

Министерство образования и науки Российской Федерации
Ярославский государственный университет им. П. Г. Демидова

В. С. Кузнецов
А. Я. Пархоменко
А. В. Проказников

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебное пособие

Рекомендовано

*Научно-методическим советом университета для студентов,
обучающихся по направлениям «Физика», «Радиофизика»,
«Электроника и наноэлектроника», «Радиотехника»,
«Инфокоммуникационные технологии и системы связи»*

Ярославль
ЯрГУ
2014

УДК 51:33(075)
ББК В171я73+В172я73
К89

Рекомендовано
*Редакционно-издательским советом университета
в качестве учебного издания. План 2014 года*

Рецензенты:

И. И. Амиров, доктор физико-математических наук;
кафедра теоретической механики ФГБОУ ВПО ЯГТУ

Кузнецов, Владимир Степанович

К89 Теория вероятностей и математическая статистика : учеб. пособие / В. С. Кузнецов, А. Я. Пархоменко, А. В. Проказников ; Яросл. гос. ун-т им. П. Г. Демидова. — Ярославль : ЯрГУ, 2014. — 100 с.

ISBN 978-5-8397-1034-4

Основы теории вероятностей и математической статистики необходимы студентам физических специальностей для качественной обработки экспериментальных результатов и при изучении таких теоретических дисциплин, как макроскопическая электродинамика, квантовая механика, статистическая физика, физическая кинетика и теория конденсированного состояния вещества.

Пособие предназначено для студентов, обучающихся по направлениям 011200.62 Физика, 011800.62 Радиофизика, 210100.62 Электроника и наноэлектроника, 210400.62 Радиотехника, 210700.62 Инфокоммуникационные технологии и системы связи (дисциплина «Теория вероятностей и математическая статистика», цикл Б2), очной формы обучения.

УДК 51:33(075)
ББК В171я73+В172я73

ISBN 978-5-8397-1034-4

©ЯрГУ, 2014

Введение

Теория вероятностей, как и другие разделы математики, имеет дело не с самими явлениями окружающего мира, а с их математическими моделями. В математической модели должны быть правильно переданы существенные стороны изучаемого явления, а несущественные — отброшены. Слишком подробное описание изучаемого явления приводит к усложнению математической модели и может значительно затруднить дальнейшее исследование. Излишнее упрощение модели может привести к неверным выводам. Насколько удачно введена модель, можно в каждом конкретном случае судить по согласованию теоретических выводов с опытом.

С подобными задачами мы часто сталкиваемся в квантовой механике, электродинамике, статистической физике и физической кинетике.

Так, например, в объеме $V = 1 \text{ см}^3$ твердого тела находится порядка $N = 10^{22}$ атомов, для строгого описания движения этих атомов в каждый момент времени нужно задавать 3×10^{22} импульсов и столько же координат. Состояние системы из N атомов, определяемое $3N$ импульсами и $3N$ координатами, называется *микроскопическим состоянием*.

В макроскопической теории состояние такой системы обычно характеризуется небольшим числом макроскопических параметров: температура T , объем V , давление p , масса системы m и так далее, причем число этих параметров значительно меньше $6N$. Состояние системы, определяемое этими параметрами, называется *макроскопическим состоянием*.

Очевидно, что одному макросостоянию соответствует несколько микросостояний. Иными словами, переход от микроскопического к макроскопическому описанию системы всегда сопровождается некоторой потерей информации и приобретением коллективных, синергетических свойств. Для вычисления макроскопических параметров на основании поведения атомов, молекул и полей в системе нам вполне достаточно знать только вероятность нахождения системы в каждом микросостоянии и уметь манипулировать с ней. Таким образом, статистическая физика и физическая кинетика полностью основаны на использовании понятий и методов теории вероятности.

Вероятностный подход является основой квантовой механики.

1. Понятие вероятности

К сожалению, специалисты по теории вероятностей не сформулировали пока единого определения понятия вероятности, удобного для практического использования и математически строгого. Математический анализ показал, что используемые на практике формулировки этого фундаментального понятия имеют разную степень надежности. Однако при выполнении определенных условий знание области применимости этих формулировок дает возможность получать правильные результаты.

1.1. «Комбинаторное» определение

Если событие может приводить к N равновозможным различным исходам и если в n случаях появляется признак A , то вероятность A есть n/N . Такое определение можно подвергнуть критике на том основании, что выражение «равновозможно» на самом деле означает «равновероятно» и в рассуждениях, таким образом, содержится порочный круг.

Пример 1.1. Пусть совершенно случайным образом отобраны 23 человека. Какова вероятность того, что по крайней мере у двоих окажется один и тот же день рождения?

Полное число возможных комбинаций из 23 дней рождения, считая, что в году 365 дней, равно: $N = 365^{23}$. Число комбинаций, в которых имеется 23 различных дня рождения, равно $N - n = 365 \times 364 \times 363 \times 362 \times \dots \times 344 \times 343$. Вероятность того, что все дни рождения различны, равна

$$\frac{N - n}{N} = \frac{365 \times 364 \times 363 \times \dots \times 343}{365^{23}}.$$

Вероятность того, что не все дни рождения различны, равна

$$\frac{n}{N} = 1 - \frac{N - n}{N} = 1 - 1 \left(1 - \frac{1}{365}\right) \left(1 - \frac{2}{365}\right) \dots \left(1 - \frac{22}{365}\right) \approx 0,5072972.$$

Таким образом, для группы, состоящей из 23 или более человек, вероятность того, что по крайней мере у двоих окажется общий день рождения, превышает 0,5. Для группы из 25 человек эта вероятность равна 0,5686997. Для группы из 50 человек вероятность равна 0,9703736. Для группы из 57 человек — 0,9901225, из 60 человек — 0,9901225.

Как правило, описание вероятностей, основанное на комбинаторном определении, сводится к комбинаторным вычислениям. Рассмотрим некоторые

примеры, позволяющие понять технику вычисления вероятностей. Напомним вначале простейшие комбинаторные формулы.

1.1.1. Правило суммы

Если X и Y — два не пересекающихся конечных множества, то число элементов, содержащихся в объединении этих множеств, равно сумме чисел элементов в каждом из них.

Условимся обозначать число элементов конечного множества X через $|X|$, тогда число элементов объединения X и Y равно: $|X \cup Y| = |X| + |Y|$.

Обычно правило суммы формулируется несколько иначе: если некоторый объект X можно выбрать n способами, а объект Y — m способами, причем любой способ выбора объектов X отличен от любого способа выбора Y , то выбор « X или Y » можно сделать $n + m$ способами.

1.1.2. Правило произведения

Будем рассматривать последовательности элементов x_1, x_2, \dots, x_k (не обязательно различных) некоторой длины k : (x_1, x_2, \dots, x_k) .

Условимся для краткости называть такие последовательности *строками*. Две строки (x_1, x_2, \dots, x_k) и (y_1, y_2, \dots, y_k) будем считать различными в том и только в том случае, если хотя бы для одного номера i (из совокупности $1, 2, \dots, k$) элемент x_i отличен от y_i .

Правило произведения может быть сформулировано следующим образом: пусть элемент x_1 может быть выбран n_1 способами; при каждом выборе x_1 элемент x_2 может быть выбран n_2 способами; при каждом выборе (x_1, x_2) элемент x_3 может быть выбран n_3 способами и т. д.; наконец, при каждом выборе $(x_1, x_2, \dots, x_{k-1})$ элемент x_k может быть выбран n_k способами. Тогда число различных строк (x_1, x_2, \dots, x_k) равно произведению $n_1 \times n_2 \times \dots \times n_k$.

1.1.3. Размещения

Часто возникает необходимость в подсчете числа строк той или иной длины, составленных из элементов данного множества X .

Вот одна из подобных задач. Пусть множество X состоит из n элементов. Рассмотрим строки (x_1, x_2, \dots, x_k) длиной k при условии, что все элементы x_1, x_2, \dots, x_k различны. Разумеется, что все это возможно, если $k \leq n$.

Любая строка такого рода называется размещением без повторения из n элементов по k или просто размещением из n элементов по k . Число всех размещений из n элементов по k обозначается A_n^k .

Чтобы найти число A_n^k , заметим, что для выбора элемента x_1 имеется n возможностей, если x_1 выбран, то для x_2 имеется $n - 1$ возможность, если выбраны x_1, x_2 , то для выбора x_3 имеется $n - 2$ возможности, и т. д. Пользуясь правилом произведения, находим отсюда, что

$$A_n^k = n \times (n - 1) \times (n - 2) \times \dots \times (n - (k - 1)) = \frac{n!}{(n - k)!}.$$

Особо важным случаем является $k = n$. Строки без повторения, составленные из всех элементов n -членного множества X , называются перестановками из n элементов и обозначаются P_n .

$$P_n = n \times (n - 1) \times (n - 2) \times \dots \times 2 \times 1 = n!$$

Добавим к этому равенство $0! \equiv 1$, которое принимается по определению.

1.1.4. Сочетания

Пусть X — множество, состоящее из n элементов. Любое подмножество Y множества X , содержащее k элементов, рассматриваемое без учета их порядка, называется сочетанием k элементов из n . Общее число таких сочетаний обозначается через C_n^k или $\binom{n}{k}$; здесь обязательно $k \leq n$.

Число сочетаний из n элементов по k равно

$$C_n^k = \binom{n}{k} = \frac{A_n^k}{k!} = \frac{n!}{k! (n - k)!}$$

Свойства:

- а) $C_n^{n-k} = n! / [(n - k)! k!] = C_n^k$;
- б) правило Паскаля $C_{n+1}^k = C_n^{k-1} + C_n^k$ для $1 \leq k \leq n$;
- в) $C_n^0 + C_n^1 + \dots + C_n^n = 2^n$.

Бином Ньютона

Из школьного курса известны формулы: $(a + b)^2 = a^2 + 2ab + b^2$ и $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$. Их обобщением является следующая формула, называемая обычно *биномом Ньютона*:

$$(a + b)^n = C_n^0 a^0 b^n + C_n^1 a^1 b^{n-1} + \dots + C_n^{n-1} a^{n-1} b^1 + C_n^n a^n b^0 = \sum_{k=0}^n C_n^k a^k b^{n-k}.$$

Обобщение предыдущей формулы

$$(a_1 + a_2 + a_3 + \dots + a_m)^n = \sum_{k_1, k_2, \dots, k_{m-1}, k_m=1}^n \frac{n!}{k_1! k_2! \dots k_m!} a_1^{k_1} a_2^{k_2} \dots a_m^{k_m},$$

при условии, что $k_1 + k_2 + \dots + k_{m-1} + k_m = n$.

1.1.5. Гипергеометрическое распределение

Пусть дана совокупность n элементов, среди которых k отмеченных (например, бракованных изделий, выигрышных билетов и т. д.). Выбирается наугад $m \leq n$ объектов. Какова вероятность того, что среди них окажется l отмеченных?

Выбирать m объектов из n можно C_n^m различными способами. l отмеченных из полного их числа k можно выбрать C_k^l способами, причем каждому такому способу соответствует C_{n-k}^{m-l} способов, добрать еще $m - l$ объектов из общего числа m , выбирать их из $n - k$ не отмеченных.

Следовательно, число способов, благоприятствующих появлению l отмеченных объектов среди m выбранных, равно $C_k^l C_{n-k}^{m-l}$ согласно теореме умножения. Поэтому искомая вероятность равна:

$$P_{n,m}(k, l) = \frac{C_k^l C_{n-k}^{m-l}}{C_n^m}.$$

Полученная вероятность называется *гипергеометрическим распределением*.

Пример 1.2. Какова вероятность выигрыша при игре в «Спортлото»?

В данном случае $n = 49$ (число наименований видов спорта), $k = 6$ (число отмеченных видов спорта), $m = 6$ (число выбранных видов спорта). Следовательно, вероятность угадывания l видов спорта равна:

$$P_{49,6}(6, l) = \frac{C_6^l C_{49-6}^{6-l}}{C_{49}^6}.$$

В том числе вероятность угадывания всех шести видов спорта — $P_{49,6}(6, 6) = 1/C_{49}^6 = 7.1511234 \times 10^{-8}$, пяти — $P_{49,6}(6, 5) = C_6^5 C_{43}^1 / C_{49}^6 = 1.844990 \times 10^{-5}$, четырех — $P_{49,6}(6, 4) = C_6^4 C_{43}^2 / C_{49}^6 = 9.686196 \times 10^{-4}$, трех — $P_{49,6}(6, 3) = C_6^3 C_{43}^3 / C_{49}^6 = 1.765040 \times 10^{-2}$.

Рассмотрим далее несколько важных задач на размещения, часто возникающих при изучении некоторых систем в квантовой механике и статистической физике.

Система Максвелла—Больцмана характеризуется как система k различных частиц, каждая из которых может находиться в одном из состояний вне зависимости от того, где при этом находятся остальные частицы. В такой системе возможны всего n^k различных размещений. Если при этом все такие размещения (состояния системы) равновероятны, то говорят о статистике Максвелла—Больцмана. Вероятность каждого состояния равна

$$P = n^{-k}.$$

Система Бозе—Эйнштейна определяется как система k тождественных частиц, каждая из которых независимо от остальных может находиться в одном из n состояний. Рассмотрим ящик, содержащий n ячеек и k частиц.

	*			**		**		* *	* *
						**			

Распределение k частиц по n ячейкам можно трактовать как всевозможные перемешивания системы из $n - 1$ перегородок и k частиц при условии, что все перегородки тождественны друг другу и все частицы также тождественны друг другу. Это будет число сочетаний из $n - 1 + k$ объектов по $n - 1$:

$$\frac{(n - 1 + k)!}{(n - 1)! k!} = C_{n-1+k}^{n-1} = C_{n-1+k}^k.$$

Если все состояния равновероятны, то говорят о статистике Бозе—Эйнштейна. Вероятность каждого состояния:

$$P = 1/C_{n-1+k}^k.$$

Система Ферми—Дирака определяется как система Бозе—Эйнштейна, в которой дополнительно действует принцип запрета (принцип Паули), требующий, чтобы в каждой ячейке находилось не более одной частицы. Занятую ячейку можно рассматривать как черную, пустую — как белую,

*		*	*			*		*	
---	--	---	---	--	--	---	--	---	--

тогда число разбрасываний k частиц по n ячейкам будет равно числу сочетаний C_n^k из k по n и вероятность каждого состояния

$$P = 1/C_n^k.$$

1.2. Частотное определение

Допустим, что при N числе измерений n_A раз наблюдают появление признака A . Если исходные события в этой последовательности взаимно независимы, то вероятность признака A определяется как

$$P(A) = \lim_{N \rightarrow \infty} \frac{n_A}{N}.$$

Иными словами, *вероятность события — число, к которому приближается частота наступления этого события с увеличением числа опытов.*

1.3. Геометрическое определение

Интуитивное представление о вероятности может быть составлено также в связи со следующим мысленным экспериментом. Пусть на отрезок $[a, b]$ длины $\ell = b - a$ наугад бросается точка. Какова вероятность того, что она попадает на отрезок $[\alpha, \beta]$, содержащийся в $[a, b]$?

Ответ в данном случае очевиден. Поскольку вероятность попасть в $[\alpha, \beta]$ не зависит от того, где именно на $[a, b]$ расположен отрезок $[\alpha, \beta]$, то искомая вероятность равна:

$$P = \frac{\beta - \alpha}{b - a}.$$

Контрольные вопросы

1. В чем смысл «комбинаторного» определения вероятности?
2. Объяснить правила сумм и произведения множеств.
3. Раскрыть специфику понятий «размещения», «перестановки» и «сочетания».
4. Привести примеры событий, описываемых полиномиальным распределением и гипергеометрическим распределением.
5. Дать определения систем, имеющих статистики Максвелла—Больцмана, Бозе—Эйнштейна и Ферми—Дирака.
6. Указать, в чём смысл частотного и геометрического определений вероятности.

2. Классическая теоретико-множественная модель

2.1. Алгебра событий

Итак, в полном соответствии с предложенными определениями вероятности мы принимаем для нее такое определение: *вероятность случайного события* — это связанное с данным событием положительное число, около которого колеблется частота наступления этого события в данных сериях опыта.

Такое определение вероятности далеко от совершенства в математическом смысле слова, но оно удобно в практическом решении задач и позволяет делать некоторые обобщения.

2.1.1. Комбинации событий

При нахождении вероятностей необходимо учитывать связи между событиями. Естественно, что при проведении какого-нибудь измерения или наблюдения можно зафиксировать все случайные исходы этих измерений или наблюдений.

В общей теоретико-вероятностной схеме для каждого эксперимента со случайным исходом должны быть указаны все элементарные исходы, отвечающие следующим требованиям.

В результате эксперимента непременно происходит один и только один из этих исходов. Каждый такой исход принято называть элементарным событием, обозначать элементарные события будем буквой ω_i . По смыслу элементарные события неразложимы на «более элементарные». Множество всех элементарных событий в теории вероятностей принято называть *пространством элементарных событий*. Пространство элементарных событий будем обозначать буквой Ω , а сами элементарные события будут *точками пространства элементарных событий*.

Событие называют случайным, если при осуществлении определенной совокупности условий S оно может либо произойти, либо не произойти. В дальнейшем вместо того, чтобы говорить «совокупность условий S осуществлена», мы будем говорить кратко: «произведено испытание». Таким образом, мы будем рассматривать событие как результат испытаний. На практике одному событию (например, A) может соответствовать несколько элементарных событий, т. е. событие A можно считать подмножеством Ω , состоящим из точек $\omega \in \Omega$, представляющих те исходы эксперимента, при которых про-

исходит событие A . Случайные события обычно обозначаются латинскими буквами A, B, C, \dots . Событие в физическом опыте, которое описывается подмножеством A , происходит только в тех случаях, когда происходит какое-либо элементарное событие из A . Совокупность всех событий также образует множество Ω .

Пример 2.1. Стрелок стреляет по мишени, разделенной на четыре области. Выстрел — это испытание. Попадание в определенную область мишени — событие.

События называются несовместимыми, если появление одного из них исключает появление других событий в одном и том же испытании.

Пример 2.2. Из ящика с деталями наудачу извлекается деталь. Появление годной детали исключает появление дефектной детали. События «появилась годная деталь» и «появилась деталь с дефектом» — несовместимы.

События называются единственно возможными, если появление в результате испытаний одного и только одного из них является достоверным событием.

Пример 2.3. Стрелок произвел выстрел по цели. Обязательно произойдет одно из следующих двух событий: попадание или промах. Эти события единственно возможные.

События называются равновозможными, если есть основания считать, что ни одно из этих событий не является более возможным, чем другие.

Пример 2.4. Появление того или иного числа очков на брошенной игральной кости есть события равновозможные.

2.2. Операции над событиями

Рассмотрим теперь операции над событиями. Эти операции являются, по существу, операциями над множествами, и только терминология в ряде случаев носит вероятностный характер.

Для наглядности удобно изображать пространство элементарных исходов некоторой областью на поверхности (рис. 1). Тогда события будут изображаться частями этой области. Пусть A и B — некоторые события, определенные на данном пространстве элементарных событий.

1. Объединением событий A и B называется такое событие $C = A \cup B$, которое состоит из тех и только тех элементарных исходов, которые принадлежат по крайней мере одному из событий A или B .

Можно сказать, что в реальном опыте событие, соответствующее $A \cup B$,

состоит в том, что произошло: A или B . Событию $A \cup B$ соответствует затемненная область на рис. 2(а).

2. Произведением или пересечением событий A и B называется событие C , состоящее из элементарных событий, принадлежащих одновременно A и B , и обозначается $C = A \cap B \equiv A \times B$. Иными словами, событие $C = A \cap B$ происходит тогда и только тогда, когда происходят и A , и B . Событию C соответствует затемненная область на рис. 2(б).

3. Разностью $C = A \setminus B$ называется событие, состоящее из элементов множества A , не принадлежащих B , т. е. событие $A \setminus B$ состоит в том, что A произошло, а B не произошло. Событию C соответствует затемненная область на рис. 2(с).

Принято называть событие Ω *достоверным событием*, а пустое множество \emptyset — *невозможным*.

Событие $\Omega \setminus A \equiv \bar{A}$ называют противоположным событию A . Событие \bar{A} означает, что A не произошло (см. рис. 1).

Очевидно, что $\bar{\bar{A}} = A$ и

$$A \setminus B = A \cap \bar{B} = \bar{B} \setminus \bar{A}.$$

Кроме того, операции \cup и \cap являются событиями взаимно дистрибутивными:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C);$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

что можно легко проверить графически, воспользовавшись рис. 2(d) и 2(e). Наконец, в теории вероятностей и ее приложениях важную роль играет так называемый принцип двойственности (законы де Моргана), который может быть выражен следующими соотношениями:

$$\overline{A \cup B} = \bar{A} \cap \bar{B};$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B}.$$

Введем еще одно понятие. Если событие A происходит всякий раз, когда происходит событие B , то будем говорить, что событие A является следствием B , и писать $B \subset A$ или $A \supset B$ [см. рис. 2(f)].

Если $A \subset B$ и $B \subset A$, то события A и B происходят или не происходят лишь одновременно. В таком случае будем писать $A = B$, множества A и B при этом совпадают.

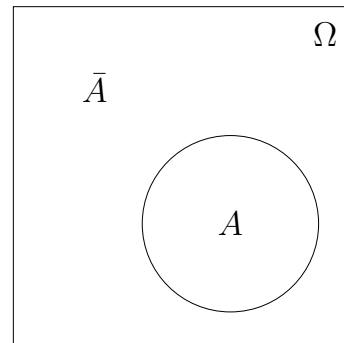


Рис. 1. Графическое представление события

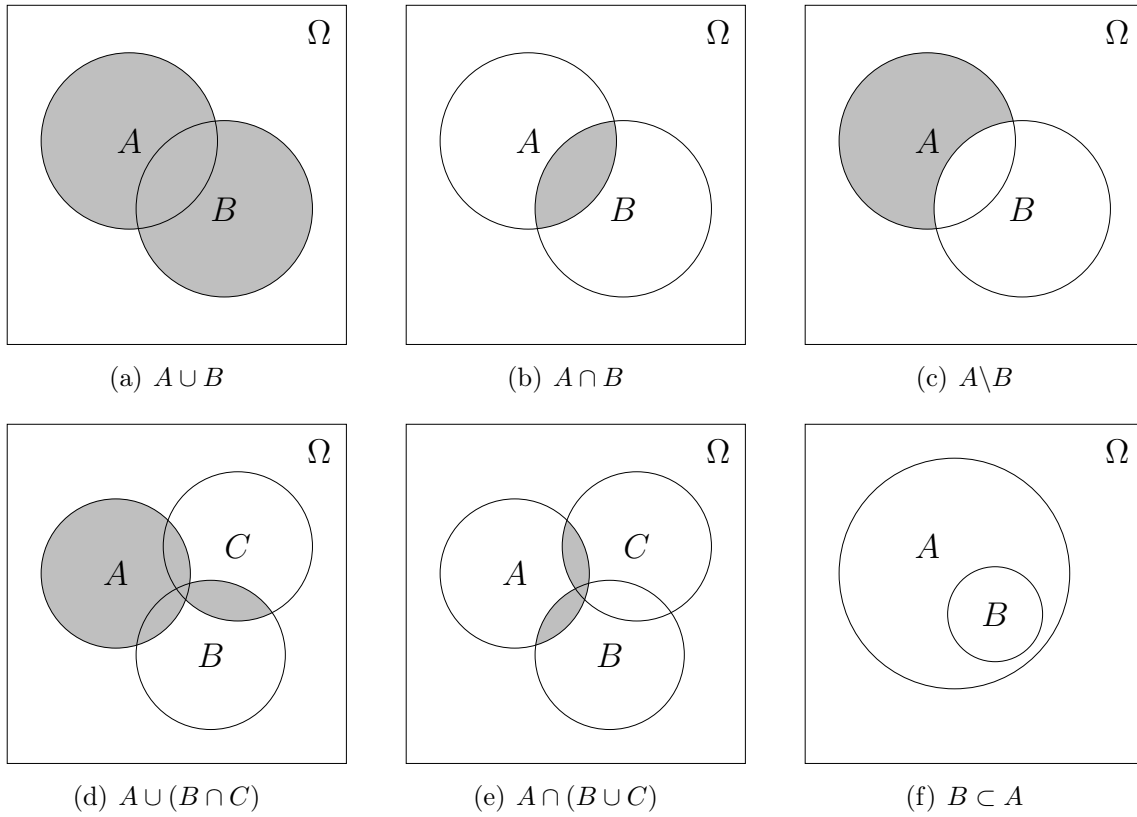


Рис. 2. Графическое представление операций над событиями

Рассмотренные свойства операций над событиями носят алгебраический характер. В теории вероятностей принимается, что из всех событий $A \in \Omega$ выбирается только такой класс \mathcal{F} событий, который удовлетворяет следующим требованиям:

1. Само пространство элементарных событий принадлежит \mathcal{F} ; $\Omega \in \mathcal{F}$.
2. Для каждой пары событий A и B из включения $A \in \mathcal{F}$ и $B \in \mathcal{F}$ следует включение $A \cap B \in \mathcal{F}$. Иными словами, класс вместе с каждой парой событий содержит их объединение.
3. Вместе с каждым событием A класс \mathcal{F} содержит противоположное событие $\bar{A} \in \mathcal{F}$. Отсюда следует:
 - а) $\emptyset \in \mathcal{F}$, поскольку $\emptyset = \bar{\Omega}$;
 - б) согласно принципу двойственности $\overline{A \cap B} = \bar{A} \cup \bar{B}$, и класс \mathcal{F} вместе с каждой парой событий A и B содержит их пересечения $A \cap B$ и разность $A \setminus B = A \cap \bar{B}$;
 - в) наименьшей системой подмножеств, образующих алгебру \mathcal{F} , очевидно, является система $\{\Omega, \emptyset\}$.

Класс множеств \mathcal{F} называется σ -алгеброй, или борелевским полем событий, если число событий $A \in \Omega$ счётно и выполняется условие:

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}.$$

2.3. Аксиоматическое построение теории вероятности

До сих пор мы исходили из представления о вероятности случайного события как о числе, которое вводили разными способами: комбинаторное, частотное и геометрическое определения. Все эти определения не обладают строгостью. Чтобы придать понятиям теории вероятностей строгость, присущую другим математическим понятиям, можно воспользоваться широко распространенным в математике аналитическим методом. Он состоит в том, что с самого начала фиксируются первичные, не подлежащие определению понятия данной теории. Их основные свойства формируются в виде ряда аксиом. После этого все положения теории выводятся из аксиом строго логическим путем, без обращения к посторонним понятиям типа «наглядность», «здравый смысл» и др. Современная аксиоматика теории вероятностей принадлежит советскому математику А. Н. Колмогорову.

Пусть дано пространство элементарных событий Ω и в нем некоторая система событий A , образующих алгебру событий. Вероятность на $\{\Omega, \mathcal{F}\}$ есть числовая функция, определенная на множествах A из \mathcal{F} и обладающая следующими свойствами:

A1. $P(A) \geq 0$ для любого $A \in \mathcal{F}$.

A2. $P(\Omega) = 1$.

A3. Если бесконечная последовательность событий $\{A_n\}$ такова, что все события попарно несовместны, т. е. $A_i \cap A_j = \emptyset$ при $i \neq j$, то

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

В широком смысле тройка $\{\Omega, \mathcal{F}, P\}$ называется вероятностным пространством. Если некоторая алгебра \mathcal{F} образует σ -алгебру, то аксиома **A3** будет выполняться автоматически и тройку $\{\Omega, \mathcal{F}, P\}$ называют просто вероятностным пространством. В заключение отметим, что введенная система аксиом является неполной и непротиворечивой.

2.3.1. Свойства вероятности

1. $P(\emptyset) = 0$, это следует из равенства $\emptyset \cup \Omega = \Omega$.
2. $P(\bar{A}) = 1 - P(A)$, так как $A \cup \bar{A} = \Omega$ и $A \cap \bar{A} = \emptyset$.
3. Если $A \subset B$, то $P(A) \leq P(B)$. Это следует из соотношения между вероятностями: $P(A) + P(\bar{A} \cap B) = P(B)$.
4. $P(A) \leq 1$.
5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, так как $A \cup B = A \cup (B \setminus (A \cap B))$, отсюда $P(A \cup B) = P(A) + P(B \setminus (A \cap B)) = P(A) + P(B) - P(A \cap B)$.
6. $P(A \cup B) \leq P(A) + P(B)$.
- 7.

$$\begin{aligned} P\left(\bigcup_{k=1}^n A_k\right) &= \sum_{k=1}^n P(A_k) - \sum_{i < k} P(A_i \cap A_k) + \sum_{j < i < k} P(A_j \cap A_i \cap A_k) - \dots \\ &+ (-1)^{n+1} P(A_1 \cap A_2 \cap \dots \cap A_n). \end{aligned}$$

8. Если A_n есть монотонно возрастающая последовательность множеств ($A_n \subset A_{n+1}$) и $A = \bigcup_{n=1}^{\infty} A_n$, то $P(A) = \lim_{n \rightarrow \infty} P(A_n)$.

2.4. Условная вероятность

Прежде чем переходить к формальному определению условной вероятности, введем ее на основании комбинаторного определения вероятности.

Предположим, что событие A может осуществляться n_A способами, событие B — n_B способами, а событие $A \cap B$ — n_{AB} способами, причем полное число всевозможных исходов равно N . Тогда вероятности: $P(A) = n_A/N$, $P(B) = n_B/N$, $P(A \cap B) = n_{AB}/N = (n_{AB}/n_A)(n_A/N) = (n_{AB}/n_A)P(A)$.

Из n_A случаев, в которых происходило событие A , в относительной доле, равной n_{AB}/n_A , происходит также и событие B .

Таким образом, n_{AB}/n_A есть вероятность B при условии, что произошло событие A . Эта вероятность записывается в виде $P(B|A)$ и называется условной вероятностью события B при условии, что произошло событие A .

Определение. Пусть задано вероятностное пространство $\{\Omega, \mathcal{F}, P\}$, в котором выбраны два произвольных события A и B . Если $P(B) > 0$, условная вероятность события A при условии, что произошло событие B , равна $P(A|B) = P(A \cap B)/P(B)$.

2.5. Независимые события

Определение. Два события A и B называются независимыми, если вероятность $P(A \cap B) = P(A)P(B)$.

Графически два независимых события A и B изображаются двумя не пересекающимися областями в пространстве элементарных событий Ω , как это изображено на рис. 3.

Покажем, если события A и B независимы, то условная вероятность события B при условии, что произошло событие A , совпадает с безусловной вероятностью события B : $P(A \cap B) = P(B|A)P(A) = P(A)P(B)$, то есть $P(B|A) = P(B)$.

Это означает, что если события A и B независимы, то сведения об осуществлении события A не меняют вероятности осуществления события B .

Пусть рассматриваются n событий A_1, A_2, \dots, A_n . Эти события называются *попарно независимыми*, если выполняется $P(A_i \cap A_j) = P(A_i)P(A_j)$ при всех значениях индексов $i, j = 1, 2, \dots, n$. Указанные события называются *независимыми в совокупности*, если имеют место соотношения:

$$P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k), \quad \text{где } i \neq j \neq k;$$

$$P(A_i \cap A_j \cap A_k \cap A_m) = P(A_i)P(A_j)P(A_k)P(A_m), \quad \text{где } i \neq j \neq k \neq m;$$

или в общем виде

$$P\left(\bigcap_{j=1}^N A_j\right) = \prod_{j=1}^N P(A_j).$$

Некоторые свойства независимых событий

1. Если $P(A) > 0$, то независимость событий A и B эквивалентна равенству $P(B|A) = P(B)$. Доказательство очевидно.

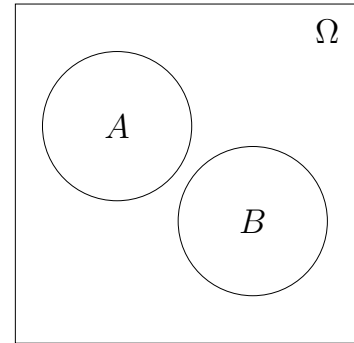


Рис. 3. Графическое представление двух независимых событий

2. Если события A и B независимы, то независимы B и \bar{A} . Действительно, $P(B \cap \bar{A}) = P(B \setminus A \cap B) = P(B) - P(A \cap B) = P(B)(1 - P(A)) = P(B)P(\bar{A})$.
3. Пусть A и B_1 , A и B_2 независимы и $B_1 \cap B_2 = \emptyset$. Тогда независимы события A и $B_1 \cup B_2$. Действительно, $P(A \cap (B_1 \cup B_2)) = P((A \cap B_1) \cup (A \cap B_2)) = P(A \cap B_1) + P(A \cap B_2) = P(A)[P(B_1) + P(B_2)] = P(A)P(B_1 \cup B_2)$.

Пример 2.5. Из колоды в 52 карты вынимается одна карта. Пусть A означает «туз», а B — «пиковая масть». Доказать, что A и B — независимые события.

Решение. $P(A) = 4/52$, $P(B) = 13/52$; вероятность извлечения пикового туза $P(A \cap B) = 1/52$, но этот же результат можно получить путем умножения $P(A)$ на $P(B)$.

Следовательно, события A и B независимы.

Пример 2.6. В предыдущей задаче в колоду добавлен джокер, т. е. карта, не имеющая масти.

Решение. $P(A) = 4/53$, $P(B) = 13/53$, $P(A \cap B) = 1/53$, в то же время $P(A)P(B) = 52/(53)^2$, следовательно, события A и B не независимы. Этот результат объясняется следующими рассуждениями. Если мы знаем, что данная карта — туз, то «равноправие» среди остальных характеристик данной карты по мастям нарушается из-за того, что туз не может быть джокером.

Пример 2.7. Независимо один от другого проводятся три опыта: бросание монеты, бросание игральной кости и вытаскивание карты из колоды. Чему равна вероятность того, что результаты этих трех опытов таковы: первого опыта — герб, второго — 5 очков и третьего опыта — туз?

Решение. Пусть событие A — выпадение герба, событие B — выпадение 5 очков и событие C — извлечение туза, тогда $P(A) = 1/2$, $P(B) = 1/6$, $P(C) = 1/13$. Требуемая вероятность

$$P(A \cap B \cap C) = \frac{1}{2} \cdot \frac{1}{6} \cdot \frac{1}{13} = \frac{1}{156}.$$

Пример 2.8. Электрические лампочки производятся на одной машине, при этом в среднем одна лампочка из тысячи оказывается бракованной. Каждая лампочка производится независимо от других. Чему равна вероятность того, что две последовательные лампочки окажутся исправными?

Решение. Пусть E и F означают события: «исправная первая лампочка» и «исправная вторая лампочка». Тогда $P(E) = P(F) = 1 - 0,001 = 0,999$ и $P(E \cap F) = P(E)P(F) = (0,999)^2 = 0,998$.

2.6. Формула полной вероятности

Пусть рассматривается некоторое событие A и требуется найти его вероятность по следующим исходным данным.

Известны вероятности $P(A|H_i)$ осуществления события A при условии, что произошло некоторое событие H_i из системы событий H_1, H_2, \dots, H_n , которые попарно несовместимы и в объединении дают все пространство элементарных исходов: $H_i \cap H_j = \emptyset$, где $i \neq j$; $\cup_i H_i = \Omega$. Кроме того, известны вероятности $P(H_i)$ этих событий. Используя операции объединения $\cup_{i=1}^n H_i = \Omega$, пересечения событий $A = A \cap \Omega$, теоремы сложения $P(A \cap B) = P(A) + P(B)$ при $A \cap B = \emptyset$ и умножения вероятностей $P(A \cap B) = P(B|A) \times P(A)$, получаем следующую цепочку равенств $P(A) = P(A \cap \Omega) = P(A \cap \bigcup_{i=1}^n H_i) = P(\bigcup_{i=1}^n A \cap H_i) = \sum_{i=1}^n P(A \cap H_i) = \sum_{i=1}^n P(A|H_i) P(H_i)$. Формула

$$P(A) = \sum_{i=1}^n P(A|H_i) P(H_i) \quad (2.1)$$

называется *формулой полной вероятности*.

Пример 2.9. На фабрике, изготавливающей болты, первая машина производит 25 %, вторая — 35 %, третья — 40 % всех изделий. Брак в их продукции составляет соответственно 5 %, 4 %, 2 %. Какова вероятность того, что случайно выбранный болт окажется дефектным?

Решение. Обозначим через A событие, состоящее в том, что случайно выбранный болт — дефектный, а через B_1, B_2, B_3 — событие, состоящее в том, что этот болт произведен соответственно первой, второй и третьей машиной. Тогда $P(A) = P(B_1) \times P(A|B_1) + P(B_2) \times P(A|B_2) + P(B_3) \times P(A|B_3) = 0,25 \times 0,05 + 0,35 \times 0,04 + 0,40 \times 0,02 = 0,0345$.

2.7. Формула Байеса

Иногда возникает задача, в определенном смысле обратная только что рассмотренной, а именно: требуется найти вероятность возникновения события H_i при условии, что произошло событие A , т. е. найти $P(H_i|A)$.

Ответить на этот вопрос можно, если воспользоваться формулой полной вероятности (2.1) и формулой умножения вероятностей:

$$P(H_i \cap A) = P(H_i|A) P(A) = P(A|H_i) P(H_i).$$

Отсюда $P(A) = P(A|H_i) P(H_i)/P(H_i|A)$, то есть

$$P(H_i|A) = \frac{P(A|H_i) P(H_i)}{\sum_{j=1}^n P(A|H_j) P(H_j)}. \quad (2.2)$$

Эта соотношение известно под названием *формулы Байеса*.

Пример 2.10. Вернемся к условию предыдущей задачи. Определим вероятность того, что случайно выбранный болт произведен первой, второй и третьей машиной:

$$P(A|B_1) = 0,25 \times 0,05/0,0345 = 125/345,$$

$$P(A|B_2) = 0,35 \times 0,04/0,0345 = 140/345,$$

$$P(A|B_3) = 0,40 \times 0,02/0,0345 = 80/345.$$

Контрольные вопросы

1. Дать определения случайных событий, несовместимых событий. Привести соответствующие схемы.
2. Дать определения объединения и пересечения множества событий. Привести соответствующие схемы.
3. Дать определения разности событий, противоположного события. Привести соответствующие схемы.
4. Какой класс событий рассматривается в теории вероятностей?
5. Перечислить свойства вероятности.
6. В чем специфика понятия условной вероятности?
7. Перечислить свойства независимых событий.
8. Вывести формулу полной вероятности.
9. Вывести формулу Байеса.

3. Последовательность независимых испытаний

3.1. Схема Бернулли

Рассмотрим дискретное вероятностное пространство $\{\Omega, \mathcal{F}, P\}$, в котором пространство элементарных событий $\Omega = \{\omega_1, \omega_2, \dots, \omega_R\}$ содержит R событий. Алгебра \mathcal{F} состоит из всех подмножеств Ω , а вероятность определена для каждого элементарного события $\omega_i \in \Omega$ ($i = 1, 2, 3, \dots, R$) равенством

$$P(\omega_i) = P_i, \quad \sum_{j=1}^R P_j = 1. \quad (3.1)$$

Пусть ω_i — случайное событие, сопоставленное исходу некоторого опыта S , а $\omega_1, \omega_2, \dots, \omega_R$ — возможные исходы этого опыта. Тогда опыту S , повторенному дважды, должно быть сопоставлено вероятностное пространство

$$\{\Omega_2, \mathcal{F}_2, P_2\} = \{\Omega, \mathcal{F}, P\} \times \{\Omega, \mathcal{F}, P\}. \quad (3.2)$$

Элементарными событиями этого пространства будут упорядоченные пары исходов $\omega_{ij} = (\omega_i, \omega_j) \in \Omega_2 = \Omega \times \Omega$, а вероятность P_2 на алгебре \mathcal{F}_2 можно задать многими способами, определяемыми условиями опыта. В том случае, когда исходы первого опыта никак не влияют на исходы второго, выполняется $P_2(\omega_i, \omega_j) = P_i P_j$. Очевидно, n раз независимо повторенному опыту S отвечает следующее вероятностное пространство:

$$\{\Omega_n, \mathcal{F}_n, P_n\} = [\{\Omega, \mathcal{F}, P\}]^n, \quad P(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_n}) = P_{i_1} P_{i_2} \dots P_{i_n}. \quad (3.3)$$

Определение. Последовательностью n независимых испытаний называется вероятностное пространство $\{\Omega_n, \mathcal{F}_n, P_n\}$, элементами событий в котором являются последовательности $(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_n})$ и вероятности элементарных событий определяются формулой:

$$P_n(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_n}) = \prod_{k=1}^n P_{i_k}. \quad (3.4)$$

Схемой Бернулли называется последовательность независимых испытаний, в каждом из которых имеется лишь два исхода, т.е. $\Omega = \{\omega_1, \omega_2\}$. Это равнозначно принятию условия, что все пространство элементарных событий, связанное с опытом S , состоит только из двух элементов ω и $\bar{\omega}$. Исход ω называют «успехом», и вероятность успеха $P(\omega)$ обозначают через p . Исход $\bar{\omega}$ называют «неудачей», вероятность которой $P(\bar{\omega}) = 1 - p = q$.

В серии из n независимых испытаний в схеме Бернулли пространство элементарных событий Ω_n состоит из 2^n элементов, причем вероятность обнаружить k успехов в серии из n испытаний равна $P(n, k) = p^k q^{n-k}$.

В качестве примера схемы Бернулли можно привести опыт с бросанием монеты или игральной кости. В последнем случае можно принять за успех ω выпадения определенного числа очков (скажем, одного очка), а $\bar{\omega}$ — невыпадение этого числа (одного очка).

По схеме Бернулли можно найти вероятность того, что в серии из n испытаний успех наступит точно k раз ($k \leq n$) при условии, что не имеет значения, когда именно в этих испытаниях будут наблюдаться k успехов. Количество событий, состоящих в наступлении k успехов в n испытаниях, равно C_n^k , а поскольку они несовместны, то

$$P_n(k) = C_n^k p^k q^{n-k}, \quad \text{где } k = 1, 2, \dots, n. \quad (3.5)$$

В частности, вероятность не иметь ни одного успеха $P_n(0) = q^n$. Следовательно, вероятность иметь хотя бы один успех $P_n(k > 0) = 1 - q^n$.

Совокупность вероятностей (3.5) с $k = 1, 2, \dots, n$ называется *биномиальным распределением*. Такое название связано с тем, что эта вероятность является общим членом бинома Ньютона:

$$(p + q)^n = \sum_{m=0}^n C_n^m p^m q^{n-m} = 1. \quad (3.6)$$

Схему Бернулли можно обобщить на случай, когда в опыте возможно не два исхода, а несколько: $\Omega = \{\omega_1, \omega_2, \dots, \omega_r\}$, причем

$$\sum_{i=1}^r P(\omega_i) = \sum_{i=1}^r p_i = 1. \quad (3.7)$$

Пусть проводится n независимых испытаний. Вероятность того, что будет получено m_1, m_2, \dots, m_r ($\sum_i m_i = n$) исходов первого, второго и т. д. видов, дается формулой:

$$P_n(m_1, m_2, \dots, m_r) = \frac{n!}{m_1! m_2! \dots m_r!} p_1^{m_1} p_2^{m_2} \dots p_r^{m_r}. \quad (3.8)$$

Эта формула называется *полиномиальным распределением*.

Пример 3.1. В семье 10 детей. Считая для упрощения, что вероятность рождения мальчика равна 0,5, найдем вероятность того, что в семье имеется 0, 1, 2, ..., 10 мальчиков.

Решение. Отметим, что в силу свойств $C_n^k = C_n^{n-k}$ и $q = p = 0,5$ имеет место равенство $P_n(k) = P_n(n - k)$. Таким образом: $P_{10}(0) = P_{10}(10) = C_{10}^0/2^{10} = 1/1024$, $P_{10}(1) = P_{10}(9) = C_{10}^1/2^{10} = 10/1024$, $P_{10}(2) = P_{10}(8) = C_{10}^2/2^{10} = 45/1024$, $P_{10}(3) = P_{10}(7) = C_{10}^3/2^{10} = 120/1024$, $P_{10}(4) = P_{10}(6) = C_{10}^4/2^{10} = 210/1024$, $P_{10}(5) = C_{10}^5/2^{10} = 252/1024$.

3.2. Наиболее вероятное число успехов

Если n фиксировано, то $P_n(k)$ в (3.2) превращается в некоторую функцию от аргумента k , принимающую значения $0, 1, 2, \dots, n$. На рис. 4 точкам соответствуют биномиальные распределения $P_n(k)$ при трех значениях $n = 11, 30, 100$ и $p = 0,7$.

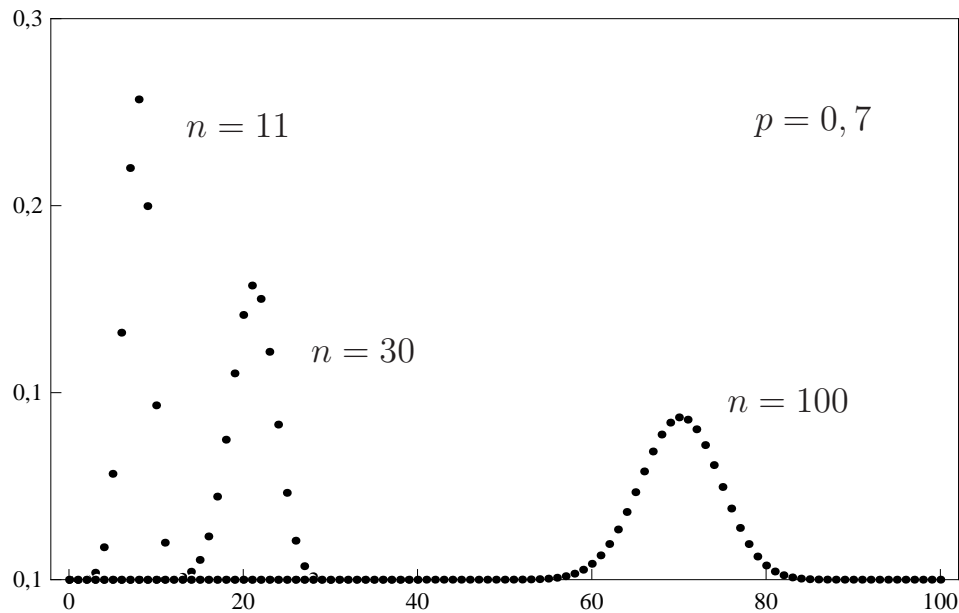


Рис. 4. Биномиальное распределение $P_n(k)$ при трех значениях $n = 11, 30, 100$ и $p = 0,7$

Особенностью этих распределений является то, что вероятность $P_n(k)$ сначала возрастает при увеличении k , достигает наибольшего значения при некотором наиболее вероятном значении $k = k_{\text{нв}}$, а затем убывает.

Кривая, вообще говоря, асимметрична, но при увеличении n форма кривой при немалых p стремится к симметричной. Практически график биномиального распределения можно считать симметричным при $n \geq 4$.

Естественно поставить вопрос: при каком значении аргумента эта функция достигает максимума? Иными словами, какое число успехов является наиболее вероятным при n испытаний?

Это значение можно оценить при больших n . Согласно формуле Стирлинга величина $n!$ при $n \rightarrow \infty$ приближенно равна:

$$n! \simeq \sqrt{2\pi n} n^n e^{-n} e^{\theta(n)}, \quad (3.9)$$

где остаточный множитель $\theta(n)$ удовлетворяет неравенству: $|\theta(n)| \leq 1/(12n)$. Отсюда можно записать:

$$P_n(k) \simeq \frac{n^n}{k^k (n-k)^{n-k}} p^k q^{n-k}. \quad (3.10)$$

Дифференцируя эту формулу по k и приравнявая производную нулю, получаем, что максимум находится при $k_{\text{нв}} = np$.

При небольших значениях n результат можно проанализировать численно. Анализ позволяет сделать заключение, что максимум достигается при таком целом числе $k_{\text{нв}}$, которое заключено в области

$$np - q < k_{\text{нв}} < np + p. \quad (3.11)$$

Таким образом, одно из ближайших к np целых чисел равно наиболее вероятному числу успехов. Или другая трактовка: np можно рассматривать в определенном смысле как среднее число успехов в n опытах.

3.3. Распределение Пуассона

Формулы биномиального распределения вероятностей приводят при больших n к очень громоздким вычислениям. Поэтому важно иметь приближенные, но зато достаточно простые формулы для вычисления соответствующих вероятностей. В частности, нередко встречаются задачи, в которых рассматривается большое число независимых испытаний, причем вероятность наступления события A при каждом отдельном испытании мала. В этом случае вероятности $P_n(k)$ могут быть приближенно вычислены по так называемой *формуле Пуассона*. Эта формула получается как предельная для биномиального распределения, когда число испытаний n стремится к бесконечности, а вероятность успеха p стремится к нулю.

Теорема Пуассона. Пусть дана последовательность $\{S_n\}$ серий независимых испытаний, состоящих соответственно из 1, 2, ..., n испытаний, и пусть вероятность p события A при каждом испытании n -й серии равна $p = \lambda/n$, где λ — параметр, не зависящий от n . Тогда вероятность $P_n(k)$ того, что число

наступлений события A в n -й серии будет равно k , при $n \rightarrow \infty$ и фиксированном значении k можно вычислить по формуле

$$P_n(k) \simeq \frac{\lambda^k}{k!} e^{-\lambda}.$$

Доказательство:

$$\begin{aligned} \lim_{n \rightarrow \infty} P_n(k) &= \lim_{n \rightarrow \infty} \frac{n!}{k! (n-k)!} p^k (1-p)^{n-k} = \\ &= \lim_{n \rightarrow \infty} \frac{n!}{k! (n-k)! n^k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\ &= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \frac{n(n-1) \cdots (n-k+1)}{n^k} = \\ &= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \lim_{n \rightarrow \infty} \left[1 \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right)\right] = \frac{\lambda^k}{k!} e^{-\lambda}. \end{aligned}$$

Распределение вероятностей, определяемое формулой

$$P(k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots, \quad \lambda > 0, \quad (3.12)$$

называется *распределением Пуассона* и является одним из важнейших в теории вероятностей. Оно описывает вероятности во многих задачах, таких как число отказов радиоэлектронной аппаратуры за время t , число распавшихся к моменту времени t радиоактивных атомов, число вызовов на телефонной станции и т. д. Очевидно,

$$\sum_{m=0}^{\infty} \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} e^{\lambda} = 1. \quad (3.13)$$

Формула Пуассона получена нами как предельная для последовательности серий независимых событий, в которой число испытаний стремится к бесконечности, а вероятность наступления события A стремится к нулю. Поэтому ясно, что если мы хотим воспользоваться приближенной формулой

$$P_n(m) \simeq \frac{\lambda^m}{m!} e^{-\lambda}, \quad m = 0, 1, 2, \dots, n, \quad \lambda = np,$$

для одной серии из n испытаний, то n должно быть велико, а вероятность p события A в этой серии мала, т. е. формула (3.12) приемлема для редких явлений.

Пример 3.2. В лотерее в среднем на 1 000 номеров разыгрывается один выигрыш. Какова вероятность, имея 100 билетов, получить не менее 2 выигрышей?

Это схема Бернулли с $n = 100$, а вероятность успеха $p = 1/1000$, так что $\lambda = np = 0,1$. Искомая вероятность:

$$P_x = \sum_{m=2}^{100} P_{100}(m) = 1 - P_{100}(0) - P_{100}(1).$$

По формуле Пуассона:

$$\begin{aligned} P_{100}(0) &\simeq P(0) = e^{-0,1} \simeq 0,9, \\ P_{100}(1) &\simeq P(1) = 0,1 e^{-0,1} \simeq 0,09, \end{aligned}$$

откуда искомая вероятность

$$P_x \simeq 1 - 0,9 - 0,09 = 0,01.$$

3.4. Предельные теоремы Муавра—Лапласа

Рассмотрим другую предельную форму биномиального распределения, предполагая, что вероятность p немалая величина $0 < p < 1$. Эта задача была решена для случая $p = 1/2$ в 1730 г. французским математиком Абрахамом де Муавром (1667–1754), а затем была обобщена для произвольного p Пьером-Симоном де Лапласом (1749–1827) в 1783 г. Две его теоремы известны как локальная и интегральная предельные теоремы. Доказательство основывается на известной из курса математического анализа формуле Стирлинга (3.9). Мы приведем эти теоремы без вывода: желающие могут найти их практически в любом учебнике по теории вероятности.

3.4.1. Локальная предельная теорема Муавра—Лапласа

Если вероятность наступления события A в n независимых испытаниях постоянна и равна p ($0 < p < 1$), то вероятность $P_n(m)$ того, что в этих испытаниях событие A наступит ровно m раз, удовлетворяет при $n \rightarrow \infty$ соотношению:

$$\lim_{n \rightarrow \infty} \sqrt{npq} P_n(m) = \varphi(x_m) = \frac{1}{\sqrt{2\pi}} e^{-x_m^2/2} \quad (3.14)$$

равномерно для всех m , для которых $x_m = (m - np)/\sqrt{npq}$ находится в каком-либо конечном интервале.

Пример 3.3. Вероятность того, что станок-автомат производит годную деталь, равна $8/9$. За смену изготовлено 280 деталей. Определить вероятность того, что среди них 20 бракованных.

Решение. $n = 280$, $k = 20$, $p = 1/9$, $q = 8/9$. Отсюда

$$\sqrt{npq} = \sqrt{280(1/9)(8/9)} \approx 5,2588,$$

и $t = (k - np) / \sqrt{npq} = [20 - 280(1/9)] / 5,2588 \approx -2,11$. Из таблицы, в которой приведены значения функции

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2},$$

находим $\varphi(-2,11) = 0,0431$, и искомая вероятность $P_{280}(20) = 0,0082$.

3.4.2. Интегральная предельная теорема Муавра—Лапласа

Если m есть число наступлений события A в серии n независимых испытаний, в каждом из которых вероятность этого события равна p , причем $0 < p < 1$, то равномерно относительно a и b ($-\infty \leq a \leq b \leq \infty$) при $n \rightarrow \infty$ имеет место соотношение:

$$P \left\{ a \leq \frac{m - np}{\sqrt{npq}} \leq b \right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-z^2/2} dz. \quad (3.15)$$

Для решения задач, требующих применения интегральной предельной теоремы Муавра—Лапласа, пользуются специальными таблицами, так как неопределенный интеграл

$$\int e^{-z^2/2} dz$$

не выражается через элементарные функции. Таблица значений для интеграла Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$$

приводится во многих книгах по теории вероятностей.

Пример 3.4. Вероятность того, что деталь не прошла ОТК, равна $0,2$. Найти вероятность того, что среди 400 случайно отобранных деталей окажется непроверенных от 70 до 100 деталей.

Решение. По условию $p = 0,2$, $q = 0,8$, $n = 400$, $k_1 = 70$, $k_2 = 100$. Воспользуемся интегральной теоремой Муавра—Лапласа

$$P_{400}(70, 100) = \Phi(b) - \Phi(a).$$

Вычислим нижний и верхний пределы интегрирования:

$$a = \frac{70 - 400 \times 0,2}{\sqrt{400 \times 0,2 \times 0,8}} = -1,25, \quad b = \frac{100 - 400 \times 0,2}{\sqrt{400 \times 0,2 \times 0,8}} = 2,5.$$

Из таблиц находим $\Phi(2,5) = 0,4938$ и $\Phi(-1,25) = -0,3944$. Отсюда

$$P_{400}(70, 100) = 0,4938 + 0,3944 = 0,8882.$$

Контрольные вопросы

1. Изложить основные положения схемы Бернулли.
2. Привести оценку наиболее вероятного числа успехов.
3. Вывести формулу распределения Пуассона.
4. Изложить смысл и формулировки локальной и интегральной предельных теорем Муавра—Лапласа.
5. Привести примеры применения предельных теорем Муавра—Лапласа.

4. Последовательность взаимосвязанных испытаний

4.1. Цепи Маркова

Проведение последовательности испытаний в схеме Бернулли основывалось на понятии независимости испытаний. Существует, однако, большой круг задач, в которых последовательно проводимые испытания не являются независимыми, а, наоборот, связаны между собой в определенного рода цепь. Один из вариантов такой цепи получил название цепи Маркова.

1. В отличие от той ситуации, которая имела место в случае схемы Бернулли, где рассматривалась последовательность из заданного числа n испытаний, мы будем считать, что проводится бесконечная последовательность испытаний.
2. Далее мы примем, что имеется m попарно несовместных событий $A_1, A_2, A_3, \dots, A_m$ (m — фиксированное число), таких, что в результате каждого испытания обязательно наступает одно из них. (В схеме Бернулли рассматриваются только два исхода — A_1 и A_2 .)
3. Наконец — и это самое существенное — мы будем считать испытания связанными между собой некоторым специфическим способом. А именно, вероятность события A_j ($j = 1, 2, \dots, m$) в любом s -м испытании зависит от результата предыдущего, $(s-1)$ -го испытания, и только от него (тем самым она не зависит от результатов более ранних испытаний: $(s-2)$ -го, $(s-3)$ -го, \dots , 1-го).

Обозначим такую условную вероятность как

$$P_{ji} = P(A_j | A_i), \quad 0 \leq P_{ji} \leq 1. \quad (4.1)$$

В этом обозначении отражен тот факт, что вероятность перехода в состояние A_j из состояния A_i на шаге s не зависит от номера s . Это свойство называется *однородностью* последовательности испытаний.

Каждое из чисел P_{ji} называют вероятностью перехода, имея в виду следующий переход: A_i (в $(s-1)$ -м испытании) $\rightarrow A_j$ (в s -м испытании). Матрицу, образованную из чисел P_{ji} , называют *матрицей перехода*:

$$\Pi = \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1m} \\ P_{21} & P_{22} & \cdots & P_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ P_{m1} & P_{m2} & \cdots & P_{mm} \end{pmatrix}. \quad (4.2)$$

По определению все элементы матрицы Π неотрицательны и сумма элементов каждого столбца равна единице:

$$\sum_{j=1}^m P_{ji} = 1, \quad i = 1, 2, 3, \dots, m. \quad (4.3)$$

Матрицы, обладающие этими свойствами, называются *стохастическими*.

Пусть задано распределение вероятностей для первого испытания:

$$P^{(1)}(A_i) = a_i, \quad i = 1, 2, 3, \dots, m, \quad a_i \geq 0; \quad \sum_{i=1}^m a_i = 1. \quad (4.4)$$

Тогда вероятность наблюдения события A_j во втором испытании будет определяться формулой:

$$P^{(2)}(A_j) = \sum_{i=1}^m P_{ji} a_i. \quad (4.5)$$

Аналогично можно получить выражение для вероятности третьего:

$$P^{(3)}(A_k) = \sum_{i=1}^m \sum_{j=1}^m P_{kj} P_{ji} a_i, \quad (4.6)$$

четвертого $P^{(4)}(A_l)$ и т. д. испытаний.

Если матрицу Π рассматривать как вероятность перехода за один шаг, то вероятность перехода за два шага $\Pi^{(2)}$:

$$P_{ki}^{(2)} \equiv \sum_{j=1}^m P_{kj} P_{ji},$$

или в матричной форме:

$$\Pi^{(2)} = \Pi^{(1)} \times \Pi^{(1)} = \left[\Pi^{(1)} \right]^2. \quad (4.7)$$

В общем случае вероятность перехода за N шагов

$$\Pi^{(N)} = \left[\Pi^{(1)} \right]^N. \quad (4.8)$$

4.2. Пределные вероятности

Следующей важной задачей является исследование поведения вероятностей $P_{ji}^{(N)}$ при неограниченном увеличении числа испытаний N . Решение этой задачи в ряде случаев может быть получено на основе теоремы Маркова.

Теорема Маркова. Пусть существует такое число ν шагов, что все вероятности $P_{ji}^{(\nu)}$ строго положительны, т. е.

$$\min_{1 \leq j \leq m} P_{ji}^{(\nu)} = \delta > 0, \quad i = 1, 2, 3, \dots, m.$$

Тогда для каждого состояния A_j существует предельная вероятность его наступления, т. е. такое число p_j , что независимо от i имеет место равенство

$$\lim_{N \rightarrow \infty} P_{ji}^{(N)} = p_j.$$

Доказательство этой теоремы довольно сложно, и мы его не приводим. Отметим, однако, что смысл содержащегося в теореме утверждения интуитивно понятен: вероятность того, что система окажется в состоянии A_j , практически не зависит от того, в каком состоянии находилась система в далеком прошлом. Это свойство цепей Маркова, называемое *эргодичностью*, действительно имеет место, хотя и не во всех случаях.

Возникает вопрос: если существует предельная вероятность, то каким образом ее можно найти?

Легко видеть, что искомые числа p_j ($j = 1, 2, \dots, m$) удовлетворяют системе из m уравнений:

$$p_j = \sum_{k=1}^m P_{jk} p_k, \quad \text{где } j = 1, 2, 3, \dots, m \quad (4.9)$$

при условии

$$\sum_{j=1}^m p_j = 1. \quad (4.10)$$

Итак, предельные вероятности должны удовлетворять системе, состоящей из уравнений (4.9) и (4.10). Более того, из (4.9) следует, что предельные вероятности существуют, если

$$\det [P_{ji} - \delta_{ij}] = 0.$$

Контрольные вопросы

1. Как формируются последовательности событий, называемые «цепи Маркова»?
2. Раскрыть смысл матрицы перехода.
3. Как определяется вероятность перехода из одного состояния в другое за n шагов в цепи Маркова?
4. В чем состоит смысл утверждения теоремы Маркова?

5. Случайные величины

5.1. Понятие случайной величины

Пусть $\{\Omega, \mathcal{F}, P\}$ — вероятностное пространство. Случайной величиной ξ называется однозначная действительная функция от случайного события ω :

$$\xi(\omega) = \xi. \quad (5.1)$$

В отличие от случайного события, являющегося качественной характеристикой случайного результата испытаний ω , случайная величина ξ характеризует результат испытания количественно. Примерами случайной величины могут служить размер обрабатываемой детали, содержание химических элементов в чугуне и т. д. Случайные величины бывают дискретными и непрерывными.

Дискретной называют такую случайную величину, которая принимает конечное или бесконечное счетное множество значений.

Непрерывной называют такую случайную величину, которая может принимать любые значения из некоторого конечного или бесконечного интервала. Число возможных значений непрерывной случайной величины бесконечно.

Пример 5.1. При бросании игральной кости возможны шесть элементарных исходов ω . Очевидно, $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$. С каждым элементарным исходом ω_i можно связать число очков. Тогда случайная величина $\xi = \xi(\omega_i) = i$ равна числу очков, выпавших на кости при бросании.

Пример 5.2. В схеме Бернулли для n независимых испытаний множество Ω состоит из элементарных событий $\omega = (x_1, x_2, \dots, x_n)$, где $x_i = 1$, если при i -м испытании произошел «успех», и $x_i = 0$ в случае «неудачи». Тогда случайная величина $\xi = \xi(\omega) = x_1 + x_2 + \dots + x_n$ равна числу успехов при n испытаниях в схеме Бернулли.

Может оказаться, что для какого-нибудь события A несколькими элементарными событиями $\omega_i \in A$ соответствует одна и та же случайная величина. Такие случайные величины будем называть *вырожденными*.

С каждым элементарным событием $\omega \in A$ можно связать случайную величину, называемую *индикатором* события A :

$$\chi_A = \chi_A(\omega) = \begin{cases} 1, & \text{если } \omega \in A, \\ 0, & \text{если } \omega \notin A. \end{cases} \quad (5.2)$$

Индикаторы обладают следующими свойствами:

а) $\chi_\emptyset = 0$;

б) $\chi_\Omega = 1$;

в) $\chi_{A \cap B} = \chi_A * \chi_B$;

г) $\chi_B = 1 - \chi_A$;

д) если события A и B несовместны ($A \cap B = \emptyset$), то $\chi_{A \cup B} = \chi_A + \chi_B$.

Пусть $\xi_1 < \xi_2 < \dots < \xi_R$ — всевозможные значения случайной величины ξ . Рассмотрим событие A_i , состоящее из всех тех элементарных событий ω , для которых случайная величина $\xi(\omega) = \xi_i$.

Так как все ξ_i различны, то события A_i и A_j ($i \neq j$) несовместны. Поскольку $\xi_1, \xi_2, \xi_3, \dots, \xi_R$ исчерпывают все возможные значения случайной величины ξ , то

$$\bigcup_{i=1}^R A_i = \Omega.$$

Таким образом, со случайной величиной ξ можно связать разбиение Ω , состоящее из множеств

$$A_i = \{\omega : \xi(\omega) = \xi_i\}, \quad i = 1, 2, 3, \dots, R, \quad (5.3)$$

в каждом из которых $\xi(\omega)$ постоянна. Такие множества образуют разбиение, порожденное случайной величиной ξ .

С помощью разбиения (5.3) случайную величину ξ можно представить в виде линейной комбинации индикаторов

$$\xi = \xi(\omega) = \sum_{i=1}^R \xi_i \chi_{A_i}. \quad (5.4)$$

Для задания случайной величины недостаточно перечислить все ее возможные значения. Необходимо также задать вероятности их появления. Законом распределения случайной величины называется всякое соотношение между возможными значениями случайной величины и соответствующими им вероятностями. Про случайную величину говорят, что она подчиняется данному закону распределения. Закон распределения дискретной случайной величины можно задать с помощью таблицы:

x_1	x_2	x_3	\dots	x_R
p_1	p_2	p_3	\dots	p_R

Верхний ряд таблицы состоит из различных чисел x_i , которые принимает случайная величина, а нижний ряд — из вероятностей, которые удовлетворяют условию:

$$p_i \geq 0, \quad \sum_{i=1}^R p_i = 1.$$

Приведем несколько законов распределения для дискретных величин.

1. *Биноминальный закон* для числа успехов m при n независимых испытаниях в схеме Бернулли:

$$P_n(\xi = m) = C_n^m p^m (1 - p)^{n-m}, \quad m = 1, 2, \dots, n.$$

2. *Гипергеометрическое распределение* — распределение числа белых шаров m при выборке n шаров из урны, содержащей M белых и $N - M$ черных шаров, без возвращения их обратно:

$$P(\xi = m) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}, \quad m = 0, 1, \dots, \min(n, M).$$

3. *Равномерное распределение*:

$$P(\xi = m) = \frac{1}{N}, \quad m = 1, 2, \dots, N.$$

До сих пор мы рассматривали вероятностное пространство на конечном множестве элементов ξ_i . Полученные выражения легко обобщить и на бесконечное, но счетное множество элементов, устремив N к бесконечности.

Приведем два распределения для вероятностного пространства со счетным множеством событий.

1. *Пуассоновское распределение*:

$$P(\xi = m) = \frac{\lambda^m}{m!} e^{-\lambda},$$

где $m = 0, 1, 2, \dots$, зависящее от параметра λ . Очевидно,

$$\sum_{m=0}^{\infty} \frac{\lambda^m}{m!} e^{-\lambda} = 1.$$

2. *Геометрическое распределение*:

$$P(\xi = m) = (1 - p) p^m,$$

зависящее от параметра $0 < p < 1$, причем очевидно, что

$$\sum_{m=1}^{\infty} (1 - p) p^m = 1.$$

5.2. Случайные величины в общей схеме

В общем случае непрерывного набора случайных событий и случайных величин прежние определение закона распределения случайной величины затруднительно, поэтому дается немного иное определение.

Определение. В произвольном вероятностном пространстве $\{\Omega, \mathcal{F}, P\}$ случайной величиной ξ называется такая функция $\xi = \xi(\omega)$ от элементарного события ω , для которого при любом численном значении x неравенство $\xi(\omega) < x$ является событием. Вероятность этого события

$$P(\xi < x) \equiv F_\xi(x) \quad (5.5)$$

называется *функцией распределения случайной величины ξ* .

Функция распределения $F_\xi(x)$ определяет вероятность на множестве значений ξ , являясь в то же время конструкцией значительно более простой, чем вероятность.

Пример 5.3. Пусть ξ — число успехов в серии из n испытаний по схеме Бернулли. Тогда соответствующая функция распределения определится равенством:

$$F_\xi(x) = \begin{cases} 0 & \text{при } x \leq 0, \\ \sum_{k < x} C_n^k p^k (1-p)^{n-k} & \text{при } 0 < x \leq n, \\ 1 & \text{при } x > n. \end{cases} \quad (5.6)$$

Пример 5.4. Если ξ распределена по закону Пуассона, то ее функция распределения $F_\xi(x)$ имеет вид:

$$F_\xi(x) = \begin{cases} 0 & \text{при } x \leq 0, \\ \sum_{k < x} \frac{\lambda^k}{k!} e^{-\lambda} & \text{при } x > 0. \end{cases} \quad (5.7)$$

Пример 5.5. Говорят, что случайная величина ξ имеет нормальное (гаусово) распределение $F_\xi(x)$, если ее функция распределения имеет вид:

$$F_\xi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-(z-\mu)^2/(2\sigma^2)} dz. \quad (5.8)$$

Перечислим основные свойства функции распределения.

1. $0 \leq F_\xi(x) \leq 1$ для всех x ;
2. $\lim_{x \rightarrow +\infty} F_\xi(x) = 1$, $\lim_{x \rightarrow -\infty} F_\xi(x) = 0$.

3. $F_\xi(x_1) \leq F_\xi(x_2)$, если $x_1 \leq x_2$. Иными словами, функция распределения $F_\xi(x)$ — неубывающая функция.
4. $\lim_{\varepsilon \rightarrow +0} F_\xi(x + \varepsilon) = F_\xi(x + 0)$, $\lim_{\varepsilon \rightarrow +0} F_\xi(x - \varepsilon) = F_\xi(x)$, т. е. функция распределения $F_\xi(x)$ непрерывна слева.

Используя введенную функцию распределения, мы можем уточнить определения дискретной и непрерывной случайных величин.

Случайная величина ξ называется *дискретной*, если она принимает конечное или счетное множество значений и для нее функция распределения $F_\xi(x)$ может быть представлена в виде:

$$F_\xi(x) = \sum_{x_k < x} P_k \equiv \sum_k P(\xi(\omega_k) < x). \quad (5.9)$$

Очевидно, что $F_\xi(x)$ не зависит от способа нумерации значений случайной величины ξ_k . Функция распределения любой дискретной величины разрывна, возрастает скачками в точках $x = x_k$. Величина скачка равна

$$F_\xi(x) \Big|_{x=x_k+\varepsilon} - F_\xi(x) \Big|_{x=x_k-\varepsilon} = P_k. \quad (5.10)$$

Случайная величина ξ называется *непрерывной*, если она может принимать любые значения из бесконечного интервала и ее функция распределения может быть представлена в виде:

$$F_\xi(x) = \int_{-\infty}^x \rho_\xi(y) dy. \quad (5.11)$$

Функция $\rho_\xi(y)$ называется *плотностью распределения вероятности* (*плотностью вероятности*) случайной величины ξ и предполагается не отрицательной $\rho_\xi(y) \geq 0$ и кусочно-непрерывной.

Плотность вероятности $\rho_\xi(y)$ полностью определяет функцию распределения $F_\xi(x)$, причем

$$\rho_\xi(x) = \frac{dF_\xi(x)}{dx}. \quad (5.12)$$

5.3. Функции от случайных величин

Пусть ξ — непрерывная случайная величина с функцией распределения $F_\xi(x)$ и плотностью $\rho_\xi(x)$.

Если $g(x)$ — числовая функция от действительного аргумента x , то можно образовать случайную величину $\eta = g(\xi)$, которая является функцией от ξ . Спрашивается, как по закону распределения случайной величины ξ найти закон распределения η ? Приведем несколько примеров, из которых станет ясно, как надо поступать в общем случае.

Пример 5.6. Рассмотрим линейную функцию $g(x) = ax + b$. Положим $\eta = a\xi + b$. Если $a > 0$, то по определению функции распределения

$$F_\eta(y) = P\{\eta < y\} = P\{a\xi + b < y\} = P\{\xi < (y - b)/a\} = F_\xi((y - b)/a).$$

Откуда

$$\rho_\eta(y) = \frac{dF_\eta(y)}{dy} = \frac{1}{a} \frac{dF_\xi(x)}{dx} \Big|_{x=(y-b)/a} = \frac{1}{a} \rho_\xi \left(\frac{y - b}{a} \right).$$

Если $a < 0$, то

$$F_\eta(y) = P\{\eta < y\} = P\{a\xi + b < y\} = P(\xi > (y - b)/a) = 1 - F_\xi((y - b)/a)$$

и

$$\rho_\eta(y) = \frac{dF_\eta(y)}{dy} = -\frac{1}{a} \frac{dF_\xi(x)}{dx} \Big|_{x=(y-b)/a} = -\frac{1}{a} \rho_\xi \left(\frac{y - b}{a} \right).$$

Объединяя эти два случая, имеем

$$\rho_\eta(y) = \frac{dF_\eta(y)}{dy} = \frac{1}{|a|} \frac{dF_\xi(x)}{dx} \Big|_{x=(y-b)/a} = \frac{1}{|a|} \rho_\xi \left(\frac{y - b}{a} \right).$$

Пример 5.7. Рассмотрим монотонно возрастающую функцию $g(x) = x^3$. Положим $\eta = \xi^3$, тогда имеем

$$F_\eta(y) = P\{\eta < y\} = P\{\xi^3 < y\} = P\{\xi < y^{1/3}\} = F_\xi(y^{1/3}).$$

Продифференцировав это равенство, получим

$$\rho_\eta(y) = \frac{dF_\eta(y)}{dy} = \frac{1}{3y^{2/3}} \frac{dF_\xi(x)}{dx} \Big|_{x=y^{1/3}} = \frac{1}{3y^{2/3}} \rho_\xi(y^{1/3}).$$

Пример 5.8. Рассмотрим монотонно убывающую функцию $g(x) = e^{-x}$.

$$\begin{aligned} F_\eta(y) &= P\{\eta < y\} = P\{e^{-\xi} < y\} = P\{\xi > -\ln y\} = \\ &= 1 - P\{\xi < -\ln y\} = 1 - F_\xi(-\ln y), \end{aligned}$$

откуда

$$\rho_\eta(y) = \frac{dF_\eta(y)}{dy} = \frac{1}{y} \frac{dF_\xi(x)}{dx} \Big|_{x=-\ln y} = \frac{1}{y} \rho_\xi(-\ln y).$$

5.4. Многомерные законы распределения

Очень часто в вероятностных моделях приходится рассматривать сразу несколько случайных величин. Например, при стрельбе по плоской мишени случайная точка попадания имеет две координаты ξ и η , которые являются случайными величинами, при антропометрических исследованиях основными параметрами человеческого тела являются вес, рост и объем груди, которые при случайном выборе человека из какой-либо группы также случайны.

В математической модели в этих случаях на вероятностном пространстве $\{\Omega, \mathcal{F}, P\}$ определены несколько случайных величин $\xi_1, \xi_2, \dots, \xi_r$, которые иногда удобно рассматривать как координаты случайной точки или компоненты случайного вектора $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_r)$ из некоторого r -мерного пространства $R^{(r)}$.

Совместным законом распределения этих случайных величин называется вероятность попадания точки $\vec{\xi}$ в r -мерное множество $\{B\}$:

$$P_{\vec{\xi}}(B) = P(\vec{\xi} \in B), \quad (5.13)$$

рассматриваемая в зависимости от множества $\{B\}$. Закон распределения случайного вектора (5.13) называется *многомерным*, или, более точно, r -мерным.

По аналогии с одномерным случаем можно ввести в рассмотрение r -мерную функцию распределения случайных величин $\vec{\xi}$:

$$F_{\vec{\xi}}(\vec{x}) = P(\xi_1 < x_1, \xi_2 < x_2, \dots, \xi_r < x_r). \quad (5.14)$$

В случае непрерывного изменения случайных величин

$$F_{\vec{\xi}}(\vec{x}) = \int_{-\infty}^{x_1} du_1 \int_{-\infty}^{x_2} du_2 \cdots \int_{-\infty}^{x_r} du_r \rho_{\vec{\xi}}(\vec{u}), \quad (5.15)$$

где $\rho_{\vec{\xi}}(\vec{u})$ — многомерная плотность вероятности. Очевидно,

$$\rho_{\vec{\xi}}(\vec{x}) = \frac{\partial^r F_{\vec{\xi}}(\vec{x})}{\partial x_1 \partial x_2 \cdots \partial x_r}. \quad (5.16)$$

Многомерное распределение определяет одномерное распределение

$$F_{\xi_1}(x_1) = F_{\vec{\xi}}(x_1, \infty, \infty, \dots, \infty). \quad (5.17)$$

5.5. Независимость случайных величин

Случайные непрерывные величины $\xi_1, \xi_2, \dots, \xi_r$ называются независимыми, если

$$F_{\vec{\xi}}(\vec{x}) = F_{\xi_1}(x_1) F_{\xi_2}(x_2) \cdots F_{\xi_r}(x_r) = \prod_{i=1}^r F_{\xi_i}(x_i). \quad (5.18)$$

Если продифференцировать r раз $F_{\vec{\xi}}(\vec{x})$, то получим

$$\rho_{\vec{\xi}}(\vec{x}) = \rho_{\xi_1}(x_1) \rho_{\xi_2}(x_2) \cdots \rho_{\xi_r}(x_r) = \prod_{i=1}^r \rho_{\xi_i}(x_i). \quad (5.19)$$

В случае дискретного распределения независимость означает

$$P(\xi_1 = x_1, \xi_2 = x_2, \dots, \xi_r = x_r) = \prod_{i=1}^r P(\xi_i = x_i). \quad (5.20)$$

Многомерное распределение определяет одномерное, а по одномерному распределению можно восстановить многомерное, если случайные величины независимы.

Теорема. Если случайные величины $\chi_1, \chi_2, \dots, \chi_r$ являются функциями $\chi_i = g_i(\vec{\xi})$ от независимых случайных величин $\xi_1, \xi_2, \dots, \xi_r$, то они также независимы.

Эта теорема доказывается достаточно просто (рекомендуется проделать самостоятельно).

5.6. Свертка распределений

Пусть ξ_1 и ξ_2 — независимые случайные величины и $\eta = g(x_1, x_2) = x_1 + x_2$. Если имеются плотности $\rho_{\xi_1}(x_1)$ и $\rho_{\xi_2}(x_2)$, то в силу независимости

$$\rho_{\xi_1, \xi_2}(x_1, x_2) = \rho_{\xi_1}(x_1) \rho_{\xi_2}(x_2).$$

Тогда

$$F_{\eta}(z) = \iint_{\xi_1 + \xi_2 \leq z} \rho_{\xi_1}(x_1) \rho_{\xi_2}(x_2) dx_1 dx_2 = \int_{-\infty}^{\infty} \rho_{\xi_1}(x_1) dx_1 \int_{-\infty}^{z-x_1} \rho_{\xi_2}(x_2) dx_2.$$

Производя во внутреннем интеграле замену $x_2 = u - x_1$ и меняя порядок интегрирования, получаем

$$F_{\eta}(z) = \int_{-\infty}^z du \int_{-\infty}^{\infty} \rho_{\xi_1}(x_1) \rho_{\xi_2}(u - x_1) dx_1.$$

Дифференцируя это равенство по z , приходим к *формуле свертки* (*формуле композиции*) для плотности:

$$\rho_{\xi_1+\xi_2}(z) = \int_{-\infty}^{\infty} \rho_{\xi_1}(x_1) \rho_{\xi_2}(z - x_1) dx_1. \quad (5.21)$$

С помощью этой формулы можно по плотности двух независимых случайных величин ξ_1 и ξ_2 находить плотность их суммы.

Контрольные вопросы

1. Привести определения дискретной и непрерывной случайных величин.
2. Раскрыть понятие функции от случайных величин.
3. Что такое «многомерные законы распределения»?
4. Какие случайные величины называются независимыми?
5. Дать определение свертки распределений случайных величин.

6. Числовые характеристики случайных величин

Согласно изложенному ранее исчерпывающей характеристикой случайной величины является ее функция распределения. Но далеко не в каждой задаче нужно знать весь закон распределения. В ряде случаев можно обойтись одним или несколькими числами, отражающими наиболее важные особенности закона распределения, например числом, имеющим смысл среднего значения случайной величины, или же числом, характеризующим средний размер отклонения случайной величины от своего среднего значения, и т. д. Такого рода числа называют *числовыми характеристиками случайной величины*. Их роль в теории вероятностей чрезвычайно велика; многие задачи удаётся решить до конца, оставляя в стороне законы распределения и оперируя только числовыми характеристиками.

6.1. Математическое ожидание

Наиболее важное место среди числовых характеристик занимает так называемое *математическое ожидание*, или, другими словами, *среднее значение* случайной величины.

6.1.1. Математическое ожидание дискретной случайной величины

Определение. Математическим ожиданием $M(\xi)$ случайной величины $\xi = \xi(\omega_k)$, заданной на дискретном вероятностном пространстве $\{\Omega, \mathcal{F}, P\}$, называется число

$$M(\xi) = \sum_{k=1} \xi(\omega_k) P(\omega_k), \quad (6.1)$$

где $P(\omega_k)$ — вероятность события $\omega_k \in \Omega$.

Если множество событий счетно, т. е. в (6.1) стоит сумма бесконечного ряда, то к определению математического ожидания мы должны добавить следующее требование: ряд (6.1) должен сходиться абсолютно. Если ряд (6.1) не сходится абсолютно, то говорят, что математическое ожидание не существует.

Пример 6.1. Найти математическое ожидание числа очков, выпавших при бросании игральной кости.

Решение. Обозначим указанную случайную величину через x . Ее закон распределения имеет вид:

x_i	1	2	3	4	5	6
P_i	1/6	1/6	1/6	1/6	1/6	1/6

Отсюда $M(\xi) = 1(1/6) + 2(1/6) + 3(1/6) + 4(1/6) + 5(1/6) + 6(1/6) = 3,5$.

Пример 6.2. Найти математическое ожидание дискретной величины m ($m = 0, 1, 2, \dots, N$) — числа успехов в опыте из N независимых испытаний в схеме Бернулли.

Решение. В схеме Бернулли число успехов m (вероятность успеха p и вероятность неудачи $q = 1 - p$) подчиняется биномиальному распределению:

$$P_N(m) = C_N^m p^m q^{N-m}.$$

Отсюда

$$M(m) = \sum_{m=0}^N m P_N(m) = \sum_{m=1}^N \frac{N! p^m q^{N-m}}{(m-1)!(N-m)!} = Np(p+q)^{N-1} = Np.$$

Пример 6.3. Найти математическое ожидание дискретной величины k ($k = 0, 1, 2, \dots$), распределенной по закону Пуассона.

Решение. Закон Пуассона задается формулой ($\lambda > 0$)

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

Отсюда

$$M(k) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda.$$

Пример 6.4. Согласно американским статистическим данным смертности вероятность того, что 25-летний человек проживет еще год, равна 0,992. Страховая компания предлагает такому человеку застраховать свою жизнь на год на сумму в 1 000 долларов. Страховой взнос равен 10 долларам. Какую прибыль ожидает получить компания?

Решение. Составим закон распределения случайной величины ξ — прибыли компании.

x_i	10	−990
P_i	0,992	0,008

Следовательно, $M(x) = 10 \times 0,992 - 990 \times 0,008 = 2$. Ожидаемая прибыль положительна, что дает возможность страховой компании продолжать дело.

6.1.2. Математическое ожидание непрерывной случайной величины

Определение. Математическим ожиданием $M(\xi)$ случайной величины $\xi = \xi(u_1, u_2, \dots, u_n) = \xi(\vec{u})$, заданной на абсолютно непрерывном вероятностном пространстве $\{\Omega, \mathcal{F}, P\}$, называется число

$$M(\xi) = \int \cdots \int \xi(\vec{u}) \rho(\vec{u}) du_1 du_2 \dots du_n, \quad (6.2)$$

если интеграл абсолютно сходится. Если интеграл не сходится абсолютно, то говорят, что математическое ожидание $M(\xi)$ не существует. Здесь $\rho(\vec{u})$ — многомерная плотность вероятности.

Заметим, что математическое ожидание можно выразить через физическую вероятность

$$M(\xi) = \int \cdots \int \xi(\vec{u}) dF(\vec{u}). \quad (6.3)$$

Пример 6.5. Для материальной точки, находящейся на отрезке $[a, b]$, найти среднее значение координаты ξ , удовлетворяющей равномерному закону распределения.

Решение. Равномерное распределение на отрезке $[a, b]$ задается следующей плотностью распределения:

$$\rho_\xi(x) = \begin{cases} 1/(b-a), & a \leq x \leq b, \\ 0, & x < a, x > b. \end{cases}$$

Среднее значение координаты ξ или математическое ожидание ξ

$$M(\xi) = \int_{-\infty}^{\infty} x \rho_\xi(x) dx = \frac{1}{b-a} \int_a^b x dx = \frac{a+b}{2}.$$

Как и следовало ожидать, $M(\xi)$ совпадает с серединой отрезка $[a, b]$.

Пример 6.6. Найти математическое ожидание случайной величины ξ , обладающей неотрицательными значениями и распределенной при $x \geq 0$ по экспоненциальному закону с плотностью

$$\rho_\xi(x) = \lambda e^{-\lambda x}.$$

Решение.

$$M(\xi) = \int_{-\infty}^{\infty} x \rho_\xi(x) dx = \lambda \int_0^{\infty} x e^{-\lambda x} dx = 1/\lambda.$$

Отсюда следует смысл параметра λ , входящего в плотность распределения, как величины, обратно пропорциональной математическому ожиданию ξ .

Пример 6.7. Найти математическое ожидание случайной величины ξ , распределенной по нормальному закону с плотностью

$$\rho_{\xi}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-a)^2/(2\sigma^2)}.$$

Решение.

$$M(\xi) = \int_{-\infty}^{\infty} x \rho_{\xi}(x) dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x e^{-(x-a)^2/(2\sigma^2)} dx = a.$$

Отсюда следует смысл величины a , входящей в плотность нормального закона распределения, — математическое ожидание величины ξ .

Пример 6.8. Найти математическое ожидание случайной величины ξ , имеющей распределение Коши ($a > 0$):

$$\rho_{\xi}(x) = \frac{1}{\pi} \frac{a}{a^2 + x^2}.$$

Решение. Вычислим интеграл:

$$\int_{-\infty}^{\infty} x \rho_{\xi}(x) dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x a dx}{a^2 + x^2} = 0.$$

Однако математическое ожидание величины ξ существует в том случае, если существует и так называемый абсолютный первый момент $M(|\xi|)$ величины ξ . Вычислим этот момент:

$$M(|\xi|) = \int_{-\infty}^{\infty} |x| \rho_{\xi}(x) dx = \frac{2}{\pi} \int_0^{\infty} \frac{x a dx}{a^2 + x^2} = \frac{a}{\pi} \log |a^2 + x^2| \Big|_0^{\infty} \rightarrow \infty.$$

Этот интеграл логарифмически расходится на верхнем пределе, а значит как абсолютный первый момент ξ , так и математическое ожидание величины ξ для такого распределения не существуют.

Свойства математического ожидания

1. $M(C) = C$, где C — постоянная.
2. $M(C\xi) = C M(\xi)$, если C — постоянная.
3. $|M(\xi)| \leq M(|\xi|)$ для любой величины ξ .

4. $M(\xi) \geq 0$, если $\xi \geq 0$.

5. Свойство аддитивности: $M(\xi_1 + \xi_2) = M(\xi_1) + M(\xi_2)$ для любых случайных величин ξ_1 и ξ_2 . Иначе говоря, математическое ожидание суммы случайных величин равно сумме математических ожиданий слагаемых.

Эти свойства следуют из соответствующих свойств интегралов или рядов.

6. Свойство мультипликативности: $M(\xi_1 \xi_2) = M(\xi_1) M(\xi_2)$, если случайные величины ξ_1 и ξ_2 независимы.

Это свойство следует из определения независимых случайных величин.

7. Свойство монотонности: $M(\xi) \geq M(\eta)$, если $\xi \geq \eta$.

Это свойство следует из того факта, что $P(\omega) \geq 0$, но поскольку $\xi(\omega) \geq \eta(\omega)$ или, что то же самое, $\xi(\omega) - \eta(\omega) \geq 0$, то $M(\xi - \eta) \geq 0$ по свойству 4 или, используя свойство 5, $M(\xi) - M(\eta) \geq 0$.

8. Математическое ожидание индикатора χ_A события A равно вероятности этого события $M(\chi_A) = P(A)$.

Это свойство получается в результате следующих преобразований:

$$M(\chi_A) = \sum_{\omega} \chi_A(\omega) P(\omega) = \sum_{\omega \in A} 1 \cdot P(\omega) + \sum_{\omega \notin A} 0 \cdot P(\omega) = \sum_{\omega \in A} P(\omega) = P(A).$$

9. Неравенство Коши—Буняковского: $|M(\xi \eta)| \leq \sqrt{M(\xi^2) M(\eta^2)}$, если величины справа конечны.

Доказательство. Поскольку событие $(\lambda \xi + \eta)^2$ неотрицательно при любом вещественном λ , то $M[(\lambda \xi + \eta)^2] \geq 0$ по свойству 4. Раскрыв скобки, получим:

$$M[(\lambda \xi + \eta)^2] = \lambda^2 M(\xi^2) + 2\lambda M(\xi \eta) + M(\eta^2) \geq 0.$$

Это неравенство будет иметь место при любых λ только том случае, если равенство не будет иметь решения при вещественных λ , т. е. $[M(\xi \eta)]^2 - M(\xi^2) M(\eta^2) \leq 0$ или $[M(\xi \eta)]^2 \leq M(\xi^2) M(\eta^2)$.

10. $P(\xi \geq 1) \leq M(\xi)$, если $\xi \geq 0$.

Доказательство. Пусть событие $A \equiv (\xi(\omega) \geq 1)$, тогда противоположное событие $\bar{A} \equiv (0 \leq \xi(\omega) < 1)$. Справедливы следующие преобразования:

$$P(\xi \geq 1) = \sum_{\omega \in A} P(\omega) \leq \sum_{\omega \in A} \xi(\omega) P(\omega) \leq \sum_{\omega \in A} \xi(\omega) P(\omega) + \sum_{\omega \in \bar{A}} \xi(\omega) P(\omega) = M(\xi).$$

11. Неравенство Чебышёва: $P(|\xi| \geq \varepsilon) \leq M(\xi^2)/\varepsilon^2$ для любого $\varepsilon > 0$, если $M(\xi^2)$ конечно.

Доказательство. Рассмотрим следующие преобразования:

$$P(|\xi| \geq \varepsilon) = P\left(\frac{|\xi|}{\varepsilon} \geq 1\right) = P(\eta \geq 1) \leq M(\eta) = \frac{M(\xi^2)}{\varepsilon^2},$$

где введена случайная величина $\eta = \xi^2/\varepsilon^2$, принимающая неотрицательные значения, и использовано свойство 10.

6.2. Дисперсия случайной величины

Определение. Дисперсией $D(\xi)$ случайной величины ξ называется число

$$D(\xi) = M[(\xi - M(\xi))^2], \quad (6.4)$$

если математическое ожидание в правой части (6.4) существует.

Дисперсия является мерой отклонения значений случайной величины от ее среднего значения. Величину $\sigma(\xi) = \sqrt{D(\xi)}$ называют средним квадратичным отклонением.

Свойства дисперсии

1. $D(\xi) \geq 0$ для любой случайной величины ξ .
2. $D(C) = 0$, если C — постоянная.
3. $D(\xi) = M(\xi^2) - [M(\xi)]^2$ для любой случайной величины ξ .
4. $D(C\xi) = C^2 D(\xi)$, если C — постоянная.
5. $D(\xi_1 + \xi_2) = D(\xi_1) + D(\xi_2)$, если случайные величины ξ_1 и ξ_2 независимы. Доказательство этих свойств дисперсии простое.
6. Неравенство Чебышёва:

$$P(|\xi - M(\xi)| \geq \varepsilon) \leq \frac{D(\xi)}{\varepsilon^2} \quad (6.5)$$

для любого $\varepsilon > 0$, если $D(\xi)$ конечно.

Доказательство. Достаточно воспользоваться свойством 11 математического ожидания, подставить вместо случайной величины ξ новую величину $\xi - M(\xi)$ и учесть определение дисперсии ξ . Часто бывает удобнее пользоваться неравенством Чебышёва в несколько измененном виде:

$$P(|\xi - M(\xi)| < \varepsilon) > 1 - \frac{D(\xi)}{\varepsilon^2}. \quad (6.6)$$

Пример 6.9. Найти дисперсию дискретной величины m — числа успехов в опыте из N независимых испытаний в схеме Бернулли, где $m = 0, 1, \dots, N$.

Решение. В схеме Бернулли число успехов m (вероятность успеха p и вероятность неудачи $q = 1 - p$) подчиняется биномиальному распределению:

$$P_N(m) = C_N^m p^m q^{N-m},$$

и математическое ожидание числа успехов $M(m) = Np$ (см. пример 6.2). Отсюда дисперсия m :

$$D(m) = \sum_{m=0}^N (m - Np)^2 P_N(m) = \sum_{m=2}^N \frac{N! p^m q^{N-m}}{(m-2)!(N-m)!} + \\ + \sum_{m=1}^N \frac{N! p^m q^{N-m}}{(m-1)!(N-m)!} - (Np)^2 = N(N-1)p^2 + Np - N^2p^2 = Npq.$$

Среднее квадратическое отклонение $\sigma(m) = \sqrt{D(m)} = \sqrt{Npq}$.

Пример 6.10. Найти дисперсию дискретной величины k ($k = 0, 1, 2, \dots$), распределенной по закону Пуассона.

Решение. Закон Пуассона задается формулой ($\lambda > 0$)

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

с параметром λ , имеющим смысл математического ожидания (см. пример 6.3). Тогда дисперсия величины k :

$$D(k) = \sum_{k=0}^{\infty} (k - \lambda)^2 \frac{\lambda^k}{k!} e^{-\lambda} = \lambda(\lambda + 1) e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} - \lambda^2 = \lambda.$$

Для такого распределения дисперсия совпадает с математическим ожиданием. Среднее квадратическое отклонение $\sigma(k) = \sqrt{D(k)} = \sqrt{\lambda}$.

Пример 6.11. Найти дисперсию случайной величины ξ , распределенной нормально с плотностью

$$\rho_{\xi}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-a)^2/(2\sigma^2)}.$$

Решение.

$$D(\xi) = M \left[(\xi - M(\xi))^2 \right] = \int_{-\infty}^{\infty} (x - a)^2 \rho_{\xi}(x) dx = \\ = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (x - a)^2 e^{-(x-a)^2/(2\sigma^2)} dx = \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2} dy = \sigma^2,$$

где использовано математическое ожидание $M(\xi) = a$ (см. пример 6.7). Тем самым выясняется смысл параметра σ , входящего в выражение для плотности нормального распределения как среднее квадратическое отклонение величины ξ .

Пример 6.12. Для материальной точки, находящейся на отрезке $[a, b]$, найти дисперсию координаты точки ξ , удовлетворяющей равномерному закону распределения.

Решение. Равномерное распределение на отрезке $[a, b]$ задается следующей плотностью распределения:

$$\rho_{\xi}(x) = \begin{cases} 1/(b-a), & a \leq x \leq b, \\ 0, & x < a, x > b, \end{cases}$$

а среднее значение координаты ξ равно $M(\xi) = (a+b)/2$ (см. пример 6.5). Найдем дисперсию ξ :

$$D(\xi) = \int_{-\infty}^{\infty} (x - M(\xi))^2 \rho_{\xi}(x) dx = \frac{1}{b-a} \int_a^b x^2 dx - [M(\xi)]^2 = \frac{(b-a)^2}{12}.$$

Среднее квадратическое отклонение $\sigma(\xi) = \sqrt{D(\xi)} = (b-a)/(2\sqrt{3})$.

Пример 6.13. Найти дисперсию случайной величины ξ , обладающей неотрицательными значениями и распределенной при $x \geq 0$ по экспоненциальному закону с плотностью

$$\rho_{\xi}(x) = \lambda e^{-\lambda x}.$$

Решение.

$$D(x) = \int_{-\infty}^{\infty} (x - M(\xi))^2 \rho_{\xi}(x) dx = \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx - 1/\lambda^2 = 1/\lambda^2,$$

где использовано математическое ожидание $M(\xi) = 1/\lambda$ (см. пример 6.6). Среднее квадратическое отклонение $\sigma(\xi) = \sqrt{D(\xi)} = 1/\lambda$ совпадает с математическим ожиданием.

6.3. Нормированные случайные величины

Определение. Случайная величина η называется *нормированной*, если ее математическое ожидание $M(\eta) = 0$, а дисперсия $D(\eta) = 1$.

От любой случайной величины ξ , имеющей конечные математическое ожидание и дисперсию, можно перейти к нормированной случайной величине η с помощью линейного преобразования:

$$\eta = \frac{\xi - M(\xi)}{\sqrt{D(\xi)}}.$$

Именно это преобразование и было сделано в доказательстве локальной теоремы Муавра—Лапласа, когда от случайной величины m — числа успехов в N испытаниях — перешли к новой величине $x_m = (m - Np)/\sqrt{Npq}$, используя значения математического ожидания $M(m) = Np$ и дисперсии $D(m) = Npq$ для биномиального распределения (см. примеры 6.2 и 6.9).

6.4. Моменты распределений

В предыдущих разделах были введены две числовые характеристики одномерного распределения — математическое ожидание и дисперсия. Они являются частным случаем целого набора числовых характеристик, получивших название *моментов распределения вероятности*.

Рассмотрение начнем с *начального момента порядка m* и *абсолютного (начального) момента порядка m* , где $m \geq 0$, которые имеют следующие определения:

$$\alpha_m = M(\xi^m), \quad \beta_m = M(|\xi|^m). \quad (6.7)$$

Для дискретной случайной величины это сводится к суммам степеней ее значений $\xi(\omega_k)$ с весами, равными их вероятностям $P(\omega_k)$:

$$\alpha_m = \sum_k [\xi(\omega_k)]^m P(\omega_k), \quad \beta_m = \sum_k |\xi(\omega_k)|^m P(\omega_k). \quad (6.8)$$

Для непрерывной случайной величины, распределение которой удобно задавать плотностью распределения $\rho_\xi(x)$, моменты (6.7) выражаются следующими интегралами:

$$\alpha_m = \int_{-\infty}^{+\infty} x^m \rho_\xi(x) dx, \quad \beta_m = \int_{-\infty}^{+\infty} |x|^m \rho_\xi(x) dx. \quad (6.9)$$

Видно, что математическое ожидание случайной величины — это начальный момент первого порядка, т. е. $M(\xi) = \alpha_1$.

Помимо начальных моментов, рассматриваются также *центральный момент порядка m* и *абсолютный центральный момент порядка m* , где $m \geq 0$:

$$\mu_m = M([\xi - M(\xi)]^m), \quad \mu_m^{(a)} = M(|\xi - M(\xi)|^m), \quad (6.10)$$

вычисляемых относительно центра распределения $M(\xi)$. Центральный момент первого порядка всегда равен нулю. Из этого определения также следует, что дисперсия случайной величины — это центральный момент второго порядка, т. е. $D(\xi) = \mu_2$. Если распределение симметрично относительно

центра, то все (существующие) центральные моменты нечетного порядка обращаются в ноль.

Для некоторых распределений удобно пользоваться *факториальным моментом порядка m* :

$$\alpha_{[m]} = M(\xi^{[m]}) = M[\xi(\xi - 1) \cdots (\xi - m + 1)], \quad (6.11)$$

и его центральным обобщением:

$$\mu_{[m]} = M\left\{[\xi - M(\xi)]^{[m]}\right\}. \quad (6.12)$$

Начальный и центральный моменты можно обобщить, определив момент порядка $m \geq 0$ случайной величины ξ относительно числа X , а именно следующее математическое ожидание $M[(\xi - X)^m]$.

Рассмотрим примеры вычисления высших моментов.

Пример 6.14. Найти третий и четвертый центральные моменты случайной величины, имеющей нормальное распределение $\xi \in N(a, \sigma^2)$.

Решение. В силу симметрии плотности распределения величины ξ относительно центра a третий центральный момент $\mu_3 = 0$. Вычислим четвертый центральный момент:

$$\mu_4 = \frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{+\infty} (x - a)^4 e^{-(x-a)^2/(2\sigma^2)} dx = \frac{4\sigma^4}{\sqrt{\pi}} \int_0^{\infty} y^{3/2} e^{-y} dy = 3\sigma^4.$$

Чтобы характеризовать отличие распределения произвольной случайной величины от нормального распределения, вводятся две другие числовые характеристики: *коэффициент асимметрии* γ_1 и *коэффициент эксцесса* γ_2 :

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}}, \quad \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3, \quad (6.13)$$

которые для нормального распределения обращаются в ноль.

Пример 6.15. Найти третий и четвертый центральные моменты дискретной случайной величины, распределенной по закону Пуассона.

Решение. Случайная величина ξ имеет множество спектральных значений $k = 0, 1, 2, \dots$, выпадающих с вероятностями $P(\xi = k) = (\lambda^k/k!) e^{-\lambda}$, где $\lambda > 0$. В этой задаче третий μ_3 и четвертый μ_4 центральные моменты удобно вычислять, используя начальные факториальные моменты $\alpha_{[n]}$:

$$\mu_3 = M[(\xi - \lambda)^3] = \alpha_{[3]} + 3(1 - \lambda)\alpha_{[2]} + (1 - 3\lambda + 3\lambda^2)\alpha_{[1]} - \lambda^3,$$

$$\begin{aligned}\mu_4 &= M[(\xi - \lambda)^4] = \alpha_{[4]} + (6 - 4\lambda) \alpha_{[3]} + \\ &+ (7 - 12\lambda + 6\lambda^2) \alpha_{[2]} + (1 - 4\lambda + 6\lambda^2 - 4\lambda^3) \alpha_{[1]} + \lambda^4,\end{aligned}$$

где учли, что $M(\xi) = \lambda$. Факториальный момент порядка n для распределения Пуассона легко вычисляется:

$$\alpha_{[n]} = \sum_{k=0}^{\infty} k(k-1) \cdots (k-n+1) P(\xi = k) = \lambda^n e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda^n.$$

После подстановки соответствующих моментов в μ_3 и μ_4 получим:

$$\mu_3 = \lambda, \quad \mu_4 = \lambda(1 + 3\lambda).$$

Коэффициенты асимметрии γ_1 и эксцесса γ_2 (6.13) в распределении Пуассона имеют очень простой вид:

$$\gamma_1 = \frac{1}{\sqrt{\lambda}}, \quad \gamma_2 = \frac{1}{\lambda},$$

где учтена дисперсия $D(\xi) = \mu_2 = \lambda$. В пределе $\lambda \rightarrow \infty$ эти коэффициенты обращаются в ноль и распределение Пуассона переходит в нормальное.

6.5. Ковариация. Коэффициент корреляции

Определение. Матрицей ковариации случайных величин $\xi_i = \xi_i(\omega)$ и $\xi_j = \xi_j(\omega)$ называется матрица, матричные элементы которой задаются математическими ожиданиями вида:

$$a_{ij} = \text{cov}(\xi_i, \xi_j) = M\{[\xi_i - M(\xi_i)][\xi_j - M(\xi_j)]\}. \quad (6.14)$$

Из определения следует, что матрица ковариации является симметричной: $a_{ij} = a_{ji}$. Используя свойства математического ожидания, ковариацию (6.14) можно свести к выражению:

$$\text{cov}(\xi_i, \xi_j) = M(\xi_i \xi_j) - M(\xi_i) M(\xi_j). \quad (6.15)$$

На главной диагонали в матрице ковариации стоят дисперсии

$$a_{ii} = D(\xi_i) = M[\xi_i - M(\xi_i)]^2,$$

поэтому они неотрицательны. Недиagonальные элементы, т. е. a_{ij} при $i \neq j$, называются *корреляционными моментами* случайных величин ξ_i и ξ_j .

В случае зависимых случайных величин 5-е свойство дисперсии обобщается следующим образом:

$$D(\xi_i + \xi_j) = D(\xi_i) + D(\xi_j) + 2 \operatorname{cov}(\xi_i, \xi_j),$$

которое можно обобщить на дисперсию от суперпозиции случайных величин:

$$D \left[\sum_{i=1}^n C_i \xi_i \right] = \sum_{i,j=1}^n C_i C_j a_{ij}, \quad (6.16)$$

где C_i — вещественные коэффициенты.

Из двух действительных случайных величин ξ_1 и ξ_2 , которые, вообще говоря, могут быть и зависимыми, можно построить комплексную случайную величину $\eta = \xi_1 + i\xi_2$ и комплексно сопряженную к ней величину $\eta^* = \xi_1 - i\xi_2$. Дисперсия η определяется по правилу:

$$D(\eta) = M \{ [\eta - M(\eta)] [\eta^* - M(\eta^*)] \} = D(\xi_1) + D(\xi_2),$$

т. е. корреляционный момент ξ_1 и ξ_2 отсутствует.

Из свойства неотрицательности дисперсии следует, что правая часть (6.16), являющаяся квадратичной формой коэффициентов C_i , также неотрицательна, а значит, неотрицательны все главные миноры матрицы ковариации a_{ij} . Если ограничиться какими-либо двумя величинами ξ_i и ξ_j ($i \neq j$) из имеющегося набора, то соответствующий им определитель

$$\begin{vmatrix} D(\xi_i) & \operatorname{cov}(\xi_i, \xi_j) \\ \operatorname{cov}(\xi_i, \xi_j) & D(\xi_j) \end{vmatrix} = D(\xi_i) D(\xi_j) - [\operatorname{cov}(\xi_i, \xi_j)]^2 \geq 0.$$

Отсюда следует неравенство, называемое *неравенство Коши—Буняковского*:

$$|\operatorname{cov}(\xi_i, \xi_j)| \leq \sqrt{D(\xi_i) D(\xi_j)}, \quad (6.17)$$

имеющее простой смысл: если у каких-либо двух случайных величин конечные дисперсии, то конечен и корреляционный момент этих величин.

Вместо корреляционного момента ξ_i и ξ_j при $i \neq j$, т. е. $\operatorname{cov}(\xi_i, \xi_j)$, удобно ввести другую количественную характеристику степени их зависимости:

$$r_{ij} = \frac{\operatorname{cov}(\xi_i, \xi_j)}{\sqrt{D(\xi_i) D(\xi_j)}}, \quad (6.18)$$

которая называется *коэффициентом корреляции*. Неравенство Коши—Буняковского (6.17) накладывает на коэффициент корреляции ограничение:

$$|r_{ij}| \leq 1.$$

В зависимости от значения коэффициента корреляции имеется несколько возможностей: если $r_{ij} = 0$, то ξ_i и ξ_j некоррелированы; если $r_{ij} > 0$, то корреляция между ξ_i и ξ_j положительна; если $r_{ij} < 0$, то корреляция между ξ_i и ξ_j отрицательна. Заметим, что некоррелированность ξ_i и ξ_j вовсе не означает их независимость. При этом справедливо обратное утверждение: две независимые величины некоррелированы $r_{ij} = 0$.

Пример 6.16. Найти коэффициент корреляции ξ_1 и ξ_2 при условии, что $\xi_2 = A\xi_1 + B$, где A и B — постоянные.

Решение. Пусть $M(\xi_1) = \mu$ и $D(\xi_1) = \sigma^2$, тогда для второй величины $M(\xi_2) = A\mu + B$ и $D(\xi_2) = A^2\sigma^2$. Найдем корреляционный момент:

$$\text{cov}(\xi_1, \xi_2) = M[(\xi_1 - \mu)(\xi_2 - A\mu - B)] = A D(\xi_1) = A\sigma^2.$$

Подставим полученные числовые характеристики в определение коэффициента корреляции (6.18):

$$r_{ij} = \frac{A\sigma^2}{\sqrt{\sigma^2 \times A^2\sigma^2}} = \frac{A}{|A|} = \begin{cases} 1, & \text{при } A > 0, \\ 0, & \text{при } A = 0, \\ -1, & \text{при } A < 0. \end{cases}$$

Отсюда видно, что если $A \neq 0$, т.е. случайная величина ξ_2 не постоянна, то коэффициент корреляции $|r_{ij}| = 1$.

Пример 6.17. Вычислить элементы матрицы ковариации непрерывной двумерной случайной величины, распределение которой задается плотностью:

$$\rho_{\xi_1\xi_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \exp \left\{ -\frac{1}{2(1-r^2)} \left[\frac{x_1^2}{\sigma_1^2} + \frac{x_2^2}{\sigma_2^2} - 2r \frac{x_1x_2}{\sigma_1\sigma_2} \right] \right\}.$$

Решение. Непосредственные расчеты дают для математических ожиданий $M(\xi_1) = M(\xi_2) = 0$ и матрицы ковариации

$$a_{ij} = \begin{pmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

Отсюда очевиден смысл параметра r , входящего в плотность распределения, — это коэффициент корреляции ξ_1 и ξ_2 . Видно также, что при $r = 0$ совместная плотность распределения превращается в произведение двух одномерных функций распределения случайных величин ξ_1 и ξ_2 с нормальными законами распределения, а именно: $\xi_1 \in N(0, \sigma_1^2)$ и $\xi_2 \in N(0, \sigma_2^2)$, т.е. случайные величины независимы в этом случае и матрица ковариации принимает

диагональный вид. Заметим, что квадратичная по x_1 и x_2 форма в показателе экспоненты плотности $\rho_{\xi_1\xi_2}(x_1, x_2)$ содержит матрицу вида:

$$b_{ij} = \frac{1}{1-r^2} \begin{pmatrix} 1/\sigma_1^2 & -r/(\sigma_1\sigma_2) \\ -r/(\sigma_1\sigma_2) & 1/\sigma_2^2 \end{pmatrix},$$

которая является обратной к матрице ковариации a_{ij} .

Контрольные вопросы

1. Привести определение математического ожидания дискретной и непрерывной случайной величины.
2. Указать свойства математического ожидания.
3. Привести определение и свойства дисперсии случайной величины.
4. Что такое нормировка случайной величины?
5. Дать определение ковариации случайной величины.

7. Закон больших чисел и центральные предельные теоремы

Знание вероятности или функции распределения вероятности случайной величины позволяет вычислить математическое ожидание и дисперсию случайной величины. Естественно, напрашивается вопрос об обратной задаче: по числовым характеристикам случайных величин можно ли получить информацию об их распределении и о характере функции распределения.

Ответ на первый вопрос о сходимости набора случайных величин $\{\xi_n\}$, где $n = 1, 2, \dots$, к некоторой величине ξ дает *закон больших чисел*, представляющий собой совокупность нескольких теорем.

На второй вопрос о сходимости распределения случайных величин $\{\xi_n\}$ к некоторому закону распределения вероятностей P_n отвечают *центральные предельные теоремы* теории вероятностей.

7.1. Закон больших чисел

Неравенство Чебышёва позволяет доказать ряд важных теорем, объединенных общим названием «закон больших чисел». Основная часть из этих теорем принадлежит самому П. Л. Чебышёву.

Теорема Чебышёва. Пусть имеется бесконечная последовательность независимых случайных величин x_1, x_2, \dots с одним и тем же математическим ожиданием $M(x_1) = M(x_2) = \dots = m$ и дисперсиями, ограниченными одной и той же постоянной C : $D(x_1) < C, D(x_2) < C, \dots$. Тогда, каково бы ни было положительное число ε , вероятность события $|(x_1 + x_2 + \dots + x_n)/n - m| < \varepsilon$ удовлетворяет условию:

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{x_1 + x_2 + \dots + x_n}{n} - m \right| < \varepsilon \right) = 1.$$

Доказательство. Введем случайную величину $S_n = (x_1 + x_2 + \dots + x_n)/n$. Из свойств математического ожидания следует

$$M(S_n) = \frac{1}{n} \sum_{i=1}^n M(x_i) = m.$$

Далее, так как величины x_1, x_2, \dots, x_n независимы, то

$$D(S_n) = \frac{1}{n^2} \sum_{i=1}^n D(x_i) \leq \frac{nC}{n^2} = \frac{C}{n}.$$

Сопоставляя полученное неравенство с неравенством Чебышёва, будем иметь

$$P(|S_n - m| < \varepsilon) > 1 - \frac{C}{n\varepsilon^2},$$

откуда и следует утверждение теоремы в пределе $n \rightarrow \infty$.

Смысл теоремы Чебышёва можно пояснить следующим примером. Пусть требуется измерить некоторую физическую величину m . В силу неизбежных при измерениях ошибок результат измерений будет случайной величиной. Обозначим эту величину через x , ее математическое ожидание $M(x) = m$ будет совпадать с измеряемой величиной, а дисперсия равна некоторой константе C , характеризующей точность измерительного прибора. Проведем n независимых измерений и обозначим величиной x_1 результат первого измерения, x_2 — результат второго измерения и т. д. Среднее арифметическое S_n является случайной величиной. Однако с увеличением числа измерений эта величина почти перестает быть случайной, она все более приближается к точному значению m . Тем самым оправдывается рекомендуемый в практической деятельности способ получения более точных результатов измерений: одна и та же величина измеряется многократно, и в качестве ее значения берется среднее арифметическое значение полученных результатов измерений.

Из теоремы Чебышёва можно получить другую теорему, которую до Чебышёва иным способом доказал Я. Бернулли.

Теорема Бернулли. Пусть μ_n — число успехов в серии из n независимых испытаний по схеме Бернулли и p — вероятность успеха в каждом отдельном испытании. Тогда при любом положительном ε

$$\lim_{n \rightarrow \infty} P(|\mu_n/n - p| < \varepsilon) = 1.$$

Доказательство. Для доказательства воспользуемся представлением μ_n в виде суммы индикаторов χ_k : $\mu_n = \chi_1 + \chi_2 + \dots + \chi_n$, где

$$\chi_k = \begin{cases} 1, & \text{если в } k\text{-ом испытании был успех,} \\ 0, & \text{если в } k\text{-ом испытании была неудача.} \end{cases}$$

Так как все индикаторы независимы и одинаково распределены, то их математические ожидания $M(\chi_k) = p$ и дисперсии $D(\chi_k) = p(1 - p) = pq$ одинаковы. Тогда для числа успехов μ_n получим:

$$M(\mu_n) = \sum_{i=1}^n M(\chi_i) = np, \quad D(\mu_n) = \sum_{i=1}^n D(\chi_i) = npq,$$

и теорема Бернулли — частный случай теоремы Чебышёва.

Проиллюстрируем полученные результаты примерами.

Пример 7.1. Оценить с помощью неравенства Чебышёва вероятность того, что отклонение случайной величины ξ от своего математического ожидания $M(\xi)$ будет меньше трех средних квадратичных отклонений.

Решение. Полагая $\varepsilon = 3\sigma(\xi)$ в неравенстве Чебышёва (6.6), получаем

$$P(|\xi - M(\xi)| < 3\sigma(\xi)) > 1 - \frac{[\sigma(\xi)]^2}{[3\sigma(\xi)]^2} = 1 - \frac{1}{9} = \frac{8}{9} \approx 0,8889.$$

Эта оценка, как известно, называется «правило трех сигм» и дает нижнюю границу вероятности попадания случайной величины в указанный интервал. Для сравнения: у случайной величины $\xi \in N(a, \sigma^2)$, распределенной по нормальному закону, эта вероятность равна $P(|\xi - a| < 3\sigma) = 0,9973$.

Пример 7.2. Глубина моря a измеряется прибором, не имеющим систематической ошибки. Среднее квадратическое отклонение измерений не превосходит 15 м. Сколько нужно сделать независимых измерений, чтобы с вероятностью, не меньшей 0,9, можно было утверждать, что среднее арифметическое этих измерений отличается от a по модулю меньше, чем на 5 метров?

Решение. Обозначим через X_i результаты n независимых измерений глубины моря, тогда $M(X_i) = a$, что означает отсутствие при измерениях систематической ошибки. Поскольку измерения производятся с одинаковой точностью, то $\sigma(X_i) = \sqrt{D(X_i)} = 15$ м. При заданном $\varepsilon = 5$ м нужно найти такое число n , которое удовлетворяет неравенству $P \geq 0,9$. Воспользуемся неравенством Чебышёва (6.6):

$$P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - a\right| < 5 \text{ м}\right) > 1 - \frac{225}{25n} \geq 0,9,$$

откуда следует, что $0,1 \geq 9/n$ или $n \geq 90$, т. е. измерение нужно проводить более 90 раз.

Пример 7.3. Вероятность, что машинистка сделает опечатку на одной странице рукописи, составляет $p = 0,2$. Оценить вероятность того, что в рукописи, содержащей 400 страниц, частота появления опечатки отличается от соответствующей вероятности по модулю меньше, чем на 0,05.

Решение. В данной задаче $n = 400$, $p = 0,2$, $q = 1 - p = 0,8$, и $\varepsilon = 0,05$. Пусть μ_n — количество опечаток в тексте, а $\nu_n = \mu_n/n$ — частота появления опечатки, тогда, воспользовавшись неравенством Чебышёва (6.6), получим:

$$P(|\nu_n - 0,2| < 0,05) > 1 - \frac{0,2 \times 0,8}{400 \times 0,05^2} = 0,84,$$

где учтено, что $D(\mu_n) = pq$.

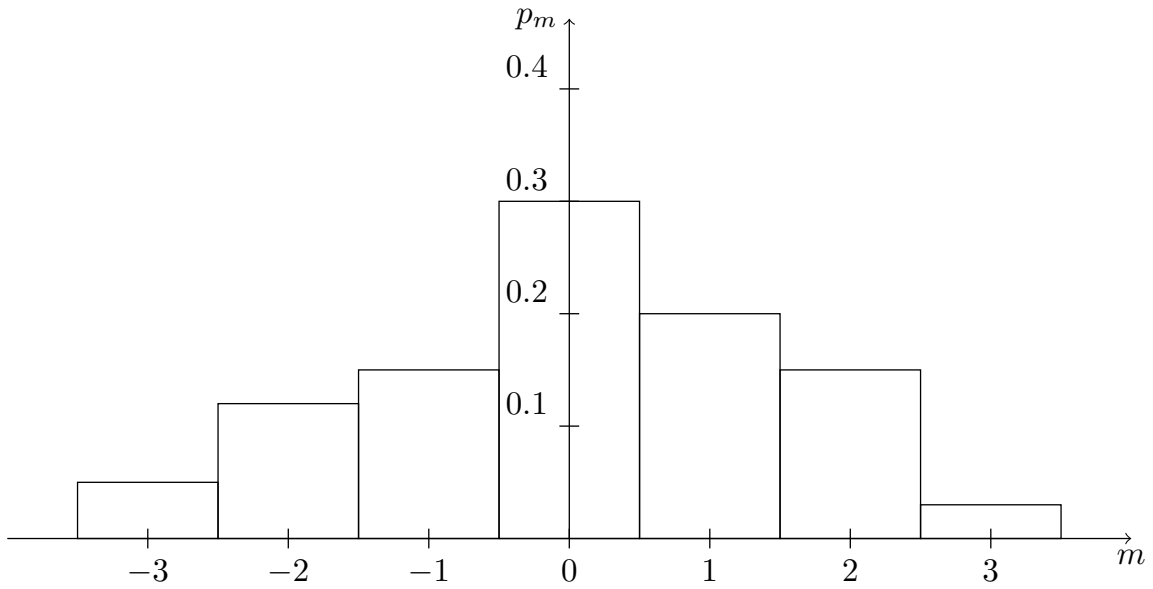


Рис. 5. Представление распределения случайной величины посредством прямоугольников единичного основания

7.2. Центральные предельные теоремы

Предварительное понимание содержания центральных предельных теорем может быть получено следующим способом.

Рассмотрим сумму независимых случайных величин $\eta_n = \xi_1 + \xi_2 + \dots + \xi_n$, где $n = 1, 2, 3, \dots$, которые принимают целочисленные значения и все имеют одинаковые распределения $P(\xi_i = m) = p_m$, $m = 0, \pm 1, \pm 2, \dots$ для всех i .

Это распределение можно представить следующим образом: построим график из прямоугольников, основание каждого прямоугольника равно 1, а высота — p_m , так что площадь равна p_m и $\sum_m p_m = 1$. В общем случае получим практически произвольный набор прямоугольников (см. рис. 5).

Рассмотрим теперь вместо η_n нормированные случайные величины

$$\eta_n^* = \frac{\eta_n - M(\eta_n)}{\sqrt{D(\eta_n)}} = \frac{\eta_n - \mu}{\sigma\sqrt{n}},$$

где $\mu = M(\eta_i)$ и $\sigma^2 = D(\eta_i)$ для всех i из набора всех n случайных величин. Значениями случайной величины η_n^* являются числа $x_n(m) = (m - n\mu)/\sqrt{n\sigma^2}$, причем $P(\eta_n^* = x_n(m)) = P(\eta_n = m)$.

Построим теперь аналогичный «график» распределения η_n^* . По оси абсцисс отложим значения $x_n(m)$, $m = 0, \pm 1, \pm 2, \dots$ и, как раньше, построим прямоугольники, площади которых равны $P(\eta_n = m)$.

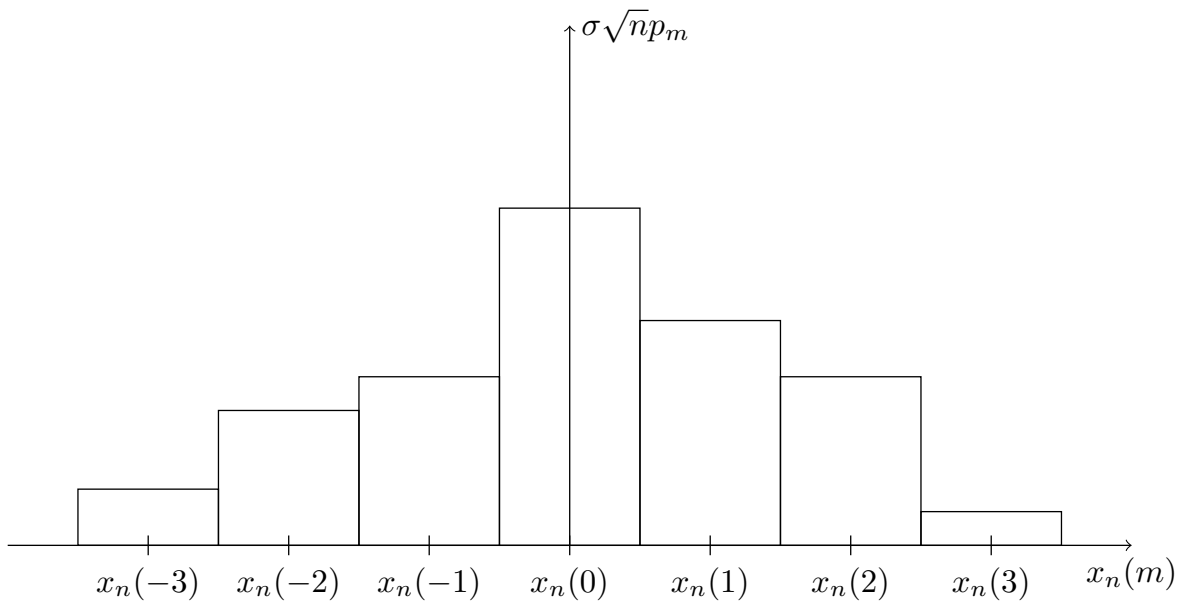


Рис. 6. Представление распределения нормированной случайной величины η_n^*

Поскольку длины основания этих прямоугольников теперь равны $1/(\sigma\sqrt{n})$, то их высоты должны быть равны $\sigma\sqrt{n}p_m$, т. е. площади остаются неизменными (см. рис. 6). Как показывает теория, при достаточно большом n верхние основания прямоугольников почти точно ложатся на фиксированную кривую

$$y(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

то есть

$$\sigma\sqrt{n} P(\eta_n^* = x_n(m)) \approx \frac{1}{\sqrt{2\pi}} e^{-x_n^2(m)/2} \quad \text{при } n \rightarrow \infty.$$

При этом естественно, что

$$P(a \leq \eta_n^* \leq b) = \sum_m P(a \leq x_n(m) \leq b) \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx.$$

Этот факт, вытекающий из интегральной предельной теоремы Муавра—Лапласа, и составляет, по существу, содержание центральных предельных теорем, которые отличаются друг от друга формулировкой различных математических условий, обеспечивающих различные обобщения теоремы. Теорему Муавра—Лапласа можно сформулировать следующим образом.

Теорема Муавра—Лапласа. Если случайные величины $\xi_1, \xi_2, \dots, \xi_n$ независимы и $P(\xi_k = 1) = 1 - P(\xi_k = 0) = p$ для любого $k = 1, 2, \dots, n$, то

при $0 < p < 1$ и $n \rightarrow \infty$ равномерно по $x \in (a, b)$:

$$P\left(a < \frac{\xi_1 + \xi_2 + \dots + \xi_n - n M(\xi_1)}{\sqrt{n D(\xi_1)}} < b\right) \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx.$$

Эта теорема позволяет сделать обобщения на другие случайные независимые величины. Приведем без доказательства две теоремы.

Теорема 1. Если случайные величины $\xi_1, \xi_2, \dots, \xi_n$ независимы, одинаково распределены и имеют конечную дисперсию, то при $n \rightarrow \infty$ равномерно по $x \in (-\infty, \infty)$:

$$P\left(\frac{\xi_1 + \xi_2 + \dots + \xi_n - na}{\sigma\sqrt{n}} < x\right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-x^2/2} dx,$$

где $a = M(\xi_k)$, $\sigma^2 = D(\xi_k)$ для любого $k = 1, \dots, n$.

Теорема 2. Пусть $\xi_1, \xi_2, \dots, \xi_n$ — независимые случайные величины, имеющие конечный третий абсолютный момент: $M(|\xi_k - M(\xi_k)|^3) < \infty$. Определим числовые характеристики $a_k = M(\xi_k)$, $b_k^2 = D(\xi_k)$, $c_k^3 = M((\xi_k - a_k)^3)$ и их суммы:

$$A_n = \sum_{k=1}^n a_k, \quad B_n^2 = \sum_{k=1}^n b_k^2, \quad C_n^3 = \sum_{k=1}^n c_k^3.$$

Если при $n \rightarrow \infty$ отношение $C_n/B_n \rightarrow 0$, то

$$P\left(\frac{\xi_1 + \xi_2 + \dots + \xi_n - A_n}{B_n} < x\right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-x^2/2} dx.$$

В заключение отметим, что нельзя пользоваться предельными теоремами в тех случаях, когда вероятность $P((\xi_1 + \xi_2 + \dots + \xi_n - A_n)/B_n < x)$ близка к единице или к нулю, поскольку «хвосты» распределения требуют очень осторожной оценки.

7.3. Применение центральных предельных теорем

7.3.1. Вероятность и частота

Оценим, насколько сильно может отличаться частота от вероятности в серии из n независимых испытаний по схеме Бернулли. Используя теорему

Муавра—Лапласа, находим:

$$\begin{aligned} P\left(\left|\frac{\eta_n}{n} - p\right| < \varepsilon\right) &= P\left(\left|\frac{\eta_n - np}{\sqrt{npq}}\right| < a\right) \\ &\approx \frac{1}{\sqrt{2\pi}} \int_{-a}^a e^{-x^2/2} dx = \Phi_0(a) - \Phi_0(-a) = 2\Phi_0(a), \end{aligned}$$

где $a = \varepsilon\sqrt{n/pq}$, η_n — число успехов в серии из n испытаний и введена функция, называемая *интегралом вероятностей*:

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz = \frac{1}{2} \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right),$$

где $\operatorname{erf}(x) = (2/\sqrt{\pi}) \int_0^x e^{-t^2} dt$ — функция ошибок. Отсюда также следует

$$P\left(\left|\frac{\eta_n - np}{\sqrt{npq}}\right| < \varepsilon\sqrt{\frac{n}{pq}}\right) = P\left(\frac{\eta_n}{n} - \varepsilon < p < \frac{\eta_n}{n} + \varepsilon\right) = \alpha.$$

Интервал значений $(\eta_n/n - \varepsilon, \eta_n/n + \varepsilon)$ называется *доверительным интервалом* для вероятности p с уровнем доверия α .

7.3.2. Среднее арифметическое

Пусть $\xi_1, \xi_2, \dots, \xi_n$ — независимые случайные величины с одинаковыми математическими ожиданиями $M(\xi_k) = \mu$ и дисперсиями $D(\xi_k) = \sigma^2$. Тогда при $n \rightarrow \infty$ вероятность

$$P(|(\xi_1 + \xi_2 + \dots + \xi_n)/n| < \varepsilon) = 2\Phi_0(\varepsilon\sqrt{n}/\sigma).$$

Из таблиц нормального распределения следует, что при $\varepsilon\sqrt{n}/\sigma = 3$ вероятность $P(|(\xi_1 + \xi_2 + \dots + \xi_n)/n| < 3\sigma/\sqrt{n}) = 0,9973$. Это так называемое «правило трех сигм».

Пример 7.4. Независимые случайные величины X_i распределены равномерно на отрезке $[0, 1]$. Найти закон распределения случайной величины

$$Y = \sum_{i=1}^{100} X_i,$$

а также вероятность того, что $55 < Y < 70$.

Решение. Условия ЦПТ соблюдаются, поэтому случайная величина Y имеет приближенно плотность распределения

$$f_Y(y) \approx \frac{1}{\sqrt{2\pi D(Y)}} e^{-[y-M(Y)]^2/[2D(Y)]}.$$

По известным формулам для математического ожидания и дисперсии в случае равномерного распределения на отрезке находим: $M(X_i) = (0 + 1)/2 = 1/2$, $D(X_i) = (1 - 0)^2/12 = 1/12$, $\sigma(X_i) = \sqrt{D(X_i)} = 1/(2\sqrt{3})$. Тогда

$$M(Y) = M\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} M(X_i) = 100 \times \frac{1}{2} = 50,$$

$$D(Y) = D\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} D(X_i) = 100 \times \frac{1}{12} = \frac{25}{3},$$

и

$$f_Y(y) \approx \frac{3}{5\sqrt{6\pi}} e^{-3(y-50)^2/50}.$$

Используя интегральную теорему Муавра—Лапласа

$$P(\alpha \leq Y \leq \beta) \approx \Phi\left(\frac{\beta - M(Y)}{\sqrt{D(Y)}}\right) - \Phi\left(\frac{\alpha - M(Y)}{\sqrt{D(Y)}}\right),$$

находим

$$P(55 < Y < 70) = \Phi(4\sqrt{3}) - \Phi(\sqrt{3}) = \Phi(6,9282) - \Phi(1,73) \approx 0,04.$$

Пример 7.5. Машинистке требуется напечатать текст из 8 000 слов, каждое из которых содержит не менее четырех букв. Вероятность сделать ошибку в любом из этих слов равна 0,01. Какова вероятность, что в напечатанном тексте будет сделано не более 90 ошибок?

Решение. Применим формулу

$$P_n(k \leq m) = \sum_{k=0}^m P_n(k) \approx \Phi_0\left(\frac{m - np}{\sqrt{npq}}\right).$$

Так как $n = 8000$, $p = 0,01$, $q = 1 - p = 0,99$, $m = 90$, то $npq = 79,2$ и

$$\frac{m - np}{\sqrt{npq}} \approx \frac{10}{8,9} \approx 1,12.$$

Используя значение $\Phi_0(1,12) = 0,8686$, получаем $P_n(k \leq 90) \approx 0,869$.

Контрольные вопросы

1. В чем состоит смысл теоремы Чебышёва?
2. Привести доказательство теоремы Чебышёва.
3. Что такое «индикатор успеха»?
4. Привести доказательство теоремы Бернулли.
5. Что утверждает теорема Муавра—Лапласа?
6. Сформулировать центральные предельные теоремы.
7. Рассказать о применении центральных предельных теорем.

8. Случайные процессы

8.1. Понятие о случайном процессе

Случайные процессы (случайные функции) — это случайные величины, зависящие от параметров, принимающих дискретное или непрерывное множество значений. В любой автоматической системе, в любой системе, связанной с управлением, приходится учитывать воздействие случайных помех, являющихся функциями времени, температуры, давления и т. д. Теория случайных процессов находит широкое применение во многих разделах физики и техники.

Определение. Пусть $\{\Omega, \mathcal{F}, P\}$ — вероятностное пространство, T — некоторое числовое множество. Действительная функция

$$\xi(t) = f(t, \omega), \quad (8.1)$$

определенная при $t \in T$, $\omega \in \Omega$, называется *случайным процессом* (случайной функцией), если $f(t, \omega)$, как функция $\omega \in \Omega$, является случайной величиной при каждом $t \in T$.

Если значение ω фиксировано, то $f(t, \omega)$, рассматриваемая как функция t , $t \in T$, называется *реализацией случайного процесса* $\xi(t)$, или *выборочной функцией*.

Если значение t фиксировано, то случайная величина $f(t, \omega)$ называется *сечением случайного процесса* $\xi(t)$ в точке $t \in T$. Величина t может меняться непрерывно или же принимать ряд дискретных значений. Если T является подмножеством множества целых чисел, то случайный процесс называется процессом с дискретным временем или случайной последовательностью, примером которой является марковская цепь.

Рассмотрим некоторые частные случаи.

8.2. Пуассоновский поток событий

Рассмотрим некоторое событие A , которое может происходить в случайные моменты времени, и пусть $\xi(t)$ — число наступлений события A в промежутке времени t . Пусть выполняются условия:

1. Случайные величины $\xi(t_i)$, где $i = 1, 2, 3, \dots$, для не пересекающихся промежутков времени независимы в совокупности.

2. Для любого промежутка времени вероятность наступления события A в этом промежутке зависит лишь от длины этого промежутка (свойство однородности процесса по времени).

3. $P(\xi(t) = 1) = \lambda t + \mathcal{O}(t)$, где $\lambda > 0$, и $P(\xi(t) > 1) = \mathcal{O}(t)$ при $t \rightarrow 0$.

Сформулированным условиям подчиняются многие реальные явления: распад радиоактивного вещества, отказы радиоэлектронной аппаратуры, вызовы на телефонной станции, запросы на обслуживание станков и т. д.

Теорема. Пусть выполнены условия 1—3. Тогда распределение случайного вектора $\{\xi(t_1), \xi(t_2), \dots, \xi(t_r)\}$ для не пересекающихся промежутков времени t_1, t_2, \dots, t_r является пуассоновским:

$$P(\xi(t_1) = k_1, \dots, \xi(t_r) = k_r) = \frac{(\lambda t_1)^{k_1}}{k_1!} \dots \frac{(\lambda t_r)^{k_r}}{k_r!} e^{-\lambda(t_1 + \dots + t_r)}.$$

Поскольку доказательство этой теоремы довольно простое, здесь мы приводить его не будем.

Случайный процесс, подчиняющийся условиям 1—3 и этой теореме, называется *процессом Пуассона* или *пуассоновским потоком событий*.

Замечание. Поскольку процесс однороден по времени, то интервалы t_1, t_2, \dots можно отсчитывать не от нуля, а от произвольного фиксированного момента времени s_0 , т. е. в роли случайных величин $\xi(t_j)$ могут выступать разности: $\xi(t_1) - \xi(s_0), \xi(t_2) - \xi(t_1), \dots$, и случайная величина $\xi(t_1) - \xi(s_0)$ имеет распределение

$$P(\xi(t_1) - \xi(s_0) = k) = \frac{[\lambda(t_1 - s_0)]^k}{k!} e^{-\lambda(t_1 - s_0)}.$$

Если дополнительно потребовать, чтобы $s_0 = 0$, то говорят, что процесс начинается в нуле.

Пример 8.1. Обозначим через τ время ожидания первого события в пуассоновском потоке событий и найдем его функцию распределения. Ясно, что $P(\tau > t) = P(\xi(t) - \xi(0) = 0) = \exp(-\lambda t)$ и, таким образом, функция распределения времени ожидания первого события равна:

$$F_\tau(t) = P(\tau < t) = 1 - e^{-\lambda t}.$$

Следовательно, плотность распределения времени ожидания τ имеет вид:

$$\rho_\tau(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Среднее время ожидания или математическое ожидание τ :

$$M(\tau) = \int_{-\infty}^{\infty} t \rho_{\tau}(t) dt = \int_0^{\infty} t \lambda e^{-\lambda t} dt = \frac{1}{\lambda},$$

т. е. λ — среднее число событий, происходящих в единицу времени.

8.3. Винеровский процесс

Винеровским процессом $\xi(t)$, где $0 \leq t < \infty$, называется случайный процесс, обладающий следующими свойствами:

1. Для любых $0 \leq t_0 < t_1 < t_2 < t_3 < \dots < t_n$ случайные величины $\xi_1 = \xi(t_1) - \xi(t_0)$, $\xi_2 = \xi(t_2) - \xi(t_1)$, \dots , $\xi_n = \xi(t_n) - \xi(t_{n-1})$ независимы.
2. Случайные величины ξ_k имеют нормальное распределение

$$\rho_{\xi_k}(x) = \frac{1}{\sqrt{2\pi\sigma^2(t_k - t_{k-1})}} e^{-x^2/[2\sigma^2(t_k - t_{k-1})]},$$

где постоянную процесса σ^2 обычно полагают равной 1.

3. Процесс начинается в нуле, т. е. $\xi(0) = 0$.

Из всего многообразия процессов рассмотрим только один.

Пример 8.2. Движение броуновской частицы. Представим себе частицу, взвешенную в однородной жидкости. Из-за столкновений с молекулами жидкости она совершает хаотическое движение. Обозначим через ξ смещение частицы по оси Ox от равновесного положения, причем из-за хаотичности движения величину ξ можно рассматривать как случайную. Очевидно,

$$M(\xi) = \int_{-\infty}^{\infty} x \rho_{\xi}(x) dx = 0, \quad D(\xi) = \int_{-\infty}^{\infty} x^2 \rho_{\xi}(x) dx = \sigma^2 t.$$

Отсюда видно, что дисперсия является линейной функцией времени t . Постоянная σ^2 называется *коэффициентом диффузии* и зависит как от параметров частицы, так и от вязкости жидкости. Аналогичными свойствами обладают смещения и по двум другим осям Oy и Oz .

8.4. Марковские процессы

Частный случай марковских процессов — конечные однородные цепи Маркова — уже был рассмотрен в разделе 4. В этой подразделе мы представим основные свойства марковских процессов с непрерывным временем.

Определение. Случайный процесс $\xi(t)$ называется *марковским*, если для любого момента времени t случайная величина $\xi(t)$ при известном значении $\xi(t_1)$ в предыдущий момент времени $t_1 < t$ не зависит от случайных величин $\xi(s)$, определенных в моменты $s < t_1$.

Иными словами, в марковских процессах (процессах без последствия) вероятностные свойства процесса в момент времени t определяются состоянием системы в момент t_1 , где $t > t_1$, но не зависят от состояний системы во все предыдущие моменты времени.

Рассмотрение марковских процессов начнем со случая, когда система имеет дискретный набор состояний $\{x_i\}$. Введем вероятность перехода

$$P_{ji}(t, s) = P(\xi(t) = x_j, \xi(s) = x_i) \geq 0,$$

которая удовлетворяет условиям:

$$P_{ji}(t, t) = \delta_{ji}, \quad \sum_j P_{ji}(t, s) = 1.$$

Из формулы для полной вероятности следует:

$$P_{ji}(t, t_0) = \sum_k P_{jk}(t, s) P_{ki}(s, t_0), \quad (8.2)$$

при $t_0 < s < t$. Если ввести в рассмотрение предельную вероятность $P_j(t)$, которая получается из $P_{ji}(t, t_0)$ в пределе $t \rightarrow \infty$, то при $t > s$ она должна удовлетворять уравнению:

$$P_j(t) = \sum_i P_{ji}(t, s) P_i(s).$$

Теорема Колмогорова. Пусть переходные вероятности $P_{ji}(t, s)$ имеют частные производные по t и s . Тогда при $s \leq t$

$$\frac{\partial P_{ji}(t, s)}{\partial t} = \sum_k A_{jk}(t) P_{ki}(t, s), \quad (8.3)$$

$$\frac{\partial P_{ji}(t, s)}{\partial s} = \sum_k P_{jk}(t, s) A_{ki}(s), \quad (8.4)$$

где

$$A_{ik}(t) = \left. \frac{\partial P_{ik}(t, s)}{\partial s} \right|_{s=t},$$

причем $A_{ik} \geq 0$ при $i \neq k$, $A_{ii} \leq 0$, и выполняется условие $\sum_i A_{ik}(t) = 0$.

Уравнения (8.3) называется *прямой системой уравнений Колмогорова*, а уравнения (8.4) — *обратной системой*.

Доказательство. Используя (8.2) и (8.3), при $t > s$ получаем:

$$\begin{aligned} \frac{\partial P_{ji}(t, s)}{\partial t} &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} [P_{ji}(t + \Delta t, s) - P_{ji}(t, s)] = \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \sum_k [P_{jk}(t + \Delta t, t) P_{ki}(t, s) - P_{ji}(t, s)] = \\ &= \lim_{\Delta t \rightarrow 0} \left[\sum_{k \neq j} \frac{P_{jk}(t + \Delta t, t) - P_{jk}(t, t)}{\Delta t} P_{ki}(t, s) + \right. \\ &\quad \left. + P_{ji}(t, s) \frac{P_{jj}(t + \Delta t, t) - 1}{\Delta t} \right] = \\ &= \sum_{k \neq j} A_{jk}(t) P_{ki}(t, s) + A_{jj}(t) P_{ji}(t, s) = \sum_k A_{jk}(t) P_{ki}(t, s). \end{aligned}$$

Здесь мы учли следующие очевидные свойства:

$$\sum_i P_{ik}(t + \Delta t, t_1) = 1, \quad \lim_{\Delta t \rightarrow 0} P_{jk}(t + \Delta t, t) = \delta_{jk}.$$

Аналогично можно получить уравнения для обратной системы (8.4).

Отправляясь от прямой системы уравнений Колмогорова (8.3), установим дифференциальные уравнения для предельных вероятностей $P_j(t)$. Если $P_i(t_0)$ — распределения в момент времени t_0 , то

$$P_j(t) = \sum_i P_{ji}(t, t_0) P_i(t_0).$$

Умножим обе части системы (8.3) на $P_i(t_0)$, сделав подстановку $s = t_0$, и просуммируем по i , тогда получим систему дифференциальных уравнений:

$$\frac{dP_j(t)}{dt} = \sum_k A_{jk}(t) P_k(t)$$

с начальными условиями $P_i(t_0) = a_i$.

Для однородного марковского процесса переходные вероятности зависят лишь от разности $(t - t_0)$. В этом случае A_{ij} — константы и система принимает линейный вид:

$$\frac{dP_j(t)}{dt} = \sum_k A_{jk} P_k(t). \quad (8.5)$$

Пример 8.3. Двусторонняя реакция. Будем считать, что система может находиться в двух состояниях: не распавшаяся ω_1 и распавшаяся ω_2 частицы. Пусть возможен как процесс распада $\omega_1 \rightarrow \omega_2$ с вероятностью перехода в единицу времени $A_{21} = \alpha$, так и обратный переход $\omega_2 \rightarrow \omega_1$ с вероятностью $A_{12} = \beta$. Из условия $\sum_{k=1}^2 A_{kj} = 0$ следует, что $A_{11} = -\alpha$ и $A_{22} = -\beta$. Тогда из (8.5) получаем:

$$\frac{dP_1(t)}{dt} = -\alpha P_1(t) + \beta P_2(t), \quad \frac{dP_2(t)}{dt} = \alpha P_1(t) - \beta P_2(t).$$

Легко заметить, что $P_1(t) + P_2(t)$ не зависит явно от времени. Зададим начальное распределение вероятностей: $P_1(0) = 1$ и $P_2(0) = 0$, т. е. распавшиеся частицы отсутствуют. Тогда

$$P_1(t) = \frac{\beta}{\alpha + \beta} + \frac{\alpha}{\alpha + \beta} e^{-(\alpha + \beta)t}, \quad P_2(t) = \frac{\alpha}{\alpha + \beta} [1 - e^{-(\alpha + \beta)t}].$$

При $t \rightarrow \infty$ вероятности принимают конечные значения:

$$P_1(t) = \frac{\beta}{\alpha + \beta}, \quad P_2(t) = \frac{\alpha}{\alpha + \beta},$$

то есть процесс эргодичен.

Если восстановление невозможно, как в случае радиоактивного распада, то $\beta = 0$ и

$$P_1(t) = e^{-\alpha t}, \quad P_2(t) = 1 - e^{-\alpha t},$$

где α — вероятность распада в единицу времени.

Пример 8.4. Пуассоновский поток требований. Пусть на некоторую систему обслуживания поступают требования так, что $\xi(t)$ — число требований за время t , которые образуют однородный марковский процесс со счетным числом состояний ω_i , где $i = 0, 1, 2, 3, \dots$. Из состояния i система может непосредственно перейти только в состояние $i + 1$. Таким образом, $A_{i+1,i} = \alpha$, а остальные $A_{ji} = 0$ при $i \neq j$. Из условия $\sum_{k=1}^2 A_{kj} = 0$ следует, что $A_{ii} = -\alpha$.

Тогда для предельной вероятности $P_j(t)$ из (8.5) с начальными условиями $P_0(0) = 1$ и $P_j(0) = 0$ при $j > 0$ получаем систему уравнений:

$$\frac{dP_0(t)}{dt} = -\alpha P_0(t), \quad \frac{dP_j(t)}{dt} = -\alpha P_j(t) + \alpha P_{j-1}(t). \quad (8.6)$$

Эта система легко решается, если ввести новую функцию $q_j(t)$:

$$P_j(t) = e^{-\alpha t} q_j(t). \quad (8.7)$$

Тогда система (8.6) преобразуется к виду:

$$\frac{dq_0(t)}{dt} = 0, \quad \frac{dq_j(t)}{dt} = \alpha q_{j-1}(t)$$

с начальными условиями $q_0(0) = 1$ и $q_j(0) = 0$. Методом итераций получаем решение этой системы:

$$q_0(t) = 1, \quad q_1(t) = \alpha t, \quad q_2(t) = \frac{(\alpha t)^2}{2}, \quad \dots, \quad q_j(t) = \frac{(\alpha t)^j}{j!},$$

и, подставив в (8.7), искомую вероятность:

$$P_j(t) = \frac{(\alpha t)^j}{j!} e^{-\alpha t},$$

то есть рассмотренный поток требований является пуассоновским.

8.5. Теория непрерывных случайных процессов

До сих пор рассматривались процессы, в которых случайная величина принимала ряд дискретных значений. Для таких процессов была сформулирована теорема А. Н. Колмогорова. Представляет интерес рассмотреть также ситуацию, когда случайная величина изменяется непрерывно. С такими задачами очень часто сталкиваются в теории газовой динамики, в теории конденсированных сред, в статистической радиофизике. Большой вклад в развитие этого направления внесли А. Н. Колмогоров, В. Феллер и А. И. Хинчин.

По аналогии с матрицей переходов для дискретных случайных величин введем функцию $F(\tau, y; t, x)$, равную вероятности того, что в момент τ случайная величина $\xi(\tau)$ примет значение, меньшее y , если известно, что в момент t ($t < \tau$) имело место равенство $\xi(t) = x$. Дополнительное знание состояний системы в более ранние, чем t , моменты времени для данных процессов не изменяет функцию $F(\tau, y; t, x)$.

Функция $F(\tau, y; t, x)$ удовлетворяет условиям:

$$1. \lim_{y \rightarrow -\infty} F(\tau, y; t, x) = 0 \text{ и } \lim_{y \rightarrow +\infty} F(\tau, y; t, x) = 1.$$

2. Функция $F(\tau, y; t, x)$ непрерывна относительно аргумента y .

Рассмотрим моменты времени t, s, τ , упорядоченные следующим образом: $t < s < \tau$. Из формулы для полной вероятности можно вывести обобщенное уравнение Маркова:

$$F(\tau, y; t, x) = \int F(\tau, y; s, z) dF(s, z; t, x).$$

Если ввести в рассмотрение плотность вероятности:

$$F(\tau, y; t, x) = \int_{-\infty}^y \rho(\tau, z; t, x) dz,$$

то обобщенное уравнение Маркова может быть записано в виде:

$$\rho(\tau, y; t, x) = \int_{-\infty}^{+\infty} \rho(\tau, y; s, z) \rho(s, z; t, x) dz.$$

Это выражение часто называют *уравнением Смолуховского*.

А. Будем считать, что существуют частные производные:

$$\frac{\partial}{\partial t} \rho(\tau, y; t, x), \quad \frac{\partial}{\partial x} \rho(\tau, y; t, x), \quad \frac{\partial^2}{\partial x^2} \rho(\tau, y; t, x),$$

при любых значениях t, x и $\tau > t$;

Б. При любых $\delta > 0$ существуют пределы

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|y-x| < \delta} (y-x) \rho(t + \Delta t, x; t, y) dy = a(t, x),$$

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|y-x| < \delta} (y-x)^2 \rho(t + \Delta t, x; t, y) dy = b(t, x).$$

Первая функция часто называется «коэффициент сноса», а вторая — «коэффициент диффузии», которые при переходе от случайной величины к случайному вектору становятся «вектором сноса» и «матрицей коэффициентов диффузии» соответственно.

В. Существуют непрерывные частные производные:

$$\frac{\partial}{\partial \tau} \rho(\tau, y; t, x); \quad \frac{\partial}{\partial y} \rho(\tau, y; t, x); \quad \frac{\partial^2}{\partial y^2} \rho(\tau, y; t, x).$$

А. Н. Колмогоровым было показано, что при выполнении этих условий и дополнительных условий, накладываемых на усредняемую функцию (обнуление самой функции и ее первых производных на бесконечности), плотность вероятности $\rho(\tau, y; t, x)$ удовлетворяет уравнению:

$$\frac{\partial}{\partial \tau} \rho(\tau, y; t, x) = -\frac{\partial}{\partial y} [a(\tau, y) \rho(\tau, y; t, x)] + \frac{1}{2} \frac{\partial^2}{\partial y^2} [b(\tau, y) \rho(\tau, y; t, x)],$$

которое называется *вторым уравнением Колмогорова*. Необходимо только отметить, что подобное уравнение было выведено до Колмогорова физиками А. Фоккером и М. Планком при обобщении теории диффузии и известно как *уравнение Фоккера—Планка*.

8.6. Стационарные случайные процессы

Еще одна категория случайных процессов, которая активно используется во многих разделах физики и технических приложениях, — это стационарные случайные процессы. В качестве примеров стационарных случайных процессов можно привести такие, как колебания напряжения или силы тока в электрической осветительной сети, случайные шумы в радиоприемнике, колебания самолета на установившемся режиме горизонтального полета, процесс качки корабля и т. п.

Часто параметром t случайного процесса $\xi(t)$ является время, в этом случае такие процессы можно различать по степени однородности их протекания во времени. Если у процесса имеется некоторая тенденция развития, то его числовые характеристики, вообще говоря, эволюционируют, т. е. изменяются при переходе от одного момента времени к другому. Такие случайные процессы принято называть *нестационарными*. Математическое описание таких процессов — достаточно сложная задача. В некоторых случаях для упрощения математической модели процесса и расчетов его характеристик нестационарный процесс можно заменить *стационарным*.

Случайный процесс $\xi(t)$ является стационарным, если его многомерная плотность вероятности $\rho_\xi(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$ зависит только от величины интервалов $t_2 - t_1, t_3 - t_1, \dots, t_n - t_1$ и не зависит от положения этих

интервалов в области изменения аргумента t_1 . Отсюда следует, в частности, что, во-первых, для стационарного процесса одномерная плотность вероятности $\rho_\xi(x; t) = \rho_\xi(x)$ не зависит от времени и, во-вторых, двумерная плотность $\rho_\xi(x_1, x_2; t_1, t_2) = \rho_\xi(x_1, x_2; u)$ зависит от разности $u = t_2 - t_1$. Как следствие, все моменты одномерного распределения постоянны. Часто бывает достаточно для определения случайного процесса как стационарного постоянства первых двух моментов:

$$M[\xi(t)] = M[\xi(0)] = a, \quad D[\xi(t)] = D[\xi(0)] = \sigma^2,$$

а также зависимости корреляционного момента реализаций процесса в моменты времени t и $t + u$

$$R_\xi(u) = M[\xi(t+u)\xi(t)] - M[\xi(t+u)]M[\xi(t)] = M[\xi(u)\xi(0)] - a^2$$

только от u . Корреляционная функция любой случайной функции обладает свойством симметрии $R_\xi(-u) = R_\xi(u)$. Если перейти к нормированному случайному процессу $\eta(t) = [\xi(t) - a]/\sigma$, то этот процесс будет характеризоваться $M[\eta(t)] = 0$, $D[\eta(t)] = 1$ и $R_\eta(u) = M[\eta(u)\eta(0)]$. Как и в случае многомерной случайной величины, взаимозависимость ошибок удобно характеризовать коэффициентом корреляции:

$$r_\xi(u) = \frac{R_\xi(u)}{R_\xi(0)} = \frac{R_\xi(u)}{D[\xi(0)]},$$

который обладает очевидным свойством $r_\xi(0) = 1$.

Пример 8.5. Рассмотрим случайный процесс вида

$$\chi(t) = \sum_{k=1}^n C_k [\xi_k \cos(\lambda_k t) + \eta_k \sin(\lambda_k t)],$$

где λ_k ($k = 1, 2, \dots, n$) — набор спектральных значений, ξ_k и η_k — случайные величины и действительные коэффициенты C_k удовлетворяют условию нормировки $\sum_{k=1}^n C_k^2 = 1$. Если ξ_k и η_k имеют числовые характеристики:

$$M(\xi_k) = M(\eta_k) = 0, \quad M(\xi_k \xi_m) = M(\eta_k \eta_m) = \delta_{km}, \quad M(\xi_k \eta_m) = 0,$$

где δ_{km} — символ Кронекера, т. е. они попарно независимы, то корреляционный момент можно привести к виду:

$$R_\chi(u) = \sum_{k=1}^n C_k^2 \cos[\lambda_k u].$$

Если учесть, что $M[\chi(t)] = 0$, то случайный процесс $\chi(t)$ является стационарным. Корреляционному моменту $R_\chi(u)$ можно сопоставить некоторую функцию распределения $F(x)$, которая будет расти скачками в точках $\pm\lambda_k$ на $C_k^2/2$, и процесс, обладающий такой функцией, назвали *процессом с дискретным спектром*.

Если $\xi(t)$ представляет собой случайный ток или напряжение в электрической цепи, то $M[\xi(u)]$ — это постоянная составляющая, а $R_\xi(u)$ — средняя мощность флуктуации этой величины.

Контрольные вопросы

1. Что называется случайным процессом?
2. Сформулировать свойства, которым подчиняется пуассоновский поток событий.
3. Написать распределение случайного вектора для пуассоновского потока событий.
4. Сформулировать свойства, которым подчиняется винеровский поток событий.
5. Какой процесс называется марковским?
6. Записать прямую и обратную систему уравнений Колмогорова.
7. Описать процесс, называемый двусторонней реакцией.
8. Описать процесс, называемый пуассоновским потоком событий.
9. Записать уравнение Смолуховского.
10. Записать уравнение Фоккера—Планка.
11. Какой процесс называется стационарным?

9. Элементы математической статистики

9.1. χ^2 -распределение

Рассмотрим следующую задачу. Значения случайной величины \mathcal{X} сосредоточены в симметричном интервале $(-\alpha, \alpha)$. Найти плотность вероятности величины $\mathcal{Y} = \mathcal{X}^2$, если плотность распределения вероятностей величины \mathcal{X} в интервале обозначена через $p_{\mathcal{X}}(x)$.

Функция $y(x) = x^2$ в интервале $(-\alpha, \alpha)$ не является монотонной, поэтому непосредственно применять выведенные ранее формулы нельзя. Как видно из рис. 7, событие $(y_1 < \mathcal{Y} < y_2)$ при $0 < y_1 < y_2 < \alpha^2$ равносильно сумме двух несовместных случайных событий $(x_1 < \mathcal{X} < x_2)$ и $(-x_2 < \mathcal{X} < -x_1)$, где $x_1 = \sqrt{y_1}$ и $x_2 = \sqrt{y_2}$. Поэтому для вероятности события $(y_1 < \mathcal{Y} < y_2)$ можно записать:

$$P(y_1 < \mathcal{Y} < y_2) = \int_{x_1}^{x_2} p_{\mathcal{X}}(x) dx + \int_{-x_2}^{-x_1} p_{\mathcal{X}}(x) dx.$$

В первом интеграле сделаем замену $x = \sqrt{y}$, а во втором — $x = -\sqrt{y}$, получим:

$$P(y_1 < \mathcal{Y} < y_2) = \int_{y_1}^{y_2} [p_{\mathcal{X}}(\sqrt{y}) + p_{\mathcal{X}}(-\sqrt{y})] \frac{dy}{2\sqrt{y}}.$$

Видно, что плотность распределения вероятностей величины \mathcal{Y} равна:

$$p_{\mathcal{Y}}(y) = [p_{\mathcal{X}}(\sqrt{y}) + p_{\mathcal{X}}(-\sqrt{y})] \frac{1}{2\sqrt{y}}, \quad \text{при } 0 < y < \alpha^2. \quad (9.1)$$

Вне интервала $(0, \alpha^2)$ величина \mathcal{Y} значений не имеет и $p_{\mathcal{Y}}(y) = 0$. В частности, когда плотность $p_{\mathcal{X}}(x)$ представляет собой четную функцию $p_{\mathcal{X}}(-x) = p_{\mathcal{X}}(x)$, формула (9.1) упрощается:

$$p_{\mathcal{Y}}(y) = \frac{p_{\mathcal{X}}(\sqrt{y})}{\sqrt{y}}, \quad \text{при } 0 < y < \alpha^2.$$

Так, если случайная величина \mathcal{X} имеет стандартное нормальное распределение с плотностью вероятностей $(-\infty < x < \infty)$:

$$p_{\mathcal{X}}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

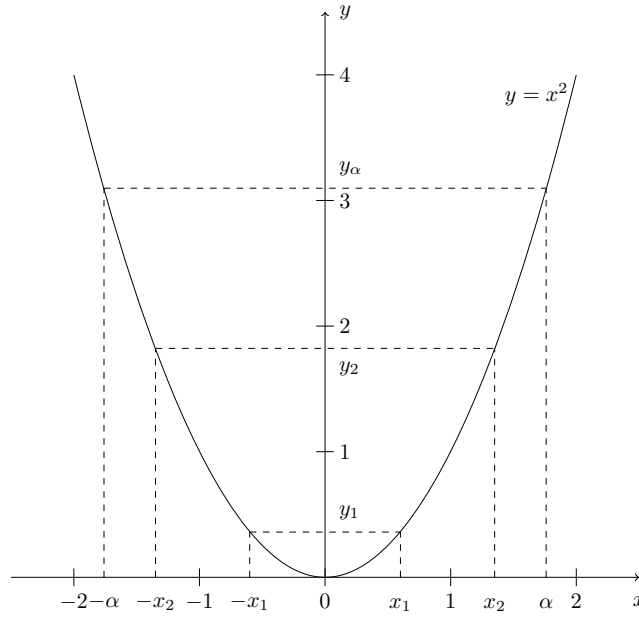


Рис. 7. Соответствие между величинами \mathcal{X} и \mathcal{Y}

то для величины $\mathcal{Y} = \mathcal{X}^2$ мы получим при $y > 0$:

$$p_{\mathcal{Y}}(y) = \frac{1}{\sqrt{2\pi y}} e^{-y/2}$$

и $p_{\mathcal{Y}}(y) = 0$ при $y \leq 0$, то есть величина \mathcal{Y} будет иметь Γ -распределение («гамма-распределение») с параметрами $\lambda = 1/2$ и $\alpha = 1/2$.

Распределением «хи-квадрат» (используется обозначение χ^2 -распределение) называется распределение суммы квадратов независимых случайных величин, следующих стандартному нормальному распределению.

Найдем плотность этого распределения. Обозначим исходные случайные величины через $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_k$, а сумму их квадратов через

$$\chi_k^2 = \mathcal{X}_1^2 + \mathcal{X}_2^2 + \dots + \mathcal{X}_k^2.$$

Число слагаемых k называется числом степеней свободы величины χ_k^2 ; сам закон распределения обозначается χ^2 ; индекс k здесь совпадает с количеством разрядов. Так как каждая величина \mathcal{X}_i ($i = 1, 2, \dots, k$) имеет стандартное нормальное распределение, то ее квадрат имеет Γ -распределение с параметрами $\lambda = 1/2$ и $\alpha = 1/2$. Но при композиции k Γ -распределений с параметрами $\lambda = 1/2$ и $\alpha = 1/2$ мы получим снова Γ -распределение с параметрами

$\lambda = 1/2$ и $\alpha = k/2$. Плотность этого распределения равна:

$$p_{\chi_k^2}(u) = \begin{cases} C_k u^{k/2-1} e^{-u/2}, & \text{при } u > 0, \\ 0, & \text{при } u \leq 0, \end{cases}$$

где $C_k = 2^{-k/2}/\Gamma(k/2)$. При $k = 4$ имеем для $u > 0$:

$$C_4 = \frac{1}{2^2\Gamma(2)} = \frac{1}{4}; \quad p_{\chi_4^2} = \frac{u}{4} e^{-u/4}.$$

Из самого определения χ^2 -распределения непосредственно следует, что сумма независимых величин, имеющих χ^2 -распределения, также будет иметь χ^2 -распределение, причем при сложении таких величин их числа степеней свободы складываются.

Покажем, что если независимые случайные величины \mathcal{X} и \mathcal{Y} имеют Γ -распределения $p(x) = C_\alpha x^{\alpha-1} e^{-\lambda x}$ при $x > 0$, где коэффициент $C_\alpha = \lambda^\alpha/\Gamma(\alpha)$, и $p(x) = 0$ при $x \leq 0$ с одинаковыми значениями параметра λ , то плотность распределения их суммы $\mathcal{Z} = \mathcal{X} + \mathcal{Y}$ также имеет Γ -распределение с тем же параметром λ и вторым параметром, равным $\alpha + \beta$.

Плотность распределения суммы $\mathcal{Z} = \mathcal{X} + \mathcal{Y}$ запишем в виде свёртки:

$$p_{\mathcal{Z}}(z) = p_{\mathcal{X}} * p_{\mathcal{Y}}(z) = \int_{-\infty}^{\infty} p_{\mathcal{X}}(x) p_{\mathcal{Y}}(z-x) dx.$$

Так как функция $p_{\mathcal{X}}(x)$ обращается в нуль при $x < 0$, а функция $p_{\mathcal{Y}}(z-x)$ — при $z-x < 0$, то есть при $x > z$, то последний интеграл отличен от 0 только при $0 < x < z$ и при этом может быть записан так:

$$p_{\mathcal{Z}}(z) = \int_0^z p_{\mathcal{X}}(x) p_{\mathcal{Y}}(z-x) dx = C_\alpha C_\beta \int_0^z e^{-\lambda x} x^{\alpha-1} e^{-\lambda(z-x)} (z-x)^{\beta-1} dx.$$

Этот интеграл преобразуем с помощью замены $x = zt$:

$$p_{\mathcal{Z}}(z) = C_\alpha C_\beta e^{-\lambda z} \int_0^1 t^{\alpha-1} z^{\alpha-1} (1-t)^{\beta-1} z^{\beta-1} z dt = C_\gamma z^{\alpha+\beta-1} e^{-\lambda z},$$

где $C_\gamma = C_\alpha C_\beta \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt = \lambda^{\alpha+\beta}/\Gamma(\alpha+\beta)$. Последнее выражение получается при использовании представления В-функции через Γ -функции:

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Таким образом, случайная величина \mathcal{Z} тоже имеет Γ -распределение с параметрами λ и $\alpha + \beta$, т.е. композиция двух Γ -распределений с одинаковыми значениями параметра λ дает снова Γ -распределение с тем же значением λ .

Закон распределения вероятностей называют *устойчивым*, если композиция таких законов есть тот же закон (отличающийся, вообще говоря, параметрами). Из проведенного анализа следует, что Γ -распределение устойчиво.

9.2. Распределение Стьюдента. Доверительный интервал

Нормальное распределение случайной величины \mathcal{X} с плотностью

$$p_{\mathcal{X}}(x; a, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-a)^2/(2\sigma^2)} \quad (9.2)$$

определяется двумя параметрами a и σ^2 , которые представляют собой центр распределения $a = M(\mathcal{X})$ и дисперсию $\sigma^2 = D(\mathcal{X})$ ($\sigma = \sqrt{D(\mathcal{X})}$ — среднее квадратическое отклонение). Оценками a и σ^2 служат среднее значение $\bar{\mathcal{X}}$ и выборочная дисперсия S^2 :

$$\bar{\mathcal{X}} = \frac{1}{n} \sum_{i=1}^n \mathcal{X}_i, \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (\mathcal{X}_i - \bar{\mathcal{X}})^2. \quad (9.3)$$

Эти величины удовлетворяют требованиям, предъявляемым к точечным оценкам параметров.

Определим эти требования на примере одного неизвестного параметра. Предположим, что функция распределения случайной величины, получаемой в результате выборки значений $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n$, зависит от параметра θ :

$$P(\mathcal{X}_i < x_i) = F_{\mathcal{X}_i}(x_i; \theta),$$

и все случайные величины \mathcal{X}_i независимы и одинаково распределены. Тогда оценкой θ_n^* называется функция $\theta_n^* = \theta_n^*(x_1, x_2, \dots, x_n)$, которая, таким образом, является случайной величиной. Естественно потребовать, чтобы при проведении большого числа опытов значение θ_n^* было близко к значению оцениваемого параметра θ . Для этого наложим на функцию $\theta_n^*(x_1, x_2, \dots, x_n)$ следующие условия:

1. *Несмещенность оценки*, т.е. при любом $n \rightarrow \infty$

$$M(\theta_n^*) = \theta. \quad (9.4)$$

2. *Состоятельность оценки*, т.е. при любом $\varepsilon > 0$ и $n \rightarrow \infty$

$$P(|\theta_n^* - \theta| < \varepsilon) \rightarrow 1. \quad (9.5)$$

Это условие можно переформулировать в терминах дисперсии оценки θ_n^* , а именно, $D(\theta_n^*) \rightarrow 0$ при $n \rightarrow \infty$.

3. Эффективность оценки

$$e(\theta_n^*) = [n I(\theta) D(\theta_n^*)]^{-1}, \quad (9.6)$$

где в предположении положительной плотности распределения $p_{\mathcal{X}}(x; \theta)$ каждого измерения вводится математическое ожидание

$$I(\theta) = M \left[\frac{\partial \ln p_{\mathcal{X}}(x; \theta)}{\partial \theta} \right]^2, \quad (9.7)$$

которое называется *фишеровским количеством информации*, содержащимся в одном наблюдении. Будем считать, что выполняются *условия регулярности*: область интегрирования не зависит от θ , интегралы по этой области можно дифференцировать под знаком интеграла и существует выражение $I(\theta)$, причем оно положительно. Для эффективности оценки получим

$$0 \leq e(\theta_n^*) \leq 1. \quad (9.8)$$

Оценка θ_n^* считается *эффективной*, если $e(\theta_n^*) = 1$.

Границы эффективности (9.8) определяются *неравенством Рао—Крамера*, которое при условиях регулярности позволяет получить нижнюю границу для дисперсии оценки θ_n^* :

$$D(\theta_n^*) \geq [n I(\theta)]^{-1}. \quad (9.9)$$

Это неравенство справедливо не только для непрерывных, но для дискретных случайных величин.

Иногда вместо эффективности оценки (9.6) удобнее использовать *асимптотическую эффективность*:

$$e_0(\theta_n^*) = [I(\theta) S^2]^{-1}, \quad (9.10)$$

где S^2 — оценка дисперсии (9.3), причем оценка θ_n^* считается *асимптотически эффективной*, если $e_0(\theta_n^*) = 1$.

Пример 9.1. Вычислить математическое ожидание, дисперсию и эффективность среднего значения $\bar{\mathcal{X}}$, определенного первым уравнением в (9.3).

Решение. В соответствии с математической моделью все случайные величины независимы в совокупности и одинаково распределены по нормальному

закону с математическим ожиданием $M(\mathcal{X}_k) = a$ и дисперсией $D(\mathcal{X}_k) = \sigma^2$. Начнем с вычисления математического ожидания и дисперсии:

$$M(\bar{\mathcal{X}}) = \frac{1}{n} \sum_{k=1}^n M(\mathcal{X}_k) = a, \quad D(\bar{\mathcal{X}}) = \frac{1}{n^2} \sum_{k=1}^n D(\mathcal{X}_k) = \frac{\sigma^2}{n}.$$

Полученные значения характеристик показывают, что среднее значение $\bar{\mathcal{X}}$ — это несмещенная и состоятельная оценка параметра a . Найдем теперь эффективность $\bar{\mathcal{X}}$, вычислив первоначально информацию Фишера (9.7):

$$I(a) = M \left[\frac{\partial \ln p_{\mathcal{X}}(x; a, \sigma^2)}{\partial a} \right]^2 = M \left[\frac{(x - a)^2}{\sigma^4} \right] = \frac{1}{\sigma^2} D \left(\frac{x - a}{\sigma} \right) = \frac{1}{\sigma^2},$$

где $p_{\mathcal{X}}(x; a, \sigma^2)$ — плотность распределения (9.2). Подставив $D(\bar{\mathcal{X}})$ и $I(a)$ в (9.6), получим эффективность $e(\bar{\mathcal{X}}) = 1$, т. е. среднее значение $\bar{\mathcal{X}}$ является эффективной оценкой параметра a .

Пример 9.2. Вычислить математическое ожидание, дисперсию и эффективность выборочной дисперсии S^2 , определенной вторым уравнением в (9.3).

Решение. Более удобно перейти к нормированным случайным величинам $\mathcal{Y}_k = (\mathcal{X}_k - a) / \sigma$, которые так же, как и \mathcal{X}_k , будут независимы в совокупности и распределены нормально $\mathcal{Y}_k \in N(0, 1)$. Согласно результатам предыдущего примера их среднее значение $\bar{\mathcal{Y}} = \sum_{k=1}^n \mathcal{Y}_k / n$ имеет $M(\bar{\mathcal{Y}}) = 0$ и $D(\bar{\mathcal{Y}}) = 1/n$. Выборочная дисперсия S^2 (9.3) в терминах \mathcal{Y}_k примет вид:

$$S^2 = \frac{\sigma^2}{n-1} \sum_{k=1}^n (\mathcal{Y}_k - \bar{\mathcal{Y}})^2 = \frac{\sigma^2}{n-1} \left[\sum_{k=1}^n \mathcal{Y}_k^2 - n\bar{\mathcal{Y}}^2 \right].$$

Математическое ожидание выборочной дисперсии вычисляется просто:

$$M(S^2) = \frac{\sigma^2}{n-1} \left[\sum_{k=1}^n D(\mathcal{Y}_k) - nD(\bar{\mathcal{Y}}) \right] = \sigma^2.$$

Видно, что S^2 — несмещенная оценка дисперсии, что было достигнуто правильным выбором коэффициента $1/(n-1)$ перед суммой квадратов. Перейдем к вычислению дисперсии S^2 :

$$D(S^2) = M(S^2 - \sigma^2)^2 = \frac{\sigma^4}{(n-1)^2} \left[\sum_{k,m=1}^n M(\mathcal{Y}_k^2 \mathcal{Y}_m^2) - 2n \sum_{k=1}^n M(\mathcal{Y}_k^2 \bar{\mathcal{Y}}^2) - \right. \\ \left. - 2(n-1) \sum_{k=1}^n M(\mathcal{Y}_k^2) + n^2 M(\bar{\mathcal{Y}}^4) + 2n(n-1) M(\bar{\mathcal{Y}}^2) + (n-1)^2 \right].$$

Если учесть, что математическое ожидание от произведения нечетного числа случайных величин обращается в ноль, то дисперсию можно записать в виде:

$$D(S^2) = \sigma^4 \left[1 - \frac{2}{n} \sum_{k=1}^n M(\mathcal{Y}_k^2) + \frac{1}{n^2} \sum_{k=1}^n M(\mathcal{Y}_k^4) + \frac{n^2 - 2n + 3}{n^2 (n-1)^2} \sum_{k \neq m=1}^n M(\mathcal{Y}_k^2) M(\mathcal{Y}_m^2) \right].$$

У нормально распределенной случайной величины \mathcal{Y}_k математические ожидания $M(\mathcal{Y}_k^2) = 1$ и $M(\mathcal{Y}_k^4) = 3$, поэтому

$$D(S^2) = \sigma^4 \left[1 - 2 + \frac{3}{n} + \frac{n^2 - 2n + 3}{n(n-1)} \right] = \frac{2\sigma^4}{n-1}.$$

$D(S^2)$ обращается в ноль при $n \rightarrow \infty$, т.е. оценка дисперсии состоятельна. Для нахождения эффективности оценки надо найти информацию Фишера:

$$I(\sigma^2) = M \left[\frac{\partial \ln p_{\mathcal{X}}(x; a, \sigma^2)}{\partial \sigma^2} \right]^2 = M \left[-\frac{1}{2\sigma^2} + \frac{(x-a)^2}{2\sigma^4} \right]^2 = \frac{1}{2\sigma^4}.$$

После подстановки $D(S^2)$ и $I(\sigma^2)$ в эффективность (9.6) получим

$$e(S^2) = \frac{n-1}{n} < 1.$$

Рассмотренная выборочная дисперсия S^2 асимптотически эффективна.

Величина $\bar{\mathcal{X}}$ имеет нормальное распределение с параметрами a и σ^2/n . Величина

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n \frac{(\mathcal{X}_i - \bar{\mathcal{X}})^2}{\sigma^2} = \mathcal{Y}_1^2 + \mathcal{Y}_2^2 + \dots + \mathcal{Y}_{n-1}^2,$$

где \mathcal{Y}_k ($k = 1, 2, \dots, n-1$) независимы и распределены нормально с параметрами $M(\mathcal{Y}_k) = 0$ и $D(\mathcal{Y}_k) = 1$. Таким образом, величина $(n-1)S^2/\sigma^2$ имеет χ^2 -распределение с $k = n-1$ степенями свободы.

Последнее утверждение легко продемонстрировать при $n = 2$:

$$\sum_{i=1}^2 \frac{(\mathcal{X}_i - \bar{\mathcal{X}})^2}{\sigma^2} = \left(\frac{\mathcal{X}_1 - \mathcal{X}_2}{\sigma\sqrt{2}} \right)^2 = \mathcal{Y}_1^2.$$

где $\bar{\mathcal{X}} = (\mathcal{X}_1 + \mathcal{X}_2)/2$. Разность $\mathcal{Y}_1 = (\mathcal{X}_1 - \mathcal{X}_2)/(\sqrt{2}\sigma)$ имеет нормальное распределение с параметрами

$$M(\mathcal{Y}_1) = \frac{a-a}{\sqrt{2}\sigma} = 0, \quad D(\mathcal{Y}_1) = \frac{\sigma^2 + \sigma^2}{(\sqrt{2}\sigma)^2} = 1,$$

а ее квадрат — χ^2 -распределение с одной степенью свободы.

Получим доверительную оценку параметра σ^2 . Зададим надежность γ (или уровень значимости $\alpha = 1 - \gamma$) и по таблицам χ^2 -распределения с числом степеней свободы k найдем такие значения u_1 и u_2 , что

$$\int_0^{u_1} p_{\chi_k^2}(u) du = \int_{u_2}^{\infty} p_{\chi_k^2}(u) du = \frac{\alpha}{2},$$

тогда вероятность события $(u_1 < (n-1)S^2/\sigma^2 < u_2)$ будет равна

$$P\left(u_1 < \frac{(n-1)S^2}{\sigma^2} < u_2\right) = \int_{u_1}^{u_2} p_{\chi_k^2}(u) du = 1 - \alpha = \gamma.$$

Разрешив двойное неравенство в вероятности относительно σ^2 , получим:

$$z_1 S^2 < \sigma^2 < z_2 S^2,$$

где $z_2 = (n-1)/u_1$ и $z_1 = (n-1)/u_2$ — верхняя и нижняя границы отношения σ^2/S^2 . Это дает доверительный интервал для параметра σ^2 с надежностью γ .

Доверительную оценку параметра a записывают обычно в форме

$$|\bar{\mathcal{X}} - a| < t \sqrt{\frac{S^2}{n}}. \quad (9.11)$$

Для расчета вероятности события (9.11) надо найти распределение вероятностей отношения

$$\mathcal{T}_n = \frac{\bar{\mathcal{X}} - a}{\sqrt{S^2/n}}. \quad (9.12)$$

Преобразуем это отношение. Величина S^2 может быть представлена в виде

$$S^2 = \frac{\sigma^2}{n-1} (\mathcal{Y}_1^2 + \mathcal{Y}_2^2 + \dots + \mathcal{Y}_{n-1}^2),$$

где все величины $\mathcal{Y}_1, \mathcal{Y}_2, \dots, \mathcal{Y}_{n-1}$ независимы и имеют стандартное нормальное распределение. Поэтому отношение (9.12) можно записать в виде:

$$\mathcal{T}_n = \frac{\sqrt{n-1} \mathcal{Y}_0}{\sqrt{\mathcal{Y}_1^2 + \mathcal{Y}_2^2 + \dots + \mathcal{Y}_{n-1}^2}},$$

где $\mathcal{Y}_0 = (\bar{\mathcal{X}} - a)/\sqrt{\sigma^2/n}$. Отсюда видно, что распределение отношения (9.12) не зависит ни от a , ни от σ^2 . Плотность этого распределения называется

распределением Стьюдента и имеет вид:

$$S(t, n) \equiv S_{n-1}(t) = \frac{\Gamma(n/2)}{\sqrt{\pi(n-1)}\Gamma((n-1)/2)} \left(1 + \frac{t^2}{n-1}\right)^{-n/2}.$$

Так как $S(t, n)$ — четная функция от t , то вероятность осуществления неравенства $|\mathcal{T}_n| < t_\gamma$ определяется интегралом

$$P(|\mathcal{T}_n| < t_\gamma) = 2 \int_0^{t_\gamma} S(t, n) dt = \gamma.$$

Воспользовавшись определением \mathcal{T}_n (9.12) и заменив неравенство в круглых скобках равносильным ему двойным неравенством, получим

$$P\left(\bar{X} - t_\gamma \sqrt{S^2/n} < a < \bar{X} + t_\gamma \sqrt{S^2/n}\right) = \gamma.$$

Итак, пользуясь распределением Стьюдента, мы нашли доверительный интервал, покрывающий неизвестный параметр a с надежностью γ .

Пример 9.3. Произведено 5 независимых измерений случайной величины $\mathcal{X} \in N(a, 20)$. Результаты наблюдений таковы: $x_1 = -25$, $x_2 = 34$, $x_3 = -20$, $x_4 = 10$ и $x_5 = 21$. Найти оценку для $a = M(\mathcal{X})$, а также построить для него 95 %-й доверительный интервал.

Решение. Находим сначала \bar{x} : $\bar{x} = (-25 + 34 - 20 + 10 + 21)/5 = 4$. Учитывая, что $\gamma = 0,95$ и $\Phi_0(t_\gamma) = \gamma/2 = 0,475$, получаем $t_\gamma = 1,96$. Тогда, используя

$$\varepsilon = \frac{t_\gamma \sigma}{\sqrt{n}} = 1,96 \times 20/\sqrt{5} \approx 17,5,$$

получаем доверительный интервал для $a = M(\mathcal{X})$:

$$\left(\bar{x} - t_\gamma \frac{\sigma}{\sqrt{n}}, \bar{x} + t_\gamma \frac{\sigma}{\sqrt{n}}\right) = (4 - 17,5; 4 + 17,5) = (-13,5; 21,5).$$

Пример 9.4. По условию предыдущего примера, считая, что случайная величина $\mathcal{X} \in N(a, \sigma^2)$, построить для неизвестного $M(\mathcal{X}) = a$ доверительный интервал. Считать $\gamma = 0,95$.

Решение. Оценка \bar{x} для $M(\mathcal{X})$ знаем: $\bar{x} = 4$. Далее находим значения:

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

и $S \approx 25,7$. Из таблицы для $\gamma = 0,95$ и $n - 1 = 4$ степеней свободы следует $t_\gamma = 2,78$. Учитывая $\varepsilon = 2,78 \times 25,7/2,24 \approx 31,9$, получаем доверительный интервал: $(-27,9; 35,9)$.

Контрольные вопросы

1. Как находится плотность распределения вероятностей квадрата случайной величины, если известна ее плотность распределения вероятностей?
2. Как находится плотность распределения вероятностей квадрата случайной величины, если известна ее плотность распределения вероятностей и она является четной функцией?
3. Дать определение «хи-квадрат» распределения.
4. Что такое «число степеней свободы» распределения «хи-квадрат»?
5. Вывести формулу, определяющую функцию плотности распределения суммы двух случайных величин, имеющих гамма-распределение с одинаковым параметром λ .
6. Рассказать о доверительной оценке параметра дисперсии для случайной величины, распределенной согласно «хи-квадрат» распределению.
7. Рассказать о доверительной оценке математического ожидания случайной величины, распределенной согласно «хи-квадрат» распределению.
8. Что такое «распределение Стьюдента»?
9. Как строится доверительный интервал параметра на основании распределения Стьюдента?

10. Сглаживание экспериментальных кривых.

Критерии согласия

10.1. Сглаживание экспериментальных зависимостей по методу наименьших квадратов

Решение задачи выравнивания или сглаживания зависит от того, что мы условимся считать «наилучшим» сглаживанием. Общепринятым считается так называемый метод наименьших квадратов, при котором требование наилучшего согласования кривой $y = \varphi(x)$ и экспериментальных точек сводится к тому, чтобы сумма квадратов отклонений экспериментальных точек от сглаживающей кривой была минимальной.

Предположим, что истинная зависимость y от x в точности выражается формулой $y = \varphi(x)$, а экспериментальные точки уклоняются от этой зависимости вследствие ошибок измерений, которые, как принято считать, подчиняются нормальному закону распределения.

Рассмотрим какое-нибудь значение аргумента x_i . Результат опыта есть случайная величина Y_i , распределенная по нормальному закону с математическим ожиданием $\varphi(x_i)$ и со средним квадратическим отклонением σ_i , характеризующим ошибку измерения. Предположим, что точность измерения во всех точках одинакова: $\sigma_1 = \sigma_2 = \dots = \sigma_n = \sigma$. Нормальный закон, по которому распределяется величина Y_i , запишется в виде:

$$f_i(y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-[y_i - \varphi(x_i)]^2 / (2\sigma^2)}.$$

В результате опыта произошло событие: случайные величины (Y_1, Y_2, \dots, Y_n) приняли конкретный набор значений (y_1, y_2, \dots, y_n) .

Поставим задачу: так подобрать математические ожидания $\varphi(x_1), \varphi(x_2), \dots, \varphi(x_n)$, чтобы вероятность этого события была максимальной.

Строго говоря, необходимо пользоваться соответствующими элементами вероятностей:

$$f_i(y_i) dy_i = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-[y_i - \varphi(x_i)]^2 / (2\sigma^2)} dy_i.$$

Найдем вероятность того, что система случайных величин (Y_1, Y_2, \dots, Y_n) примет значения, лежащие в интервале $(y_i, y_i + dy_i)$, где $i = 1, 2, \dots, n$. Так как опыты независимы, то эта вероятность равна произведению элементов

вероятностей для всех значений i :

$$\prod_{i=1}^n f_i(y_i) dy_i = K \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\sum_{i=1}^n [y_i - \varphi(x_i)]^2 / (2\sigma^2)}, \quad (10.1)$$

где K — коэффициент.

Экспонента в вероятности (10.1) всегда меньше единицы и имеет наибольшее значение, когда ее показатель минимален по абсолютной величине:

$$\frac{1}{2\sigma^2} \sum_{i=1}^n [y_i - \varphi(x_i)]^2 = \min.$$

Отсюда получим требование метода наименьших квадратов: для того чтобы данная совокупность значений (y_1, y_2, \dots, y_n) была наивероятнейшей, нужно выбрать функцию $\varphi(x)$ так, чтобы сумма квадратов отклонений наблюдаемых значений y_i от $\varphi(x_i)$ была минимальной:

$$\sum_{i=1}^n [y_i - \varphi(x_i)]^2 = \min. \quad (10.2)$$

Перейдем теперь к задаче определения значений параметров (a, b, c, \dots) , исходя из принципа наименьших квадратов (10.2). Пусть из каких-либо соображений выбран вид функции $y = \varphi(x)$, зависящей от нескольких числовых параметров: a, b, c, \dots , которые считаются свободными и подлежат определению методом наименьших квадратов. Из соображений удобства дальнейшего изложения запишем y как функцию не только x , но и параметров a, b, c, \dots , т. е. $y = \varphi(x; a, b, c, \dots)$. Для выполнения условия минимальности (10.2) необходимо подобрать a, b, c, \dots так, чтобы функция в левой части (10.2) была экстремальна по этим параметрам. Продифференцируем сумму в условии (10.2) по a, b, c, \dots и приравняем эти частные производные нулю:

$$\sum_{i=1}^n [y_i - \varphi(x; a, b, c, \dots)] \frac{\partial \varphi}{\partial a} \bigg|_{x=x_i} = 0,$$

$$\sum_{i=1}^n [y_i - \varphi(x; a, b, c, \dots)] \frac{\partial \varphi}{\partial b} \bigg|_{x=x_i} = 0$$

и т. д. по всем остальным фиксируемым параметрам. Решение этой системы уравнений позволяет найти значения параметров, обеспечивающих экстремум суммы (10.2).

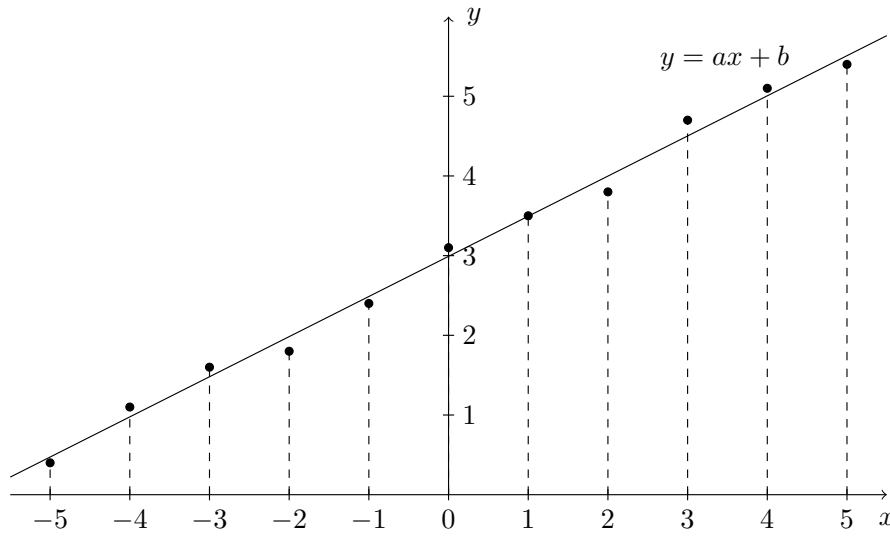


Рис. 8. Пример подгонки данных под линейную зависимость

Рассмотрим частный случай, когда функция $\varphi(x)$ — линейная функция, т. е. $y = \varphi(x; a, b) = ax + b$. Пример подгонки конкретных данных под линейную зависимость рассмотрен в конце раздела, а эмпирические данные и подогнанная под них прямая линия изображены на рис. 8. Рассмотренная выше процедура минимизации приводит к системе двух уравнений:

$$\sum_{i=1}^n [y_i - (ax_i + b)] x_i = 0, \quad \sum_{i=1}^n [y_i - (ax_i + b)] = 0,$$

где учли значения частных производных: $\partial\varphi/\partial a|_{x=x_i} = x_i$ и $\partial\varphi/\partial b|_{x=x_i} = 1$. Раскрыв скобки, просуммировав и поделив оба уравнения на n , получим:

$$\frac{1}{n} \sum_{i=1}^n x_i y_i - \frac{a}{n} \sum_{i=1}^n x_i^2 - \frac{b}{n} \sum_{i=1}^n x_i = 0, \quad \frac{1}{n} \sum_{i=1}^n y_i - \frac{a}{n} \sum_{i=1}^n x_i - b = 0.$$

Суммы представляют собой статистические моменты:

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n x_i &= m_x^*, & \frac{1}{n} \sum_{i=1}^n x_i^2 &= \alpha_2^*[X], \\ \frac{1}{n} \sum_{i=1}^n y_i &= m_y^*, & \frac{1}{n} \sum_{i=1}^n x_i y_i &= \alpha_{1,1}^*[X, Y], \end{aligned} \quad (10.3)$$

в терминах которых система уравнений приводится к виду:

$$\alpha_{1,1}^*[X, Y] - a \alpha_2^*[X] - b m_x^* = 0, \quad m_y^* - a m_x^* - b = 0.$$

Выразим b из второго уравнения $b = m_y^* - a m_x^*$ и подставим в первое:

$$\alpha_{1,1}^*[X, Y] - a \alpha_2^*[X] - m_x^* (m_y^* - a m_x^*) = 0.$$

Разрешая последнее уравнение относительно параметра a , имеем:

$$a = \frac{\alpha_{1,1}^*[X, Y] - m_x^* m_y^*}{\alpha_2^*[X] - (m_x^*)^2}.$$

Это выражение упрощается после подстановки центральных моментов:

$$\alpha_{1,1}^*[X, Y] - m_x^* m_y^* = K_{xy}^*, \quad \alpha_2^*[X] - (m_x^*)^2 = D_x^*, \quad (10.4)$$

откуда следуют окончательные выражения для a и b :

$$a = \frac{K_{xy}^*}{D_x^*}, \quad b = m_y^* - \frac{K_{xy}^*}{D_x^*} m_x^*, \quad (10.5)$$

где

$$K_{xy}^* = \frac{1}{n} \sum_{i=1}^n (x_i - m_x^*) (y_i - m_y^*), \quad D_x^* = \frac{1}{n} \sum_{i=1}^n (x_i - m_x^*)^2.$$

Таким образом, задача решена

$$y = \frac{K_{xy}^*}{D_x^*} x + m_y^* - \frac{K_{xy}^*}{D_x^*} m_x^*,$$

или в ином виде

$$y - m_y^* = \frac{K_{xy}^*}{D_x^*} (x - m_x^*).$$

Пример 10.1. Аппроксимировать данные, приведенные в таблице, линейной зависимостью $y(x) = ax + b$.

x_i	-5	-4	-3	-2	-1	0	1	2	3	4	5
y_i	0,4	1,1	1,6	1,8	2,4	3,1	3,5	3,8	4,7	5,1	5,4

Решение. Непосредственное вычисление статистических моментов (10.3) дает следующие значения:

$$m_x^* = 0, \quad m_y^* = 2,991, \quad \alpha_2^*[X] = 10, \quad \alpha_{11}^*[X, Y] = 5,036,$$

или в терминах величин (10.4):

$$K_{xy}^* = 5,036, \quad D_x^* = 10.$$

После подстановки в (10.5) получим оценки параметров линейного уравнения:

$$a = 0,5036, \quad b = 2,991.$$

Эмпирические точки и аппроксимирующая прямая приведены на рис. 8.

10.2. Критерии согласия

В этом разделе рассмотрим вопрос, связанный с проверкой правдоподобия гипотез, а именно — вопрос о согласованности теоретического и статистического распределений.

Допустим, что распределение выравнено с помощью теоретической кривой $f(x)$. Как бы хорошо ни была подобрана теоретическая кривая, между нею и статистическим распределением неизбежны расхождения. Вопрос — насколько они существенны, то есть являются ли случайными или следствием плохой подгонки? Для ответа на этот вопрос служат критерии согласия.

Идея состоит в следующем. На основании данного статистического материала надо проверить гипотезу H , состоящую в том, что случайная величина X подчиняется некоторому определенному закону распределения. Этот закон может быть задан в той или иной форме, например в виде функции распределения $F(x)$, или плотности распределения $f(x)$, или же в виде совокупности вероятностей p_i , где p_i — вероятность того, что случайная величина X попадает в пределы i -го разряда.

Так как функция распределения $F(x)$ является наиболее общей и определяет собой любую другую, будем формулировать гипотезу H как состоящую в том, что случайная величина X имеет функцию распределения $F(x)$.

Для того чтобы принять гипотезу или ее отвергнуть, рассмотрим некоторую величину U , называемую мерой и характеризующую степень расхождения теоретического и статистического распределений.

Величина U может быть выбрана различными способами: например, в качестве U можно взять сумму квадратов отклонений теоретических вероятностей p_i от соответствующих частот p_i^* , или же сумму тех же квадратов с некоторыми коэффициентами, называемыми «весами», или же максимальное отклонение статистической функции распределения $F^*(x)$ от теоретической $F(x)$ и т. д. Очевидно, что это есть некоторая случайная величина. Закон распределения этой случайной величины зависит от закона распределения величины X , над которой производятся опыты, и от числа опытов n . Если гипотеза H верна, то закон распределения U определяется законом распределения величины X , т. е. функцией $F(x)$, и числом n .

Пусть закон распределения нам известен. В результате серии опытов обнаружено, что выбранная нами мера расхождения приняла некоторое значение u . Спрашивается, можно ли объяснить это случайными причинами или же это расхождение слишком велико?

Для ответа на этот вопрос предположим, что гипотеза H верна, и вычислим в этом предположении вероятность того, что за счет случайных причин, связанных с недостаточным объемом опытного материала, мера расхождения U окажется не меньше, чем наблюдаемое нами на опыте значение u , то есть вычислим вероятность события $(U \geq u)$.

Если эта вероятность весьма мала, то гипотезу H следует отвергнуть как маловероятную, если же эта вероятность значительна, следует признать, что экспериментальные данные не противоречат гипотезе H .

Возникает вопрос о том, каким способом следует выбирать меру расхождения U ? Оказывается, что при некоторых способах ее выбора закон распределения величины U обладает весьма простыми свойствами и при достаточно большом n практически не зависит от функции $F(x)$. Именно такими мерами расхождения и пользуются в математической статистике в качестве критериев согласия.

Рассмотрим один из наиболее часто применяемых критериев согласия — так называемый «критерий χ^2 » («критерий хи-квадрат» Пирсона).

Пусть произведено n независимых опытов, в каждом из которых случайная величина X приняла определенное значение. Результаты опытов сведены в таблицу в k разрядов.

I_i	(x_1, x_2)	(x_2, x_3)	\dots	(x_k, x_{k+1})
p_i^*	p_1^*	p_2^*	\dots	p_k^*

Требуется проверить, согласуются ли экспериментальные данные с гипотезой о том, что случайная величина X имеет данный закон распределения (заданную $F(x)$ или $f(x)$); назовем этот закон распределения «теоретическим».

Зная теоретический закон распределения, можно найти теоретические вероятности попадания случайной величины в каждый из разрядов: p_1, p_2, \dots, p_k . Проверяя согласованность теоретического и статистического распределений, мы будем исходить из расхождений между теоретическими вероятностями p_i и наблюдаемыми частотами p_i^* . Естественно выбрать в качестве меры расхождения между теоретическим и статистическим распределением сумму квадратов отклонений $(p_i^* - p_i)$, взятых с некоторыми «весами» c_i :

$$U = \sum_{i=1}^k c_i (p_i^* - p_i)^2. \quad (10.6)$$

Веса вводятся из-за неравнозначности различных разрядов. Действительно, одно и то же по абсолютной величине отклонение $(p_i^* - p_i)$ может быть мало-значительным, если сама вероятность p_i велика, и очень значительным, если

она мала. Поэтому естественно «веса» c_i брать обратно пропорциональными вероятностям разрядов p_i .

К. Пирсон показал, что если положить:

$$c_i = \frac{n}{p_i}, \quad (10.7)$$

то при больших n закон распределения величины U обладает весьма простыми свойствами: он практически не зависит от функции распределения $F(x)$ и от числа разрядов k , а именно: при увеличении n этот закон приближается к так называемому « χ^2 -распределению».

При выборе коэффициентов c_i согласно (10.7) мера расхождения обычно обозначается χ^2 :

$$U \equiv \chi^2 = n \sum_{i=1}^k \frac{(p_i^* - p_i)^2}{p_i}. \quad (10.8)$$

Для удобства вычислений, чтобы не иметь дела с дробными величинами с большим числом нулей, можно внести n под знак суммы и, учитывая, что $p_i^* = m_i/n$, где m_i — число значений в i -м разряде, приведем формулу (10.8) к виду:

$$U \equiv \chi^2 = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i}. \quad (10.9)$$

χ^2 -распределение зависит от параметра r (или k), называемого числом «степеней свободы» распределения.

Число «степеней свободы» r равно числу разрядов k минус число независимых условий — «связей», наложенных на частоты p_i^* . Примерами таких условий могут быть:

$$\sum_{i=1}^k p_i^* = 1, \quad (10.10)$$

если мы требуем только того, чтобы сумма частот была равна единице (это требование накладывается во всех случаях);

$$\sum_{i=1}^k \tilde{x}_i p_i^* = m_x, \quad (10.11)$$

если мы подбираем теоретическое распределение с тем условием, чтобы совпадали теоретическое и статистическое средние значения;

$$\sum_{i=1}^k (\tilde{x}_i - m_x^*)^2 p_i^* = D_x, \quad (10.12)$$

если мы требуем, кроме того, совпадения теоретической и статистической дисперсий и т. д.

Для χ^2 -распределения составлены таблицы. Пользуясь этими таблицами, можно для каждого значения χ^2 и числа степеней свободы r найти вероятности P того, что величина, распределенная по закону χ^2 , превзойдет это значение.

Входами являются: значение вероятности P и число степеней свободы r . Числа, стоящие в таблице, представляют собой соответствующие значения χ^2 .

Контрольные вопросы

1. Раскрыть основной смысл процедуры сглаживания экспериментальных зависимостей.
2. Какие величины в процедуре сглаживания экспериментальных зависимостей подбираются из условия максимальной, а какие — из условия минимальности?
3. Как реализуется процедура минимизации?
4. Продемонстрировать процедуру сглаживания экспериментальных зависимостей для случая линейной зависимости.
5. Раскрыть основной смысл исследования правдоподобия гипотез на основании «критериев согласия».
6. Каким способом выбирается мера?
7. Рассказать о критерии «хи-квадрат» Пирсона.
8. Какова плотность распределения «хи-квадрат»?
9. Каков выбор коэффициентов в критерии «хи-квадрат»?
10. Каким образом устанавливается правдоподобие на основании критерия «хи-квадрат»?

Задачи

1. Какова вероятность угадывания шифра из 4 цифр автоматической камеры хранения?
2. Вынимаются наугад 2 кости домино. Какова вероятность того, что их можно приставить друг к другу?
3. Четверо господ сдают шляпы в гардероб, затем берут их наугад. Какова вероятность того, что k человек получают свои шляпы ($0 \leq k \leq 4$)?
4. В семье трое детей. Будем считать, что вероятности рождения мальчика и девочки одинаковы. Какова вероятность того, что в семье все трое — мальчики?
5. Бросаются 2 игральные кости. Какова вероятность того, что сумма выпавших очков равняется 7, больше 7, меньше 7, что на первой кости выпало больше очков, чем на второй?
6. Пять человек входят в лифт девятиэтажного дома. Какова вероятность того, что все они выйдут на 6-м этаже, что все они выйдут на одном каком-то этаже, что все они выйдут на разных этажах?
7. В урне находится 10 шаров белого цвета и 4 шара красного цвета. Вынимаются один за другим 2 шара без возвращения. Какова вероятность того, что оба они красные? Рассмотреть эту задачу по схеме случайного выбора «с возвращением».
8. Из колоды в 36 карт одна за другой вынимаются (по схеме без возвращения) 6 карт. Какова вероятность того, что среди вынутых карт будет 4 туза, менее двух тузов, по крайней мере 2 туза?
9. В аудитории находятся k человек. Какова вероятность того, что по крайней мере у двух из них дни рождения совпадут?
10. Дважды бросается пара игровых костей. Какова вероятность того, что число очков, выпавших на каждой кости при первом бросании, не совпадет с числом очков, выпавших на этих же костях при втором бросании?
11. Задумано пятизначное число без повторяющихся цифр. Какова вероятность его угадать, если известно, что оно четное или делится на 5?

12. Куб, все грани которого окрашены, распилен на тысячу кубиков одинакового размера, которые затем тщательно перемешаны. Найти вероятность того, что наудачу извлеченный кубик будет иметь окрашенных граней: а) одну; б) две; в) три.
13. N человек ($N > 2$) случайным образом рассаживаются за круглым столом. Найти вероятность того, что два фиксированных лица A и B окажутся рядом.
14. Имеются две урны. В первой урне a белых и b черных шаров, во второй — c белых и d черных. Из каждой урны вынимается по шару. Найти вероятность того, что оба шара будут белыми.
15. Из полной колоды карт (52 листа, 4 масти) вынимается сразу несколько карт. Сколько карт нужно вынуть для того, чтобы с вероятностью, большей, чем 0,5, утверждать, что среди них по крайней мере две карты будут одной и той же масти?
16. По схеме случайного выбора «с возвращением» из множества целых чисел $\Omega = \{1, 2, \dots, N\}$ выбираются два числа ξ_1 и ξ_2 . Найти вероятность $P(\xi_1^2 + \xi_2^2 \leq N^2)$ в пределе, когда $N \rightarrow \infty$.
17. Случайная точка A равномерно распределена в правильном треугольнике с вершинами $(-a, 0)$, $(a, 0)$ и $(0, a\sqrt{3})$. Найти вероятность того, что квадрат с центром в точке A и сторонами длины x , параллельными осям координат, целиком содержится в этом треугольнике.
18. Случайная точка $A = (\xi_1, \xi_2)$ равномерно распределена в единичном квадрате с вершинами $(0, 0)$, $(1, 0)$, $(0, 1)$ и $(1, 1)$. Пусть η — число действительных корней трехчлена $f_{\xi_1 \xi_2}(x) = x^3/3 - \xi_1^2 x + \xi_2$. Найти вероятности событий $P(\eta = 1)$ и $P(\eta = 3)$.
19. Однородный прямой круговой цилиндр с высотой H и радиусом основания R бросают на горизонтальную плоскость. Найти вероятность того, что цилиндр упадет на боковую поверхность. При каком соотношении между значениями H и R вероятности упасть на боковую поверхность и на основания будут одинаковы.
20. Батарея из M орудий ведет огонь по группе, состоящей из L целей, где $M \leq L$. Наводчики орудий выбирают себе цели последовательно, случайным образом, при условии, что никакие два орудия стрелять по одной

и той же цели не могут. Найти вероятность того, что будут обстреляны цели с номерами $1, 2, \dots, M$.

21. Батарея, состоящая из M орудий, ведет огонь по группе из L самолетов ($M \leq L$). Каждый наводчик орудия выбирает себе цель случайно и независимо от других. Найти вероятность того, что все M орудий будут стрелять по одной и той же цели.
22. Имеется m радиолокационных станций, следящих за космическими объектами, каждая из которых за один цикл обзора обнаруживает объект с вероятностью p (независимо от других циклов и от других станций). За время T каждая станция успевает сделать n циклов. Найти вероятность следующих событий: A — объект будет обнаружен хотя бы одной станцией; B — объект будет обнаружен каждой из станций.
23. Для повышения надежности работы прибора он дублируется другим точно таким же прибором, надежность (вероятность безотказной работы) каждого прибора равна p . При выходе из строя первого прибора происходит мгновенное переключение на второй, причем надежность переключающего устройства равна единице. Определить надежность всей системы.
24. Техническое устройство, состоящее из K узлов, работало в течение некоторого времени t . За это время первый узел мог выйти из строя с вероятностью q_1 , второй — с вероятностью q_2 и так далее. Наладчик, вызванный для осмотра устройства, обнаруживает и устраняет неисправность каждого узла, если она имеется, с вероятностью p , а с вероятностью $1 - p$ объявляет узел исправным. Найти вероятность того, что после осмотра наладчиком хотя бы один узел устройства будет неисправным.
25. Группа студентов состоит из N_1 отличников, N_2 хорошо успевающих и N_3 занимающихся слабо. Отличники на предстоящем экзамене могут получить только отличные оценки. Хорошо успевающие студенты могут получить с равной вероятностью хорошие и отличные оценки. Слабо занимающиеся могут получить с равной вероятностью хорошие, удовлетворительные и неудовлетворительные оценки. Для сдачи экзамена вызывается наугад один студент. Найти вероятность того, что он получит хорошую или отличную оценку.

26. Оценить с помощью неравенства Чебышёва вероятность того, что: а) при бросании монеты 500 раз число выпаданий герба будет заключено между 200 и 300; б) при бросании 10 игральные костей сумма очков отклонится от математического ожидания меньше чем на 0,8.
27. Дисперсия каждой из данных независимых случайных величин не превышает 5. Найти число этих величин, при котором вероятность отклонения их средней арифметической от математического ожидания менее чем на 0,1 превысит 0,9.
28. Оценить вероятность того, что при бросании монеты 500 раз частота появления герба отклонится от вероятности появления герба при одном бросании по модулю менее чем на 0,1.
29. Стрелок попадает при выстреле в мишень в десятку с вероятностью 0,5, в девятку — 0,3, в восьмерку — 0,1, в семерку — 0,1. Стрелок сделал 100 выстрелов. Какова вероятность того, что он набрал не менее 940 очков?
30. Приживаются в среднем 70 % числа посаженных саженцев. Сколько нужно посадить саженцев, чтобы с вероятностью не меньше 0,9 ожидать, что отклонение числа прижившихся саженцев от их математического ожидания не превышало по модулю 40? Решить задачу с помощью неравенства Чебышёва.

Литература

1. Вентцель, Е. С. Теория вероятностей / Е. С. Вентцель. — М. : КноРус, 2010.
2. Чистяков, В. П. Курс теории вероятностей / В. П. Чистяков. — М. : Дрофа, 2008.
3. Гмурман, В. Е. Теория вероятностей и математическая статистика / В. Е. Гмурман. — М. : Юрайт-Издат, 2013.
4. Боровков, А. А. Теория вероятностей / А. А. Боровков. — М. : Эдиториал УРСС, 2009.
5. Гнеденко, Б. В. Курс теории вероятностей / Б. В. Гнеденко. — М. : Либроком, 2011.
6. Феллер, В. Введение в теорию вероятностей и ее приложения. В 2 т. / В. Феллер. — М. : Либроком, 2010.
7. Зубков, А. М. Сборник задач по теории вероятностей / А. М. Зубков, Б. А. Севастьянов, В. П. Чистяков. — СПб. : Лань, 2009.
8. Прохоров, А. В. Задачи по теории вероятностей. Основные понятия. Пределные теоремы. Случайные процессы / А. В. Прохоров, В. Г. Ушаков, Н. Г. Ушаков. — М. : КДУ, 2009.
9. Румшинский, Л. З. Элементы теории вероятностей / Л. З. Румшинский. — М. : Наука, 1976.
10. Пытьев, Ю. П. Курс теории вероятностей и математической статистики для физиков / Ю. П. Пытьев, И. А. Шишмарев. — М. : Изд-во МГУ, 1983.
11. Солодовников, А. С. Теория вероятностей / А. С. Солодовников. — М. : Просвещение, 1983.
12. Справочник по теории вероятностей и математической статистике / ред. В. С. Королюк, Н. И. Портенко, А. В. Скороход, А. Ф. Турбин. — М. : Наука, 1985.
13. Вентцель, Е. С. Теория вероятностей / Е. С. Вентцель, Л. А. Овчаров. — М. : Наука, 1973.

14. Розанов, Ю.Л. Теория вероятностей, случайные процессы и математическая статистика / Ю.Л. Розанов. — М.: Наука, 1989.
15. Севастьянов, Б.А. Курс теории вероятностей и математической статистики / Б.А. Севастьянов. — М.: Наука, 1982.
16. Лозв, М. Теория вероятностей / М. Лозв. — М.: Иностранная литература, 1962.
17. Афанасьев, В.В. Теория вероятностей / В.В. Афанасьев. — М.: ВЛАДОС, 2007.
18. Проказников, А.В. Теория вероятностей и математическая статистика / А.В. Проказников, В.С. Кузнецов. — Ярославль: ЯрГУ, 2003.

Оглавление

Введение	3
1. Понятие вероятности	4
2. Классическая теоретико-множественная модель	10
3. Последовательность независимых испытаний	20
4. Последовательность взаимосвязанных испытаний	28
5. Случайные величины	32
6. Числовые характеристики случайных величин	41
7. Закон больших чисел и центральные предельные теоремы .	55
8. Случайные процессы	64
9. Элементы математической статистики	75
10. Сглаживание экспериментальных кривых.	
Критерии согласия	85
Задачи	93
Литература	97

Учебное издание

Кузнецов Владимир Степанович
Пархоменко Александр Яковлевич
Проказников Александр Владимирович

**Теория вероятностей
и математическая статистика**

Учебное пособие

Редактор, корректор М. Э. Левакова
Компьютерная верстка А. Я. Пархоменко

Подписано в печать 05.07.2014. Формат 60 × 84/16.

Усл. печ. л. 5,81. Уч.-изд. л. 5,0.

Тираж 103 экз. Заказ

Оригинал макет подготовлен
в редакционно-издательском отделе ЯрГУ.

Ярославский государственный университет им. П. Г. Демидова.
150000 Ярославль, ул. Советская, 14.