mathcal{F}(a). \tag{3.14}$$

Представим событие $(-\infty,\, a)$ в виде следующего объединения:

$$(-\infty,\, a) = (-\infty,\, a-1) \cup \Big[a-1,\, a-\frac{1}{2}\Big) \cup \dots \cup \Big[a-\frac{1}{k},\, a-\frac{1}{k+1}\Big) \dots$$

Обозначим компоненты этого объединения через A_1, A_2, $\ldots$, соответственно. Вычислим вероятностную функцию от обеих частей равенства:

$$\mathcal{F}(a) = \mathsf{P}_\xi(A_1) + \sum_{n=2}^{\infty} \mathsf{P}_\xi(A_k).$$

Правая часть этого равенства является сходящимся к $\mathcal{F}(a)$ числовым рядом, поскольку вероятностная функция обладает свойством счетной аддитивности. Частичные суммы S_k $(k = 1, 2, \ldots)$ этого ряда равны, соответственно, $\mathcal{F}(a - 1/k)$, откуда следует, что

$$\lim_{k\to\infty} \mathcal{F}\Big(a - \frac{1}{k}\Big) = \mathcal{F}(a).$$

Рассмотрим теперь произвольное положительное число ε. Найдем соответствующий ему номер K такой, что $\mathcal{F}(a) - \mathcal{F}(a - 1/K) < \varepsilon$. По числу $a - 1/K$ подберем номер N такой, что при всех $n > N$ выполнено неравенство: $x_n > a - 1/K$. Тогда в силу монотонности функции $\mathcal{F}(x)$ будет выполнено неравенство: $\mathcal{F}(a) - \mathcal{F}(x_n) < \varepsilon$ для всех этих n. Соотношение (3.14) доказано.

Для доказательства четвертого свойства представим множество $\mathbb{R}$ в виде счетного объединения:

$$\mathbb{R} = (-\infty,\, 1) \cup [1,\, 2) \cup \cdots \cup [n,\, n+1) \ldots$$

Вычислим вероятностную функцию от этого объединения. С одной стороны, ее значение равно единице, с другой — пределу частичных сумм ряда, которые равны $\mathcal{F}(n)$. Рассмотрим теперь произвольную последовательность чисел x_n, стремящуюся к бесконечности, и положительное число ε. При достаточно больших значениях n члены последовательности становятся больше некоторого числа K такого, что $1 - \mathcal{F}(K) < \varepsilon$. Следовательно, $1 - \mathcal{F}(x_n) < \varepsilon$ для членов последовательности с такими номерами. Что и требовалось доказать.

Пятое свойство докажите в качестве упражнения, воспользовавшись равенством (3.7).

Следствие. Скачок функции $\mathcal{F}(x)$ в произвольной точке a равен вероятности события $\xi = a$, то есть

$$\lim_{x\to a+0} \mathcal{F}(x) - \lim_{x\to a-0} \mathcal{F}(x) = \mathsf{P}_\xi(\{a\}). \tag{3.15}$$

Если функция распределения непрерывна, то вероятность любого события вида $\xi = a$ равна нулю.

Доказательство. Левосторонний предел в точке a существует и равен $\mathcal{F}(a)$, как показано при доказательстве теоремы 3.4. Сравним равенства (3.12) и (3.15). Для того, чтобы доказать справедливость последнего, достаточно установить, что правосторонний предел функции $\mathcal{F}(x)$ в точке a существует для любой последовательности чисел x_n, сходящейся к точке a справа. Это делается точно так, как при доказательстве свойств 3) и 4) функции $\mathcal{F}(x)$ в теореме 3.5.

Равенство нулю вероятности события $\xi = a$ при непрерывной функции распределения вытекает непосредственно из формулы (3.15).

Задача 3.18. *Могут ли следующие функции являться функциями распределения для каких-нибудь случайных величин?*

$$\text{а)} \qquad \mathcal{F}(x) = \frac{1}{x^2+1}\,;$$

$$\text{б)} \qquad \mathcal{F}(x) = \begin{cases} 0 & \text{при } x < 0, \\ 1 & \text{при } x \geqslant 1; \end{cases}$$

$$\text{в)} \qquad \mathcal{F}(x) = \begin{cases} 0 & \text{при } x \leqslant 0, \\ x/2 & \text{при } 0 < x \leqslant 2, \\ 1 & \text{при } x > 2. \end{cases}$$

В случае а) ответ отрицательный, поскольку данная функция убывает при положительных значениях x. Не выполнены второе и четвертое необходимые условия, указанные в теореме 3.4. В случае б) не выполнено третье условие. В случае в) функция обладает необходимыми свойствами 1) – 5), перечисленными в теореме 3.4, но это еще не дает нам права утверждать, что перед нами — функция распределения какой-то случайной величины. Исходя из опыта решения задач, мы можем описать случайную величину, имеющую такую функцию распределения. Пусть результатом эксперимента является число ξ из промежутка $[0,\,2]$. Пусть на $\mathbb{R}$ построена геометрическая модель, в которой вероятность попадания значения ξ в точку равна нулю, а вероятность попадания значения ξ в промежуток $[a,\,b]$, содержащийся в $[0,\,2]$, определяется через отношение длин: $(b-a)/2$. Тогда функция распределения имеет вид в).

Итак, остается невыясненным вопрос: является ли функция $\mathcal{F}(x)$ функцией распределения какой-нибудь случайной величины, если она обладает всеми необходимыми свойствами, перечисленными в теореме 3.4? Ответом на этот вопрос является следующая теорема.

Теорема 3.5. *Пусть дана функция $\mathcal{F}(x)$, определенная на $\mathbb{R}$ и обладающая свойствами 1) – 5), перечисленными в теореме 3.4. Тогда существует неотрицательная счетно-аддитивная функция* $\mathsf{P}(A)$, *определенная на* $\mathfrak{B}$, *такая, что* $\mathsf{P}(\mathbb{R}) = 1$ *и* $\mathsf{P}(\,(-\infty, x)\,) = \mathcal{F}(x)$.

Доказательство этой теоремы основано на некоторых фактах из теории меры Лебега. Здесь мы не приводим это доказательство, поскольку мера Лебега обычно не изучается студентами нематематических специальностей.

Определив по функции $\mathcal{F}(x)$ вероятностную функцию $\mathsf{P}(A)$ на $\mathfrak{B}$, мы можем указать случайную величину ξ, имеющую такое распределение на $\mathbb{R}$: пусть ξ отображает $\mathbb{R}$ взаимно однозначно на $\mathbb{R}$, то есть результатом некоторого абстрактного эксперимента является действительное число, а вероятности событий из $\mathfrak{B}$ вычисляются с помощью $\mathsf{P}(A)$. Случайная величина с заданной функцией распределения не является, строго говоря, единственной. (Подумайте, почему.)

В заключение приведем классификацию случайных величин, основанную на свойствах соответствующих им функций распределения на $\mathbb{R}$.

Определение. Случайная величина ξ называется *непрерывной*, если для нее существует распределение на $\mathbb{R}$, которому соответствует непрерывная функция распределения $\mathcal{F}(x)$. Распределение также называется непрерывным.

Случайная величина ξ называется *дискретной*, если для нее существует распределение на $\mathbb{R}$ с кусочно-постоянной функцией $\mathcal{F}$. Распределение в этом случае называется дискретным.

Все остальные случайные величины, имеющие вероятностные модели на $\mathbb{R}$, относятся к *смешанному типу*. Распределения называются распределениями смешанного типа.

Проанализировав рис. 3.2–3.4, приходим к выводу, что случайная величина, рассмотренная в первом примере, является непрерывной, во втором — дискретной, в третьем — смешанного типа.

Замечание. Ранее уже давалось определение дискретной случайной величины как величины, принимающей с ненулевой вероятностью не более, чем счетное число значений. Функция распределения такой случайной величины, очевидно, является кусочно-постоянной. В качестве упражнения докажите, что если случайная величина ξ имеет кусочно-постоянную функцию распределения $\mathcal{F}(x)$, то ξ является дискретной в смысле «старого» определения. Воспользуйтесь известным в теории функций действительной переменной фактом: монотонная на $\mathbb{R}$ функция может иметь не более, чем счетное число разрывов первого рода.

3.6. Плотность распределения вероятности. Абсолютно непрерывные случайные величины

В этом разделе мы рассмотрим так называемые *абсолютно непрерывные* функции распределения, широко применяемые на практике. Для этих функций можно легко определять, в какой из промежутков равной длины значения соответствующей случайной величины попадают с большей вероятностью, в какой — с меньшей.

Определение абсолютно непрерывной функции. Если функция $f(x)$ может быть представлена в виде

$$f(x) = f(c) + \int\limits_c^x p(x)\, dx \tag{3.16}$$

при всех x, принадлежащих $[c,\, d]$, то она называется *абсолютно непрерывной* на $[c,\, d]$.

Как известно из курса математического анализа, интеграл в формуле (3.16) обязательно существует, если функция $p(x)$ непрерывна на указанном отрезке или ограничена на нем и имеет конечное число точек разрыва. (В более общих случаях интеграл определяется по Лебегу.) Все абсолютно непрерывные функции непрерывны. Непрерывные же функции не обязательно обладают свойством абсолютной непрерывности. Можно привести пример непрерывной функции, не обладающей свойством абсолютной непрерывности. Важнейшее свойство абсолютно непрерывных на $[c,\, d]$ функций следующее:

$$\frac{df(x)}{dx} = p(x)\ , \tag{3.17}$$

во всех точках непрерывности $p(x)$.

Определение плотности распределения вероятности. Если функция распределения $\mathcal{F}(x)$ случайной величины ξ может быть представлена в виде несобственного интеграла

$$\mathcal{F}(x) = \int\limits_{-\infty}^x p(x)\, dx, \tag{3.18}$$

то подынтегральная функция $p(x)$ называется *функцией плотности распределения вероятности* случайной величины ξ.

(В технических приложениях иногда используется более общее определение плотности, которое распространяется и на случайные величины, имеющие разрывные функции распределения. При этом плотность понимается как обобщенная функция, значения которой в точках разрыва $\mathcal{F}(x)$ определяются с помощью дельта-функции Дирака.)

Пусть функция распределения удовлетворяет условию (3.18). Тогда для любых чисел c и x выполнено равенство:

$$\mathcal{F}(x) - \mathcal{F}(c) = \int\limits_c^x p(x)\,dx,$$

следовательно, $\mathcal{F}(x)$ абсолютно непрерывна на $\mathbb{R}$. Поэтому *функции распределения, удовлетворяющие равенству* (3.18) *и сами случайные величины, имеющие такие функции распределения, в теории вероятностей называются абсолютно непрерывными.*

Простейшее достаточное условие существования функции плотности распределения сформулируем в виде теоремы.

Теорема 3.6. *Если функция распределения $\mathcal{F}(x)$ непрерывна на $\mathbb{R}$ и имеет непрерывную производную всюду, за исключением разве лишь конечного числа точек, то справедливо соотношение (3.18), в котором функция $p(x)$ совпадает с $\mathcal{F}'(x)$ во всех точках ее непрерывности. В точках разрыва производной $\mathcal{F}(x)$ функцию $p(x)$ можно доопределить произвольно.*

Доказательство. В рассматриваемом случае выполнены условия существования интеграла от $\mathcal{F}'(x)$ на любом промежутке вида (c, x). Этот интеграл равен $\mathcal{F}(x) - \mathcal{F}(c)$. Переходя к пределу при c стремящемся к $-\infty$, получаем соотношение (3.18), в котором под интегралом стоит $\mathcal{F}'(x)$. Это соотношение остается справедливым и для любой другой подынтегральной функции $p(x)$, совпадающей с $\mathcal{F}'(x)$ во всех точках ее непрерывности.

Рассмотрим функции распределения, приведенные в задачах 3.15–3.17. Выясним, какие из них могут быть представлены в виде интеграла (3.18). Если это возможно, найдем аналитический вид соответствующих функций плотности распределения и построим их графики.

В задаче 3.15 рассматривалась следующая функция распределения:

$$\mathcal{F}(x) = \begin{cases} 0 & \text{при } \ x \leqslant 0, \\ \mathsf{P}_\xi(\,[0,x)\,) = 2x & \text{при } \ 0 \leqslant x \leqslant 1/2, \\ 1 & \text{при } \ x > 1/2. \end{cases}$$

Эта функция является непрерывной. Производная от нее существует везде, кроме точек $x = 0$ и $x = 1/2$. В точках существования производная непрерывна. Ее аналитическое задание имеет вид:

$$\mathcal{F}\,'(x) = \begin{cases} 0 & \text{при } \ x < 0, \\ 2 & \text{при } \ 0 < x < 1/2, \\ 0 & \text{при } \ x > 1/2. \end{cases}$$

По теореме 3.6 получаем, что $p(x) = \mathcal{F}\,'(x)$ при всех x, кроме $x = 0$, $x = 1/2$. В этих точках можно, по желанию, доопределить функцию $p(x)$ какими-нибудь числами. В теории вероятностей принято доопределять такую функцию плотности следующим образом:

$$p(x) = \begin{cases} 0 & \text{при } \ x < 0, \\ 2 & \text{при } \ 0 \leqslant x \leqslant 1/2, \\ 0 & \text{при } \ x > 1/2. \end{cases}$$

Такое определение значений $p(x)$ в точках разрыва связано с тем, что события вида $\xi = 0$ и $\xi = 1/2$ не являются, в принципе, невозможными, хотя вероятность, соответствующая им в математической модели, равна нулю. График функции $p(x)$ изображен на рис. 3.5.

Рассмотрим теперь функцию распределения из задачи 3.16:

$$\mathcal{F}(x) = \begin{cases} 0 & \text{при } x \leqslant 1, \\ 3/7 & \text{при } 1 < x \leqslant 2, \\ 6/7 & \text{при } 2 < x \leqslant 3, \\ 1 & \text{при } x > 3. \end{cases}$$

Эта функция не является непрерывной, тем более, абсолютно непрерывной. Следовательно, она не может быть представлена с помощью интеграла (3.18). Функция плотности распределения в данном случае не

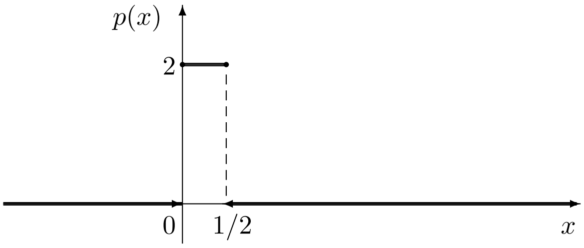


Рис. 3.5

существует. Тем не менее функция распределения имеет производную во всех точках, кроме точек разрыва. Это показывает, что равенство (3.18) не равносильно дифференцируемости функции $\mathcal{F}(x)$ при почти всех x. Аналогичный результат получаем, рассматривая $\mathcal{F}(x)$ из задачи 3.17.

Итак, здесь изо всех непрерывных случайных величин мы выделяем те случайные величины, которые имеют функцию плотности распределения вероятности. Особое внимание к таким величинам обусловлено тем, что при обработке конкретного статистического материала, зачастую сперва удается подобрать стандартную функцию плотности распределения случайной величины, связанной с проводимым экспериментом, а уже по этой плотности найти функцию распределения.

Если случайная величина ξ, имеет плотность распределения $p(x)$, то вероятность попадания ее значений в промежутки $[a,\, b]$, $(a,\, b]$, $[a,\, b)$, $(a,\, b)$, очевидно, равна

$$\mathsf{P}_\xi(\,(a,\, b)\,) = \int\limits_a^b p(x)\,dx. \tag{3.19}$$

Геометрически это означает, что вероятность попадания значений ξ в промежуток $(a,\, b)$ равна площади криволинейной трапеции под графиком $p(x)$, соответствующей этому промежутку. Следовательно, вероятность попадания в $(a,\, b)$ тем больше, чем больше площадь под графиком, соответствующая этому отрезку.

Выделим два важных свойства всех функций плотности.

Теорема 3.7. *Пусть функция $p(x)$ является функцией плотности распределения какой-то случайной величины ξ. Тогда $p(x) \geqslant 0$ во всех* *точках непрерывности $p(x)$, и выполнено равенство:*

$$\int\limits_{-\infty}^{+\infty} p(x)\,dx = 1. \tag{3.20}$$

Доказательство. Пусть точка x_0 является точкой непрерывности функции $p(x)$, но $p(x_0) < 0$. Можно показать, что в точке x_0 существует производная функции $\mathcal{F}(x)$, которая равна $p(x_0)$. Такого быть не может, поскольку непрерывная функция $\mathcal{F}(x)$ нигде не убывает.

Для доказательства свойства (3.20) достаточно воспользоваться свойствами функции распределения $\mathcal{F}(x)$.

В качестве простого упражнения докажите теперь следующую теорему.

Теорема 3.8. *Пусть функция $p(x)$ неотрицательна, интегрируема на любом промежутке действительной оси (в смысле Лебега или, в частности, Римана), причем выполнено равенство (3.20). Тогда эта функция является функцией плотности для некоторой случайной величины.*

Задача 3.19. *При каких значениях параметра A функция*

$$p(x) = \frac{A}{1+x^2}$$

является функцией плотности распределения некоторой абстрактной случайной величины?

Из теоремы 3.7 следует, что значение A должно быть неотрицательно. Кроме того, формула (3.20) дает возможность вычислить единственное возможное в этом случае значение параметра A. Действительно,

$$1 = \int\limits_{-\infty}^{+\infty} \frac{A}{1+x^2}\,dx = A\pi,$$

следовательно, $A = 1/\pi$. Применяя формулу (3.18), получаем, что при таком значении A функция распределения имеет вид:

$$\mathcal{F}(x) = \frac{1}{\pi}\left(\operatorname{arctg} x + \frac{\pi}{2}\right).$$

Эта функция удовлетворяет условиям теоремы 3.5 предыдущего раздела, а значит является функцией распределения некоторой случайной величины.

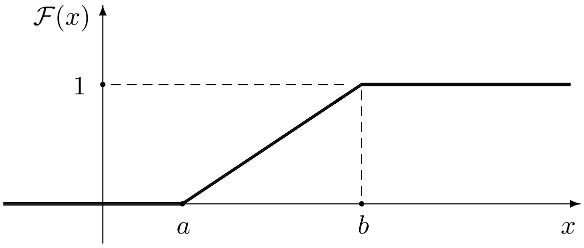


Рис. 3.6. Функция равномерного распределения

3.7. Примеры абсолютно непрерывных распределений

В этом разделе мы рассмотрим несколько конкретных видов распределений, имеющих функцию плотности. Такие распределения характерны для многих случайных величин, часто встречающихся на практике. Каким образом удалось доказать, что в рассматриваемых ниже примерах законы распределения именно таковы? Для этого привлекались специальные методы математической статистики, основанные на так называемых «предельных теоремах» теории вероятностей. Здесь мы примем без доказательств, что случайные величины, описанные в примерах имеют конкретные распределения.

Равномерное распределение на отрезке. Пусть случайная величина ξ может принимать значения только из данного отрезка $[a, b]$, причем вероятность попадания значений ξ в любую часть этого отрезка, имеющую длину α, равна отношению $\alpha/(b-a)$. Легко доказать, что функция распределения такой величины задается следующим образом:

$$\mathcal{F}(x) = \begin{cases} 0 & \text{при } x \leqslant a, \\ \mathsf{P}_\xi(\,[a,x)\,) = \dfrac{x-a}{b-a} & \text{при } a \leqslant x \leqslant b, \\ 1 & \text{при } x > b. \end{cases} \tag{3.21}$$

График этой функции изображен на рис. 3.6. Такая функция распределения уже встречалась нам в задаче 3.15. В предыдущем параграфе мы показали, как найти функцию плотности с помощью производной от функции $\mathcal{F}(x)$. В рассматриваемом случае функция плотности имеет

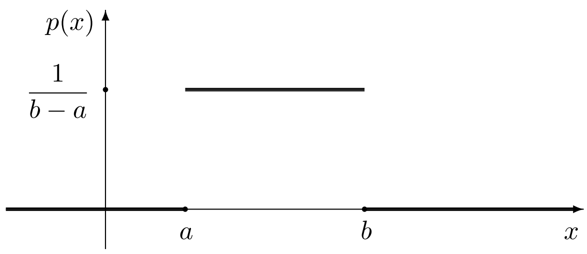


Рис. 3.7. Плотность равномерного распределения

вид:

$$p(x) = \begin{cases} 0 & \text{при} \ \ x < a, \\ \dfrac{1}{b-a} & \text{при} \ \ a \leqslant x \leqslant b, \\ 0 & \text{при} \ \ x > b. \end{cases} \qquad (3.22)$$

Ее график изображен на рис. 3.7. Напомним, что значения функции $p(x)$ в точках a и b не являются значениями производной от $\mathcal{F}(x)$.

Если распределение случайной величины имеет функцию плотности вида (3.22), то оно называется *равномерным на отрезке* $[a,\, b]$. Числа a и b называются *параметрами* равномерного распределения. Они однозначно определяют равномерное распределение.

Задача 3.20. *Известно, что автобус номер 13 приходит на остановку с интервалом ровно в 13 минут. Гражданин X подошел к остановке в некоторый случайный момент времени. Предполагая, что время ожидания им автобуса имеет равномерное распределение на отрезке* $[0,\, 13]$*, вычислить вероятность того, что гражданин X будет ждать автобус не более двух минут.*

Для вычисления вероятности можно воспользоваться формулой (3.21) при $a = 0$, $b = 13$. Полученная вероятность равна $2/13$. Тот же результат мы получим из формулы (3.19). Для вычисления интеграла можно воспользоваться его геометрическим смыслом: вероятность равна площади прямоугольника с высотой $1/13$ и длиной основания 2, находящегося под графиком $p(x)$.

Задача 3.21. *В результате многолетних наблюдений установлено, что вероятность прождать автобус номер 102 на остановке от*

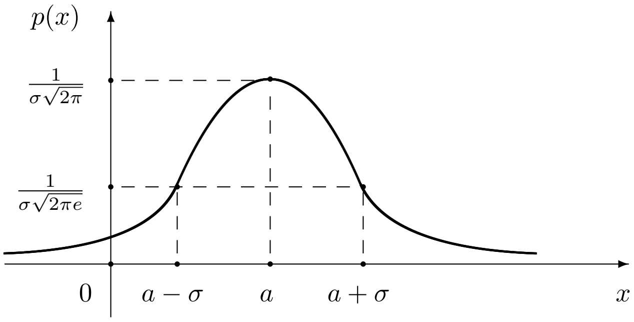


Рис. 3.8. Кривая Гаусса

пяти до 25 минут равна 0.2. *Предполагая, что время ожидания распределено равномерно, найти максимально возможное время ожидания* T.

Найдем минимальное из всех возможных значений величины $1/T$. Для этого вычислим высоту прямоугольника, имеющего площадь 0.2 и длину основания $25 - 5 = 20$. Высота равна 0.01. Вся площадь под графиком $p(x)$ равна 1. Следовательно, максимальная длина отрезка $[0,\,T]$ равна 100. Максимально возможное время ожидания равно 100 минутам.

Нормальное распределение на прямой. Если функция плотности распределения случайной величины имеет вид

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-0.5\,((x-a)/\sigma)^2}, \tag{3.23}$$

где числа a и σ — параметры, $\sigma > 0$, то соответствующее распределение называется *нормальным распределением на прямой*.

Рассматриваемая функция $p(x)$ удовлетворяет условиям теоремы 3.8. Действительно, она неотрицательна и непрерывна. Справедливость условия (3.20) вытекает из известной формулы Пуассона

$$\int\limits_{-\infty}^{\infty} e^{-0.5\,t^2}\,dt = \sqrt{2\pi},$$

если в итеграле сделать замену переменной $t = (x-a)/\sigma$. График функции $p(x)$ имеет вид колокола и называется *кривой Гаусса* (см. рис. 3.8).

Функция вида (3.23) отражает предельные свойства биномиального распределения.

Задача 3.22. *Монета подбрасывается сто раз. Случайная величина ξ — число выпавших гербов. Считая распределение ξ биномиальным с параметрами $n = 100$, $p = 0.5$, вычислить вероятности событий $\xi = m$ для m от 1 до 100. Сравнить полученные числа со значениями функции $p(x)$ вида (3.23) при $a = np = 50$, $\sigma = \sqrt{npq} = 5$, где $q = 1 - p$. Построить график функции распределения ξ и сравнить его с функцией $\mathcal{F}(x)$, определенной по формуле (3.18).*

Для расчета вероятности события $\xi = m$, которое ранее обозначалось символом $E_{n,m}$, воспользуемся формулой Бернулли (2.14):

$$\mathsf{P}(E_{n,m}) = C_n^m\, p^m\, q^{n-m}.$$

По этой формуле при $n = 100$, $m = 0$ имеем $\mathsf{P}(E_{100,0}) = q^n = 0.5^{100}$. Для остальных m расчеты вероятности событий $E_{n,m}$ проделаем, используя отношение:

$$\frac{\mathsf{P}(E_{n,m+1})}{\mathsf{P}(E_{n,m})} = \frac{C_n^{m+1}p^{m+1}q^{n-m+1}}{C_n^m p^m q^{n-m}} = \frac{(n-m)p}{(m+1)q} = \frac{100-m}{m+1}$$

(см. табл. 3.1). Вероятности событий $E_{100,m}$ и $E_{100,101-m}$ совпадают, поскольку $p = q = 0.5$. Поэтому имеет смысл анализировать только вероятности, соответствующие значениям m от 0 до 50. При $m < 36$ все вероятности меньше, чем 10^{-3}. Вероятности, соответствующие значениям m от 36 до 50, округленные до трех верных цифр, помещены во второй столбец таблицы.

В третий столбец помещены значения функции $p(x)$ при тех же m. Значения во втором и третьем столбцах таблицы практически совпадают. Это иллюстрирует следующую предельную теорему.

Теорема 3.9 (локальная теорема Муавра—Лапласа). *Пусть вероятность p постоянна, $q = 1 - p$, $0 \leqslant m \leqslant n$, $x = (m - np)/\sqrt{npq}$. Тогда*

$$\lim_{n\to\infty} \sqrt{npq}\,\mathsf{P}(E_{n,m}) \Big/ \frac{1}{\sqrt{2\pi}}\, e^{-0.5\, x_n^2} = 1$$

равномерно для всех m, для которых x находится в каком-либо конечном интервале.

m	$P(E_{100,m})$	$p(m)$
36	$1.56 \cdot 10^{-3}$	$1.58 \cdot 10^{-3}$
37	$2.70 \cdot 10^{-3}$	$2.72 \cdot 10^{-3}$
38	$4.47 \cdot 10^{-3}$	$4.48 \cdot 10^{-3}$
39	$7.11 \cdot 10^{-3}$	$7.09 \cdot 10^{-3}$
40	$1.08 \cdot 10^{-2}$	$1.08 \cdot 10^{-2}$
41	$1.58 \cdot 10^{-2}$	$1.58 \cdot 10^{-2}$
42	$2.22 \cdot 10^{-2}$	$2.22 \cdot 10^{-2}$
43	$3.00 \cdot 10^{-2}$	$2.99 \cdot 10^{-2}$
44	$3.90 \cdot 10^{-2}$	$3.88 \cdot 10^{-2}$
45	$4.84 \cdot 10^{-2}$	$4.84 \cdot 10^{-2}$
46	$5.80 \cdot 10^{-2}$	$5.79 \cdot 10^{-2}$
47	$6.66 \cdot 10^{-2}$	$6.66 \cdot 10^{-2}$
48	$7.35 \cdot 10^{-2}$	$7.37 \cdot 10^{-2}$
49	$7.80 \cdot 10^{-2}$	$7.82 \cdot 10^{-2}$
50	$7.96 \cdot 10^{-2}$	$7.98 \cdot 10^{-2}$

Таблица 3.1

Доказательство теоремы опускаем. Из приведенной теоремы при больших n получаем приближенную формулу:

$$\mathsf{P}(E_{n,m}) \approx \frac{1}{\sqrt{2\pi npq}}\, e^{-0.5\, x^2}; \quad x = \frac{m - np}{\sqrt{npq}}.$$

Эмпирически установлено, что эта приближенная формула дает хороший результат, когда $npq \geqslant 10$.

Построим теперь график функции биномиального распределения, которую обозначим $\mathcal{F}_\xi(x)$. Рассматриваемая функция является кусочно-постоянной (ступенчатой). Длина ступенек равна единице, высота скачков равна вероятностям событий $E_{100,m}$ (см. рис. 3.9). В тех же осях построим график функции нормального распределения с параметрами $a = 50$, $\sigma = 5$ (рис. 3.10, плавная линия). Обозначим эту функцию $\mathcal{F}(x)$. Графики функций дискретного и нормального распределения близки. Полученный результат иллюстрирует вторую предельную теорему.

Теорема 3.10 (интегральная теорема Муавра—Лапласа). *Если вероятность p постоянна, то для любого числа x справедливо со-*

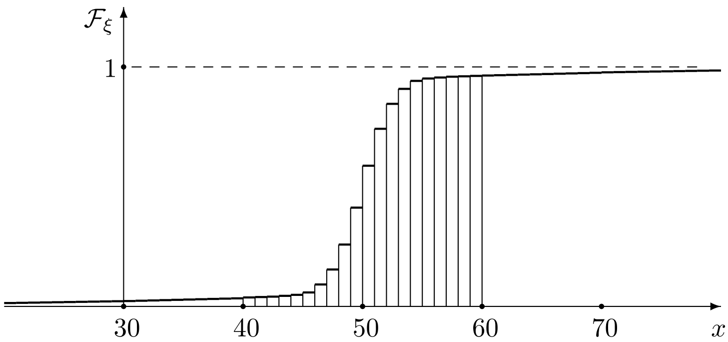


Рис. 3.9. Функция биномиального распределения

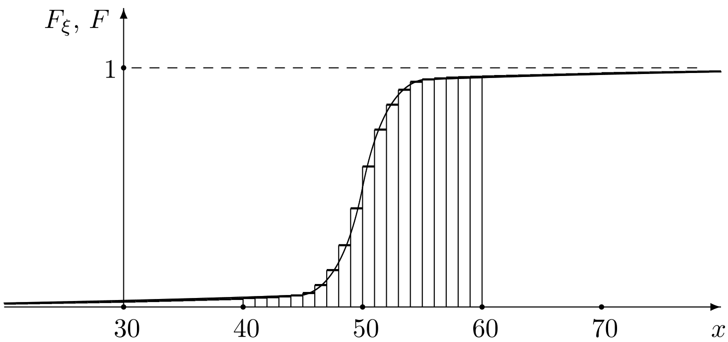


Рис. 3.10. Функция нормального распределения

отношение

$$\lim_{n\to\infty} \mathsf{P}\left(\sum_{m\leqslant np+x\sqrt{npq}} E_{n,m}\right) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-0.5\,t^2}\,dt.$$

Доказательство также опускаем. Из этой теоремы Муавра—Лапласа следует, что при увеличении параметра n величина

$$\max_{x\in\mathbb{R}} |\mathcal{F}(x) - \mathcal{F}_\xi(x)|$$

будет стремиться к нулю. Поэтому можно утверждать, что в рассматриваемом примере нормальное распределение является *предельным* для биномиального. Аналогично, любое другое биномиальное распределение стремится к нормальному.

Функция нормального распределения определена на всем множестве $\mathbb{R}$, а функция биномиального распределения — только при $x > 0$. Но это не мешает рассматривать нормальный закон в качестве предельного для биномиального закона, поскольку вероятность события $\xi < 0$, вычисленная в соответствии с нормальным законом практически равна нулю (см. таблицу).

Оказывается, что многие случайные величины, являющиеся ошибками измерений при помощи приборов, имеют нормальный закон распределения. Этому закону подчиняются также многие случайные величины, являющиеся физиологическими характеристиками живых организмов определенного вида, например, рост или вес. С помощью предельных теорем теории вероятности можно показать, что, если случайная величина ξ имеет нормальный закон распределения, то параметр a является пределом (в некотором вероятностном смысле) среднего арифметического значений ξ, наблюдавшихся в n опытах, при условии, что n стремится к бесконечности. Поэтому при подборе параметров нормального закона, наиболее близко описывающего экспериментальные данные, в качестве a выбирают среднее значение ξ, наблюдавшееся при большом количестве опытов. Параметр σ характеризует среднее квадратическое отклонение наблюдавшихся значений от a. Примем это без доказательства.

Задача 3.23. *На станке изготовлено 250 деталей. При проверке этих деталей было измерено абсолютное (то есть по модулю) отклонение диаметров деталей от заданного числа. Данные измерений помещены в таблицу:*

Номер промежутка (i)	1	2	3	4	5
Интервал (Δ_i) (в мк)	$[0, 5)$	$[5, 10)$	$[10, 15)$	$[15, 20)$	$[20, 25]$
Число деталей (N_i)	15	75	100	50	10
Частота ($N_i/250$)	0.06	0.30	0.40	0.20	0.04

Выдвинуть гипотезу о характере распределения случайной величины ξ — абсолютного отклонения диаметра детали от заданного числа.

Прежде всего, заметим, что если существует вероятностная модель на $\mathbb{R}$ для рассматриваемой случайной величины, то ее функция распределения должна быть непрерывна. Действительно, из практических соображений следует, что бесконечно малым изменениям промежутка $(-\infty, x)$ соответствуют бесконечно малые изменения частот попадания значений ξ в промежуток. Вероятность получить при измерении абсолютного отклонения величину, строго равную числу a (с точностью, например, до 10^{-100}), практически, равна нулю. Ведь при измерении отклонений всегда имеется погрешность.

Допустим, что распределение имеет функцию плотности. Чтобы иметь представление об «очертаниях» графика функции плотности, составим специальную диаграмму, опираясь на опытные данные. Такая диаграмма называется в статистике *гистограммой*. Алгоритм ее построения таков:

1) результаты опыта (значения ξ) записываются в порядке неубывания;

2) множество наблюдавшихся значений разбивается на непересекающиеся промежутки Δ_i (вторая строка таблицы);

3) определяется число N_i значений (с учетом кратности), попадающих в каждый промежуток (третья строка таблицы);

4) вычисляется относительная частота попаданий в каждый промежуток: N_i/N, где N — общее число наблюдавшихся значений ξ (у нас $N = 250$);

5) на оси абсцисс строятся указанные выше промежутки;

6) на каждом из построенных промежутков как на основании строится прямоугольник, площадь которого равна относительной частоте.

Построенная гистограмма является аналогом функции плотности: сумма площадей прямоугольников, очевидно, равна единице. Высота i-го прямоугольника получается делением относительной частоты попадания значений ξ в промежуток Δ_i на длину этого промежутка. В статистике такое число называют *плотностью относительной частоты* на промежутке Δ_i. Обозначим его p_i. При увеличении числа измерений можно

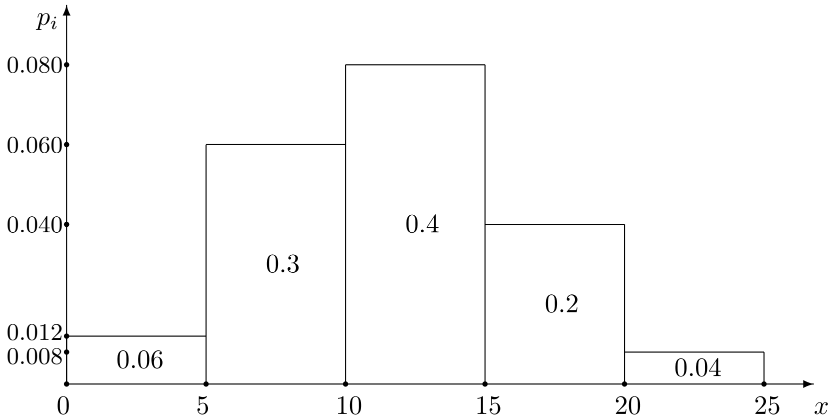


Рис. 3.11. Гистограмма

строить гистограмму с промежутками меньшей длины, в которых тем не менее будет содержаться достаточно много значений ξ. Верхняя линия, ограничивающая гистограмму, будет стремиться к абсолютно непрерывной кривой, «очертания» которой похожи на колокол. Это заметно уже на рис. 3.11. Естественная гипотеза: распределение ξ является нормальным.

Займемся теперь чисто технической проблемой: как вычислить вероятность попадания значений ξ в заданный промежуток $(x_1,\, x_2)$, если известно, что распределение является нормальным. Из формулы (3.19) в нашем частном случае вытекает основная расчетная формула

$$\mathsf{P}(\,(x_1, x_2)\,) = \frac{1}{\sigma\,\sqrt{2\pi}} \int\limits_{x_1}^{x_2} e^{-0.5\,((x-a)/\sigma)^2}\, dx. \tag{3.24}$$

Как известно, первообразная от функции вида (3.23) не является элементарной функцией. Поэтому для вычислений по формуле (3.24) применяется следующий прием. Рассмотрим «вспомогательную» функцию Лапласа:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{x} e^{-0.5\,t^2}\, dt.$$

Произведя в интеграле из формулы (3.24) замену переменной $(x-a)/\sigma =$

t, получим, что

$$\mathsf{P}(\,(x_1, x_2)\,) = \Phi\,\left(\frac{x_2 - a}{\sigma}\right) - \Phi\,\left(\frac{x_1 - a}{\sigma}\right). \tag{3.25}$$

Функция Лапласа не зависит от конкретных параметров a и σ. Они учитываются в формуле (3.25). Для функции Лапласа с помощью методов приближенного интегрирования составлены таблицы значений на промежутке [0, 5] с разной степенью точности. Вариант такой таблицы помещен в разделе «Приложение» нашего учебного пособия. Очевидно, что функция Лапласа является нечетной, следовательно, нет необходимости помещать в таблицу ее значения при отрицательных x. При $x > 0.5$ значения $\Phi(x)$ с большой точностью (до 10^{-5}) совпадают с числом 0.5, поскольку

$$\lim_{x\to+\infty} \Phi(x) = 0.5. \tag{3.26}$$

В дальнейшем, при решении задач мы, для краткости, будем записывать это свойство как $\Phi(+\infty) = 0.5$.

Простым, но очень важным свойством любой нормально распределенной случайной величины является так называемое

Правило трех сигм: *с вероятностью не менее, чем* 0.997, *значения случайной величины, имеющей нормальне распределение, содержатся в промежутке* $[a - 3\sigma,\ a + 3\sigma]$.

Справедливость этого правила вытекает из формулы (3.25) и того факта, что $\Phi(3) - \Phi(-3) = 2\,\Phi(3) \approx 0.997$.

Задача 3.24. *Вес одного заряда охотничьего пороха должен равняться* 2.3 *грамма. Этот заряд взвешивается на весах. В то время, как стрелка весов показывает нужный вес, этот вес, на деле, представляет собой случайную величину* ξ *с нормальным законом распределения, поскольку весы «не без греха». Среднее значение* ξ *(параметр* a*) равно нужному весу:* 2.3 *грамма. Среднее квадратическое отклонение (параметр* σ*) равно* 0.15 *грамма. Определить вероятность повреждения ружья, если максимально допустимый вес порохового заряда равен* 2.5 *граммам.*

Повреждение ружья наступит, если величина заряда превысит значение 2.5 грамма. Этому соответствует событие $\xi > 2.5$. Вычислим его вероятность, используя формулы (3.25) и (3.26):

$$\mathsf{P}((2.5, +\infty)) = \Phi\,(+\infty) - \Phi\,(\,(2.5 - 2.3)/0.15\,) \approx 0.5 - 0.41 = 0.09.$$

Нужное здесь значение $\Phi(1.3(3))$ приближенно подсчитано с помощью таблицы из раздела «Приложения».

Задача 3.25. *Какой ширины должно быть поле допуска, чтобы с вероятностью не более* 0.0027 *получалась деталь с проверяемым размером вне поля допуска, если случайные отклонения размера от середины поля допуска подчиняются закону нормального распределения с параметрами* $a = 0$ *и* $\sigma = 5$ *микрон?*

Здесь нас интересует событие вида $|\xi - a| > r$, где через r обозначена половина ширины (радиус) поля допуска. Вероятность этого события обозначим через $\mathsf{P}(|\xi - a| > r)$. Она должна удовлетворять неравенству

$$\mathsf{P}(|\xi - a| > r) \leqslant 0.0027,$$

которое равносильно неравенству

$$1 - \mathsf{P}(|\xi - a| > r) = \mathsf{P}(\,[a - r,\, a + r]\,) > 0.9973.$$

Применяя «правило трех сигм», получаем, что $r \geqslant 15$ микрон, следовательно, ширина допуска должна быть не менее 30 микрон.

В качестве дополнения к решению последней задачи рассмотрим метод нахождения числа r из неравенства

$$\mathsf{P}(\,[a - r,\, a + r]\,) > p.$$

Воспользовавшись формулой (3.25), получаем, что

$$\mathsf{P}(\,[a - r,\, a + r]\,) = 2\Phi(r/\sigma) > p.$$

Подберем по таблице значений $\Phi(x)$ наименьшее число x_0, такое что $\Phi(x_0) \geqslant p/2$. Остается лишь решить неравенство $r/\sigma \geqslant x_0$, откуда $r \geqslant x_0\sigma$.

Задача 3.26. *Рост случайно выбранного жителя густонаселенного острова Литтл — случайная величина, имеющая нормальный закон распределения с параметрами* $a = 140$ *см и* $\sigma = 5$ *см. Будем говорить, что рост* h *жителя острова не является значимым, если выполнено одно из двух условий:*

1) жители с ростом больше или равным h *составляют менее 1% всего населения;*

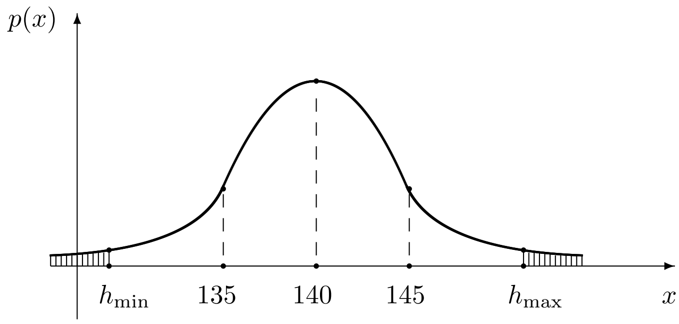


Рис. 3.12

2) жители с ростом меньше или равным h составляют менее 1% всего населения. Остальные значения роста назовем значимыми.

Какова вероятность, что рост случайно выбранного островитянина является значимым? Является ли рост 127 см значимым? Каков наибольший значимый рост? Чему равен наименьший значимый рост?

В силу условий 1) и 2), незначимыми объявлены «самые маленькие» и «самые большие», редко встречающиеся значения роста. Кавычки подчеркивают, что сравнение проводится нестандартно. Жители с незначимым ростом составляют менее 2% населения, но их доля в общем количестве островитян может быть, теоретически, сколь угодно близка к 2%. Следовательно, естественно считать, что вероятность незначимого роста равна 0.02 (2%), а вероятность значимого роста равна 0.98 (98%). Незначимый рост должен принадлежать либо промежутку вида $(h_{\max}, +\infty)$, либо промежутку вида $(-\infty, h_{\min})$, где $h_{\max}$, $h_{\min}$ — наибольший и наименьший значимый рост. Этим двум числам соответствуют точки на числовой оси, симметричные относительно центра $x = a = 140$ (см). На рис. 3.12 каждая из заштрихованных областей имеет площадь, равную вероятности 0.01 (1%). Ответить на второй вопрос будет легко, если определены значения $h_{\max}$ и $h_{\min}$. Найдем их с помощью функции Лапласа из соотношений:

$$\mathsf{P}(\,[h_{\max}, +\infty)\,) = 0.5 - \Phi(\,(h_{\max} - 140)/5\,) = 0.1;$$

$$\mathsf{P}(\,(-\infty, h_{\min}]\,) = \Phi(\,(h_{\min} - 140)/5\,) + 0.5 = 0.1.$$

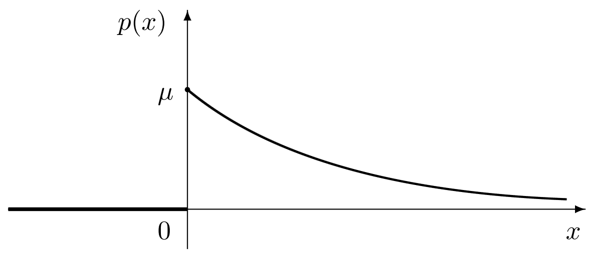


Рис. 3.13. Плотность экспоненциального распределения

Приближенные значения таковы: $h_{\max} \approx 151.75$ (см), $h_{\min} \approx 128.25$ (см). Следовательно, рост 127 см не является значимым.

В заключение отметим, что рост островитянина, строго говоря — дискретная случайная величина. Поэтому подчиняться нормальному закону она может лишь с некоторой, в данном случае очень малой, погрешностью.

Экспоненциальное (показательное) распределение на прямой. Если функция плотности распределения случайной величины имеет вид

$$p(x) = \begin{cases} 0 & \text{при } \ x < 0, \\ \mu\, e^{-\mu x} & \text{при } \ x \geqslant 0, \end{cases} \tag{3.27}$$

где число μ — положительный параметр, то соответствующее распределение называется *экспоненциальным распределением на прямой*.

График функции плотности экспоненциального распределения имеет вид, изображенный на рис. 3.13. Методами математической статистики установлено, что экспоненциальное распределение имеет срок службы различных механических устройств, время безотказной работы отдельных элементов этих устройств, величина предельной нагрузки на испытываемую деталь, и многие другие случайные величины.

Проинтегрировав функцию плотности (3.27) на промежутке $(-\infty, x)$, получим аналитическое выражение для функции экспоненциального распределения:

$$\mathcal{F}(x) = \begin{cases} 0 & \text{при } \ x \leqslant 0, \\ 1 - e^{-\mu x} & \text{при } \ x \geqslant 0. \end{cases} \tag{3.28}$$

График $\mathcal{F}(x)$ изображен на рис. 3.14. Число μ является единственным

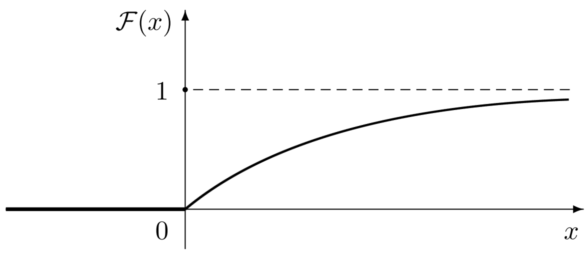


Рис. 3.14. Функция экспоненциального распределения

параметром экспоненциального закона распределения. Можно показать, что величина $1/\mu$ имеет такой же «статистический смысл», как и параметр a для нормального закона: к этому числу при большом числе опытов «стремится» среднее наблюдаемое значение случайной величины, имеющей соответствующее распределение.

Задача 3.27. *Срок «службы» китайских кроссовок — случайная величина, имеющая экспоненциальный закон распределения. Известно, что средний срок службы составляет один год. Найти параметр μ и вычислить вероятность того, что кроссовки прослужат не менее двух лет.*

Как говорилось выше, $1/\mu$ — это средний срок службы кроссовок, следовательно, у нас $\mu = 1$. Для вычисления вероятности события $\xi \geqslant 2$ воспользуемся формулой (3.28).

$$\mathsf{P}(\,[2, +\infty)\,) = 1 - \mathcal{F}(2) = 1 - (1 - e^{-2}) = e^{-2} \approx 0.14.$$

3.8. Числовые характеристики абсолютно непрерывных случайных величин

В этом разделе мы определим основные числовые характеристики абсолютно непрерывных случайных величин и выясним, чему они равны в случаях, когда плотности имеют стандартный вид.

Определение математического ожидания абсолютно непрерывной случайной величины. Пусть абсолютно непрерывная случайная величина ξ имеет известую плотность распределения $p(x)$. Математическим ожиданием M_ξ случайной величины ξ называется несобственный интеграл вида

$$M_\xi = \int\limits_{-\infty}^{+\infty} xp(x)\,dx. \qquad (3.29)$$

Если интеграл расходится, то говорят, что математическое ожидание не существует.

В качестве упражнений докажите, что стандартные абсолютно непрерывные случайные величины имеют следующие значения математического ожидания:
1) равномерный закон: $M_\xi = (a+b)/2$,
2) нормальный закон: $M_\xi = a$,
3) экспоненциальный закон: $M_\xi = 1/\mu$.
В случае, когда закон распределения не является стандартным, можно найти математическое ожидание по определению.

Определение дисперсии абсолютно непрерывной случайной величины. Пусть абсолютно непрерывная случайная величина ξ имеет известую плотность распределения $p(x)$. Пусть существует математическое ожидание M_ξ. Дисперсией D_ξ случайной величины ξ называется несобственный интеграл:

$$D_\xi = \int\limits_{-\infty}^{+\infty} (x - M_\xi)^2 p(x)\,dx. \qquad (3.30)$$

Если интеграл расходится, говорят, что дисперсия не существует. Для вычисления дисперсии можно применить более удобную формулу

$$D_\xi = \int\limits_{-\infty}^{+\infty} x^2 p(x)\,dx - M_\xi^2. \qquad (3.31)$$

Равносильность формул (3.30) и (3.31) доказывается с помощью преобразований подынтегральной функции в (3.30) и определения (3.29). Приведем, далее, без доказательства формулы для вычисления дисперсии случайных величин, имеющих стандартные плотности распределения.
1) равномерный закон: $D_\xi = (b-a)^2/12$,
2) нормальный закон: $D_\xi = \sigma^2$,
3) экспоненциальный закон: $D_\xi = 1/\mu^2$.

Легко заметить, что дисперсия измеряется не в таких единицах, как математическое ожидание: единицы измерения возводятся в квадрат. Это не всегда удобно. Для единообразия единиц измерения из дисперсии извлекают квадратный корень.

Определение среднего квадратического отклонения. Средним квадратическим отклонением случайной величины ξ называется квадратный корень из ее дисперсии: $\sigma_\xi = \sqrt{D_\xi}$.

Математическое ожидание, как и для дискретных случайных величин, является «предсказанием» среднего арифметического наблюдавшихся значений изучаемой случайной величины (то есть среднего значения). Дисперсия — «предсказанием» среднего квадрата отклонения от среднего значения.

Рассмотрим для абсолютно непрерывных величин проблему определения минимального интервала вида $(M_\xi - r,\ M_\xi + r)$, в который значения ξ попадают с заданной вероятностью γ. Если закон распределения случайной величины известен, то такой интервал можно точно определить, интегрируя плотность по симметричному интервалу с центром в M_ξ и переменным радиусом r. Для многих случайных величин (в частности, для нормального распределения) такой метод опирается на численное интегрирование. С помощью теоремы Чебышева (теорема 3.1) можно получить простую, но приближенную оценку интервала вида $(M_\xi - r,\ M_\xi + r)$, в который значения ξ попадают с вероятностью заведомо большей, чем γ. Этот метод не требует знания закона распределения, а только знание M_ξ, D_ξ и γ.

Для приближенного определения радиуса окрестности математического ожидания, в которую значения случайной величины попадают с вероятностью больше, чем γ, надо правую часть неравенства (3.6) приравнять γ.

Если в неравенство (3.6) вместо r подставить $3\ \sigma_\xi$, то соответствующая вероятность будет не меньше, чем $8/9$. Этот результат любопытно сравнить с «точным» результатом для нормального закона, который равен 0.997.

В заключение параграфа рассмотрим «абстрактную» задачу, решение которой использует определения и свойства функции распределения, плотности распределения, математического ожидания и дисперсии абсолютно непрерывной случайной величины.

Задача 3.28. *Плотность распределения случайной величины* X

имеет вид:

$$p(x) = \begin{cases} \dfrac{x+1}{2} & \textit{при} \ \ -1 \leqslant x \leqslant 0, \\ \dfrac{C-x}{2C} & \textit{при} \ \ 0 \leqslant x \leqslant C, \\ 0 & \textit{при} \ \ x < -1 \ \ \textit{или} \ \ x > C. \end{cases}$$

где число C — положительный параметр.

Требуется найти значение параметра C, числовые характеристики случайной величины X, функцию распределения и вероятность того, что $X < 0.5\ C$. Кроме того требуется построить графики функции распределения и плотности распределения.

Решение задачи начнем с определения значения параметра C. Воспользуемся формулой (3.20):

$$1 = \int\limits_{-\infty}^{+\infty} p(x)\, dx = \int\limits_{-\infty}^{-1} 0\, dx + \int\limits_{-1}^{0} \frac{x+1}{2}\, dx + \int\limits_{0}^{C} \frac{C-x}{2C}\, dx + \int\limits_{C}^{+\infty} 0\, dx =$$

$$= 0 + \frac{1}{4} + \frac{C}{4} + 0 = \frac{C+1}{4}.$$

Из полученного равенства вытекает, что $C = 3$. Тот же результат можно получить, опираясь на геометрический смысл интеграла. Общая площадь под графиком функции $p(x)$ равна единице. Площадь над лучами $(-\infty, -1)$, $(C, +\infty)$ равна нулю (поскольку там $p(x) = 0$). Отрезку $[-1, 0]$ соответствует площадь треугольника, равная 0.25. Отрезку $[0, C]$ соответствует площадь, равная $0.25\ C$. Следовательно, $1 = 0.25 + 0.25\ C$, то есть $C = 3$.

Формула плотности распределения вероятности имеет вид:

$$p(x) = \begin{cases} \dfrac{x+1}{2} & \text{при} \ \ -1 \leqslant x \leqslant 0, \\ \dfrac{3-x}{6} & \text{при} \ \ 0 \leqslant x \leqslant 3, \\ 0 & \text{при} \ \ x < -1 \ \ \text{или} \ \ x > 3. \end{cases}$$

График этой функции приведен на рис. 3.15.

С помощью функции плотности найдем числовые характеристики случайной величины X по формулам (3.29), (3.31):

$$M_X = \int_{-\infty}^{+\infty} xp(x)\,dx = \int_{-\infty}^{-1} 0\,dx + \int_{-1}^{0} \frac{x(x+1)}{2}\,dx + \int_{0}^{3} \frac{x(3-x)}{6}\,dx + \int_{3}^{+\infty} 0\,dx =$$

$$= 0 - \frac{1}{12} + \frac{9}{12} + 0 = \frac{2}{3},$$

$$D_X = \int_{-\infty}^{+\infty} x^2 p(x)\,dx - M_X^2 =$$

$$= \int_{-\infty}^{-1} 0\,dx + \int_{-1}^{0} \frac{x^2\cdot(x+1)}{2}\,dx + \int_{0}^{3} \frac{x^2\cdot(3-x)}{6}\,dx + \int_{3}^{+\infty} 0\,dx - \frac{4}{9} =$$

$$= 0 + \frac{1}{24} + \frac{27}{24} + 0 - \frac{4}{9} = \frac{13}{18}.$$

Интегрируя функцию плотности на промежутке $(-\infty,\, x)$, получим формулы для функции распределения $\mathcal{F}(x)$:

$$\mathcal{F}(x) = \begin{cases} 0 & \text{при} \;\; x < -1, \\ \dfrac{(x+1)^2}{4} & \text{при} \;\; -1 \leqslant x \leqslant 0, \\ 1 - \dfrac{(3-x)^2}{12} & \text{при} \;\; 0 \leqslant x \leqslant 3, \\ 1 & \text{при} \;\; x > 3. \end{cases}$$

График функции распределения изображен на рис. 3.16. Осталось вычислить вероятность того, что $X < 1.5$. Это можно сделать с помощью функции распределения:

$$\mathsf{P}(X < 1.5) = \mathcal{F}(1.5) = 1 - (1.5)^2/12 = 13/16.$$

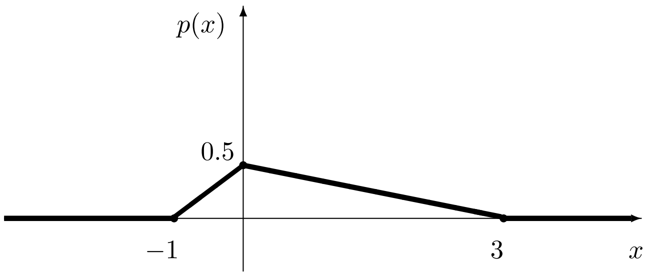


Рис. 3.15. Плотность распределения вероятности

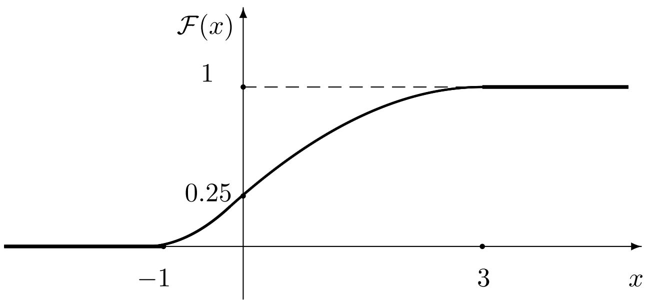


Рис. 3.16. Функция распределения вероятности

3.9. Функции от одной случайной величины

В этом параграфе мы рассмотрим случайные величины, являющиеся функциями от других, возможно «стандартных», случайных величин. Мы покажем, как найти закон распределения функции и ее числовые характеристики, если известен закон распределения аргумента.

Определение функции от случайной величины. Пусть задана некоторая случайная величина ξ с областью определения Ω и множеством значений V_ξ, на котором, в свою очередь, определена обычная функция $y(x)$ от одной действительной переменной x. Тогда случайная величина Y, отображающая Ω в множество действительных чисел по правилу $Y = y(\xi)$, называется *функцией от случайной величины* ξ.

Задача 3.29. *Стрелок попадает по мишени с вероятностью* 0.8 *при каждом выстреле, независимо от предыдущих выстрелов. Стрельба ведется до первого попадания. Случайная величина* Y *- число сделанных промахов. Найти закон распределения* Y *и ее числовые характеристики.*

Рассмотрим стандартную случайную величину ξ - число сделанных выстрелов до первого «успеха». Эта величина имеет геометрический закон распределения, в котором

$$\mathsf{P}(\xi = k) = 0.2^{k-1} \cdot 0.8, \qquad M_\xi = 1/0.8 = 1.25, \qquad D_\xi = 0.2/0.8^2 = 0.3125.$$

Очевидно, что $Y = y(\xi) = \xi - 1$. Следовательно, Y имеет множество значений $\{0, 1, 2, \dots\}$, причем $Y = m$ тогда и только тогда, когда $\xi = m + 1$. Отсюда вытекает, что закон распределения Y задается формулой $\mathsf{P}(Y = m) = 0.2^m \cdot 0.8$. Из определения математического ожидания вытекает, что

$$\begin{aligned} M_Y = M_{\xi-1} = (x_1 - 1) \cdot p_1 + (x_2 - 1) \cdot p_2 + \dots = \\ = M_\xi - (p_1 + p_2 + \dots) = M_\xi - 1 = 0.25. \end{aligned}$$

Из определения дисперсии вытекает, что $D_Y = D_\xi = 0.3125$.

Задача 3.30. *Пусть случайная величина* ξ *принимает равновозможно любое целое значение от* -5 *до* 5, *включительно. Других значений нет. Случайная величина* Y *связана с* ξ *функциональной зависимостью:* $Y = y(\xi) = \xi^2 - \xi - 6$. *Найти закон распределения* Y*, ее математическое ожидание и дисперсию.*

Поскольку в множестве значений ξ содержится всего 11 точек, $\mathsf{P}(\xi = a) = 1/11$, для любого целого a от -5 до 5. Величина Y принимает значение 24 тогда и только тогда, когда $\xi = -5$, следовательно, $\mathsf{P}(Y = 24) = 1/11$. Каждому из остальных значений Y должна соответствовать вероятность 2/11, так как каждое такое значение Y получается из двух различных значений ξ, а именно, $Y = 0$ тогда и только тогда, когда $\xi = 3$ или $\xi = -2$; $Y = 14$ при $\xi = -4$ или $\xi = 5$; $Y = 6$ при $\xi = -3$ или $\xi = 4$; $Y = -4$ при $\xi = -1$ или $\xi = 2$; $Y = -6$ при $\xi = 1$ или $\xi = 2$. Ряд распределения Y имеет вид:

Значения Y	-6	-4	0	6	14	24
Вероятности	2/11	2/11	2/11	2/11	2/11	1/11

В общем случае ряд распределения функции y от дискретной случайной величины ξ (с известным распределением) можно найти по следующим правилам:

1) определить множество всех значений Y функции $y(\xi)$ (оно, очевидно, является конечным или счетным);

2) для каждого значения Y функции $y(\xi)$ найти множество X всех значений ξ, от которых получаем именно такое значение функции (то есть найти полный прообраз Y);

3) каждому значению Y функции $y(\xi)$ приписать вероятность его прообраза X.

Зная ряд распределения функции, легко можно подсчитать ее математическое ожидание и дисперсию. Так, например, в задаче 3.30

$$M_Y = 24 \cdot \frac{1}{11} + 14 \cdot \frac{2}{11} + 6 \cdot \frac{2}{11} - 4 \cdot \frac{2}{11} - 6 \cdot \frac{2}{11} = 4,$$

$$D_Y = 24^2 \cdot \frac{1}{11} + 14^2 \cdot \frac{2}{11} + 6^2 \cdot \frac{2}{11} + (-4)^2 \cdot \frac{2}{11} + (-6)^2 \cdot \frac{2}{11} - 16 = 88.$$

Очевидно, что те же значения математического ожидания и дисперсии можно было получить, не вникая в закон распределения Y, а пользуясь только законом распределения ξ. Сформулируем правила сразу в общем виде:

$$M_Y = \sum_i y(x_i) \cdot p_i, \tag{3.32}$$

$$D_Y = \sum_i y(x_i)^2 \cdot p_i - {M_Y}^2. \tag{3.33}$$

В формулах (3.32) значения ξ обозначены через x_i, а соответствующие им вероятности — через p_i.

Заметим, что из существования математического ожидания и дисперсии для данной дискретной случайной величины ξ, не вытекает существование математического ожидания и дисперсии для функции от этой случайной величины (для бесконечных таблиц распределения).

Можно доказать, что если существуют математическое ожидание и дисперсия функции от абсолютно непрерывной случайной величины, то они вычисляются по формулам:

$$M_Y = \int\limits_{-\infty}^{+\infty} y(x)p(x)\,dx, \tag{3.34}$$

$$D_Y = \int\limits_{-\infty}^{+\infty} (y(x) - M_Y)^2 p(x)\,dx = \int\limits_{-\infty}^{+\infty} y(x)^2 p(x)\,dx - {M_Y}^2. \tag{3.35}$$

Из формул (3.32) – (3.35) вытекают полезные следствия (докажите самостоятельно в качестве упражнения):

а) если $y(x) = kx + b$, то для $Y = y(\xi)$ имеем $M_Y = kM_\xi + b$ и $D_Y = k^2 D_\xi$;

б) если $y(x) = x^2$, то для $Y = y(\xi)$ имеем $M_Y = D_\xi + {M_\xi}^2$;

в) если $y(x) = y_1(x) + y_2(x)$, то для $Y = y(\xi)$ имеем $M_Y = M_1 + M_2$, где M_1 — математическое ожидание для $y_1(\xi)$, а M_2 — для $y_2(\xi)$ (предполагаем, что M_1 и M_2 существуют).

Задача найти закон распределения функции от абсолютно непрерывной случайной величины сводится к проблеме определения ее типа и отыскания ее функции распределения. С общими принципами решения этой задачи можно ознакомиться в [2, 4, 9]. Здесь мы рассмотрим два примера, иллюстрирующих эти общие принципы.

Задача 3.31. *Случайная величина* ξ *имеет равномерное распределение на отрезке* $[0,\,1]$. *Найти закон распределения величины* $Y = 2\xi + 3$ *и ее числовые характеристики.*

Вычислим сперва математическое ожидание и дисперсию Y по следствиям из формул (3.34), (3.35): $M_Y = 2\,M_\xi + 3 = 2 \cdot 0.5 + 3 = 4$, $D_Y = 4\,D_\xi = 1/3$.

Закон распределения Y найдем из следующих соображений:

$$\mathcal{F}_Y(x) = \mathsf{P}(Y < x) = \mathsf{P}(2\xi + 3 < x) = \mathsf{P}(\xi < 0.5(x-3)) = \mathcal{F}_\xi(0.5(x-3));$$

$$p_Y(x) = \frac{d\mathcal{F}_Y(x)}{dx} = 0.5 p_\xi(0.5(x-3)) = \begin{cases} 0 & \text{при } \ x < 3, \\ 1/2 & \text{при } \ 3 \leqslant x \leqslant 5, \\ 0 & \text{при } \ x > 5. \end{cases}$$

Последняя формула вытекает из того, что число x принадлежит отрезку $[3,\, 5]$ тогда и только тогда, когда число $0.5(x-3)$ принадлежит отрезку $[0,\, 1]$.

Задача 3.32. *Случайная величина ξ имеет равномерное распределение на отрезке* $[0,\, 1]$. *Найти закон распределения величины $Y = \xi^2$ и ее числовые характеристики.*

Как было сказано выше, $M_Y = D_\xi + {M_\xi}^2 = 1/12 + 1/4 = 1/3$. Вычислим ту же величину с помощью формулы (3.34): $M_Y = \int\limits_0^1 x^2\, dx = 1/3$.

Дисперсию Y определим по формуле (3.35): $D_Y = \int\limits_0^1 x^4\, dx - 1/9 = 4/45$.

Найдем плотность распределения Y:

$$\mathcal{F}_Y(x) = \mathsf{P}(\xi^2 < x) = \begin{cases} \mathsf{P}(\xi < \sqrt{x}) = \mathcal{F}_\xi(\sqrt{x}) = \sqrt{x} & \text{при } \ 0 \leqslant x \leqslant 1, \\ 0 & \text{при } \ x < 0, \\ 1 & \text{при } \ x \geqslant 1; \end{cases}$$

$$p_Y(x) = \frac{d\mathcal{F}_Y(x)}{dx} = \begin{cases} 0 & \text{при } \ x < 0, \\ \dfrac{1}{2\sqrt{x}} & \text{при } \ 0 < x < 1, \\ 0 & \text{при } \ x > 1. \end{cases}$$

Вычислим с помощью плотности Y математическое ожидание и дисперсию Y:

$$M_Y = \int\limits_0^1 \frac{x}{2\sqrt{x}}\, dx = \frac{1}{3}, \qquad D_Y = \int\limits_0^1 \frac{x^2}{2\sqrt{x}}\, dx - {M_Y}^2 = \frac{4}{45}.$$

3.10. Задачи для самостоятельного решения

3.10.1. Первый уровень сложности

1. Время ожидания автобуса T имеет равномерное распределение вероятности на отрезке $[0, 8]$ (минут). Найти плотность распределения, функцию распределения, математическое ожидание, дисперсию T. Построить графики плотности и функции распределения. Вычислить $\mathsf{P}(2 < T < 5)$.

2. Масса случайно выбранного из одной партии пакета с крупой имеет нормальный закон распределения вероятности. Математическое ожидание равно 0.9 кг, среднее квадратическое отклонение равно 0.005 кг. Запишите формулу плотности, постройте график плотности, подсчитайте вероятность того, что масса пакета отклоняется от 0.9 кг (в ту или иную сторону) не более, чем на 10 грамм.

3. Время бесперебойной работы прибора имеет экспоненциальный закон распределения вероятности. Математическое ожидание равно 14 месяцам. Запишите формулу плотности, постройте график плотности, подсчитайте вероятность того, что прибор проработает не более 15 месяцев без перебоев.

4. Диаметр круглого мокрого пятна, которое получается при попадании дождевой капли на асфальт, — случайная величина D. Она имеет нормальный закон распределения с параметрами $a = 5$ мм и $\sigma = 1$ мм. Найдите закон распределения случайной величины L, равной длине окружности пятна, постройте график плотности распределения вероятности для L, подсчитайте вероятность того, что $L < 15$ мм.

5. Погрешность ξ измерения длины отрезка с помощью линейки имеет равномерный закон распределения на $[-0.5, 0.5]$ (миллиметров). С помощью такой линейки измеряют длину a стороны квадрата. Получено среднее значение длины стороны, равное 11 мм. Случайная величина S — вычисляемая по формуле $S = a^2$ площадь квадрата. Приняв $M_a = 11$, $a = 11 + \xi$, найдите закон распределения S, математическое ожидание и дисперсию S.

6. Давление в скороварке подчинено нормальному закону распределения вероятности. Среднее давление равно пяти атмосферам. Среднее квадратическое отклонение равно 0.7 атмосферы. Клапан срабатывает при давлении, превышающем семь атмосфер. Найдите вероятность срабатывания клапана.

7. Морковка считается первосортной, если ее кончик не разветвлен и максимальный диаметр сечения, перпендикулярного вертикальной оси, находится в диапазоне от 15 до 30 миллиметров. Пусть на данном поле разветвленные морковки попадаются с вероятностью 0.2, независимо от их размеров. Средний максимальный диаметр морковки равен 20 мм, среднее квадратическое отклонение его равно 3 мм. Подсчитайте процент морковок первого сорта на этом поле. (Он равен вероятности того, что случайно выбранная морковка первосортна, увеличенной в сто раз.)

8. Может ли функция

$$\mathcal{F}(x) = \begin{cases} 0 & \text{при } x < 0, \\ 1 & \text{при } x \geqslant 0 \end{cases}$$

быть функцией распределения вероятности?

9. Определите, к какому типу относится функция распределения

$$\mathcal{F}(x) = \begin{cases} 0 & \text{при } x \leqslant 1, \\ 0.1 & \text{при } 1 < x \leqslant 2, \\ 0.5(x-1) & \text{при } 2 < x \leqslant 3, \\ 1 & \text{при } x > 3. \end{cases}$$

10. Математическое ожидание случайной величины X равно пяти, среднее квадратическое отклонение равно двум. Может ли X иметь равномерный закон распределения на отрезке $[0, 10]$? Может ли X иметь показательный закон распределения?

3.10.2. Второй уровень сложности

1. Случайная величина X имеет равномерный закон распределения на отрезке $[a, b]$. При этом $\mathsf{P}(X \geqslant 2) = 0.1$ и $\mathsf{P}(X \leqslant 0) = 0.4$. Найдите значения параметров a и b.

2. Случайная величина X имеет равномерный закон распределения на отрезке $[a, b]$. При этом $\mathsf{P}(X \leqslant 0) = 0.4$ и $\mathsf{P}(X \leqslant 2) = 0.8$. Найдите значения параметров a и b.

3. Случайная величина X имеет равномерный закон распределения на отрезке $[a, b]$. При этом $\mathsf{P}(1 \leqslant X \leqslant 2) = 0.2$ и $\mathsf{P}(4 \leqslant X \leqslant 6) = 0.4$. Найдите значения параметров a и b.

4. Случайная величина X имеет равномерный закон распределения на отрезке $[a, b]$. При этом $\mathsf{P}(1 \leqslant X \leqslant 2) = 0.2$ и $\mathsf{P}(4 \leqslant X \leqslant 6) = 0.6$. Найдите значения параметров a и b.

5. Случайная величина X имеет равномерный закон распределения на отрезке $[a, b]$. При этом $\mathsf{P}(1 \leqslant X \leqslant 2) = 0.1$ и $\mathsf{P}(4 \leqslant X \leqslant 6) = 0.2$. Найдите значения параметров a и b.

6. Случайная величина Y имеет экспоненциальный закон распределения, $M_Y = 0.5$, $Z = 3Y$. Найдите плотность распределения Z и величину M_Z.

7. Случайная величина Y имеет экспоненциальный закон распределения, $M_Y = 0.5$, $Z = 3Y + 1$. Найдите плотность распределения Z и величину M_Z .

8. Плотность распределения случайной величины Y имеет вид:

$$p_Y(x) = \begin{cases} 0 & \text{при } |x| > 1, \\ 1/3 & \text{при } -1 \leqslant x \leqslant 0, \\ C & \text{при } 0 < x \leqslant 1. \end{cases}$$

Определите значение C. Постройте график функции распределения, Вычислите вероятность того, что $Y \geqslant 0.5$. Подсчитайте математическое ожидание и дисперсию Y.

9. Плотность распределения случайной величины Y имеет вид:

$$p_Y(x) = \begin{cases} 0 & \text{при } |x-1| > 1, \\ C|x-1| & \text{при } 0 \leqslant x \leqslant 2. \end{cases}$$

Определите значение C. Постройте график функции распределения, найдите ее аналитический вид. Вычислите вероятность того, что $Y \leqslant 0.5$. Подсчитайте математическое ожидание и дисперсию Y.

10. Плотность распределения случайной величины Y имеет вид:

$$p_Y(x) = \begin{cases} 0 & \text{при } |x-1.5| > 1.5, \\ C(1-x/3) & \text{при } 0 \leqslant x \leqslant 3. \end{cases}$$

Определите значение C. Постройте график функции распределения, найдите ее аналитический вид. Вычислите вероятность того, что $Y \geqslant 0.8$. Подсчитайте математическое ожидание и дисперсию Y.

11. Плотность распределения случайной величины Y имеет вид:

$$p_Y(x) = \begin{cases} 0 & \text{при} \ \ x < 1, \\ Ce^{1-x} & \text{при} \ \ x \geqslant 1. \end{cases}$$

Определите значение C. Постройте график функции распределения, Вычислите вероятность того, что $Y \leqslant 0.7$. найдите ее аналитический вид. Подсчитайте математическое ожидание и дисперсию Y.

12. Плотность распределения случайной величины Y имеет вид:

$$p_Y(x) = \begin{cases} 0 & \text{при} \ \ |x| > 1, \\ C(|x| + 1/4) & \text{при} \ \ |x| \leqslant 1. \end{cases}$$

Определите значение C. Постройте график функции распределения, найдите ее аналитический вид. Вычислите вероятность того, что $Y \geqslant 0.3$. Подсчитайте математическое ожидание и дисперсию Y.

13. Вес мешка с сахаром имеет нормальный закон распределения с параметрами: $a = 50$, $\sigma = 0.5$ (килограмм). Закупочная цена за мешок равна 1000 рублей. При закупке мешок не взвешивается. Сахар продаем по 21 рублю за килограмм. С какой вероятностью потерпим убыток? С какой вероятностью получим прибыль не менее 50 рублей?

14. Вес пакета с орехами имеет равномерное распределение на отрезке от 50 до 60 грамм. Закупочная цена орехов была равна 100 рублям за килограмм. Какую минимальную цену за пакет надо назначить, чтобы с вероятностью не менее 0.9 получить не менее рубля прибыли за пакет?

15. Вес пакета с орехами имеет нормальный закон распределения с параметрами $a = 30$, $\sigma = 3$ (грамм). Закупочная цена орехов была равна 100 рублям за килограмм. Какую минимальную цену за пакет надо назначить, чтобы с вероятностью не менее 0.9 получить не менее рубля прибыли за пакет?

16. На координатной плоскости начерчена четвертая часть окружности с центром в начале координат и единичным радиусом. Точка на окружности задается с помощью случайного угла X между радиусом и осью абсцисс (против часовой стрелки). Величина угла имеет равномерный закон распределения на промежутке $[0, \pi/2]$. Найдите закон распределения для проекции точки на ось ординат.

Приложение А. Таблица значений нормированной функции Лапласа

Ниже приведена таблица значений функции

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-t^2/2} dt$$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.0000	1.00	0.3413	2.00	0.4772
0.05	0.0199	1.05	0.3531	2.10	0.4821
0.10	0.0398	1.10	0.3643	2.20	0.4861
0.15	0.0596	1.15	0.3749	2.30	0.4893
0.20	0.0793	1.20	0.3849	2.40	0.4918
0.25	0.0987	1.25	0.3944	2.50	0.4938
0.30	0.1179	1.30	0.4032	2.60	0.4953
0.35	0.1368	1.35	0.4115	2.70	0.4965
0.40	0.1554	1.40	0.4192	2.80	0.4974
0.45	0.1736	1.45	0.4265	2.90	0.4981
0.50	0.1915	1.50	0.4332	3.00	0.4986
0.55	0.2088	1.55	0.4394	3.20	0.4993
0.60	0.2257	1.60	0.4452	3.40	0.4997
0.65	0.2422	1.65	0.4505	3.60	0.4998
0.70	0.2580	1.70	0.4554	3.80	0.4999
0.75	0.2734	1.75	0.4599	4.00	0.5000
0.80	0.2881	1.80	0.4641	5.00	0.5000
0.85	0.3023	1.85	0.4678	$+\infty$	0.5000
0.90	0.3159	1.90	0.4713		
0.95	0.3289	1.95	0.4744		

Приложение Б. Первичная обработка опытных данных при изучении случайной величины

Б.1. Введение

Проведение эксперимента, запись его результатов, оформление результатов эксперимента в виде графиков или таблиц, определение числовых характеристик наблюдаемых явлений — все это необходимые составляющие исследовательской работы в различных областях человеческой деятельности, от выведения новых сортов растений до анализа педагогического воздействия на группу людей.

Предположим, что проведен эксперимент, в процессе которого несколько раз измерены значения некоторой случайной величины X. Полученные значения записаны в таблицу. Можно считать, что вы провели наблюдение за случайной величиной и получили так называемые «наблюдавшиеся значения». Вполне возможно, что не все значения, которые могла бы принимать величина X, удалось получить опытным путем. Поэтому принято называть наблюдавшиеся значения «выборочными». С какой целью проводится наблюдение? С целью найти подходящую математическую модель (закон распределения вероятности), с помощью которой можно было бы «предсказывать поведение» случайной величины: каково множество ее значений, какие значения наблюдаются чаще, какие — реже, чему равно среднее ожидаемое значение, дисперсия. Для достижения указанной цели необходимо решить последовательно несколько задач.

Б.2. Построение вариационного ряда

Прежде всего определим тип изучаемой случайной величины: дискретный или непрерывный (смешанный тип мы не рассматриваем). Это можно сделать на интуитивном уровне. Если множество значений случайной величины X конечно или счетно, то она является дискретной. Если же множество значений X несчетно, причем для любого числа x вероятности того, что $X < x$ и $X < x + \varepsilon$ близки при малых значениях ε, то величина X является непрерывной. Для дискретной случайной величины построим так называемый «дискретный вариационный ряд», для непрерывной — «интервальный вариационный ряд».

Дискретный вариационный ряд представляет собой таблицу — аналог закона распределения дискретной случайной величины. Для его построения необходимо выполнить следующие операции.

1. *Ранжирование.* Результаты наблюдений над случайной величиной расположим в порядке неубывания.

2. *Выделение вариантов.* Из полученной последовательности выделим подпоследовательность, все числа в которой различны. Каждое из этих значений называется «вариантом».

Пример. Пусть исходные результаты имели вид:

Номер опыта по посеву 20 семян	Число взошедших семян из 20
1	18
2	13
3	18
4	15
5	15

Результатом ранжирования является последовательность: 13, 15, 15, 18, 18. Варианты: 13, 15, 18.

3. *Определение абсолютных частот.* Найдем число повторений каждого варианта в исходном наборе чисел. Оно называется «абсолютной частотой» данного варианта. В приведенном выше примере абсолютная частота значения 13 равна единице, абсолютные частоты значений 15 и 18 равны двум.

4. *Определение относительных частот.* Абсолютную частоту каждого варианта разделим на общее число опытов. Полученный результат, называется «относительной частотой» данного варианта. В нашем примере относительные частоты равны 0.2, 0.4, 0.4. Их сумма, очевидно, всегда должна быть равна единице. Если при вычислении относительных частот возникает необходимость округления дробей, то оно проводится таким образом, чтобы сумма приближенных результатов также равнялась единице. Проще всего этого можно добиться, приписав последнему

варианту относительную частоту, полученную вычитанием из единицы суммы всех относительных частот для других вариантов.

Произведя операции 1 – 4, поместим их результаты в сводную таблицу, которая и называется «дискретный вариационный ряд»:

Номер по порядку	Значение варианта	Абсолютная частота	Относительная частота
1	13	1	0.2
2	15	2	0.4
3	18	2	0.4

Определение. Дискретным вариационным рядом называется ранжированная совокупность вариантов с соответствующими им абсолютными и относительными частотами.

Интервальный вариационный ряд дает представление о законе распределения случайной величины непрерывного типа, которая даже при очень большом количестве наблюдений и измерений практически не повторяет свои значения. Если для таких величин составить дискретный вариационный ряд, он будет иметь очень много строк, но абсолютные частоты большинства вариантов будут равны единице. По такому ряду трудно выявить закономерности появления тех или иных значений случайной величины. Аналогичная ситуация возникает, когда чувствительность приборов, с помощью которых производятся измерения случайной величины, не позволяет различать между собой ее значения в пределах каких-то промежутков. В подобных случаях следует строить интервальный вариационный ряд.

Определение. Интервальным вариационным рядом называется упорядоченная совокупность непересекающихся промежутков изменения случайной величины, объединением которых является промежуток, включающий все ее значения, с соответствующими абсолютными и относительными частотами попаданий в каждый промежуток.

Пусть, например, в результате измерения длин листьев дерева с точностью до 0.1 мм получилась совокупность из 50 чисел:

60.6	54.9	40.1	58.1	46.5	42.9	33.7	44.8	37.1	57.9
40.3	38.5	59.2	52.4	29.7	28.3	30.4	50.7	54.4	42.4
44.9	44.3	22.1	52.6	51.4	46.8	32.4	43.3	48.6	62.8
48.4	56.4	52.6	41.4	47.6	44.6	39.2	50.4	32.9	54.7
34.4	32.8	35.9	41.4	50.2	35.1	33.0	57.5	32.7	32.9

Для построения интервального ряда необходимо определить промежутки, на которые разбивается весь промежуток изменения длины листа: от наименьшего числа $A = 22.1$ до наибольшего числа $B = 62.8$. Обычно стараются выделить от пяти до двадцати промежутков одинаковой длины, в каждый из которых попадает не менее пяти значений случайной величины (ясно, что тогда общее количество обрабатываемых чисел должно быть не менее 35). Выделенные промежутки имеют вид:

$$[x_1,\ x_2),\ [x_2,\ x_3), \ldots, [x_{k-1},\ x_k).$$

Их левые и правые границы должны удовлетворять следующим требованиям:

1) длина промежутков (если они равны) может быть определена по формуле Стерджеса: $h = (B - A)/(1 + 3.322 \lg(n))$, где n — количество обрабатываемых чисел (в примере $n = 50$); длина может быть определена подбором так, чтобы абсолютные частоты были не менее пяти;

2) $x_1 \leqslant A$ и $A - x_1 \leqslant h$;

3) $x_k > B$ и $x_k - B \leqslant h$. Часто случается, что числа $x_1, \ldots, x_k$ удовлетворяют требованиям 1 – 3, но абсолютные частоты попадания в некоторые, особенно крайние, промежутки очень малы. Тогда по несколько крайних промежутков объединяют в один, длины больше, чем h.

Следует помнить, что допустимы разные способы разбиения на промежутки!

В качестве примера составим интервальный ряд для указанных выше 50 чисел. Выберем $x_1 = 22$. По формуле Стерджеса получается h приближенно равное шести, следовательно, промежутки будут иметь вид:

$$[22, 28),\ [28, 34),\ [34, 40),\ [40, 46),\ [46, 52),\ [52, 58),\ [58, 64).$$

Сопоставив им частоты, получим интервальный ряд:

Номер по порядку	Левая граница	Правая граница	Абсолютная частота	Относительная частота
1	22	28	1	0.02
2	28	34	10	0.2
3	34	40	6	0.12
4	40	46	11	0.22
5	46	52	9	0.18
6	52	58	9	0.18
7	58	64	4	0.08

В первом промежутке содержится всего одно число. Этот промежуток объединим со вторым. Абсолютные частоты сложим, относительные так же сложим. Получим окончательный интервальный ряд:

Номер по порядку	Левая граница	Правая граница	Абсолютная частота	Относительная частота
1	22	34	11	0.22
2	34	40	6	0.12
3	40	46	11	0.22
4	46	52	9	0.18
5	52	58	9	0.18
6	58	64	4	0.08

Б.3. Определение числовых характеристик случайной величины (точечные оценки)

Простейшей из числовых характеристик случайной величины является выборочное (синоним — статистическое) среднее.

Определение. Выборочным средним $\widetilde{M}$ наблюдавшихся значений случайной величины X называется частное от деления суммы всех этих значений на их число (то есть, среднее арифметическое значений):

$$\widetilde{M} = \frac{X_1 + X_2 + X_3 + \ldots + X_n}{n}. \tag{Б.1}$$

В формуле (Б.1) величина n — число опытов (измерений), $X_1, \ldots, X_n$ — результаты этих опытов. Числа $X_1, \ldots, X_n$ не обязательно различны. Если данные наблюдений (опытов) уже представлены в виде дискретного вариационного ряда, где $x_1, \ldots, x_k$ — варианты ($k \leqslant n$), а $a_1, \ldots, a_k$ — соответствующие им абсолютные частоты, то выборочное среднее может быть записано в виде:

$$\widetilde{M} = \frac{x_1a_1 + x_2a_2 + \ldots + x_ka_k}{n}.$$

Через относительные частоты $Q_1, \ldots, Q_k$ число $\widetilde{M}$ выражается по формуле

$$\widetilde{M} = x_1Q_1 + x_2Q_2 + \cdots + x_kQ_k. \tag{Б.2}$$

Для интервального вариационного ряда выборочное среднее вычисляется приближенно по формуле (Б.2), в которой теперь числа $x_1, \ldots, x_k$ — средние точки интервалов этого ряда, а числа $Q_1, \ldots, Q_k$ — относительные частоты попадания значений случайной величины в соответствующие промежутки.

Примеры

1) Если при пяти измерениях величина принимала значения: 13, 15, 18, 18, 15, то ее среднее статистическое значение равно

$$(13 + 15 + 18 + 18 + 15)/5 = 15.8$$

(см. формулу (Б.1)).

2) Для дискретного вариационного ряда

Номер	Значение	Абсолютная частота	Относительная частота
1	0	8	8/60
2	1	17	17/60
3	2	16	16/60
4	3	10	10/60
5	4	6	6/60
6	5	2	2/60
7	7	1	1/60

Выборочное среднее можно вычислить по формуле (Б.2) :

$$\widetilde{M} = 1\cdot 17/60+2\cdot 16/60+3\cdot 10/60+4\cdot 6/60+5\cdot 2/60+7\cdot 1/60 = 120/60 = 2.$$

3) Для интервального вариационного ряда

Номер по порядку	Левая граница	Правая граница	Абсолютная частота	Относительная частота
1	22	34	11	0.22
2	34	40	6	0.12
3	40	46	11	0.22
4	46	52	9	0.18
5	52	58	9	0.18
6	58	64	4	0.08

Определим сначала средние точки каждого промежутка:

$$x_1 = (34 + 22)/2 = 28; \quad x_2 = (40 + 34)/2 = 37; \quad x_3 = (46 + 40)/2 = 43;$$
$$x_4 = (52 + 46)/2 = 49; \quad x_5 = (58 + 52)/2 = 55; \quad x_6 = (64 + 58)/2 = 61.$$

Все числа, кроме первого, получаются прибавлением $h/2$ к левой границе интервала. Выборочное среднее вычислим по формуле (Б.2):

$$\widetilde{M} = 28 \cdot 0.22 + 37 \cdot 0.12 + 43 \cdot 0.22 + 49 \cdot 0.18 + 55 \cdot 0.18 + 61 \cdot 0.08 = 43.66\ .$$

Значение его равно среднему арифметическому наблюдавшихся значений случайной величины с некоторой погрешностью, поэтому, если есть возможность найти точное значение $\widetilde{M}$, не прибегая к нахождению середин промежутков ряда, то надо воспользоваться этой возможностью.

Напомним, что выборочное среднее является статистическим аналогом математического ожидания случайной величины.

Как известно из теории вероятностей, дисперсия случайной величины является мерой рассеивания ее значений вокруг математического ожидания. Статистическим аналогом дисперсии является выборочная дисперсия.

Определение. Выборочной (синоним — статистической) дисперсией значений случайной величины X называется среднее арифметическое квадратов отклонений наблюдаемых значений этой величины от их выборочного среднего:

$$\widetilde{D} = \frac{(X_1 - \widetilde{M})^2 + \ldots + (X_n - \widetilde{M})^2}{n},$$

где n — число опытов, $\widetilde{M}$ — выборочное среднее. Если данные наблюдений представлены в виде дискретного вариационного ряда, причем $x_1, \ldots, x_k$ — варианты, а $a_1, \ldots, a_k$ — соответствующие им абсолютные частоты, то число $\widetilde{D}$ можно также получить из формулы

$$\widetilde{D} = \frac{a_1(x_1 - \widetilde{M})^2 + \ldots + a_k(x_k - \widetilde{M})^2}{n}.$$

С помощью относительных частот дисперсию можно вычислить следующим образом:

$$\widetilde{D} = Q_1(x_1 - \widetilde{M})^2 + \ldots + Q_k \cdot (x_k - \widetilde{M})^2. \tag{Б.3}$$

Для интервального вариационного ряда может использоваться формула (Б.3), где в качестве чисел $x_1, \ldots, x_k$ выбраны середины промежутков.

Важное свойство выборочной дисперсии: так же, как и дисперсия случайной величины, выборочная дисперсия может вычисляться по формуле:

$$\widetilde{D} = \widetilde{M}(X^2) - \widetilde{M}(X)^2,$$

где $\widetilde{M}(X^2)$ — среднее арифметическое (то есть, выборочное среднее) квадратов наблюдавшихся значений случайной величины, $\widetilde{M}(X)$ — среднее арифметическое самих значений этой величины.

Примеры.

1) Если при пяти измерениях величина принимала значения: 13, 15, 18, 18, 15, то ее выборочное среднее равно $(13+15+18+18+15)/5 = 15.8$, а выборочная дисперсия равна $((13-15.8)^2+2(15-15.8)^2+2(18-15.8)^2)/5 = 3.76$.

2) Для дискретного вариационного ряда

Номер	Значение	Абсолютная частота	Относительная частота
1	0	8	8/60
2	1	17	17/60
3	2	16	16/60
4	3	10	10/60
5	4	6	6/60
6	5	2	2/60
7	7	1	1/60

Выборочное среднее равно

$$\widetilde{M} = 1\cdot 17/60+2\cdot 16/60+3\cdot 10/60+4\cdot 6/60+5\cdot 2/60+7\cdot 1/60 = 120/60 = 2\ .$$

Выборочная дисперсия может быть получена из формулы (Б.3):

$$\widetilde{D} = (2-0)^2\cdot 8/60 + (2-1)^2\cdot 17/60 + (2-2)^2\cdot 16/60 + (2-3)^2\cdot 10/60+ \\ + (2-4)^2\cdot 6/60 + (2-5)^2\cdot 2/60 + (2-7)^2\cdot 1/60 = 2.1.$$

3) Для интервального вариационного ряда

Номер по порядку	Левая граница	Правая граница	Абсолютная частота	Относительная частота
1	22	34	11	0.22
2	34	40	6	0.12
3	40	46	11	0.22
4	46	52	9	0.18
5	52	58	9	0.18
6	58	64	4	0.08

Определим сначала средние точки каждого промежутка:

$$x_1 = (34+22)/2 = 28; \quad x_2 = (40+34)/2 = 37; \quad x_3 = (46+40)/2 = 43;$$
$$x_4 = (52+46)/2 = 49; \quad x_5 = (58+52)/2 = 55; \quad x_6 = (64+58)/2 = 61.$$

Все числа, кроме, первого получаются прибавлением $h/2$ к левой границе интервала. Выборочное среднее равно

$$\widetilde{M} = 28 \cdot 0.22 + 37 \cdot 0.12 + 43 \cdot 0.22 + 49 \cdot 0.18 + 55 \cdot 0.18 + 61 \cdot 0.08 = 43.66\,.$$

Выборочная дисперсия, вычисленная по формуле (Б.3), равна

$$\begin{aligned}\widetilde{D} &= (43.66-28)^2 \cdot 0.22 + (43.66-37)^2 \cdot 0.12 + (43.66-43)^2 \cdot 0.22 + \\ &+ (43.66-49)^2 \cdot 0.18 + (43.66-55)^2 \cdot 0.18 + (43.66-61)^2 \cdot 0.08 = \\ &= 111.7.\end{aligned}$$

Выборочная дисперсия обладает одним существенным недостатком: если среднее арифметическое выражается в тех же единицах, что и значения случайной величины, то, как следует из формул, задающих дисперсию, последняя выражается уже в квадратных единицах. Этого недостатка можно избежать, взяв в качестве меры рассеивания арифметический квадратный корень из дисперсии.

Определение. Выборочным средним квадратическим отклонением называется арифметический квадратный корень из выборочной дисперсии.

Б.4. Построение полигона или гистограммы

Дискретному вариационному ряду, составленному по результатам эксперимента, можно дать наглядную геометрическую интерпретацию. Для этого в прямоугольной системе координат строят точки с координатами:

$$(x_1, Q_1),\ (x_2, Q_2), \ldots, (x_i, Q_i), \ldots, (x_k, Q_k),$$

где x_i — значение i-го варианта, а Q_i — соответствующая относительная частота. Отмеченные точки последовательно (слева направо) соединяют отрезками прямой линии. Полученная ломаная называется «полигон (синоним — многоугольник) относительных частот». Указанные выше точки называются вершинами полигона.

Заметим, что сумма ординат вершин полигона должна быть равна 1.

Пример.

Рассмотрим дискретный вариационный ряд, заданный таблицей:

Номер	Значение	Абсолютная частота	Относительная частота
1	0	8	$8/60 = 0.1(3)$
2	1	17	$17/60 = 0.28(3)$
3	2	16	$16/60 = 0.2(6)$
4	3	10	$10/60 = 0.1(6)$
5	4	6	$6/60 = 0.1$
6	5	2	$2/60 = 0.0(3)$
7	7	1	$1/60 = 0.01(6)$

Координаты вершин будущего полигона: (0, 0.133), (1, 0.283), (2, 0.267), (3, 0.167), (4, 0.1), (5, 0.033), (7, 0.017). Относительные частоты здесь округлены, но их сумма равна единице. На рис. Б.1 изображен график полигона относительных частот, построенный по данному вариационному ряду.

Для графического изображения интервального вариационного ряда служит специальная столбчатая диаграмма, которая называется «гистограмма». Для ее построения в прямоугольной системе координат на оси абсцисс откладывают отрезки, изображающие промежутки интервального вариационного ряда, и на этих отрезках, как на основаниях, строят прямоугольники. Площадь каждого прямоугольника должна равняться относительной частоте попадания наблюдавшихся значений случайной величины в соответствующий промежуток, лежащий в основании этого прямоугольника. Исходя из этого условия можно легко вычислить высоту прямоугольника: она равна частному от деления относительной частоты на длину промежутка варьирования.

Очевидно, что сумма площадей всех построенных таким образом прямоугольников равна единице, поскольку она равна сумме всех относительных частот интервального ряда. Гистограмма является статистическим аналогом графика функции плотности распределения.

В качестве примера построим гистограмму для интервального ряда:

Номер по порядку	Левая граница	Правая граница	Абсолютная частота	Относительная частота
1	22	34	11	0.22
2	34	40	6	0.12
3	40	46	11	0.22
4	46	52	9	0.18
5	52	58	9	0.18
6	58	64	4	0.08

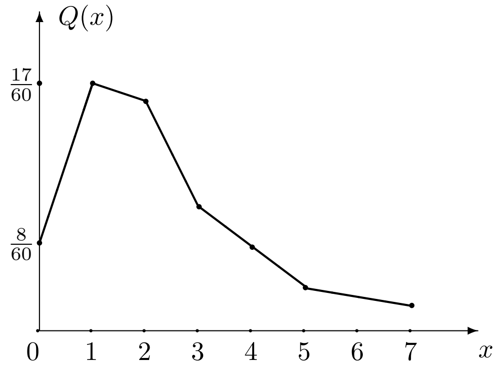


Рис. Б.1

Произведем подготовительные вычисления. Их результаты внесем в таблицу:

Номер по порядку	Левая граница	Правая граница	Длина основания прямоуг.	Высота прямоугольника
1	22	34	$34 - 22 = 12$	$0.22/12 = 0.018(3)$
2	34	40	$40 - 34 = 6$	$0.12/6 = 0.02$
3	40	46	$46 - 40 = 6$	$0.22/6 = 0.03(6)$
4	46	52	$52 - 46 = 6$	$0.18/6 = 0.03$
5	52	58	$58 - 52 = 6$	$0.18/6 = 0.03$
6	58	64	$64 - 58 = 6$	$0.08/6 = 0.01(3)$

Высоты прямоугольников округлим до трех знаков после десятичной точки. Все необходимые сведения о гистограмме поместим в таблицу:

Номер столбца	Левая граница основания	Правая граница основания	Высота прямоугольника
1	22	34	0.018
2	34	40	0.02
3	40	46	0.037
4	46	52	0.03
5	52	58	0.03
6	58	64	0.013

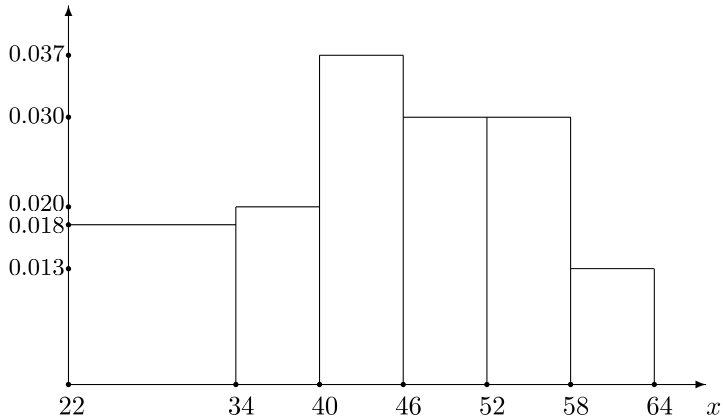


Рис. Б.2

Гистограмма, построенная по данному вариационному ряду, изображена на рис. Б.2.

Б.5. Построение графика выборочной функции распределения

Как известно, закон распределения случайной величины можно задать с помощью функции распределения: $F(x) = \mathsf{P}(X < x)$. Введем выборочный аналог функции $F(x)$.

Определение. Выборочной функцией распределения называется функция $\widetilde{F}(x)$, задающая для каждого значения x относительную частоту попадания наблюдавшихся значений случайной величины в открытый интервал $(-\infty, x)$ (т. е. события $X < x$).

Свойство статистической устойчивости частоты, обоснованное теоремой Бернулли, оправдывает целесообразность использования выборочной функции распределения для оценивания неизвестной функции распределения вероятностей между значениями изучаемой случайной величины.

Различаются два случая определения значений выборочной функции распределения: дискретный и интервальный — соответствующие представлению опытных данных в виде дискретного или интервального вариационного ряда.

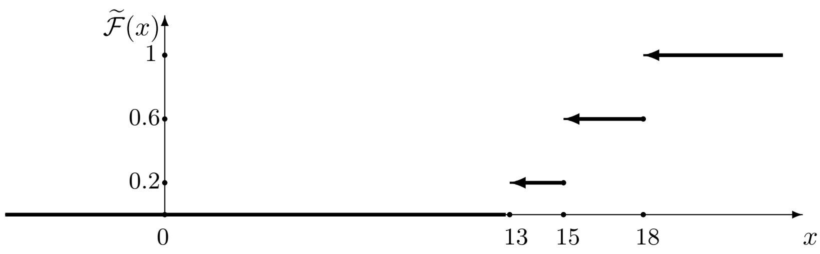


Рис. Б.3

Дискретная выборочная функция распределения Пусть результаты экспериментов записаны в виде дискретного вариационного ряда. Тогда относительная частота события $X < x$ (см. определение) равна сумме относительных частот, соответствующих вариантам, меньшим, чем число x. Она может быть легко подсчитана.

В качестве примера рассмотрим дискретный ряд:

Номер по порядку	Значение варианта	Абсолютная частота	Относительная частота
1	13	1	0.2
2	15	2	0.4
3	18	2	0.4

Определенная по этому ряду выборочная функция распределения имеет вид:

$$\widetilde{F}(x) = \begin{cases} 0, & \text{для} \quad x \leqslant 13, \\ 0.2, & \text{для} \quad 13 < x \leqslant 15, \\ 0.2 + 0.4 = 0.6, & \text{для} \quad 15 < x \leqslant 18, \\ 1, & \text{для} \quad x > 18. \end{cases}$$

График рассматриваемой функции $\widetilde{F}(x)$ полностью определяется координатами ее точек разрыва: (13, 0), (15, 0.2), (18, 0.6). Действительно, зная эти координаты, можно восстановить весь график (рис. Б.3). Для этого надо проделать следующие построения:

1) из самой левой точки разрыва влево по оси OX построить луч ($\widetilde{F} = 0$);

2) из второй точки разрыва провести влево отрезок, параллельный оси OX до пересечения с вертикалью $x = 13$;

3) из третьей точки разрыва провести влево отрезок, параллельный оси ОХ до пересечения с вертикалью $x = 15$;

4) из самой правой точки разрыва вправо провести луч, параллельный ОХ, на котором $\widetilde{F} = 1$.

Здесь мы пользуемся тем, что функция $\widetilde{F}$ кусочно-постоянна.

Интервальная выборочная функция распределения. Если результаты наблюдений представлены в виде интервального вариационного ряда, то бывает невозможно точно установить каждое отдельное наблюдавшееся значение случайной величины и его частоту, поскольку по ходу опытов фиксировалось лишь попадание значений в определенные промежутки. В таком случае применяется несколько иной, чем выше, принцип построения выборочной функции распределения.

Пусть дан интервальный ряд:

Номер по порядку	Левая граница	Правая граница	Абсолютная частота	Относительная частота
1	22	34	11	0.22
2	34	40	6	0.12
3	40	46	11	0.22
4	46	52	9	0.18
5	52	58	9	0.18
6	58	64	4	0.08

Используя приведенную в нем информацию о наблюдавшихся значениях случайной величины, определим значения выборочной функции распределения в тех точках, где это возможно:

1) $\widetilde{F}(x) = 0$ при всех $x \leqslant 22$;

2) $\widetilde{F}(34) = 0.22$;

3) $\widetilde{F}(40) = 0.22 + 0.12 = 0.34$;

4) $\widetilde{F}(46) = 0.22 + 0.12 + 0.22 = 0.56$;

5) $\widetilde{F}(52) = 0.56 + 0.18 = 0.74$;

6) $\widetilde{F}(58) = 0.74 + 0.18 = 0.92$;

7) $\widetilde{F}(64) = 1$;

8) $\widetilde{F}(x) = 1$ при $x > 64$.

Перечисленные значения определяют $\widetilde{F}(x)$ не полностью. Действительно, в промежутках, указанных в интервальном ряду, наблюдалось по нескольку разных значений случайной величины, поэтому значение $\widetilde{F}(x)$ не может оставаться постоянным внутри этих промежутков (сравните с дискретным случаем). Если подопытная случайная величина может, теоретически, принимать любые значения от некоторого числа A

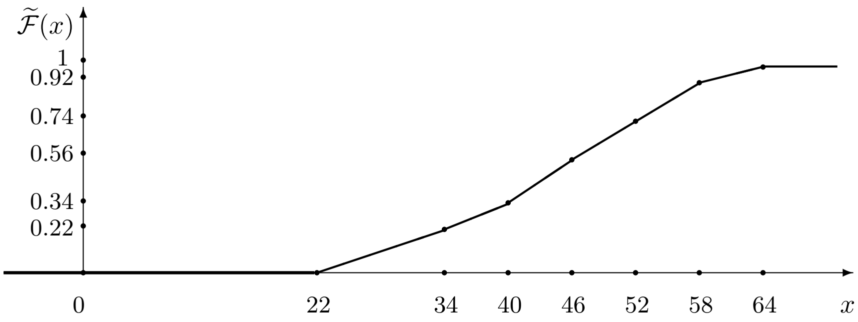


Рис. Б.4

(меньше 22) до B (больше 64), причем вероятность попадания значений в промежуток непрерывно зависит от длины промежутка, то функция распределения такой величины должна быть непрерывна. Интервальная же выборочная функция распределения должна служить статистическим аналогом непрерывной функции распределения, поэтому за ее график примем ломаную линию, соединяющую точки с известными координатами, перечисленными в пунктах 1) – 8) (см. рис. Б.4).

Б.6. Задания для лабораторных работ

1. Изучение дискретной случайной величины

Для наблюдавшихся 50 значений дискретной случайной величины постройте дискретный вариационный ряд, график выборочной функции распределения. Подсчитайте выборочное среднее, дисперсию и среднее квадратическое отклонение. Попробуйте выдвинуть гипотезу о законе распределения случайной величины. Сравните относительные частоты наблюдавшихся значений с гипотетическими вероятностями.

Варианты

1. Число земляничных карамелек в пакетике конфет «Бон Пари — фруктовая смесь»:

4	11	8	11	12	11	13	15	3	11
11	7	14	6	6	9	15	5	2	5
6	14	13	6	13	8	9	5	14	8
9	13	11	6	9	11	14	14	4	2
4	10	13	9	7	10	4	12	6	15

2. Число жевательных резинок, которые выдает автомат на пять рублей:

9	7	8	9	12	9	7	10	9	11
8	9	10	8	9	10	7	9	10	8
10	8	9	6	9	9	11	7	6	6
11	5	8	8	10	6	9	9	8	9
9	6	5	10	8	8	8	10	6	6

(На автомате написано, что он выдает от пяти до двенадцати резинок.)

3. Биолог решил узнать, сколько рыб водится в пруду. Для этого он поймал десять рыб, пометил их и выпустил обратно в пруд. Затем он произвел 50 контрольных отловов по десять рыб и подсчитал, сколько отмеченных попалось в каждом контрольном отлове:

3	1	5	4	3	4	2	3	2	1
3	2	2	2	3	2	4	0	0	3
5	1	1	3	4	4	3	3	3	3
3	0	1	1	4	3	2	3	2	2
2	1	4	3	1	4	3	5	2	1

(После каждого отлова рыбу отпускали в пруд.) Как Вы думаете, сколько рыб в пруду?

4. Из полка новобранцев выбрали 50 человек и проверили их умение стрелять из пистолета. Каждый должен был стрелять по мишени до первого попадания. Инструктор подсчитал число промахов для каждого:

1	9	0	2	9	0	3	1	4	1
0	0	0	4	3	1	3	5	4	4
12	2	0	5	1	1	11	0	7	0
1	1	2	12	0	0	1	0	4	1
1	1	8	3	8	4	0	0	1	11

5. 50 студентов сдавали тест по дискретным случайным величинам (до первого успеха). Преподаватель записал число попыток у каждого студента:

1	7	1	3	3	4	1	4	4	2
4	1	2	1	7	2	2	1	1	1
1	2	1	1	1	2	4	1	4	1
3	6	2	1	3	2	3	1	1	3
1	1	1	2	5	1	4	6	1	1

6. Кот Вася жил на даче 50 дней. Каждое утро он ходил на охоту за мышами и приносил всю свою добычу на крыльцо, чтобы показать хозяйке. Она записала, по сколько мышек он приносил:

2	1	2	2	1	1	2	2	2	0
2	4	4	0	2	4	1	3	3	3
3	0	2	2	1	3	1	2	0	3
6	2	2	0	1	1	2	1	1	1
1	1	1	0	1	2	2	0	0	3

7. Для каждого из 50 сухарей с орехами подсчитали, сколько в него попало изюминок:

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

8. В коробке лежат десять шариков. Некоторые из них — черные, остальные — белые. Ровно 50 раз проделали один и тот же опыт: достали три шарика, подсчитали количество вынутых черных шариков, шарики положили обратно. Черных шариков было:

0	1	2	2	2	2	1	1	1	0
2	2	2	2	0	2	1	2	1	2
1	1	1	2	1	1	2	1	1	0
1	1	1	1	1	0	1	0	0	1
1	1	1	1	1	2	0	0	1	1

Сколько черных шариков в коробке?

9. У 50 людей племени «Ньяму-Ньяму», достигших совершеннолетия, подсчитали количество зубов:

6	15	10	14	11	12	10	13	11	13
14	19	16	13	14	11	11	14	11	19
12	15	10	14	14	13	16	13	13	11
9	12	12	14	14	13	16	17	16	13
12	12	13	14	10	13	12	15	17	16

10. Воины племени «Дили-Били» носят на шее ожерелья из зубов. Чем больше зубов в ожерелье, тем знатнее воин. У 50 случайно выбранных воинов этнограф подсчитал число зубов в ожерелье:

15	23	19	20	18	14	15	20	19	20
18	15	19	20	23	17	19	14	20	22
19	19	16	19	16	25	19	22	14	16
22	15	14	18	23	14	21	19	17	19
23	18	20	19	23	12	18	17	15	22

2. Изучение непрерывной случайной величины

Для наблюдавшихся значений непрерывной случайной величины постройте интервальный вариационный ряд, гистограмму и график выборочной функции распределения. Подсчитайте выборочное среднее, дисперсию и среднее квадратическое отклонение. По виду гистограммы попробуйте выдвинуть гипотезу о законе распределения вероятности. Подсчитайте гипотетические вероятности для каждого интервала и сравните с относительными частотами.

Варианты

1. Кофейный автомат наливает порцию кофе (в граммах):

95	97	100	104	103	99	103	101	102	95
103	102	98	104	99	101	104	95	100	104
103	95	105	103	98	102	105	96	97	103
108	100	91	100	99	107	95	105	97	96
103	100	102	103	93	98	101	99	102	99

2. Время ожидания справки от банкомата в городе Faraway (в минутах):

14	21	14	20	21	14	24	27	18	16
9	20	15	14	23	13	21	26	16	19
25	15	27	13	14	25	15	17	22	23
17	15	20	19	18	20	27	14	27	16
22	18	14	25	19	17	21	20	17	14

3. Время, за которое поезд проходит весь маршрут «Москва — Петушки» (в часах):

7.5	5.8	5.2	6.0	4.5	4.2	7.6	6.5	6.0	8.7
6.3	4.1	4.0	5.6	6.4	4.5	5.6	6.7	5.1	6.5
6.4	3.5	7.6	6.6	4.1	5.4	3.7	7.9	9.2	4.3
7.8	4.9	4.9	5.6	5.5	6.3	6.0	5.2	5.6	6.9
5.6	6.4	6.0	3.5	4.6	5.8	6.1	6.4	3.9	7.9

4. Время (в месяцах), которое прослужил каждый из 50 новых компьютеров (с гарантией два года) до первой поломки:

13.5	0.4	11.8	23.1	10.4	1.6	23.4	26.9	78.2	15.5
34.0	38.7	5.8	45.8	8.7	16.7	4.7	19.6	6.9	9.2
90.4	39.6	8.4	11.2	28.9	10.8	63.1	3.4	8.5	21.4
1.7	4.4	5.7	72.7	104.4	12.0	1.3	18.7	11.6	10.2
56.3	18.4	1.2	97.5	11.5	16.0	99.5	87.3	13.5	5.0

5. Задержка отправки автобуса №468 (в минутах) по сравнению с расписанием:

7.6	12.2	0.9	0.8	9.3	5.5	16.5	12.3	18.0	19.9
5.5	17.1	16.7	2.8	17.9	18.3	16.5	1.4	0.1	6.4
6.6	1.2	16.9	15.2	10.5	15.1	14.2	5.4	18.0	3.6
7.6	1.8	19.6	4.9	15.6	11.8	13.1	2.2	17.5	2.9
4.3	5.0	7.8	0.8	6.9	7.3	8.7	17.5	19.6	13.4

6. Диаметр мыльного пузыря (в миллиметрах):

96	57	133	66	86	90	96	47	165	155
115	111	102	198	144	95	54	148	130	152
85	156	169	119	101	106	142	69	106	73
85	74	23	80	137	64	49	22	61	136
83	77	53	101	132	89	68	137	66	110

7. Время, которое провели у телевизора 50 школьников в течение одного выходного дня (в часах):

5.3	10.0	7.8	7.6	7.1	3.5	4.2	1.9	5.9	7.0
4.3	5.9	5.6	5.0	8.1	5.4	6.0	6.2	6.3	6.3
6.6	11.3	5.0	7.5	3.9	5.7	2.6	9.0	7.8	6.6
2.4	5.5	8.9	7.7	6.7	3.8	8.5	6.1	6.4	4.0
2.9	6.3	4.3	4.3	5.9	2.1	4.8	6.3	7.5	4.1

8. Приращение веса 50 студентов после рождественских каникул (в килограммах):

-1.0	-0.4	-0.4	2.1	5.2	5.5	-1.4	1.4	1.8	-0.2
-0.5	3.8	1.3	2.5	3.2	-0.5	4.3	-0.1	1.4	3.8
3.7	0.6	3.9	0.4	0.3	0.1	0.0	5.0	2.6	5.3
2.3	2.4	-1.5	3.8	-0.7	0.9	5.2	5.5	1.4	0.8
1.1	-0.2	-0.6	1.3	1.5	3.9	-1.8	3.5	3.1	-1.5

9. Время ответа на экзамене по математике у 50 студентов (в минутах):

17.8	17.7	14.6	12.6	14.9	19.6	14.5	15.8	14.1	18.4
11.8	12.0	10.7	14.0	19.8	10.2	13.0	18.2	19.6	19.3
16.3	10.5	13.3	18.3	11.8	16.6	15.0	14.3	15.7	12.7
13.9	17.3	16.6	15.7	14.4	13.4	18.3	18.2	18.7	11.8
11.1	18.4	11.0	15.5	13.5	19.0	11.2	13.6	18.7	12.2

10. Приращение веса 50 студентов после зимней сессии (в килограммах):

-0.1	1.2	-0.6	-0.9	-4.3	-0.1	-0.1	2.5	-2.7	-1.6
0.1	-1.0	-3.0	0.6	-0.1	0.3	-4.2	0.4	-0.3	0.3
-2.4	-4.2	-1.9	0.8	-2.5	-0.2	-2.4	-2.5	-0.8	-3.5
0.7	-3.5	1.8	-0.7	0.1	-3.0	-3.2	-0.5	-3.3	-2.2
-5.5	-0.6	-1.9	-4.2	-3.8	-1.7	1.0	-1.4	-2.4	-2.0

Литература

1. *Агапов Г.И.* Задачник по теории вероятностей. М.: Наука, 1994.
2. *Вентцель Е.С., Овчаров Л.А.* Теория вероятностей и ее инженерные приложения. М.: Наука, 1988.
3. *Вентцель Е.С., Овчаров Л.А.* Задачи и упражнения по теории вероятностей. М.: ACADEMIA, 2003.
4. *Гмурман В.Е.* Теория вероятностей и математическая статистика. М.: Высшая школа, 2003.
5. *Гмурман В.Е.* Руководство к решению задач по теории вероятностей и математической статистике. М.: Высшая школа, 2003.
6. *Кемени Дж., Снелл Дж., Томпсон Дж.* Введение в конечную математику. М.: Изд-во иностранной литературы, 1963.
7. *Колмогоров А.Н.* Основные понятия теории вероятностей. М.: Наука, 1974.
8. *Максимов Ю.Д., Куклин Б.А., Хватов Ю.А.* Теория вероятностей. Контрольные задания с образцами решений. СПб.: Изд-во СПбГТУ, 2000.
9. *Поволоцкий А.И., Чурилова М.Ю.* Теория вероятностей. СПб.: Образование, 1996.
10. *Скворцов В.В.* Теория вероятностей? — Это интересно! М.: Мир, 1993.
11. *Солодовников А.С.* Теория вероятностей. М.: Просвещение, 1978.
12. *Тихомиров С.Р.* Расчетные задания по теории вероятностей. СПб.: Нестор, 1999.