

Министерство образования и науки Российской Федерации



Уральский государственный экономический университет

Н. В. Коржавина, Н. П. Петров, С. Н. Петрова

**ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА
ДЛЯ ФИНАНСИСТОВ И МЕНЕДЖЕРОВ**

Рекомендовано
Учебно-методическим советом
Уральского государственного экономического университета
в качестве учебного пособия

Екатеринбург
2014

УДК 519.2(07)

ББК 22.172я7

К 66

Рецензенты:

кафедра теории и методики обучения физике,
технологии и мультимедийной дидактики
Уральского государственного педагогического университета
(протокол № 7 от 10 марта 2014 г.)

кандидат педагогических наук, доцент кафедры
информационного-коммуникационных технологий в образовании
Уральского государственного педагогического университета
Е. Б. Стариченко

Коржавина, Н. В.

К 66 Теория вероятностей и математическая статистика для финансистов и менеджеров [Текст] : учеб. пособие / Н. В. Коржавина, Н. П. Петров, С. Н. Петрова ; М-во образования и науки Рос. Федерации, Урал. гос. экон. ун-т. – Екатеринбург : [Изд-во Урал. гос. экон. ун-та], 2014. – 42 с.

В учебном пособии подробно рассмотрены основные понятия теории вероятностей и математической статистики. Пособие составлено в соответствии с требованиями ФГОС по направлениям 080100.62 «Экономика» и «080200.62 «Менеджмент. Содержит структурированный теоретический материал, примеры решения типовых задач, задания для самостоятельного решения, вопросы для проверки усвоения теоретического материала, список рекомендуемой литературы.

Предназначено для студентов бакалавриата дневной и заочной форм обучения, в том числе с использованием дистанционных технологий.

УДК 519.2(07)

ББК 22.172я7

© Коржавина Н. В., Петров Н. П.,
Петрова С. Н., 2014

© Уральский государственный
экономический университет,
2014

ОГЛАВЛЕНИЕ

Введение.....	4
1. Случайное событие, его частота и вероятность.....	5
2. Основные формулы комбинаторики.....	9
3. Алгебра событий.....	12
4. Теоремы сложения и умножения вероятностей.....	19
5. Формула полной вероятности.....	23
6. Повторение испытаний. Схема Бернулли.....	26
7. Дисперсия и среднеквадратическое отклонение.....	28
8. Нормальное распределение.....	31
9. Центральная предельная теорема.....	37
Библиографический список.....	42

ВВЕДЕНИЕ

Современная теория вероятностей, подобно другим разделам математики, состоит из результатов, выводимых логическим путем из некоторых основных утверждений, или аксиом, и приложений к ситуациям в реальной жизни, относительно которых предполагается, что они согласуются с аксиомами. Трудность теории вероятностей заключается в том, что объекты, составляющие предмет ее изучения, носят гораздо более общий характер и поэтому не столь наглядны, как, например, объекты геометрии или механики. Теория вероятностей занимается изучением событий и их вероятностей, представляемых числами, заключенными в интервале от 0 до 1. В случае исторически знаменитых задач, связанных с азартными играми, можно интуитивно понять, как должна быть сформулирована соответствующая математическая задача. Такая задача обычно имеет следующий вид: заданы вероятности некоторых элементарных событий; требуется вычислить вероятность какого-нибудь более сложного события, связанного с элементарными событиями некоторым простым образом.

Не стоит думать, что там, где речь идет о случайных событиях, бесполезно искать какие-то закономерности – случай он и есть случай. Существует несколько групп случайных явлений, в которых закономерности уже обнаружены и изучены, оценивать и сравнивать прогноз развития событий в этом случае можно и нужно. Само понятие «вероятность» нередко определяют как количественную меру возможности реализации интересующего нас случайного события. Правда, знание вероятности благоприятного исхода – это еще не выигрыш сам по себе, а лишь взвешивание возможностей.

Мы ежедневно принимаем многие решения в условиях неопределенности. Принято различать неопределенность и риск. Риск – это когда можно сказать, что человек знает, на что он идет, шансы известны, вероятности оценены. Конечно, не всякую неопределенность можно превратить в риск. Но там, где это несложно сделать, это может оказать реальную помощь в принятии решения.

Этот курс предназначен для менеджеров и экономистов и содержит базовые знания в области теории вероятностей и математической статистики в качестве основы для принятия решения в условиях неопределенности. Учебное пособие построено на основе ФГОС по дисциплине «Теория вероятностей и математическая статистика» и имеет выраженную практическую направленность.

1. СЛУЧАЙНОЕ СОБЫТИЕ, ЕГО ЧАСТОТА И ВЕРОЯТНОСТЬ

Исторически первой группой случайных событий, которые были исследованы математиками, были азартные игры. Сами правила игры предполагали равноправие участников перед судьбой, при всем мастерстве игрока многое зависело от «везения». Рассмотрим несколько примеров из этой области, так как они хорошо иллюстрируют некоторые возможности аппарата теории вероятностей, который пригодится в дальнейшем для других случаев. В условиях игры (в карты, в кости) математики и игроки связывают понятие вероятности выигрыша с шансом получения выигрышной комбинации по сравнению с общим числом всех возможных комбинаций.

Введем основные понятия. *Случайными* называют такие события, которые могут произойти или не произойти при осуществлении совокупности условий, связанных с возможностью появления данных событий. Случайные события обозначают буквами $A, B, C, \dots$. Каждое осуществление рассматриваемой совокупности условий называется *опытом*.

Численная мера степени объективной возможности наступления случайного события называется *вероятностью*. Вероятность случайного события A обозначается через $P(A)$.

Достоверному событию, т.е. событию, которое должно произойти при каждом испытании, приписывается вероятность $P(A)=1$.

Невозможному событию, т.е. событию, которое не может произойти ни при одном испытании, приписывается вероятность $P(A)=0$.

Вероятность случайного события заключена между нулем и единицей: $0 \leq P(A) \leq 1$. В некоторых простейших случаях вероятность случайного события может быть определена заранее (как говорят в математике, *a priori*).

Все мыслимые взаимоисключающие результаты опыта называются элементарными событиями (элементарными исходами). Несколько событий в данном опыте называются *равновозможными*, если появление какого-либо из них не более возможно, чем любого другого.

Возможные результаты каждого из однородных событий могут быть представлены в виде n единственно возможных, несовместных друг с другом и равновозможных исходов, т.е. кроме этих n исходов не может быть никаких других, никакие два из них не могут произойти одновременно и есть основания считать, что любой из них не является более возможным, чем другие. Если из этих n единственно возможных, несовместных и равновозможных случаев m случаев связаны с наступлением события A (или,

как говорят в теории вероятностей, «благоприятствуют» A), то за вероятность события A принимается отношение m к n :

$$P(A) = \frac{m}{n}.$$

Пример 1.1. При падении монеты существует два возможных результата: выпадет герб или выпадет решетка. Оба результата равновероятны, т.е. вероятность того, что монета останется лежать гербом вверх, равна 50% (или $1/2$), с такой же вероятностью выпадет другая сторона. Какой бы результат ни загадал игрок как благоприятный для себя, его шанс выиграть и вероятность проиграть одинаковы.

Пример 1.2. При бросании игральной кости существует уже шесть возможных элементарных исходов (количество выпавших очков может меняться от 1 до 6). Если игральная кость имеет правильную форму, все шесть результатов равновероятны. Другими словами, вероятность того, что при единственном выбрасывании кости выпадет, например, шесть очков, равна $1/6$. Если только эта цифра считается выигрышем при данном броске, шансов выиграть у игрока в три раза меньше, чем в прошлый раз. Если мы хотим «уравнять» шансы при бросании игральной кости с шансом выиграть при бросании монеты, нам надо изменить правила игры, например, считать выигрышем выпадение любого четного числа. Поскольку игральная кость имеет три грани с четными числами и три грани с нечетными, шансы выиграть и проиграть при единственном броске у нас будут одинаковыми (вероятность выигрыша станет равной $1/2$, т.е. такой же, как при бросании монеты).

Задание 1.1

При подбрасывании монеты вероятность выиграть составляет $1/2$, а при бросании игральной кости выигрышем считается выпадение цифры шесть (вероятность выигрыша $1/6$).

Чему равна вероятность проигрыша в каждом случае?

Значит ли это, что играть в кости менее выгодно?

Обсуждение

В каждом из двух случаев сумма вероятностей выиграть и проиграть должна составить 100%, или единицу, поскольку ничейный вариант в этих ситуациях невозможен.

Это означает, что при бросании монеты вероятность проиграть равна $1/2$, а при бросании игральной кости – $5/6$.

А вот вопрос о «выгодах» предложения поиграть в кости *по сравнению* с предложением бросить монету не так прост, как это кажется. Оставив на минуту в стороне азартные игры, обсудим одну важную для бизнеса

проблему. Решение о «выгодности» любого предпринимаемого нами действия, очевидно, зависит не только от нашей оценки риска данного предприятия, но и от величины предполагаемого выигрыша по сравнению с нашими ставками.

Чем меньше шансов получить выигрыш, тем больше должна быть величина этого выигрыша по сравнению со ставкой, чтобы сделать игру привлекательной для потенциальных игроков. Забота о привлекательности условий игры, конечно, распространяется только на те случаи, когда игроки принимают решение об участии в процессе добровольно и осмысленно. Так, чем рискованнее финансовые вложения, тем большую прибыль мы ожидаем получить в результате. Когда соотношение «риск/прибыльность» кажется нам неподходящим, мы ищем возможности покинуть «игру». Поэтому при любой оценке бизнес-проекта оценка рисков не менее важна, чем оценка прибыльности, по сути, это неотъемлемая часть финансово-экономического анализа.

Возвращаясь к нашему заданию, пришло время обсудить финансовые условия игры. Какой именно выигрыш покажется нам справедливым и почему? Если при бросании монеты участвуют два игрока, сделавшие одинаковые ставки, причем выигравший забирает все, то возможный выигрыш в такой игре должен вдвое превышать исходную ставку. Менее очевидный случай – бросание кости. Должен ли выигрыш в шесть раз превышать ставку игрока, и откуда возьмется эта сумма, если игроков по-прежнему только двое? Вот если бы игроков было шестеро, и каждый поставил бы на разную цифру, то при одинаковых исходных ставках получилась бы вполне справедливая игра. Выигравший забрал бы в шесть раз больше, чем поставил, но шансы каждого игрока выиграть были бы одинаковыми.

Если же играют двое, причем один выигрывает только при выпадении цифры «шесть», значит, второй выигрывает при любой другой ситуации, и его шансы на выигрыш в пять раз выше. Само по себе это не означает, что игра «нечестная», просто справедливые правила должны потребовать от второго игрока сделать исходную ставку, которая будет в пять раз выше, чем ставка первого игрока.

Определение 1.1. Совокупность всех возможных результатов опыта в теории вероятностей называется *пространством элементарных исходов*, будем обозначать это пространство греческой буквой W . Элементарные исходы обозначаются как w_i , где i может принимать значения от одного до максимума по числу возможных вариантов результата опыта. Для наглядности W изображают в виде некоторой области на плоскости, а элементарные исходы w_i – точками в этой области. Мы будем пользоваться матема-

тическим обозначением $W=\{w_i, i=1, \dots\}$ для описания того факта, что пространство элементарных исходов W образуется совокупностью всех элементарных исходов w_i .

Комбинаторика – это раздел математики, в котором изучаются вопросы о том, сколько различных комбинаций, подчиненных тем или иным условиям, можно составить из заданных объектов. Основы комбинаторики очень важны для оценки вероятностей случайных событий, т.к. именно они позволяют подсчитать принципиально возможное количество различных вариантов развития событий.

Определение 1.2. Элементарные исходы могут образовывать группы, каждая из которых называется *событием*. Событие A , принадлежащее пространству W , наступает тогда и только тогда, когда наступает один из элементарных исходов w_i , входящих в A .

Пример 1.3. В примере 1.2 событием можно считать факт выпадения четной цифры при бросании кости. Это событие наступает, когда выпадает или цифра два, или четыре, или шесть (при трех элементарных исходах из шести возможных). Мы будем пользоваться математическим обозначением $A=\{w_i, i=1, \dots\}$ для описания того факта, что событие образуется некоторой группой элементарных исходов (напомним, что событие содержится внутри пространства элементарных исходов, является его подмножеством).

В данном примере пространство элементарных исходов W состоит из следующих элементарных исходов:

$$w_1=\{1\}, w_2=\{2\}, w_3=\{3\}, w_4=\{4\}, w_5=\{5\}, w_6=\{6\}.$$

Событие A (выпадение четной цифры) можно записать как $A=\{w_2, w_4, w_6\}$. Иногда говорят, что элементарные исходы w_2, w_4 и w_6 *благоприятны* для наступления события A , в то время как w_1, w_3 и w_5 , напротив, *неблагоприятны* для него.

Сортировка всего пространства элементарных исходов на благоприятные для интересующего нас события и неблагоприятные очень важна для нашей оценки вероятности реализации этого события.

Задание 1.2. Из трех цифр 1, 2, 3 наудачу берутся две цифры и составляется двузначное число. Выпишите все элементарные исходы этого опыта. Составьте из элементарных исходов события A, B и C , такие что:

$$A=\{\text{число содержит цифру } 1\};$$

$$B=\{\text{число содержит только нечетные цифры}\};$$

$$C=\{\text{число четное}\}.$$

Запишите эти события, используя соответствующие математические обозначения.

Обсуждение

Пространство элементарных исходов W состоит из следующих элементарных исходов:

$$w_1=\{12\}, w_2=\{21\}, w_3=\{13\}, w_4=\{31\}, w_5=\{23\}, w_6=\{32\}.$$

Тогда:

$$A=\{w_1, w_2, w_3, w_4\};$$

$$B=\{w_3, w_4\};$$

$$C=\{w_1, w_6\}.$$

Отметим, что само пространство элементарных исходов W также представляет собой событие, происходящее всегда (при любом элементарном исходе w). События, от которых в данном эксперименте «никуда не деться», называются достоверными событиями.

События называются *независимыми*, если реализация одного из них не оказывает никакого влияния на вероятность реализации другого. К этому важному свойству мы еще вернемся в следующих разделах. События называются *несовместными*, если они не могут произойти одновременно. Тоже очень важное свойство, мы встретимся с ним еще не раз.

Пример 1.4. Элементарные исходы являются несовместными событиями при однократном опыте. Например, если мы бросили монету один раз, выпадет или решетка или герб, но, конечно, что-нибудь одно.

В простейшем случае (две стороны одной монеты, шесть граней кубика) вероятность реализации каждого элементарного исхода несложно посчитать исходя из простого здравого смысла. Задача существенно усложняется, если этих элементарных исходов очень много и их приходится определенным образом комбинировать для составления интересующего нас события. Именно в таких более сложных ситуациях на помощь приходит математика. В следующем параграфе мы познакомимся с некоторыми правилами специального раздела математики – *комбинаторики*.

2. ОСНОВНЫЕ ФОРМУЛЫ КОМБИНАТОРИКИ

Пусть имеется k групп элементов, причем i -я группа состоит из n_i элементов. Выберем по одному элементу из каждой группы. Тогда общее число N способов, которыми можно произвести такой выбор, определяется соотношением $N = n_1 \cdot n_2 \cdot n_3 \cdot \dots \cdot n_k$. Поясним это правило на простом примере.

Пример 2.1. Пусть имеется две группы элементов, причем первая группа состоит из n_1 элементов, а вторая – из n_2 элементов. Сколько различных пар элементов можно составить из этих двух групп таким образом, чтобы в паре было по одному элементу от каждой группы? Допустим, мы взяли первый элемент из первой группы и, не меняя его, перебрали все

возможные пары, меняя только элементы из второй группы. Таких пар для этого элемента можно составить n_2 . Затем мы берем второй элемент из первой группы и также составляем для него все возможные пары. Таких пар тоже будет n_2 . Так как в первой группе всего n_1 элемент, всего возможных вариантов будет $n_1 \cdot n_2$.

Пример 2.2. Сколько трехзначных четных чисел можно составить из цифр 0, 1, 2, 3, 4, 5, 6, если цифры могут повторяться?

Решение: $n_1=6$ (так как в качестве первой цифры можно взять любую цифру из 1, 2, 3, 4, 5, 6), $n_2=7$ (так как в качестве второй цифры можно взять любую цифру из 0, 1, 2, 3, 4, 5, 6), $n_3=4$ (так как в качестве третьей цифры можно взять любую цифру из 0, 2, 4, 6).

Итак, $N = n_1 \cdot n_2 \cdot n_3 = 6 \cdot 7 \cdot 4 = 168$.

В том случае, когда все группы состоят из одинакового числа элементов, т.е. $n_1=n_2=\dots n_k=n$, можно считать, что каждый выбор производится из одной и той же группы, причем элемент после выбора снова возвращается в группу. Тогда число всех способов выбора равно n^k . Такой способ выбора носит название *выборки с возвращением*.

Пример 2.3. Сколько всего четырехзначных чисел можно составить из цифр 1, 5, 6, 7, 8?

Решение. Для каждого разряда четырехзначного числа имеется пять возможностей, значит $N = 5 \cdot 5 \cdot 5 \cdot 5 = 5^4 = 625$.

Рассмотрим множество, состоящие из n элементов. Это множество будем называть *генеральной совокупностью*.

Определение 2.1. Размещением из n элементов по m называется любой упорядоченный набор из m различных элементов, выбранных из генеральной совокупности в n элементов.

Пример 2.4. Различными размещениями из трех элементов $\{1, 2, 3\}$ по два будут наборы (1, 2), (2, 1), (1, 3), (3, 1), (2, 3), (3, 2). Размещения могут отличаться друг от друга как элементами, так и их порядком.

Число размещений обозначается A_n^m и вычисляется по формуле

$$A_{m,n} = n(n-1)(n-2)\dots(n-m+1) = \frac{n!}{(n-m)!}$$

Замечание: $n!=1 \cdot 2 \cdot 3 \cdot \dots \cdot n$ (читается: «эн факториал»), кроме того, полагают, что $0!=1$.

Пример 2.5. Сколько существует двузначных чисел, в которых цифра десятков и цифра единиц различные и нечетные?

Решение. Так как нечетных цифр пять, а именно 1, 3, 5, 7, 9, то эта задача сводится к выбору и размещению на две разные позиции двух из пяти различных цифр, т.е. указанных чисел будет:

$$A_5^2 = 5 \cdot 4 = 20.$$

Определение 2.2. Сочетанием из n элементов по m называется любой неупорядоченный набор из m различных элементов, выбранных из генеральной совокупности в n элементов.

Пример 2.6. Для множества $\{1, 2, 3\}$ сочетаниями являются $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$.

Число сочетаний обозначается C_n^m и вычисляется по формуле

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

Пример 2.7. Сколькими способами читатель может выбрать две книги из шести имеющихся?

Решение. Число способов равно числу сочетаний из шести книжек по две, т.е. равно:

$$C_6^2 = \frac{6!}{2!4!} = \frac{5 \cdot 6}{2} = 15.$$

Определение 2.3. Перестановкой из n элементов называется любой упорядоченный набор этих элементов.

Пример 2.8. Всевозможными перестановками множества, состоящего из трех элементов $\{1, 2, 3\}$, являются: $(1, 2, 3)$, $(1, 3, 2)$, $(2, 3, 1)$, $(2, 1, 3)$, $(3, 2, 1)$, $(3, 1, 2)$.

Число различных перестановок из n элементов обозначается P_n и вычисляется по формуле $P_n = n!$

Пример 2.9. Сколькими способами семь книг разных авторов можно расставить на полке в один ряд?

Решение. Это задача о числе перестановок семи разных книг. Имеется $P_7 = 7! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 = 5040$ способов осуществить расстановку книг.

Обсуждение. Мы видим, что число возможных комбинаций можно посчитать по разным правилам (перестановки, сочетания, размещения), причем результат получится различный, так как принцип подсчета и сами формулы различаются. Внимательно посмотрев на определения, можно заметить, что результат зависит от нескольких факторов одновременно.

Во-первых, от того, из какого количества элементов мы можем комбинировать их наборы (насколько велика генеральная совокупность элементов).

Во-вторых, от того, какой величины наборы элементов нам нужны.

И, в-третьих, важно знать, является ли для нас существенным порядок элементов в наборе. Поясним последний фактор на следующем примере.

Пример 2.10. На родительском собрании присутствует 20 человек. Сколько существует различных вариантов состава родительского комитета, если в него должны войти 5 человек?

Решение. В этом примере нас не интересует порядок фамилий в списке комитета. Если в результате в его составе окажутся одни и те же люди, то по смыслу для нас это один и тот же вариант. Поэтому мы можем воспользоваться формулой для подсчета числа *сочетаний* из 20 элементов по 5. Иначе будут обстоять дела, если каждый член комитета изначально отвечает за определенное направление работы. Тогда при одном и том же списочном составе комитета внутри него возможно $5!$ вариантов *перестановок*, которые имеют значение. Количество разных (и по составу, и по сфере ответственности) вариантов определяется в этом случае числом *размещений* из 20 элементов по 5.

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1. Сколько трехзначных четных чисел можно составить из цифр 0, 1, 2, 3, 4, 5, 6, если цифры могут повторяться?
2. Сколько существует пятизначных чисел, которые одинаково читаются слева направо и справа налево?
3. В классе десять предметов и пять уроков в день. Сколькими способами можно составить расписание на один день?
4. Сколькими способами можно выбрать 4 делегата на конференцию, если в группе 20 человек?
5. Сколькими способами можно разложить восемь различных писем по восьми различным конвертам, если в каждый конверт кладется только одно письмо?
6. Из трех математиков и десяти экономистов надо составить комиссию, состоящую из двух математиков и шести экономистов. Сколькими способами это можно сделать?

3. АЛГЕБРА СОБЫТИЙ

Представим, что у нас проводится эксперимент с пространством из n элементарных исходов, которые *равновероятны*. Элементарные исходы являются *несовместными* событиями (напомним, что несовместные события – это те, которые не могут произойти одновременно), поэтому вероятность каждого из них равна $1/n$. Допустим, нас интересует событие A , которое наступает только при реализации *благоприятных* элементарных

12

исходов, количество последних m ($m < n$). Тогда, согласно классическому определению, вероятность такого события

$$P(A) = \frac{m}{n}.$$

Для любого события A справедливо неравенство: $0 < P(A) < 1$.

Пример 3.1. Лотерея состоит из 1000 билетов, среди них 200 выигрышных. Наугад вынимается один билет из 1000. Чему равна вероятность того, что этот билет выигрышный?

Решение. Различных исходов в этом примере 1000 ($n = 1000$). В интересующее нас событие A входят 200 исходов ($m = 200$). Таким образом,

$$P(A) = \frac{200}{1000} = 0,2.$$

Пример 3.2. В коробке лежат 200 белых, 100 красных и 50 зеленых шаров. Наудачу вынимается один шар. Чему равны вероятности получить шар белого, красного или зеленого цвета?

Решение. Рассмотрим события:

$A = \{\text{вынули белый шар}\},$

$B = \{\text{вынули красный шар}\},$

$C = \{\text{вынули зеленый шар}\}.$

$n = 350,$

тогда

$$P(A) = \frac{200}{350} = \frac{4}{7};$$

$$P(B) = \frac{100}{350} = \frac{2}{7};$$

$$P(C) = \frac{50}{350} = \frac{1}{7}.$$

Пример 3.3. Бросается игральная кость. Чему равны вероятности следующих событий:

$A = \{\text{выпала грань с 6 очками}\},$

$B = \{\text{выпала грань с четным числом очков}\},$

$C = \{\text{выпала грань с числом очков, делящимся на 3}\}?$

Решение. $n = 6$. Событию A благоприятствует один исход, событию B – три исхода, событию C – два исхода. Таким образом,

$$P(A) = \frac{1}{6};$$

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

$$P(C) = \frac{2}{6} = \frac{1}{3}.$$

Иногда в задачах число элементарных исходов бывает так велико, что выписать их все не представляется возможным. Поэтому применяются формулы из комбинаторики (см. раздел 2).

Пример 3.4. Из колоды в 36 карт вытаскивают три. Какова вероятность того, что среди вынутых карт нет десятков?

Решение. В этом примере элементарным исходом является случайный набор из трех карт. Общее число элементарных исходов равно $N = C_{36}^3$, элементарные исходы считаем равновероятными. Благоприятных исходов (количество возможных наборов по три карты из той же колоды, но без десятков) $m = C_{32}^3$. Таким образом, вероятность события A {вынуто 3 карты

из 36 и среди них нет десятков}: $P(A) = \frac{C_{32}^3}{C_{36}^3} = \frac{248}{357}$.

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1. Бросаются одновременно две игральные кости. Найти вероятности следующих событий: A – сумма выпавших очков равна 8; B – произведение выпавших очков равно 8.

2. В конверте среди 100 фотокарточек находится одна разыскиваемая. Из конверта наудачу извлечены 10 карточек. Найти вероятность того, что среди них окажется разыскиваемая.

3. Набирая номер телефона, абонент забыл последние три цифры и, помня лишь, что эти цифры различны, набрал их наудачу. Найти вероятность того, что номер набран правильно.

При оценке вероятности наступления какого-либо случайного события очень важно предварительно хорошо представлять, зависит ли вероятность наступления интересующего нас события от того, как развиваются остальные события. В случае классической схемы, когда все исходы равновероятны, мы уже можем оценить значения вероятности интересующего нас отдельного события самостоятельно. Мы можем сделать это даже в том случае, если событие является сложной совокупностью нескольких элементарных исходов. А если несколько случайных событий происходит одновременно или последовательно? Как это влияет на вероятность реализации интересующего нас события?

Если я несколько раз кидаю игральную кость и хочу, чтобы выпала «шестерка», а мне все время не везет, значит ли это, что надо увеличивать ставку, потому что, согласно теории вероятностей, мне вот-вот должно повезти? Увы, теория вероятностей не утверждает ничего подобного. Ни кости, ни карты, ни монетки *не умеют запоминать*, что они продемонстрировали нам в прошлый раз. Им совершенно не важно, в первый или в десятый раз сегодня я испытываю свою судьбу. Каждый раз, когда я повторяю бросок, я знаю только одно: и на этот раз вероятность выпадения «шестерки» снова равна одной шестой. Конечно, это не значит, что нужная мне цифра не выпадет никогда. Это означает лишь то, что мой проигрыш после первого броска и после любого другого броска – независимые события.

События А и В называются *независимыми*, если реализация одного из них никак не влияет на вероятность другого события. Например, вероятности поражения цели первым из двух орудий не зависят от того, поразило ли цель другое орудие, поэтому события «первое орудие поразило цель» и «второе орудие поразило цель» независимы.

Если два события А и В независимы и вероятность каждого из них известна, то вероятность одновременного наступления события А и события В (обозначается АВ) можно посчитать, воспользовавшись следующей теоремой.

Теорема умножения вероятностей для независимых событий

$P(AB) = P(A) \cdot P(B)$ – вероятность одновременного наступления двух независимых событий равна произведению вероятностей этих событий.

Пример 3.5. Вероятности попадания в цель при стрельбе первого и второго орудий соответственно равны: $p_1 = 0,7$; $p_2 = 0,8$. Найти вероятность попадания при одном залпе обоими орудиями одновременно.

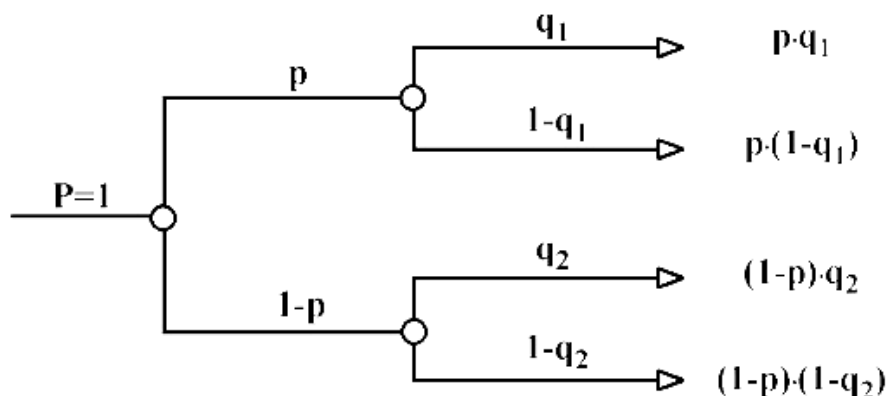
Решение. Как мы уже видели, события А (попадание первого орудия) и В (попадание второго орудия) независимы, т.е. $P(AB) = P(A) \cdot P(B) = p_1 \cdot p_2 = 0,56$.

Что произойдет с нашими оценками, если исходные события не являются независимыми? Давайте немного изменим предыдущий пример.

Пример 3.6. Два стрелка на соревнованиях стреляют по мишеням, причем если один из них стреляет метко, то соперник начинает нервничать, и его результаты ухудшаются. Как превратить эту житейскую ситуацию в математическую задачу и наметить пути ее решения? Интуитивно понятно, что надо каким-то образом разделить два варианта развития событий, составить, по сути дела, два сценария, две разные задачи. В первом случае, если соперник промахнулся, сценарий будет благоприятным для нервного спортсмена и его меткость будет выше. Во втором случае, если

соперник прилично реализовал свой шанс, вероятность поразить мишень для второго спортсмена снижается.

Для разделения возможных сценариев (их также называют гипотезами) развития событий мы будем часто использовать схему «дерева вероятностей». Эта схема похожа по смыслу на дерево решений, с которым вам, наверное, уже приходилось иметь дело. Каждая ветка представляет собой отдельный сценарий развития событий, только теперь она имеет собственное значение так называемой *условной вероятности* (q_1, q_2, q_1-1, q_2-1).

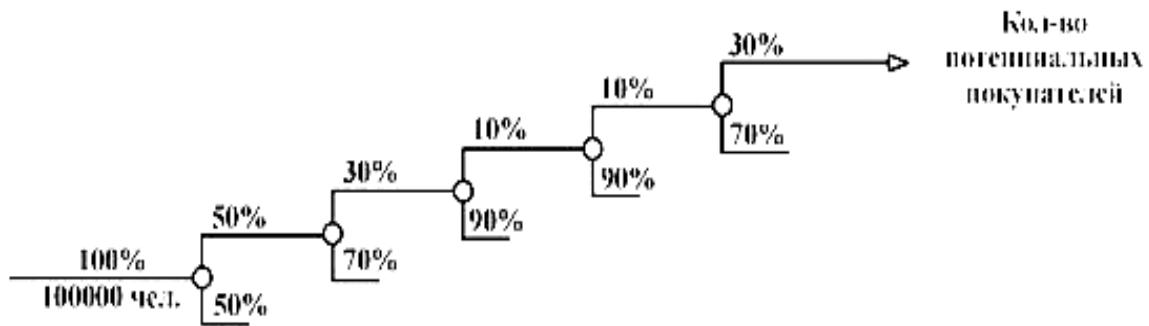


Эта схема очень удобна для анализа последовательных случайных событий.

Остается выяснить еще один немаловажный вопрос: откуда берутся исходные значения вероятностей в *реальных ситуациях*? Ведь не с одними же монетами и игральными костями работает теория вероятностей? Обычно эти оценки берутся из статистики, а когда статистические сведения отсутствуют, мы проводим собственное исследование. И начинать его нам часто приходится не со сбора данных, а с вопроса, какие сведения нам вообще нужны.

Пример 3.7. Допустим, нам надо оценить в городе с населением 100 тыс. жителей объем рынка для нового товара, который не является предметом первой необходимости, например, для бальзама по уходу за окрашенными волосами. Рассмотрим схему «дерева вероятностей». При этом значение вероятности на каждой «ветке» нам надо приблизительно оценить. Итак, наши оценки емкости рынка:

- 1) из всех жителей города женщин 50%;
- 2) из всех женщин только 30% красят волосы часто;
- 3) из них только 10% пользуются бальзамами для окрашенных волос;
- 4) из них только 10% могут набраться смелости попробовать новый товар;
- 5) из них 70% обычно покупает все не у нас, а у наших конкурентов.



Решение. По закону перемножения вероятностей, определяем вероятность интересующего нас события $A = \{\text{житель города покупает у нас этот новый бальзам}\} = 0,00045$.

Умножим это значение вероятности на число жителей города. В результате имеем всего 45 потенциальных покупательниц, а если учесть, что одного пузырька этого средства хватает на несколько месяцев, получается не слишком оживленная торговля.

И все-таки польза от наших оценок есть.

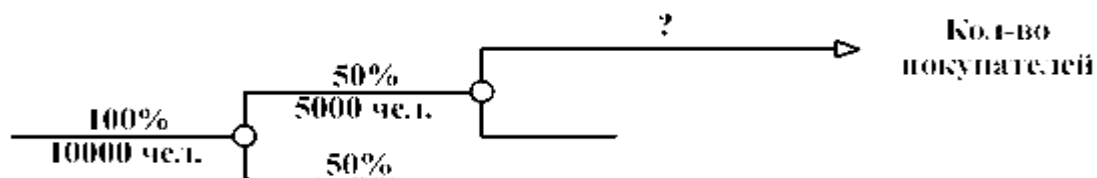
Во-первых, мы можем сравнивать прогнозы разных бизнес-идей, на схемах у них будут разные «развилки», и, конечно, значения вероятности тоже будут разные.

Во-вторых, как мы уже говорили, случайная величина не потому называется случайной, что она совсем ни от чего не зависит. Просто ее *точное* значение заранее не известно. Мы знаем, что среднее количество покупателей может быть увеличено (например, с помощью рекламы нового товара). Так что имеет смысл сосредоточить усилия на тех «развилках», где распределение вероятностей нас особенно не устраивает, на тех факторах, на которые мы в состоянии повлиять.

Рассмотрим еще один количественный пример исследования покупательского поведения.

Пример 3.8. За день продовольственный рынок посещает в среднем 10000 чел. Вероятность того, что посетитель рынка заходит в павильон молочных продуктов, равна $1/2$. Известно, что в этом павильоне в среднем продается в день 500 кг различных продуктов. Можно ли утверждать, что средняя покупка в павильоне весит всего 100 г?

Обсуждение. Конечно, нельзя. Понятно, что не каждый, кто заходил в павильон, в результате что-то там купил.

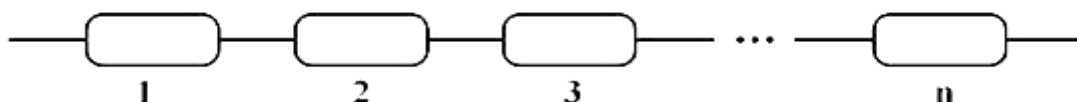


Как показано на схеме, чтобы ответить на вопрос о среднем весе покупки, мы должны найти ответ на вопрос, какова вероятность того, что человек, зашедший в павильон, что-нибудь там купит. Если таких данных в нашем распоряжении не имеется, а нам они нужны, придется их получить самим, понаблюдав некоторое время за посетителями павильона. Допустим, наши наблюдения показали, что только пятая часть посетителей павильона что-то покупает.

Как только эти оценки нами получены, задача становится уже простой. Из 10000 чел., пришедших на рынок, 5000 зайдут в павильон молочных продуктов, покупок будет только 1000. Средний вес покупки равен 500 г. Интересно отметить, что для построения полной картины происходящего логика условных «ветвлений» должна быть определена на каждом этапе нашего рассуждения так же четко, как если бы мы работали с «конкретной» ситуацией, а не с вероятностями.

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1. Пусть есть электрическая цепь, состоящая из n последовательно соединенных элементов, каждый из которых работает независимо от остальных.



Известна вероятность p невыхода из строя каждого элемента. Определите вероятность исправной работы всего участка цепи (событие А).

2. Студент знает 20 из 25 экзаменационных вопросов. Найдите вероятность того, что студент знает предложенные ему экзаменатором три вопроса.

3. Производство состоит из четырех последовательных этапов, на каждом из которых работает оборудование, для которого вероятности выхода из строя в течение ближайшего месяца равны соответственно p_1 , p_2 , p_3 и p_4 . Найдите вероятность того, что за месяц не случится ни одной остановки производства из-за неисправности оборудования.

4. ТЕОРЕМЫ СЛОЖЕНИЯ И УМНОЖЕНИЯ ВЕРОЯТНОСТЕЙ

Классическое определение вероятности $P(A)$ события A как отношения числа благоприятных элементарных исходов m к числу всех элементарных исходов n предполагает, что все элементарные исходы *равновероятны*. Однако это условие выполняется далеко не всегда, поэтому мы введем еще одно определение вероятности – статистическое (или частотное).

Как оценить вероятность интересующего нас события, если в процессе испытания элементарные исходы вовсе не обязаны быть равновероятными? Строго говоря, необходимо было бы много раз проделать интересующий нас опыт и узнать частоту реализации различных элементарных исходов. В пределе, при увеличении числа испытаний, отношение числа m реализованных событий A к общему количеству испытаний n и будет определять вероятность $P(A) = m/n$. Важно понимать, что статистический подход не противоречит классическому, а лишь расширяет границы возможного применения аппарата теории вероятностей. Поэтому все приемы, которые вы уже освоили в рамках классической схемы, можно будет использовать и в дальнейшем. Для решения практических задач нам понадобятся следующие важные теоремы.

1. Теорема сложения вероятностей для несовместных событий: $P(A + B) = P(A) + P(B)$ – вероятность наступления в результате эксперимента *хотя бы одного* из двух несовместных событий равна сумме вероятностей этих событий.

Обсуждение. Напомним, что события A и B называются *несовместными*, если в результате опыта они не могут произойти одновременно. (Пожалуйста, не путайте их с независимыми событиями, которые мы обсуждали в прошлом разделе. Независимые события могут спокойно сосуществовать друг с другом.)

Пример 4.1. По статистике, в прошлом году 10% жителей нашего города встретили Новый год в отъезде, 40% ходили в гости или в ресторан, оставаясь в городе, остальные встречали Новый год дома. Считая, что эта тенденция сохранится, посчитайте вероятность того, что житель нашего города встретит Новый год дома.

Решение. Здесь можно смело пользоваться теоремой сложения вероятностей, так как события встречи Нового года в разных местах одним и тем же человеком несовместны. Поэтому все, кто встретит Новый год в гостях или в другом городе (они составят вместе 40%+10%), не смогут встретить его дома. Принимая общее число жителей города за 100%, найдем, что 50% в прошлый раз оставалось дома. Полагая, что эти же пропорции сохранятся и в этом году, найдем, что вероятность встретить Но-

вый год дома для жителя нашего города равна $P = 0,5$. (Заметим, что в данном случае нам было удобно посчитать сначала вероятность обратного события, а потом вычесть результат из 100%.)

Что произойдет с нашими оценками, если исходные события не являются несовместными? Давайте немного изменим предыдущий пример.

Пример 4.2. Владелец фирмы частных такси хочет сделать прогноз количества клиентов на новогоднюю ночь. Пусть, по его сведениям, в прошлом году Новый год встретили дома 50%, в компании друзей или родственников, но не выезжая из города – 80%, в отъезде были 10%. Почему у него получилось в сумме больше 100%?

Видимо, каких-то жителей он посчитал больше одного раза. Скорее всего, тех, кто сидел дома, но одновременно принимал друзей или родственников, которые пришли к нему в гости. Поскольку эти события не являются несовместными, просто складывая вероятности, он завышает свои оценки.

Впрочем, это относится не только к оценке вероятности события, но и к решению любых задач на подсчет элементов объединения двух множеств путем сложения. Если множества частично перекрываются, сумма их элементов будет больше, чем реальное количество элементов, поскольку при арифметическом сложении элементы этого «перекрытия» мы невольно посчитаем дважды, и как входящие в первое множество, и как входящие во второе. Выход здесь один: мы должны заметить, что множества частично «перекрываются», посчитать число элементов в их общей части и вычесть это число из суммы (так как при суммировании мы его посчитали дважды).

В случае подсчета вероятности события C , которое наступает или при наступлении события A , или при наступлении события B , если A и B не являются несовместными, можно воспользоваться следующей теоремой.

Общая теорема сложения вероятностей

$P(C) = P(A) + P(B) - P(AB)$, где $P(AB)$ – вероятность одновременного наступления событий A и B .

Пример 4.3. В урне 10 белых, 15 черных, 20 синих и 25 красных шаров. Вынули один шар. Какова вероятность того, что вынутый шар белый? черный? синий? красный? белый или черный? синий или красный? белый, черный или синий?

Решение. Имеем

$$n = 10 + 15 + 20 + 25 = 70, P(B) = \frac{10}{70} = \frac{1}{7}, P(Ч) = \frac{15}{70} = \frac{3}{14},$$
$$P(C) = \frac{20}{70} = \frac{2}{7}, P(K) = \frac{25}{70} = \frac{5}{14}.$$

Применяем аксиому сложения вероятностей:

$$P(B + Ч) = P(B) + P(Ч) = \frac{1}{7} + \frac{3}{14} = \frac{5}{14};$$

$$P(C + K) = P(C) + P(K) = \frac{2}{7} + \frac{5}{14} = \frac{9}{14};$$

$$P(B + Ч + C) = 1 - P(K) = 1 - \frac{5}{14} = \frac{9}{14}.$$

Пример 4.4. В первом ящике 2 белых и 10 черных шаров. Во втором ящике 8 белых и 4 черных шара. Из каждого ящика вынули по шару. Какова вероятность, что оба шара белые?

Решение. В данном случае речь идет о совмещении событий A и B , где событие A – появление белого шара из первого ящика, событие B – появление белого шара из второго ящика. При этом A и B – независимые события. Имеем:

$$P(A) = \frac{2}{12} = \frac{1}{6}, P(B) = \frac{8}{12} = \frac{2}{3}.$$

Применяем теорему умножения вероятностей:

$$P(AB) = P(A) \cdot P(B) = \frac{1}{6} \cdot \frac{2}{3} = \frac{1}{9}.$$

Пример 4.5. В условиях предыдущей задачи определить вероятность того, что один из вынутых шаров белый, а другой черный.

Решение

Событие A – появление белого шара из первого ящика.

Событие B – появление белого шара из второго ящика.

Событие C – появление черного шара из первого ящика ($C = \bar{A}$).

Событие D – появление черного шара из второго ящика ($D = \bar{B}$).

$$P(A) = \frac{1}{6}, P(B) = \frac{2}{3}, P(C) = P(\bar{A}) = 1 - \frac{1}{6} = \frac{5}{6},$$

$$P(D) = P(\bar{B}) = 1 - \frac{2}{3} = \frac{1}{3}.$$

Определим вероятность того, что шар, вынутый из первого ящика, белый, а из второго ящика – черный:

$$P(AD) = P(A) \cdot P(D) = \frac{1}{6} \cdot \frac{1}{3} = \frac{1}{18}.$$

Определим вероятность того, что шар, вынутый из первого ящика, черный, а из второго ящика – белый:

$$P(BC) = P(B) \cdot P(C) = \frac{2}{3} \cdot \frac{5}{6} = \frac{5}{9}.$$

Определим теперь вероятность того, что шар, вынутый из одного ящика (безразлично из первого или второго), будет белым, а шар, вынутый из другого ящика, – черным. Применяем теорему сложения вероятностей:

$$P = P(AD) + P(BC) = \frac{1}{18} + \frac{5}{9} = \frac{11}{18}.$$

Пример 4.6. Три стрелка стреляют по цели. Вероятность попадания в цель для первого стрелка равна 0,75, для второго – 0,8, для третьего – 0,9. Определить вероятность того, что все три стрелка одновременно попадут в цель.

Решение. Имеем

$$P(A) = 0,75, \quad P(B) = 0,8, \quad P(C) = 0,9;$$

$$P(ABC) = P(A) \cdot P(B) \cdot P(C) = 0,75 \cdot 0,8 \cdot 0,9 = 0,54.$$

Пример 4.7. В условиях предыдущей задачи определить вероятность того, что в цель попадет хотя бы один стрелок.

Решение. $P(\overline{A}) = 1 - 0,75 = 0,25$ (вероятность промаха первого стрелка); $P(\overline{B}) = 1 - 0,8 = 0,2$ (вероятность промаха второго стрелка); $P(\overline{C}) = 1 - 0,9 = 0,1$ (вероятность промаха третьего стрелка); тогда $P(\overline{ABC})$ – вероятность одновременного промаха всех трех стрелков – определится следующим образом:

$$P(\overline{ABC}) = P(\overline{A}) \cdot P(\overline{B}) \cdot P(\overline{C}) = 0,25 \cdot 0,2 \cdot 0,1 = 0,005.$$

Но событием, противоположным событию $\overline{ABC}$, является событие, заключающееся в поражении цели хотя бы одним стрелком. Следовательно, искомая вероятность найдется по формуле

$$P = 1 - P(\overline{ABC}), \text{ т.е. } P = 1 - 0,005 = 0,995.$$

Пример 4.8. В ящике a белых и b черных шаров. Какова вероятность того, что из двух вынутых шаров один белый, а другой черный? (Вынутый шар в урну не возвращается.)

Решение

Событие A – появление белого шара при первом вынимании.

Событие B – появление черного шара при втором вынимании.

Событие C – появление черного шара при первом вынимании.

Событие D – появление белого шара при втором вынимании.

Вычислим вероятность того, что первый вынутый шар белый, а второй – черный:

$$P_1 = P(A) \cdot P(B / A) = \frac{a}{a+b} \cdot \frac{b}{a+b-1} = \frac{ab}{(a+b)(a+b-1)}.$$

Найдем вероятность того, что первый вынутый шар черный, а второй – белый:

$$P_2 = P(C) \cdot P(D / C) = \frac{b}{a+b} \cdot \frac{a}{a+b-1} = \frac{ab}{(a+b)(a+b-1)}.$$

Таким образом, вероятность того, что один из вынутых шаров белый, а другой – черный, определится по теореме сложения:

$$P = P_1 + P_2, \text{ т.е. } P = \frac{2ab}{(a+b)(a+b-1)}.$$

ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ

1. Из колоды в 36 карт наудачу вынимается одна. Какова вероятность того, что будет вынута карта пик или дама?

2. На полке стоят 15 книг, из них пять в переплете. Наудачу берут три книги. Найти вероятность того, что хотя бы одна из взятых книг в переплете (событие A).

5. ФОРМУЛА ПОЛНОЙ ВЕРОЯТНОСТИ

В данном параграфе мы рассмотрим более сложные задачи, имеющие большое практическое значение. Эти задачи встречаются на практике в случае, если имеются несколько возможных сценариев развития событий (несколько вероятных гипотез). Известны как величины вероятности реализации каждого из сценариев, так и вероятность наступления интересующего нас события для любого из этих сценариев, а нас интересует, какова полная (совокупная) вероятность наступления интересующего нас события.

Предположим, что в результате опыта может произойти одно из n несовместных событий (гипотез) $H_1, H_2, \dots, H_n$. Пусть также имеется некоторое событие A и известны $P(H_i)$ – вероятность гипотезы, $P(A/H_i)$ –

условная вероятность события А при этой гипотезе. Тогда вероятность события А вычисляется по формуле полной вероятности:

$$P(A) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + \dots + P(H_n) \cdot P(A/H_n),$$

или

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A/H_i).$$

Пример 5.1. Из 40 деталей 10 изготовлены в первом цехе, 25 – во втором, а остальные – в третьем. Первый и третий цехи дают продукцию отличного качества с вероятностью 0,9, второй цех – с вероятностью 0,7. Какова вероятность того, что взятая наудачу деталь будет отличного качества?

Решение. Обозначим событие $A = \{\text{выбрана деталь отличного качества}\}$, $H_i = \{\text{выбранная деталь изготовлена в } i\text{-м цехе}\}$, $i = 1, 2, 3$. Тогда

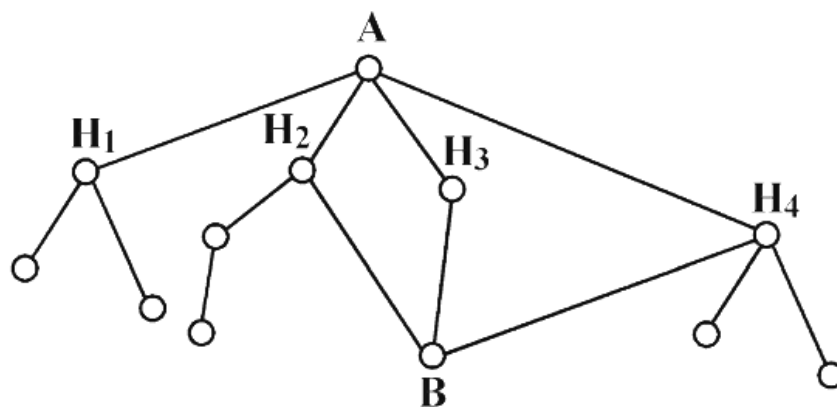
$$P(H_1) = \frac{10}{40} = \frac{1}{4}; \quad P(H_2) = \frac{25}{40} = \frac{5}{8}; \quad P(H_3) = \frac{5}{40} = \frac{1}{8}.$$

По условию задачи $P(A/H_1) = P(A/H_3) = 0,9$, $P(A/H_2) = 0,7$. По формуле полной вероятности находим искомую вероятность:

$$P(A) = \sum_{i=1}^3 P(H_i)P(A|H_i) = \frac{1}{4} \cdot 0,9 + \frac{5}{8} \cdot 0,7 + \frac{1}{8} \cdot 0,9 = 0,775.$$

Пример 5.2. На рисунке изображена схема дорог. Найти вероятность того, что турист, вышедший из пункта А, попадет в пункт В, если на развилке он наугад выбирает любую дорогу (кроме обратной).

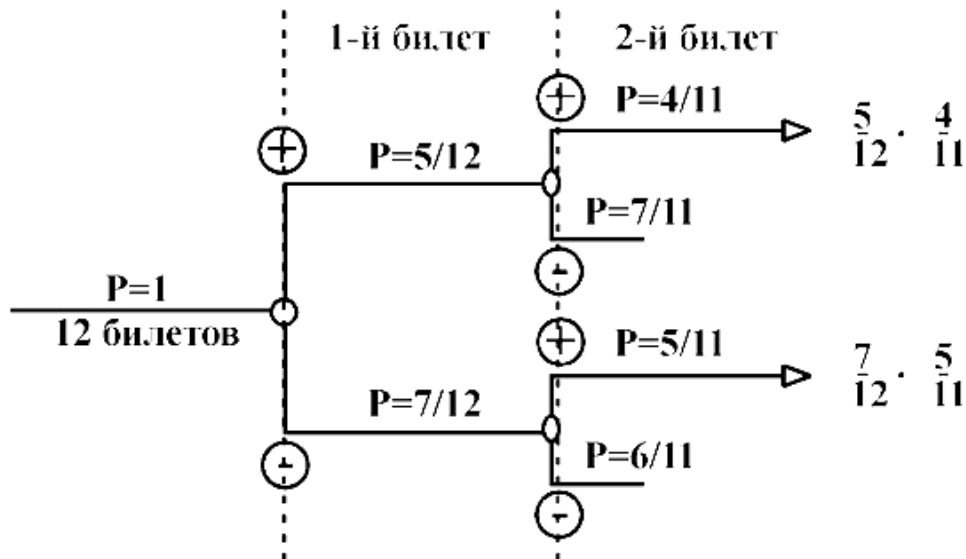
Решение. Обозначим $H_i = \{\text{приход туриста в пункт } H_i\}$, $i = 1, 2, 3, 4$. Поскольку, выйдя из пункта А, он выбирает любую дорогу наугад, то $P(H_i) = 1/4$, $i = 1, 2, 3, 4$.



Исходя из схемы дорог, определяем, что $P(B/H_1) = 0$; $P(B/H_2) = 1/2$; $P(B/H_3) = 1$; $P(B/H_4) = 1/3$. Таким образом, по формуле полной вероятности

$$P(B) = \sum_{i=1}^3 P(H_i)P(B|H_i) = \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot \frac{1}{3} = \frac{11}{24} \approx 0,458.$$

Пример 5.3. Из двенадцати лотерейных билетов пять выигрышных. Билеты вытягиваются по одному без возвращения. Какова вероятность того, что во второй раз вытянут выигрышный билет?



Решение. Как обычно, вдоль каждой ветви «дерева вероятностей» значения вероятностей перемножаются, а затем значения на концах нужных веток между собой складываются. В результате получаем ответ:

$$P(B) = \frac{5}{12} \cdot \frac{4}{11} + \frac{7}{12} \cdot \frac{5}{11} = \frac{5}{12}.$$

ЗАДАНИЕ 1

В отделе найма персонала проводится тестирование на вакантную руководящую должность. Тест составлен из двух производственных ситуаций, не связанных между собой логически. По каждой ситуации предлагается три примера дальнейших действий, из которых надо выбрать один наилучший.

Вероятность того, что претендент знает ответ на первую часть теста, равна P_1 , вероятность того, что он знает ответ на вторую часть, равна P_2 .

Допустим, что в случае, когда претендент не знает ответа, он принимает решение произвольно выбирать из трех предлагаемых вариантов наугад.

Нарисуйте дерево вероятностей и посчитайте вероятности для разных сценариев развития событий. Прокомментируйте свои оценки.

Какова вероятность того, что испытуемый знает ответы на обе части теста?

Какова полная вероятность, что испытуемый ответит правильно на обе части теста?

Можете ли вы предложить свои способы повышения достоверности результатов такого тестирования?

6. ПОВТОРЕНИЕ ИСПЫТАНИЙ. СХЕМА БЕРНУЛЛИ

Практические задачи, связанные с оценкой вероятности наступления события в результате нескольких равноценных попыток, могут анализироваться с применением формулы Бернулли или (при большом количестве таких попыток) с применением приближенной формулы Пуассона. Для работы с этим материалом вам снова потребуется знание основ комбинаторики.

Схема Бернулли состоит в следующем: производится последовательность испытаний, в каждом из которых вероятность наступления определенного события A одна и та же и равна p . Испытания предполагаются независимыми (считается, что вероятность появления события A в каждом из испытаний не зависит от того, появилось или не появилось это событие в других испытаниях). Наступление события A обычно называют успехом, а ненаступление – неудачей. Обозначим вероятность неудачи $q = 1 - P(A) = (1 - p)$. Вероятность того, что в n независимых испытаниях успех наступит ровно m раз, выражается *формулой Бернулли*:

$$P_{m,n} = C_n^m p^m q^{n-m}, \text{ где } q = 1 - p.$$

Вероятность $P_n(m)$ при данном n сначала увеличивается при увеличении m от 0 до некоторого значения m_0 , а затем уменьшается при изменении m от m_0 до n .

Поэтому m_0 называют *наивероятнейшим числом* наступлений успеха в опытах. Это число m_0 заключено между числами $np - q$ и $np + p$ (или, что то же самое, между числами $n(p+1) - 1$ и $n(p+1)$). Если число $np - q$ – целое число, то наивероятнейших чисел два: $np - q$ и $np + p$.

Важное замечание. Если $np - q < 0$, то наивероятнейшее число выигрышей равно нулю.

Пример 6.1. Игральная кость бросается 4 раза. При каждом броске нас интересует событие $A = \{\text{выпала шестерка}\}$.

Решение. Здесь четыре испытания, и так как кубик симметричен, то $p = P(A) = 1/6$, $q = 1 - p = 5/6$.

Вероятность того, что в 4 независимых испытаниях успех наступит ровно m раз ($m < 4$), выражается формулой Бернулли:

$$P_4(m) = C_4^m \left(\frac{1}{6}\right)^m \left(\frac{5}{6}\right)^{4-m}, \quad m = 0, \dots, 4.$$

Посчитаем эти значения и запишем их в таблицу.

m	0	1	2	3	4
$P(m)$	0,482253	0,096451	0,01929	0,003858	0,000772

Самое вероятное число успехов в нашем случае $m_0 = 0$.

Пример 6.2. Вероятность появления успеха равна $3/5$. Найти наивероятнейшее число наступлений успеха, если число испытаний равно 19, 20.

Решение. При $n = 19$ находим $np - q = 19 \cdot \frac{3}{5} - \frac{2}{5} = 11$.

Таким образом, максимальная вероятность достигается для двух значений m_0 , равных 11 и 12. Эта вероятность равна $P_{19}(11) = P_{19}(12) = 0,1797$. При $n = 20$ максимальная вероятность достигается только для одного значения m_0 , так как

$$np - q = 20 \cdot \frac{3}{5} - \frac{2}{5} = 12 - \frac{2}{5}$$

не является целым числом. Наивероятнейшее число наступлений успеха m_0 равно 12. Вероятность его появления равна $P_{20}(12) = 0,1797$. Совпадение чисел $P_{20}(12)$ и $P_{19}(12)$ вызвано лишь сочетанием значений n и p и не имеет общего характера.

На практике в случае, когда n велико, а p мало (обычно $p < 0,1$; $npq < 10$), вместо формулы Бернулли применяют приближенную формулу Пуассона

$$P_n(m) \approx \frac{\lambda^m}{m!} \cdot e^{-\lambda}, \quad \text{где} \quad \lambda = np.$$

Пример 6.3. Радиоаппаратура состоит из 1000 элементов. Вероятность отказа одного элемента в течение года равна 0,002. Какова вероятность отказа двух элементов за год? Какова вероятность отказа не менее двух элементов за год?

Решение. Будем рассматривать работу каждого элемента как отдельное испытание. Обозначим $A = \{\text{отказ элемента за год}\}$.

$$P(A) = p = 0,002, \quad \lambda = np = 1000 \cdot 0,002 = 2.$$

По формуле Пуассона

$$P_{1000}(2) \approx \frac{2^2}{2!} \cdot e^{-2} = 2 \cdot e^{-2} \approx 0,2707.$$

Обозначим через $P_{1000}(>2)$ вероятность отказа не менее двух элементов за год. Переходя к противоположному событию, вычислим $P_{1000}(>2)$ как

$$P_{1000}(\geq 2) = 1 - P_{1000}(0) - P_{1000}(1) \approx 1 - \frac{2^2}{0!} \cdot e^{-2} - \frac{2^2}{1!} \cdot e^{-2} = 1 - 3 \cdot e^{-2} \approx 0,594.$$

7. ДИСПЕРСИЯ И СРЕДНЕКВАДРАТИЧЕСКОЕ ОТКЛОНЕНИЕ

Рассмотрим пример. Фирма имеет две палатки, торгующие горячей выпечкой, которая выпекается на месте из полуфабрикатов. В таблице приводится примерная сводка ежедневной выручки каждой из палаток за неделю (р.).

Дни недели	Понедельник	Вторник	Среда	Четверг	Пятница	Суббота	Воскресенье
Палатка 1	205	268	258	218	341	1515	1397
Палатка 2	759	801	670	599	633	420	301

Различие в ежедневной выручке в основном связано с расположением палаток. Палатка 1 находится в парке отдыха, в то время как палатка 2 расположена напротив школы и вблизи проходной крупного НИИ. Владелец фирмы решил выплачивать ежемесячную премию продавцам той палатки, которая даст в этом месяце большую выручку. При распределении премии выяснилась удивительная вещь: выигрыш в этом «соревновании» зависел только от количества выходных в месяце. Не хотелось бы приводить большое количество цифр за весь месяц в целом, но и без этого видно, что если бы владельцу фирмы пришла в голову идея ежедневного премирования победителя какой-то фиксированной суммой, «Палатка выходного дня» могла бы рассчитывать на премии в два с половиной раза реже, хотя недельная выручка от нее больше. В таких условиях более разумное соревнование могло бы быть основано на осреднении показателей за неделю. Допустим, недельные показатели практически совпали. Как оценить, какая из палаток полезнее для фирмы, если по каким-то причинам фирме необходимо продать одну из них? Если выручка практически совпадает, владелец, по-видимому, поинтересуется стабильностью работы торговой точки. Вины продавцов в этом нет, но если оборудование работает два дня

в неделю на износ, а в остальное время больше простоев, выход из строя такого оборудования более вероятен. Пусть в один (случайным образом выпавший) день в неделю идет сильный дождь и на улицах мало прохожих. Падение выручки особенно резко заметно, когда такой дождливый день совпадает с одним из выходных.

Для сравнения можно представить спортсменов, которые имеют равные шансы выиграть, но один из них выступает ровнее. Скорее всего, именно он и будет принят в состав сборной. Но вот еще один вопрос: а не делает ли эта самая нестабильная палатка работу фирмы в целом более стабильной, прекрасно дополняя работу палатки 2? Давайте выдвинем это утверждение в качестве гипотезы и попробуем его доказать или опровергнуть. Чтобы оценить эту проблему количественно, надо прежде всего просуммировать дневную выручку обеих палаток.

Дни недели	Понедельник	Вторник	Среда	Четверг	Пятница	Суббота	Воскресенье
Палатки 1+2	964	1069	928	817	974	1935	1698

То, что мы описали общими словами как «нестабильность работы», в статистике называется *характеристикой рассеивания*. К этим характеристикам относятся такие показатели, как *дисперсия* и *среднее квадратическое отклонение*. Покажем на предыдущем примере, как определяются эти понятия. Посчитаем сначала среднее арифметическое выручки для каждой палатки отдельно и для обеих палаток вместе (осреднение проводим за семь дней):

$$X_{cp.1} = 600 \text{ р.}, X_{cp.2} = 598 \text{ р.}, X_{cp.1+2} = 1198 \text{ р.}$$

Чтобы сравнить разброс значений, посчитаем для обеих палаток дневные отклонения выручки от их собственного среднего значения.

Дни недели	Понедельник	Вторник	Среда	Четверг	Пятница	Суббота	Воскресенье	Всего
Палатка 1	-395	-332	-342	-383	-259	915	797	0
Палатка 2	161	203	72	1	35	-178	-297	0
Палатки 1+2	-234	-129	-270	-382	-224	737	500	0

Чтобы измерить, насколько одна палатка «нестабильнее» другой, хочется сложить всю строку за неделю и получить общее отклонение за весь

отчетный период. Но этого делать нельзя, мы сами так построили эти показатели, что, сложив, получим ноль (с точностью до погрешности округления – среднее арифметическое величина не обязательно целая). Чтобы избежать этого обнуления, нам надо, чтобы каждое отклонение от среднего арифметического «лишилось» своего знака. Для этого возводят каждую величину в квадрат и лишь затем суммируют весь ряд значений.

Чтобы не зависеть от периода осреднения, делят полученную сумму квадратов на число слагаемых (в нашем случае по-прежнему на семь). Такая величина называется *дисперсией*.

Торговая точка	Дисперсия, р. ²	Среднее квадратическое отклонение, р.
Палатка 1	295522	543,6
Палатка 1	27633	166,2
Палатки 1+2	161938	402,4

Мы видим, что дисперсия действительно очень показательная величина. У «Палатки выходного дня» она выше более чем в десять раз. Дисперсию можно посчитать в *Excel* автоматически, даже не считая предварительно среднее арифметическое, программа сделает это сама. Для этого, находясь в файле *Excel*, нажмите в верхнем меню кнопку f_x . Затем выберите среди функций тип «СТАТИСТИЧЕСКИЕ» и из предложенного перечня в окошке – ДИСПРА. Затем, по подсказке, поставив курсор в поле «Число 1», проведите мышью вдоль строки с набранными значениями. Этот вид подсчета называется «вычисление смещенной дисперсии по генеральной совокупности».

Дисперсией часто пользуются, но более удобная характеристика носит название *среднее квадратическое отклонение* (обычно обозначается греческой буквой омега ω). Среднее квадратическое отклонение – это квадратный корень из дисперсии, он удобен тем, что имеет ту же размерность, что и исходные величины. Так, в нашем случае дисперсия имела бы размерность «рубли в квадрате», в то время как среднее квадратическое отклонение получается просто и привычно, в рублях. В нашем примере видно, что суммарная дисперсия и среднее квадратическое отклонение у двух палаток вместе все-таки выше, чем у одной первой палатки, причем среднее квадратическое отклонение выше более чем в два раза. Значит, наша гипотеза о «повышенной стабильности суммы» за счет присутствия второй палатки несостоятельна.

Иногда вместо среднего арифметического употребляют другие характерные величины, если это по каким-то причинам лучше описывает выборку. Так, если расставить выборку по возрастанию (или убыванию) той

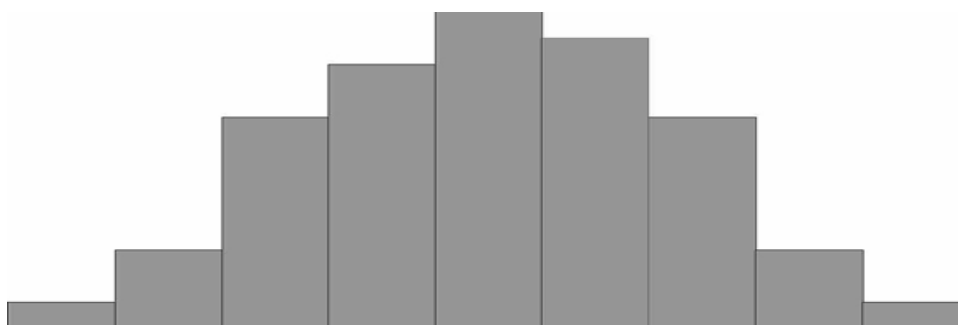
величины, которой мы интересуемся, то *медиана* – это то, что будет ровно посередине «строя». Например, если мы расположим по порядку длительности интервалы времени: *секунда, минута, час, сутки и неделя*, то медианой будет *час*.

Еще одно понятие для замены среднего – *мода*. Само название позволяет легко запомнить это определение. Если мы выстроим по порядку все пары обуви на складе по размеру, то самый ходовой размер будет модой. Мода – это то, что непременно должны учитывать производители упаковок и фасовщики. Если бы большинство людей покупало за один раз стакан молока, молочные пакеты не были бы литровыми.

В следующем параграфе мы начнем работать со случайными величинами, имеющими нормальное распределение, и эти понятия нам снова встретятся.

8. НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Если выйти на улицу любого города и случайным образом выбранных прохожих спросить о том, какой у них рост, вес, возраст, доход, и т.п., а потом построить график любой из этих величин, например, роста... Но не будем спешить, а посмотрим, как можно построить такой график. Сначала мы просто запишем результаты своего исследования. Потом мы рассортируем всех людей по группам, так чтобы каждый попал в свой диапазон роста, например, «от 180 до 181 включительно». После этого мы должны посчитать количество людей в каждой подгруппе-диапазоне, это будет частота попадания роста жителей города в данный диапазон. Обычно эту часть удобно оформить в виде таблички. Если затем эти частоты построить по оси y , а диапазоны отложить по оси x , можно получить так называемую гистограмму, упорядоченный набор столбиков, ширина которых равна, в данном случае, одному сантиметру, а высота будет равна той частоте, которая соответствует каждому диапазону роста. Если вам попалось достаточно много жителей, то ваша схема будет выглядеть примерно так.



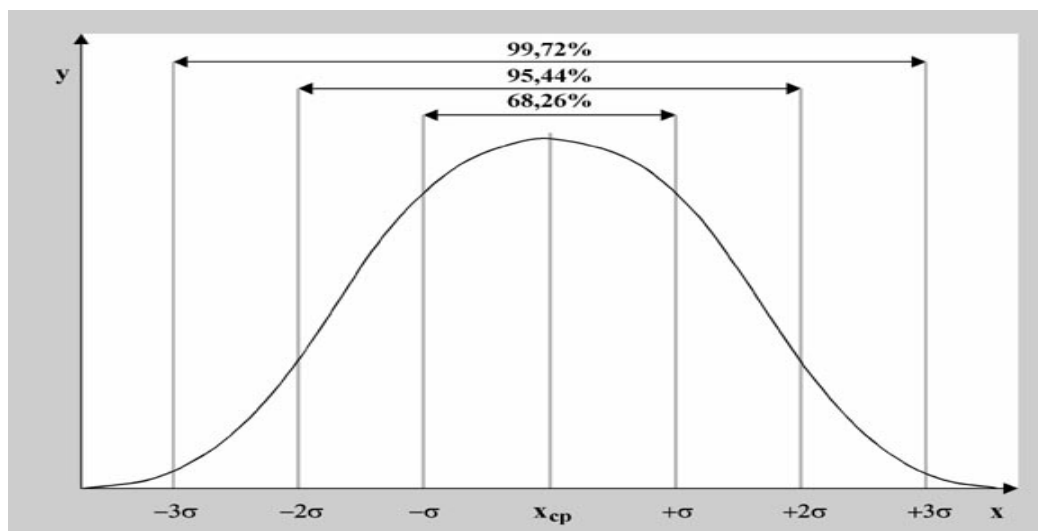
Дальше можно уточнить задачу. Каждый диапазон разбить на десять, жителей рассортировать по росту с точностью до миллиметра. Диаграмма станет глаже, но уменьшится по высоте, «оплывет» вниз, так как в каждом маленьком диапазоне количество жителей уменьшается. Чтобы избежать этого, просто увеличим масштаб по вертикальной оси в 10 раз. Если гипотетически повторить эту процедуру несколько раз, будет вырисовываться та знаменитая колоколообразная фигура, которая характерна для нормального (или Гауссова) распределения. В результате относительная частота встречаемости каждого конкретного диапазона роста может быть посчитана как отношение площади «ломтика» под кривой, приходящегося на этот диапазон, к площади подо всей кривой. Стандартизированные кривые нормального распределения, значения функций которых приводятся в таблицах книг по статистике, всегда имеют суммарную площадь под кривой, равную единице. Это связано с тем, что, как вы помните из курса теории вероятностей, вероятность достоверного события всегда равна 100% (или единице), а для любого человека иметь хоть какое-то значение роста – достоверное событие. А вот вероятность того, что рост произвольного человека попадет в определенный выбранный нами диапазон, будет зависеть от трех факторов.

Во-первых, от величины такого диапазона – чем точнее наши требования, тем меньше вероятность, что нам повезет. Во-вторых, от того, насколько «популярен» выбранный нами рост. Напомним, что мода – самое часто встречающееся значение роста. Кстати, для нормального распределения мода, медиана и среднее значение совпадают. Кривая нормального распределения симметрична относительно среднего значения. И, в-третьих, вероятность попадания роста в определенный диапазон зависит от характеристики рассеивания случайной величины. Отчасти это связано с единицами измерения (представьте, что мы бы измеряли людей в дюймах, а не в миллиметрах, но сами люди и их рост были бы теми же). Но дело не только в этом. Просто, некоторые процессы кучнее группируются возле среднего значения, в то время как другие более разбросаны. Например, рост собак и рост домашних кошек имеют разный разброс значений, их кривые нормального распределения будут выглядеть по-разному (напомним еще раз, что площадь под обеими кривыми будет единичной). Так, кривая для роста кошек будет более узкой и высокой, а для роста собак кривая будет ниже и шире.

Для характеристики разброса конечного ряда данных в предыдущем разделе мы использовали величину среднего квадратического отклонения. Аналогичная величина используется для характеристики кривой нормаль-

ного распределения. Она обозначается буквой s и называется в этом случае стандартным отклонением. Это очень важная величина для кривой нормального распределения. Кривая нормального распределения полностью задана, если известно среднее значение X_{cp} и отклонение s .

Кроме того, любой житель города с вероятностью 68% попадет в диапазон роста $X_{cp} \pm s$, с вероятностью 95% – в диапазон $X_{cp} \pm 2s$, и с вероятностью 99,7% – в диапазон $X_{cp} \pm 3s$.



Для вычисления других значений вероятности, которые могут вам понадобиться, можно воспользоваться приведенной таблицей.

После этого раздела вам нужно будет выполнить второе письменное задание.

ЗАДАНИЕ 2

В городе модельное агентство приглашает на работу красивых девушек от 18 до 25 лет, имеющих рост не ниже 1,75 метра, с местной пропиской. Можно считать, что распределение роста и возраста среди жителей города подчиняется нормальному закону. Средний рост женщин данной возрастной категории считайте равным 166 см, отклонение 5 см. Средний возраст 33 года, отклонение 14 лет. Вероятность того, что девушка красива, составляет 10%.

Нарисуйте дерево вероятностей и ответьте, пожалуйста, на следующие вопросы:

1. Какова вероятность нужного роста?
2. Какова вероятность нужного возраста?
3. Какой процент жителей города может претендовать на работу в агентстве?

Таблица вероятности попадания случайной величины
в отмеченный (заштрихованный) диапазон

σ				
0,0	50,00	50,00	100,00	0,00
0,1	46,02	53,98	92,04	7,97
0,2	42,07	57,93	84,14	15,85
0,3	38,21	61,79	76,42	23,58
0,4	34,46	65,54	68,92	31,29
0,5	30,85	69,15	61,70	38,30
0,6	27,43	72,57	54,86	45,15
0,7	24,20	75,80	48,40	51,61
0,8	21,19	78,81	42,38	57,63
0,9	18,41	81,59	36,82	63,19
1,0	15,87	84,13	31,74	68,17
1,1	13,57	86,43	27,14	72,87
1,2	11,51	88,49	23,02	76,99
1,3	9,68	90,32	19,36	80,64
1,4	8,08	91,92	16,16	83,85
1,5	6,68	93,32	13,36	86,64
1,6	5,48	94,52	10,96	89,04
1,7	4,46	95,54	8,92	91,08
1,8	3,59	96,41	7,18	92,81
1,9	2,87	97,13	5,74	94,25
2,0	2,28	97,72	4,56	95,45
2,1	1,79	98,21	3,58	96,43
2,2	1,39	98,61	2,78	97,22
2,3	1,07	98,93	2,14	97,85
2,4	0,82	99,18	1,64	98,36
2,5	0,62	99,38	1,24	98,76
2,6	0,47	99,53	0,94	99,07
2,7	0,35	99,65	0,70	99,31
2,8	0,26	99,74	0,52	99,49
2,9	0,19	99,81	0,38	99,63
3,0	0,14	88,86	0,28	99,73

Пусть у нас имеется большое количество предметов с нормальным распределением некоторых характеристик (например, полный склад однотипных овощей, размер и вес которых варьируется). Вы хотите знать средние

характеристики всей партии товара, но у вас нет ни времени, ни желания измерять и взвешивать каждый овощ. Вы понимаете, что в этом нет необходимости. Сколько штук надо было бы взять на выборочную проверку?

Прежде чем дать несколько полезных для этой ситуации формул, напомним некоторые обозначения. Во-первых, если бы мы все-таки промерили весь склад овощей (это множество элементов называется генеральной совокупностью), то мы узнали бы со всей доступной нам точностью среднее значение веса всей партии. Назовем это среднее значение $X_{ср.ген}$ – генеральным средним. Мы уже знаем, что нормальное распределение определяется полностью, если известно его среднее значение и отклонение s . Правда, пока мы ни $X_{ср.ген}$, ни генеральной совокупности не знаем. Мы можем только взять некоторую выборку, замерить нужные нам значения и посчитать для этой выборки как среднее значение $X_{ср.выб}$, так и среднее квадратическое отклонение $S_{выб}$. Известно, что если наша выборочная проверка содержит большое количество элементов (обычно n больше 30), и они взяты действительно случайным образом, то s генеральной совокупности почти не будет отличаться от $S_{выб}$. Кроме того, для случая нормального распределения мы можем пользоваться следующими формулами:

с вероятностью 95%

$$X_{ср.ген} = X_{ср.выб} \pm 1,96 \frac{\sigma}{\sqrt{n}};$$

с вероятностью 99%

$$X_{ср.ген} = X_{ср.выб} \pm 2,58 \frac{\sigma}{\sqrt{n}};$$

в общем виде с вероятностью $P(t)$

$$X_{ср.ген} = X_{ср.выб} \pm t \frac{\sigma}{\sqrt{n}}.$$

Связь значения t со значением вероятности $P(t)$, с которой мы хотим знать доверительный интервал, можно взять из следующей таблицы.

$P(t)$	0,683	0,950	0,954	0,990	0,997
t	1,00	1,96	2,00	2,58	3,00

Таким образом, мы определили, в каком диапазоне находится среднее значение для генеральной совокупности (с данной вероятностью). Если у нас нет достаточно большой выборки, мы не можем утверждать, что генеральная совокупность имеет $s = S_{выб}$. Кроме того, в этом случае проблема-

точно близость выборки к нормальному распределению. В этом случае также пользуются $S_{выб}$ вместо s в формуле

$$X_{ср.ген} = X_{ср.выб} \pm t \frac{\sigma}{\sqrt{n}}.$$

Но значение t для фиксированной вероятности $P(t)$ будет зависеть от количества элементов в выборке n . Чем больше n , тем ближе будет полученный доверительный интервал к значению $X_{ср.ген}$. Значения t в этом случае берутся из другой таблицы (t -критерий Стьюдента), которую мы приводим ниже.

Значения t -критерия Стьюдента для вероятности 0,95 и 0,99

n	P		n	P	
	0,95	0,99		0,95	0,99
2	12,71	63,66	18	2,11	2,90
3	4,30	9,93	19	2,10	2,88
4	3,18	5,84	20	2,093	2,861
5	2,78	4,60	25	2,064	2,797
6	2,57	4,03	30	2,045	2,756
7	2,45	3,71	35	2,032	2,720
8	2,37	3,50	40	2,022	2,708
9	2,31	3,36	45	2,016	2,692
10	2,26	3,25	50	2,009	2,679
11	2,23	3,17	60	2,001	2,662
12	2,20	3,11	70	1,996	2,649
13	2,18	3,06	80	1,991	2,640
14	2,16	3,01	90	1,987	2,633
15	2,15	2,98	100	1,984	2,627
16	2,13	2,95	120	1,980	2,617
17	2,12	2,92	∞	1,960	2,576

Пример 8.1. Из работников фирмы случайным образом отобрано 30 чел. По выборке оказалось, что средняя зарплата (в месяц) составляет 10 тыс. р. при среднем квадратическом отклонении 3 тыс. р. С вероятностью 0,99 определить среднюю зарплату в фирме.

Решение. По условию имеем $n = 30$, $X_{cp.} = 10000$, $S = 3000$, $P = 0,99$. Для нахождения доверительного интервала воспользуемся формулой, соответствующей критерию Стьюдента. По таблице для $n = 30$ и $P = 0,99$ находим $t = 2,756$, следовательно,

$$30000 - 2,756 \cdot \frac{5000}{\sqrt{30}} < X_{cp.ген} < 30000 + 2,756 \cdot \frac{5000}{\sqrt{30}},$$

т.е. искомый доверительный интервал $27484 < X_{cp.ген} < 32516$.

Итак, с вероятностью 0,99 можно утверждать, что интервал (27484; 32516) содержит внутри себя среднюю зарплату в фирме. Мы надеемся, что вы будете пользоваться этим методом, при этом не обязательно, чтобы при вас каждый раз была таблица. Подсчеты можно проводить в *Excel* автоматически. Находясь в файле *Excel*, нажмите в верхнем меню кнопку f_x . Затем выберите среди функций тип «статистические» и из предложенного перечня в окошке – СТЬЮДРАСПОБР. Затем, по подсказке, поставив курсор в поле «вероятность», наберите значение обратной вероятности (в нашем случае вместо вероятности 0,95 надо набирать вероятность 0,05). Видимо, электронная таблица составлена так, что результат отвечает на вопрос, с какой вероятностью мы можем ошибиться. Аналогично в поле «степень свободы» введите значение $(n-1)$ для своей выборки.

9. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

Мы уже знаем, что нормальное распределение – особенное. Некоторые его свойства мы сможем использовать и для распределений, которые, строго говоря, нормальными не назовешь. Задача, которую мы рассмотрим в этом разделе, имеет чрезвычайно важное значение для бизнеса. Это задача о диагностировании тенденций к изменению показателей. Удобство использования нормального распределения некоторых случайных величин и особые возможности, которые закон нормального распределения предоставляет исследователю, породили ряд теорем, которые позволяют пользоваться этими свойствами, даже если генеральная совокупность представляет собой «не вполне нормальное распределение».

Центральная предельная теорема имеет несколько формулировок, мы не будем их здесь полностью приводить и доказывать. Для нас важно знать только то, что в большинстве случаев среднее арифметическое вы-

борки, взятой из генеральной совокупности (напомним, что это среднее арифметическое – тоже случайная величина), ложится на нормальное распределение гораздо лучше, чем исходная генеральная совокупность. Другими словами, если мы возьмем несколько выборок из генеральной совокупности, то средние арифметические величины этих выборок будут представлять собой новую случайную величину с практически нормальным распределением. Именно эта теорема и позволит нам проверять так называемые статистические гипотезы, т.е. делать заключение о наличии тенденции к изменению показателей деятельности, которые сами по себе, являясь случайными величинами, имеют право на некоторый разброс.

Пример 9.1. Фирма поместила информацию о своей продукции в каталоге. Был указан один из двух номеров телефона отдела продаж, на который и раньше поступали звонки потенциальных покупателей. Другой номер телефона в каталоге не упоминался. За два месяца до выхода каталога и в течение двух месяцев после было зарегистрировано определенное количество звонков на эти телефоны (см. таблицу). Как нам определить, подействовала ли информация, данная в каталоге, или мы имеем дело со случайным оживлением на рынке, а деньги на рекламу потрачены напрасно?

Телефон	До рекламы	После рекламы	Ожидаемая величина (для сравнения)
Из каталога	216	305	274,5
Другой	142	150	180,5
Всего	358	455	455

Решение. Последний столбец в таблице – ожидаемые величины. Это наши оценки, сделанные из предположения, что ничего не изменилось и реклама не оказала никакого действия, т.е. произошло общее оживление на рынке и больше ничего, а пропорции между числом звонков на оба телефона должны сохраниться в точности. {Ожидаемая величина для телефона из каталога}=455 216/358=274,5 {Ожидаемая величина для другого телефона}=455 142/358=180,5. Наше предположение, о том, что реклама не оказала никакого воздействия на изменение числа покупателей, носит название *нулевой гипотезы*. *Альтернативная гипотеза* заключается в предположении о наличии такого влияния. Наша задача – выбрать более *достоверную* из этих двух гипотез. Чтобы оценить, насколько значимы отклонения реальной ситуации от ожидания по нулевой гипотезе, для обоих телефонов мы должны посчитать величины

$$\chi^2 = \left[\frac{(\text{Полученное значение} - \text{Ожидаемая величина})^2}{\text{Ожидаемая величина}} \right],$$

поставить их в таблицу и просуммировать.

Телефон	До рекламы	После рекламы	Ожидаемая величина (для сравнения)	χ^2
Из каталога	216	305	274,5	3,38
Другой	142	150	180,5	5,15
Всего	358	455	455	8,53

Дальнейшие наши действия – определить, с какой вероятностью посчитанные отклонения «ложатся» на соответствующую кривую. Для такой оценки можно воспользоваться значениями так называемого χ^2 -критерия Пирсона. Обычно эти значения задаются в виде стандартных таблиц в книгах по статистике (см. ниже).

Теперь несколько слов о том, как пользоваться этой таблицей. Буквы *d.f.* означают число степеней свободы. Чтобы посчитать степени свободы, нужно просто взять в таблице с исходными данными количество строк n и столбцов m и посчитать величину $(n-1)(m-1)$. Это и будет количество степеней свободы в каждом конкретном случае. При этом строки и столбцы берутся только для самих исходных данных, ни строка суммирования (всего), ни столбец подсчета ожидаемых величин при определении степени свободы не учитывается. В нашем случае $d.f. = (2-1)(2-1) = 1$, это означает, что степень свободы равна 1, и в таблице χ^2 мы должны пользоваться соответствующей строкой (верхней). Теперь о столбцах этой таблицы. Цифры 0,99; 0,95 и т.д. означают, что величины отклонений χ^2 , стоящие в этих столбцах с вероятностью 0,99; 0,95 и т.д., возникли случайно. В нашем примере вероятность случайного происхождения отклонения составляет менее 0,01 (меньше одного шанса из ста!). Мы вполне можем считать, что реклама оказала воздействие. Обратите внимание, что критерий χ^2 не говорит категорически, что случайность тут невозможна, просто вероятность этого очень мала. Другими словами, если мы отбросим нулевую гипотезу и выберем альтернативную, то вероятность ошибки будет меньше одного процента.

Если вы будете пользоваться этим методом, совсем не нужно считать каждый раз вручную все отклонения. Подсчеты можно проводить в *Excel* автоматически. Сначала запишите известные вам показатели в виде таблицы. Затем посчитайте в *Excel* столбец ожидаемых величин. После этого нажмите в верхнем меню кнопку f_x . Затем выберите среди функций тип «статистические» и из предложенного перечня в окошке – ХИ2ТЕСТ. Затем, по подсказке, поставив курсор в поле «ожидаемый интервал», выделите мышью столбец ожидаемых значений (но не захватывайте сумму в нижней строке).

d.f.	P					
	0,99	0,95	0,90	0,10	0,05	0,01
1	0,0002	0,004	0,02	2,71	3,84	6,64
2	0,02	0,10	0,21	4,61	5,99	9,21
3	0,12	0,35	0,58	6,25	7,82	11,34
4	0,30	0,71	1,06	7,78	9,49	13,28
5	0,55	1,15	1,61	9,24	11,07	15,09
6	0,87	1,64	2,20	10,65	12,59	16,81
7	1,24	2,17	2,83	12,02	14,06	18,48
8	1,65	2,73	3,49	13,36	15,51	20,09
9	2,09	3,33	4,17	14,68	16,92	21,67
10	2,56	3,94	4,87	15,99	18,31	23,21
11	3,05	4,58	5,58	17,28	19,68	24,72
12	3,57	5,23	6,30	18,55	21,03	26,22
13	4,11	5,89	7,04	19,81	22,36	27,68
14	4,66	6,57	7,79	21,06	23,69	29,14
15	5,23	7,26	8,55	22,31	25,00	30,58
16	5,81	7,96	9,31	23,54	26,30	32,00
17	6,41	8,67	10,09	24,77	27,59	33,41
18	7,02	9,39	10,86	25,99	28,87	34,81
19	7,63	10,12	11,65	27,20	30,14	36,19
20	8,26	10,85	12,44	28,41	31,41	37,57
21	8,90	11,59	13,24	29,62	32,67	38,93
22	9,54	12,34	14,04	30,81	33,92	40,29
23	10,20	13,09	14,85	32,01	35,17	41,64
24	10,86	13,85	15,66	33,19	36,42	43,98
25	11,52	14,61	16,47	34,38	37,65	44,31
26	12,20	15,37	17,29	35,56	38,89	45,64
27	12,88	16,15	18,11	36,74	40,11	46,96
28	13,56	16,93	18,94	37,92	41,34	48,28
29	14,26	17,71	19,77	39,09	42,56	49,59
30	14,95	18,49	20,60	40,26	43,77	50,89
40	22,16	26,51	29,05	51,81	55,76	63,69
50	29,71	34,76	37,69	63,17	67,51	76,15
100	70,07	77,93	82,36	118,50	124,34	135,81

Аналогично в поле «фактический интервал» введите массив из столбика фактических данных после рекламы. Программа сама посчитает граничную вероятность того, что отклонение было случайным. Так в нашем варианте более точное значение вероятности составляет примерно 0,0035. В таблице мы попали по значению χ^2 между столбцами и посчитать вероятность с такой точностью не смогли. Видимо, для того чтобы вы привыкли пользоваться подобными оценками, имеет смысл обсудить вопрос о «степени свободы». Что это такое и какие степени свободы вообще могут быть? Понятно, что оценка значимости происходящих изменений может происходить только при наличии данных, как полученных при гипотетическом воздействии этих изменений, так и свободных от изменений. В качестве заведомо не подверженных изменениям данных в нашем примере выступали показания числа звонков на оба телефона до публикации каталога. Кроме того, для дополнительной объективности данных мы использовали один телефон как неизвестный в рекламе. Это позволило нам исключить возможное влияние сезонных изменений спроса или другие подобные факторы.

В других ситуациях мы можем сравнивать динамику спроса на один товар с динамикой спроса на другой, если идет целевая раскрутка этого товара или же товар входит в моду. И в этой ситуации свойства нормального распределения помогут нам сделать вывод о значимости происходящих изменений.

После этого раздела вам нужно будет выполнить третье письменное задание.

ЗАДАНИЕ 3

В одной и той же торговой палатке, сменяясь через неделю, работают два разных продавца (А и В). Таблица со значениями недельной выручки (тыс. р.) приводится за 8 последних недель.

Неделя	1	2	3	4	5	6	7	8
Продавец А	119		93		89		94	
Продавец В		132		91		102		105

Ответьте, пожалуйста, на следующие вопросы:

1. Какой продавец работал лучше с точки зрения суммарной выручки?
2. Можно ли утверждать, что это отклонение было случайным?
3. С какой достоверностью сделано ваше утверждение?
4. Какая нулевая и альтернативная гипотезы вами были приняты?
5. Какая степень свободы содержится в приведенных данных?
6. Прокомментируйте результаты своих количественных оценок и оформите их в виде таблицы.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Вентцель Е.С.* Теория вероятностей. – М.: Высшая школа, 2005.
2. *Гмурман В.Е.* Руководство к решению задач по теории вероятностей и математический статистике. – М.: Высшее образование, 2009.
3. *Гмурман В.Е.* Теория вероятностей и математическая статистика. – М.: Высшее образование, 2009.
4. *Кремер Н.Ш.* Теория вероятностей и математическая статистика. – М.: ЮНИТИ-ДАНА, 2007.
5. *Теория вероятностей и математическая статистика* / под ред. В.И. Ермакова. – М.: ИНФРА-М, 2004.

Учебное издание

Коржавина Наталья Валерьевна,
Петров Николай Петрович,
Петрова Светлана Николаевна

**ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА
ДЛЯ ФИНАНСИСТОВ И МЕНЕДЖЕРОВ**

Учебное пособие

Редактор и корректор *Л. В. Матвеева*

Поз. 103. Подписано в печать 09.09.2014.
Формат бумаги 60 × 84 1/16. Гарнитура Times New Roman. Бумага офсетная.
Печать плоская. Уч.-изд. л. 1,8. Усл. печ. л. 2,44. Заказ 547. Тираж 180 экз.
Издательство Уральского государственного экономического университета
620144, г. Екатеринбург, ул. 8 Марта/Народной Воли, д. 62/45

Отпечатано с готового оригинал-макета
в подразделении оперативной полиграфии
Уральского государственного экономического университета