Министерство образования и науки
Государственное образовательное учреждение высшего профессионального образования
Национальный исследовательский Томский политехнический университет

О. Л. Крицкий, А. А. Михальчук, А. Ю. Трифонов, М. Л. Шинкеев

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

для технических университетов

Теория вероятностей

Учебное пособие

Издательство ТПУ

Томск 2010

УДК 581
К14

К14 **Крицкий О. Л., Михальчук А. А., Трифонов А. Ю., Шинкеев М. Л.** Теория вероятностей и математическая статистика для технических университетов. I. Теория вероятностей: Учебное пособие. — Томск: Изд-во ТПУ, 2010. — 212 с.

Настоящее пособие представляет собой изложение первой части курса «Теория вероятностей и математическая статистика» и содержит материал по разделу «Теория вероятностей» этого курса. Оно содержит теоретический материал в объёме, предусмотренном ныне действующей программой курса высшей математики для инженерно-физических специальностей и специальности «Прикладная математика» технических университетов. Теоретический курс дополнен индивидуальными заданиями для самостоятельного решения по каждому разделу. Предлагаемое пособие может быть полезно студентам старших курсов, магистрантам и аспирантам, специализирующимся в области прикладной и финансовой математики, теоретической и математической физики.

Пособие предназначено для студентов инженерно-физических специальностей, студентов, специализирующимся в области финансовой математики, и студентов, обучающихся в системе элитного технического образования.

УДК 581

Рекомендовано к печати Редакционно-издательским советом
Томского политехнического университета

Работа частично поддержана
АВЦП Министерства образования и науки РФ № 2.1.1/3436;
Федеральным агентством по науке и инновациям России
по контракту № 02.740.11.0238

Рецензенты

Доктор физико-математических наук, профессор Томского государственного
педагогического университета, г. Томск
Осетрин К.Е.

Доктор физико-математических наук, профессор Томского государственного
университета, г. Томск
Багров В.Г.

Содержание

Введение

Во многих вопросах науки и техники, производства и экономики, военного дела и управления приходится встречаться с ситуациями, исход которых не поддается точному прогнозированию. Так, например,

- при стрельбе из орудия нельзя предсказать, что посланный снаряд попадёт в цель;
- при бросании монеты нельзя предсказать, что она упадёт гербом вверх;
- при передаче сообщений (например, телеграфных знаков) по каналу связи могут появиться помехи, которые исказят сообщение или полностью подавят его;
- нельзя предсказать цену акций в заданный момент времени;
- аппаратура может отказать на любом этапе своей работы из-за отказа каких-либо деталей, но мы не можем предсказать, произойдёт это или нет;
- при покупке лотерейного билета нельзя предсказать, выпадет ли выигрыш на купленный билет.

Такого рода примеров можно привести очень много. Что общего в подобных ситуациях?

Дело в том, что в каждом приведённом примере на исход влияет очень большое число разного рода причин, законы действия которых неизвестны. Так, в первом примере — при стрельбе из орудия — можно назвать некоторые из таких причин: скорость и направление ветра, осадки, температура воздуха и т.д.

Все это приводит к тому, что результат опыта не определяется однозначно; говорят, что результат такого опыта случаен.

Выявлением и изучением закономерностей в случайных явлениях занимается специальная область математики — теория вероятностей и её многочисленные ответвления. Предметом теории вероятностей является изучение вероятностных закономерностей массовых однородных случайных событий.

Строят эту теорию дедуктивно, исходя из некоторых аксиом и определений.

Методы теории вероятностей по своей природе приспособлены для исследования массовых явлений, т.е. таких явлений, которые многократно повторяются или которые можно воспроизвести сколько угодно раз при сохранении неизменным основного комплекса условий. Эти методы не дают возможности предсказать исход отдельного случайного явления, но позволяют предсказать суммарный результат большого числа однородных случайных явлений.

ЧАСТЬ I

Теория вероятностей

ГЛАВА 1

Введение в теорию вероятностей

1. Алгебра событий

Современное построение теории вероятностей как раздела математики основывается на аксиоматическом подходе и опирается на элементарные понятия теории множеств. Теория вероятностей оперирует с действительными или мысленными опытами, имеющими случайный исход.

Одним из основных понятий теории вероятностей является случайное событие или просто событие.

Событие — первичное понятие в теории вероятностей в том смысле, что его нельзя определить через другие, более элементарные понятия.

Под событием понимается всякий факт, который в результате опыта (испытания, эксперимента) может произойти или не произойти.

♦ Под *испытанием* (*опытом*, *экспериментом*) будем понимать осуществление какого-либо определённого комплекса условий, которые могут быть воспроизведены сколь угодно большое число раз (реально или мысленно).

События обозначаются прописными буквами латинского алфавита: A, B, C и т.д.

Приведём примеры событий:

A — попадание в цель при выстреле (опыт — произведение выстрела, событие — попадание в цель);

B — появление герба при бросании монеты;

C — выход из строя детали сотового телефона при испытании.

Условимся различать составные (или разложимые) и элементарные (неразложимые) события.

Во избежание неясностей при описании случайных явлений, результатов опыта или наблюдений необходимо формализовать эти описания. С этой целью вводится множество Ω элементарных исходов эксперимента (пространство элементарных событий).

♦ Совокупность (множество) всех мыслимых исходов опыта, таких, что в результате испытания может произойти один и только один из них, назовем *пространством элементарных событий* и будем обозначать его Ω. Каждый элемент $\omega \in \Omega$ (отдельный исход опыта) будем называть *элементарным событием.*

Пространство элементарных событий является первоначальным понятием при построении математической модели явлений, имеющих случайную (стохастическую) природу.

♦ Любое подмножество множества элементарных событий называется *событием* (или *случайным событием*).

В дальнейшем утверждение «A есть подмножество Ω» записывается в виде $A \subset \Omega$. Говорят, что событие A наступило, если результатом испытания явился исход ω, принадлежащий множеству A ($\omega \in A$).

Рассмотрим ряд примеров, поясняющих выбор множества Ω.

Пример 1.1. При однократном подбрасывании монеты пространство исходов

Ω состоит из двух точек:

$$\Omega = \{Г, Р\},$$

где Г – «герб», Р – «решётка» (просторечное «решка»).

Мы исключаем возможности типа «монета стала на ребро», «монета исчезла» и т.д.

Пример 1.2. Подбрасывание игральной кости один раз. В этом опыте естественно выбрать

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\},$$

где через ω_k обозначен исход опыта, заключающийся в выпадении k очков. В этом опыте возможны шесть исключающих друг друга исходов. Пусть нас интересует событие A — выпадение чётного числа очков:

$$A = \{\omega_2, \omega_4, \omega_6\}, \quad A \subset \Omega.$$

В частности, можно рассматривать событие Ω (ведь каждое множество есть свое собственное подмножество).

Пример 1.3 (задача о встрече). Два студента условились встретиться в назначенном месте между одиннадцатью и двенадцатью часами и ждать друг друга в течение 10 мин (но не долее, чем до истечения часа). Каждый из них наудачу и независимо от другого выбирает момент своего прихода. Описать пространство Ω и событие C, означающее, что встреча состоится.

Решение. Обозначим через x время прихода одного студента и через y время прихода второго. Каждый из них может прийти в любой момент между 11 и 12 часами, поэтому

$$11 \leqslant x \leqslant 12 \quad \text{и} \quad 11 \leqslant y \leqslant 12.$$

В системе координат выберем за начало отсчёта число 11, тогда пространство элементарных событий Ω представляет собой множество точек квадрата

$$\Omega = \left\{ (x, y) : \begin{array}{l} 11 \leqslant x \leqslant 12, \\ 11 \leqslant y \leqslant 12. \end{array} \right\}$$

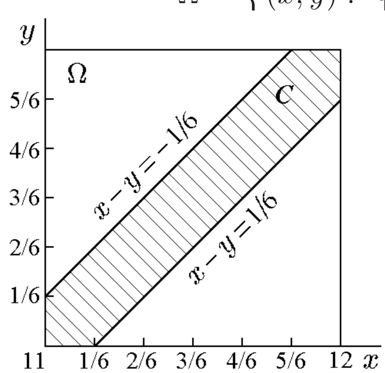


Рис. 1.

Не любая точка квадрата определяет ситуацию, благоприятствующую встрече, например точка (11⅙, 11½). Первый пришел в 11 час 10 мин, второй – в 11 час 30 мин, но первый уже ушел, прождав его условленные 10 мин.

Какие же исходы опыта будут благоприятствовать встрече? Те, для которых разность между моментами прихода студентов окажется не более 10 мин, т.е.

$$|x - y| \leqslant 1/6.$$

Построим область $-1/6 \leqslant x - y \leqslant 1/6$; её граничными линиями будут

$$x - y = -1/6;$$
$$x - y = 1/6.$$

Множество исходов опыта, благоприятствующих встрече, есть область C, заштрихованная на рис. 1.

♦ Событие Ω называется *достоверным*, если оно не может не произойти в условиях данного опыта или явления.

Ко всему пространству элементарных событий добавляется еще пустое множество $\emptyset$. Это множество тоже рассматривается как событие и называется невозможным событием.

♦ Событие $\emptyset = \{\omega \in \emptyset\}$ называется *невозможным*, если оно не может произойти при выполнении условий данного опыта.

Пример достоверного события — выпадение не более 6 очков при бросании игральной кости, пример невозможного события — выпадение 7 очков при бросании игральной кости.

Заметим, что пространство элементарных событий в одном и том же опыте можно задавать по-разному; например, при случайном выборе точки на плоскости её положение можно задавать как декартовыми координатами (x, y), так и полярными (ρ, φ).

♦ Событие $\overline{A}$ называется *противоположным* событию A или его *дополнением*, если оно происходит тогда и только тогда, когда не происходит событие A:

$$\overline{A} = \Omega - A.$$

♢ Для наглядности воспользуемся диаграммами Венна–Эйлера. Будем изображать область Ω в виде прямоугольника, каждая точка которого соответствует элементарному исходу, а событиям — некоторые области внутри Ω. Тогда событию $\overline{A}$, противоположному событию A, можно сопоставить заштрихованную на рис. 2 область.

Очевидно, что событие $\overline{A}$ состоит из элементов множества Ω, которые не входят в множество A (см. рис. 2). Событие, противоположное достоверному, является невозможным, т.е. $\overline{\Omega} = \emptyset$. Событие $\overline{\overline{A}}$, противоположное к противоположному, есть событие A ($\overline{\overline{A}} = A$).

♦ Событие C называется *суммой* или *объединением двух событий* A и B, если оно происходит лишь тогда, когда происходит по крайней мере одно из событий A или B:

$$C = A + B = A \cup B.$$

Определение можно распространить на любое конечное число слагаемых.

На диаграмме Венна–Эйлера сумме двух событий $A + B$ можно сопоставить заштрихованную на рис. 3 область.

♢ Из определения вытекают следующие свойства операции сложения:

$$\begin{aligned}
&A + B = B + A;\\
&(A + B) + C = A + (B + C);\\
&A + A = A;\\
&A + \Omega = \Omega;\\
&A + \emptyset = A;\\
&A + \overline{A} = \Omega.
\end{aligned}$$

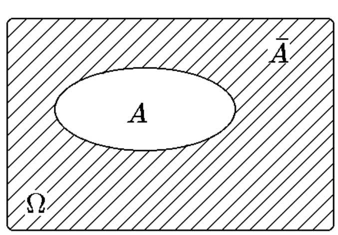


Рис. 2.

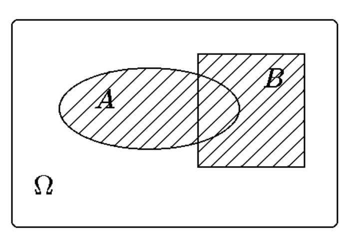


Рис. 3.

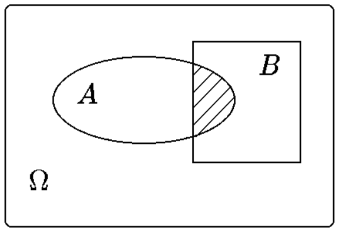


Рис. 4.

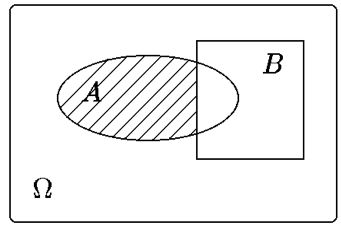


Рис. 5.

♦ Событие C называется *произведением двух событий* A и B, если оно происходит тогда, когда происходят и событие A, и событие B:

$$C = A \cdot B = A \cap B.$$

Таким образом, произведение событий A и B есть множество, состоящее из элементов, принадлежащих как множеству A, так и множеству B. На диаграмме Венна–Эйлера произведению двух событий $A \cap B$ сопоставим заштрихованную область на рис. 4.

◊ Из определения вытекают следующие свойства операции умножения:

$$\begin{aligned} &A \cdot B = B \cdot A; \\ &A \cdot (B \cdot C) = (A \cdot B) \cdot C; \\ &A \cdot A = A; \\ &A \cdot \Omega = A; \\ &A \cdot \emptyset = \emptyset; \\ &A \cdot \overline{A} = \emptyset. \end{aligned}$$

♦ *Разностью двух событий* A и B называется такое событие C, которое происходит лишь тогда, когда происходит событие A, а событие B не происходит:

$$C = A - B = A \setminus B.$$

Событию C сопоставляется заштрихованная область на рис. 5.

Из определения следует, что $A - B = A\overline{B}$, то есть разность событий можно свести к их произведению.

♦ Говорят, что *событие* A *влечёт событие* B ($A \subset B$), если при наступлении события A неизбежно наступает событие B (см. рис. 6).

В этом случае множество A является подмножеством B. Так, при бросании игральной кости событие A — «выпадет два очка» влечёт событие B — «выпадет чётное число очков».

♦ Если $A \subset B$, а $B \subset A$, то A и B называются *эквивалентными* (*равносильными*), что записывается как $A = B$.

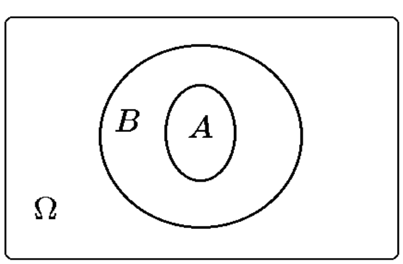


Рис. 6.

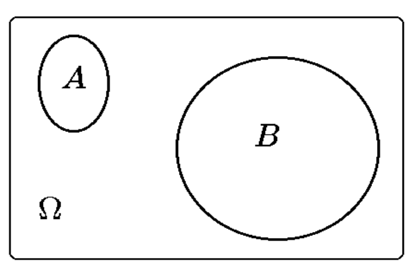


Рис. 7.

Например, при бросании двух костей события: A — «выпадет чётная сумма» и B — «на каждой грани выпадут очки одной чётности» являются эквивалентными.

♦ События A и B *несовместны*, если $A \cdot B = \emptyset$ (см. рис. 7). В частности, несовместными событиями являются противоположные.

♦ События $A_1, A_2, \ldots, A_n$ образуют *полную группу событий*, если хотя бы одно из них обязательно случится в результате опыта.

Иными словами, события $A_1, \ldots, A_n$, $n \leqslant \infty$, образуют полную группу, если их объединение есть достоверное событие:

$$\bigcup_{i=1}^{n} A_i = \Omega \quad \text{или} \quad \sum_{i=1}^{n} A_i = \Omega.$$

Помимо отмеченных выше, операции над событиями обладают следующими свойствами:

1. $(A + B)C = AC + DC$.
2. $AB + C = (A + C)(B + C)$.
3. $\overline{A + B} = \overline{A} \cdot \overline{B}$.
4. $\overline{A \cdot B} = \overline{A} + \overline{B}$.

Доказательства этих свойств дословно повторяют доказательства соответствующих свойств операций над множествами в теории множеств.

Из установленных свойств операций над событиями следует, что для любых событий A и B

$$A = A\Omega = A(B + \overline{B}) = AB + A\overline{B}. \tag{1.1}$$

Эта формула дает разложение любого события A на два непересекающихся события.

Если $B \subset A$, то $AB = B$, и из (1.1) следует $A = B + A\overline{B}$ или $A = B \cup A\overline{B}$.

Пример 1.4. Установить, справедливо ли соотношение $A - (B - C) = (A - B) + C$.

Решение. Преобразуем правую и левую части, используя свойства операций с событиями: $D = (A - B) + C = A\bar{B} + C$, аналогично $E = A - (B - C) = A(\overline{B - C}) = A\overline{B\bar{C}} = A(\bar{B} + C) = A\bar{B} + AC$. Следовательно, $E \subset D$, так как $AC \subset C$, а соотношение $C \subset AC$ в общем случае не справедливо, то $D \neq E$, и исходное утверждение имеет вид $A - (B - C) \supset (A - B) + C$.

Пример 1.5. Дана система из трех блоков a_1, a_2, b (см. рис. 8), где a_1 и a_2 — дублирующие блоки (схема последовательного соединения блока b с подсистемой a параллельно соединенных блоков a_1 и a_2).

Записать события, состоящие в том, что система 1) исправна, 2) неисправна.

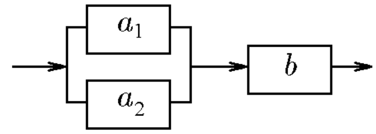


Рис. 8.

Решение. Введём обозначения: A_1 — событие, состоящее в том, что блок a_1 исправен; A_2 — событие, состоящее в том, что блок a_2 исправен; B — событие, состоящее в том, что блок b исправен; S — событие, состоящее в том, что система исправна; S_a — событие, состоящее в том, что подсистема a исправна. Разобьем систему на две подсистемы a и b. Подсистема a дублирующих блоков исправна в том случае, когда исправен хотя бы один из блоков a_k $(k = 1, 2)$ исправен, то есть $S_a = A_1 + A_2$. Для исправности системы необходима исправность подсистемы a и блока b, то есть $S = BS_a$. Таким образом, $S = B(A_1 + A_2)$. Аналогичные рассуждения могут привести к $\bar{S} = \bar{A}_1 \cdot \bar{A}_2 + \bar{B}$.

Заметим, что $\bar{S}$ можно построить по S с помощью свойств противоположных событий.

♦ Множество $\mathcal{F}$, состоящее из подмножеств множества Ω, называется *алгеброй событий*, если
1) множество $\mathcal{F}$ содержит достоверное событие, т.е $\Omega \in \mathcal{F}$;
2) вместе с событием A множество $\mathcal{F}$ содержит противоположное ему событие $\overline{A}$;
3) вместе с любым конечным набором событий множество $\mathcal{F}$ содержит объединение этих событий.

Условия 1–3 называют аксиомами алгебры событий.

Пример 1.6. Два стрелка стреляют по мишени. Пусть событие A — попал в цель первый стрелок, событие B — попал в цель второй стрелок. Записать на языке алгебры событий следующие события: C — оба стрелка попали в цель; D — оба стрелка промахнулись; E — попал в цель ровно один стрелок; F — хотя бы один стрелок попал в цель.

Решение. Событие C заключается в одновременном осуществлении событий A и B, что по определению является произведением этих событий, следовательно $C = AB$.

Событие D заключается в том, что не произойдёт ни событие A, ни событие B, то есть в одновременном осуществлении событий, противоположных событиям A и B, что есть произведение противоположных событий, следовательно, $D = \bar{A}\bar{B}$.

Событие E произойдёт, если в цель попадёт первый стрелок, а второй при этом не попадёт, либо если попадёт в цель второй стрелок, а первый при этом промахнётся. Событие, заключающееся в том, что первый стрелок попал, а второй не попал в цель, есть произведение события A и события, противоположного событию B, то есть это событие $A\bar{B}$. Аналогично событие, заключающееся в том, что второй стрелок попал в цель, а первый не попал, есть $\bar{A}B$. Поскольку нас устраивает любое из этих событий, то событие E есть сумма этих событий: $E = A\bar{B} + \bar{A}B$.

Событию F удовлетворяют следующие элементарные события: попал в цель только первый стрелок, попал в цель только второй стрелок, оба стрелка попали в цель. Следовательно, F есть сумма этих событий: $F = A\bar{B} + \bar{A}B + AB$. Кроме того, можно заметить, что событие, противоположное к событию F, есть событие «ни один стрелок не попал в цель»; следовательно, событие F можно выразить через противоположное событие следующим образом: $F = \overline{\bar{A}\bar{B}}$. И, наконец, событие F — попал в цель хотя бы один стрелок — по определению есть сумма событий A и B: $F = A + B$. Используя свойства операций над событиями, легко убедиться, что все формулы тождественны. Последняя запись является хоть и самой простой, но не самой удобной, поскольку в ней мы выразили событие через сумму совместных событий. Как правило, если будет идти речь о наступлении хотя бы одного события из группы событий, удобнее всего выразить это событие через противоположное.

♦ Алгебра событий $\mathcal{F}$, замкнутая относительно счётных объединений, называется *σ-алгеброй.*

Алгебра событий обладает следующими свойствами.

Свойство 1. Невозможное событие принадлежит множеству $\mathcal{F}$ ($\emptyset \in \mathcal{F}$).

Доказательство. Согласно аксиоме 1, $\Omega \in \mathcal{F}$, следовательно, $\emptyset = \overline{\Omega} \in \mathcal{F}$ в силу аксиомы 2.

Свойство 2. Если $A_1, A_2, \dots \in \mathcal{F}$, то $\prod\limits_i A_i \in \mathcal{F}$.

Доказательство. Согласно свойствам операций над событиями, $\prod\limits_i A_i = \overline{\sum\limits_i \overline{A}_i}$, а в силу аксиом 2 и 3 справедливо $\overline{\sum\limits_i \overline{A}_i} \in \mathcal{F}$.

Таким образом, применение счётного числа операций над событиями, таких как объединение, пересечение, дополнение, к множествам из $\mathcal{F}$ снова даёт множество из $\mathcal{F}$, или, как говорят, множество $\mathcal{F}$ замкнуто относительно этих операций.

Множество $\mathcal{F}$ будем также называть *полем событий*, связанных с данным испытанием.

Пример 1.7. Подбрасывается игральная кость. Привести примеры событий, являющихся и не являющихся σ-алгебрами.

Решение. Пусть $\Omega = \{1, 2, 3, 4, 5, 6\}$, где k-ый исход означает, что выпало k очков. Тогда, например, следующие множества $\mathcal{F}$ являются σ-алгебрами:
1) $\mathcal{F} = \{\Omega, \emptyset\}$ — тривиальная σ-алгебра;
2) $\mathcal{F} = \{\Omega, \emptyset, \{1, 3, 5\}, \{2, 4, 6\}\}$;
3) $\mathcal{F}$ — множество всех подмножеств множества Ω.
Примером множества, не являющегося σ-алгеброй, может служить, например, множество $\mathcal{F} = \{\Omega, \emptyset, \{1, 2\}, \{3, 4\}, \{5, 6\}\}$, так как, в частности, $\{1, 2\} \cup \{3, 4\} = \{1, 2, 3, 4\} \notin \mathcal{F}$.

2. Частота события. Свойства частоты события

Естественно сравнивать события по тому, как часто каждое из них появляется при повторении данного опыта. Если при повторении опыта одно событие появляется чаще, чем другое, то говорят, что первое *вероятнее* второго. Ясно, что для сравнения событий необходимо предположить, что данный опыт можно производить сколько угодно (в принципе, неограниченное число) раз.

♦ *Частотой* (*относительной частотой*) события называется отношение числа его появлений к числу всех произведённых опытов.

Таким образом, частота события A определяется формулой

$$\mu_n(A) = \frac{m}{n} = \frac{m_n(A)}{n},$$

где n — общее число испытаний; $m = m_n(A)$ — число появлений события A в n испытаниях.

Опыт показывает, что при многократном повторении испытаний частота $\mu_n(A)$ случайного события обладает устойчивостью. В качестве иллюстрации рассмотрим данные по проверке симметричности монеты. Пусть m_n — число выпадений герба в n испытаниях, так что m_n/n — частота выпадения герба.

В следующей таблице приведены результаты появления герба при подбрасывании монеты, экспериментально полученные разными исследователями, начиная с XVIII в. Их фамилии помещены в первом столбце таблицы, во втором столбце указано число подбрасываний, а в третьем — частота появления герба.

Бюффон	4040	0,507
Де Морган	4092	0,5005
Джевонс	20480	0,5068
Романовский	80640	0,4923
Пирсон	24000	0,5005
Феллер	10000	0,4979

Таким образом, мы видим, что при большом числе бросаний монеты частота появления герба мало отличается от числа 0,5, т.е. обладает устойчивостью.

Отметим основные свойства частоты.

Свойство 1. Частота любого события представляет собой неотрицательное число, не превосходящее 1, причём частота невозможного события равна нулю, а частота достоверного события равна 1:

$$\begin{gathered} 0 \leqslant \mu_n(A) \leqslant 1, \\ \mu_n(\emptyset) = 0, \quad \mu_n(\Omega) = 1. \end{gathered} \tag{2.1}$$

Свойство 2. Частота появления одного из несовместных событий, безразлично какого именно, равна сумме их частот. Это следует непосредственно из того, что число появлений сложного события, состоящего из нескольких несовместных событий, равно сумме чисел появления каждого из этих событий

$$\mu_n(A+B) = \mu_n(A) + \mu_n(B) = \frac{m_n(A)}{n} + \frac{m_n(B)}{n}.$$

Очевидно, что если частота события в данной серии опытов равна нулю (или единице), то из этого не следует, что событие невозможно (достоверно). Так, например, если при пяти бросаниях монеты герб не появился ни разу, то из этого не следует, что появление герба невозможно.

Устойчивость частоты события дает основание считать, что с каждым событием связано некоторое число — *вероятность этого события*, — около которого стремится стабилизироваться его частота. Например, частота появления герба при бросании монеты, очевидно, должна стабилизироваться около 1/2. Следовательно, вероятность появления герба равна 1/2.

Вероятность события A обозначается $P(A)$. Это, конечно, не исключает применения сокращенных обозначений, например $P(A) = p$ и т.п.

Понятие вероятности события является первичным и поэтому не нуждается в определении. Оно представляет собой результат абстрагирования, необходимый для построения любой теории. Отвлекаясь от сложных и несущественных колебаний частоты при неограниченном повторении опытов и оставляя основную, существенную закономерность, наблюдаемую в данном явлении, — устойчивость частоты, — мы и вводим абстрактное понятие вероятности события.

Вероятность события в данном опыте — его объективная характеристика. Она имеет вполне определённое значение независимо от того, собираемся мы производить опыты или нет.

♦ Под *статистической вероятностью события A* понимают предел, к которому стремится частота события при неограниченном увеличении числа испытаний.

◇ Таким образом, под статистической вероятностью события понимается число, относительно которого частота события $\mu(A)$ устойчиво колеблется и к которому приближается с незначительными отклонениями, носящими случайный незакономерный характер, при неограниченном увеличении числа испытаний. Математическое содержание понятия «устойчиво колеблется» будет рассмотрено более подробно в главе «Предельные теоремы. Закон больших чисел».

◇ Из определения следует, что если опытным путем установлена относительная частота некоторого события A, то полученное число можно принять за приближённое значение вероятности этого события при условии, что n достаточно велико. Тем не менее, такое определение вероятности оказывается неудобным. Во-первых, всю последовательность частот получить невозможно. Кроме того, относительная частота $\mu(A)$, как мы уже отмечали, обнаруживает устойчивость именно потому, что данному событию A присуща некоторая вероятность $P(A)$. То есть именно устойчивость относительной частоты события порождена вероятностью события, а не наоборот, как в данном определении.

3. Аксиомы теории вероятностей

При аксиоматическом подходе к определению вероятности последняя задается перечислением ее свойств. Простейшие свойства вероятности определяются естественными свойствами частоты $\mu(A)$, указанными выше.

Пусть каждому событию ставится в соответствие некоторое число, называемое *вероятностью события.*

♦ *Вероятностью* называется числовая функция $P(A)$, заданная на множестве событий, образующих σ-алгебру $\mathcal{F}$, если выполняются следующие аксиомы.

Аксиома 1. *Вероятность любого события A неотрицательна:*

$$0 \leqslant P(A). \tag{3.1}$$

Аксиома 2. *Вероятность достоверного события равна единице:*

$$P(\Omega) = 1. \tag{3.2}$$

Аксиома 3 (сложения вероятностей). *Если $A_1, A_2, \ldots, A_n, \ldots$ — несовместные события, то*

$$P\biggl(\sum_{i=1}^{\infty} A_i\biggr) = \sum_{i=1}^{\infty} P(A_i). \tag{3.3}$$

Ныне принятое аксиоматическое определение вероятности было введено в 1933 г. А.Н. Колмогоровым.

Аксиомы теории вероятностей позволяют вычислять вероятности любых событий (подмножеств пространства Ω) с помощью вероятностей элементарных событий. Вопрос о том, как определить вероятности элементарных событий, при этом не рассматривается. На практике они определяются либо из соображений, связанных с симметрией опыта (например, для симметричной игральной кости естественно считать одинаково вероятным выпадение каждой из граней), либо же на основе опытных данных (частот).

◊ Заметим, что вероятность события A, определённая аксиомами 1–3, задается не на пространстве Ω, а на некоторой σ-алгебре событий, определённой на Ω. Можно показать, что существуют множества $A \subset \Omega$, для которых нельзя определить вероятность, которая удовлетворяла бы аксиомам 1–3. Поэтому в дальнейшем мы будем рассматривать только те множества $A \subset \Omega$, для которых мы можем определить вероятность.

♦ Тройка $R = \langle \Omega, \mathcal{F}, P \rangle$, где Ω — пространство элементарных исходов, $\mathcal{F}$ — σ-алгебра его подмножеств, а P — вероятностная мера на $\mathcal{F}$, называется *вероятностным пространством.*

Итак, вероятность есть функция P: $\mathcal{F} \to R$, удовлетворяющая условиям аксиом 1–3, или, как говорят, нормированная (вероятностная) мера, заданная на множестве $\mathcal{F}$.

◊ Можно показать, что аксиома 3 эквивалентна двум следующим аксиомам.

Аксиома 4. *Если A и B несовместны, то $P(A + B) = P(A) + P(B)$.*

Аксиома 5. *Если* $A_1 \supset A_2 \supset A_3 \supset \dots \supset A_n \supset \dots$ *и* $A = \bigcap\limits_{i=1}^{\infty} A_i$ *или* $A_1 \subset A_2 \subset A_3 \subset \dots \subset A_n \subset \dots$ *и* $A = \sum\limits_{i=1}^{\infty} A_i$, *то* $P(A) = \lim\limits_{n\to\infty} P(A_n)$.

4. Свойства вероятности

Рассмотрим основные свойства вероятности.

Свойство 1. Вероятность невозможного события равна нулю:

$$P(\emptyset) = 0. \tag{4.1}$$

Действительно, $\Omega = \Omega + \emptyset$, а события Ω и $\emptyset$ несовместны: $\Omega\emptyset = \emptyset$. Тогда, согласно третьей аксиоме теории вероятностей,

$$P(\Omega) = P(\Omega + \emptyset) = P(\Omega) + P(\emptyset).$$

Отсюда следует, что $P(\emptyset) = 0$, так как, согласно аксиоме 2, $P(\Omega) = 1$.

Свойство 2. Для любого события A вероятность противоположного события $\overline{A}$ выражается равенством

$$P(\overline{A}) = 1 - P(A). \tag{4.2}$$

Действительно, $\Omega = A + \overline{A}$, а события A и $\overline{A}$ несовместны: $A\overline{A} = \emptyset$. Следовательно,

$$P(\Omega) = P(A + \overline{A}) = P(A) + P(\overline{A})$$

или

$$1 = P(A) + P(\overline{A}).$$

Свойство 3. Если событие A влечёт за собой событие B, т.е. $A \subset B$, то вероятность события C, где C — разность событий B и A, определяется соотношением

$$P(C) = P(B \setminus A) = P(B) - P(A).$$

Действительно, если $A \subset B$, то событие B можно представить в виде суммы несовместных событий $B = A + (B \setminus A)$. Тогда

$$P(B) = P(A) + P(B \setminus A),$$

откуда следует, что (см. рис. 6)

$$P(B \setminus A) = P(B) - P(A).$$

Свойство 4. Если событие A влечёт за собой событие B, т.е. $A \subset B$, то вероятность события A не может быть больше вероятности события B, т.е. $P(A) \leqslant P(B)$.

Действительно, в силу предыдущего свойства, если $A \subset B$, то $P(A) = P(B) - P(B \setminus A)$. Но, согласно аксиоме 1,

$$P(B \setminus A) \geqslant 0,$$

откуда следует, что (см. рис. 6)

$$P(A) \leqslant P(B).$$

Свойство 5. Вероятность любого события заключена между нулем и единицей:

$$0 \leqslant P(A) \leqslant 1,$$

Справедливость этого утверждения непосредственно следует из аксиом 1 и 2 и свойства 4.

Свойство 6 (теорема сложения вероятностей). Вероятность суммы любых двух событий равна сумме вероятностей этих событий минус вероятность их совместного появления:

$$P(A+B) = P(A) + P(B) - P(AB). \tag{4.3}$$

Действительно, событие $A+B$ можно представить как сумму несовместных событий: $A+B = B + (A \setminus (AB))$ (см. рис. 9). Тогда

$$P(A+B) = P(B) + P(A \setminus (AB)).$$

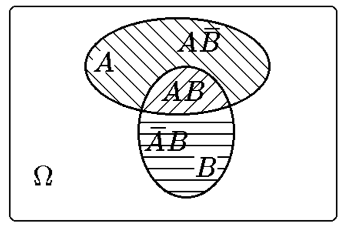


Рис. 9.

Но $AB \subset A$. Следовательно, согласно свойству 3 (см. также рис. 5),

$$P(A \setminus AB) = P(A) - P(AB).$$

В частности, если события A и B несовместны, то $P(AB) = P(\emptyset) = 0$ и

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

Свойство 7. Вероятность суммы событий не превосходит сумму вероятностей этих событий:

$$P(A+B) \leqslant P(A) + P(B). \tag{4.4}$$

Справедливость соотношения (4.4) следует непосредственно из предыдущего свойства с учетом аксиомы 1.

Соотношение (4.3) может быть обобщено на любое количество событий.

Свойство 8 (общее правило сложения вероятностей). Вероятность суммы n событий $A_1, A_2, \ldots, A_n$ может быть вычислена по формуле

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{1\leqslant i<j}^{n} P(A_iA_j) + \sum_{1\leqslant i<j<k}^{n} P(A_iA_jA_k) - \\ - \sum_{1\leqslant i<j<k<l}^{n} P(A_iA_jA_kA_l) + \ldots + (-1)^{n-1} P(A_1A_2\cdots A_n). \tag{4.5}$$

Соотношение (4.5) доказывается методом математической индукции. Оно справедливо для $n=2$ в силу (4.3). Предположим теперь, что оно справедливо для суммы $n-1$ событий, и докажем его справедливость для суммы n событий.

Для суммы $n-1$ событий $A_2, A_3, \ldots, A_n$ имеем

$$P\left(\sum_{i=2}^{n} A_i\right) = \sum_{i=2}^{n} P(A_i) - \sum_{2\leqslant i<j}^{n} P(A_iA_j) + \sum_{2\leqslant i<j<k}^{n} P(A_iA_jA_k) - \ldots$$

Для суммы $n-1$ событий $A_1A_2, A_1A_3, \ldots, A_1A_n$ по той же формуле имеем

$$P\left(\sum_{i=2}^{n} A_1A_i\right) = \sum_{i=2}^{n} P(A_1A_i) - \sum_{2\leqslant i<j}^{n} P(A_1A_iA_j) + \sum_{2\leqslant i<j<k}^{n} P(A_1A_iA_jA_k) - \ldots$$

Тогда, представив сумму n событий в виде суммы двух событий: A_1 и $\sum\limits_{i=2}^{n} A_i$, получим

$$P\left(\sum_{i=1}^{n} A_i\right) = P\left(A_1 + \sum_{i=2}^{n} A_i\right) = P(A_1) + P\left(\sum_{i=2}^{n} A_i\right) - P\left(A_1 \sum_{i=2}^{n} A_i\right) =$$

$$= P(A_1) + \sum_{i=2}^{n} P(A_i) - \sum_{2\leqslant i<j}^{n} P(A_iA_j) + \sum_{2\leqslant i<j<k}^{n} P(A_iA_jA_k) - \ldots$$

$$-\left[\sum_{i=2}^{n} P(A_1A_i) - \sum_{2\leqslant i<j}^{n} P(A_1A_iA_j) + \sum_{2\leqslant i<j<k}^{n} P(A_1A_iA_jA_k) - \ldots\right] =$$

$$= \sum_{i=1}^{n} P(A_i) - \sum_{1\leqslant i<j}^{n} P(A_iA_j) + \sum_{1\leqslant i<j<k}^{n} P(A_iA_jA_k) - \ldots + (-1)^{n-1} P(A_1A_2\cdots A_n).$$

Таким образом, справедливость соотношения (4.5) доказана.

◊ В частности, для трёх событий из (4.5) следует

$$P(A+B+C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).$$

5. Классическое определение вероятности

Как уже отмечалось, аксиомы и теоремы теории вероятностей позволяют вычислять вероятности любых событий с помощью вероятностей элементарных событий, но они не позволяют определить вероятность самих элементарных событий. В тех случаях, когда элементарные события обладают свойством симметрии в том смысле, что все элементарные события находятся в одинаковом отношении к условиям, которые определяют характер испытания, можно использовать классическое определение вероятности. Например, бросание игральной кости или монеты обладает свойством симметрии по отношению к выпадению того или иного числа очков на кости или той или иной стороны монеты. Таким же свойством симметрии обладают правильно организованная жеребьёвка или тираж лотереи.

Для опытов, обладающих симметрией возможных исходов, применяется способ непосредственного вычисления вероятностей событий в так называемой *схеме случаев* (иначе — *схеме урн*). Этот способ основан на допущении о равновероятности (равновозможности) элементарных событий.

♦ Несколько событий $A_1, A_2, \ldots, A_n$ называются *равновозможными*, если в силу симметрии условий опыта относительно этих событий вероятности их одинаковы:

$$P(A_1) = P(A_2) = \ldots = P(A_n).$$

Если в каком-либо опыте пространство элементарных событий Ω можно представить в виде полной группы несовместных и равновозможных событий $\omega_1, \omega_2, \ldots, \omega_n$:

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\},$$

то такие события называются *случаями* (или *шансами*), а про сам опыт говорят, что он *сводится к схеме случаев* (*схеме урн*).

♦ Событие ω_i называется *благоприятствующим* событию A, если наступление события ω_i влечёт за собой наступление события A:

$$\omega_i \in A.$$

Так как случаи (события) $\omega_1, \omega_2, \dots, \omega_n$ образуют полную группу событий, то

$$\sum_{i=1}^{n} \omega_i = \Omega. \tag{5.1}$$

Так как элементарные события $\omega_1, \omega_2, \dots, \omega_n$ несовместны, то, согласно аксиоме сложения вероятностей,

$$\sum_{i=1}^{n} P(\omega_i) = P\left(\sum_{i=1}^{n} \omega_i\right) = P(\Omega) = 1. \tag{5.2}$$

Так как элементарные события $\omega_1, \omega_2, \dots, \omega_n$ равновозможны, то вероятность каждого из них одна и та же и равна $1/n$:

$$P(\omega_1) = P(\omega_2) = \ldots = P(\omega_n) = \frac{1}{n}. \tag{5.3}$$

Отсюда непосредственно следует так называемая классическая формула для вероятности события $A = \sum\limits_{i=1}^{m} \omega_i$:

$$P(A) = \sum_{i=1}^{m} P(\omega_i) = P\left(\sum_{i=1}^{m} \omega_i\right) = \frac{m}{n}.$$

♦ Если опыт сводится к схеме случаев, то вероятностью события называют *отношение числа m благоприятных этому событию исходов к общему числу n всех равновозможных исходов испытания*:

$$P(A) = \frac{m}{n}. \tag{5.4}$$

Здесь m — число случаев, благоприятствующих событию A; n — общее число случаев.

Формула (5.4), принимавшаяся когда-то за определение вероятности, при современном аксиоматическом подходе есть следствие аксиомы сложения вероятностей. Ее можно записать так:

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

Это классическое определение вероятности, $|A|$ и $|\Omega|$ есть обозначение числа элементов *конечных* множеств A, Ω.

Таким образом, будем говорить, что имеется *классический случайный опыт*, если

1) пространство Ω конечно,

2) элементарные события равновозможны.

Прежде чем рассматривать примеры на вычисление вероятностей по классическому определению, нужно научиться вычислять $|A|$ и $|\Omega|$, т.е. число элементов множества A и число элементов множества Ω.

Пример 5.1. В корзине два белых и три чёрных шара. Вынимается наугад один шар. Найти вероятность того, что он белый.

Решение. Общее число элементарных исходов $n = 5$ (можно достать любой из пяти шаров). Среди этих исходов благоприятствуют событию (будет вынут белый шар) два исхода. Следовательно, искомая вероятность: $P(A) = 2/5$.

Пример 5.2. Бросают 2 игральные кости. Найти вероятности следующих событий: A — сумма числа очков не превосходит 5; B — произведение числа очков не превосходит 4; C — произведение числа очков делится на 8.

Решение. Используем классическую формулу вероятности. Определим общее число исходов: поскольку в случае подбрасывания одной кости имеем 6 исходов, то в случае подбрасывания двух костей имеем $n = 6 \cdot 6 = 36$ исходов. Найдём число благоприятных исходов.

Множество исходов, благоприятных событию A, состоит из 10 исходов:

$$\{(1,1),(1,2),(1,3),(1,4),(2,1),(2,2),(2,3),(3,1),(3,2),(4,1)\}$$

Соответственно, вероятность того, что сумма числа очков не превосходит 5: $P(A) = m_A/n = 10/36$.

Множество исходов, благоприятных событию B, состоит из 8 исходов:

$$\{(1,1),(1,2),(1,3),(1,4),(2,1),(2,2),(3,1),(4,1)\}.$$

Соответственно, вероятность того, что произведение числа очков не превосходит 4: $P(B) = m_B/n = 8/36$.

Множество исходов, благоприятных событию C, состоит из 5 исходов:

$$\{(2,4),(4,2),(4,4),(4,6),(6,4)\}.$$

Соответственно, вероятность того, что произведение числа очков делится на 8: $P(C) = m_C/n = 5/36$.

Для подсчёта числа исходов в более сложных случаях используют правила комбинаторики, необходимые сведения из которой мы приведём для полноты изложения.

Комбинаторные формулы

1. Правило умножения (основное правило комбинаторики)

Пусть имеется k различных множеств, первое из которых содержит n_1 элементов, 2-е — n_2 элементов, k-е — n_k элементов.

Будем составлять всевозможные новые множества, взяв по одному элементу из каждого данного. Очевидно, что каждое новое множество будет содержать k элементов. Оказывается, что общее число таких новообразованных множеств будет

$$N = n_1 \cdot n_2 \cdot n_3 \cdots n_k.$$

Доказывается методом полной математической индукции.

Пример 5.3. Опыт состоит в подбрасывании трёх игральных костей. Сколько различных равновозможных исходов содержит пространство Ω этого испытания?

Решение. Число исходов при подбрасывании каждой из костей равно 6. По правилу умножения общее число исходов при подбрасывании трёх костей $N = 6^3 = 6 \cdot 6 \cdot 6 = 216$. Множество Ω в этом случае можно описать так:

$$\Omega = \{(1,1,1),(1,1,2),\ldots,(6,6,6)\}.$$

Опыт классический, так как исходы равновозможны.

2. Размещения

Пусть из множества, содержащего n различных элементов, отбираются m элементов и размещаются в порядке их появления. Получающиеся таким образом упорядоченные комбинации называются *размещениями* из n элементов по m элементов, а полное число таких выборок определяется формулой

$$A_n^m = n(n-1)(n-2)\cdots(n-m+1) = \frac{n!}{(n-m)!}.$$

3. Перестановки

Пусть имеется n различных элементов, из которых формируются выборки, отличающиеся порядком элементов. Получающиеся таким образом упорядоченные комбинации называются *перестановками* из n элементов, а полное число таких выборок определяется формулой (частный случай размещений при $m = n$)

$$P_n = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1 = n!$$

4. Сочетания

Пусть из множества, состоящего из n элементов, отбираются m различных элементов без учета порядка их появления. Получающиеся таким образом комбинации называются сочетаниями из n элементов по m элементов, а полное число таких выборок определяется формулой

$$C_n^m = \frac{A_n^m}{P_m} = \frac{n(n-1)(n-2)\cdots(n-m+1)}{m!} = \frac{n!}{m!(n-m)!}.$$

Пример 5.4. В корзине находятся восемь белых и шесть чёрных шаров. Наудачу вынимаются пять шаров. Найти вероятности следующих событий: A — все вынутые шары белые; B — среди вынутых три белых и два чёрных шара; D — среди вынутых хотя бы три белых шара.

Решение. Опыт классический, так как исходы равновозможны. Общее число элементарных исходов n равно числу способов, которыми можно выбрать из 14 шаров по 5 различных шаров, т.е. $n = C_{14}^5$. Число благоприятных событию A исходов равно числу способов, которыми можно выбрать из 8 белых шаров по 5 различных шаров, т.е. $m_A = C_8^5$. Тогда

$$P(A) = \frac{m_A}{n} = \frac{C_8^5}{C_{14}^5} = \frac{8!}{5!3!}\frac{5!9!}{14!} = \frac{4}{143}.$$

При подсчёте благоприятных исходов для события B учтём, что 3 белых шара из 8 имеющихся можно получить C_8^3 способами. При этом для каждой выбранной комбинации из трёх белых шаров два шара из 6 имеющихся чёрных можно выбрать C_6^2 способами. Тогда, согласно основному правилу комбинаторики, $m_B = C_8^3 C_6^2$ и

$$P(B) = \frac{m_B}{n} = \frac{C_8^3 C_6^2}{C_{14}^5} = \frac{60}{143}.$$

Для события D благоприятны исходы, когда выборка содержит либо 3, либо 4, либо 5 белых шаров. Поэтому $m_D = C_8^3 C_6^2 + C_8^4 C_6^1 + C_8^5 C_6^0$ и

$$P(D) = \frac{m_D}{n} = \frac{C_8^3 C_6^2 + C_8^4 C_6^1 + C_8^5 C_6^0}{C_{14}^5} = \frac{94}{143}.$$

Пример 5.5. Числа $1, 2, \ldots, 9$ записываются в случайном порядке. Какова вероятность, что числа 1 и 2 будут рядом в порядке возрастания?

Решение. Общее число исходов n равно числу перестановок из 9 элементов, т.е. $n = P_9 = 9!$. Число 1 в этой последовательности из 9 чисел может занимать 8 различных позиций (с 1-ой по 8-ую), при этом число 2 должно занимать соседнюю позицию. При каждом фиксированном положении чисел 1 и 2, оставшиеся 7 позиций могут быть заняты оставшимися 7-ью числами, как очевидно, $7!$ способами (число перестановок из 7 элементов). Следовательно,

$$P(A) = \frac{8P_7}{P_9} = \frac{8 \cdot 7!}{9!} = \frac{1}{9}.$$

Пример 5.6. Наудачу взятый телефонный номер состоит из 5 цифр. Какова вероятность, что в нём все цифры различные?

Решение. Так как на каждом из пяти мест в пятизначном номере может стоять любая из цифр: $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$, то общее количество различных пятизначных номеров равно $n = 10^5$. Номера, у которых все цифры различные, — это размещения из 10 элементов по 5, то есть $m = A_{10}^5$. Следовательно,

$$P(A) = \frac{m}{n} = \frac{A_{10}^5}{10^5} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{10^5} = 0{,}3024.$$

Пример 5.7. В партии из N изделий имеется $M < N$ бракованных. Предположим, что вынимаются наугад n изделий. Определить вероятность того, что среди них будет хотя бы одно бракованное.

Решение. Исход опыта — n извлечённых наугад изделий из N. Исходы равновозможны, т.е. $|\Omega| = C_N^n$.

Событие A — хотя бы одно изделие бракованное. Нужно найти, сколько исходов входит в событие A.

Вместо события A удобнее рассмотреть противоположное событие $\overline{A}$ — среди n извлечённых изделий нет ни одного бракованного. Тогда $|\overline{A}| = C_{N-M}^n$ и

$$P(\overline{A}) = \frac{C_{N-M}^n}{C_N^n},$$

а по второму свойству вероятности

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

т.е.

$$P(A) = 1 - \frac{C_{N-M}^n}{C_N^n}.$$

Пример 5.8. Партия товара состоит k_1 изделий 1-го сорта, k_2 изделий 2-го сорта, k_3 изделий 3-го сорта и k_4 изделий 4-го сорта ($\sum\limits_{i=1}^{4} k_i = N$). Для контроля наудачу выбираются M изделий. Определить вероятность того, что среди них ровно m_1 изделий 1-го сорта, m_2 изделий 2-го сорта, m_3 изделий 3-го сорта и m_4 изделий 4-го сорта.

Решение. Общее число элементарных исходов n равно числу способов, которыми можно выбрать из имеющихся N изделий M различных изделий, т.е. $n = C_N^M$. m_1 изделий 1-го сорта можно выбрать из имеющихся k_1 изделий 1-го сорта $C_{k_1}^{m_1}$ способами, m_2 изделий 2-го сорта можно выбрать из k_2 имеющихся изделий 2-го сорта $C_{k_2}^{m_2}$ способами и т.д. Тогда общее число благоприятных исходов, согласно основному правилу комбинаторики, равно $m_A = C_{k_1}^{m_1} C_{k_2}^{m_2} C_{k_3}^{m_3} C_{k_4}^{m_4}$ и соответствующая вероятность

$$P(A) = \frac{m_A}{n} = \frac{C_{k_1}^{m_1} C_{k_2}^{m_2} C_{k_3}^{m_3} C_{k_4}^{m_4}}{C_N^M}. \tag{5.5}$$

6. Геометрическое определение вероятности

Классическое определение вероятности основано на рассмотрении конечного числа равновероятных элементарных событий. Но как вычислить вероятность события, если Ω имеет бесконечное множество элементов, как, например, в «задаче о встрече» (см. пример 1.3)?

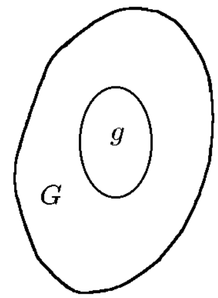


Рис. 10.

Существует формальное определение вероятности для опытов с бесконечным числом исходов. В подобных случаях пространство элементарных исходов может быть некоторой областью G (которую для определённости будем считать расположенной в $\mathbb{R}^2$, т.е. $G \subset \mathbb{R}^2$), и в ней содержится другая область g (рис. 10) с определённой мерой (длиной, площадью, объёмом). Под событием A можно понимать исходы, входящие в область g. Пусть, например, наугад брошена «точка» M. Какова вероятность того, что «точка» M попадёт в область g, являющуюся частью области G?

Хотя каждое из множеств G и g содержит бесконечное количество точек, естественно положить, что вероятность попадания в множество G больше и притом во столько раз, во сколько площадь S_G области G превышает площадь S_g области g. Считая все допустимые способы попадания равновозможными, естественно полагать, что искомая вероятность равна

$$P(A) = \frac{S_g}{S_G}.$$

В общем случае множества G и g могут иметь разную размерность ($\mathbb{R}^1$, $\mathbb{R}^3$ и др.), но приведённая формула сохраняет свой смысл с той лишь разницей, что множества в общем случае оцениваются мерой (длиной, площадью, объёмом).

Таким образом, в общем случае

$$P(A) = \frac{\operatorname{mes} g}{\operatorname{mes} G}. \tag{6.1}$$

$\Diamond$ Мера множества — обобщение понятия длины отрезка, площади фигуры, объёма фигуры или массы тела при некотором заданном пространственном (поверхностном, линейном) распределении плотности массы.

При этом предполагается, что g и G изображаются некоторыми множествами евклидова пространства $\mathbb{R}^k$, имеющими меру, и равновеликие множества считаются равновероятными.

Проверим выполнение аксиом.

1. $0 \leqslant P(A)$, так как длина, площадь или объём — положительные величины.

2. Аксиома 2 также справедлива, так как

$$P(\Omega) = \frac{\operatorname{mes} G}{\operatorname{mes} G} = 1.$$

3. Справедливость аксиомы 3 следует из свойств аддитивности длины, площади, объёма: мера суммы непересекающихся областей равна сумме их мер.

Таким образом, все аксиомы выполняются.

Пример 6.1. Для условий примера 1.3 («задача о встрече») найти вероятность события C — два студента встретятся.

Решение. Пространство Ω и событие C описаны в примере 1.3 (см. рис. 1). Применив геометрическое определение, получим

$$P(C) = \frac{\operatorname{mes}\{(x,y) \,|\, |x-y| \leqslant 1/6\}}{\operatorname{mes}\left\{(x,y) \,\middle|\, \begin{array}{l} 0 \leqslant x \leqslant 1 \\ 0 \leqslant y \leqslant 1 \end{array}\right\}} = \frac{1-(5/6)^2}{1^2} = \frac{11}{36}.$$

Пример 6.2. Случайным образом в интервале $(0,1)$ выбираются два числа: a и b. Найти вероятность следующих событий: A: $a+b<1$; B: $ab<1/2$.

Решение. Выберем декартову систему координат, и на оси Ox будем откладывать число a, а на оси Oy — число b. По условиям задачи $0<a<1$, $0<b<1$. Очевидно, что множеству элементарных исходов (область G) при таком подходе будет соответствовать квадрат со стороной, равной единице: $\operatorname{mes} G = S_G = 1$.

Областью g, благоприятствующей событию A, будет являться множество точек данного квадрата, для которых (заштрихованная на рис. 11 область) $\operatorname{mes} g = S_g = 1/2$. Следовательно,

$$P(A) = \frac{S_g}{S_G} = \frac{1/2}{1} = \frac{1}{2}.$$

Областью g, благоприятствующей событию B, будет являться множество точек квадрата, для которых $xy < 1/2$ (заштрихованная на рис. 12 область):

$$S_g = \frac{1}{2} + \int\limits_{1/2}^{1} \frac{1}{2x}dx = \frac{1}{2} + \frac{1}{2}\ln x\Big|_{1/2}^{1} = \frac{1}{2} + \frac{1}{2}\ln 2 \approx 0{,}85.$$

Следовательно,

$$P(B) = \frac{S_g}{S_G} = \frac{(1+\ln 2)/2}{1} \approx \frac{0{,}85}{1} = 0{,}85.$$

Пример 6.3. В отрезке единичной длины наудачу выбираются две точки. Определить вероятность того, что расстояние между точками меньше $1/2$.

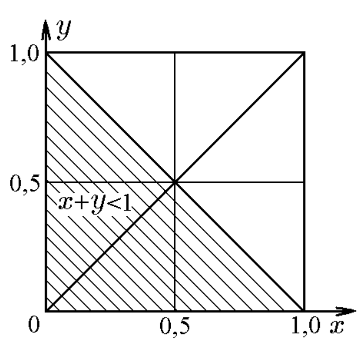


Рис. 11.

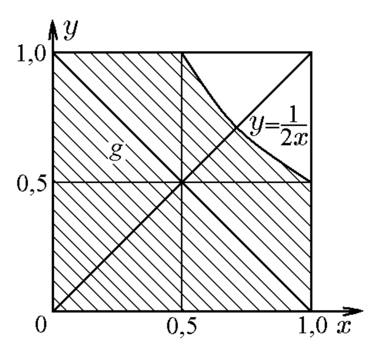


Рис. 12.

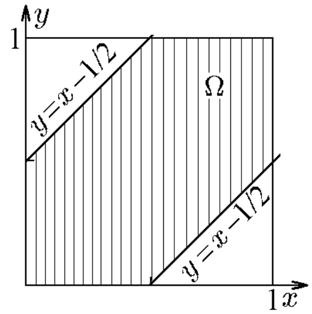


Рис. 13.

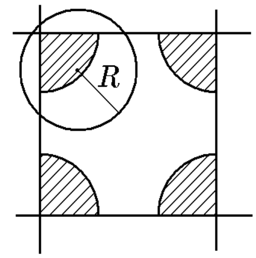


Рис. 14.

Решение. Пусть x — расстояние от конца отрезка до первой точки, а y — расстояние от того же конца отрезка до второй точки. Тогда пространство элементарных исходов: $\Omega = \{(x, y) \in \mathbb{R}^2 | 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1\}$. Рассматриваемому событию «расстояние между точками меньше $1/2$» соответствует множество A точек множества Ω, для которых $|x-y| < 1/2$. Тогда, согласно геометрическому определению вероятности, вероятность этого события (см. рис. 13)

$$P = \frac{S_A}{S_\Omega} = \frac{1 - 1/2 \cdot 1/2}{1} = \frac{3}{4}.$$

Пример 6.4. Плоскость разграфлена на квадраты сеткой параллельных линий с шагом 3 см. На плоскость брошена монета диаметром 2 см. Какова вероятность, что монета пересечёт четыре квадрата?

Решение. Будем характеризовать положение монеты положением её центра. Очевидно, что в качестве Ω можно взять множество точек квадрата, в которые попал центр монеты. Монета пересечёт четыре квадрата, если центр монеты будет находиться на расстоянии, меньшем радиуса монеты $R = 1$ см, от одной из четырех вершин квадрата (множество A соответствует заштрихованной фигуре на рис. 14). Тогда вероятность этого события

$$P = \frac{S_A}{S_\Omega} = \frac{4 \cdot 1/4 \cdot \pi R^2}{9} = \frac{\pi}{9}.$$

Проведём качественное сравнение сформулированных выше определений вероятности.

Каждое из определений вероятности возникло в результате решения конкретных прикладных задач и, следовательно, отражает закономерности явлений реального мира. К тому же каждое из них обладает своими положительными качествами, но и не лишено недостатков. Узость классического определения связана с выделением конечной полной группы равновероятных событий, что не всегда позволяет представить каждое событие в виде разложения по событиям этой группы.

Наоборот, геометрическое определение теряет смысл в случае конечного множества G: оно оказывается неприменимым, когда равновеликие множества не будут равновероятными. Статистическое определение предполагает повторение одного и того же опыта любое число раз, что практически можно считать неосуществимым, а сама вероятность к тому же не вычисляется – находится лишь некоторое приближение к ней.

Аксиоматическое построение теории вероятностей, данное академиком А.Н. Колмогоровым, объединило различные подходы к определению вероятности и в значительной степени способствовало развитию этой науки.

Однако ни аксиоматический, ни классический, ни геометрический, ни статистический подходы к определению вероятности не исчерпывают реальное понятие «вероятность», а являются лишь приближениями ко всё более полному его раскрытию. Предположение о том, что при данных условиях данному событию отвечает некоторая вероятность, является гипотезой, которая в каждой отдельной задаче требует проверки и обоснования.

7. Условная вероятность. Независимость событий

Рассмотрим пример. Предположим, что эксперимент состоит в трёхкратном подбрасывании монеты. Тогда пространство элементарных событий имеет следующую структуру:

$$\Omega = \{\text{ГГГ, ГГР, ГРГ, ГРР, РГГ, РГР, РРГ, РРР}\},$$

т.е. $|\Omega| = 8$. Исходы равновозможны, т.е. $P(\omega_i) = 1/8$ для всех $\omega_i \in \Omega$, $i = \overline{1, 8}$.

Пусть A – событие, состоящее в том, что герб появится ровно один раз. Событие A – подмножество Ω, $A = \{\text{ГРР, РГР, РРГ}\}$. Воспользовавшись классическим определением вероятности, найдём

$$P(A) = \frac{|A|}{|\Omega|} = \frac{3}{8}.$$

Предположим теперь, что об исходе опыта известно, что произошло событие B – число выпавших гербов нечётно:

$$B = \{\text{ГГГ, ГРР, РГР, РРГ}\}.$$

Как изменится вероятность события A при этой дополнительной информации? Или скажем так: чему равна вероятность события A при условии, что произошло событие B?

Число исходов, благоприятствующих событию B, равно 4, т.е. $|B| = 4$. Событию A благоприятствуют три исхода, входящих в событие B. Естественно новую вероятность события A при условии, что произошло событие B, принять

$$P(A/B) = \frac{3}{4}.$$

Рассмотрим теперь более общий пример.

Пусть имеется классическое пространство Ω:

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\};$$
$$|\Omega| = n, \quad P(\omega_i) = \frac{1}{n}, \quad \omega_i \in \Omega, \quad i = \overline{1, n},$$

и пусть A, B – некоторые события, причём $|A| = r$, $|B| = m$ и $|A \cdot B| = k$.

Условную вероятность $P(A/B)$ по аналогии с предыдущим можно вычислить так:

$$P(A/B) = \frac{k}{m} = \frac{k/n}{m/n} = \frac{P(A \cdot B)}{P(B)},$$

а

$$P(B/A) = \frac{k}{r} = \frac{k/n}{r/n} = \frac{P(A \cdot B)}{P(A)}.$$

♦ Вероятность события A, вычисленная при условии, что имело место событие B, называется *условной вероятностью*, обозначается $P(A/B)$ и определяется соотношениями

$$P(A/B) = \frac{P(A \cdot B)}{P(B)}, \tag{7.1}$$

$$P(B/A) = \frac{P(A \cdot B)}{P(A)}. \tag{7.2}$$

Из формул (7.1) и (7.2) следует теорема умножения вероятностей.

Теорема 7.1 (теорема умножения вероятностей). *Вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную при условии, что первое событие имело место:*

$$P(A \cdot B) = P(A)P(B/A) = P(B)P(A/B). \tag{7.3}$$

Утверждение теоремы 7.1 может быть обобщено на случай нескольких событий.

Теорема 7.2 (общая теорема умножения вероятностей). *Вероятность произведения нескольких событий равна произведению вероятностей этих событий, причём вероятность каждого следующего по порядку события вычисляется при условии, что все предыдущие имели место:*

$$P(A_1A_2 \cdots A_n) = P(A_1)P(A_2/A_1) \cdots P(A_n/A_1A_2 \cdots A_{n-1}). \tag{7.4}$$

При этом соответствующие условные вероятности должны быть определены.

Доказательство следует непосредственно из предыдущей теоремы. Соотношение (7.4) доказывается методом математической индукции.

Пример 7.1. В корзине находятся два белых и три чёрных шара. Вынимают два шара. Найти вероятность того, что оба шара белые.

Решение. Обозначим события: A — вынуты два белых шара, A_1 — первый вынутый шар белый, A_2 — второй вынутый шар белый. Тогда, в соответствии с классическим определением вероятности, вероятность вынуть первым белый шар $P(A_1) = 2/5$. После того, как из корзины был извлечён белый шар, в корзине осталось 4 шара, среди которых один белый. Следовательно, вероятность вынуть вторым белый шар, при условии, что первым был вынут белый шар, $P(A_2/A_1) = 1/4$. Следовательно,

$$P(A) = P(A_1)P(A_2/A_1) = \frac{2}{5}\frac{1}{4} = 0{,}1.$$

♢ Условная вероятность обладает всеми свойствами вероятности, поскольку её логично рассматривать как обычную вероятность, определённую на другом пространстве элементарных событий (роль достоверного события Ω играет событие B). Проверим выполнение аксиом теории вероятности для условной вероятности:

1) $P(A/B) \geqslant 0$ для всех A, т.е. аксиома 1 выполняется;

2) если $A = B$, то

$$P(B/B) \equiv \frac{P(B \cdot B)}{P(B)} = \frac{P(B)}{P(B)} = 1,$$

т.е. аксиома 2 также выполняется;

3) пусть A_1 и A_2 — такие события, что $A_1 \cdot A_2 = \emptyset$ и $P(B) > 0$. Найдём

$$P((A_1 + A_2)/B) \stackrel{\text{def}}{=} \frac{P((A_1 + A_2)B)}{P(B)} = \frac{P(A_1 B + A_2 B)}{P(B)} =$$
$$= \frac{P(A_1 B)}{P(B)} + \frac{P(A_2 B)}{P(B)} \stackrel{\text{def}}{=} P(A_1/B) + P(A_2/B),$$

т.е.

$$P((A_1 + A_2)/B) = P(A_1/B) + P(A_2/B),$$

и аксиома 3 выполняется.

Пример 7.2. Среди 25 экзаменационных билетов имеется 5 «счастливых» и 20 «несчастливых». Студенты подходят за билетами один за другим по очереди. У кого больше вероятность вытащить счастливый билет: у того, кто подошел за билетом первым, или у того, кто подошел вторым?

Решение. Опишем пространство элементарных событий Ω:

$$\Omega = \{1\text{С}2\text{Н},\ 1\text{Н}2\text{С},\ 1\text{С}2\text{С},\ 1\text{Н}2\text{Н}\} = \{\omega_1, \omega_2, \omega_3, \omega_4\}$$

(например, элемент 1С2Н будем понимать так: 1-й студент получил счастливый билет, а 2-й — несчастливый). Воспользовавшись «теоремой умножения вероятностей» (7.3), введём вероятности на пространстве Ω:

$$P(\omega_1) = P(1\text{С}2\text{Н}) = P(1\text{С})P(2\text{Н}/1\text{С}) = \frac{5}{25}\frac{20}{24};$$
$$P(\omega_2) = P(1\text{Н}2\text{С}) = P(1\text{Н})P(2\text{С}/1\text{Н}) = \frac{20}{25}\frac{5}{24};$$
$$P(\omega_2) = P(1\text{С}2\text{С}) = P(1\text{С})P(2\text{С}/1\text{С}) = \frac{5}{25}\frac{4}{24};$$
$$P(\omega_2) = P(1\text{Н}2\text{Н}) = P(1\text{Н})P(2\text{Н}/1\text{Н}) = \frac{20}{25}\frac{19}{24}.$$

Проверим:

$$P(\Omega) = \sum_{i=1}^{4} P(\omega_i) = \frac{1}{25 \cdot 24}(5 \cdot 20 + 20 \cdot 5 + 5 \cdot 4 + 20 \cdot 19) = 1.$$

Событие A — 1-й студент получает счастливый билет:

$$A = \{1\text{С}2\text{Н},\ 1\text{С}2\text{С}\} = \{\omega_1, \omega_3\}.$$

Согласно аксиоме 3,

$$P(A) = P(\omega_1) + P(\omega_3) = \frac{20 \cdot 5}{25 \cdot 24} + \frac{5 \cdot 4}{25 \cdot 24} = \frac{1}{5}.$$

Событие B — 2-й студент получает счастливый билет:

$$B = \{1\text{Н}2\text{С},\ 1\text{С}2\text{С}\} = \{\omega_2, \omega_3\}.$$

Согласно аксиоме 3,

$$P(B) = P(\omega_2) + P(\omega_3) = \frac{20 \cdot 5}{25 \cdot 24} + \frac{5 \cdot 4}{25 \cdot 24} = \frac{1}{5}.$$

Итак, $P(A) = P(B) = 1/5$.

♦ События A и B называются *независимыми*, если вероятность каждого из них не зависит от наступления или ненаступления другого:

$$P(A/B) = P(A) \quad \text{и} \quad P(B/A) = P(B). \tag{7.5}$$

Теорема 7.3. *События A и B независимы тогда и только тогда, когда вероятность их одновременного появления равна произведению вероятностей этих событий*

$$P(AB) = P(A)P(B). \tag{7.6}$$

Доказательство следует непосредственно из теоремы умножения вероятностей (7.3) и определения (7.5).

Теорема 7.4. *Если события A и B независимы, то события A и $\overline{B}$, события $\overline{A}$ и B, события $\overline{A}$ и $\overline{B}$ также независимы.*

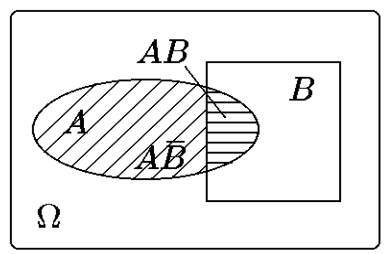


Рис. 15.

Доказательство. Докажем первое утверждение теоремы.

Так как события A и B независимы, то, согласно предыдущей теореме, $P(AB) = P(A)P(B)$. Поскольку (см. рис. 15)

$$AB + A\overline{B} = A, \quad (AB)(A\overline{B}) = \emptyset,$$

то, согласно аксиоме 3, запишем

$$P(AB) + P(A\overline{B}) = P(A),$$

т.е.

$$P(A\overline{B}) = P(A) - P(AB).$$

Воспользовавшись для $P(AB)$ соотношением (7.6), получим

$$P(A\overline{B}) = P(A) - P(A)P(B).$$

Отсюда следует

$$P(A\overline{B}) = P(A) - P(A)[1 - P(\overline{B})]$$

или

$$P(A\overline{B}) = P(A)P(\overline{B}),$$

что и требовалось доказать.

Второе и третье утверждения теоремы доказываются аналогично.

Пример 7.3. Испытание заключается в подбрасывании игральной кости. Событие A заключается в том, что выпало два очка; B — выпало чётное число очков. Найти $P(A)$, $P(B)$, $P(A/B)$, $P(B/A)$ и выяснить, зависимы или нет события A и B.

Решение. В примере рассматривается классическая вероятность. По формуле (5.4) находим $P(A) = 1/6$, $P(B) = 1/2$. Если событие B наступило, то выпало либо два, либо четыре, либо шесть очков. Следовательно, $P(A/B) = 1/3$. Если событие A наступило, то это событие влечёт событие B, следовательно, $P(B/A) = 1$. Так как $P(A/B) \neq P(A)$ или $P(B/A) \neq P(B)$, то события A и B зависимы.

♦ События $A_1, A_2, \ldots, A_n$ называются независимыми в совокупности, если для любых k из них ($k \leqslant n$) выполняется соотношение

$$P(A_{i_1}A_{i_2}\cdots A_{i_k}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k}) = \prod_{j=1}^{k} P(A_{i_j}). \qquad (7.7)$$

♦ Если соотношение (7.7) выполняется только при $k = 2$, то события называются *попарно независимыми*.

Пример 7.4. Два стрелка независимо друг от друга стреляют по мишени. Вероятность попадания в цель для первого стрелка равна 0,7, для второго 0,8. Определить вероятности следующих событий: а) ровно один стрелок попадёт в цель; б) хотя бы один из стрелков попадёт в цель.

Решение. Пусть событие A — попал в цель первый стрелок, событие B — попал второй стрелок.

а) Событие C — ровно один стрелок попадёт в цель — есть событие $A\bar{B} + \bar{A}B$. Так как события $A\bar{B}$ и $\bar{A}B$ несовместны, а события A и B независимы, то $P(C) = P(A)P(\bar{B}) + P(\bar{A})P(B) = 0{,}7 \cdot 0{,}2 + 0{,}8 \cdot 0{,}3 = 0{,}38$.

б) Событие D — попал в цель хотя бы один стрелок — можно представить в виде суммы событий A и B: $D = A + B$. Следовательно, $P(D) = P(A) + P(B) - P(AB)$. Так как события A и B независимы, то $P(D) = P(A) + P(B) - P(A)P(B) = 0{,}7 + 0{,}8 - 0{,}56 = 0{,}94$.

Как мы уже отмечали ранее, когда речь идёт о появлении хотя бы одного события в серии испытаний, приведённые выше прямые вычисления $P(D)$ являются не самыми удобными, особенно если число испытаний велико. Лучше в такой ситуации перейти к противоположному событию, определив $\bar{D} = \bar{A}\bar{B}$; тогда с учетом (4.2) запишем

$$P(D) = 1 - P(\bar{D}) = 1 - P(\bar{A}\bar{B}) = 1 - P(\bar{A})P(\bar{B}) = 1 - 0{,}3 \cdot 0{,}2 = 0{,}94.$$

Пример 7.5. В двух партиях товара, соответственно, 86% и 32% доброкачественных изделий. Наудачу выбирают по одному изделию из каждой партии. Какова вероятность обнаружить среди них: а) два бракованных; б) одно доброкачественное и одно бракованное; в) хотя бы одно бракованное?

Решение. В соответствии с условием задачи, вероятности выбрать доброкачественное изделие из каждой партии равны соответственно $p_1 = 0{,}86$ и $p_2 = 0{,}32$. Воспользовавшись теоремами сложения и умножения вероятностей, найдём
а) вероятность обнаружить два бракованных:

$$P = (1 - p_1)(1 - p_2) = 0{,}14 \cdot 0{,}68 = 0{,}0952;$$

б) вероятность обнаружить одно доброкачественное и одно бракованное:

$$P = p_1(1 - p_2) + p_2(1 - p_1) = 0{,}86 \cdot 0{,}68 + 0{,}14 \cdot 0{,}32 = 0{,}6296;$$

в) вероятность обнаружить хотя бы одно бракованное:

$$P = 1 - p_1p_2 = 1 - 0{,}86 \cdot 0{,}32 = 0{,}7248.$$

Пример 7.6. Монету бросают 10 раз. Найти вероятность того, что хотя бы один раз появится орёл.

Решение. Событие «в десяти испытаниях хотя бы один раз появится орёл» является противоположным событию «в десяти испытаниях ни разу не появится орёл». Так как вероятность того, что орёл не появится в одном испытании, $q = p = 1/2$, то вероятность того, что хотя бы один раз появится орёл в 10 испытаниях

$$P = 1 - q^{10} = 1 - \left(\frac{1}{2}\right)^{10} = \frac{1023}{1024}.$$

Пример 7.7. На отрезке $[0; 1]$ выбираются случайным образом 5 точек с координатами $x_1, x_2, \ldots, x_5$. Какова вероятность того, что минимальная из координат точек больше $1/5$?

Решение. Рассмотрим событие $\{\min(x_1, x_2, \ldots, x_5) > 1/5\}$. Данное событие эквивалентно событию $\{$координата каждой точки больше $1/5\}$. Вероятность этого события, согласно теореме умножения,

$$P(\min(x_1, x_2, \ldots, x_5) > 1/5) = P(x_1 > 1/5, x_2 > 1/5, \ldots, x_5 > 1/5) = \\ = P^5(x_1 > 1/5) = (4/5)^5.$$

Пример 7.8. Урна содержит 6 шаров с номерами от 1 до 6. Шары извлекаются по одному без возвращения. Найти вероятность того, что хотя бы у одного шара совпадут номер шара и порядковый номер его извлечения. Найти предельное значение вероятности, если число шаров в урне стремится к бесконечности.

Решение. Обозначим через A_k, $k = \overline{1, 6}$, событие «шар с номером k извлечён k-ым по порядку», а через A событие «хотя бы у одного шара совпал номер с порядковым номером извлечения». На первый взгляд кажется, что вероятность события проще найти, перейдя к противоположному событию, как это обычно делается, когда речь идёт о появления хотя бы одного события. Однако противоположное событие — «ни у одного из шаров не совпал номер с порядковым номером извлечения» — оказывается связанным с еще более трудоемкими расчётами вероятности, чем прямое. Поэтому найдём вероятность данного события по определению как вероятность суммы событий:

$$A_1 + A_2 + \ldots + A_6.$$

Поскольку события A_k, $k = \overline{1, 6}$, совместны, то по общему правилу сложения вероятностей

$$P(A) = \sum_{i=1}^{6} P(A_i) - \sum_{i \leqslant i < j \leqslant 6} P(A_iA_j) + \sum_{i \leqslant i < j < k \leqslant 6} P(A_iA_jA_k) - \ldots - P(A_1A_2A_3A_4A_5A_6).$$

Найдём $\sum\limits_{i=1}^{6} P(A_i)$. Очевидно, что $P(A_1) = 1/6$. Событие A_2 произойдёт, если первый вынутый шар не будет шаром с номером 2, а второй вынутый шар будет являться шаром с номером 2, следовательно, его вероятность по теореме умножения вероятностей: $P(A_2) = 5/6 \cdot 1/5 = 1/6$. Аналогично можно получить для всех остальных событий: $P(A_3) = \ldots = P(A_6) = 1/6$. Тогда

$$\sum_{i=1}^{6} P(A_i) = 6 \cdot \frac{1}{6} = 1.$$

Далее можно заметить, что вероятности всех произведений событий попарно, всех произведений событий по три и т.д. также равны между собой. То есть достаточно найти вероятность одного из таких произведений и определить общее количество слагаемых в каждой сумме. В результате получим:

$$\sum_{i\leqslant i<j\leqslant 6} P(A_iA_j) = C_6^2 P(A_iA_j) = \frac{6\cdot 5}{2!}\cdot\frac{1}{6}\cdot\frac{1}{5} = \frac{1}{2!};$$

$$\sum_{i\leqslant i<j<k\leqslant 6} P(A_iA_jA_k) = C_6^3 P(A_iA_jA_k) = \frac{6\cdot 5\cdot 4}{3!}\cdot\frac{1}{6}\cdot\frac{1}{5}\cdot\frac{1}{4} = \frac{1}{3!};$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots;$$

$$P(\underbrace{A_iA_j\cdots A_k}_{6\ \text{множителей}}) = \frac{1}{6}\cdot\frac{1}{5}\cdot\frac{1}{4}\cdot\frac{1}{3}\cdot\frac{1}{2} = \frac{1}{6!}.$$

Окончательно получим

$$P(A) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \frac{1}{5!} - \frac{1}{6!}.$$

Проведя подобные вычисления при условии, что число шаров в урне стремится к бесконечности, в пределе получим

$$P(A) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \frac{1}{5!} - \frac{1}{6!} + \ldots = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} = 1 - \frac{1}{e}.$$

Пример 7.9. Два стрелка по очереди стреляют по мишени. Вероятность попадания для первого при каждом выстреле $1/3$, для второго — $2/3$. Выигрывает тот, кто первым попадёт в мишень. Какова вероятность, что первый выиграет не позднее своего 3-го выстрела? Каковы вероятности выигрыша для каждого при сколь угодно длительной стрельбе?

Решение. Введём обозначения: A_k — первый попал в мишень при k-ом выстреле, B_k — второй попал при k-ом выстреле. По условию $P(A_k) = p_1 = 1/3$, т.е. $q_1 = 2/3$; $P(B_k) = p_2 = 2/3$. Событие A — «первый выиграл не позднее 3-го выстрела» — через A_k и B_k выразится следующим образом: $A = A_1 + \bar{A}_1\bar{B}_1A_2 + \bar{A}_1\bar{B}_1\bar{A}_2\bar{B}_2A_3$. Тогда вероятность этого события

$$P(A) = p_1 + q_1q_2p_1 + q_1^2q_2^2p_1 = 1/3 + 2(1/3)^3 + 4(1/3)^5 \approx 0{,}424.$$

При сколь угодно длительной стрельбе вероятность выигрыша для первого:

$$P(A) = p_1 + q_1q_2p_1 + q_1^2q_2^2p_1 + \ldots = \sum_{k=0}^{\infty} p_1(q_1q_2)^k = \frac{p_1}{1-q_1q_2} = \frac{3}{7} \approx 0{,}429.$$

◊ Из попарной независимости событий не следует независимость группы событий в совокупности.

Проиллюстрируем это утверждение следующими примерами.

Пример 7.10 (Бернштейна). Пусть на плоскость бросают правильный тетраэдр, три грани которого окрашены соответственно в красный (К), синий (С) и зелёный (З), а на четвертую нанесены все три цвета. События К, С, З — тетраэдр упал на плоскость гранью, окраска которой содержит красный, синий, зелёный цвет соответственно. Проверить, являются ли события К, С, З попарно независимыми и независимыми в совокупности.

Решение. Красный (соответственно синий, зелёный) цвет присутствует на двух гранях из четырех. Поэтому

$$P(\text{К}) = P(\text{С}) = P(\text{З}) = \frac{2}{4} = \frac{1}{2}.$$

Три, а значит, и два цвета присутствуют только на одной грани из четырех. Следовательно,

$$P(\text{КС}) = \frac{1}{4}, \quad P(\text{КЗ}) = \frac{1}{4}, \quad P(\text{СЗ}) = \frac{1}{4},$$

т.е. события К, С, З попарно независимы, но

$$P(\text{КСЗ}) = \frac{1}{4} \neq P(\text{К})P(\text{С})P(\text{З}) = \frac{1}{8},$$

т.е. события К, С, З не являются независимыми в совокупности.

Пример 7.11. Точка с координатой ξ выбирается наудачу на отрезке $[0, 3]$, и независимо от нее точка с координатой η выбирается наудачу на отрезке $[0, 2]$. Проверить, являются ли три события: $A = \{\xi + \eta < 2\}$, $B = \{1 < \xi < 5/2\}$ и $C = \{\eta < 1\}$, независимыми в совокупности.

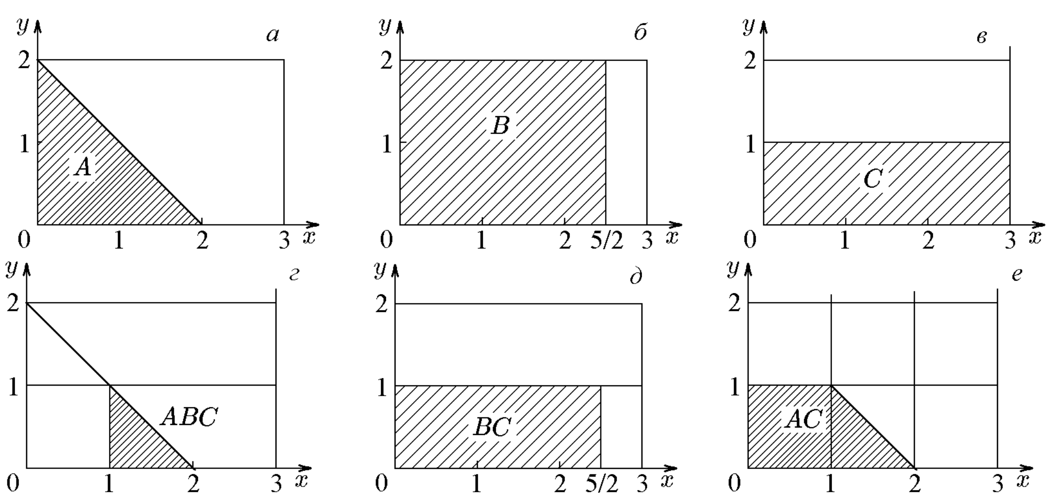


Рис. 16.

Решение. Пространство элементарных исходов Ω есть множество точек плоскости (ξ, η), для которых $0 \leqslant \xi \leqslant 3$, $0 \leqslant \eta \leqslant 2$ и $S_\Omega = 6$. Событию $A = \{\xi+\eta < 2\}$ отвечает множество точек, ограниченных прямыми $x = 0$, $y = 0$, $y = 2 - x$ (рис.

16,*а*): $S_A = 2$ и $P(A) = 1/3$. Событию $B = \{1 < \xi < 5/2\}$ отвечает область, заштрихованная на рис. 16,*б*: $S_B = 3$ и $P(B) = 1/2$. Событию $C = \{\eta < 1\}$ отвечает область, заштрихованная на рис. 16,*в*: $S_C = 3$ и $P(C) = 1/2$. События независимы в совокупности, если для любого сочетания из них вероятность произведения событий равна произведению вероятностей. Найдём $P(ABC)$. Из рис. 16,*г* видно, что $S_{ABC} = 1/2$ и $P(ABC) = 1/12 = P(A)P(B)P(C)$. Однако из этого равенства еще не следует независимость событий. Необходимо проверить все вероятности появлений всех пар событий. Находим $P(AB) = P(A) = 1/3$ (см. рис. 16,*а*), $P(BC) = 5/12$ (рис. 16,*д*), $P(AC) = 1/4$ (рис. 16,*е*). Видим, что $P(AB) \neq P(A)P(B)$, $P(AC) \neq P(A)P(C)$; следовательно, события в совокупности зависимы. Заметим, что $P(BC) = P(B)P(C)$, то есть события B и C попарно независимы.

Пример 7.12. Система S состоит из двух независимых подсистем S_a и S_b (см. рис. 17). Неисправность хотя бы одной подсистемы ведет к неисправности всей системы (подсистемы соединены последовательно). Каждая подсистема состоит из двух независимых дублирующих блоков a_k и b_k $(k = 1, 2)$ (схема параллельного соединения блоков в подсистемах).

Найти надежность системы — вероятность того, что система будет исправна в течение некоторого времени, если известно, что надежности блоков a_k и b_k равны соответственно 0,8 и 0,9.

Решение. Введём обозначения: A_k — событие, состоящее в том, что блок a_k исправен; B_k — событие, состоящее в том, что блок b_k исправен; S — событие, состоящее в том, что система исправна; S_a — событие, состоящее в том, что подсистема a исправна; S_b — событие, состоящее в том, что подсистема b исправна. Разобьем систему на две подсистемы a и b. Подсистема a дублирующих блоков исправна в том случае, когда исправен хотя бы один из блоков a_k $(k = 1, 2)$, то есть $S_a = A_1 + A_2$ — сумма двух совместных независимых событий. Следовательно $P(S_a) = P(A_1) + P(A_2) - P(A_1)P(A_2) = 0{,}8 + 0{,}8 - 0{,}64 = 0{,}96$. Аналогично $P(S_b) = P(B_1) + P(B_2) - P(B_1P(B_2) = 0{,}9 + 0{,}9 - 0{,}81 = 0{,}99$. Для исправности системы необходима исправность подсистем a и b, то есть $S = S_aS_b$. Таким образом, $P(S) = P(S_a)P(S_b) = 0{,}9504$.

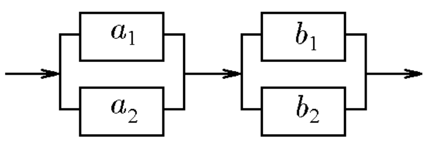


Рис. 17.

8. Формула полной вероятности. Формулы Байеса

В тех случаях, когда вероятность события A найти трудно, но легко вычисляются (или даны по условию) вероятности этого события относительно полной группы попарно несовместных событий, применяют формулу полной вероятности. Эта формула является основным средством при подсчёте вероятности сложных событий с использованием условных вероятностей.

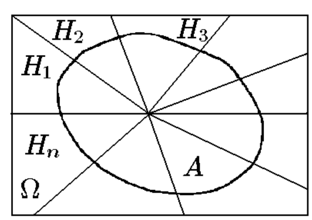


Рис. 18.

Теорема 8.1. *Пусть* $\Omega = H_1 + H_2 + \ldots + H_n$, *причём* $H_iH_j = \emptyset$ *при* $i \neq j$ *и* A *— произвольное событие,* $A \in \Omega$ (*см. рис. 18*), *тогда имеет место следующая формула:*

$$P(A) = \sum_{k=1}^{n} P(H_k)P(A/H_k). \tag{8.1}$$

Соотношение (8.1) *называется формулой полной вероятности.*

Доказательство. С учётом свойств операций над событиями

$$A = A(H_1 + H_2 + \ldots + H_n) = AH_1 + AH_2 + \ldots + AH_n.$$

Согласно аксиоме 3 теории вероятностей,

$$P(A) = P(AH_1) + P(AH_2) + \ldots + P(AH_n).$$

Воспользовавшись теоремой умножения, получим

$$P(A) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) + \ldots + P(H_n)P(A/H_n)$$

или

$$P(A) = \sum_{k=1}^{n} P(H_k)P(A/H_k),$$

что и требовалось доказать.

Следствие 8.1.1. Пусть выполняются условия предыдущей теоремы и $P(A) > 0$. Тогда

$$P(H_k/A) = \frac{P(H_k)P(A/H_k)}{\sum\limits_{j=1}^{n} P(H_j)P(A/H_j)}. \tag{8.2}$$

Соотношения (8.2) называются *формулами Байеса.*

Доказательство. По определению

$$P(H_k/A) = \frac{P(H_kA)}{P(A)}. \tag{8.3}$$

По теореме умножения

$$P(H_kA) = P(H_k)P(A/H_k), \tag{8.4}$$

а по формуле полной вероятности

$$P(A) = \sum_{j=1}^{n} P(H_j)P(A/H_j). \tag{8.5}$$

Подставив (8.4) и (8.5) в (8.3), получим соотношение (8.2).

$\Diamond$ События $H_1, H_2, \ldots, H_n$ часто называют *гипотезами*, так как заранее не известно, с каким из них произойдёт событие A. Формулы Байеса (формулы гипотез) позволяют переоценить вероятности событий (гипотез) $H_1, H_2, \ldots, H_n$ в связи с появлением события A, иными словами — вычислить апостериорные вероятности по априорным.

Пример 8.1. На заводе, изготавливающем микросхемы, 1-я линия производит 25%, 2-я — 35%, а 3-я — 40% всех изделий. Брак в их продукции составляет соответственно 5, 4 и 2%. Каковы вероятности того, что случайно выбранная микросхема окажется дефектной? Какова вероятность того, что случайно выбранная микросхема произведена на первой, второй и третьей линии, если она оказалась дефектной?

Решение. Обозначим через A событие, состоящее в том, что случайно выбранная микросхема дефектна, а через H_1, H_2, H_3 — события, состоящие в том, что эта микросхема произведёна на первой, второй и третьей линиях (см. рис. 19).

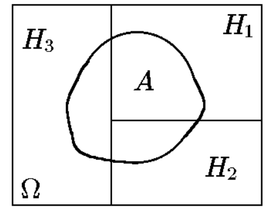


Рис. 19.

Ясно, что
1. $H_1 + H_2 + H_3 = \Omega$;
2. $H_i H_j = \emptyset$ при $i \neq j$;
3. $A \subset \Omega$.

Следовательно, можно воспользоваться формулой полной вероятности

$$P(A) = \sum_{k=1}^{3} P(H_k) P(A/H_k).$$

По условию

$$\begin{aligned} P(H_1) &= 0{,}25 \quad \text{и} \quad P(A/H_1) = 0{,}05; \\ P(H_2) &= 0{,}35 \quad \text{и} \quad P(A/H_2) = 0{,}04; \\ P(H_3) &= 0{,}40 \quad \text{и} \quad P(A/H_3) = 0{,}02. \end{aligned}$$

Тогда вероятность того, что выбран дефектный болт, равна

$$P(A) = 0{,}25 \cdot 0{,}05 + 0{,}35 \cdot 0{,}04 + 0{,}40 \cdot 0{,}02 = 0{,}0345.$$

По формулам Байеса (8.2) имеем

$$\begin{aligned} P(H_1/A) &= \frac{0{,}25 \cdot 0{,}05}{0{,}0345} = 0{,}36232; \\ P(H_2/A) &= \frac{0{,}35 \cdot 0{,}04}{0{,}0345} = 0{,}40580; \\ P(H_3/A) &= \frac{0{,}40 \cdot 0{,}02}{0{,}0345} = 0{,}23188. \end{aligned}$$

$\Diamond$ Сравните эти вероятности с вероятностями соответствующего события, если не известно, что произошло событие A (т.е. до опыта).

Пример 8.2. В корзине находится один шар — с равной вероятностью белый или чёрный. В корзину опустили белый шар, и после перемешивания извлекли один шар. Он оказался белым. Какова вероятность, что в корзине остался белый шар?

Решение. Пусть гипотеза H_1 состоит в том, что в корзине исходно находится белый шар, гипотеза H_2 — в корзине находится чёрный шар. Так как с равной вероятностью в корзине может находиться как белый, так и чёрный шар, то $P(H_1) = P(H_2) = 1/2$. После того как в корзину был опущен белый шар, вероятность вынуть белый шар (событие A) в предположении справедливости гипотезы H_1 есть $P(A/H_1) = 1$. Аналогично вероятность вынуть белый шар в предположении гипотезы H_2: $P(A/H_2) = 1/2$. Следовательно, по формуле полной вероятности

$$P(A) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) = \frac{1}{2} \cdot 1 + \frac{1}{2}\frac{1}{2} = \frac{3}{4}.$$

Тогда вероятность, что в корзине остался белый шар (то есть верна гипотеза H_1)

$$P(H_1/A) = \frac{P(H_1)P(A/H_1)}{P(A)} = \frac{1/2}{3/4} = \frac{2}{3}.$$

Пример 8.3. В магазин поступают однотипные изделия с трёх заводов, причём первый завод поставляет 50% изделий, второй — 30%, а третий — 20% изделий. Среди изделий 1-го завода первосортных 70%, второго — 80%, третьего — 90%. Куплено одно изделие. Оно оказалось первосортным. Определить вероятность того, что купленное изделие выпущено первым заводом.

Решение. Воспользуемся формулами полной вероятности и Байеса. Введём гипотезы H_1, H_2, H_3 — купленное изделие изготовлено соответственно первым, вторым и третьим заводами. По условию $P(H_1) = 0{,}5$; $P(H_2) = 0{,}3$; $P(H_2) = 0{,}2$. Условные вероятности события A — изделие первосортное — равны, соответственно, $P(A/H_1) = 0{,}7$; $P(A/H_2) = 0{,}8$; $P(A/H_3) = 0{,}9$. Тогда по формуле полной вероятности вероятность того, что купленное изделие окажется первосортным, определится соотношением

$$P(A) = \sum_{k=1}^{3} P(H_k)P(A/H_k) = 0{,}5 \cdot 0{,}7 + 0{,}3 \cdot 0{,}8 + 0{,}2 \cdot 0{,}9 = 0{,}77.$$

Вероятность того, что купленное изделие выпущено первым заводом, если оно оказалось первосортным, найдём по формуле Байеса:

$$P(H_1/A) = \frac{P(H_1)P(A/H_1)}{P(A)} = \frac{0{,}5 \cdot 0{,}7}{0{,}77} \approx 0{,}455.$$

Пример 8.4. Прибор состоит из двух независимых узлов a и b, соединенных последовательно (см. рис. 20). При этом неисправность хотя бы одного узла ведет к неисправности прибора. Прибор в результате испытания вышел из строя (отказал). Найти вероятность того, что отказал только узел a, если известно, что надежности узлов a и b равны p_1 и p_2 соответственно.

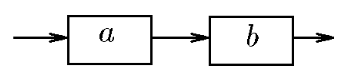


Рис. 20.

Решение. Введём обозначения: A — событие, состоящее в том, что узел a исправен; B — событие, состоящее в том, что узел b исправен; S — событие, состоящее в том, что прибор исправен. Введём гипотезы, характеризующие возможные совокупные состояния узлов: $H_1 = \bar{A}B$ — отказал только узел a; $H_2 = A\bar{B}$ — отказал только узел b; $H_3 = \bar{A}\bar{B}$ — отказали оба узла; $H_4 = AB$ — оба узла исправны. Имеем $P(H_1) = (1-p_1)p_2$; $P(H_2) = p_1(1-p_2)$; $P(H_3) = (1-p_1)(1-p_2)$; $P(H_4) = p_1p_2$. При этом $P(\bar{S}/H_1) = 1$; $P(\bar{S}/H_2) = 1$; $P(\bar{S}/H_3) = 1$; $P(\bar{S}/H_4) = 0$. Тогда, согласно (8.2), находим

$$P(H_1/\bar{S}) = \frac{P(H_1)P(\bar{S}/H_1)}{\sum\limits_{k=1}^{4} P(H_k)P(\bar{S}/H_k)} =$$

$$= \frac{(1-p_1)p_2 \cdot 1}{(1-p_1)p_2 \cdot 1 + p_1(1-p_2) \cdot 1 + (1-p_1)(1-p_2) \cdot 1 + p_1p_2 \cdot 0} = \frac{(1-p_1)p_2}{1-p_1p_2}.$$

Пример 8.5. Два стрелка стреляют по мишени, делая по одному выстрелу. Вероятность попадания для первого стрелка 0,8, для второго 0,4. После стрельбы в мишени обнаружена только одна пробоина. Воспользовавшись формулой полной вероятности, найти вероятность того, что попал первый стрелок.

Решение. Введём обозначения событий: C — попал в цель только один стрелок, A — первый стрелок попал в цель, B — второй стрелок попал в цель. Тогда $C =$

$A\bar{B} + \bar{A}B$. То есть, можно считать, что событие может наступить в результате осуществления двух гипотез: $H_1 = A\bar{B}$ — попал в цель только первый стрелок, $H_2 = \bar{A}B$ — попал в цель только второй стрелок. Имеем $P(H_1) = 0{,}8 \cdot 0{,}6 = 0{,}48$; $P(H_2) = 0{,}4 \cdot 0{,}2 = 0{,}08$; $P(C/H_1) = 1$; $P(C/H_2) = 1$; $P(C) = 0{,}48 + 0{,}08 = 0{,}56$. Следовательно,

$$P(H_1/C) = \frac{P(H_1)P(C/H_1)}{P(C)} = \frac{0{,}48 \cdot 1}{0{,}56} = \frac{6}{7}.$$

Пример 8.6. В первой урне 10 белых и 8 чёрных шаров, во второй 5 белых и 9 чёрных. Из первой урны во вторую переложены 6 шаров, затем из второй извлечён один шар. Определить вероятность того, что выбранный из второй урны шар — белый.

Решение. Очевидно, что вероятность вынуть белый шар из второй урны (событие A) зависит от того, какие шары были переложены. Поэтому естественным казалось бы ввести гипотезы, связанные с тем, какие шары были переложены из первой урны во вторую. Таких гипотез будет, очевидно, семь. Однако проще рассмотреть систему гипотез, связанную с тем, из какой изначально урны был вытащенный шар. Обозначим через H_1 гипотезу — вытащенный шар был из первой урны, $P(H_1) = 6/20$, $P(A/H_1) = 10/18$; через H_2 — вытащенный шар был из второй урны, $P(H_2) = 14/20$, $P(A/H_2) = 5/14$. Тогда

$$P(A) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) = \frac{5}{12}.$$

9. Схема Бернулли

В теории вероятностей большое значение имеет простая схема случайных опытов, называемая схемой Бернулли.

9.1. Формула Бернулли

♦ *Испытаниями Бернулли* (*схемой Бернулли*) называются независимые испытания с двумя исходами и вероятностью «успеха», не меняющейся от испытания к испытанию.

Рассмотрим два примера.

1. Пусть n раз бросают монету. Нас интересует появление герба — «успех», тогда появление решки — «неудача». Вероятность «успеха» равна 0,5 в каждом из n испытаний.

Пространство Ω можно описать таким образом:

$$\Omega = \left\{\underbrace{00\ldots0}_{n}; \underbrace{10\ldots0}_{n}; \underbrace{01\ldots0}_{n}; \ldots; \underbrace{00\ldots1}_{n}; \underbrace{110\ldots0}_{n}; \ldots; \underbrace{111\ldots1}_{n}\right\},$$

где n — длина цепочки, «успех» обозначен через 1, «неудача» — через 0.

2. Стрельба в цель n одинаково метких стрелков (попадание — «успех»):

$$\Omega = \{00\ldots0,\ 10\ldots0,\ 01\ldots0,\ 00\ldots1,\ \ldots,\ 111\ldots1\}.$$

♢ И в первом, и во втором примерах элементами пространства Ω являются цепочки из нулей и единиц длиной n. Это справедливо и для всех испытаний по схеме Бернулли. Пусть вероятность успеха равна p, а неудачи $q = 1 - p$. Тогда вероятность элементарного события $\omega \in \Omega$, содержащего μ единиц (успехов), согласно теореме умножения,

$$P(\omega) = p^{\mu}q^{n-\mu}, \qquad \omega \in \Omega, \quad 0 \leqslant p \leqslant 1, \quad q = 1 - p.$$

Теорема 9.1. *Если μ — число успехов в n испытаниях Бернулли, то*

$$P\{\mu = m\} = P_n(m) = C_n^m p^m q^{n-m}, \tag{9.1}$$

где $m = \overline{0, n}$; p — вероятность «успеха» в отдельном испытании. Числа $P_n(m)$ (9.1) называются биномиальными вероятностями, а формула (9.1) называется формулой Бернулли.

Доказательство. Рассмотрим те элементарные исходы в Ω, у которых число успехов μ равно m:

$$\{\mu = m\} = \{\omega | \mu = m\}. \tag{9.2}$$

Любой элемент ω из (9.2) имеет одну и ту же вероятность

$$p^m q^{n-m}.$$

А сколько всего таких элементов ω? Элементы ω в (9.1) различаются только расположением нулей и единиц. Количество элементов однозначно определяется выбором m мест для единиц из n возможных. Этот выбор можно осуществить C_n^m способами. Следовательно,

$$P\{\mu = m\} = P\{\omega | \mu = m\} = C_n^m p^m q^{n-m},$$

т.е. приходим к (9.1), что и требовалось доказать.

Заметим, что

$$\sum_{m=0}^{n} P\{\mu = m\} = \sum_{m=0}^{n} C_n^m p^m q^{n-m} = \sum_{m=0}^{n} C_n^m p^m (1-p)^{n-m} =$$
$$= (1-p)^n + C_n^1 p(1-p)^{n-1} + C_n^2 p^2 (1-p)^{n-2} + \ldots + p^n = [(1-p)+p]^n = 1,$$

т.е. $P_n(m)$ — члены разложения бинома $(p+q)^n$.

Отметим некоторые частные случаи.

1. Вероятность того, что событие A не наступит ни разу в n испытаниях, равна

$$P_n(0) = q^n.$$

2. Вероятность того, что событие A наступит n раз в n испытаниях, равна

$$P_n(n) = p^n.$$

3. Вероятность того, что событие A произойдёт один раз в n испытаниях, равна

$$P_n(1) = npq^{n-1}.$$

4. Вероятность того, что событие A наступит хотя бы один раз в n испытаниях, равна

$$\sum_{m=1}^{n} P_n(m) = 1 - q^n \tag{9.3}$$

как вероятность события, противоположного событию, что A не произойдёт ни разу.

5. Вероятность того, что событие A произойдёт в n испытаниях не менее r раз, но не более k раз, равна

$$P(A) = \sum_{m=r}^{k} P_n(m).$$

Пример 9.1. Вероятность попадания в цель при одном выстреле равна 0,6. Какова вероятность того, что 8 выстрелов дадут 5 попаданий?

Решение. Здесь $n = 8$, $m = 5$, $p = 0{,}6$, $q = 0{,}4$. Воспользовавшись формулой (9.1), имеем

$$P_8(5) = C_8^5 p^5 q^3 = \frac{8!}{5!(8-5)!} 0{,}6^5 \cdot 0{,}4^3 \approx 0{,}28.$$

Пример 9.2. Для поражения цели нужно не менее 3-х попаданий снаряда. Найти вероятность поражения цели при 7 выстрелах, если вероятность попадания при одном выстреле равна 1/3.

Решение. Пусть событие A — цель поражена. Тогда, обозначив через P вероятность m попаданий в n выстрелах, запишем

$$P(A) = P_7(3) + P_7(4) + P_7(5) + P_7(6) + P_7(7) = 1 - [P_7(0) + P_7(1) + P_7(2)] =$$
$$= 1 - \left[\left(\frac{2}{3}\right)^7 + 7\frac{1}{3} \cdot \left(\frac{2}{3}\right)^6 + \frac{7 \cdot 6}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^5\right] = 0{,}43.$$

Пример 9.3. В среднем 20% пакетов акций на аукционах продаются по первоначально заявленной цене. Найти вероятность того, что из 9 пакетов акций в результате торгов по первоначально заявленной цене:

1) не будут проданы 5 пакетов;
2) будут проданы: а) менее 2 пакетов; б) не более 2; в) хотя бы 2 пакета.

Решение. 1) Вероятность того, что пакет акций не будет продан по первоначально заявленной цене, $p = 1 - 0{,}2 = 0{,}8$. По формуле Бернулли (9.1) имеем $P_9(5) = C_9^5 0{,}8^5 0{,}2^4 = 0{,}066$.

2) По условию $p = 0{,}2$. Тогда

2.а) $P_9(m < 2) = P_9(0) + P_9(1) = C_9^0 0{,}2^0 0{,}8^9 + C_9^1 0{,}2^1 0{,}8^8 = 0{,}436$;

2.б) $P_9(m \leqslant 2) = P_9(0) + P_9(1) + P_9((2) =$
$= C_9^0 0{,}2^0 0{,}8^9 + C_9^1 0{,}2^1 0{,}8^8 + C_9^2 0{,}2^2 0{,}8^7 = 0{,}738$;

2.в) $P_9(m \geqslant 2) = 1 - P_9(m < 2) = 1 - 0{,}436 = 0{,}564$.

Пример 9.4. Монета брошена 10 раз. Найти вероятность того, что хотя бы один раз появится орёл.

Решение. Так как в нашем случае $p = q = 1/2$, то вероятность того, что хотя бы один раз появится орёл в 10 испытаниях определится формулой (9.3) при $n = 10$:

$$P(A) = 1 - q^{10} = 1 - \left(\frac{1}{2}\right)^{10} = \frac{1023}{1024}.$$

Пример 9.5. Сколько раз надо сыграть в «Спортлото», чтобы с вероятностью, большей 1/2, угадать хотя бы один раз 6 номеров из 49.

Решение. Вероятность угадать 6 номеров из 49 в одном розыгрыше можно определить по классической формуле:

$$p = \frac{1}{C_{49}^6} = \frac{43!6!}{49!} \approx 0{,}72 \cdot 10^{-8}.$$

Вероятность хотя бы одного такого события в n розыгрышах (событие A) выразим через вероятность противоположного события: $P(A) = 1 - P(\bar{A})$. Вероятность, что в n испытаниях ни разу не будет угадано 6 номеров: $P(\bar{A}) = (1-p)^n =$

$(1 - 0{,}72 \cdot 10^{-8})^n = 0{,}999999928^n$. Тогда $P(A) = 1 - P(\bar{A}) = 1 - 0{,}999999928^n$. Так как по условиям задачи данная вероятность должна быть больше $1/2$, то получаем уравнение для нахождения n:

$$1 - 0{,}999999928^n > 0{,}5, \quad \text{т.е.} \quad 0{,}999999928^n < 0{,}5.$$

Прологарифмировав, получим $n \ln(0{,}999999928) < \ln(0{,}5)$ или

$$n > \frac{\ln(0{,}5)}{\ln(0{,}999999928)} \approx 9627043{,}8.$$

Пример 9.6. Монету подбрасывают до тех пор, пока орёл не выпадет 4 раза. Найти вероятность того, что решка при этом появится ровно 2 раза.

Решение. Поскольку по условию задачи в последнем испытании должен появиться орёл, то в предыдущих испытаниях должны 3 раза выпасть орёл и 2 раза решка. Тогда интересующее нас событие можно представить в виде произведения двух независимых событий: A — «в серии из пяти подбрасываний орёл выпадет 3 раза» и B — «при шестом подбрасывании выпадет орёл». Вероятность события A находим по формуле Бернулли: $P(A) = C_5^3(1/2)^3(1/2)^2$, а вероятность события B, очевидно, $P(B) = 1/2$. Тогда по теореме умножения $P = P(A)P(B) = C_5^3(1/2)^6 \approx 0{,}156$.

◊ Если каждое испытание имеет k исходов, вероятности которых $p_1, p_2, \ldots, p_k$, $\sum\limits_k p_k = 1$, то вероятность того, что в n испытаниях первый исход появится m_1 раз, второй исход появится m_2 раз и т.д., определится по формуле

$$P = \frac{n!}{m_1! m_2! \cdots m_k!} p_1^{m_1} p_2^{m_2} \cdots p_k^{m_k}. \tag{9.4}$$

Доказательство формулы аналогично доказательству для случая двух исходов.

9.2. Наивероятнейшее число «успехов» в схеме Бернулли

Часто необходимо знать, при каком m вероятность принимает наибольшее значение, т.е. требуется найти наивероятнейшее число m_0 наступления события A в данной серии опытов. Число успехов m_0, которому соответствует наибольшая вероятность в испытаниях по схеме Бернулли, называется *наивероятнейшим числом успехов*.

Теорема 9.2. *Пусть $m_0 = [(n+1)p]$ — наибольшее целое число, не превосходящее $(n+1)p$. Тогда $P_n(k)$ принимает наибольшее значение при $k = m_0$.*

Доказательство. Заметим, что

$$\frac{P_n(k)}{P_n(k-1)} = \frac{C_n^k p^k q^{n-k}}{C_n^{k-1} p^{k-1} q^{n-k+1}} = \frac{n-k+1}{k}\frac{p}{q} =$$
$$= 1 + \frac{(n-k+1)p - kq}{kq} = 1 + \frac{(n+1)p - k}{kq}.$$

Поэтому, если $k < (n+1)p$, то $P_n(k) > P_n(k-1)$, т.е. с возрастанием k вероятности $P_n(k)$ возрастают; если же $k > (n+1)p$, то $P_n(k) < P_n(k-1)$ и с возрастанием k вероятности $P_n(k)$ убывают. Значит, число m_0 должно удовлетворять двойному неравенству

$$np - q \leqslant m_0 \leqslant np + p,$$

где $m_0 \in \mathbb{Z}$ — целое число.

Пусть $m_0 = [(n+1)p]$ — наибольшее целое число, не превосходящее $(n+1)p$. Тогда $P_n(k)$ принимает наибольшее значение при $k = m_0$.

Заметим, что сегмент $[np-q, np+p]$, в котором лежит m_0, имеет длину

$$(np+p) - (np-q) = p + q = 1.$$

Поэтому, если координата какого-либо из его концов не является целым числом, то в этом сегменте лежит единственное целое число и m_0 определено однозначно. В том случае, когда координаты обоих концов — целые числа, имеются два наивероятнейших значения: $m_0^{(1)} = np - q$ и $m_0^{(2)} = np + p$.

Пример 9.7. Вероятность приема сигнала радистом высокого класса равна 0,7. Передано пять сигналов. Определить наиболее вероятное число принятых сигналов.

Решение. Применима схема Бернулли: $n = 5$; $p = 0{,}7$; $q = 0{,}3$. Наиболее вероятным числом принятых сигналов будет то число m_0, при котором вероятность $P_n(m_0)$ будет наибольшей:

$$np + p = 5 \cdot 0{,}7 + 0{,}7 = 4{,}2;$$
$$np - q = 5 \cdot 0{,}7 - 0{,}3 = 3{,}2.$$

Наиболее вероятное значение m_0 лежит на сегменте $[3{,}2; 4{,}2]$ и, следовательно, равно 4.

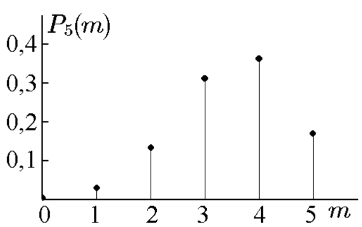


Рис. 21.

Приведём расчёт вероятностей $P_5(m)$ при $m = 1, 2, 3, 4, 5$:

$$P_5(0) = 1 \cdot 0{,}3^5 = 0{,}00243;$$
$$P_5(1) = 5 \cdot 0{,}7 \cdot 0{,}3^4 = 0{,}02835;$$
$$P_5(2) = 10 \cdot 0{,}49 \cdot 0{,}027 = 0{,}13230;$$
$$P_5(3) = 10 \cdot 0{,}343 \cdot 0{,}09 = 0{,}30870;$$
$$P_5(4) = 5 \cdot 0{,}2401 \cdot 0{,}3 = 0{,}36015;$$
$$P_5(5) = 1 \cdot 0{,}16807 = 0{,}16807.$$

Наибольшее значение вероятности отвечает $m = 4$ (рис. 21). Следовательно, наиболее вероятное число сигналов, принятых радистом, равно 4.

Пример 9.8. Вероятность изготовления стандартной детали на автоматическом станке равна 0,9. Определить вероятность того, что из 9 наудачу взятых ...ртными. Найти наиве- ...чных деталей.

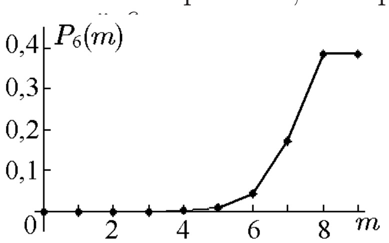


Рис. 22.

Решение. Условие задачи соответствует схеме Бернулли. «Успех» — появление стандартной детали. Тогда $n = 9$, $p = 0{,}9$, $q = 0{,}1$. Подсчитаем вероятность:

$$P_9(6) = C_9^6 \cdot 0{,}9^6 \cdot 0{,}1^3 \approx 0{,}0446$$

и наиболее вероятное число стандартных деталей из 9 взятых наугад:

$$m = [(n+1)p] = [10 \cdot 0{,}9] = [9] = 9.$$

Таким образом, наиболее вероятных значений два: $P_9(9) = P_9(8) \approx 0{,}397$ (рис. 22).

9.3. Предельные теоремы Пуассона и Муавра–Лапласа

Пример 9.9. Вероятность того, что изделие некоторого производства окажется бракованным, равна 0,005. Чему равна вероятность того, что из 10000 наудачу взятых деталей бракованных окажется: а) ровно 40; б) не более 70.

Решение. По формуле Бернулли

$$\text{а)}\quad P_{10000}(40) = C_{10000}^{40} \cdot 0{,}995^{9960} \cdot 0{,}005^{40}.$$

В случае б) воспользуемся аксиомой 3:

$$\text{б)}\quad P_{10000}(m \leqslant 70) = \sum_{m=0}^{70} P_{10000}(m) = \sum_{m=0}^{70} C_{10000}^{m} \cdot 0{,}995^{10000-m} \cdot 0{,}005^{m}.$$

$\Diamond$ Непосредственное вычисление этих вероятностей очень трудоемко. Возникает вопрос об отыскании простых приближенных формул для вероятностей $P_n(m)$ и $\sum\limits_{m=s}^{k} P_n(m)$ при больших n. Затруднения при вычислениях возникают также при малых значениях p и q.

Теорема 9.3 (Пуассона). *Если $n \to \infty$ и $p \to 0$ так, что $np \to \lambda$, $0 < \lambda < \infty$, то*

$$P_n(m) = C_n^m p^m q^{n-m} \to P_m(\lambda) = \frac{\lambda^m}{m!} e^{-\lambda} \tag{9.5}$$

при любом постоянном целом m.

Доказательство. Совершим предельный переход в формуле Бернулли (9.1):

$$\begin{aligned}
\lim_{n\to\infty} P_n(m) &= \lim_{n\to\infty} C_n^m p^m q^{n-m} = \lim_{n\to\infty} \frac{n!}{m!(n-m)!} p^m q^{n-m} = \\
&= \lim_{n\to\infty} \frac{n(n-1)(n-2)\cdots(n-m+1)}{m!} \Big(\frac{\lambda}{n}\Big)^m \Big(1 - \frac{\lambda}{n}\Big)^{n-m} = \\
&= \frac{\lambda^m}{m!} \lim_{n\to\infty} \frac{n(n-1)(n-2)\cdots(n-m+1)}{n^m} \lim_{n\to\infty} \Big(1 - \frac{\lambda}{n}\Big)^{n-m} = \\
&= \frac{\lambda^m}{m!} \lim_{n\to\infty} e^{-\lambda(n-m)/n} = \frac{\lambda^m e^{-\lambda}}{m!} = P_m,
\end{aligned}$$

что и требовалось доказать.

$\Diamond$ Таким образом, при больших n и малых p можно воспользоваться приближенной формулой

$$P_n(m) \approx \frac{\lambda^m}{m!} e^{-\lambda}, \quad \lambda = np. \tag{9.6}$$

Функция $P_m(\lambda)$ табулирована.

Если мало q, то пуассоновским приближением можно воспользоваться для вычисления числа неудач.

Пример 9.10. Завод отправил на базу 5000 доброкачественных деталей. Вероятность того, что в пути изделие повредится, равна 0,0002. Найти вероятность того, что на базу прибудут 3 негодных изделия.

Решение. Условие задачи соответствует схеме Бернулли: $n = 5000$, $p = 0{,}0002$, $\lambda = 1$ (рис. 23)

$$P_{5000}^3 \approx \frac{\lambda^3}{3!}e^{-\lambda} \approx 0{,}06131.$$

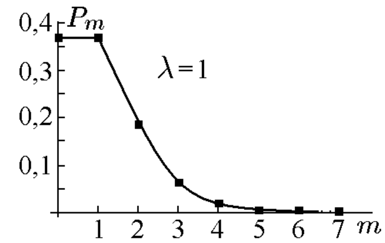


Рис. 23.

Формулу Пуассона (9.6) можно также применять вместо формулы Бернулли, если число испытаний велико и точно не известно, но известно среднее число λ появлений события в этой серии испытаний.

Пример 9.11. Наборщик делает в среднем по одной опечатке на страницу. Предположив, что вероятность опечатки каждого символа одинакова и не зависит от опечаток других символов, найти вероятность того, что на данной странице будет не более двух опечаток, а также вероятность того, что в книге из ста страниц нет страницы, содержащей более двух опечаток.

Решение. Очевидно, что вероятность того или иного числа опечаток на странице определяется формулой Бернулли, однако нам неизвестно ни точное число n символов на данной странице, ни вероятность p опечатки каждого символа. Но, поскольку число символов на странице велико, а вероятность опечатки одного символа мала, то эту вероятность можно приближенно вычислить по формуле Пуассона (9.6) с параметром $\lambda = 1$ (среднее число опечаток на странице). Тогда вероятность того, что страница содержит не более двух опечаток

$$P(m \leqslant 2) = P(m = 0) + P(m = 1) + P(m = 2) \approx \frac{1^0e^{-1}}{0!} + \frac{1^1e^{-1}}{1!} + \frac{1^2e^{-1}}{2!} =$$
$$= e^{-1}\Big(1 + 1 + \frac{1}{2}\Big) = \frac{5}{2e} \approx 0{,}92.$$

Вероятность, что из ста страниц нет страницы, содержащей более двух опечаток: $P = p^{100} \approx (0{,}92)^{100} \approx 2 \cdot 10^{-4}$.

$\Diamond$ Если оба параметра p и q нельзя считать малыми, используются теоремы Муавра–Лапласа (локальная и интегральная).

Введём функции $\varphi(x)$ и $\Phi(x)$ следующими соотношениями:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \quad x \in \mathbb{R};$$

$$\Phi(x) = \int\limits_0^x \varphi(u)du = \frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-u^2/2}du, \quad x \in \mathbb{R}.$$

Функция $\varphi(x)$ — чётная, функция $\Phi(x)$ — нечётная, для них составлены таблицы (см. рис. 24).

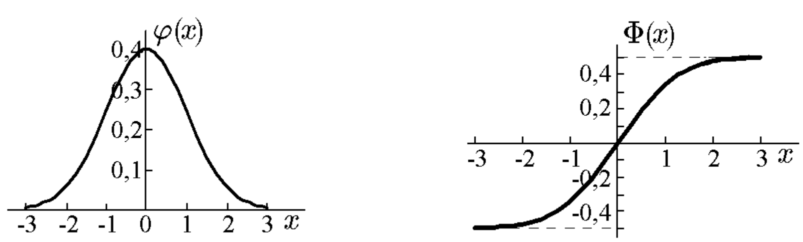


Рис. 24.

Теорема 9.4 (локальная теорема Муавра–Лапласа). *Если* $n \to \infty$, *а* p *постоянна* $(0 < p < 1)$, *то*

$$P_n(m) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{npq}}\varphi(x), \tag{9.7}$$

где

$$x = \frac{m - np}{\sqrt{npq}}.$$

Доказательство приведём позднее в разделе «Теоремы Муавра–Лапласа и Пуассона».

Таким образом, когда p и q не малы, при больших n можно воспользоваться приближенной формулой

$$P_n(m) \approx \frac{1}{\sqrt{npq}}\varphi(x).$$

Подставив сюда x из (9.7), получим

$$P_n(m) \approx \frac{1}{\sqrt{npq}}\frac{1}{\sqrt{2\pi}}\exp\left\{-\left[\frac{m-np}{\sqrt{2npq}}\right]^2\right\}. \tag{9.8}$$

Теорема 9.5 (интегральная теорема Муавра–Лапласа). *Если* p *постоянна* $(0 < p < 1)$ *и* n *достаточно велико, то*

$$P_n(m_1 \leqslant m \leqslant m_2) \xrightarrow[n\to\infty]{} \Phi(x_2) - \Phi(x_1), \tag{9.9}$$

где

$$x_1 = \frac{m_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{m_2 - np}{\sqrt{npq}}. \tag{9.10}$$

◊ Из теоремы следует, что, когда p и q не малы, при больших n можно воспользоваться формулой

$$P_n(m_1 \leqslant m \leqslant m_2) \approx \Phi(x_2) - \Phi(x_1).$$

Подставив сюда x_1 и x_2 из соотношения (9.10), получим

$$P_n(m_1 \leqslant m \leqslant m_2) \approx \Phi\Big(\frac{m_2 - np}{\sqrt{npq}}\Big) - \Phi\Big(\frac{m_1 - np}{\sqrt{npq}}\Big). \tag{9.11}$$

Пример 9.12. Вероятность поражения цели при одном выстреле равна 0,4. Найти вероятность того, что что цель будет поражена а) 200 раз; б) от 200 до 300 раз в серии из 600 выстрелов.

Решение. Условие задачи соответствует схеме Бернулли: $n = 600$; $p = 0{,}4$.

а) Для $P_{600}(200)$ найдём приближение, используя локальную теорему Муавра–Лапласа (9.8):

$$P_{600}(200) \approx \frac{1}{\sqrt{600\cdot 0{,}4\cdot 0{,}6}}\varphi\Big(\frac{200 - 600\cdot 0{,}4}{\sqrt{npq}}\Big) = \frac{1}{12}\varphi\Big(\frac{40}{12}\Big) =$$
$$= 0{,}083\varphi(3{,}33) \approx 0{,}083\cdot 0{,}00154 \approx 0{,}0001.$$

б) Воспользуемся интегральной теоремой Муавра–Лапласа (9.11):

$$P_{600}(200 \leqslant m \leqslant 300) \approx \Phi\Big(\frac{300 - 240}{\sqrt{144}}\Big) - \Phi\Big(\frac{200 - 240}{\sqrt{144}}\Big) = \Phi\Big(\frac{60}{12}\Big) - \Phi\Big(-\frac{40}{12}\Big) =$$
$$= \Phi(5) + \Phi(3{,}33) = 0{,}5 + 0{,}4996 = 0{,}9996.$$

Пример 9.13. По результатам проверок налоговыми инспекциями установлено, что в среднем каждое второе малое предприятие региона имеет нарушение финансовой дисциплины. Найти вероятность того, что из 1000 зарегистрированных в регионе малых предприятий имеют нарушения финансовой дисциплины: а) 480 предприятий; б) не менее 480.

Решение. По условию $p = 0{,}5$. а) Так как можно считать, что n велико ($n \gg 1$), то применим формулу Муавра–Лапласа (9.8):

$$x = \frac{m - np}{\sqrt{npq}} = \frac{480 - 1000 \cdot 0{,}5}{\sqrt{1000 \cdot 0{,}5 \cdot 0{,}5}} = -1{,}265.$$

Тогда

$$P_n(m) \approx \frac{1}{\sqrt{1000 \cdot 0{,}5 \cdot 0{,}5}} \varphi(-1,265) = \frac{\varphi(1,265)}{\sqrt{250}} = \frac{0{,}1793}{\sqrt{250}} = 0{,}0113.$$

б) Необходимо найти $P_{1000}(m \geqslant 480) = P_{1000}(480 \leqslant m \leqslant 1000)$. Воспользуемся соотношением (9.10):

$$x_1 = \frac{480 - 1000 \cdot 0{,}5}{\sqrt{1000 \cdot 0{,}5 \cdot 0{,}5}} = -1{,}265, \qquad x_2 = \frac{1000 - 1000 \cdot 0{,}5}{\sqrt{1000 \cdot 0{,}5 \cdot 0{,}5}} = 31{,}6;$$

Тогда

$$P_{1000}(480 \leqslant m \leqslant 1000) = \Phi(31,6) - \Phi(-1,265) \approx 0{,}5 + 0{,}397 \approx 0{,}897.$$

Вычисление значения $\Phi(1{,}265)$ с помощью пакета Statistica приведено в третьей части настоящего учебного пособия.

Пример 9.14. В страховой компании 10 тыс. клиентов. Страховой взнос каждого клиента составляет 500 руб. При наступлении страхового случая, вероятность которого по имеющимся данным и оценкам экспертов можно считать равной $p = 0{,}005$, страховая компания обязана выплатить клиенту страховую сумму 50 тыс. руб. На какую прибыль может рассчитывать страховая компания с надёжностью 0,95?

Решение. Прибыль компании составляет разность между суммарным взносом всех клиентов и суммой, выплаченной n_0 клиентам при наступлении страхового случая, т.е.

$$\Pi = 500 \cdot 10 - 50 \cdot n_0 = 50(100 - n_0) \text{ тыс. руб.}$$

Для определения n_0 применим интегральную формулу Муавра–Лапласа:

$$P_{10000}(0 \leqslant m \leqslant n_0) = \Phi(x_2) - \Phi(x_1) = 0{,}95,$$

где m — число клиентов, которым будет выплачена страховая сумма;

$$x_1 = \frac{0 - 10000 \cdot 0{,}005}{\sqrt{10000 \cdot 0{,}005 \cdot 0{,}995}} = -7{,}09; \quad x_2 = \frac{n_0 - 10000 \cdot 0{,}005}{\sqrt{10000 \cdot 0{,}005 \cdot 0{,}995}}.$$

Отсюда $\Phi(x_2) = \Phi(x_1) + 0{,}95 = \Phi(-7{,}09) + 0{,}95 \approx -0{,}5 + 0{,}95 \approx 0{,}45$. Из таблицы значений функции Лапласа следует, что $x_2 \approx 1{,}645$. Тогда $n_0 = 50 + x_2\sqrt{49{,}75} \approx 50 + 1{,}645\sqrt{49{,}75} \approx 61{,}6$ и $\Pi \approx 50(100 - 61{,}6) \approx 1920$, т.е. с надёжностью 0,95 ожидаемая прибыль составит 1,92 млн. руб.

Пример 9.15. Из 100 конденсаторов за время T выходят из строя 4 конденсатора. Для контроля выбирают 5 конденсаторов. Найти вероятность того, что среди них за время T выйдет из строя ровно 1 конденсатор, используя классическое определение вероятности, формулу Бернулли, формулу Пуассона и локальную теорему Лапласа.

Решение. 1) Используя классическое определение вероятности, по формуле (5.5) находим

$$P_1 = \frac{C_{k_1}^{m_1} C_{k_2}^{m_2}}{C_N^M} = \frac{C_4^1 C_{96}^4}{C_{100}^5} = \frac{4 \cdot 3321960}{75287520} \approx 0{,}1765.$$

2) По формуле Бернулли (9.1)

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

где $m = m_1 = 1$, $n = M = 5$, $p = k_1/N = 0{,}04$, находим

$$P_2 = c_5^1 \cdot 0{,}04 \cdot (0{,}96)^2 \approx 5 \cdot \cdot 0{,}04 \cdot 0{,}84935 \approx 0{,}1699.$$

3) По формуле Пуассона (9.5):

$$P \approx \frac{\lambda^m}{m!} e^{-\lambda},$$

где $\lambda = np = 0{,}2$, находим

$$P_3 = \frac{(0{,}2)^1}{1} e^{-0{,}2} \approx 0{,}2 \cdot 0{,}81873 \approx 0{,}16375.$$

4) Используя локальную теорему Лапласа, по формуле (9.7):

$$P \approx \frac{\varphi(x)}{\sqrt{npq}},$$

находим

$$x = \frac{m - np}{\sqrt{npq}} = \frac{1 - 0{,}2}{\sqrt{5 \cdot 0{,}4 \cdot 0{,}96}} \approx \frac{0{,}8}{0{,}438178} \approx 1{,}822574$$

и

$$P_4 \approx \frac{\varphi(1{,}8257)}{0{,}438178} \approx \frac{0{,}7535}{0{,}4382} \approx 0{,}1720.$$

♢ Интегральная теорема Муавра–Лапласа позволяет оценить близость частоты и вероятности. Пусть p — вероятность успеха в схеме Бернулли и μ — общее число успехов. Частотой успеха называют отношение μ/n. Оценим вероятность события $A = (|\mu/n - p| < \delta)$.

Событие A можно представить в виде

$$\Big(\Big|\frac{\mu}{n} - p\Big| < \delta\Big) = \Big(\frac{|\mu - np|}{n} < \delta\Big) =$$
$$= (|\mu - np| < \delta n) = (-\delta n + np < \mu < \delta n + np).$$

Если n достаточно велико, то

$$P\Big(\Big|\frac{\mu}{n} - p\Big| < \delta\Big) = P(-\delta n + np < \mu < \delta n + np) \approx$$

$$\approx \Phi\Big(\frac{\delta n + np - np}{\sqrt{npq}}\Big) - \Phi\Big(\frac{-\delta n + np - np}{\sqrt{npq}}\Big) = 2\Phi\Big(\frac{\delta n}{\sqrt{npq}}\Big) = 2\Phi\Big(\delta\sqrt{\frac{n}{pq}}\Big).$$

Итак,

$$P\Big(\Big|\frac{\mu}{n} - p\Big| < \delta\Big) = 2\Phi\Big(\delta\sqrt{\frac{n}{pq}}\Big).$$

Пример 9.16. Вероятность приёма некоторого сигнала равна 0,72. Определить, каково должно быть общее количество принятых сигналов, чтобы частота приёма этого сигнала отличалась от вероятности его приёма не более чем на 0,1 с надёжностью 0,95.

Решение. Условия задачи удовлетворяют схеме Бернулли, в которой $p = 0{,}72$; $\delta = 0{,}1$; n неизвестно, но известно, что

$$P\Big(\Big|\frac{\mu}{n} - p\Big| < \delta\Big) = 0{,}95$$

или

$$2\Phi\Big(\delta\sqrt{\frac{n}{pq}}\Big) = 0{,}95;$$

$$2\Phi\Big(0{,}1\sqrt{\frac{n}{0{,}72\cdot 0{,}28}}\Big) = 0{,}95;$$

$$\Phi\Big(0{,}1\sqrt{\frac{n}{0{,}72\cdot 0{,}28}}\Big) = \frac{0{,}95}{2} = 0{,}475.$$

По таблице находим

$$0{,}1\sqrt{\frac{n}{0{,}72\cdot 0{,}28}} = 1{,}96,$$

т.е.

$$n = (1{,}96)^2\cdot 0{,}72\cdot 0{,}28\cdot 100 \approx 77{,}45.$$

Так как n должно быть целым, то общее количество принятых сигналов должно быть не менее 77.

ГЛАВА 2

Случайные величины

10. Одномерная случайная величина

10.1. Одномерная случайная величина. Функция распределения одномерной случайной величины

До сих пор мы имели дело со случайными событиями. Событие является качественной характеристикой результата случайного опыта. Но случайный результат опыта можно характеризовать и количественно, если ввести соответствие между элементарными исходами и некоторыми числами. Количественной характеристикой случайного опыта является случайная величина.

Принято определять случайную величину как числовую функцию, определяемую на множестве элементарных событий.

♦ Числовая функция $\xi = \xi(\omega)$ со значениями в $\mathbb{R}$, определённая на пространстве элементарных событий Ω, называется *случайной величиной*.

Вообще говоря, если мы имеем дело с вероятностями, определёнными на некоторой σ-алгебре событий $\mathcal{F}$, то случайная величина $\xi(\omega)$ должна быть не произвольной функцией, отображающей Ω в $\mathbb{R}$, а σ-измеримой функцией. То есть для любого $x \in \mathbb{R}$ множество исходов ω, для которых $\xi(\omega) < x$, должно принадлежать σ-алгебре событий $\mathcal{F}$.

Мы будем обозначать случайные величины греческими буквами ξ, η, ζ, μ и т.д., а принимаемые ими значения — строчными латинскими буквами (в англо-американской и иногда в отечественной литературе случайные величины обозначаются прописными латинскими буквами X, Y, Z).

При таком определении на случайные величины распространяются все правила действий с обычными функциями: их можно складывать, вычитать, перемножать и т.д.

Пример 10.1. В модели двукратного подбрасывания монеты с пространством исходов $\Omega = \{ГГ, ГР, РГ, РР\}$ или $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ определим случайную величину $\xi = \xi(\omega)$ с помощью таблицы

ω	ГГ	ГР	РГ	РР
$\xi(\omega)$	2	1	1	0

Здесь $\xi(\omega)$ по своему смыслу есть не что иное, как число «гербов», отвечающих исходу ω. Поскольку в рассматриваемом случае Ω состоит из конечного числа точек, то множество значений $\{0, 1, 2\} \subset \mathbb{R}$ случайной величины ξ также конечно.

Пример 10.2. Производятся выстрелы до первого попадания в цель. Обозначим через единицу попадание, через нуль промах. Тогда $\Omega = \{1; 01; 001; 0001; \ldots\}$. Множество исходов счётно. Случайная величина ξ — число выстрелов до первого попадания в цель. Множество её возможных значений счётно:

$$\xi \in \{1, 2, 3, 4, \ldots\}.$$

Пример 10.3. Время безотказной работы сотового телефона — случайная величина, множество значений которой заполняет интервал $[0, \infty[$.

Пример 10.4. Цена акции — случайная величина ξ. Область изменения случайной величины $]0, +\infty[$.

◊ Любое событие может быть охарактеризовано случайной величиной. Справедливо и обратное: каждое значение случайной величины ξ можно трактовать как событие, причём различным значениям ξ соответствуют непересекающиеся события.

Случайная величина считается определённой, если известен закон её распределения.

♦ *Законом распределения случайной величины* называется всякое соотношение, устанавливающее связь между возможными значениями или множеством значений случайной величины и соответствующими вероятностями.

◊ В законе распределения содержится основная и важная информация о случайной величине. Из него можно получить практически все возможные сведения о случайной величине.

◊ В зависимости от возможных значений, принимаемых случайной величиной, и характера закона распределения действительные случайные величины можно разделить на три группы: дискретные, непрерывные и непрерывно-дискретные (смешанные).

Итак, мы приходим к выводу, что для полной характеристики той или иной случайной величины как таковой необходимо и достаточно знать:

1) перечень всех возможных значений этой случайной величины,

2) вероятности, соответствующие этим значениям (или множеству значений).

Наиболее общей формой закона распределения случайной величины является функция распределения.

♦ *Функцией распределения случайной величины* ξ называется функция $F_\xi(x)$, для каждого вещественного значения x равная вероятности события $\xi < x$, где случайная величина принимает значения меньше x, т.е.

$$F(x) = F_\xi(x) = P(\xi < x). \tag{10.1}$$

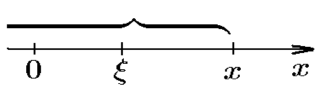


Рис. 25.

Определение функции распределения допускает простую геометрическую интерпретацию. Если рассматривать случайную величину как случайную точку ξ оси Ox, которая в результате эксперимента может занять то или иное положение, то функция распределения есть вероятность того, что случайная точка ξ попадёт левее точки x (рис. 25).

◊ Функцию распределения $F(x)$ иногда называют также интегральной функцией распределения или интегральным законом распределения.

Функция распределения — универсальная характеристика случайной величины. Она существует для всех случайных величин: как прерывных, так и непрерывных.

10.2. Свойства функции распределения одномерной случайной величины

Рассмотрим основные свойства функции распределения одномерной случайной величины.

Свойство 1. Функция распределения $F(x)$ есть неубывающая неотрицательная функция своего аргумента, т.е. если $x_2 > x_1$, то $F(x_2) \geqslant F(x_1) \geqslant 0$.

В соответствии с первой аксиомой вероятностей, $F(x) = P(\xi < x) \geqslant 0$. Пусть $x_1 < x_2$. Так как событие $\xi < x_1$ влечёт событие $\xi < x_2$, то $P(\xi < x_1) \leqslant P(\xi < x_2)$, т.е. $F(x_1) \leqslant F(x_2)$.

Свойство 2. На минус бесконечности функция распределения равна нулю:

$$F(-\infty) = \lim_{x\to-\infty} F(x) = 0.$$

Действительно, событие $\xi < -\infty$ является невозможным. Следовательно, $P(\xi < -\infty) = P(\emptyset) = 0$. Тогда по определению $F(-\infty) = 0$.

Свойство 3. На плюс бесконечности функция распределения равна единице:

$$F(+\infty) = \lim_{x\to+\infty} F(x) = 1. \tag{10.2}$$

Действительно, событие $\xi < +\infty$ является достоверным. Следовательно, $P(\xi < \infty) = P(\Omega) = 1$. Тогда по определению $F(\infty) = 1$.

Свойство 4. Функция распределения непрерывна слева (рис. 26):

$$\lim_{x\to x_0-0} F(x) = F(x_0).$$

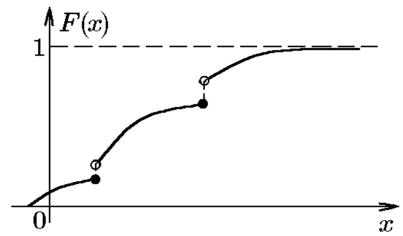


Рис. 26.

Действительно, пусть $\{x_n\}$ — возрастающая последовательность, сходящаяся к x_0. Тогда $\{\xi < x_1\} \subset \{\xi < x_2\} \subset \ldots \subset \{\xi < x_n\} \subset \{\xi < x_{n+1}\} \subset \ldots$ и $\sum\limits_{n=1}^{\infty}\{\xi < x_n\} = \{\xi < x_0\}$. Следовательно, в соответствии с аксиомой 5 теории вероятностей получим

$$\lim_{x\to x_0-0} F(x) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} P(\xi < x_n) = P(\xi < x_0) = F(x_0).$$

Свойство 5. Вероятность того, что случайная величина примет значения из заданного интервала, равна приращению функции распределения в этом интервале:

$$P(a \leqslant \xi < b) = F(b) - F(a).$$

Так как событие $(\xi < b)$ можно представить в виде суммы двух несовместных событий: $(\xi < b) = (\xi < a) + (a \leqslant \xi < b)$, то, согласно третьей аксиоме теории вероятности, $P(\xi < b) = P(\xi < a) + P(a \leqslant \xi < b)$. Следовательно, $P(a \leqslant \xi < b) = P(\xi < b) - P(\xi < a) = F(b) - F(a)$.

♢ В частности, вероятность $P(\xi = x_0)$ того, что случайная величина примет заданное значение, определяется соотношением

$$P(\xi = x_0) = F(x_0 + 0) - F(x_0).$$

Из этого соотношения следует, что если функция распределения непрерывна в точке x_0, то $P(\xi = x_0) = 0$.

♢ График функции распределения $F(x)$ в общем случае представляет собой график неубывающей функции, значения которой начинаются от нуля и доходят до единицы, причём в отдельных точках функция может иметь скачки (разрывы) (рис. 26).

Таким образом, универсальной характеристикой, одинаково пригодной как для дискретных, так и для непрерывных одномерных случайных величин, является функция распределения вероятностей $F(x)$, определяющая вероятность того, что случайная величина ξ примет значение меньше некоторого числа x.

10.3. Дискретная одномерная случайная величина

Наиболее простой вид имеют законы распределения так называемых дискретных величин.

⧫ Случайная величина называется *дискретной*, если множество её возможных значений конечно или счётно (счётное множество – такое множество, элементы которого могут быть пронумерованы).

Примерами могут служить:

1. Число вызовов, поступающих оператору сотовой связи в течение суток. Случайная величина ξ может принимать следующее множество значений:

$$x_1 = 0,\ x_2 = 1,\ x_3 = 2,\ \ldots$$

2. Число дефектных изделий в партии из n штук. Возможные значения случайной величины ξ следующие: 0, 1, 2, ..., n.

Пусть дискретная случайная величина ξ имеет конечное множество возможных значений

$$x_1,\ x_2,\ \ldots,\ x_n.$$

В результате опыта величина ξ примет одно из этих значений, т.е. произойдёт одно из событий

$$\{\xi = x_1\},\ \{\xi = x_2\},\ \ldots,\ \{\xi = x_n\}.$$

Рассматриваемые события несовместны, так как случайная величина в результате эксперимента может принять только одно значение, и образуют полную группу событий (никаких других событий, кроме указанных, в результате опыта произойти не может), т.е.

$$\{\xi = x_1\} + \{\xi = x_2\} + \ldots + \{\xi = x_n\} = \Omega.$$

Обозначим вероятность событий через

$$P(\xi = x_i) = P_i.$$

Тогда

$$\sum_{i=1}^{n} P(\xi = x_i) = \sum_{i=1}^{n} P_i = 1.$$

Последнее соотношение остаётся справедливым и при $n \to \infty$.

Простейшей формой закона распределения дискретной величины ξ является *ряд распределения* — таблица, в верхней строке которой перечислены все значения случайной величины $x_1, x_2, \ldots, x_i, \ldots$ в порядке их возрастания, а в нижней — соответствующие им вероятности

ξ	x_1	x_2	$\ldots$	x_i	$\ldots$
P	P_1	P_2	$\ldots$	P_i	$\ldots$

.

Графическое изображение ряда распределения называется *многоугольником распределения* (рис. 27).

Для дискретной случайной величины ξ, которая может принимать значения $x_1, x_2, \ldots, x_n$, функция распределения имеет вид

$$F(x) = P(\xi < x) = \sum_{x_i < x} P(\xi = x_i).$$

Суммирование распространяется на все возможные значения, которые по своей величине меньше аргумента x. Из определения следует, что функция распределения дискретной случайной величины разрывна и возрастает скачком при переходе через точки возможных значений $x_1, x_2, \ldots, x_n$, причём величина скачка равна вероятности соответствующего значения (рис. 28).

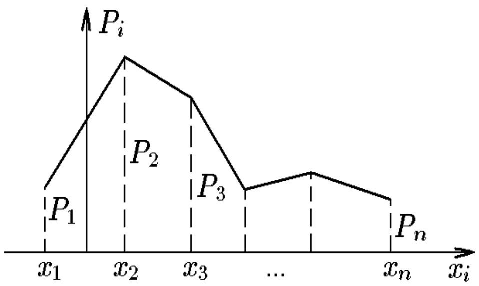


Рис. 27.

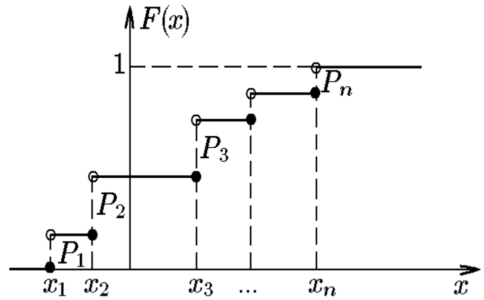


Рис. 28.

Пример 10.5. Дан ряд распределения случайной величины ξ:

ξ	0	1	2	3
P	0,1	0,3	0,2	0,4

Найти функцию распределения $F(x)$ этой случайной величины.

Решение. По определению функции распределения найдём, что

если $x \leqslant 0$, то $F(x) = P(\xi < x) = 0$;

если $0 < x \leqslant 1$, то $F(x) = P(\xi < x) = P(\xi = x_1) = p_1 = 0{,}1$;

если $1 < x \leqslant 2$, то $F(x) = P(\xi < x) = P(\xi = x_1) + P(\xi = x_2) = P_1 + P_2 = 0{,}1 + 0{,}3 = 0{,}4$;

если $2 < x \leqslant 3$, то $F(x) = P(\xi < x) = P(\xi = x_1) + P(\xi = x_2) + P(\xi = x_3) = P_1 + P_2 + P_3 = 0{,}1 + 0{,}3 + 0{,}2 = 0{,}6$;

если $x > 3$, то $F(x) = P(\xi < x) = P(\xi = x_1) + P(\xi = x_2) + P(\xi = x_3) + P(\xi = x_4) = P_1 + P_2 + P_3 + P_4 = 0{,}1 + 0{,}3 + 0{,}2 + 0{,}4 = 1{,}0$.

10.4. Непрерывные случайные величины. Плотность вероятности и функция распределения

Множество возможных значений дискретной случайной величины конечно или счётно; недискретные случайные величины характеризуются тем, что их множество возможных значений не счётно.

♦ Случайная величина ξ называется *непрерывной*, если для неё существует такая неотрицательная кусочно-непрерывная функция $f(x)$, что функция распределения случайной величины ξ удовлетворяет равенству

$$F(x) = \int\limits_{-\infty}^{x} f(t)dt. \tag{10.3}$$

Функция $f(x)$ называется *плотностью распределения вероятностей*, или кратко – плотностью распределения.

Примеры непрерывных случайных величин:

1) длительность телефонного разговора $t \in [0, \infty[$;

2) ошибка измерения;

3) величина входного сигнала в радиотехническом устройстве;

4) отклонение сопротивления резистора от номинального – область возможных значений $]-\infty, \infty[$;

5) цена акции или опциона.

Свойства плотности и функции распределения вероятностей непрерывной случайной величины

Свойство 1. Плотность распределения вероятностей является неотрицательной функцией ($f(x) \geqslant 0$) для всех $x \in]-\infty, \infty[$.

Доказательство непосредственно следует из определения.

Свойство 2. Если случайная величина имеет непрерывную плотность распределения вероятности, то её функция распределения также непрерывна.

Действительно, непрерывность функции распределения следует из представления

$$F_\xi(x) = \int\limits_{-\infty}^{x} f_\xi(t)\,dt$$

и непрерывности интеграла как функции переменного верхнего предела.

Свойство 3. Если функция $f(x)$ непрерывна в точке x_0, то

$$F'(x_0) = f(x_0).$$

Доказательство. В случае непрерывной функции $f(x)$ функция $F(x)$ дифференцируема и

$$\frac{dF(x)}{dx} = \frac{d}{dx}\int\limits_{-\infty}^{x} f(t)dt = f(x).$$

Свойство 4. Справедливо соотношение

$$P(a \leqslant \xi < b) = \int\limits_{a}^{b} f(x)dx. \tag{10.4}$$

Согласно свойству 5 функции распределения вероятности,

$$P(a \leqslant \xi < b) = F(b) - F(a)$$

для всех a, b, $a < b$. Из определения непрерывной случайной величины

$$F(b) = \int\limits_{-\infty}^{b} f(x)dx = \int\limits_{-\infty}^{a} f(x)dx - \int\limits_{a}^{b} f(x)dx,$$

т.е.

$$F(b) = F(a) + \int\limits_{a}^{b} f(x)dx,$$

откуда следует

$$\int\limits_{a}^{b} f(x)dx = F(b) - F(a).$$

Свойство 5. Справедливо соотношение

$$\int\limits_{-\infty}^{\infty} f(x)dx = 1,$$

Действительно, по свойству 3 функций распределения (10.2): $F(\infty) = 1$, следовательно,

$$F(\infty) = \int\limits_{-\infty}^{\infty} f(x)dx = 1.$$

Свойство 6. Справедливо представление

$$f(x) \approx \frac{P(x \leqslant \xi < x + \Delta x)}{\Delta x} = \frac{F(x + \Delta x) - F(x)}{\Delta x},$$

чем и объясняется термин «плотность вероятности».

Действительно, если $f(x)$ непрерывна, то по теореме о среднем

$$P(x \leqslant \xi < x + \Delta x) = \int\limits_{x}^{x+\Delta x} f(x)dt = f(c)\Delta x,$$

где $\Delta x > 0$, $x \leqslant c \leqslant x + \Delta x$.

Геометрически $P(x \leqslant \xi < x + \Delta x)$ при малых Δx приближенно равна площади элементарного прямоугольника, опирающегося на отрезок Δx (см. рис. 29)

$$P(x \leqslant \xi < x + \Delta x) = f(x)\Delta x + \beta(x)$$

и

$$f(c) = f(x) + \alpha(x),$$

где $\alpha(x) \to 0$ и $\beta(x) \to 0$ при $\Delta x \to 0$. Откуда и следует справедливость утверждения.

Свойство 7. Вероятность того, что непрерывная случайная величина ξ примет фиксированное значение a, равна нулю:

$$P(\xi = a) = \int\limits_{a}^{a} f(x)dx = 0.$$

Непосредственно следует из соотношения (10.4).

$\Diamond$ Событие $\xi = a$ возможно, но имеет вероятность, равную нулю. Соответственно этому, имеется противоположное событие, означающее попадание случайной точки ξ в любую точку действительной оси за исключением точки a и являющееся не достоверным, но имеющим вероятность, равную единице.

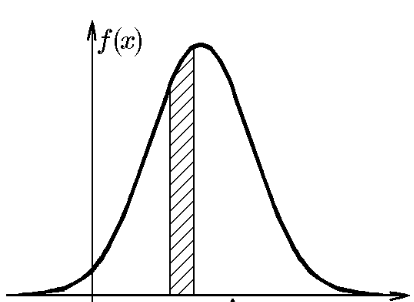


Рис. 29.

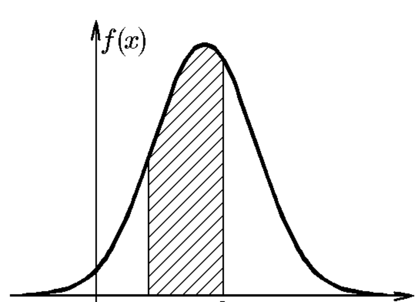


Рис. 30.

◊ Вероятность попадания любой случайной величины (дискретной или непрерывной) на участок оси от a до b (включая a и не включая b) выражается формулой (10.4):

$$P(a \leqslant \xi < b) = F(b) - F(a).$$

Так как для непрерывной случайной величины $P\{\xi = a\} = 0$, знак равенства $(a \leqslant \xi)$ можно отбросить и

$$P(a < \xi < b) = F(b) - F(a).$$

Геометрически эта вероятность равна площади криволинейной трапеции, ограниченной прямыми $x = a$, $x = b$, осью Ox и кривой $y = f(x)$ (см. рис. 30).

Формула (10.3) имеет физическую аналогию. Вероятность $P\{a < \xi < b\}$ можно трактовать как массу неоднородного стержня с линейной плотностью распределения массы $f(x)$, причём масса всего стержня равна единице, так как

$$\int\limits_a^b f(x)dx = 1$$

как вероятность достоверного события.

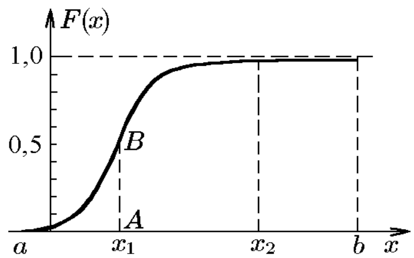


Рис. 31.

◊ Пусть построен график функции распределения $F(x)$ некоторой случайной величины ξ (см. рис. 31). Ордината $|AB|$ графика при $\xi = x_1$ геометрически изображает вероятность того, что случайная величина примет какое-нибудь значение, меньшее x_1, а не вероятность того, что она примет значение x_1. Ранее установлено, что вероятность того, что непрерывная случайная величина примет определённое значение x_1, равна нулю.

Отсюда, в частности, следует, что для непрерывных случайных величин вероятности

$$P(x_1 < \xi < x_2), \quad P(x_1 \leqslant \xi < x_2), \quad P(x_1 < \xi \leqslant x_2), \quad P(x_1 \leqslant \xi \leqslant x_2)$$

совпадают и равны $F(x_2) - F(x_1)$.

Если непрерывная случайная величина ξ может принимать только значения в интервале от a до b (где a и b — некоторые постоянные), то её функция распределения равна нулю для всех значений $x \leqslant a$ и единице для всех значений $x > b$, поскольку события $\xi < x$ для любого значения $x < a$ являются в этом случае невозможными, а для любого значения $x > b$ — достоверными.

Сравним свойства $F(x)$ и $f(x)$:

$F(x) = P(\xi < x)$	$f(x) = F'(x)$
1. $F(x)$ – неубывающая функция, т.е. $F(x_2) \geqslant F(x_1)$ при $x_2 > x_1$	1. $f(x) \geqslant 0$
2. $F(-\infty) = \lim\limits_{x\to-\infty} F(x) = 0$	2. $F(x) = \int\limits_{-\infty}^{x} f(t)dt$
3. $F(+\infty) = \lim\limits_{x\to+\infty} F(x) = 1$	3. $\int\limits_{-\infty}^{\infty} f(x)dx = 1$
4. $P(x_1 \leqslant \xi < x_2) = F(x_2) - F(x_1)$	4. $P(x_1 \leqslant \xi < x_2) = \int\limits_{x_1}^{x_2} f(x)dx$

Пример 10.6. Дана плотность распределения $f(x)$ случайной величины ξ:

$$f(x) = \begin{cases} A\cos x, & |x| \leqslant \pi/2; \\ 0, & |x| > \pi/2. \end{cases}$$

Требуется
1) найти коэффициент A;
2) построить график плотности распределения $f(x)$;
3) найти функцию распределения $F(x)$ и построить ее график;
4) найти вероятность попадания величины ξ в интервал от 0 до $\pi/4$.

Решение. 1) Для определения коэффициента A воспользуемся условием нормировки плотности распределения:

$$\int\limits_{-\infty}^{\infty} f(x)dx = \int\limits_{-\pi/2}^{\pi/2} A\cos x\,dx = 2A = 1,$$

откуда $A = 1/2$.

2) График плотности $f(x)$ представлен на рис. 32.

3) По определению,

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

Так как на интервале $]-\infty, -\pi/2[$ функция $f(x) = 0$, то $F(x) = 0$.

На интервале $]-\pi/2, \pi/2[$ функция $f(x) = 0{,}5\cos x$, следовательно, на этом интервале

$$F(x) = \int\limits_{-\infty}^{-\pi/2} f(t)dt + \int\limits_{-\pi/2}^{x} f(t)dt = 0 + \int\limits_{-\pi/2}^{x} \frac{1}{2}\cos t\,dt = \frac{1}{2}(\sin x + 1).$$

Для интервала $]\pi/2, \infty[$ запишем

$$F(x) = \int\limits_{-\infty}^{-\pi/2} f(t)dt + \int\limits_{-\pi/2}^{\pi/2} f(t)dt + \int\limits_{\pi/2}^{x} f(t)dt = \int\limits_{-\pi/2}^{\pi/2} \frac{1}{2}\cos t\,dt = 1.$$

Следовательно,

$$F(x) = \begin{cases} 0, & x < -\pi/2; \\ \dfrac{1}{2}(\sin x + 1), & |x| \leqslant \pi/2; \\ 1, & x > \pi/2. \end{cases}$$

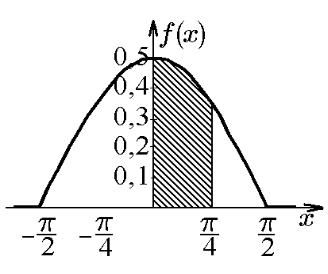


Рис. 32.

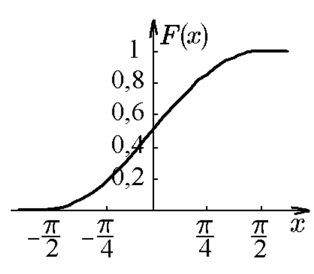


Рис. 33.

График функции $F(x)$ изображен на рис. 33.

4) Вероятность попадания величины ξ на участок от 0 до $\pi/4$ находим по формуле

$$P\Big(0 \leqslant \xi < \frac{\pi}{4}\Big) = \int\limits_0^{\pi/4} \frac{1}{2}\cos x\,dx = \frac{1}{2}\sin x\Big|_0^{\pi/4} = \frac{\sqrt{2}}{4}$$

либо по формуле $P(\alpha \leqslant \xi < \beta) = F(\beta) - F(\alpha)$:

$$P\Big(0 \leqslant \xi < \frac{\pi}{4}\Big) = \frac{1}{2}\Big(\sin\frac{\pi}{4} + 1\Big) - \frac{1}{2}(\sin 0 + 1) = \frac{\sqrt{2}}{4}.$$

С геометрической точки зрения, вероятность $p = \sqrt{2}/4$ равна площади заштрихованной на рис. 32 области.

11. Системы случайных величин. Функция распределения системы случайных величин

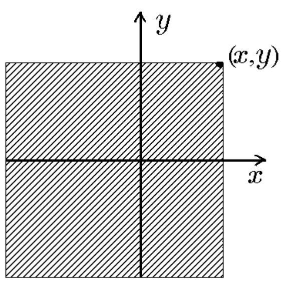


Рис. 34.

◊ В практических задачах часто приходится иметь дело с системами случайных величин $(\xi, \eta, \zeta, \ldots)$. Свойства системы случайных величин не исчерпываются свойствами её отдельных составляющих, они включают также зависимости между этими величинами. При рассмотрении вопросов, связанных с системами случайных величин, удобно пользоваться геометрической интерпретацией системы. Например, для двух случайных величин $\{\xi, \eta\}$ вектор $\vec{\zeta} = (\xi, \eta)$ — случайный вектор. Для наглядности будем рассматривать системы двух случайных величин. Обобщение на случай произвольной размерности не представляет труда.

♦ Совокупность случайных величин $\{\xi, \eta\}$, заданных на одном вероятностном пространстве, будем называть *случайным вектором* $\vec{\zeta} = (\xi, \eta)$ или *системой случайных величин*.

♦ *Функцией распределения системы двух случайных величин* ξ и η называется функция $F(x, y)$, в каждой точке $(x, y) \in \mathbb{R}^2$ равная вероятности события $(\xi < x, \eta < y)$:

$$F(x, y) = F_{\xi,\eta}(x, y) = P(\xi < x, \eta < y).$$

Если рассматривать случайную величину $\vec{\zeta}$ как случайную точку плоскости xOy, то функция распределения $F(x, y)$ есть вероятность того, что точка с координатами (ξ, η) принадлежит квадранту $\xi < x$, $\eta < y$ (см. рис. 34).

◊ Аналогично определяется функция распределения системы n случайных величин.

11.1. Свойства функции распределения двумерной случайной величины

Свойство 1. Функция $F(x, y)$ — неубывающая функция своих аргументов.

Доказательство аналогично соответствующему доказательству для одномерной случайной величины.

Свойство 2. Справедливы соотношения $F(x, -\infty) = F(-\infty, y) = F(-\infty, -\infty) = 0$.

Действительно, $F(-\infty, y) = P(\xi < -\infty, \eta < y) = 0$, так как событие $\xi < -\infty$ невозможно.

Свойство 3. Справедливы соотношения $F(x,\infty) = F_\xi(x)$, $F(\infty, y) = F_\eta(y)$, где $F_\xi(x)$ и $F_\eta(y)$ — функции распределения величин ξ и η соответственно.

Действительно, $F(\infty, y) = P(\xi < \infty, \eta < y) = P(\eta < y) = F_\eta(y)$, так как событие $\xi < \infty$ — достоверное.

Свойство 4. Справедливо соотношение $F(\infty,\infty) = 1$.

Действительно, $F(\infty,\infty) = P(\xi < \infty, \eta < \infty) = 1$ как вероятность достоверного события.

Свойство 5. Функция $F(x, y)$ непрерывна по каждой из переменных x, y слева.

Доказательство аналогично доказательству соответствующего свойства одномерной случайной величины.

Свойство 6. Вероятность попадания случайной точки (x, y) в прямоугольник $\alpha \leqslant x < \beta$, $\gamma \leqslant y < \delta$:

$$P(\alpha \leqslant x < \beta, \gamma \leqslant y < \delta) = F(\beta, \delta) - F(\beta, \gamma) - F(\alpha, \delta) + F(\alpha, \gamma)$$

(по аналогии с одномерной случайной величиной условимся левую и нижнюю границы включать в интервал, а правую и верхнюю не включать).

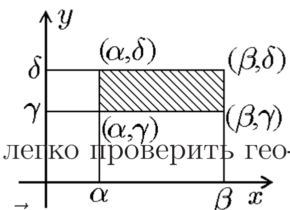


Рис. 35.

Справедливость этого утверждения легко проверить геометрически (рис. 35).

♦ Говорят, что случайный вектор $\vec{\zeta} = \{\xi, \eta\}$ имеет *дискретное распределение*, если множество пар значений его компонент (x_i, y_i) конечно или счётно, причём

$$\sum_i \sum_j P(\xi = x_i, \eta = y_j) = 1.$$

♦ Таблица, содержащая перечень возможных комбинаций значений компонент дискретного случайного вектора $\{\xi, \eta\}$ и соответствующих им вероятностей, называется *таблицей совместного распределения* величин ξ и η.

Если известен закон распределения дискретного случайного вектора (таблица распределения), то из него могут быть получены одномерные законы распределения его компонент (обратное в общем случае неверно):

$$P(\xi = x_i) = \sum_j P(\xi = x_i, \eta = y_j), \qquad P(\eta = y_j) = \sum_i P(\xi = x_i, \eta = y_j).$$

♦ Говорят, что случайный вектор $\vec{\zeta} = (\xi, \eta)$ имеет *непрерывное распределение*, если существует неотрицательная функция $f_{\xi,\eta}(x, y)$ такая, что для любой точки $(x, y) \in \mathbb{R}^2$ функция распределения $F_{\xi\eta}(x, y)$ случайного вектора $\{\xi, \eta\}$ представима в виде

$$F_{\xi,\eta}(x, y) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f_{\xi,\eta}(\xi, \eta)\, d\xi d\eta. \tag{11.1}$$

Функция $f_{\xi,\eta}(x, y)$ называется *плотностью совместного распределения* величин ξ и η. Плотность совместного распределения обладает свойствами, аналогичными свойствам плотности распределения одномерной величины.

♢ Геометрически функцию $f(x, y)$ можно изобразить поверхностью, которая называется *поверхностью распределения*.

11.2. Свойства плотности распределения двумерной случайной величины

Свойство 1. Плотность распределения является неотрицательной функцией: $f(x,y) \geqslant 0$.

Доказательство непосредственно следует из определения.

Свойство 2. Справедливо соотношение

$$\int\limits_{-\infty}^{\infty} dx \int\limits_{-\infty}^{\infty} dy\; f_{\xi,\eta}(x,y) = 1.$$

Доказательство аналогично доказательству соответствующего свойства одномерной случайной величины.

Свойство 3. Справедливо соотношение

$$f_{\xi,\eta}(x,y) = \frac{\partial^2 F_{\xi,\eta}(x,y)}{\partial x \partial y}.$$

Доказательство непосредственно следует из определения (11.1).

♢ Исходя из данного свойства и определения производной, плотность совместного распределения вероятности можно было бы определить как

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \lim_{\Delta y \to 0} \frac{1}{\Delta y}\Big(\frac{\partial F(x,y+\Delta y)}{\partial x} - \frac{\partial F(x,y)}{\partial x}\Big) =$$
$$= \lim_{\Delta y \to 0} \frac{1}{\Delta y}\Big(\lim_{\Delta x \to 0} \frac{F(x+\Delta x, y+\Delta y) - F(x, y+\Delta y)}{\Delta x} -$$
$$- \lim_{\Delta x \to 0} \frac{F(x+\Delta x, y) - F(x,y)}{\Delta x}\Big) =$$
$$= \lim_{\substack{\Delta x \to 0\\ \Delta y \to 0}} \frac{F(x+\Delta x, y+\Delta y) - F(x+\Delta x, y) - F(x, y+\Delta y) + F(x,y)}{\Delta x \Delta y} =$$
$$= \lim_{\substack{\Delta x \to 0\\ \Delta y \to 0}} \frac{P(x \leqslant \xi < x+\Delta x, y \leqslant \eta < y+\Delta y)}{\Delta x \Delta y}.$$

Свойство 4. Вероятность попадания в область D равна объёму цилиндрического тела, направляющей которого является граница области D, а образующая параллельна оси Oz, ограниченного поверхностью распределения $z = f(x,y)$ и опирающегося на эту область:

$$P\big((\xi,\eta) \in D\big) = \iint\limits_D f(x,y) dx\, dy. \tag{11.2}$$

Данное соотношение вытекает непосредственно из предыдущего свойства, позволяющего определить вероятность попадания случайных величин ξ, η в бесконечно малую область $dx\, dy$ как $f(x,y) dx\, dy$. В частности, для прямоугольной области

$$P(\alpha \leqslant \xi < \beta, \gamma \leqslant \eta < \delta) = \int\limits_\alpha^\beta \int\limits_\gamma^\delta f(x,y) dx\, dy. \tag{11.3}$$

Свойство 5. Если случайный вектор $\vec{\zeta} = (\xi, \eta)$ имеет непрерывное распределение, то и каждая из величин ξ и η имеет непрерывное распределение:

$$f_\xi(x) = \int\limits_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dy; \qquad f_\eta(y) = \int\limits_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dx. \tag{11.4}$$

Доказательство. По определению $F_\xi(x) = P(\xi < x)$, при этом величина η может принимать любые значения. Следовательно,

$$F_\xi(x) = F_{\xi,\eta}(x,\infty) = \int\limits_{-\infty}^{x} \Big(\int\limits_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dy \Big) dx.$$

Таким образом, ξ — непрерывная случайная величина с плотностью

$$f_\xi(x) = \int\limits_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dy.$$

Аналогично доказывается второе соотношение из (11.4).

11.3. Независимые случайные величины

♦ Случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ называются *независимыми*, если для любых множеств $\mathcal{D}_1 \subset \mathbb{R}$, $\mathcal{D}_2 \subset \mathbb{R}$, $\ldots, \mathcal{D}_n \subset \mathbb{R}$ имеет место равенство

$$P(\xi_1 \in \mathcal{D}_1, \xi_2 \in \mathcal{D}_2, \ldots, \xi_n \in \mathcal{D}_n) = P(\xi_1 \in \mathcal{D}_1)P(\xi_2 \in \mathcal{D}_2) \cdots P(\xi_n \in \mathcal{D}_n).$$

Можно показать, что это определение равносильно следующему.

♦ Случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ называются *независимыми*, если для любых $x_1, x_2, \ldots, x_n$ имеет место равенство

$$F(x_1, x_2, \ldots, x_n) = F_{\xi_1}(x_1)F_{\xi_2}(x_2) \cdots F_{\xi_n}(x_n). \tag{11.5}$$

Для систем случайных величин, имеющих дискретное или непрерывное распределение, последнее определение можно переформулировать так.

♦ Случайные величины $\xi_1, \xi_2, \ldots, \xi_n$, имеющие дискретное распределение, независимы, если для любых $x_1, x_2, \ldots, x_n$ имеет место равенство

$$P(\xi_1 = x_1, \xi_2 = x_2, \ldots, \xi_n = x_n) = P(\xi_1 = x_1)P(\xi_2 = x_2) \cdots P(\xi_n = x_n).$$

♦ Случайные величины $\xi_1, \xi_2, \ldots, \xi_n$, имеющие непрерывное распределение, независимы, если для любых $x_1, x_2, \ldots, x_n$ имеет место равенство

$$f_{\vec{\xi}}(x_1, x_2, \ldots, x_n) = f_{\xi_1}(x_1)f_{\xi_2}(x_2) \cdots f_{\xi_n}(x_n). \tag{11.6}$$

Заметим, что в случае независимых случайных величин закон их совместного распределения полностью определяется одномерными законами распределения этих величин.

♦ *Условной плотностью распределения* величины ξ, входящей в систему $\vec{\zeta} = (\xi, \eta)$, при условии, что величина η приняла определённое значение y, называется функция

$$f_\xi(x/y) = \frac{f(x,y)}{f_\eta(y)}. \tag{11.7}$$

Пример 11.1. Два стрелка независимо друг от друга стреляют по мишени. Описать закон распределения системы случайных величин (ξ, η) — числа попаданий в мишень, если вероятность попадания в мишень для первого стрелка 0,7, для второго 0,6 (каждый из стрелков производит только один выстрел).

Решение. Так как события $(\xi = x_i)$, $(\eta = y_i)$ независимы, то вероятности p_{ij} могут быть найдены по формуле произведения вероятностей независимых событий:

$$p_{11} = P(\xi = 0, \eta = 0) = 0{,}3 \cdot 0{,}4 = 0{,}12;$$
$$p_{12} = P(\xi = 0, \eta = 1) = 0{,}3 \cdot 0{,}6 = 0{,}18;$$
$$p_{21} = P(\xi = 1, \eta = 0) = 0{,}7 \cdot 0{,}4 = 0{,}28;$$
$$p_{22} = P(\xi = 1, \eta = 1) = 0{,}7 \cdot 0{,}6 = 0{,}42.$$

Закон распределения представим в виде таблицы с двумя входами.

Таблица 1

Закон распределения системы случайных величин (ξ, η) — числа попаданий в мишень

$\xi \diagdown \eta$	0	1	$\sum$
0	0,12	0,18	0,3
1	0,28	0,42	0,7
$\sum$	0,4	0,6	1

Отметим, что сумма всех вероятностей $\sum\limits_{i,j=1}^{2} p_{ij} = 1$, сумма вероятностей по строкам равна вероятностям соответствующих значений величины ξ: $\sum\limits_{j=1}^{2} p_{ij} = p_i$, а сумма вероятностей по столбцам равна вероятностям соответствующих значений величины η: $\sum\limits_{j=1}^{2} p_{ij} = p_j$.

Пример 11.2. Двумерная случайная величина имеет плотность совместного распределения

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

график которой изображен на рис. 36. Найти функцию распределения $F(x, y)$, плотности распределения компонент $f(x)$, $f(y)$ и $P(|\xi| < 1, |\eta| < 1)$.

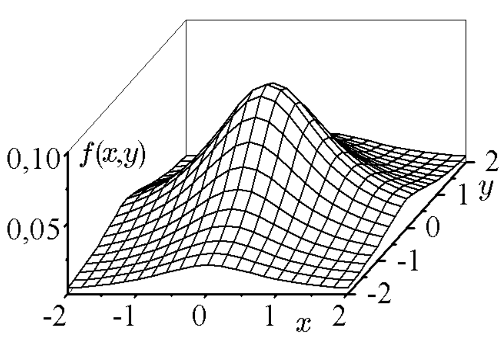


Рис. 36.

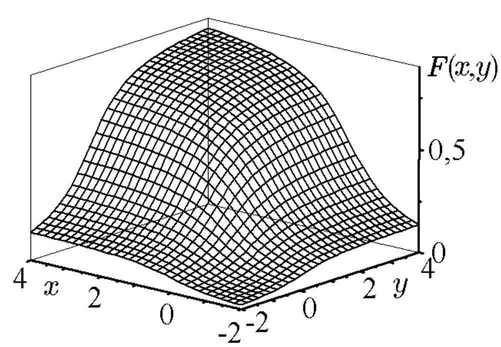


Рис. 37.

Решение. Воспользуемся формулой (11.1):

$$F_{\xi,\eta}(x,y) = \frac{1}{\pi^2}\int\limits_{-\infty}^{y}\Big(\frac{1}{1+y^2}\int\limits_{-\infty}^{x}\frac{dx}{1+x^2}\Big)dy = \frac{1}{\pi^2}\int\limits_{-\infty}^{y}\frac{1}{1+y^2}\Big(\operatorname{arctg} x + \frac{\pi}{2}\Big)dy =$$

$$= \Big(\frac{1}{\pi}\operatorname{arctg} x + \frac{1}{2}\Big)\frac{1}{\pi}\int\limits_{-\infty}^{y}\frac{dy}{1+y^2} = \Big(\frac{1}{\pi}\operatorname{arctg} x + \frac{1}{2}\Big)\Big(\frac{1}{\pi}\operatorname{arctg} y + \frac{1}{2}\Big).$$

График полученной функции распределения изображен на рис. 37. Плотности распределения каждой из компонент найдём по формуле (11.4):

$$f_\xi(x) = \int\limits_{-\infty}^{\infty} f(x,y)dy = \frac{1}{\pi^2}\frac{1}{1+x^2}\int\limits_{-\infty}^{+\infty}\Big(\frac{1}{1+y^2}\Big)dy = \frac{1}{\pi^2}\frac{1}{1+x^2}\pi = \frac{1}{\pi}\frac{1}{1+x^2};$$

$$f_\eta(y) = \frac{1}{\pi}\frac{1}{1+y^2}.$$

Заметим, что, согласно (11.6), (11.7), компоненты ξ и η являются независимыми и их найденные плотности распределения совпадают с условными.

Вероятность попадания в квадрат найдём по формуле (11.3):

$$P(|\xi|<1, |\eta|<1) = \frac{1}{\pi^2}\int\limits_{-1}^{+1}\frac{dx}{1+x^2}\int\limits_{-1}^{+1}\frac{dy}{1+y^2} = \frac{1}{\pi^2}\Big(\operatorname{arctg} x\Big|_{-1}^{+1}\Big)\Big(\operatorname{arctg} y\Big|_{-1}^{+1}\Big) = \frac{1}{4}.$$

Пример 11.3. Двумерная случайная величина $\vec{\zeta} = (\xi,\eta)$ равномерно распределена внутри треугольника с вершинами в точках $O(0,0)$, $A(1,0)$, $B(0,1)$ (рис. 38). Найти плотность распределения $f(x,y)$, плотности распределения каждой из компонент $f_\xi(x)$, $f_\eta(y)$, условные плотности распределения каждой из компонент $f_\xi(x/y)$, $f_\eta(y/x)$ и вероятность $P(\xi \geqslant 0{,}5,\ \eta < 1)$.

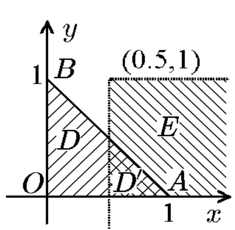


Рис. 38.

Решение. Так как система $\vec{\zeta} = (\xi,\eta)$ равномерно распределена внутри треугольника, то $f(x,y) = C$, если точка (x,y) принадлежит треугольнику OAB, и $f(x,y) = 0$ в остальных точках. Найдём C из условия нормировки:

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

Следовательно,

$$\frac{1}{C} = \iint\limits_D dx\,dy,$$

где область D — область, ограниченная сторонами треугольника OAB. Отсюда следует

$$\iint\limits_D dx\,dy = \frac{1}{2},$$

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

Плотности распределения величин ξ и η находим по формулам (свойство 5)
$$f_\xi(x) = \int\limits_{-\infty}^{\infty} f(x,y)dy = \begin{cases} \int\limits_0^{1-x} 2dy = 2y\Big|_0^{1-x} = 2(1-x), & 0 < x \leqslant 1, \\ 0, & x \notin [0,1[; \end{cases}$$
$$f_\eta(y) = \int\limits_{-\infty}^{\infty} f(x,y)dx = \begin{cases} \int\limits_0^{1-y} 2dx = 2x\Big|_0^{1-y} = 2(1-y), & 0 < y \leqslant 1, \\ 0, & x \notin [0,1[. \end{cases}$$

Условные плотности вероятностей выразим через плотность распределения и полученные плотности вероятностей величин ξ и η:
$$f_\xi(x/y) = \frac{f(x,y)}{f_\eta(y)} = \frac{2}{2(1-y)} = \frac{1}{1-y}, \quad 0 < y < 1, \quad 0 < x < 1-y;$$
$$f_\eta(y/x) = \frac{f(x,y)}{f_\xi(x)} = \frac{2}{2(1-x)} = \frac{1}{1-x}, \quad 0 < x < 1, \quad 0 < y < 1-x.$$

Вероятность $P(\xi \geqslant 0{,}5, \eta < 1)$ есть вероятность того, что случайная величина ξ примет значения, лежащие в области E (см. рис. 38):
$$P(\xi \geqslant 0{,}5, \eta < 1) = \int\limits_{0,5}^{\infty}\int\limits_{-\infty}^{1} f(x,y)dx\,dy = \iint\limits_{D'} 2dx\,dy,$$
где область D' — пересечение областей D и E, т.е. часть треугольника OAB, ограниченная слева прямой $x = 0{,}5$ (см. рис. 38). Следовательно,
$$P(\xi \geqslant 0{,}5;\ \eta < 1) = \int\limits_{0,5}^{1}\int\limits_{0}^{1-x} 2dx\,dy = -(1-x)^2\Big|_{0,5}^{1} = 0{,}25.$$

11.4. Функции случайных величин

Часто при проведении измерений мы получаем информацию не о самой случайной величине, а о некоторой функции от нее. Поэтому необходимо проводить операции с функциями от случайных величин.

Пусть ξ — случайная величина с известным законом распределения $F_\xi(x)$, а случайная величина $\eta = \varphi(\xi)$, где $\varphi(x)$ — неслучайная функция. Требуется определить закон распределения $F_\eta(x)$ случайной величины η.

Законы распределения непрерывных случайных величин ξ и $\eta = \varphi(\xi)$ взаимосвязаны в силу следующих свойств.

Теорема 11.1. *Пусть ξ — непрерывная случайная величина, имеющая функцию распределения $F_\xi(x)$ и плотность распределения $f_\xi(x)$, а функция $\varphi\colon \mathbb{R} \to \mathbb{R}$ дифференцируема и монотонна. Тогда случайная величина $\eta = \varphi(\xi)$ имеет плотность распределения*
$$f_\eta(x) = |[\varphi^{-1}(x)]'| f_\xi(\varphi^{-1}(x)). \tag{11.8}$$

Доказательство. Пусть для определённости $\varphi(x)$ — возрастающая функция. Так как $\varphi(x)$ монотонна и дифференцируема, то существует обратная к ней функция $\varphi^{-1}(x)$, которая будет также возрастающей и дифференцируемой. Тогда для функции распределения величины $\eta = \varphi(\xi)$ запишем

$$F_\eta(x) = F_{\varphi(\xi)}(x) = P(\varphi(\xi) < x) = P(\xi < \varphi^{-1}(x)) = F_\xi(\varphi^{-1}(x)) = \int\limits_{-\infty}^{\varphi^{-1}(x)} f_\xi(t)dt.$$

Проведём в интеграле замену переменных: $t = \varphi^{-1}(z)$, $dt = [\varphi^{-1}(z)]'dz$; $t_1 = -\infty$, $z_1 = -\infty$; $t_2 = \varphi^{-1}(x)$, $z_2 = x$, и получим

$$F_\eta(x) = \int\limits_{-\infty}^{x} f_\xi(\varphi^{-1}(z))[\varphi^{-1}(z)]'dz.$$

Так как $f_\xi(\varphi^{-1}(z))[\varphi^{-1}(z)]' > 0$, то по определению эта функция является плотностью распределения случайной величины η: $f_\eta(x) = [\varphi^{-1}(x)]' f_\xi(\varphi^{-1}(x))$, что и требовалось доказать. Доказательство для случая монотонно убывающей функции $\varphi(x)$ аналогично.

Пример 11.4. Найти распределение случайной величины η, гармонически (квадратично и кубично) зависящей от другой случайной величины ξ:

$$\text{а) } \eta = \sin\xi; \qquad \text{б) } \eta = \xi^2; \qquad \text{в) } \eta = \xi^3,$$

где ξ — положительно определенная случайная величина.

Решение. а) В соотношении (11.8) положим

$$\varphi(x) = \sin x, \qquad \varphi^{-1}(x) = \arcsin x, \qquad (\varphi^{-1}(x))' = \frac{1}{\sqrt{1-x^2}}$$

(последнее справедливо в соответствующих областях определения рассматриваемых функций). Тогда

$$f_\eta(x) = \frac{f_\xi(\arcsin x)}{\sqrt{1-x^2}}. \tag{11.9}$$

Аналогично из (11.8) найдём для $\varphi(x) = x^2$:

$$f_\eta(x) = \frac{1}{2\sqrt{x}} f_\xi(\sqrt{x}), \tag{11.10}$$

и для $\varphi(x) = x^3$:

$$f_\eta(x) = \frac{1}{3\sqrt[3]{x^2}} f_\xi(\sqrt[3]{x}). \tag{11.11}$$

Следствие 11.1.1. Если $\eta = a\xi + b$, то

$$f_\eta(x) = \frac{1}{|a|} f_\xi\Big(\frac{x-b}{a}\Big).$$

Доказательство непосредственно следует из (11.8).

Если ξ и $\eta = \varphi(\xi)$ — дискретные случайные величины, то ряд распределения величины η может быть получен непосредственно путем подсчёта вероятностей значений величины η через вероятности значений величины ξ по формуле

$$P(\eta = y_k) = \sum_{i:(\varphi(x_i)=y_k)} P(\xi = x_i), \tag{11.12}$$

т.е. суммирование проводится по всем значениям индекса i, для которых $\varphi(x_i) = y_k$.

Пример 11.5. Дан ряд распределения случайной величины ξ. Построить ряд распределения случайной величины $\eta = \xi^2$.

ξ	-1	0	1
P	$1/8$	$5/8$	$1/4$

Решение. Составим таблицу значений η:

ξ	-1	0	1
η	1	0	1

Видим, что η может принимать два значения: 0 и 1, причём, очевидно, $P(\eta = 0) = 5/8$, а вероятность значения $\eta = 1$ можно подсчитать по теореме сложения вероятностей: $P(\eta = 1) = 1/8 + 1/4 = 3/8$. В результате ряд распределения η имеет вид

ξ	0	1
η	$5/8$	$3/8$

Обобщим формулу (11.8) на многомерный случай. Пусть $\vec{\xi}$ — k-мерная непрерывная случайная величина с плотностью распределения $f_\xi(\vec{x}) = f_\xi(x_1, x_2, \ldots, x_k)$. Пусть многомерная случайная величина $\vec{\eta}$ определена как вектор-функция вида $\vec{\eta} = g(\vec{\xi})$, где $\eta_i = g_i(\xi_1, \ldots, \xi_k)$, $i = \overline{1, k}$. Как и ранее, пусть соответствие между $\vec{\xi}$ и $\vec{\eta}$ взаимно однозначно и существует обратное преобразование $\vec{\xi} = g^{-1}(\vec{\eta})$, такое что $\xi_i = g_i^{-1}(\eta_i, \eta_2, \ldots, \eta_k)$, $i = \overline{1, k}$. Тогда совместная плотность вероятности случайных величин η_i, $i = \overline{1, k}$, равна

$$f_\eta(x_1, \ldots, x_k) = f_\xi(g_1^{-1}(\vec{x}), g_2^{-1}(\vec{x}), \ldots, g_k^{-1}(\vec{x}))|J(\vec{x})|, \tag{11.13}$$

где $\vec{x} = \{x_1, x_2, \ldots, x_k\}$ — вектор аргументов; $J(\vec{x})$ — якобиан преобразования вида

$$J(\vec{x}) = \frac{D(g^{-1}(\vec{x}))}{D(\vec{x})} = \frac{D(g_1^{-1}(\vec{x}), g_2^{-1}(\vec{x}), \ldots, g_k^{-1}(\vec{x}))}{D(x_1, x_2, \ldots, x_k)} = \begin{vmatrix} \dfrac{\partial g_1^{-1}(\vec{x})}{\partial x_1} & \ldots & \dfrac{\partial g_1^{-1}(\vec{x})}{\partial x_k} \\ \ldots & \ldots & \ldots \\ \dfrac{\partial g_k^{-1}(\vec{x})}{\partial x_1} & \ldots & \dfrac{\partial g_k^{-1}(\vec{x})}{\partial x_k} \end{vmatrix}.$$

В частности, при линейном преобразовании $\vec{\eta} = A\vec{\xi} + \vec{\mu}$, где A — невырожденная матрица, а $\vec{\mu}$ — некоторый вектор сдвига, имеем

$$\vec{\xi} = A^{-1}(\vec{\eta} - \vec{\mu}), \qquad g^{-1}(\vec{x}) = A^{-1}(\vec{x} - \vec{\mu}), \qquad \frac{D(g_1^{-1}(\vec{x}))}{D(\vec{x})} = \det(A^{-1}) = \frac{1}{\det A}$$

и, используя (11.13), получим

$$f_\eta(\vec{x}) = \frac{1}{|\det A|} f_\xi(A^{-1}(\vec{x} - \vec{\mu})). \tag{11.14}$$

♦ Пусть $\{\xi, \eta\}$ — случайный вектор и $\zeta = \varphi(\xi, \eta)$, где $\varphi(u, v)$ — неслучайная функция, тогда функцию распределения случайной величины ζ можно, очевидно, определить как

$$F_\zeta(z) = P(\zeta < z) = \sum_{\varphi(x_i, y_j) < z}\sum P(\xi = x_i, \eta = y_j)$$

для вектора с дискретным распределением и как

$$F_\zeta(z) = P(\zeta < z) = \iint\limits_{\varphi(x,y)<z} f_{\xi,\eta}(x, y) dx\, dy$$

для вектора с непрерывным распределением.

11.5. Композиция случайных величин

Одной из наиболее важных задач при рассмотрении функций случайного вектора является так называемая задача композиции — задача определения закона суммы случайных величин по известным законам распределения самих величин.

Теорема 11.2 (формула свертки). *Если случайные величины ξ и η независимы и имеют непрерывные распределения с плотностями $f_\xi(x)$ и $f_\eta(y)$, то*

$$f_{\xi+\eta}(z) = \int\limits_{-\infty}^{\infty} f_\xi(x) f_\eta(z - x) dx = \int\limits_{-\infty}^{\infty} f_\eta(y) f_\xi(z - y) dy. \tag{11.15}$$

Доказательство. По определению

$$F_{\xi+\eta}(z) = P(\xi + \eta < z) = \iint\limits_{\xi+\eta<z} f_{\xi\eta}(x, y) dx\, dy = \iint\limits_{\xi+\eta<z} f_\xi(x) f_\eta(y) dx\, dy =$$

$$= \int\limits_{-\infty}^{\infty} dx \int\limits_{-\infty}^{z-x} f_\xi(x) f_\eta(y) dy.$$

Сделаем в интеграле замену переменных: $t = x + y$. Тогда

$$F_{\xi+\eta}(z) = \int\limits_{-\infty}^{\infty} dx \int\limits_{-\infty}^{z} f_\xi(x) f_\eta(t - x) dx = \int\limits_{-\infty}^{z} dt \int\limits_{-\infty}^{\infty} f_\xi(x) f_\eta(t - x) dx =$$

$$= \int\limits_{-\infty}^{z} \left(\int\limits_{-\infty}^{\infty} f_\xi(x) f_\eta(t - x) dx\right) dt = \int\limits_{-\infty}^{z} f_{\xi+\eta}(t) dt.$$

Следовательно,

$$f_{\xi+\eta}(z) = \int\limits_{-\infty}^{\infty} f_\xi(x) f_\eta(z - x) dx,$$

что и требовалось доказать.

Аналогичная теорема имеет место и для дискретных случайных величин.

Теорема 11.3. *Если случайные величины ξ и η независимы, то*

$$P(\xi + \eta = z_k) = \sum_i P(\xi = x_i)P(\eta = z_k - x_i). \tag{11.16}$$

Доказательство аналогично доказательству предыдущей теоремы.

Нетрудно заметить, что функция распределения разности случайных величин ξ и η определяется соотношением

$$f_{\xi-\eta}(z) = \int\limits_{-\infty}^{\infty} f_\xi(x) f_\eta(z + x) dx.$$

♦ Распределение называется *устойчивым по суммированию*, если сумма двух величин, имеющих одно и то же распределение, имеет то же распределение.

◊ Наряду с суммой и разностью случайных величин в приложениях могут возникать их произведение и частное. Так, плотность распределения произведения $\zeta = \xi\eta$ определяется соотношением (см., например, [14], где рассмотрено большое количество задач по преобразованию функции плотности распределения вероятности)

$$f_\zeta(x) = \int\limits_{-\infty}^{0} \frac{1}{t} f_\xi(t) f_\eta\Big(\frac{x}{t}\Big) dt - \int\limits_{0}^{\infty} \frac{1}{t} f_\xi(t) f_\eta\Big(\frac{x}{t}\Big) dt, \tag{11.17}$$

а плотность распределения частного — соотношением

$$f_\zeta(x) = \int\limits_{-\infty}^{0} t f_\xi(tx) f_\eta(t) dt - \int\limits_{0}^{\infty} t f_\xi(tx) f_\eta(x) dt. \tag{11.18}$$

12. Числовые характеристики случайной величины

Законы распределения случайной величины полностью описывают случайную величину с вероятностной точки зрения. Однако эти функции не всегда известны, да и во многих задачах они не нужны. Зачастую достаточно знать только основные числовые параметры, характеризующие случайную величину, такие как интервал, в котором находится большинство значений случайной величины; среднее значение, относительно которого группируются значения случайной величины; разброс значений случайной величины относительно среднего значения и т.д. Такие характеристики называются *числовыми характеристиками случайной величины*.

12.1. Математическое ожидание случайной величины

Рассмотрим дискретную случайную величину ξ, имеющую возможные значения $\{x_k\}_{k=1}^n$ с вероятностями $\{P_k\}_{k=1}^n$ (для случайной величины, принимающей счётное множество значений, $n = \infty$).

♦ *Математическим ожиданием*, или *средним значением*, дискретной случайной величины ξ называется число

$$M(\xi) = \sum_{i=1}^{n} x_i P_i. \tag{12.1}$$

Математическое ожидание случайной величины может и не существовать, если при $n = \infty$ соответствующая сумма (12.1) расходится. Поэтому предполагается, что при $n = \infty$ ряд (12.1) сходится абсолютно.

Рассмотрим непрерывную случайную величину ξ с плотностью вероятности $f(x)$.

♦ *Математическим ожиданием* непрерывной случайной величины ξ называется число, равное интегралу

$$M(\xi) = \int_{-\infty}^{\infty} x f(x) dx, \tag{12.2}$$

если он сходится.

В противном случае считают, что случайная величина ξ не имеет математического ожидания.

Формула (12.2) имеет такой же вероятностный смысл, как и для дискретной случайной величины: математическое ожидание одномерной случайной величины — это некоторое её среднее значение.

◊ В литературе наряду с (12.1), (12.2) используются обозначения

$$M(\xi) = M\xi = m_\xi = M[\xi] = E_\xi = \bar{\xi} = \langle \xi \rangle.$$

В тех случаях, когда возможные значения случайной величины располагаются не по всей оси Ox, а принадлежат интервалу $[a, b]$, можно записать

$$M(\xi) = \bar{\xi} = \int_a^b x f(x) dx.$$

Из определения вытекают следующие свойства математического ожидания, доказательство которых будем проводить на примере либо дискретной, либо непрерывной случайной величины.

Свойство 1. Математическое ожидание постоянной C равно этой постоянной: $M(C) = C$.

Доказательство. Постоянную C можно рассматривать как случайную величину ξ, которая может принимать только одно значение C с вероятностью, равной единице. Поэтому

$$M(\xi) = C \cdot 1 = C.$$

Свойство 2. Для произвольной функции $\eta = \varphi(\xi)$ случайного аргумента ξ

$$M(\eta) = M(\varphi(\xi)) = \begin{cases} \sum\limits_{i=1}^{n} \varphi(x_i) P_i & \text{для дискретной случайной величины;} \\ \int\limits_{-\infty}^{\infty} \varphi(x) f(x) dx & \text{для непрерывной случайной величины.} \end{cases} \tag{12.3}$$

Доказательство. Пусть функция $\varphi(\xi)$ принимает значения $c_1, c_2, \ldots$ с вероятностями $P(\varphi(\xi) = c_m) = \sum\limits_{k:\varphi(x_k)=c_m} P(\xi = x_k)$. Тогда (см. (11.12))

$$M(\eta) = M(\varphi(\xi)) = \sum_m c_m P(\varphi(\xi) = c_m) = \sum_m c_m \sum_{k:(\varphi(x_k)=c_m)} P(\xi = x_k) =$$

$$= \sum_m \sum_{k:(\varphi(x_k)=c_m)} \varphi(x_k) P(\xi = x_k) = \sum_i \varphi(x_i) P(\xi = x_i).$$

Аналогично для непрерывной случайной величины из (11.8) запишем

$$M(\varphi(\xi)) = M(\eta) = \int\limits_{-\infty}^{\infty} y f_\eta(y) dy = \int\limits_{-\infty}^{\infty} y |[\varphi^{-1}(y)]'| f_\xi(\varphi^{-1}(y)) dy.$$

Сделав замену переменных $y = \varphi(x)$, $x = \varphi^{-1}(y)$, $dx = |\varphi^{-1}(y)|' dy$, получим (12.3).

Свойство 2 может быть обобщено и на функцию случайного вектора. Так, для функции двух случайных величин $\varphi(\xi, \eta)$ запишем

$$M(\varphi(\xi,\eta)) = \begin{cases} \sum\limits_i \sum\limits_j \varphi(x_i, y_j) P(\xi = x_i, \eta = y_j) & \text{для дискретного вектора;} \\ \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \varphi(x,y) f_{\xi,\eta}(x,y) dx\, dy & \text{для непрерывного вектора.} \end{cases}$$

Свойство 3. Постоянный множитель можно выносить за знак математического ожидания, т.е. математическое ожидание произведения постоянной величины на случайную величину равно произведению этой постоянной на математическое ожидание случайной величины:

$$M(C\xi) = CM(\xi). \tag{12.4}$$

Доказательство. Положим $\varphi(x) = Cx$ и воспользуемся предыдущим свойством:

$$M(C\xi) = \sum_{i=1}^n C x_i P_i = C \sum_{i=1}^n x_i P_i = CM(\xi).$$

Свойство 4. Математическое ожидание суммы нескольких случайных величин равно сумме математических ожиданий этих величин:

$$M(\xi_1 + \xi_2 + \ldots + \xi_n) = M(\xi_1) + M(\xi_2) + \ldots + M(\xi_n). \tag{12.5}$$

Доказательство проведём для суммы двух непрерывных величин. Согласно свойству 2,

$$M(\xi + \eta) = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} (x + y) f_{\xi,\eta}(x,y) dx\, dy =$$

$$= \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} x f_{\xi,\eta}(x,y) dx\, dy + \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} y f_{\xi,\eta}(x,y) dx\, dy =$$

$$= \int_{-\infty}^{\infty} x f_\xi(x)dx + \int_{-\infty}^{\infty} y f_\eta(y)dy = M(\xi) + M(\eta).$$

Здесь мы воспользовались тем, что

$$f_\xi(x) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dy \quad \text{и} \quad f_\eta(y) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dx.$$

Таким образом, утверждение доказано.

Свойство 5. Математическое ожидание суммы постоянной и случайной величин равно сумме постоянной величины и математического ожидания случайной величины:

$$M(\xi + C) = \sum_{i=1}^{n}(x_i + C)P_i = \sum_{i=1}^{n} x_i P_i + C\sum_{i=1}^{n} P_i = M(\xi) + C. \tag{12.6}$$

Доказательство. Заметим, что постоянную C можно рассматривать как случайную величину ξ, которая может принимать только одно значение C с вероятностью, равной единице. Поэтому (12.6) является частным случаем соотношения (12.5).

Свойство 6. Математическое ожидание произведения двух независимых случайных величин равно произведению математических ожиданий этих величин:

$$M(\xi\eta) = M(\xi)M(\eta). \tag{12.7}$$

Доказательство проведём для непрерывных величин. По определению,

$$M(\xi\eta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{\xi,\eta}(x,y)dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_\xi(x) f_\eta(y)dx\,dy =$$

$$= \int_{-\infty}^{\infty} x f_\xi(x)dx \int_{-\infty}^{\infty} y f_\eta(y)dy = M(\xi)M(\eta),$$

что и требовалось доказать.

♦ Случайная величина $\hat{\xi} = \xi - M(\xi)$ (отклонение случайной величины от её математического ожидания) называется *центрированной случайной величиной*.

Центрирование случайной величины равносильно переносу начала координат в среднюю, «центральную» точку, абсцисса которой равна математическому ожиданию.

Отметим, что $M(\hat{\xi}) = 0$. В самом деле,

$$M(\hat{\xi}) = M(\xi - M(\xi)) = M(\xi) - M(\xi) = 0.$$

Математическое ожидание является важнейшей из характеристик положения. Среди прочих характеристик положения выделяют моду и медиану случайной величины.

♦ *Модой* x_{mod} дискретной случайной величины называют ее наиболее вероятное значение.

♦ *Медианой случайной величины* ξ называется такое значение x_{med}, для которого $P(\xi < x_{\text{med}}) = P(\xi \geqslant x_{\text{med}})$, то есть медиана является решением уравнения $F(x_{\text{med}}) = 1/2$.

Пример 12.1. Найти математическое ожидание $M(\xi)$ дискретной случайной величины ξ, заданной законом распределения

ξ	-5	2	3	4
P	0,4	0,3	0,1	0,2

Решение. По определению, $M(\xi) = -5 \cdot 0{,}4 + 2 \cdot 0{,}3 + 3 \cdot 0{,}1 + 4 \cdot 0{,}2 = -0{,}3$.

12.2. Дисперсия случайной величины и её свойства

Математическое ожидание случайной величины характеризует её среднее значение, дисперсия же служит характеристикой разброса значений случайной величины относительно математического ожидания.

♦ *Дисперсией* случайной величины ξ называется математическое ожидание квадрата отклонения случайной величины от её среднего значения и обозначается

$$D(\xi) = D\xi = M(\xi - M(\xi))^2. \tag{12.8}$$

Из определения (12.8) следует, что для непрерывной случайной величины ξ дисперсия определяется интегралом

$$D(\xi) = \int\limits_{-\infty}^{\infty} (\xi - M(\xi))^2 f(x)dx, \tag{12.9}$$

если он сходится, а для дискретной случайной величины — суммой

$$D(\xi) = \sum_{i=1}^{n} (x_i - M(\xi))^2 P(\xi = x_i). \tag{12.10}$$

Дисперсия $D(\xi)$ имеет размерность квадрата случайной величины, для характеристики рассеивания же удобнее использовать величину, размерность которой совпадает с размерностью случайной величины.

♦ *Среднеквадратичным отклонением* (*стандартным отклонением*, или *стандартом*) случайной величины ξ называется корень квадратный из дисперсии:

$$\sigma = \sqrt{D(\xi)}.$$

Если воспользоваться свойствами математического ожидания, то из (12.8) получим

$$M(\xi - M(\xi))^2 = M(\xi^2 - 2\xi M\xi + (M\xi)^2) = M\xi^2 - 2M\xi M\xi + (M\xi)^2.$$

Отсюда следует

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

Свойства дисперсии

Свойство 1. Для любой случайной величины ξ дисперсия $D(\xi) \geqslant 0$.

Справедливость утверждения следует непосредственно из определения.

Свойство 2. Если C — постоянная, то $D(C) = 0$.

Действительно,

$$D(C) = M([C - M(C)]^2) = M((C - C)^2) = M(0) = 0.$$

Свойство 3. Если C — постоянная, то $D(C\xi) = C^2 D\xi$, $D(\xi + C) = D(\xi)$.

Действительно, согласно свойствам математического ожидания,

$$D(C\xi) = M((C\xi)^2) - M^2(C\xi) = C^2 M(\xi^2) - C^2 M^2(\xi) = C^2 D(\xi).$$

Аналогично

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

Свойство 4. Если случайные величины ξ и η независимы, то

$$D(\xi + \eta) = D\xi + D\eta.$$

Доказательство. Согласно свойствам математического ожидания,

$$D(\xi + \eta) = M((\xi + \eta)^2) - M^2(\xi + \eta) = M(\xi^2 + 2\xi\eta + \eta^2) - [M(\xi) + M(\eta)]^2 = \\ = M(\xi^2) + 2M(\xi\eta) + M(\eta^2) - M^2(\xi) - 2M(\xi)M(\eta) - M^2(\eta) = \\ = M(\xi^2) - M^2(\xi) + M(\eta^2) - M^2(\eta) = D(\xi) + D(\eta)$$

12.3. Моменты случайной величины и их свойства

Помимо математического ожидания и дисперсии, в качестве числовых характеристик случайных величин применяются моменты. На практике чаще всего применяют начальные и центральные моменты.

♦ *Начальным моментом* k-го порядка случайной величины ξ называется математическое ожидание величины ξ^k, т.е.

$$\alpha_k(\xi) = M(\xi^k).$$

Для дискретной случайной величины начальный момент определяется соотношением

$$\alpha_k = \sum_{i=1}^{n} x_i^k P_i,$$

а для непрерывной случайной величины ξ начальный момент k-го порядка равен интегралу

$$\alpha_k = \int_{-\infty}^{\infty} x^k f(x)dx, \qquad k = \overline{0, \infty}.$$

Начальный момент первого порядка случайной величины есть ни что иное, как математическое ожидание: $\alpha_1 = M(\xi)$.

♦ Число $\mu_k(\xi) = M(\xi - M(\xi))^k$ называется *центральным моментом* k-го порядка. Для дискретной случайной величины центральный момент k-го порядка вычисляется по формуле

$$\mu_k = \mu_k(\xi) = \sum_{i=1}^{n} [x_i - M(\xi)]^k P_i.$$

Центральный момент k-го порядка непрерывной случайной величины равен интегралу

$$\mu_k = \int\limits_{-\infty}^{\infty} [x - M(\xi)]^k f(x) dx, \qquad k = \overline{0, \infty}.$$

♢ Нетрудно заметить, что $\mu_1 = 0$, $\mu_2 = D(\xi)$.

♦ *Начальным моментом* порядков k, s системы случайных величин (ξ, η) называется математическое ожидание

$$\alpha_{k,s} = M(\xi^k \eta^s).$$

Для системы величин, имеющих дискретное распределение,

$$\alpha_{k,s} = \sum_i \sum_j x_i^k y_j^s P(\xi = x_i, \eta = y_j);$$

для системы величин, имеющих непрерывное распределение,

$$\alpha_{k,s} = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} x^k y^s f_{\xi,\eta}(x, y) dx\, dy.$$

Первые начальные моменты совпадают с математическими ожиданиями величин ξ и η:

$$\alpha_{1,0} = M[\xi^1 \eta^0] = M\xi;$$
$$\alpha_{0,1} = M[\xi^0 \eta^1] = M\eta.$$

Действительно, например, для непрерывных величин:

$$\alpha_{1,0} = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} x^1 y^0 f_{\xi,\eta}(x, y) dx\, dy = \int\limits_{-\infty}^{\infty} x\, dx \int\limits_{-\infty}^{\infty} f_{\xi,\eta}(x, y) dy = \int\limits_{-\infty}^{\infty} x f_\xi(x) dx = M(\xi).$$

♦ *Центральным моментом* порядков k, s системы случайных величин (ξ, η) называется математическое ожидание

$$\mu_{k,s} = M\big([\xi - M(\xi)]^k [\eta - M(\eta)]^s\big). \tag{12.11}$$

Очевидно, что

$$\mu_{k,s} = \sum_i \sum_j [x_i - M(\xi)]^k [y_j - M(\eta)]^s P(\xi = x_i, \eta = y_j)$$

для системы величин, имеющих дискретное распределение, и

$$\mu_{k,s} = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} [x - M(\xi)]^k [y - M(\eta)]^s f_{\xi,\eta}(x, y) dx\, dy$$

для системы величин, имеющих непрерывное распределение.

Можно заметить, что вторые центральные моменты $\mu_{2,0}$ и $\mu_{0,2}$ есть дисперсии величин ξ и η: $\mu_{2,0} = D(\xi)$, $\mu_{0,2} = D(\eta)$.

♢ По определению, плотность вероятности однозначно определяет все моменты. Справедливо и обратное утверждение.

Теорема 12.1. *Пусть $f(x)$ и $g(x)$ — плотности распределения вероятности случайных величин ξ и ζ и*

$$\alpha_k = \int\limits_{-\infty}^{\infty} f(x)x^k dx, \quad \beta_k = \int\limits_{-\infty}^{\infty} g(x)x^k dx.$$

Тогда если $\alpha_k = \beta_k$, $k = \overline{0, \infty}$, и функция $\varphi(x) = f(x) - g(x)$ разложима в ряд Тейлора, т.е.

$$\varphi(x) = f(x) - g(x) = \sum_{k=0}^{\infty} C_k x^k,$$

то $f(x) = g(x)$.

Доказательство. Рассмотрим интеграл

$$I = \int\limits_{-\infty}^{\infty} \big(f(x) - g(x)\big)^2 dx.$$

Очевидно, что $I > 0$ и

$$I = \int\limits_{-\infty}^{\infty} f(x)(f(x) - g(x))dx - \int\limits_{-\infty}^{\infty} g(x)(f(x) - g(x))dx = \sum_{k=1}^{\infty} C_k\alpha_k - \sum_{k=1}^{\infty} C_k\beta_k = 0.$$

Следовательно, $f(x) = g(x)$.

Если случайная величина распределена симметрично относительно математического ожидания, то все центральные моменты нечётных порядков равны нулю. Для характеристики степени отклонения распределения от симметричного используют центральный момент третьего порядка.

♦ Величина

$$A = A(\xi) = \frac{\mu_3}{\sigma_\xi^3}$$

называется *коэффициентом асимметрии* или коэффициентом скошенности.

Коэффициент асимметрии характеризует степень асимметрии распределения случайной величины относительно ее математического ожидания. Для симметричных распределений $A = 0$. Если пик графика функции $f(x)$ смещен в сторону малых значений («хвост» на графике функции $f(x)$ справа), то $A > 0$. В противном случае $A < 0$ (см. рис. 39).

Четвертый центральный момент служит для характеристики островершинности распределения.

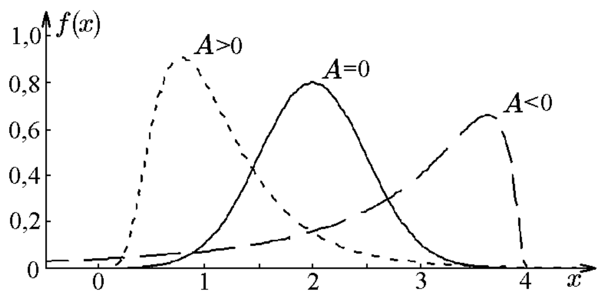


Рис. 39.

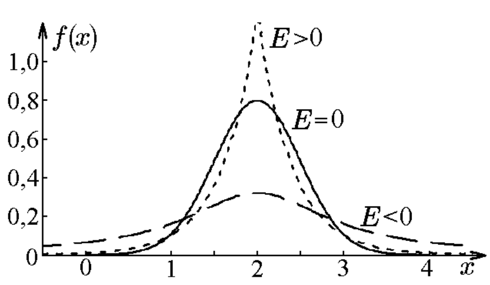


Рис. 40.

♦ Величина

$$E = E(\xi) = \frac{\mu_4}{\sigma_\xi^4} - 3$$

называется *эксцессом* случайной величины или *коэффициентом островершинности.*

Эксцесс показывает, насколько распределение отличается от так называемого нормального распределения, для которого $E = 0$.

Коэффициент эксцесса является мерой остроты графика функции плотности распределения $f(x)$ (см. рис. 40).

♦ *Квантилью* порядка β, отвечающей заданной вероятности β, распределения непрерывной случайной величины ξ называется действительное число τ_β, удовлетворяющее уравнению $P(\xi < \tau_\beta) = \beta$ (см. рис. 41).

Квантиль порядка $1/2$ является *медианой.*

Значения $\tau_{0,75}$ и $\tau_{0,25}$ называются соответственно верхней и нижней *квартилями.* Квартильный размах, равный разности верхней и нижней квартилей, представляет собой интервал вокруг медианы, в который случайная величина ξ попадает с вероятностью 0,5.

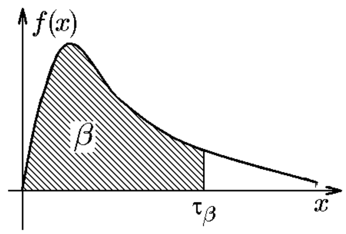


Рис. 41.

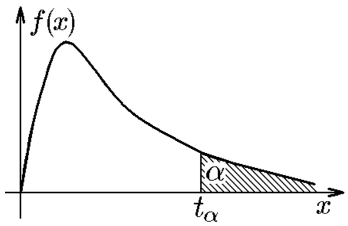


Рис. 42.

♦ *Критической точкой* порядка α распределения непрерывной случайной величины ξ называется действительное число t_α, удовлетворяющее уравнению $P(\xi \geqslant t_\alpha) = \alpha$ (см. рис. 42).

Очевидно, что квантиль порядка β совпадает с критической точкой порядка $\alpha = 1 - \beta$, поскольку

$$P(\xi \geqslant t_\alpha) = 1 - P(\xi \leqslant t_\alpha) = \alpha.$$

12.4. Ковариация и ее свойства

Рассмотрим систему случайных величин

$$\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_n).$$

♦ *Корреляционным моментом* или *ковариацией* случайных величин ξ_i и ξ_j называется величина

$$K_{ij} = \operatorname{cov}(\xi_i \xi_j) = M\big[(\xi_i - M(\xi_i))(\xi_j - M(\xi_j))\big].$$

Нетрудно заметить, что

$$K_{ii} = D_i, \quad K_{ij} = K_{ji}.$$

Корреляционные моменты и дисперсию удобно записывать в виде матрицы

$$\operatorname{cov}(\vec{\xi}) = A(\vec{\xi}) = \|K_{ij}\|_{n\times n} = M[(\vec{\xi} - M(\vec{\xi}))(\vec{\xi} - M(\vec{\xi}))^\intercal]. \tag{12.12}$$

Матрица (12.12) называется *ковариационной матрицей системы* или *матрицей ковариаций.*

Для системы случайных величин (ξ, η), имеющих дискретное распределение, корреляционный момент определяется соотношением

$$K_{\xi\eta} = \sum_i \sum_j (x_i - M(\xi))(y_j - M(\eta))p_{ij},$$

а для системы непрерывно распределённых случайных величин — соотношением

$$K_{\xi\eta} = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} (x - M(\xi))(y - M(\eta))f(x, y)dx\, dy.$$

Свойства ковариации

Свойство 1. Для ковариации случайных величин ξ и η справедлива формула

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

Свойство 2. Если случайные величины ξ и η независимы, то $\operatorname{cov}(\xi, \eta) = 0$. Обратное утверждение в общем случае неверно, т.е. из того, что $K_{\xi\eta} = 0$, не следует, что величины ξ и η независимы.

Свойство 3. Справедливо соотношение $M(\xi\eta) = M(\xi)M(\eta) + \operatorname{cov}(\xi, \eta)$.

Свойство 4. Справедливо соотношение $D(\xi + \eta) = D(\xi) + D(\eta) + 2\operatorname{cov}(\xi, \eta)$.

Ковариация зависит от размерности случайных величин ξ и η, поэтому удобнее ввести безразмерную величину и использовать ее в качестве характеристики взаимного влияния случайных величин.

♦ Величина

$$\rho(\xi, \eta) = \rho_{\xi,\eta} = \frac{\operatorname{cov}(\xi, \eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}}$$

(при условии, что $D(\xi)$, $D(\eta)$ существуют и не равны нулю) называется *коэффициентом корреляции* случайных величин ξ, η .

♦ Матрица $r = \|\rho_{ij}\|$, где

$$\rho_{ij} = \frac{K_{ij}}{\sigma_i \sigma_j}, \quad \sigma_i = \sqrt{D_i} = \sqrt{K_{ii}}, \quad i, j = \overline{1, n},$$

называется *корреляционной матрицей системы.*

Свойства коэффициента корреляции

Ограничим рассмотрение свойствами коэффициента корреляции для двух случайных величин.

Свойство 1. Справедливо соотношение $|\rho_{\xi,\eta}| \leqslant 1$.

Доказательство. Воспользуемся свойствами математического ожидания и дисперсии:

$$\rho(\xi, \eta) = \frac{\operatorname{cov}(\xi, \eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}} = \frac{M(\hat{\xi}\hat{\eta})}{\sigma_\xi \sigma_\eta} = M\Big(\frac{\hat{\xi}}{\sigma_\xi}\frac{\hat{\eta}}{\sigma_\eta}\Big) = M(\xi_0 \eta_0),$$

где

$$\xi_0 = \frac{\hat{\xi}}{\sigma_x} = \frac{\xi - M(\xi)}{\sigma_x}, \quad \eta_0 = \frac{\eta - M(\eta)}{\sigma_y}$$

— стандартизированные случайные величины. Теперь воспользуемся неравенством $ab \leqslant (a^2+b^2)/2$ и свойством монотонности: если $\xi(\omega_i) \leqslant \eta(\omega_i)$ для всех $\omega_i \in \Omega$, то $M(\xi) \leqslant M(\eta)$. Получим

$$M(\xi_0\eta_0) \leqslant M\Big(\frac{1}{2}[\xi_0^2 + \eta_0^2]\Big) = \frac{1}{2}[M(\xi_0^2) + M(\eta_0^2)] = \frac{1}{2}[D(\xi_0) + D(\eta_0)] = 1.$$

Аналогично, используя неравенство $ab \geqslant -(a^2+b^2)/2$, можно показать, что $\rho(\xi,\eta) \geqslant -1$.

Свойство 2. Если величины ξ и η связаны линейной зависимостью: $\eta = a\xi + b$, то коэффициент корреляции $\rho_{\xi,\eta} = 1$, если $a > 0$, и $\rho_{\xi,\eta} = -1$, если $a < 0$.

Доказательство. Имеем $\rho_{\xi,\eta} = K_{\xi,\eta}/(\sigma_\xi\sigma_\eta)$. Используя свойства математического ожидания и дисперсии, найдём

$$\begin{gathered}K_{\xi,\eta} = M(\hat{\xi}\hat{\eta}) = M([\xi - M(\xi)][\eta - M(\eta)]) = \\ = M([\xi - M(\xi)][a\xi + b - M(a\xi + b)]) = aM([\xi - M(\xi)]^2) = aD(\xi); \\ \sigma_\eta = \sqrt{D(\eta)} = \sqrt{D(a\xi + b)} = \sqrt{a^2 D(\xi)} = |a|\sigma_\xi.\end{gathered}$$

Следовательно,

$$\rho_{\xi,\eta} = \frac{K_{\xi,\eta}}{\sigma_\xi\sigma_\eta} = \frac{aD(\xi)}{|a|\sigma_\xi^2} = \frac{a}{|a|} = \begin{cases} 1, & a > 0; \\ -1, & a < 0. \end{cases}$$

Свойство 3. Если коэффициент корреляции $\rho = \pm 1$, то величины ξ и η линейно зависимы.

Доказательство. Пусть $\rho_{\xi,\eta} = 1$, но это означает, что в формуле $M(\xi_0\eta_0) \leqslant M((\xi_0^2 + \eta_0^2)/2)$ стоит знак равенства, или $M((\xi_0 - \eta_0)^2) = 0$, что возможно лишь в том случае, если величина ξ_0 с вероятностью 1 равна величине η_0. Следовательно,

$$\frac{\xi - M(\xi)}{\sigma_\xi} = \frac{\eta - M(\eta)}{\sigma_\eta}, \quad \eta = \frac{\sigma_\eta}{\sigma_\xi}\xi - \frac{\sigma_\eta}{\sigma_\xi}M(\xi) + M(\eta) = a\xi + b.$$

♦ Случайные величины ξ и η называются *некоррелированными*, если $\rho_{\xi,\eta} = 0$, *положительно коррелированными*, если $\rho_{\xi,\eta} > 0$, и *отрицательно коррелированными*, если $\rho_{\xi,\eta} < 0$.

Таким образом, коэффициент корреляции характеризует степень линейной зависимости между величинами ξ и η, причём если $\rho_{\xi,\eta} > 0$, то возрастанию ξ соответствует возрастание в среднем (условного математического ожидания) η, а если $\rho_{\xi,\eta} < 0$, то возрастанию ξ соответствует убывание в среднем η.

Свойства матрицы ковариаций

Свойство 1. Матрица ковариаций $A(\vec{\xi})$ случайного вектора $\vec{\xi}$ неотрицательно определена.

Доказательство. Пусть $X = \{x_1, x_2, \ldots, x_n\}$ — произвольный вектор, $\vec{\xi}$ — случайный центрированный вектор, $A = \operatorname{cov}(\vec{\xi})$. Тогда

$$X^\intercal AX = X^\intercal M(\vec{\xi}\vec{\xi}^\intercal)X = M(X^\intercal\vec{\xi}(X^\intercal\vec{\xi})^\intercal) = M(X^\intercal\vec{\xi})^2 \geqslant 0,$$

что и требовалось доказать.

Заметим, что равенство возможно лишь в случае, если $X^\intercal\vec{\xi} = 0$ с вероятностью единица, т.е. при условии, что величины ξ_i связаны линейной зависимостью.

Из этого свойства вытекает, что определитель матрицы ковариаций и все её собственные значения неотрицательны. Причем $\det A = 0$ тогда и только тогда, когда координаты $\vec{\xi}$ связаны линейной зависимостью.

Заметим также, что из данного свойства матрицы ковариаций легко получить рассмотренные выше свойства коэффициента корреляции. Действительно, пусть $\vec{\xi} = \{\xi_1, \xi_2\}$, тогда

$$A = \operatorname{cov}(\vec{\xi}) = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix}.$$

Так как $\det A \geqslant 0$, то $K_{11}K_{22} - K_{12}^2 \geqslant 0$ или $\rho_{12}^2 \leqslant 1$, причем $|\rho_{12}| = 1$ тогда и только тогда, когда $\det A = 0$, т.е. если ξ_1 и ξ_2 линейно зависимы.

Свойство 2. Если вектор $\vec{\eta}$ получен путем линейного преобразования вектора $\vec{\xi}$: $\vec{\eta} = C\vec{\xi} + \vec{\alpha}$, где C — произвольная матрица размера $n \times n$, то

$$\operatorname{cov}(\vec{\eta}) = C\operatorname{cov}(\vec{\xi})C^\intercal.$$

Доказательство. По определению,

$$\operatorname{cov}(\vec{\eta}) = M((\vec{\eta} - M(\vec{\eta}))(\vec{\eta} - M(\vec{\eta}))^\intercal).$$

Найдём $M(\vec{\eta})$:

$$M(\vec{\eta}) = M(C\vec{\xi} + \vec{\alpha}) = CM(\vec{\xi}) + \vec{\alpha}.$$

Тогда

$$\begin{gathered}\operatorname{cov}(\vec{\eta}) = M((C\vec{\xi} + \vec{\alpha} - CM(\vec{\xi}) - \vec{\alpha})(C\vec{\xi} + \vec{\alpha} - CM(\vec{\xi}) - \vec{\alpha})^\intercal) = \\ = M((C(\vec{\xi} - M(\vec{\xi})))(C(\vec{\xi} - M(\vec{\xi})))^\intercal) = \\ = CM((\vec{\xi} - M(\vec{\xi}))(\vec{\xi} - M(\vec{\xi}))^\intercal)C^\intercal = C\operatorname{cov}(\vec{\eta})C^\intercal.\end{gathered}$$

Пример 12.2. Закон распределения системы случайных величин (ξ, η) задан в виде таблицы

$\xi\diagdown\eta$	1	2
-1	0,3	0,2
3	0,1	0,4

Определить математические ожидания и дисперсии для каждой из величин, входящих в систему, а также вычислить корреляционный момент величин ξ и η и коэффициент корреляции.

Решение. Найдём математические ожидания величин ξ и η:

$$M(\xi) = \sum_{i=1}^{2}\sum_{j=1}^{2} x_i p_{ij} = -1 \cdot 0{,}3 + (-1) \cdot 0{,}2 + 3 \cdot 0{,}1 + 3 \cdot 0{,}4 = 1;$$

$$M(\eta) = \sum_{i=1}^{2}\sum_{j=1}^{2} y_j p_{ij} = 1 \cdot 0{,}3 + 1 \cdot 0{,}1 + 2 \cdot 0{,}2 + 2 \cdot 0{,}4 = 1{,}6.$$

Дисперсии найдём по формулам

$$D(\xi) = M(\xi^2) - M^2(\xi), \qquad D(\eta) = M(\eta^2) - M^2(\eta).$$

Для этого определим

$$M(\xi^2) = \sum_{i=1}^{2}\sum_{j=1}^{2} x_i^2 p_{ij} = (-1)^2 \cdot 0{,}3 + (-1)^2 \cdot 0{,}2 + 3^2 \cdot 0{,}1 + 3^2 \cdot 0{,}4 = 5;$$

$$M(\eta^2) = \sum_{i=1}^{2}\sum_{j=1}^{2} y_j^2 p_{ij} = 1^2 \cdot 0{,}3 + 1^2 \cdot 0{,}1 + 2^2 \cdot 0{,}2 + 2^2 \cdot 0{,}4 = 2{,}8.$$

Следовательно,

$$D(\xi) = 5 - 1^2 = 4; \qquad D(\eta) = 2{,}8 - 1{,}6^2 = 0{,}24.$$

Корреляционный момент найдём по формуле

$$K_{\xi,\eta} = M(\xi\eta) - M(\xi)M(\eta),$$

для этого определим

$$M(\xi\eta) = \sum_{i=1}^{2}\sum_{j=1}^{2} x_i y_j p_{ij} = (-1) \cdot 1 \cdot 0{,}3 + (-1) \cdot 2 \cdot 0{,}2 + 3 \cdot 1 \cdot 0{,}1 + 3 \cdot 2 \cdot 0{,}4 = 2.$$

Отсюда

$$K_{\xi,\eta} = 2 - 1{,}6 \cdot 1 = 0{,}4; \qquad \rho(\xi, \eta) = \frac{0{,}4}{\sqrt{4}\sqrt{0{,}24}} \approx 0{,}408.$$

Пример 12.3. Проводится проверка качества партии изделий. Из отобранных 5 изделий ξ изделий отвечают стандарту; среди них η ($\eta \leqslant 3$) — высшего качества. Совместное распределение системы (ξ, η) задано двумерной таблицей

$\eta_j \backslash \xi_i$	0	1	2	3	4	5
0	0,02	0,08	0,1	0,06	0,04	0,02
1	0	0,03	0,07	0,09	0,1	0,12
2	0	0	0,02	0,04	0,06	0,08
3	0	0	0	0,01	0,02	0,04

Найти одномерные законы распределения каждой из величин ξ и η. Определить математические ожидания, среднеквадратичные отклонения величин ξ и η и их

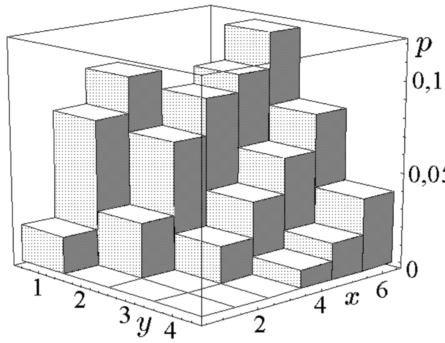


Рис. 43.

Решение. Проверим выполнение условия нормировки:

$$\sum_{m=1}^{4}\sum_{n=1}^{6} p_{mn} = 1.$$

Графиком вероятностей распределения является гистограмма системы (ξ, η) (рис. 43).

Возможен плоский вариант гистограмм — наложение построчных распределений (см. рис. 44).

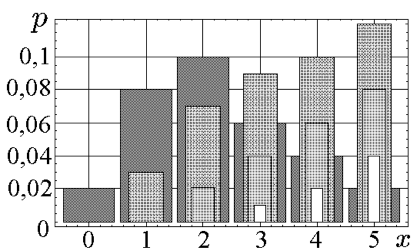


Рис. 44.

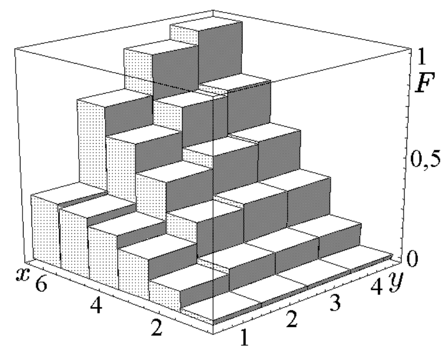


Рис. 45.

Функция распределения вектора $\vec{\zeta} = (\xi, \eta)$ определяется формулой

$$F_{\xi,\eta}(x, y) = \sum_{y_m<y}\sum_{x_n<x} p_{mn},$$

где суммирование распространяется на все m, для которых $y_m < y$, и все n, для которых $x_n < x$ (табл. 2 и рис. 45)

Таблица 2

Функция распределения системы случайных величин $\vec{\zeta} = (\xi, \eta)$ при проверке качества партии изделий

$y_j \backslash x_i$	0	1	2	3	4	5
0	0,02	0,10	0,20	0,26	0,30	0,32
1	0,02	0,13	0,30	0,45	0,59	0,73
2	0,02	0,13	0,32	0,51	0,71	0,93
3	0,02	0,13	0,32	0,52	0,74	1,00

Найдём одномерные законы распределения по каждой из величин ξ и η и построим графики одномерных законов распределения каждой из величин.

Для построения ряда распределения по ξ элементы матрицы распределения системы (ξ, η) (табл. 2) суммируем по столбцам (распределение по ξ см. в табл. 3 и на рис. 46).

Таблица 3

Ряд распределения случайной величины ξ при проверке качества партии изделий

x_i	0	1	2	3	4	5
p_i	0,02	0,11	0,19	0,20	0,22	0,26

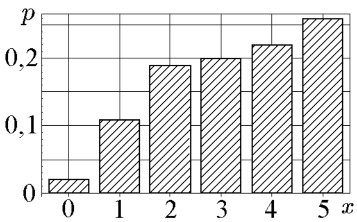


Рис. 46.

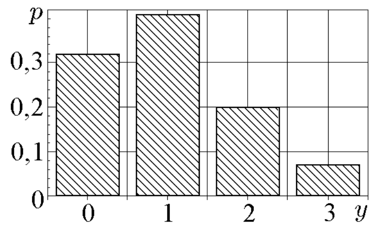


Рис. 47.

Для построения ряда распределения по η элементы матрицы распределения системы (ξ, η) (табл. 12.4.) суммируем по строкам (распределение по η см. в табл. 4 и на рис. 47).

Таблица 4

Ряд распределения случайной величины η при проверке качества партии изделий

y_i	0	1	2	3
p_i	0,32	0,41	0,20	0,07

Числовые характеристики найдём, имея данные табл. 3, 4. Математические ожидания:

$$M(\xi) = \sum_i x_i p_i = 0 \cdot 0{,}02 + 1 \cdot 0{,}11 + 2 \cdot 0{,}19 + 3 \cdot 0{,}2 + 4 \cdot 0{,}22 + 5 \cdot 0{,}26 = 3{,}27;$$

$$M(\eta) = \sum_i \sum_j y_j p_{ij} = \sum_j y_j p_j = 0 \cdot 0{,}32 + 1 \cdot 0{,}041 + 2 \cdot 0{,}2 + 3 \cdot 0{,}07 = 1{,}02.$$

Дисперсии и среднеквадратичные отклонения:

$$D(\xi) = \sum_i \sum_j [x_i - M(\xi)]^2 p_{ij} = \sum_i [x_i - M(\xi)]^2 p_i = \sum_i x_i^2 p_i - M^2(\xi) = 1{,}9971;$$

$$\sigma_\xi = \sqrt{D(\xi)} = 1{,}41319;$$

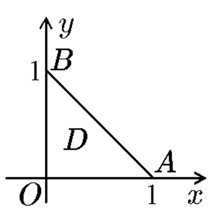
$$D(\eta) = \sum_i \sum_j [y_j - M(\eta)]^2 p_{ij} = \sum_j [y_j - M(\eta)]^2 p_j = \sum_j y_j^2 p_j - M^2(\eta) = 0{,}7996;$$

$$\sigma_\eta = \sqrt{D(\eta)} = 0{,}894204;$$

$$K_{\xi\eta} = \sum_i \sum_j [x_i - M(\xi)][y_j - M(\eta)] p_{ij} = \sum_i \sum_j x_i y_j p_{ij} - M(\xi)M(\eta) = 0{,}6346.$$

Коэффициент корреляции системы двух случайных величин:

$$\rho_{\xi,\eta} = \frac{K_{\xi,\eta}}{\sigma_\xi \sigma_\eta} = 0{,}502185.$$



В треугольнике с вершинами $O(0,0)$, $A(1,0)$, выбирают наудачу точку. Пусть ξ, η — коор-ки. Найти коэффициент корреляции между η.

Рис. 48.

Решение. Так как все точки внутри треугольника равноправны, то плотность совместного распределения $f_{\xi\eta}(x,y)$ величин ξ и η постоянна:

$$f_{\xi\eta}(x,y) = \begin{cases} 2, & \text{если } (x,y) \in D; \\ 0, & \text{если } (x,y) \notin D. \end{cases}$$

Тогда

$$M(\xi) = M(\eta) = 2\iint\limits_D x\,dx\,dy = 2\int\limits_0^1 x\,dx \int\limits_0^{1-x} dy = \frac{1}{3};$$

$$M(\xi\eta) = 2\iint\limits_D xy\,dx\,dy = 2\int\limits_0^1 x\,dx \int\limits_0^{1-x} y\,dy = \frac{1}{12};$$

$$D(\xi) = D(\eta) = 2\iint\limits_D \Big(x - \frac{1}{3}\Big)^2 dx\,dy = \frac{1}{18};$$

$$\operatorname{cov}(\xi,\eta) = \frac{1}{12} - \frac{1}{9} = -\frac{1}{36}$$

и коэффициент корреляции

$$\rho(\xi,\eta) = \frac{-1/36}{\sqrt{1/18}\sqrt{1/18}} = -\frac{1}{2}.$$

Коэффициент корреляции получился отрицательным, что можно было предсказать, так как (см. рис. 48) чем больше ξ, тем меньше в среднем η.

Пример 12.5. По плотности вероятностей

$$f(x,y) = \frac{a}{(1+x^6)(1+y^6)}$$

двумерной случайной величины $\vec{\zeta} = (\xi,\eta)$ определить коэффициент a, функцию распределения $F(x,y)$, математические ожидания и дисперсии системы случайных величин и их корреляционный момент. Найти одномерные законы распределения каждой из величин.

Решение. Коэффициент a определяется из свойства нормировки:

$$1 = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(x,y)\,dx\,dy = a\int\limits_{-\infty}^{+\infty} \frac{dx}{1+x^6} \int\limits_{-\infty}^{+\infty} \frac{dy}{1+y^6} = a\Big(\frac{2\pi}{3}\Big)^2.$$

Применим стандартную процедуру интегрирования дробно-рациональной функции:

$$\int \frac{dt}{1+t^6} = I(t) + C = \frac{1}{12}\Big[4\operatorname{arctg} t + 2\operatorname{arctg}(2t+\sqrt{3}) + 2\operatorname{arctg}(2t-\sqrt{3}) + \\ + \sqrt{3}\ln\Big|\frac{t^2+\sqrt{3}t+1}{t^2-\sqrt{3}t+1}\Big|\Big] + C.$$

Таким образом, $a = 9/4\pi^2$, а плотность вероятностей двумерной случайной величины принимает вид (см. рис. 49)

$$f(x,y) = \frac{9}{4\pi^2}\frac{1}{(1+x^6)(1+y^6)}.$$

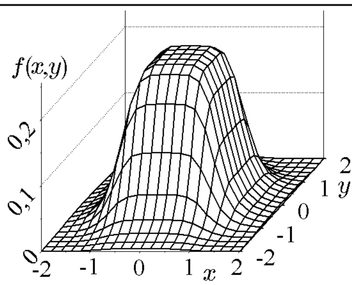


Рис. 49.

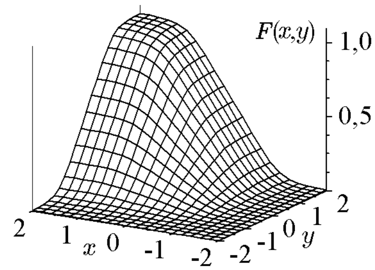


Рис. 50.

Функцию распределения $F(x,y)$ находим по формуле (11.1):

$$F(x,y) = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} f(x,y)\,dx\,dy = \frac{9}{4\pi^2}\Big(I(x)+\frac{\pi}{3}\Big)\Big(I(y)+\frac{\pi}{3}\Big) =$$

$$= \Big(\frac{3}{2\pi}I(x)+\frac{1}{2}\Big)\Big(\frac{3}{2\pi}I(y)+\frac{1}{2}\Big).$$

Здесь мы воспользовались тем, что $I(\infty)=\pi/3$. График функции распределения $F(x,y)$ изображен на рис. 50.

Плотности распределения каждой из компонент найдём по формуле (11.2):

$$f_\xi(x) = \int\limits_{-\infty}^{\infty} f(x,y)dy = \frac{9}{4\pi^2}\frac{1}{1+x^6}\int\limits_{-\infty}^{+\infty}\Big(\frac{1}{1+y^6}\Big)dy = \frac{9}{4\pi^2}\frac{1}{1+x^6}\frac{2\pi}{3} = \frac{3}{2\pi}\frac{1}{1+x^6}$$

и аналогично

$$f_\eta(y) = \frac{3}{2\pi}\frac{1}{1+y^6}.$$

Заметим, что, согласно (11.6), (11.7), компоненты ξ и η являются независимыми и их найденные плотности распределения совпадают с условными.

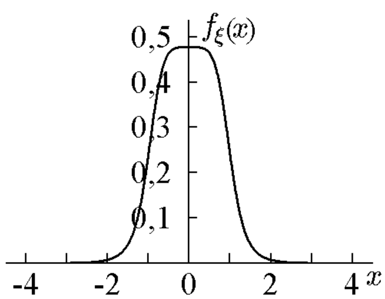


Рис. 51.

График плотности распределения $f_\xi(x)$ изображен на рис. 51.

Далее находим числовые характеристики случайной величины:

$$M(\xi) = \int\limits_{-\infty}^{+\infty} x f_\xi(x)dx = \frac{3}{2\pi}\int\limits_{-\infty}^{+\infty}\frac{x}{1+x^6}dx = 0.$$

Аналогично

$$M(\eta)=0.$$

При вычислении математического ожидания мы воспользовались тем, что определённый интеграл от нечетной функции с симметричными пределами равен нулю. Непосредственное вычисление интеграла приводит к тому же результату, поскольку

$$\int\frac{t\,dt}{1+t^6} = \frac{1}{12}\Big[2\sqrt{3}\,\mathrm{arctg}(2t-\sqrt{3}) - 2\sqrt{3}\,\mathrm{arctg}(2t+\sqrt{3}) +$$

$$+\ln\Big|\frac{(t^2+1)^2}{(t^2+\sqrt{3}t+1)(t^2-\sqrt{3}t+1)}\Big|\Big] + C.$$

Для дисперсий получим

$$D(\xi) = \int\limits_{-\infty}^{+\infty} [x - M(\xi)]^2 f_\xi(x)\, dx = \frac{3}{2\pi} \int\limits_{-\infty}^{+\infty} \frac{x^2}{1+x^6} dx = \frac{3}{2\pi} \frac{\operatorname{arctg}(x^3)}{3} \Big|_{-\infty}^{+\infty} = \frac{1}{2}$$

и аналогично

$$D(\eta) = \frac{1}{2}.$$

Тогда $\sigma = \sqrt{D(\xi)} = 1/\sqrt{2} \approx 0{,}707$ и корреляционный момент в силу независимости случайных величин ξ и η равен нулю: $K_{\xi\eta} = 0$.

Пример 12.6. Дана плотность вероятности системы случайных величин

$$f(x,y) = \begin{cases} a\sin(x+y) & \text{при } 0 \leqslant x \leqslant \pi/2;\ 0 \leqslant y \leqslant \pi/2; \\ 0 & \text{при любых других } x \text{ и } y; \end{cases}$$

Определить коэффициент a, функцию распределения системы, математические ожидания и дисперсии системы случайных величин и их корреляционный момент. Найти одномерные законы распределения каждой из величин.

Решение. На основании свойства нормировки определяем коэффициент a:

$$a = 1 \Big/ \Bigg[\int\limits_0^{\pi/2} \int\limits_0^{\pi/2} \sin(x+y) dy\, dx \Bigg],$$

где

$$\int\limits_0^{\pi/2} \int\limits_0^{\pi/2} \sin(x+y) dy\, dx = -\int\limits_0^{\pi/2} \cos(x+y)\Big|_0^{\pi/2} dx = -\int\limits_0^{\pi/2} \Big[\cos\Big(x+\frac{\pi}{2}\Big) - \cos x\Big] dx = 2.$$

Тогда $a = 1/2$.

Плотность вероятности системы случайных величин $f(x,y)$ примет вид

$$f(x,y) = \begin{cases} \dfrac{1}{2}\sin(x+y) & \text{при } 0 \leqslant x \leqslant \pi/2;\ 0 \leqslant y \leqslant \pi/2; \\ 0 & \text{при любых других } x \text{ и } y; \end{cases}$$

График функции $f(x,y)$ приведен на рис. 52.

Функция распределения вероятности системы случайных величин $F_{\xi,\eta}(x,y)$ определяется следующими условиями: для $\{x \leqslant 0, y \leqslant \pi/2\} \cap \{x \leqslant \pi/2, y \leqslant 0\}$

$$F(x,y) = 0;$$

для $\{0 < x \leqslant \pi/2, 0 < y \leqslant \pi/2\}$

$$F(x,y) = \int\limits_0^x \int\limits_0^y \sin(\xi+\eta) d\xi\, d\eta = -\int\limits_0^y [\cos(x+\eta) - \cos\eta] d\eta =$$

$$= \frac{1}{2}[\sin x + \sin y - \sin(x+y)];$$

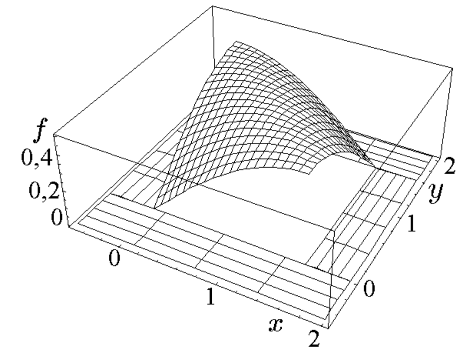


Рис. 52.

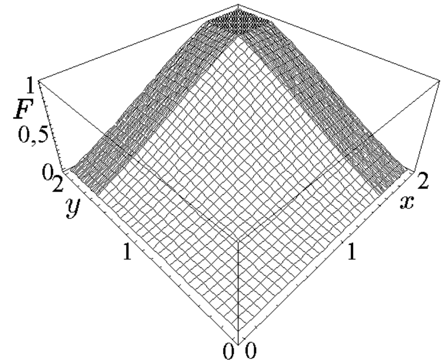


Рис. 53.

для $\{0 < x \leqslant \pi/2, \pi/2 < y < \infty\}$

$$F(x,y) = \frac{1}{2}(\sin x + 1 - \cos x);$$

для $\{\pi/2 < x < \infty, 0 < y \leqslant \pi/2\}$

$$F(x,y) = \frac{1}{2}(\sin y + 1 - \cos y);$$

для $\{x > \pi/2, y > \pi/2\}$

$$F(x,y) = 1.$$

График функции $F_{\xi,\eta}(x,y)$ приведен на рис. 53.

Определяем плотности распределения вероятностей по каждой переменной:

$$f_\xi(x) = \int\limits_0^{\pi/2} f_{\xi,\eta}(x,\eta)d\eta = \frac{1}{2}\int\limits_0^{\pi/2}\sin(x+\eta)d\eta = -\frac{1}{2}\cos(x+\eta)\Big|_0^{\pi/2} = \frac{1}{2}\Big(\cos x + \frac{1}{2}\sin x\Big).$$

Аналогично

$$f_\eta(x) = \int\limits_{-\infty}^{\infty} f_{\xi,\eta}(\xi,y)d\xi = \int\limits_0^{\pi/2} f_{\xi,\eta}(\xi,y)d\xi = \frac{1}{2}\int\limits_0^{\pi/2}\sin(\xi+y)d\xi = \frac{1}{2}(\cos y + \sin y).$$

График функции f_ξ приведен на рис. 54.

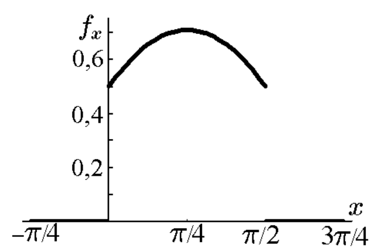


Рис. 54.

Далее найдём числовые характеристики системы случайных величин (ξ,η):

$$M(\xi) = \int\limits_0^{\pi/2}\int\limits_0^{\pi/2} x f_{\xi,\eta}(x,y)dx\,dy = 0{,}7854; \quad M(\eta) = \int\limits_0^{\pi/2}\int\limits_0^{\pi/2} y f_{\xi,\eta}(x,y)dx\,dy = 0{,}7854.$$

$$D(\xi) = \int\limits_0^{\pi/2}\int\limits_0^{\pi/2} [x - M(\xi)]^2 f_{\xi,\eta}(x,y)dx\,dy = 0{,}187647;$$

$$D(\eta) = \int\limits_0^{\pi/2}\int\limits_0^{\pi/2} [y - M(\eta)]^2 f_{\xi,\eta}(x,y)dx\,dy = 0{,}187647;$$

$$\sigma_\xi = \sqrt{D(\xi)} = 0{,}433182; \qquad \sigma_\eta = \sqrt{D(\eta)} = 0{,}433182;$$

$$K_{\xi\eta} = \int\limits_0^{\pi/2}\int\limits_0^{\pi/2} [x - M(\xi)][y - M(\eta)] f_{\xi,\eta}(x,y)dx\,dy = -0{,}0460539;$$

$$\rho_{\xi\eta} = \frac{K_{\xi\eta}}{\sigma_\xi \sigma_\eta} = -0{,}245429.$$

♦ Пусть дана система величин (ξ, η). Математическое ожидание величины η, вычисленное при условии, что величина ξ приняла определённое значение, называется *условным математическим ожиданием величины* η.

Для системы дискретных величин

$$M(\eta/\xi = x_i) = \sum_j y_j P(\eta = y_j/\xi = x_i);$$

для системы непрерывных величин

$$M(\eta/\xi = x) = \int\limits_{-\infty}^{\infty} y f_y(y/x) dy.$$

Аналогично может быть определено условное математическое ожидание величины ξ.

Условное математическое ожидание $M(\eta/\xi)$, как это следует из определения, есть некоторая функция аргумента x, которая характеризует зависимость среднего значения величины η от величины ξ. Эту функцию называют *регрессией случайной величины* η *на случайную величину* ξ, а график этой функции называют *кривой регрессии*.

Предположим, что мы хотим описать зависимость в среднем η от ξ с помощью некой функциональной зависимости $\eta = \varphi(\xi, a, b, c, \ldots)$, где a, b, c — некоторые параметры. Будем называть такую зависимость *моделью регрессии* η на ξ. Функцию $y = \varphi(x, a, b, c, \ldots)$ называют *наилучшим приближением* η в смысле метода наименьших квадратов или *среднеквадратичной регрессией* η *на* ξ, если величина $M([\eta - \varphi(\xi, a, b, c, \ldots)]^2)$ принимает наименьшее значение как функция параметров $a, b, c, \ldots$.

Рассмотрим модель линейной среднеквадратичной регрессии $\varphi(\xi, a, b) = a\xi + b$. Найдём, при каких значениях параметров a и b функция $F(a, b) = M([\eta - \varphi(\xi)]^2)$ принимает наименьшее значение. Используя свойства математического ожидания, можно записать

$$F(a,b) = M([\eta - \varphi(\xi,a,b)]^2) = M([\eta - a\xi - b]^2) =$$
$$= M(\eta^2) + a^2 M(\xi^2) + b^2 - 2aM(\xi\eta) - 2bM(\eta) + 2abM(\xi).$$

Исследуем эту функцию на экстремум. Найдём

$$\frac{\partial F(a,b)}{\partial a} = 2aM(\xi^2) - 2M(\xi\eta) + 2bM(\xi), \quad \frac{\partial F(a,b)}{\partial b} = 2b - 2M(\eta) + 2aM(\xi).$$

Приравняв $\partial F/\partial a = 0$, $\partial F/\partial b = 0$, получим систему

$$\begin{cases} aM(\xi^2) + bM(\xi) = M(\xi\eta); \\ aM(\xi) + b = M(\eta), \end{cases}$$

решив которую, найдём параметры a и b:

$$a = \frac{M(\xi\eta) - M(\xi)M(\eta)}{M(\xi^2) - M^2(\xi)} = \frac{\operatorname{cov}(\xi\eta)}{D(\xi)} = \rho_{\xi,\eta}\frac{\sigma_\eta}{\sigma_\xi},$$

$$b = \frac{M(\xi^2)M(\eta) - M(\xi)M(\xi\eta)}{M(\xi^2) - M^2(\xi)} = \frac{D(\xi)M(\eta) - M(\xi)\operatorname{cov}(\xi,\eta)}{D(\xi)} =$$

$$= M(\eta) - M(\xi)\rho_{\xi,\eta}\frac{\sigma_\eta}{\sigma_\xi}.$$

Воспользовавшись достаточным условием экстремума, легко убедиться, что функция $F(a, b)$ достигает минимума. Следовательно, уравнение линейной среднеквадратичной регрессии η на ξ имеет вид

$$y = M(\eta) + \rho_{\xi,\eta}\frac{\sigma_\eta}{\sigma_\xi}\{x - M(\xi)\}. \tag{12.13}$$

При найденных значениях a и b значение функции $F(a, b) = \sigma_\eta^2(1 - \rho_{\xi,\eta}^2)$ называется остаточной дисперсией случайной величины η относительно величины ξ. Она характеризует ошибку линейной регрессии.

12.5. Характеристическая и производящая функции

♦ *Характеристической функцией* $\gamma_\xi(t)$ случайной величины ξ называется комплекснозначная функция действительного аргумента t, определяемая как математическое ожидание случайной величины $e^{it\xi}$:

$$\gamma(t) = \gamma_\xi(t) = M(e^{it\xi}). \tag{12.14}$$

Для дискретной случайной величины:

$$\gamma(t) = \sum_{k=1}^{n} e^{itx_k} p_k, \tag{12.15}$$

для непрерывной случайной величины:

$$\gamma(t) = \int\limits_{-\infty}^{\infty} e^{itx} f(x)dx. \tag{12.16}$$

◊ Как видим, для непрерывной случайной величины характеристическая функция есть обратное преобразование Фурье $\gamma(t) = \sqrt{2\pi}F^{-1}_{x\to t}f(x)$ плотности вероятности распределения. Зная характеристическую функцию, можно с помощью преобразования Фурье найти плотность распределения:

$$f(x) = \frac{1}{\sqrt{2\pi}}F_{t\to x}\gamma(t) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} \gamma(t)e^{-itx}dt.$$

Использование характеристических функций позволяет применять для решения многих вероятностных задач теорию преобразования Фурье, хорошо разработанную в математическом анализе (см., например, [3]).

Свойства характеристической функции

Свойство 1. Характеристическая функция удовлетворяет соотношениям

$$\gamma(0) = 1, \quad |\gamma(t)| \leqslant 1. \tag{12.17}$$

Доказательство. Действительно, по определению

$$\gamma(0) = \int\limits_{-\infty}^{\infty} f(x)dx = 1.$$

С другой стороны,

$$|\gamma(t)| \leqslant \int\limits_{-\infty}^{\infty} |f(x)e^{itx}|\,dx = \int\limits_{-\infty}^{\infty} f(x)|e^{itx}|\,dx = \int\limits_{-\infty}^{\infty} f(x)dx = 1.$$

Свойство 2. Характеристическая функция случайной величины $\eta = a\xi + b$ имеет вид

$$\gamma_\eta(t) = \gamma_\xi(at)e^{ibt}. \tag{12.18}$$

Доказательство. По определению,

$$\gamma_\eta(t) = M(e^{it\eta}) = M(e^{it(a\xi+b)}) = M(e^{ita\xi}e^{itb}) = e^{itb}M(e^{ita\xi}) = \gamma_\xi(at)e^{itb}.$$

Здесь мы воспользовались тем, что постоянный множитель можно выносить за знак математического ожидания.

Свойство 3. Если существует конечное математическое ожидание случайной величины ξ^k, то

$$\gamma^{(k)}(0) = i^k M(\xi^k). \tag{12.19}$$

Доказательство. Пусть

$$\gamma(t) = \int\limits_{-\infty}^{\infty} f(x)e^{itx}dx.$$

Продифференцируем это выражение по t:

$$\gamma'(t) = i\int\limits_{-\infty}^{\infty} xf(x)e^{itx}dx,$$

следовательно,

$$\gamma'(0) = i\int\limits_{-\infty}^{\infty} xf(x)dx = iM(\xi).$$

Аналогично можно получить выражения и для производных высших порядков.

С помощью свойства 3 можно вычислять начальные моменты случайной величины по формуле

$$\alpha_k = M(\xi^k) = \frac{1}{i^k}\gamma^{(k)}(0). \tag{12.20}$$

Свойство 4. Если случайная величина ξ имеет конечные моменты любого порядка, то разложение ее характеристической функции $\gamma(t)$ в ряд Маклорена имеет вид

$$\gamma(t) = \sum_{k=0}^{\infty} \frac{i^k \alpha_k}{k!} t^k. \tag{12.21}$$

Доказательство. Справедливость утверждения непосредственно следует из предыдущего свойства.

Свойство 5. Характеристическая функция суммы независимых величин равна произведению характеристических функций этих величин:

$$\gamma_{\xi+\eta}(t) = \gamma_\xi(t)\gamma_\eta(t). \tag{12.22}$$

Доказательство. По определению,

$$\gamma_{\xi+\eta}(t) = M(e^{it(\xi+\eta)}) = M(e^{it\xi}e^{it\eta}) = M(e^{it\xi})M(e^{it\eta}) = \gamma_\xi(t)\gamma_\eta(t),$$

что и требовалось доказать. Здесь мы воспользовались соотношением (12.7).

По индукции убеждаемся, что если случайная величина ζ является суммой N независимых случайных величин: $\zeta = \sum\limits_{j=1}^{N} \xi_j$, то её характеристическая функция определяется соотношением

$$\gamma_\zeta(t) = \prod_{j=1}^{N} \gamma_{\xi_j}(t).$$

♦ *Производящей функцией начальных моментов* случайной величины ξ называется функция $\alpha(t)$, если

$$\alpha(t) = M(e^{t\xi}).$$

♦ *Производящей функцией центральных моментов* случайной величины ξ называется функция $\mu(t)$, если

$$\mu(t) = M(e^{t[\xi - M(\xi)]}).$$

Приведём основные свойства производящей функции, которые аналогичны свойствам характеристической функции.

Свойство 1. Если существуют конечные моменты $\alpha_k = M(\xi^k)$, $\mu_k = M(\xi^k)$, то

$$\alpha_k = \alpha^{(k)}(0), \qquad \mu_k = \mu^{(k)}(0). \tag{12.23}$$

Доказательство. Пусть распределение случайной величины ξ дискретно:

$$\alpha(t) = M(e^{t\xi}) = \sum_{i=1}^{n} e^{tx_i} P_i.$$

Продифференцируем это соотношение по t:

$$\alpha'(t) = \sum_{i=1}^{n} x_i e^{tx_i} P_i$$

и найдём $\alpha'(0)$:

$$\alpha'(0) = \sum_{i=1}^{n} x_i P_i = M(\xi) = \alpha_1.$$

Вычислив производные более высоких порядков, получим (12.23).

Доказательство для непрерывных распределений случайных величин аналогично.

Это свойство позволяет вычислять моменты случайной величины по формуле (12.23), аналогичной формуле (12.20).

Свойство 2. Если для случайной величины ξ существуют моменты любого порядка, то

$$\alpha(t) = \sum_{k=0}^{\infty} \frac{\alpha_k}{k!} t^k, \qquad \mu(t) = \sum_{k=0}^{\infty} \frac{\mu_k}{k!} t^k. \tag{12.24}$$

Доказательство следует непосредственно из предыдущего свойства и определения ряда Тейлора.

Свойство 3. Если $\alpha_\xi(t)$ — производящая функция случайной величины ξ, то производящая функция случайной величины $\eta = a\xi + b$ равна

$$\alpha_\eta(t) = \alpha_\xi(at) e^{bt}.$$

Доказательство аналогично доказательству свойства 2 характеристических функций.

♦ *Производящей функцией начальных моментов* $\alpha_{k,s}$ системы случайных величин (ξ, η) называется математическое ожидание

$$\alpha(t, \tau) = M(e^{t\xi + \tau\eta}).$$

♦ *Производящей функцией центральных моментов* $\mu_{k,s}$ системы случайных величин (ξ, η) называется математическое ожидание

$$\mu(t, \tau) = M(e^{t[\xi - M(\xi)] + \tau[\eta - M(\eta)]}).$$

Теорема 12.2. *Функции $\alpha(t, \tau)$ и моменты $\alpha_{k,s}$ связаны соотношением*

$$\alpha(t, \tau) = \sum_{k,s} \frac{1}{k!s!} \alpha_{k,s} t^k \tau^s.$$

Доказательство полностью соответствует доказательству аналогичной теоремы для одномерного случая.

♢ Пусть ξ и η — произвольные случайные величины. Тогда дисперсия суммы этих величин

$$\begin{aligned} D(\xi + \eta) = M(\xi^2) + 2M(\xi\eta) + M(\eta^2) - M^2(\xi) - 2M(\xi)M(\eta) - M^2(\eta) = \\ = D(\xi) + D(\eta) + 2[M(\xi\eta) - M(\xi)M(\eta)]. \end{aligned}$$

Если величина $M(\xi\eta) - M(\xi)M(\eta)$ не равна нулю, то очевидно, что величины ξ и η в этом случае зависимы (обратное утверждение не верно!). Следовательно, дисперсию можно использовать в качестве индикатора зависимости случайных величин ξ и η.

13. Биномиальное распределение и его частные случаи

13.1. Вырожденное распределение

♦ Говорят, что случайная величина ξ имеет *вырожденное распределение* с параметром a, если она может принять единственное значение $\xi = a$ с вероятностью 1. Математическое ожидание вырожденного распределения $M(\xi) = a$, дисперсия $D(\xi) = 0$.

13.2. Распределение Бернулли

♦ Говорят, что случайная величина ξ имеет *распределение Бернулли* с параметром p, и обозначают $\xi \in B_p$, если она может принимать только два значения: 1 и 0 с вероятностями p и $q = 1 - p$ соответственно, т.е. случайная величина ξ есть число появлений некоторого события в одном испытании по схеме Бернулли с вероятностью успеха p.

Характеристическая функция распределения Бернулли имеет вид

$$\gamma(t) = qe^{it\cdot 0} + pe^{it\cdot 1} = q + pe^{it}.$$

Числовые характеристики найдём из определений:

$$M(\xi) = 0 \cdot q + 1 \cdot p = p,$$
$$D(\xi) = (0 - p)^2 q + (1 - p)^2 p = p^2 q + q^2 p = pq(p + q) = pq.$$

♢ Вырожденное распределение можно рассматривать как распределение Бернулли с $q = 0$.

13.3. Биномиальное распределение

Среди законов распределения для дискретных случайных величин наиболее распространенным является *биномиальный*, частным случаем которого является распределение Бернулли. Биномиальное распределение имеет место в следующих условиях.

♦ Говорят, что случайная величина ξ *распределена по биномиальному закону* с параметрами n и p, и обозначают $\xi \in B_{n,p}$, если она может принять одно из n своих возможных значений с вероятностью

$$P(\xi = m) = P_n(m) = C_n^m p^m q^{n-m}, \quad m = \overline{0, n}, \quad q = 1 - p,$$

т.е. случайная величина ξ есть число появлений некоторого события в n испытаниях по схеме Бернулли при вероятности появления события в одном испытании, равной p.

Биномиальное распределение можно записать в виде ряда

$\xi = m$	0	1	2	...	i	...	n
$P_n(m)$	q^n	$C_n^1 pq^{n-1}$	$C_n^2 p^2 q^{n-2}$	...	$C_n^i p^i q^{n-i}$	...	p^n

♢ Многоугольники биномиального распределения для частных случаев $p = 0{,}1$; $p = 0{,}5$; $p = 0{,}9$ при $n = 20$ приведены на рис. 55.

Функция распределения случайной величины, распределённой по биномиальному закону, определяется соотношением

$$F(x) = \begin{cases} 0 & \text{при} \quad x \leqslant 0; \\ \sum\limits_{0 \leqslant m < x} P_n(m) & \text{при} \quad 0 < x \leqslant n; \\ 1 & \text{при} \quad x > n. \end{cases}$$

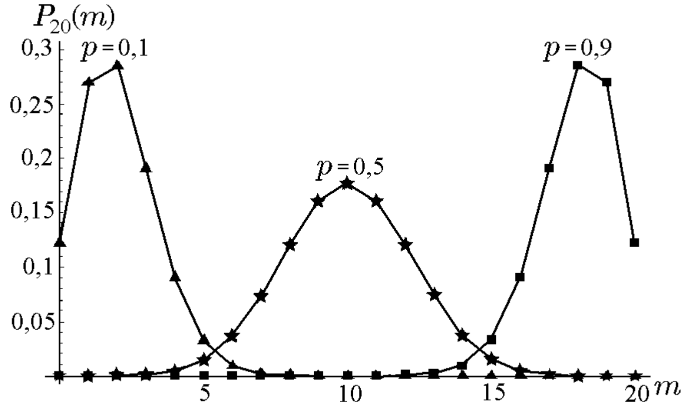


Рис. 55. Многоугольники биномиального распределения

Пример 13.1. Производятся 3 независимых выстрела по цели. Вероятность попадания при каждом выстреле равна 0,6. Найти ряд и функцию распределения числа попаданий.

Решение. Величина ξ имеет биномиальное распределение с параметрами $n = 3$, $p = 0{,}6$, $q = 0{,}4$. Случайная величина ξ (число попаданий в цель) может принять следующие значения: $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $x_3 = 3$. Эти значения случайная величина ξ принимает с вероятностями p_0, p_1, p_2, p_3, которые, в соответствии с формулой Бернулли, равны

$$\begin{aligned}
&p_0 = q^3 = 0{,}064;\\
&p_1 = C_3^1 p^1 q^2 = 3\cdot 0{,}6\cdot 0{,}4^2 = 0{,}288;\\
&p_2 = C_3^2 p^2 q = 3\cdot 0{,}6^2\cdot 0{,}4 = 0{,}432;\\
&p_3 = p^3 = 0{,}6^3 = 0{,}216.
\end{aligned}$$

Из вычисленных значений p_i, $i = \overline{0,3}$, видно, что наиболее вероятно попадание в цель двумя пулями, в то время как промах при всех выстрелах маловероятен.

Ряд распределений имеет следующий вид:

x_i	0	1	2	3
p_i	0,064	0,288	0,432	0,216

По определению, функция распределения

$$F(x) = P(\xi < x) = \sum_{x_i<x} P(\xi = x_i).$$

При $x \leqslant 0$ $F(x) = P(\xi < 0) = 0$;
при $0 < x \leqslant 1$ $F(x) = P(\xi = x_0 = 0) = 0{,}064$;
при $1 < x \leqslant 2$ $F(x) = P(\xi = 0) + P(\xi = 1) = 0{,}352$;
при $2 < x \leqslant 3$ $F(x) = \sum\limits_{i=0}^{2} P(\xi = x_i) = 0{,}784$;
при $x > 3$ $F(x) = \sum\limits_{i=0}^{3} P(\xi = x_i) = \sum\limits_{i=1}^{3} P_i = 1.$

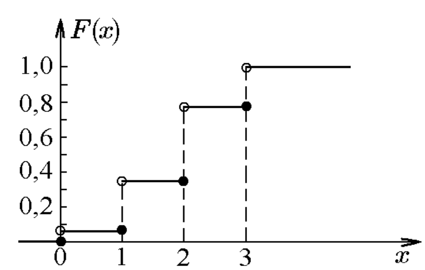


Рис. 56.

График функции распределения представлен на рис. 56.

Найдём математическое ожидание и дисперсию случайной величины, имеющей биномиальное распределение, используя производящую функцию. Для биномиального распределения по определению

$$\alpha(t) = M(e^{t\xi}) = \sum_{m=0}^{n} e^{mt} P_n(m) = \sum_{m=0}^{n} e^{mt} C_n^m p^m q^{n-m} =$$
$$= \sum_{m=0}^{n} C_n^m (pe^t)^m q^{n-m} = (q + pe^t)^n.$$

Согласно свойствам производящей функции, $\alpha_1 = \alpha'(0)$, $\alpha_2 = \alpha''(0)$, т.е.

$$\alpha'(t) = n(q + pe^t)^{n-1} pe^t, \quad \alpha'(0) = np;$$
$$\alpha''(t) = np[(n-1)(q + pe^t)^{n-2} pe^{2t} + (q + pe^t)^{n-1} e^t], \quad \alpha''(0) = np[(n-1)p + 1].$$

Тогда $M(\xi) = \alpha_1 = np$, $D(\xi) = \alpha_2 - \alpha_1^2 = npq$, $\sigma(\xi) = \sqrt{D(\xi)} = \sqrt{npq}$.

Математическое ожидание и дисперсию можно найти иначе, используя свойства математического ожидания $M(\xi)$ и дисперсии $D(\xi)$.

Искомую случайную величину рассмотрим как сумму случайных величин:

$$\xi = \xi_1 + \xi_2 + \ldots + \xi_n.$$

Ясно, что ξ_i попарно независимы. Каждая ξ_i может принимать только два значения: нуль, если событие A не наступит в i-м испытании ($i = \overline{1, n}$), и единица, если событие A наступит. Вероятности этих событий равны соответственно $q = 1 - p$ и p. Следовательно, случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ независимы и одинаково распределены по закону Бернулли с параметром p.

Поэтому математическое ожидание каждой из случайных величин ξ_i ($i = \overline{1, n}$) есть

$$M(\xi_i) = p,$$

а дисперсия

$$D(\xi_i) = pq$$

для любого i. С учётом свойств математического ожидания и дисперсии суммы одинаково распределённых независимых случайных величин находим

$$M(\xi) = M(\xi_1 + \xi_2 + \ldots + \xi_n) = M(\xi_1) + M(\xi_2) + \ldots + M(\xi_n) = np;$$
$$D(\xi) = D(\xi_1 + \xi_2 + \ldots + \xi_n) = D(\xi_1) + D(\xi_2) + \ldots + D(\xi_n) = npq.$$

Заметим также, что в данном случае коэффициент асимметрии определяется соотношением

$$A = \frac{q - p}{\sqrt{npq}} = \frac{1 - 2p}{\sqrt{np(1-p)}},$$

а эксцесс

$$E = \frac{1 - 6pq}{npq} = \frac{6p^2 - 6p + 1}{np(1-p)}.$$

Покажем, что биномиальное распределение устойчиво по суммированию.

Теорема 13.1. *Если ξ и η независимы и $\xi \in B_{n,p}$, $\eta \in B_{m,p}$, то*

$$\zeta = \xi + \eta \in B_{n+m,p}.$$

Здесь $B_{n,p}$ — биномиальное распределение с параметрами n и p.

Для доказательства в данном случае достаточно вспомнить, что случайная величина, распределённая по биномиальному закону, есть число успехов в серии испытаний Бернулли.

Пример 13.2. Случайная величина ξ представляет число бракованных деталей из выборки в 50 штук. Вероятность брака каждой детали $p = 0{,}06$. Найти $M(\xi)$, $D(\xi)$ и σ числа бракованных деталей в выборке.

Решение. Случайная величина ξ имеет биномиальное распределение. Следовательно,

$$M(\xi) = np = 50 \cdot 0{,}06 = 3 \text{ детали};$$
$$D(\xi) = npq = 50 \cdot 0{,}05 \cdot 0{,}94 = 2{,}82,$$

так как $q = 1 - p = 1 - 0{,}06 = 0{,}94$;

$$\sigma = \sqrt{npq} = \sqrt{2{,}82} = 1{,}6793.$$

Пример 13.3. Случайная величина μ — относительная частота появления события в n испытаниях по схеме Бернулли. Найти её математическое ожидание, дисперсию и среднеквадратичное отклонение.

Решение. Пусть случайная величина ξ — число появлений события в n испытаниях по схеме Бернулли. Тогда $\mu = \xi/n$. Используя свойства математического ожидания и дисперсии, находим

$$M(\mu) = M\Big(\frac{\xi}{n}\Big) = \frac{1}{n}M(\xi) = \frac{1}{n}np = p,$$
$$D(\mu) = D\Big(\frac{\xi}{n}\Big) = \frac{1}{n^2}D(\xi) = \frac{1}{n^2}npq = \frac{pq}{n}, \quad \sigma_\mu = \sqrt{\frac{pq}{n}}.$$

Полученные значения подтверждают факт устойчивости относительной частоты: ее значения группируются вокруг вероятности события ($M(\mu) = p$) и разброс их относительно этого значения тем меньше, чем больше n: ($\sigma_\mu = \sqrt{pq/n}$).

14. Геометрическое распределение

♦ Говорят, что случайная величина ξ имеет *геометрическое распределение* с параметром p, и пишут $\xi \in G_p$, если она может принять одно из своих возможных значений m с вероятностью $P(\xi = m) = P_m = pq^{m-1}$, $m = \overline{1, \infty}$, $q = 1 - p$.

♢ Таким образом, случайная величина ξ есть число попыток до первого «успеха», включая удавшуюся, или номер первого успешного испытания в схеме Бернулли с бесконечным числом испытаний и вероятностью успеха в одном испытании, равной p.

Вероятности P_m для последовательных значений m образуют геометрическую прогрессию с первым членом p и знаменателем q.

Числовые характеристики геометрического распределения вычисляются стандартным образом:

$$M(\xi) = \sum_{m=1}^{\infty} mpq^{m-1} = p\sum_{m=1}^{\infty}(q^m)' = p\Big(\sum_{m=1}^{\infty} q^m\Big)' = p\Big(\frac{q}{1-q}\Big)' = p\frac{1}{(1-q)^2} = \frac{1}{p};$$

$$M(\xi^2) = \sum_{m=1}^{\infty} m^2 p q^{m-1} = p\Big(\sum_{m=1}^{\infty} m q^m\Big)' = p\Big(q \sum_{m=1}^{\infty} m q^{m-1}\Big)' = p\Big(q \sum_{m=1}^{\infty} (q^m)'\Big)' =$$
$$= p\bigg(q\Big(\sum_{m=1}^{\infty} q^m\Big)'\bigg)' = p\Big(q\Big(\frac{q}{1-q}\Big)'\Big)' = p\Big(\frac{q}{(1-q)^2}\Big)' = p\frac{1+q}{(1-q)^3} = \frac{1+q}{p^2};$$
$$D(\xi) = M(\xi^2) - M^2(\xi) = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}.$$

Таким образом, для геометрического распределения ξ:

$$M(\xi) = \frac{1}{p}, \quad D(\xi) = \frac{q}{p^2}.$$

Аналогично можно найти коэффициент асимметрии

$$A(\xi) = \frac{2-p}{\sqrt{1-p}}$$

и эксцесс

$$E(\xi) = \frac{6-6p+p^2}{1-p}.$$

Пример 14.1. Монету подбрасывают до первого появления орла. Найти среднее число подбрасываний и дисперсию числа подбрасываний.

Решение. Число подбрасываний до первого появления орла — случайная величина ξ — распределено, очевидно, по геометрическому закону с параметром $p = 1/2$. Следовательно, $m = 1/p = 2$, $D = q/p^2 = 2$.

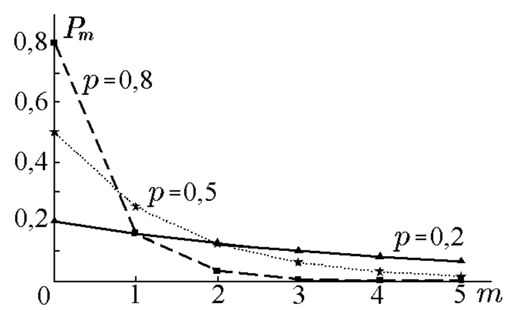


Рис. 57.

$\lozenge$ На практике рассматривают также случайную величину $\eta = \xi - 1$ — число «безуспешных» попыток, — распределённую по геометрическому закону:

$$P(\eta = m) = P_m = pq^m; \quad m = \overline{0, \infty}; \quad 0 \leqslant p \leqslant 1; \quad q = 1-p.$$

Многоугольники распределения для случайной величины η, распределённой по геометрическому закону с параметрами $p = 0{,}2$; $p = 0{,}5$; $p = 0{,}8$ приведены на рис. 57.

Основные числовые характеристики распределения случайной величины η выражаются через параметр распределения p следующим образом:

$$M(\eta) = \frac{1-p}{p}, \quad D(\eta) = \frac{1-p}{p^2},$$
$$A(\eta) = \frac{2-p}{\sqrt{1-p}}, \quad E(\eta) = \frac{6-6p+p^2}{1-p}.$$

15. Распределение Пуассона

15.1. Числовые характеристики распределения Пуассона

При решении многих практических задач приходится иметь дело с дискретными случайными величинами, распределёнными по закону Пуассона.

♦ Дискретная случайная величина ξ называется *распределённой по закону Пуассона* с параметром $\lambda > 0$ и обозначается $\xi \in \Pi_\lambda$, если она может принимать одно из своих возможных значений k ($\{\xi = k\}$) с вероятностью

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

Многоугольники распределения для случайной величины, распределённой по закону Пуассона с различными значениями параметра λ, приведены на рис. 58.

Заметим, что

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

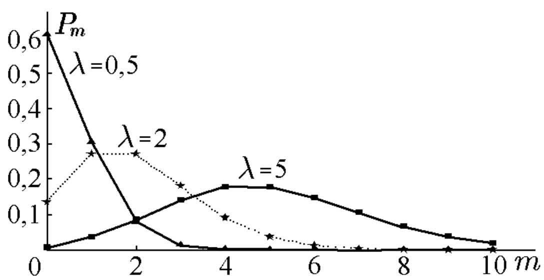


Рис. 58.

Множество значений случайной величины ξ счётно. Пуассоновское распределение является предельным для биномиального при $p \to 0$, $n \to \infty$, если $np = \lambda = \text{const}$. Этим распределением можно пользоваться, если производится большое число независимых опытов, в каждом из которых событие A происходит с малой вероятностью.

Примеры случайных величин, распределённых по закону Пуассона:

1) число отказов сложной аппаратуры за время t;
2) число вызовов, поступивших оператору сотовой связи за определённый промежуток времени;
3) число электронов, вылетающих с катода за время t;
4) число опечаток в корректуре и т.д.

Найдём математическое ожидание $M(\xi)$, дисперсию $D(\xi)$ и среднеквадратичное отклонение σ:

$$M(\xi) = \sum_{k=0}^{\infty} x_k p_k = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}e^{\lambda} = \lambda,$$

так как

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

Таким образом, $M(\xi) = \lambda$, т.е. параметр λ в распределении Пуассона есть математическое ожидание случайной величины. Воспользовавшись свойством дисперсии

$$D(\xi) = M(\xi^2) - (M(\xi))^2,$$

определим теперь $D(\xi)$:

$$M(\xi^2) = \sum_{k=0}^{\infty} k^2\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k=1}^{\infty} k\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}[(k-1)+1]\frac{\lambda^{k-1}}{(k-1)!} =$$

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

Так как каждая из сумм равна e^{λ}, то

$$M(\xi^2) = \lambda^2 e^{-\lambda} e^{\lambda} + \lambda e^{-\lambda} e^{\lambda} = \lambda^2 + \lambda.$$

Следовательно,

$$D(\xi) = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Таким образом, дисперсия случайной величины, имеющей распределение Пуассона, численно равна её математическому ожиданию.

Тогда

$$\sigma = \sqrt{\lambda}.$$

Итак, $M(\xi) = \lambda$, $D(\xi) = \lambda$, $\sigma = \sqrt{\lambda}$.

$\lozenge$ Заметим, что аналогично можно найти $A(\xi) = 1/\sqrt{\lambda}$, $E(\xi) = 1/\sqrt{\lambda}$.

Покажем, что распределение Пуассона устойчиво по суммированию.

Теорема 15.1. *Если ξ и η независимы и $\xi \in \Pi_{\lambda}$, $\eta \in \Pi_{\mu}$, то*

$$\zeta = \xi + \eta \in \Pi_{\lambda+\mu}.$$

Здесь Π_{λ}, Π_{μ}, $\Pi_{\lambda+\mu}$ — распределения Пуассона с параметрами $\lambda, \mu, \lambda+\mu$ соответственно.

Доказательство. Воспользуемся формулой (11.16):

$$P(\xi + \eta = k) = \sum_{m=0}^{k} P(\xi = k-m) = \sum_{m=0}^{k} \frac{\lambda^m e^{-\lambda}}{m!} \frac{\mu^{k-m} e^{-\mu}}{(k-m)!} =$$
$$= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{m=0}^{k} \frac{k!}{m!(k-m)!} \lambda^m \mu^{k-m} = \frac{e^{-(\lambda+\mu)}}{k!} \sum_{m=0}^{k} C_k^m \lambda^m \mu^{k-m} = \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^k}{k!},$$

что и требовалось доказать.

Пример 15.1. Найти характеристическую функцию $\gamma(t)$ дискретной случайной величины ξ, распределённой по закону Пуассона: $P(\xi = m) = \lambda^m e^{-\lambda}/m!$, $m = \overline{0, \infty}$.

Решение. По определению,

$$\gamma(t) = \sum_{m=0}^{\infty} e^{itm} P(\xi = m) = \sum_{m=0}^{\infty} \frac{e^{itm} \lambda^m e^{-\lambda}}{m!} = e^{-\lambda} \sum_{m=0}^{\infty} \frac{(\lambda e^{it})^m}{m!}.$$

Сделав замену переменной $\lambda e^{it} = z$, получим

$$\gamma(t) = e^{-\lambda} \sum_{m=0}^{\infty} \frac{z^m}{m!} = e^{-\lambda} e^{z} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}.$$

Следовательно,

$$\gamma(t) = e^{\lambda(e^{it}-1)}. \tag{15.1}$$

15.2. Распределение Пуассона как предельный случай биномиального закона распределения

Распределение Пуассона можно получить как предельный случай биномиального, когда число испытаний велико, а вероятность появления события в одном испытании очень мала, причём среднее число «успехов» в бесконечной серии испытаний остается постоянным. В этом смысле закон Пуассона можно интерпретировать как закон редких событий.

Пусть случайная величина ξ распределяется по биномиальному закону $B_{n,p}$. Рассмотрим предел биномиальной вероятности при $n \to \infty$, $p \to 0$, причём так, что $np = \lambda = \mathrm{const}$ (подобная ситуация, как мы убедимся в дальнейшем, реализуется в простейшем потоке событий):

$$\begin{gathered}\lim_{n\to\infty} P_n(m) = \lim_{n\to\infty} C_n^m p^m q^{n-m} = \\ = \lim_{n\to\infty} \frac{n(n-1)(n-2)\cdots(m-n+1)}{m!}\Big(\frac{\lambda}{n}\Big)^m \Big(1-\frac{\lambda}{n}\Big)^{n-m} = \\ = \frac{\lambda^m}{m!}\lim_{n\to\infty}\frac{n(n-1)(n-2)\cdots(m-n+1)}{n^m}\lim_{n\to\infty}\Big(1-\frac{\lambda}{n}\Big)^{n-m} = \\ = \frac{\lambda^m}{m!}\lim_{n\to\infty} e^{-\lambda(n-m)/n} = \frac{\lambda^m e^{-\lambda}}{m!} = P_m,\end{gathered}$$

т.е. вероятность в таком пределе является пуассоновской.

Следовательно, когда n велико, а p мало, можно использовать для приближённого вычисления биномиальных вероятностей формулу Пуассона.

Пример 15.2. Пекарь выпекает 160 кексов, кладя в тесто 300 изюминок. Какова вероятность, что случайно выбранный кекс не будет содержать изюминок?

Решение. Пусть ξ — число попавших в кекс изюминок. Очевидно, что ξ имеет биномиальное распределение с параметрами $p = 1/160$ и $n = 300$. Следовательно,

$$P(\xi = 0) = C_{300}^0\Big(\frac{1}{160}\Big)^0\Big(1-\frac{1}{160}\Big)^{300} = \Big(\frac{159}{160}\Big)^{300} \approx 0{,}152455.$$

Поскольку p мало, а n достаточно велико, эту задачу можно было бы решить и с помощью формулы Пуассона. Среднее число изюминок, приходящихся на один кекс, $\lambda = 300/160 = 1{,}875$, следовательно,

$$P(\xi = 0) \approx P_0 = \frac{e^{-1{,}875}(1{,}875)^0}{0!} =\approx 0{,}153355.$$

15.3. Простейший (пуассоновский) поток событий

Пуассоновскому распределению подчиняется так называемый «простейший поток событий», то есть последовательность однородных событий, наступающих одно за другим в случайные моменты времени. Примеры: поток вызовов, поступающих операторам сотовой связи, на пункт неотложной медицинской помощи, поток сбоев на ЭВМ, поток α-частиц, испускаемых радиоактивным источником и другие.

Поток называется *простейшим* (пуассоновским), если он обладает свойствами

1) стационарности, то есть вероятность попадания того или иного числа событий на любой участок времени длиной t не зависит от положения этого участка на оси времени, а зависит только от его длины t;

2) ординарности, то есть события возникают поодиночке, а не группами по два, по три и т.д., или вероятность появления двух и более событий в малый промежуток времени длиной τ пренебрежимо мала по сравнению с вероятностью появления в этот промежуток одного события;

3) отсутствия последствий, то есть «будущее» потока не зависит от его прошлого, или вероятность попадания того или иного числа событий на заданный участок оси времени не зависит от того, сколько событий попало на любой другой, не пересекающийся с ним участок.

♦ Среднее число событий потока α, появляющихся в единицу времени, называется *интенсивностью потока.*

Пусть в течение времени t действует простейший поток событий интенсивностью α. Требуется вычислить вероятность того, что за время t наступит ровно m событий (случайная величина ξ — число событий за время t, примет значение, равное m).

Разобьем отрезок времени t на n частей (n достаточно велико). Обозначим $\tau = t/n$. По свойству стационарности среднее число событий на отрезке τ есть $\alpha\tau = \alpha t/n$. Так как τ мало, то в соответствии со свойством ординарности будем считать, что на отрезке τ событие не может произойти более одного раза. Следовательно, число событий на отрезке τ подчинено распределению Бернулли с вероятностью успеха $p = \alpha t/n$.

Все n промежутков можно рассматривать как серию испытаний, результатом каждого из которых является наступление события с вероятностью $p = \alpha t/n$ или ненаступление с вероятностью $q = 1 - p = 1 - \alpha t/n$. Так как все n опытов независимы (по свойству 3), то вероятность того, что за время t наступит ровно m событий, определяется формулой Бернулли, причём тем точнее, чем больше n, т.е

$$P(\xi = m) = \lim_{n\to\infty} P_n(m) = \lim_{n\to\infty} C_n^m \Big(\frac{\alpha t}{n}\Big)^m \Big(1 - \frac{\alpha t}{n}\Big)^{n-m}.$$

Так как $p = \alpha t/n \to 0$ при $n \to \infty$, причём $np = n\alpha t/n = \alpha t = \lambda$, то данная вероятность при $n \to \infty$ стремится к вероятности, определяемой формулой Пуассона

$$P(\xi = m) = \lim_{n\to\infty} C_n^m \Big(\frac{\alpha t}{n}\Big)^m \Big(1 - \frac{\alpha t}{n}\Big)^{n-m} = P_m = \frac{e^{-\alpha t}(\alpha t)^m}{m!}.$$

Следовательно, поток событий на интервале t подчинен закону Пуассона с параметром $\lambda = \alpha t$ (α — интенсивность потока).

Пример 15.3. Среднее число вызовов, поступающих оператору сотовой связи в одну минуту, равно 2. Найти вероятность того, что за 5 минут поступит 1) 2 вызова, 2) менее двух вызовов, 3) не менее двух вызовов. Поток вызовов предполагается простейшим.

Решение. По условию $\alpha = 2$, $t = 5$, $m = 2$. Таким образом, $\lambda = 10$, и по формуле Пуассона находим:

1) $P_2 = \dfrac{10^2 e^{-10}}{2!} \approx 0{,}00225$;

2) $P(m < 2) = P_0 + P_1 = \dfrac{10^0 e^{-10}}{0!} + \dfrac{10^1 e^{-10}}{1!} \approx 0{,}000495$;

3) $P(m \geqslant 2) = 1 - P(m < 2) = 1 - 0{,}000495 \approx 0{,}999505$.

Заметим, что параметр t можно рассматривать не только как длину временного интервала, но и как длину интервала на произвольной координатной оси или как площадь некоторой области в пространстве $\mathbb{R}^2$, или как объём произвольной области в пространстве $\mathbb{R}^n$.

Пример 15.4. Средняя плотность болезнетворных микробов в одном кубическом метре воздуха равна 150. Берется на пробу 2 кубических дециметра воздуха. Найти вероятность того, что в них будет обнаружен хотя бы один микроб.

Решение. Предположим, что: 1) вероятность присутствия того или иного числа микробов в выделенном объёме зависит только от величины объёма; 2) в элементарном объёме dv не может содержаться более одного микроба. Но тогда случайная величина ξ — число микробов в выделенном объёме — будет иметь распределение Пуассона. Для 2 кубических дециметров параметр λ этого распределения — среднее число микробов — будет равен $\lambda = 2 \cdot 150 \cdot 10^{-3} = 0{,}3$. Тогда вероятность обнаружить хотя бы один микроб: $P = 1 - P(\xi = 0) = 1 - e^{-0{,}3} \approx 0{,}259$.

Пример 15.5. Деревья в лесу растут в случайных точках, которые образуют пуассоновское поле с плотностью 0,04 деревьев/м^2. Случайная величина ξ — расстояние от произвольной точки в лесу до ближайшего дерева. Найти закон распределения случайной величины ξ и среднее расстояние до ближайшего дерева.

Решение. Найдём функцию распределения ξ. По определению, $F_\xi(x) = P(\xi < x)$. Но событие $\{\xi < x\} = \{$расстояние до ближайшего дерева меньше $x\}$ равносильно событию «на расстоянии x от выбранной точки есть хотя бы одно дерево» или событию «в круге с площадью πx^2 есть хотя бы одно дерево». Поскольку число деревьев на участке леса площадью S, по условию, имеет распределение Пуассона с параметром $\lambda = 0{,}04S$, то получим

$$F_\xi(x) = 1 - e^{-0{,}04\pi x^2}, \quad x > 0.$$

Тогда плотность распределения случайной величины ξ

$$f_\xi(x) = F'_\xi(x) = 0{,}08\pi x e^{-0{,}04\pi x^2}, \quad x > 0.$$

Соответственно, среднее расстояние до ближайшего дерева равно

$$M(\xi) = \int\limits_0^\infty x f_\xi(x) dx = 2{,}5 \text{ м}.$$

16. Гипергеометрическое распределение

Случайная величина ξ имеет *гипергеометрическое распределение* с $N = 2, 3, \ldots$; $M = 1, 2, \ldots < N$; $n = 1, 2, \ldots < N$, если она принимает конечное множество натуральных значений $\xi = m$, соответственно, с вероятностями

$$P(\xi = m) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n},$$

где $m = \overline{0, \min(M, n)}$.

Пользуясь языком «схемы урн», можно сказать, что гипергеометрическое распределение возникает при следующих условиях. В урне находится N шаров, из которых ровно M белых. Пусть одновременно (или один за другим без возврата) извлечены n шаров. Найти вероятность того, что среди этих извлечённых шаров будет m белых шаров.

Таким образом, случайную величину ξ можно трактовать как число успехов в n последовательных испытаниях при условии, что известно общее число M возможных успехов в N возможных испытаниях.

На практике гипергеометрическое распределение применяется при решении задач, связанных с контролем продукции.

Для того чтобы найти числовые характеристики случайной величины, распределённой по гипергеометрическому закону, представим ξ в виде

$$\xi = \xi_1 + \xi_2 + \ldots + \xi_n, \tag{16.1}$$

где ξ_j — число успехов в j-м испытании. Очевидно, что каждая из величин ξ_j может принять одно из двух значений: 0 или 1, поэтому

$$M(\xi_j) = 0 \cdot P(\xi_j = 0) + 1 \cdot P(\xi_j = 1) = P(\xi_j = 1).$$

Случайные величины ξ_j зависимы между собой, но для каждой из них вероятность $P(\xi_i = 1)$ одна и та же и равна

$$P(\xi_j = 1) = P(\xi_1 = 1) = \frac{M}{N}.$$

Действительно, для k испытаний общее число исходов равно $N(N-1)\cdots(N-k+1)$. Благоприятными для события $\{\xi_k = 1\}$ являются исходы, дающие в k-м испытании успех. Общее число таких исходов, согласно теореме умножения, равно $M(N-1)\cdots(N-k+1)$. Тогда

$$P(\xi_k = 1) = \frac{M(N-1)\cdots(N-k+1)}{N(N-1)\cdots(N-k+1)} = \frac{M}{N}.$$

Так как математическое ожидание суммы равно сумме математических ожиданий, то

$$M(\xi) = M(\xi_1 + \xi_2 + \ldots + \xi_n) = \sum_{i=1}^{n} M(\xi_i) = n\frac{M}{N}.$$

Для вычисления дисперсии воспользуемся формулой $D(\xi) = M(\xi^2) - M^2(\xi)$. $M(\xi^2)$ найдём также через математическое ожидание как сумму

$$M(\xi^2) = M(\xi_1 + \xi_2 + \ldots + \xi_n)^2 = \sum_{i=1}^{n} M(\xi_i^2) + 2\sum_{1 \leqslant i < j}^{n} M(\xi_i\xi_j).$$

Так как ξ_j — число успехов в j-м испытании, то

$$\begin{aligned} &M(\xi_i^2) = 0^2 \cdot P(\xi_i = 0) + 1^2 \cdot P(\xi_i = 1) = P(\xi_i = 1) = \frac{M}{N};\\ &M(\xi_i, \xi_j) = 0 \cdot 0 \cdot P(\xi_i = 0, \xi_j = 0) + 0 \cdot 1 \cdot P(\xi_i = 0, \xi_j = 1) +\\ &\qquad +1 \cdot 0 \cdot P(\xi_i = 1, \xi_j = 0) + 1 \cdot 1 \cdot P(\xi_i = 1, \xi_j = 1) = P(\xi_i = 1, \xi_j = 1). \end{aligned}$$

Можно показать, что для любых пар (ξ_i, ξ_j), $i \neq j$, вероятность $P(\xi_i = 1, \xi_j = 1)$ одна и та же. Найдём эту вероятность для (ξ_1, ξ_2):

$$P(\xi_1 = 1, \xi_2 = 1) = \frac{M}{N}\frac{M-1}{N-1}.$$

Тогда

$$M(\xi^2) = n\frac{M}{N} + 2C_n^2\frac{M}{N}\frac{M-1}{N-1} = n\frac{M}{N} + n(n-1)\frac{M}{N}\frac{M-1}{N-1}$$

и

$$D(\xi) = M(\xi^2) - M^2(\xi) = \frac{nM(N-M)(N-n)}{(N-1)N^2}.$$

В данном случае интерес представляют также числовые характеристики: коэффициент асимметрии

$$A = \frac{\sqrt{N-1}(N-2M)(N-2n)}{(N-2)\sqrt{Mn(N-M)(N-n)}},$$

определяющий степень асимметричности многоугольника распределения, и эксцесс

$$E = -3 + \frac{(N-1)N^2}{Mn(N-3)(N-2)(N-M)(N-n)} \times$$
$$\times\Big\{N(N+1) - 6n(N-n) + \frac{1}{N^2}[3M(N-M)(6n(N-n) + N^2(n-2) - Nn^2)]\Big\},$$

определяющий степень островершинности многоугольника распределения.

Пример 16.1. Из партии, содержащей N изделий, среди которых имеется M стандартных, выбраны случайным образом n изделий для проверки их качества. Построить ряд распределений случайного числа $\xi = m$ стандартных изделий, содержащихся в выборке.

Решение. Искомая вероятность вычисляется по формуле

$$P(\xi = m) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}.$$

Пример 16.2. Имеются 7 электрических ламп, среди которых 3 неисправные, на вид неотличимые от новых. Наугад берутся 4 лампы и вставляются в 4 патрона. Найти ряд распределения и построить многоугольник распределения числа радиоламп ξ, которые будут работать. Найти ее числовые характеристики.

Решение. Величина ξ имеет гипергеометрическое распределение с параметрами $N = 7$; $n = 4$; $M = 4$. Строим ряд распределения величины ξ:

$$P(\xi = 1) = \frac{C_4^1 C_3^3}{C_7^4} = \frac{4}{35}; \qquad P(\xi = 2) = \frac{C_4^2 C_3^2}{C_7^4} = \frac{18}{35};$$
$$P(\xi = 3) = \frac{C_4^3 C_3^1}{C_7^4} = \frac{12}{35}; \qquad P(\xi = 4) = \frac{C_4^4 C_3^0}{C_7^4} = \frac{1}{35}.$$

Ряд распределения случайной величины ξ (значения вероятностей P_m округлены до трех знаков после запятой) будет иметь вид

$\xi = m$	1	2	3	4
P_m	0,114	0,514	0,343	0,029

Найдём числовые характеристики случайной величины ξ:

$$M(\xi) \approx 2{,}286; \quad D(\xi) \approx 0{,}490; \quad A \approx 0{,}041; \quad E \approx -0{,}283.$$

В частности, коэффициент асимметрии $A = 0{,}041 > 0$ указывает на скошенность многоугольника распределения влево, а коэффициент эксцесса $E = -0{,}283 < 0$ подчёркивает туповершинность многоугольника распределения, который изображен на рис. 59.

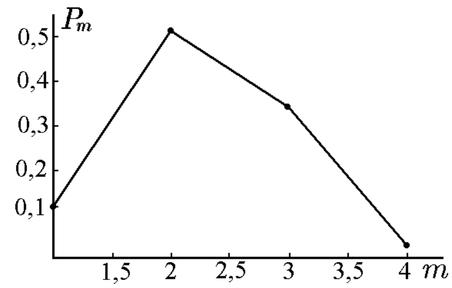


Рис. 59.

Если M и N велики, а число испытаний n много меньше M и N, то каждое из этих испытаний почти не влияет на вероятность успеха в последующих испытаниях.

Теорема 16.1. *Если $N \to \infty$, $M \to \infty$ так, что $M/N = p = \text{const}$, то для любых n, $0 \leqslant m \leqslant n$,*

$$P(\xi = m) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n} \to C_n^m p^m q^{n-m},$$

где $p = M/N$, $q = 1 - p$.

Доказательство. Рассмотрим предел вероятности случайного события $\{\xi = m\}$ при $N \to \infty$, $M \to \infty$ и учтем, что $M/N = p = \text{const}$. Заметим предварительно, что

при $M \to \infty$ $\qquad C_M^m = \dfrac{M(M-1)\cdots(M-m+1)}{m!} \sim \dfrac{M^m}{m!};$

при $N \to \infty$ $\qquad C_N^n \sim \dfrac{N^n}{n!};$

при $M \to \infty$, $N \to \infty$ $\qquad C_{N-M}^{n-m} \sim \dfrac{(N-M)^{n-m}}{(n-m)!}.$

Тогда

$$\lim_{\substack{N\to\infty\\ M\to\infty}} \frac{C_M^m C_{N-M}^{n-m}}{C_N^n} = \frac{n!}{m!(n-m)!} \lim_{\substack{N\to\infty\\ M\to\infty}} \frac{M^m (N-M)^{n-m}}{N^n} =$$
$$= C_n^m \lim_{\substack{N\to\infty\\ M\to\infty}} \left(\frac{M}{N}\right)^m \left(\frac{N-M}{N}\right)^{n-m} = C_n^m p^m (1-p)^{n-m} = P_n(m).$$

Таким образом, гипергеометрическое распределение в пределе $N \to \infty$, $M \to \infty$ переходит в биномиальное с параметрами n и $p = M/N$. Следовательно, при больших N, M для приближенного вычисления вероятностей случайных величин, распределённых по гипергеометрическому закону, можно использовать формулу Бернулли.

17. Равномерное распределение

♦ Случайная величина ξ *равномерно распределена* на отрезке $[a, b]$ и обозначена $\xi \in U_{a,b}$, если плотность её распределения имеет вид

$$f(x) = \begin{cases} C & \text{при} \quad x \in [a, b], \\ 0 & \text{при} \quad x \notin [a, b], \end{cases}$$

где $C = \text{const}$ и $a < b$.

Из условия, что

$$\int\limits_{-\infty}^{\infty} f(x)dx = 1,$$

находим C:

$$\int\limits_{-\infty}^{\infty} f(x)dx = \int\limits_{-\infty}^{a} 0dx + \int\limits_{a}^{b} Cdx + \int\limits_{b}^{\infty} 0dx = C(b-a) = 1.$$

Отсюда

$$C = \frac{1}{b-a}.$$

Следовательно, плотность вероятности имеет вид (см. рис. 60)

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{при} \quad x \in [a,b], \\ 0 & \text{при} \quad x \notin [a,b]. \end{cases}$$

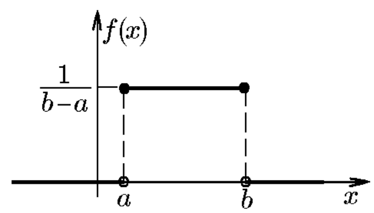


Рис. 60.

Найдём функцию распределения $F(x)$ для равномерного распределения на интервале $[a,b]$. Согласно определению,

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

1. Если $x < a$, то

$$F(x) = \int\limits_{-\infty}^{a} 0dt = 0.$$

2. Если $a \leqslant x \leqslant b$, то

$$F(x) = \int\limits_{-\infty}^{a} 0dt + \int\limits_{a}^{x} \frac{dt}{b-a} = \frac{x-a}{b-a}.$$

3. Если $x > b$, то

$$F(x) = \int\limits_{-\infty}^{a} 0dt + \int\limits_{a}^{b} \frac{dt}{b-a} + \int\limits_{b}^{x} 0dt = 1.$$

Итак (см. рис. 61),

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

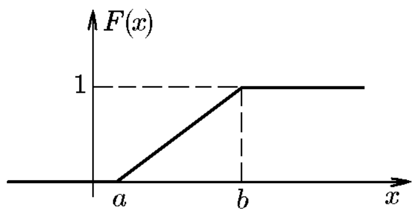


Рис. 61.

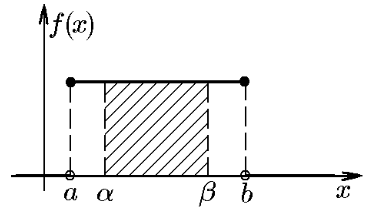


Рис. 62.

Вероятность попадания случайной величины ξ, имеющей равномерное распределение, на участок $[\alpha, \beta]$, который является частью участка $[a, b]$ (рис. 62), определяется по формуле

$$P\Big(\alpha \leqslant \xi \leqslant \beta\Big) = \frac{\beta - \alpha}{b - a}.$$

Найдём основные числовые характеристики случайной величины, имеющей равномерное распределение. Для математического ожидания получим

$$M(\xi) = \int\limits_{-\infty}^{\infty} xf(x)dx = \int\limits_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\frac{x^2}{2}\Big|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}.$$

Итак, математическое ожидание равномерного распределения равно координате середины интервала $[a, b]$. Дисперсию случайной величины ξ находим по формуле

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

но

$$M(\xi^2) = \frac{1}{b-a}\int\limits_a^b x^2dx = \frac{1}{b-a}\frac{x^3}{3}\Big|_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{b^2+ab+a^2}{3}.$$

Тогда

$$D(\xi) = \frac{1}{3}(b^2+ab+a^2) - \frac{1}{4}(a+b)^2 = \frac{(b-a)^2}{12},$$

откуда

$$\sigma = \sqrt{D(\xi)} = \frac{b-a}{2\sqrt{3}}.$$

Аналогичными вычислениями можно получить $A(\xi) = 0$ и $E(\xi) = -1{,}2$, что характеризует равномерное распределение (см. рис. 62) как симметричное и туповершинное.

Равномерное распределение характерно для ошибок грубых измерений. Если измерение какой-либо величины производится с точностью до целых делений шкалы, без определения на глаз доли деления, то ошибка измерения может иметь любое значение, не превосходящее по абсолютной величине половину деления шкалы, причём нет никаких оснований считать вероятности разных значений различными. Более того, можно с уверенностью сказать, что при большом числе таких измерений все значения ошибки в пределах от минус половины деления до плюс половины деления будут встречаться одинаково часто. Поэтому ошибка грубых измерений, производимых с точностью до целых делений шкалы, представляет собой случайную величину, равномерно распределённую в пределах от $-\delta/2$ до $\delta/2$, где δ — цена деления шкалы.

Пример 17.1. Случайная величина ξ равномерно распределена на отрезке $[-1, 3]$. Найти плотность вероятности случайной величины $\eta = \xi^2$.

Решение. Так как ξ равномерно распределена на отрезке $[-1, 3]$, то её функция распределения имеет вид

$$F_\xi(x) = \begin{cases} 0, & x \leqslant -1; \\ \dfrac{x+1}{4}, & -1 < x \leqslant 3; \\ 1, & x > 3. \end{cases}$$

Найдём функцию распределения величины η. По определению, $F_\eta(x) = P(\eta < x) = P(\xi^2 < x)$.

1) Пусть $x \leqslant 0$. Очевидно, что $F_\eta(x) = P(\xi^2 < x) = 0$.

2) Пусть $0 < x \leqslant 1$. Тогда

$$F_\eta(x) = P(\eta < x) = P(\xi^2 < x) = P(-\sqrt{x} < \xi < \sqrt{x}) = F_\xi(\sqrt{x}) - F_\xi(-\sqrt{x}) = \\ = \frac{\sqrt{x}+1}{4} - \frac{-\sqrt{x}+1}{4} = \frac{1}{2}\sqrt{x}.$$

3) Пусть $1 < x \leqslant 9$. Тогда

$$F_\eta(x) = P(\xi^2 < x) = P(\xi < \sqrt{x}) = F_\xi(\sqrt{x}) = \frac{\sqrt{x}+1}{4}.$$

4) Пусть $x > 9$. Очевидно, что в этом случае $F_\eta(x) = 1$. Таким образом,

$$F_\eta(x) = \begin{cases} 0, & x \leqslant 0; \\ \dfrac{1}{2}\sqrt{x}, & 0 < x \leqslant 1; \\ \dfrac{\sqrt{x}+1}{4}, & 1 < x \leqslant 9; \\ 1, & x > 9; \end{cases} \qquad f_\eta(x) = F'_\eta(x) = \begin{cases} 0, & x \leqslant 0; \\ \dfrac{1}{4\sqrt{x}}, & 0 < x \leqslant 1; \\ \dfrac{1}{8\sqrt{x}}, & 1 < x \leqslant 9; \\ 0, & x > 9. \end{cases}$$

Можно было решить данную задачу и иначе, воспользовавшись формулой для преобразования плотности распределения: если ξ — непрерывная случайная величина, имеющая плотность распределения $f_\xi(x)$, а функция $\eta = \varphi(\xi)$ дифференцируема и монотонна на множестве, где $f_\xi(x) > 0$, то случайная величина $\eta = \varphi(\xi)$ имеет плотность распределения $f_\eta(x) = |[\varphi^{-1}(x)]'| f_\xi(\varphi^{-1}(x))$.

В нашем случае функция $\eta = \xi^2$ немонотонна на множестве $]-1, 3[$. Однако, учитывая свойства плотности равномерно распределённой случайной величины ξ и свойства отображения $\eta = \xi^2$, можно получить, что

$f_\eta(x) = 2(\sqrt{x})' f_\xi(\sqrt{x})$, если $x \in]0, 1[$ (в интервал $]0, 1[$ отображаются два интервала $]-1, 0[$ и $]0, 1[$);

$f_\eta(x) = (\sqrt{x})' f_\xi(\sqrt{x})$, если $x \in]1, 9[$ (в интервал $]1, 9[$ отображается один интервал $]1, 3[$; $f_\eta(x) = 0$ для всех остальных значений x.

Учитывая, что $(\sqrt{x})' f_\xi(\sqrt{x}) = 1/8\sqrt{x}$, для указанных интервалов получим ту же формулу, что и выше.

Пример 17.2. Случайные величины ξ и η независимы; ξ равномерно распределена на $]0, 1[$; η имеет распределение Бернулли с параметром $p = 1/3$. Найти функцию распределения и, если существует, плотность распределения случайной величины $\zeta = \xi + \eta$.

Решение. Запишем функцию распределения величины ξ и ряд распределения величины η:

$$F_\xi = \begin{cases} 0, & x \leqslant 0; \\ x, & 0 < x \leqslant 1; \\ 1, & x > 1; \end{cases}$$

η	0	1
p	2/3	1/3

Заметим, что $\zeta = \xi + \eta$ может принимать значения от 0 до 2. Найдём функцию распределения ζ по определению.

Пусть $z \leqslant 0$. Тогда, очевидно, $F_\zeta(z) = 0$.

Пусть $0 < z \leqslant 1$. В этом случае $\zeta = \xi + \eta$ можно быть меньше z лишь тогда, когда величина η принимает значение, равное нулю. Таким образом,

$$F_\zeta(z) = P(\zeta < z) = P(\xi + \eta < z) = P(\eta = 0, \xi < z) =$$
$$= P(\eta = 0)P(\xi < z) = \frac{2}{3}F_\xi(z) = \frac{2}{3}z.$$

Пусть $1 < z \leqslant 2$. В этом случае, если $\eta = 0$, то сразу выполняется $\zeta < z$. Если же $\eta = 1$, то условие $\xi + \eta < z$ выполняется при $\xi < z - 1$. Таким образом,

$$F_\zeta(z) = P(\xi + \eta < z) = P((\eta = 0) \text{ или } (\eta = 1, \xi < z - 1)) =$$
$$= P(\eta = 0) + P(\eta = 1)P(\xi < z - 1) = \frac{2}{3} + \frac{1}{3}(z - 1).$$

Пусть $z > 2$. Тогда $F_\zeta(z) = 1$.

Итак,

$$F_\zeta(z) = \begin{cases} 0, & z \leqslant 0; \\ \dfrac{2}{3}z, & 0 < z \leqslant 1; \\ \dfrac{1}{3}(z+1), & 1 < z \leqslant 2; \\ 1, & z > 2. \end{cases}$$

Поскольку $F_\zeta(z)$ непрерывна, то ζ — непрерывная случайная величина. Следовательно,

$$f_\zeta(z) = F'_\zeta(z) = \begin{cases} 2/3, & z \in]0,1[; \\ 1/3, & z \in]1,2[; \\ 0, & z \notin]0,2[. \end{cases}$$

Приведём еще одно утверждение, которое имеет исключительно важное значение для моделирования случайных величин.

Теорема 17.1. *Если случайная величина ξ имеет непрерывную и строго монотонную функцию распределения $F_\xi(x)$, то величина $\eta = F_\xi(\xi)$ равномерно распределена на отрезке $[0,1]$, т.е. $\eta \in U_{0,1}$.*

Доказательство. Заметим, что величина $\eta \in [0,1]$ и в силу взаимно однозначного соответствия между ξ и η имеет также непрерывную функцию распределения $F_\eta(x)$. Пусть $0 < x < 1$, тогда $F_\eta(x) = P(\eta < x) = P(F_\xi(\xi) < x) = P(\xi < \xi_0)$, где ξ_0 определяется из условия $F_\xi(\xi_0) = x$. Но $P(\xi < \xi_0) = F_\xi(\xi_0) = x$. Таким образом,

$$F_\eta(x) = \begin{cases} 0, & x \leqslant 0; \\ x, & 0 < x \leqslant 1; \\ 1, & x > 1 \end{cases},$$

т.е. $F_\eta(x)$ есть функция распределения $U_{0,1}$.

Следствие 17.1.1. Если случайная величина ξ имеет непрерывную и строго монотонную функцию распределения $F_\xi(x)$, а величина $\eta \in U_{0,1}$, то $\xi = F_\xi^{-1}(\eta)$, где $F_\xi^{-1}(x)$ — функция, обратная к $F_\xi(x)$.

18. Экспоненциальное (показательное) распределение

♦ Случайная величина ξ распределена по *экспоненциальному закону* с параметром $\lambda > 0$ и обозначена $\xi \in E_\lambda$, если её плотность распределения вероятностей задаётся формулой

$$f(x) = \begin{cases} 0 & \text{при} \quad x < 0, \\ \lambda e^{-\lambda x} & \text{при} \quad x \geqslant 0. \end{cases} \tag{18.1}$$

Функция распределения случайной величины, распределённой по показательному закону,

$$F(x) = \begin{cases} 0 & \text{при} \quad x < 0, \\ 1 - e^{-\lambda x} & \text{при} \quad x \geqslant 0. \end{cases} \tag{18.2}$$

Экспоненциальное распределение реализуется в теории обслуживания, при этом ξ, например, — время ожидания при техническом обслуживании, и теории надёжности, где ξ, например, — срок службы электронной аппаратуры, компьютерной техники.

Показательное распределение тесно связано с простейшим (пуассоновским) потоком событий: интервал времени T между двумя соседними событиями в простейшем потоке имеет показательное распределение с параметром, равным интенсивности потока.

Найдём основные числовые характеристики показательного распределения. Для математического ожидания получим

$$M(\xi) = \int\limits_{-\infty}^{\infty} x f(x) dx = \lambda \int\limits_{0}^{\infty} x e^{-\lambda x} dx.$$

Вычислим интеграл по частям, положив $U = x$, $dV = e^{-\lambda x}dx$, $dU = dx$, $V = (-1/\lambda)e^{-\lambda x}$, и получим

$$M(\xi) = -\frac{x}{\lambda} e^{-\lambda x}\Big|_0^\infty - \Big(-\frac{1}{\lambda}\Big)\int\limits_0^\infty e^{-\lambda x} dx = \Big(-\frac{1}{\lambda}\Big) e^{-\lambda x}\Big|_0^\infty = \frac{1}{\lambda}.$$

Аналогично для дисперсии $D(\xi) = M(\xi^2) - [M(\xi)]^2$ получим

$$M(\xi^2) = \int\limits_{-\infty}^{\infty} x^2 f(x) dx = \lambda \int\limits_{0}^{\infty} x^2 e^{-\lambda x} dx.$$

Проинтегрируем по частям, положив $U = x^2$, $dV = e^{-\lambda x}dx$, $dU = 2x\,dx$, $V = (-1/\lambda)e^{-\lambda x}$, и получим

$$M(\xi^2) = \lambda \Big\{ x^2 \Big(-\frac{x}{\lambda}\Big) e^{-\lambda x}\Big|_0^\infty - \int\limits_0^\infty \Big(-\frac{2}{\lambda}\Big) e^{-\lambda x} 2x\,dx \Big\} = 2\int\limits_0^\infty x e^{-\lambda x} dx = \frac{2}{\lambda} M(\xi) = \frac{1}{\lambda^2}.$$

Следовательно,

$$M(\xi) = \frac{1}{\lambda}; \qquad D(\xi) = \frac{1}{\lambda^2}. \tag{18.3}$$

Аналогично можно получить $A(\xi) = 2$ и $E(\xi) = 6$.

Пример 18.1. Время безотказной работы персонального компьютера (ПК) — случайная величина T, имеющая показательное распределение с параметром $\lambda = 5$ (физический смысл величины λ — среднее число отказов в единицу времени, не считая простоев ПК для ремонта). Известно, что ПК уже проработал без отказов время τ. Найти при этих условиях плотность и функцию распределения времени, которое проработает ПК после момента τ до ближайшего отказа. Вычислить вероятность того, что ПК не выйдет из строя в течение среднего времени безотказной работы.

Решение. Число отказов ПК можно рассматривать как пуассоновский (простейший) поток событий, который обладает свойством отсутствия последствий. Поэтому вероятность появления хотя бы одного отказа на участке от τ до $\tau + t$ не зависит от того, появлялись ли отказы ранее момента τ. Следовательно, подставив $\lambda = 5$ в соотношения (18.1) и (18.2), получим

$$f(t) = \begin{cases} 0 & \text{при } t < 0, \\ 5e^{-5t} & \text{при } t \geqslant 0; \end{cases}$$

$$F(t) = \begin{cases} 0 & \text{при } t < 0, \\ 1 - e^{-5t} & \text{при } t \geqslant 0. \end{cases}$$

Графики плотности и функции полученного показательного распределения изображены на рис. 63.

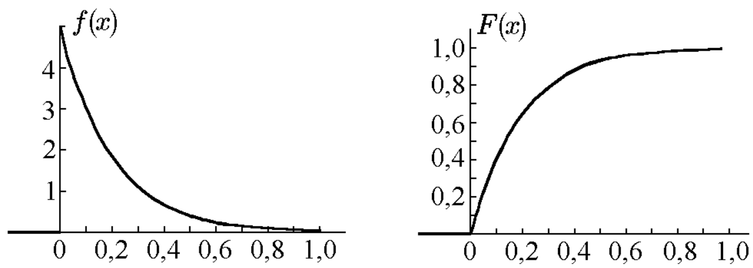


Рис. 63.

Среднее время безотказной работы ПК $M(\xi) = 1/\lambda = 0{,}2$. Тогда $P(0 \leqslant t \leqslant M(\xi)) = F(M(\xi)) - F(0) = 1 - e^{-1} \approx 0{,}63$.

Пример 18.2. Система состоит из двух независимых блоков, средний срок службы которых, соответственно, 3 и 5 лет. Найти функцию распределения случайной величины — срока службы системы — и средний срок службы системы, если сроки службы блоков распределены по показательному закону (для выхода системы из строя достаточно выхода из строя одного блока).

Решение. Пусть ξ — срок службы системы, ξ_1, ξ_2 — сроки службы блоков. Функция распределения срока службы системы $F_\xi(x)$, по определению, есть вероятность события, заключающегося в том, что система проработает меньше x лет. Так как система выйдет из строя в том случае, если хотя бы один блок выйдет из строя, то, по определению, событие $\{\xi < x\}$ есть сумма событий

$\{\xi_1 < x\}$ и $\{\xi_2 < x\}$. Поскольку события $\{\xi_1 < x\}$ и $\{\xi_2 < x\}$ совместны, по общей теореме сложения вероятностей получим

$$\begin{gathered}F_\xi(x) = P(\xi < x) = P(\xi_1 < x \text{ или } \xi_2 < x) = \\ = P(\xi_1 < x) + P(\xi_2 < x) - P(\xi_1 < x \text{ и } \xi_2 < x) = \\ = P(\xi_1 < x) + P(\xi_2 < x) - P(\xi_1 < x)P(\xi_2 < x) = F_{\xi_1}(x) + F_{\xi_2}(x) - F_{\xi_1}(x)F_{\xi_2}(x).\end{gathered}$$

Для случайной величины ξ, распределённой по показательному закону $F_\xi(x) = 1 - e^{-\alpha x}$, $x > 0$, $\alpha = 1/M(\xi)$, где $M(\xi)$ — математическое ожидание случайной величины. В нашем случае $M(\xi_1) = 3$, $M(\xi_2) = 5$. Тогда

$$F_\xi(x) = (1 - e^{-x/3}) + (1 - e^{-x/5}) - (1 - e^{-x/3})(1 - e^{-x/5}) = 1 - e^{-8x/15}, \quad x > 0.$$

Получили, что ξ также распределена по показательному закону с параметром $\alpha = 8/15$. Следовательно, средний срок службы системы $M(\xi) = 1/\alpha = 15/8 = 1{,}875$ лет.

Пример 18.3. Продолжительность работы микросхемы — случайная величина, распределённая по показательному закону с параметром $\alpha = 0{,}01$ ч$^{-1}$. Испорченную микросхему немедленно заменяют новой. Какова вероятность того, что за 100 часов микросхему придется заменять больше двух раз?

Решение. Поскольку продолжительность работы микросхемы — случайная величина, распределённая по показательному закону, то поток отказов микросхем является простейшим с интенсивностью $\alpha = 0{,}01$ мс./ч. Тогда, используя формулу Пуассона с параметром $\lambda = 0,01 \cdot 100 = 1$, найдём вероятность того, что число отказов ξ будет больше двух:

$$\begin{gathered}P(\xi > 2) = 1 - P(\xi \leqslant 2) = 1 - P(\xi = 0) - P(\xi = 1) - P(\xi = 2) = \\ = 1 - \frac{1^0 e^{-0}}{0!} - \frac{1^1 e^{-1}}{1!} - \frac{1^2 e^{-2}}{2!} \approx 0{,}323.\end{gathered}$$

Пример 18.4. Случайная величина ξ распределена по показательному закону с параметром α. Найти плотность распределения случайной величины $\eta = \xi^2$.

Решение. Функция $z = x^2$ монотонна для всех $x > 0$. Обратная функция и её производная имеют вид $x = \sqrt{z}$, $x' = 1/(2\sqrt{z})$. Учтя, что $f_\xi(x) = \alpha e^{-\alpha x}$, и подставив в формулу (21.1), получим

$$f_\eta(x) = \frac{\alpha}{2\sqrt{x}} e^{-\alpha\sqrt{x}}, \quad x > 0.$$

Пример 18.5. Случайные величины ξ и η независимы; ξ имеет равномерное распределение на $]1, 3[$, η — показательное распределение с параметром $\lambda = 1$. Найти плотность распределения случайной величины $\zeta = \xi + \eta$.

Решение. Поскольку ξ и η имеют непрерывные распределения, воспользуемся формулой свертки. Имеем

$$f_\xi(x) = \begin{cases} 1/2, & x \in]1, 3[; \\ 0, & x \notin]1, 3[; \end{cases}$$

$$f_\eta(y) = \begin{cases} e^{-y}, & y \in]0, \infty[; \\ 0, & y \notin]0, \infty[. \end{cases}$$

Заметим, что $\zeta = \xi + \eta$ принимает значения от 1 до ∞, следовательно, $f_\zeta(z) = 0$ для $z < 1$.

Пусть $1 < z < 3$. Тогда, чтобы выполнялись условия $f_\xi(x) > 0$, $f_\eta(y) = f_\eta(z-x) > 0$, необходимо $1 < x < z$. С учетом (11.15) запишем

$$f_\zeta(z) = \int_{-\infty}^{\infty} f_\xi(x) f_\eta(z-x)dx = \int_1^z \frac{1}{2}e^{-(z-x)}dx = \frac{1}{2}(1 - e^{1-z}).$$

Пусть $z > 3$. Тогда $f_\xi(x)f_\eta(y) > 0$ для $1 < x < 3$, следовательно,

$$f_\zeta(z) = \int_{-\infty}^{\infty} f_\xi(x) f_\eta(z-x)dx = \int_1^3 \frac{1}{2}e^{-(z-x)}dx = \frac{1}{2}e^{x-z}\Big|_1^3 = \frac{1}{2}(e^{3-z} - e^{1-z}).$$

Окончательно получим

$$f_\zeta(z) = \begin{cases} 0, & z < 1; \\ \dfrac{1}{2}(1 - e^{1-z}), & 1 < z < 3; \\ \dfrac{1}{2}(e^{3-z} - e^{1-z}), & z > 3. \end{cases}$$

19. Нормальное распределение

19.1. Плотность вероятности

Среди распределений непрерывных случайных величин центральное место занимает нормальное распределение, или закон Гаусса.

Случайная величина ξ *нормально распределена* или *подчиняется закону распределения Гаусса*, если её плотность распределения $f(x)$ имеет вид

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-m)^2/2\sigma^2}, \quad x \in \mathbb{R}, \tag{19.1}$$

где m, σ — параметры распределения, причём $\sigma > 0$.

Принадлежность ξ к нормальному распределению с параметрами m и σ обозначается обычно $\xi \in N_{m,\sigma^2}$ или $\xi \in N(m, \sigma^2)$.

Рассмотрим некоторые особенности функции $f(x)$ и построим её график:

1) функция $f(x)$ определена для всех $x \in \mathbb{R}$;

2) график $f(x)$ симметричен относительно прямой $x = m$;

3) функция $f(x)$ имеет максимум в точке $x = m$, равный

$$\max_{x=m} f(x) = \frac{1}{\sigma\sqrt{2\pi}};$$

4) график $f(x)$ имеет две точки перегиба: $x = m \pm \sigma$; ординаты точек перегиба одинаковы и равны

$$f(m \pm \sigma) = \frac{1}{\sigma\sqrt{2\pi e}};$$

5) ось Ox служит горизонтальной асимптотой графика функции $f(x)$:

$$\lim_{x\to\pm\infty} f(x) = 0;$$

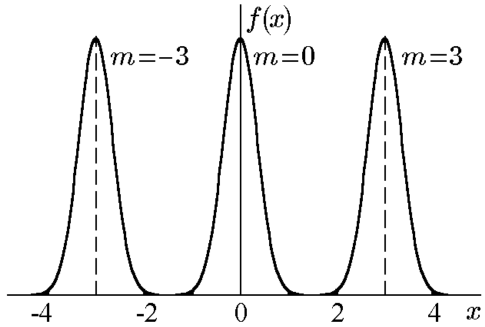


Рис. 64.

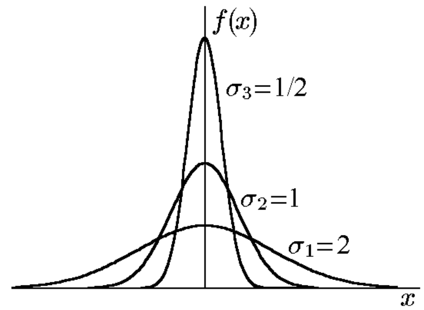


Рис. 65.

6) при постоянном σ с изменением m кривые плотности не меняют форму, но сдвинуты относительно оси Oy вправо при $m > 0$ или влево при $m < 0$ (см. рис. 64); с возрастанием же параметра σ кривая становится более пологой, сжимается к оси Ox и растягивается вдоль неё (см. рис. 65).

Примерами случайных величин, имеющих нормальное распределение, могут служить:

1) ошибки измерений;

2) величина помехи на входе радиоприемного устройства в данный момент;

3) отклонение при стрельбе;

4) многие параметры радиодеталей (ёмкость, индуктивность, сопротивление) при массовом производстве;

5) отклонения действительных размеров деталей, обработанных на станке, от номинальных.

Проверим, что

$$\int\limits_{-\infty}^{\infty} f(x)dx = 1. \tag{19.2}$$

Действительно, сделав замену переменной $(x - m)/\sigma\sqrt{2} = t$, получим $dx = \sigma\sqrt{2}dt$ и

$$\int\limits_{-\infty}^{\infty} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-(x-m)^2/2\sigma^2} dx = \frac{1}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} e^{-t^2} dt.$$

При изучении несобственных интегралов было установлено, что

$$\int\limits_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

Тогда

$$\int\limits_{-\infty}^{\infty} f(x)dx = \frac{1}{\sqrt{\pi}}\sqrt{\pi} = 1.$$

Отметим, что ввиду особой важности рассматриваемого закона значения плотности $f(x)$ при $m = 0$ и $\sigma = 1$, т.е. значения функции

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \tag{19.3}$$

затабулированы.

Для общего случая

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\Big(\frac{x-m}{\sigma}\Big)^2\right]=\frac{1}{\sigma}\varphi\Big(\frac{x-m}{\sigma}\Big),$$

поэтому значения $f(x)$ также могут быть найдены с помощью таблицы значений $\varphi(x)$.

19.2. Функция распределения

Функция распределения нормального закона имеет вид

$$F(x)=\int\limits_{-\infty}^{x}f(t)dt=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{x}e^{-(t-m)^2/2\sigma^2}dt. \tag{19.4}$$

Интеграл, стоящий в правой части, не может быть сведён к элементарным функциям в конечном виде, однако приводится к специальным функциям, которые табулируются.

Обозначим

$$z=\frac{t-m}{\sigma},\quad dz=\frac{1}{\sigma}dt,\quad t_1=x,\quad z_1=\frac{x-m}{\sigma},\quad t_2=-\infty,\quad z_2=-\infty.$$

Тогда

$$F(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{(x-m)/\sigma}e^{-z^2/2}dz=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{0}e^{-z^2/2}dz+\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{(x-m)/\sigma}e^{-z^2/2}dz. \tag{19.5}$$

Положим в первом интеграле $z=-y$, тогда

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{0}e^{-z^2/2}dz=-\frac{1}{\sqrt{2\pi}}\int\limits_{\infty}^{0}e^{-y^2/2}dy=\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{\infty}e^{-y^2/2}dy=\Phi(\infty)=\frac{1}{2}$$

на основании (19.2). Здесь

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{x}e^{-z^2/2}dz \tag{19.6}$$

есть функция Лапласа, для которой составлены таблицы значений. Второе слагаемое в (19.5) отличается от (19.6) только верхним пределом интегрирования. Поэтому функцию распределения нормального закона можно представить в виде

$$F(x)=\frac{1}{2}+\Phi\Big(\frac{x-m}{\sigma}\Big). \tag{19.7}$$

Функция Лапласа часто встречается в расчётах. Поэтому мы приведём некоторые её свойства (без доказательства):

1. $\Phi(0)=0$;
2. $\Phi(-x)=-\Phi(x)$;

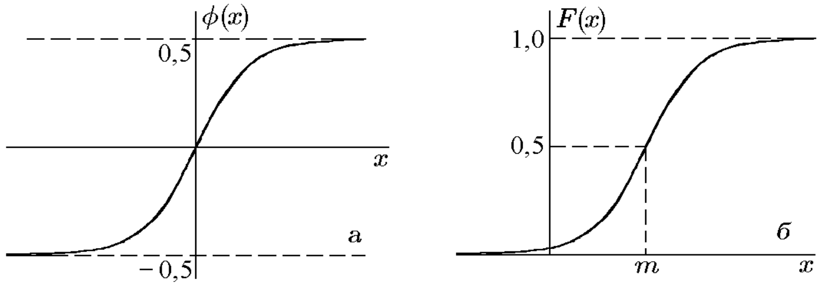


Рис. 66.

3. $\Phi(\infty) = \dfrac{1}{2}$;

4. $\dfrac{d\Phi(x)}{dx} = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2} > 0$ для всех $x \in]-\infty, \infty[$.

Функция Лапласа монотонно возрастает (рис. 66,*а*).

График функции распределения нормальной случайной величины $F(x)$ получается из графика функции Лапласа при помощи сдвига вдоль оси Oy на $1/2$ в положительном направлении (рис. 66,*б*).

Покажем, что нормальное распределение является устойчивым по суммированию.

Теорема 19.1. *Если ξ и η независимы и $\xi \in N(m_1, \sigma_1^2)$, $\eta \in N(m_2, \sigma_2^2)$, то*

$$\zeta = \xi + \eta \in N(m_1 + m_2, \sigma_1^2 + \sigma_2^2).$$

Доказательство. Докажем теорему для $\xi \in N(0,1)$, $\eta \in N(0,1)$ (доказательство для общего случая аналогично). По формуле свертки (11.15) имеем

$$f_{\xi+\eta}(z) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{-x^2/2} e^{-(z-x)^2/2} dx = \frac{1}{2\pi} e^{-z^2/2} \int\limits_{-\infty}^{\infty} e^{-x^2+zx} dx =$$

$$= \frac{1}{2\pi} e^{-z^2/2} \int\limits_{-\infty}^{\infty} e^{-(x-z/2)^2+z^2/4} dx =$$

$$= \frac{1}{2\pi} e^{-z^2/4} \int\limits_{-\infty}^{\infty} e^{-(x-z/2)^2} dx = \frac{1}{2\pi} e^{-z^2/4} \sqrt{\pi} = \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-z^2/2\cdot 2}.$$

Полученное выражение есть плотность нормального распределения с параметрами $m = 0$ и $\sigma^2 = 2$, что и требовалось доказать.

19.3. Числовые характеристики

Определим математическое ожидание, дисперсию и среднеквадратичное отклонение нормального распределения.

Математическое ожидание определяется соотношением

$$M(\xi) = \int\limits_{-\infty}^{\infty} x f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} x e^{-(x-m)^2/2\sigma^2} dx.$$

Сделаем замену переменных:

$$\frac{x-m}{\sigma}=z, \quad x=\sigma z+m, \quad dx=\sigma dz.$$

Тогда

$$M(\xi)=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}(\sigma z+m)e^{-z^2/2}\sigma dz=\frac{\sigma}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}ze^{-z^2/2}dz+\frac{m}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}e^{-z^2/2}dz.$$

Первый интеграл справа равен нулю как интеграл от нечётной функции в симметричных пределах, а второй интеграл есть известный интеграл Эйлера–Пуассона

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}e^{-z^2/2}dz=1. \tag{19.8}$$

Поэтому

$$M(\xi)=m. \tag{19.9}$$

Таким образом, параметр m в нормальном законе распределения равен математическому ожиданию случайной величины.

Дисперсия нормально распределённой случайной величины

$$D(\xi)=\int\limits_{-\infty}^{\infty}(x-M(\xi))^2f(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}(x-m)^2e^{-(x-m)^2/2\sigma^2}dx.$$

Произведя в интеграле ту же замену переменных, что и в предыдущем случае, и приняв во внимание, что пределы интегрирования не меняются, получим

$$D(\xi)=\frac{\sigma^2}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}z^2e^{-z^2/2}dz=\frac{\sigma^2}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}z\big(ze^{-z^2/2}dz\big).$$

Положив $U=z$, $ze^{-z^2/2}dz=dV$ и проинтегрировав по частям, найдём

$$D(\xi)=\frac{\sigma^2}{\sqrt{2\pi}}\Big[-ze^{-z^2/2}\Big|_{-\infty}^{\infty}+\int\limits_{-\infty}^{\infty}e^{-z^2/2}dz\Big].$$

Так как первое слагаемое в квадратных скобках равно нулю, а второе – уже известный интеграл (19.8), равный $\sqrt{2\pi}$, то

$$D(\xi)=\sigma^2. \tag{19.10}$$

Следовательно, второй параметр σ в выражении для плотности нормально распределённой случайной величины есть среднеквадратичное отклонение (стандартное отклонение этой случайной величины)

$$\sigma=\sqrt{D(\xi)}. \tag{19.11}$$

Аналогично вычисляются центральные моменты k-го порядка:

$$\mu_k = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} (x-m)^k e^{-(x-m)^2/2\sigma^2} dx.$$

Обозначим $x - m = \sqrt{2}\sigma t$, тогда $dx = \sqrt{2}\sigma\, dt$ и

$$\mu_k = \frac{1}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} (\sqrt{2}\sigma)^k t^k e^{-t^2} dt.$$

Вычислим интеграл по частям. Положим $U = t^{k-1}$, $dV = te^{-t^2}dt$, тогда $dU = (k-1)t^{t-2}dt$ и $V = -e^{t^2}/2$, а

$$\mu_k = \frac{(\sqrt{2}\sigma)^k}{2\sqrt{\pi}} \left(-t^{k-1}e^{-t^2}\Big|_{-\infty}^{\infty} + (k-1)\int\limits_{-\infty}^{\infty} t^{k-2}e^{-t^2}dt \right) = \frac{(\sqrt{2}\sigma)^k}{2\sqrt{\pi}}(k-1)\int\limits_{-\infty}^{\infty} t^{k-2}e^{-t^2}dt =$$

$$= (k-1)\sigma^2 \frac{1}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} (\sqrt{2}\sigma)^{k-2} t^{k-2} e^{-t^2} dt = (k-1)\sigma^2 \mu_{k-2}.$$

Таким образом, для центральных моментов нормальной случайной величины справедлива следующая рекуррентная формула:

$$\mu_k = (k-1)\sigma^2 \mu_{k-2}. \tag{19.12}$$

Так как нормальное распределение симметрично относительно математического ожидания, то все центральные моменты нечётных порядков равны нулю, следовательно, коэффициент асимметрии $A = 0$. Для вычисления эксцесса найдём по рекуррентной формуле (19.12) при $k = 4$ четвёртый центральный момент: $\mu_4 = 3\sigma^2\mu_2 = 3\sigma^4$, откуда

$$E = \frac{\mu_4}{\sigma^4} - 3 = 0.$$

В этом смысле кривая плотности нормального распределения является эталонной: $A = 0$, $E = 0$.

19.4. Вероятность попадания нормальной случайной величины в заданный интервал

Известно, что если случайная величина задана плотностью вероятности $f(x)$, то вероятность попадания величины ξ на участок $]\alpha, \beta[$ равна

$$P(\alpha \leqslant \xi < \beta) = \int\limits_{\alpha}^{\beta} f(x)dx = F(\beta) - F(\alpha).$$

Для нормальной случайной величины

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

Поэтому

$$P(\alpha \leqslant \xi < \beta) = \frac{1}{2} + \Phi\Big(\frac{\beta - m}{\sigma}\Big) - \frac{1}{2} - \Phi\Big(\frac{\alpha - m}{\sigma}\Big) = \Phi\Big(\frac{\beta - m}{\sigma}\Big) - \Phi\Big(\frac{\alpha - m}{\sigma}\Big),$$

где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-t^2/2} dt$$

— функция Лапласа. Следовательно,

$$P(\alpha \leqslant \xi < \beta) = \Phi\Big(\frac{\beta - m}{\sigma}\Big) - \Phi\Big(\frac{\alpha - m}{\sigma}\Big). \tag{19.13}$$

Пример 19.1. Случайная величина ξ распределена по закону $N(1, 4)$. Найти $P(2 \leqslant \xi < 3)$.

Решение. Имеем $m = 1$, $\sigma = 2$, $\alpha = 2$, $\beta = 3$. Следовательно, согласно формуле (19.13),

$$P(2 \leqslant \xi < 3) = \Phi\Big(\frac{3-1}{2}\Big) - \Phi\Big(\frac{2-1}{2}\Big) \approx 0{,}3413 - 0{,}1915 = 0{,}1498.$$

19.5. Правило трёх сигма

С помощью величин m и σ можно вывести два очень полезных правила:
1) 95% распределения лежит между значениями $m - 2\sigma$ и $m + 2\sigma$;
2) более 99% распределения заключено между $m - 3\sigma$ и $m + 3\sigma$.

Последнее правило известно под названием «правила трёх сигма».

Из формулы (19.13) следует

$$\begin{gathered} P(|\xi - m| < \lambda\sigma) = P(-\lambda\sigma < \xi - m < \lambda\sigma) = \\ = P(m - \lambda\sigma < \xi < m + \lambda\sigma) = F(m + \lambda\sigma) - F(m - \lambda\sigma) = \\ = \Phi\Big(\frac{m + \lambda\sigma - m}{\sigma}\Big) - \Phi\Big(\frac{m - \lambda\sigma - m}{\sigma}\Big) = \Phi(\lambda) - \Phi(-\lambda). \end{gathered}$$

Учитывая, что $\Phi(x)$ — нечётная функция ($\Phi(-x) = -\Phi(x)$), будем иметь

$$P(|\xi - m| < \lambda\sigma) = 2\Phi(\lambda). \tag{19.14}$$

Используя таблицу для $\Phi(x)$, в частности, получим

$$\text{при } \lambda = 1 \quad P(|\xi - m| < \sigma) = 2\Phi(1) = 0{,}6827;$$

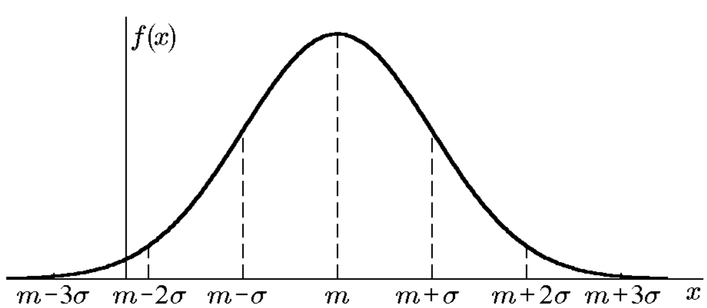


Рис. 67.

при $\lambda = 2$ $P(|\xi - m| < 2\sigma) = 2\Phi(2) = 0{,}9545$;
при $\lambda = 3$ $P(|\xi - m| < 3\sigma) = 2\Phi(3) = 0{,}9973$;
при $\lambda = 4$ $P(|\xi - m| < 4\sigma) = 2\Phi(4) = 0{,}99994$.

Вероятность того, что случайная величина ξ меньше чем на 3σ отличается от своего математического ожидания m, составляет 0,9973, т.е. почти равна единице. Таким образом, «практически почти достоверно», т.е. с уверенностью 99,73% распределённая по нормальному закону с параметрами (m, σ) случайная величина ξ принимает лишь значения в пределах «трёх сигмовых границ», т.е. в промежутке $(m - 3\sigma, m + 3\sigma)$.

Отклонение по модулю больше, чем 3σ, встречается в среднем три раза из тысячи.

19.6. Стандартное нормальное распределение

Нормальное распределение $N(m, \sigma^2)$, где $m = 0$, а $\sigma = 1$, называется *стандартным нормальным распределением*.

Плотность стандартного распределения для всех $x \in \mathbb{R}$ задаётся формулой

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

а функция распределения

$$F_{0,1}(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2} dt. \tag{19.15}$$

Нетрудно заметить, что

$$F_{0,1}(x) = \Phi(x) + \frac{1}{2}, \tag{19.16}$$

поэтому для случайной величины ξ, имеющей нормальное распределение $N(m, \sigma^2)$, запишем

$$F_\xi(x) = F_{m,\sigma^2}(x) = F_{0,1}\Big(\frac{x - m}{\sigma}\Big);$$
$$P(\alpha < \xi < \beta) = F_{0,1}\Big(\frac{\beta - m}{\sigma}\Big) - F_{0,1}\Big(\frac{\alpha - m}{\sigma}\Big);$$
$$P(|\xi - m| < \delta) = 2F_{0,1}\Big(\frac{\delta}{\sigma}\Big) - 1.$$

19.7. Характеристическая функция нормальной величины

Найдём характеристическую функцию нормальной величины. Для стандартной нормальной величины

$$\gamma(t) = M(e^{it\xi}) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{itx} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{itx - x^2/2} dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-(x^2 - 2itx)/2} dx = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-(x^2 - 2itx - t^2)/2 - t^2/2} dx =$$

$$= \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \int\limits_{-\infty}^{\infty} e^{-(x-it)^2/2} dx.$$

Сделав замену переменных $z = (x - it)/\sqrt{2}$, получим

$$\gamma(t) = \frac{1}{\sqrt{\pi}} e^{-t^2/2} \int\limits_{-\infty}^{\infty} e^{-z^2} dz = \frac{1}{\sqrt{\pi}} e^{-t^2/2} \sqrt{\pi} = e^{-t^2/2}.$$

Переход от интегрирования по прямой $y = -it/\sqrt{2}$ к интегрированию по прямой $y = 0$ оправдывается аналитичностью подынтегральной функции в части плоскости, ограниченной этими прямыми, и возможностью деформировать контур интегрирования в области аналитичности, согласно теореме Коши. Итак,

$$\gamma(t) = e^{-t^2/2}. \tag{19.17}$$

Так как для случайной величины $\eta = m\xi + b$ характеристическая функция есть $\gamma_\eta(t) = \gamma_\xi(mt)e^{ibt}$, то, положив $\eta = \sigma\xi + m$, с учетом (12.18) получим характеристическую функцию произвольного нормального распределения $N(m, \sigma^2)$:

$$\gamma(t) = e^{-\sigma^2 t^2/2} e^{imt} = e^{imt - \sigma^2 t^2/2}. \tag{19.18}$$

Нормальное распределения непрерывной случайной величины имеет очень широкое распространение в случайных природных явлениях, так как ему подчиняется распределение случайной величины, представленной в виде суммы слабо зависимых случайных величин. На практике очень часто встречаются случайные величины, образующиеся именно в результате суммирования многих случайных слагаемых, сравнимых по степени своего влияния на их сумму.

Рассмотрим далее некоторые важные распределения, порождённые нормальным распределением непрерывной случайной величины.

19.8. Интеграл ошибок

Для полноты изложения приведём сведения о специальных функциях, связанных с функцией Лапласа и иногда используемых в литературе по теории вероятностей и математической статистике.

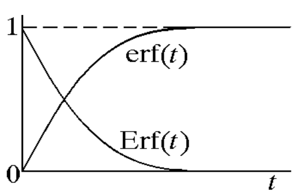


Рис. 68.

♦ *Интегралом (функцией) вероятности*, или *интегралом ошибок*, называется функция

$$\mathrm{erf}\,(t) = \frac{2}{\sqrt{\pi}} \int\limits_0^t e^{-x^2} dx; \tag{19.19}$$

$$\mathrm{erf}\,(t) = \frac{2}{\sqrt{\pi}} \int\limits_0^t e^{-x^2} dx = 2\Phi(t\sqrt{2}). \tag{19.20}$$

Непосредственно из определения следует, что функция $\mathrm{erf}\,(t)$ непрерывна и монотонно возрастает от нуля до $\mathrm{erf}\,(\infty) = 1$, так как

$$\mathrm{erf}\,(\infty) = \frac{2}{\sqrt{\pi}} \int\limits_0^\infty e^{-x^2} dx = \frac{1}{\sqrt{\pi}} \Gamma\Big(\frac{1}{2}\Big) = 1. \tag{19.21}$$

Наряду с функцией $\operatorname{erf}(t)$ рассматривают функцию

$$\operatorname{Erf}(t) = 1 - \operatorname{erf}(t) = \frac{2}{\sqrt{\pi}}\Big[\frac{\sqrt{\pi}}{2} - \int\limits_0^t e^{-x^2}dx\Big]$$

или с учётом (19.21)

$$\operatorname{Erf}(t) = \frac{2}{\sqrt{\pi}}\Big[\int\limits_0^\infty e^{-x^2}dx - \int\limits_0^t e^{-x^2}dx\Big] = \frac{2}{\sqrt{\pi}}\int\limits_t^\infty e^{-x^2}dx. \tag{19.22}$$

Функция $\operatorname{Erf}(t)$ непрерывна и монотонно убывает от $\operatorname{Erf}(0) = 1$ до $\operatorname{Erf}(\infty) = 0$ (см. рис. 68).

Пример 19.2. Найти представление интеграла ошибок в виде ряда по степеням t.

Решение. Согласно (19.19), имеем

$$\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}}\int\limits_0^t \sum_{k=0}^\infty (-1)^k \frac{x^{2k}}{k!}\,dx = \frac{2}{\sqrt{\pi}}\sum_{k=1}^\infty (-1)^{k+1}\frac{t^{2k-1}}{(2k-1)(k-1)!}.$$

Заметив, что

$$\sum_{k=1}^\infty (-1)^{k+1}\frac{t^{2k-1}}{(2k-1)(k-1)!} = \sum_{l=0}^\infty \frac{(-1)^l t^{2l}}{l!}\sum_{s=0}^\infty \frac{2^s t^{2s+1}}{(2s+1)!!},$$

получим

$$\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}}e^{-t^2}\sum_{s=0}^\infty \frac{2^s t^{2s+1}}{(2s+1)!!}.$$

20. Многомерное нормальное распределение

♦ Пусть $\vec{\xi} = \{\xi_1, \xi_2, \ldots, \xi_n\}$ — случайный вектор, имеющий непрерывное распределение; $\vec{\mu}$ — вектор математических ожиданий величин $\xi_1, \xi_2, \ldots, \xi_n$, а A — матрица ковариаций величин $\xi_1, \xi_2, \ldots, \xi_n$. Говорят, что $\vec{\xi}$ *распределён нормально* с параметрами $\vec{\mu}$ и A (обозначают $\vec{\xi} \in N(\vec{\mu}, A)$), если его плотность распределения имеет вид

$$f_{\vec{\xi}}(\vec{x}) = \Big(\frac{1}{2\pi}\Big)^{n/2}\frac{1}{\sqrt{\det A}}\exp\Big[-\frac{1}{2}(\vec{x}-\vec{\mu})^\intercal A^{-1}(\vec{x}-\vec{\mu})\Big].$$

Пусть случайный вектор $\{\xi, \eta\}$ имеет двумерное нормальное распределение и матрица ковариаций системы $\{\xi, \eta\}$ имеет вид

$$A = \begin{pmatrix} \sigma_\xi^2 & K_{\xi\eta} \\ K_{\xi\eta} & \sigma_\eta^2 \end{pmatrix}.$$

Тогда определитель матрицы ковариаций есть $\det A = \sigma_\xi^2\sigma_\eta^2 - K_{\xi\eta}^2 = \sigma_\xi^2\sigma_\eta^2(1-\rho_{\xi\eta}^2)$ и обратная матрица для матрицы ковариаций:

$$A^{-1} = \frac{1}{\sigma_\xi^2\sigma_\eta^2(1-\rho_{\xi\eta}^2)}\begin{pmatrix} \sigma_\eta^2 & -K_{\xi\eta} \\ -K_{\xi\eta} & \sigma_\xi^2 \end{pmatrix}.$$

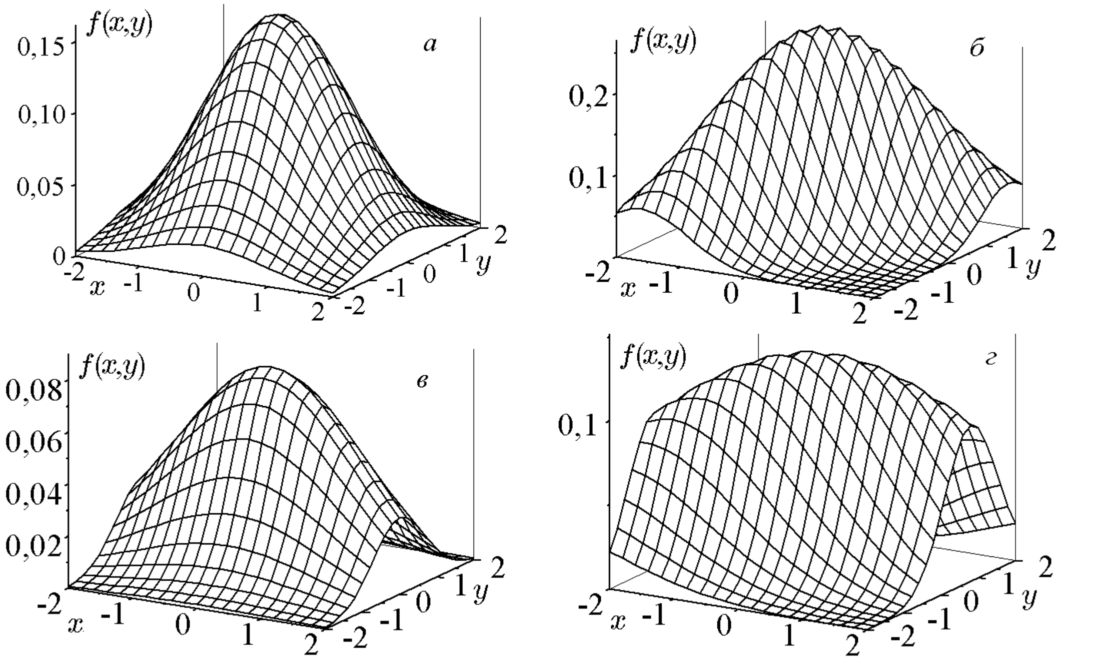


Рис. 69. Графики плотностей двумерного нормального распределения при $m_x = m_y = 0$ и $\sigma_x = \sigma_y = 1$, $\rho = 0$ (*а*); $\sigma_x = \sigma_y = 1$, $\rho = 0{,}8$ (*б*); $\sigma_x = 2$, $\sigma_y = 1$, $\rho = 0$ (*в*); $\sigma_x = 2$, $\sigma_y = 1$, $\rho = 0{,}8$ (*г*)

Пусть

$$\vec{\mu} = \begin{pmatrix} m_\xi \\ m_\eta \end{pmatrix}, \qquad \vec{x} = \begin{pmatrix} x \\ y \end{pmatrix},$$

тогда

$$(\vec{x} - \vec{\mu})^\intercal A^{-1} (\vec{x} - \vec{\mu}) = \frac{1}{\sigma_\xi^2 \sigma_\eta^2 (1 - \rho_{\xi\eta}^2)} \begin{pmatrix} x - m_\xi \\ y - m_\eta \end{pmatrix}^\intercal \begin{pmatrix} \sigma_\eta^2 & -K_{\xi\eta} \\ -K_{\xi\eta} & \sigma_\xi^2 \end{pmatrix} \begin{pmatrix} x - m_\xi \\ y - m_\eta \end{pmatrix} =$$

$$= \frac{1}{\sigma_\xi^2 \sigma_\eta^2 (1 - \rho_{\xi\eta}^2)} [(x - m_\xi)^2 \sigma_\eta^2 - 2(x - m_\xi)(y - m_\eta) K_{\xi\eta} + (y - m_\eta)^2 \sigma_\xi^2] =$$

$$= \frac{1}{1 - \rho_{\xi\eta}^2} \Big[\Big(\frac{x - m_\xi}{\sigma_\xi} \Big)^2 - 2\rho_{\xi\eta} \frac{x - m_\xi}{\sigma_\xi} \frac{y - m_\eta}{\sigma_\eta} + \Big(\frac{y - m_\eta}{\sigma_\eta} \Big)^2 \Big].$$

♦ Двумерная случайная величина $\{\xi, \eta\}$ *распределена нормально*, если её плотность вероятности $f_{\xi,\eta}(x, y)$ имеет вид

$$f_{\xi,\eta}(x, y) = \frac{1}{2\pi \sigma_\xi \sigma_\eta \sqrt{1 - \rho_{\xi\eta}^2}} \exp\Big\{ -\frac{1}{2(1 - \rho_{\xi\eta}^2)} \Big[\Big(\frac{x - m_\xi}{\sigma_\xi} \Big)^2 - $$

$$ - 2\rho_{\xi\eta} \frac{x - m_\xi}{\sigma_\xi} \frac{y - m_\eta}{\sigma_\eta} + \Big(\frac{y - m_\eta}{\sigma_\eta} \Big)^2 \Big] \Big\},$$

где m_ξ, m_η — математические ожидания величин ξ и η, σ_ξ, σ_η — их среднеквадратичные отклонения, а $\rho_{\xi\eta}$ — коэффициент корреляции между ξ и η, $x \in \mathbb{R}$, $y \in \mathbb{R}$. Графики плотности двумерного нормального распределения для различных значений коэффициентов приведены на рис. 69.

Свойства многомерного нормального распределения

Свойство 1. Для того чтобы компоненты многомерной нормальной случайной величины были независимы, необходимо и достаточно, чтобы корреляция между ними была равна нулю.

Доказательство. Необходимость вытекает непосредственно из того, что для любых случайных величин из их независимости следует некоррелированность. Докажем достаточность. Пусть ξ_i, $i = \overline{1,n}$, некоррелированы. Тогда A — диагональная матрица с диагональными элементами $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$. Соответственно обратная матрица — также диагональная с элементами $1/\sigma_1^2, 1/\sigma_2^2, \ldots, 1/\sigma_n^2$, а $\det A = \sigma_1^2 \sigma_2^2 \cdots \sigma_n^2$. Тогда

$$(\vec{x} - \vec{\mu})^\intercal A^{-1} (\vec{x} - \vec{\mu}) = \sum_{i=1}^{n} \frac{(x_i - \mu_i)^2}{\sigma_i^2}$$

или

$$f_\xi(x_1, x_2, \ldots, x_n) = \frac{(2\pi)^{-n/2}}{\sigma_1 \sigma_2 \cdots \sigma_n} \exp\left[\sum_{i=1}^{n} (x_i - \mu_i)^2 \sigma_i^2\right] = f_{\xi_1}(x_1) f_{\xi_2}(x_2) \cdots f_{\xi_n}(x_n),$$

т.е. случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ независимы.

Таким образом, для компонент многомерного нормального случайного вектора некоррелированность равносильна независимости (вспомним, что для произвольных величин из некоррелированности не следовала в общем случае независимость). Отметим также, что, согласно определению, не любая совокупность нормальных случайных величин может образовывать нормальный вектор.

Свойство 2. Если $\vec{\xi} \in N(\vec{\mu}, A)$, то любой подвектор вектора $\vec{\xi}$ будет нормально распределённой случайной величиной.

Докажем это свойство для двумерной случайной величины.

Теорема 20.1. *Если вектор $\{\xi, \eta\}$ имеет двумерное нормальное распределение*

$$\{\xi, \eta\} \in N\left[\begin{pmatrix} m_\xi \\ m_\eta \end{pmatrix}, \begin{pmatrix} \sigma_\xi^2 & K_{\xi\eta} \\ K_{\xi\eta} & \sigma_\eta^2 \end{pmatrix}\right],$$

то величины ξ и η также распределены нормально: $\xi \in N(m_\xi, \sigma_\xi^2)$, $\eta \in N(m_\eta, \sigma_\eta^2)$.

Доказательство. По свойству плотности распределения

$$f_\xi(x) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x, y) dy.$$

Сделаем замену $u = (y - m_\eta)/(\sigma_\eta)$, получим

$$f_\xi(x) = \frac{1}{2\pi\sigma_\xi\sqrt{1 - \rho_{\xi\eta}^2}} \int_{-\infty}^{\infty} \exp\Big\{ -\frac{1}{2(1 - \rho_{\xi\eta}^2)} \Big[\Big(\frac{x - m_\xi}{\sigma_\xi}\Big)^2 - 2\rho_{\xi\eta} u \frac{x - m_\xi}{\sigma_\xi} + u^2\Big]\Big\} du.$$

Выделим полный квадрат в показателе экспоненты:

$$f_\xi(x) = \frac{1}{2\pi\sigma_\xi\sqrt{1 - \rho_{\xi\eta}^2}} \int_{-\infty}^{\infty} \exp\Big\{ -\frac{1}{2(1 - \rho_{\xi\eta}^2)} \Big[\Big(u - \rho_{\xi\eta}\frac{x - m_\xi}{\sigma_\xi}\Big)^2 + (1 - \rho_{\xi\eta}^2)\Big(\frac{x - m_\xi}{\sigma_\xi}\Big)^2\Big]\Big\} du =$$

$$= \frac{1}{2\pi\sigma_\xi\sqrt{1 - \rho_{\xi\eta}^2}} \exp\Big[-\frac{(x - m_\xi)^2}{2\sigma_\xi}\Big] \int_{-\infty}^{\infty} \exp\Big[-\frac{1}{2(1 - \rho_{\xi\eta}^2)}\Big(u - \rho_{\xi\eta}\frac{x - m_\xi}{\sigma_\xi}\Big)^2\Big] du.$$

Положим

$$t = \Big(u - \rho_{\xi\eta}\frac{x - m_\xi}{\sigma_\xi}\Big) \Big/ \sqrt{1 - \rho_{\xi\eta}^2},$$

получим

$$f_\xi(x) = \frac{1}{2\pi\sigma_\xi}\exp\Big[-\frac{(x-m_\xi)^2}{2\sigma_\xi}\Big]\int\limits_{-\infty}^{\infty}\exp\Big[-\frac{t^2}{2}\Big]dt = \frac{1}{\sqrt{2\pi}\sigma_\xi}\exp\Big[-\frac{(x-m_\xi)^2}{2\sigma_\xi}\Big],$$

что и требовалось доказать.

Найдём условные плотности вероятностей. По определению,

$$f_\xi(x/y) = \frac{f_{\xi,\eta}(x,y)}{f_\eta(y)} = \Big\{\frac{1}{\sqrt{2\pi}\sigma_\eta}\exp\Big[-\frac{(y-m_\eta)^2}{2\sigma_\eta}\Big]\Big\}^{-1}\frac{1}{2\pi\sigma_\xi\sigma_\eta\sqrt{1-\rho_{\xi\eta}^2}}\times$$

$$\times\exp\Big\{-\frac{1}{2(1-\rho_{\xi\eta}^2)}\Big[\Big(\frac{x-m_\xi}{\sigma_\xi}\Big)^2 - 2\rho_{\xi\eta}\frac{x-m_\xi}{\sigma_\xi}\frac{y-m_\eta}{\sigma_\eta} + \Big(\frac{y-m_\eta}{\sigma_\eta}\Big)^2\Big]\Big\} =$$

$$= \frac{1}{\sigma_\xi\sqrt{2\pi}\sqrt{1-\rho_{\xi\eta}^2}}\exp\Big\{-\frac{1}{2(1-\rho_{\xi\eta}^2)}\Big[\Big(\frac{x-m_\xi}{\sigma_\xi}\Big)^2 - 2\rho_{\xi\eta}\frac{x-m_\xi}{\sigma_\xi}\frac{y-m_\eta}{\sigma_\eta} + \rho_{\xi\eta}^2\Big(\frac{y-m_\eta}{\sigma_\eta}\Big)^2\Big]\Big\} =$$

$$= \frac{1}{\sigma_\xi\sqrt{2\pi}\sqrt{1-\rho_{\xi\eta}^2}}\exp\Big\{-\frac{1}{2\sigma_\xi^2(1-\rho_{\xi\eta}^2)}\Big[x - \Big(m_\xi + \rho_{\xi\eta}\frac{\sigma_\xi}{\sigma_\eta}(y-m_\eta)\Big)\Big]^2\Big\}.$$

Таким образом, если случайная величина η приняла значение y, то случайная величина ξ имеет нормальное распределение с параметрами

$$M[\xi/y] = m_\xi + \rho_{\xi\eta}\frac{\sigma_\xi}{\sigma_\eta}(y-m_\eta), \qquad D[\xi/y] = \sigma_\xi^2(1-\rho_{\xi\eta}^2).$$

Свойство 3. Пусть $\vec\xi$ — многомерная случайная величина, распределённая нормально с параметрами $\vec\mu$ и A ($\vec\xi \in N(\vec\mu, A)$). Тогда многомерная случайная величина $\vec\eta$, полученная линейным преобразованием $\vec\eta = C\vec\xi + \vec d$, также является нормально распределённой, причем $\vec\eta \in N(C\vec\mu + \vec d, CAC^\intercal)$.

Доказательство. Пусть $\vec\xi \in N(\vec\mu, A)$, тогда для многомерной случайной величины $\vec\eta = C\vec\xi + \vec d$ по формуле (11.14) получим

$$f_\eta(\vec x) = \frac{(2\pi)^{-n/2}}{|\det C|\sqrt{\det A}}\exp\Big[-\frac{1}{2}[C^{-1}(\vec x - \vec d) - \vec\mu]^\intercal A^{-1}[C^{-1}(\vec x - \vec d) - \vec\mu]\Big].$$

Представим коэффициент при $\vec\mu$ как $C^{-1}C$ и вынесем C^{-1} как общий множитель за скобки, получим

$$f_\eta(\vec x) = \frac{(2\pi)^{-n/2}}{|\det C|\sqrt{\det A}}\exp\Big[-\frac{1}{2}[C^{-1}(\vec x - \vec d) - C^{-1}C\vec\mu]^\intercal A^{-1}[C^{-1}(\vec x - \vec d) - C^{-1}C\vec\mu]\Big] =$$

$$= \frac{(2\pi)^{-n/2}}{|\det C|\sqrt{\det A}}\exp\Big\{-\frac{1}{2}\{C^{-1}[\vec x - (\vec d + C\vec\mu)]\}^\intercal A^{-1}\{C^{-1}[\vec x - (\vec d + C\vec\mu)]\}\Big] =$$

$$= \frac{(2\pi)^{-n/2}}{|\det C|\sqrt{\det A}}\exp\Big\{-\frac{1}{2}[\vec x - (\vec d + C\vec\mu)]^\intercal (C^{-1})^\intercal A^{-1}C^{-1}[\vec x - (\vec d + C\vec\mu)]\Big] =$$

$$= \frac{(2\pi)^{-n/2}}{|\det C|\sqrt{\det A}}\exp\Big\{-\frac{1}{2}[\vec x - (\vec d + C\vec\mu)]^\intercal (CAC^\intercal)^{-1}[\vec x - (\vec d + C\vec\mu)]\Big],$$

что есть плотность многомерной нормальной случайной величины с параметрами $C\vec\mu + \vec d$ и $CAC^\intercal$.

Заметим, что если $\vec{\xi} \in N(\vec{0}, E)$, т.е. компоненты вектора $\vec{\xi}$ независимы и имеют стандартное нормальное распределение, то $\vec{\eta} = C\vec{\xi} + \vec{\mu}$ будет иметь нормальное распределение с параметрами $M(\vec{\eta}) = \vec{\mu}$ и $A(\vec{\eta}) = CC^{\intercal}$. Таким образом, любую нормальную случайную величину $\vec{\eta}$ с параметрами $\vec{\mu}$ и A можно получить путем линейного преобразования $\vec{\eta} = C\vec{\xi} + \vec{\mu}$ величины $\vec{\xi} \in N(\vec{0}, E)$, где матрицы C и A связаны соотношением $A = CC^{\intercal}$.

21. Распределения, связанные с нормальным

21.1. Логнормальное распределение

♦ Непрерывная случайная величина ξ имеет *логнормальное распределение*, если $\eta = \ln \xi$ подчинена нормальному закону $\eta \in N(m, \sigma^2)$.

Из определения логнормального распределения следует, что если случайная величина η распределена нормально, то $\xi = e^{\eta}$ распределена логнормально. Таким образом, логнормальному распределению подчиняется распределение случайной величины, представленной в виде произведения слабо зависимых случайных величин, сравнимых по порядку их влияния.

Из соотношения (11.8) при $\varphi(x) = e^x$ для плотности распределения вероятностей получим (см. рис. 70)

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(\ln x - \mu)^2/2\sigma^2}, \quad \mu = \ln m, \quad 0 < x < \infty. \tag{21.1}$$

Логнормальное распределение часто встречается в экономике, например в формуле Блэка–Скоулза цены европейских опционов (см. также [20]).

Найдём числовые характеристики случайной величины, распределённой по логнормальному закону. По определению,

$$M(\xi) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_0^{\infty} e^{-(\ln x - \mu)^2/2\sigma^2} dx.$$

Сделаем замену переменных: $t = (\ln x - \mu)/\sigma$. Тогда $x = e^{\mu} e^{\sigma t}$, $dx = \sigma e^{\mu} e^{\sigma t} dt$; $t = -\infty$ при $x = 0$ и $t = \infty$ при $x = \infty$. Получим

$$M(\xi) = \frac{e^{\mu}}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{\sigma t} e^{-t^2/2} dt = \frac{e^{\mu} e^{\sigma^2/2}}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-(t-\sigma)^2/2} d(t - \sigma) = e^{\mu} e^{\sigma^2/2} = m e^{\sigma^2/2}.$$

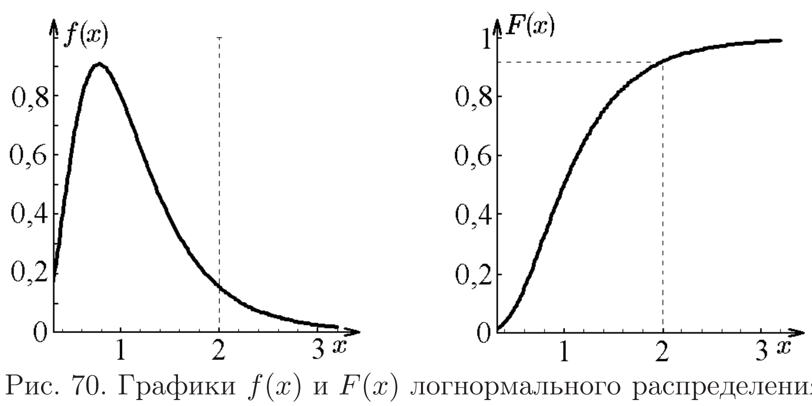


Рис. 70. Графики $f(x)$ и $F(x)$ логнормального распределения

Вычислим далее $M(\xi^2)$:

$$M(\xi^2) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^\infty x e^{-(\ln x-\mu)^2/2\sigma^2} dx.$$

Сделав ту же замену переменных: $t = (\ln x - \mu)/\sigma$, получим

$$M(\xi^2) = \frac{e^{2\mu}}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{2\sigma t} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} e^{2\mu} e^{2\sigma^2} \int_{-\infty}^\infty e^{-(t-2\sigma)^2/2} d(t-2\sigma) = e^{2\mu} e^{2\sigma^2}.$$

И, наконец, найдём дисперсию:

$$D(\xi) = M(\xi^2) - M^2(\xi) = e^{2\mu}e^{2\sigma^2} - e^{2\mu}e^{\sigma^2} = e^{2\mu}e^{\sigma^2}(e^{\sigma^2}-1) = m^2 e^{\sigma^2}(e^{\sigma^2}-1).$$

Медиана логнормального распределения совпадает с математическим ожиданием: $x_{\text{med}} = m$, мода этого распределения определяется выражением $x_{\text{mod}} = me^{-\sigma^2}$, а для коэффициента асимметрии и эксцесса получим

$$A = (e^{\sigma^2}+2)\sqrt{(e^{\sigma^2}-1)},$$
$$E = (e^{\sigma^2}-1)(e^{3\sigma^2}+3e^{2\sigma^2}+6e^{\sigma^2}+6).$$

Логнормальное распределение используется для описания распределения доходов, банковских вкладов, месячной зарплаты и т.д.

21.2. Гамма-распределение

♦ Случайная величина ξ называется *гамма-распределённой* с параметрами $\alpha > 0$, $\lambda > 0$, если распределение её плотности вероятности имеет вид

$$f(x) = \begin{cases} \dfrac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x} & \text{при } x \geqslant 0, \\ 0 & \text{при } x < 0, \end{cases} \tag{21.2}$$

где $\Gamma(\alpha)$ — гамма-функция (см. [2]), и обозначается $\xi \in \Gamma_{\lambda,\alpha}$ (рис. 71).

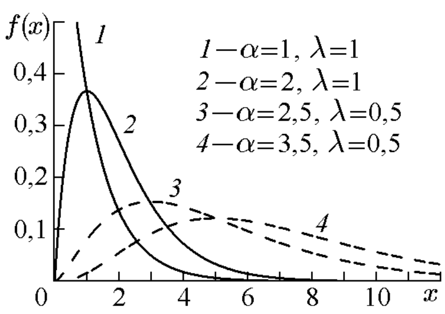


Рис. 71.

♦ При $\alpha = 1$ случайная величина ξ называется *распределённой по показательному закону*, а при $\alpha = k$, $k = \overline{2,\infty}$, — *распределённой по закону Эрланга*.

Найдём числовые характеристики гамма-распределения. По определению, математическое ожидание запишется как

$$M(\xi) = \frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda^\alpha x^\alpha e^{-\lambda x} dx.$$

Сделав замену переменных $\lambda x = t$, $dx = dt/\lambda$, получим

$$M(\xi) = \frac{1}{\lambda\Gamma(\alpha)} \int_0^\infty t^\alpha e^{-t} dt = \frac{\Gamma(\alpha+1)}{\lambda\Gamma(\alpha)} = \frac{\alpha}{\lambda}.$$

Следовательно,

$$M(\xi) = \frac{\alpha}{\lambda}. \tag{21.3}$$

Вычислим дисперсию. По определению,

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

но

$$M(\xi^2) = \frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda^\alpha x^{\alpha+1} e^{-\lambda x} dx.$$

Сделав замену переменных $\lambda x = t$, $dx = dt/\lambda$, запишем

$$M(\xi^2) = \frac{1}{\lambda^2 \Gamma(\alpha)} \int_0^\infty t^{\alpha+1} e^{-t} dt = \frac{\Gamma(\alpha+2)}{\lambda^2 \Gamma(\alpha)} = \frac{\alpha(\alpha+1)}{\lambda^2}.$$

Следовательно, для дисперсии получим

$$D(\xi) = \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2}, \tag{21.4}$$

а для среднеквадратичного отклонения найдём

$$\sigma = \frac{\sqrt{\alpha}}{\lambda}. \tag{21.5}$$

Заметим также, что $A(\xi) = 2/\sqrt{\lambda}$, $E(\xi) = 6/\alpha$.

Вычислим теперь производящую функцию центрированных моментов:

$$\mu(t) = \frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda^\alpha x^{\alpha-1} e^{-\lambda x} e^{t(x-\alpha/\lambda)} dx.$$

Сделаем замену переменных $(\lambda - t)x = u$ и $dx = du/(\lambda - t)$ и получим

$$\mu(t) = \frac{1}{\Gamma(\alpha)} \Big(\frac{\lambda}{\lambda - t}\Big)^\alpha \int_0^\infty u^{\alpha-1} e^{-u} du\, e^{-\alpha t/\lambda}$$

или

$$\mu(t) = \Big(1 - \frac{t}{\lambda}\Big)^{-\alpha} e^{-\alpha t/\lambda}. \tag{21.6}$$

Аналогично найдём характеристическую функцию гамма-распределения:

$$\gamma(t) = \Big(1 - \frac{it}{\lambda}\Big)^{-\alpha}. \tag{21.7}$$

Свойства Г-распределения

Свойство 1. Распределение $\Gamma_{\lambda,1}$ есть показательное распределение с параметром λ.

Действительно, если $\xi \in \Gamma_{\lambda,1}$, то

$$f_\xi(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0; \\ 0, & x \leqslant 0 \end{cases}$$

есть плотность распределения случайной величины, распределённой по показательному закону с параметром λ.

Свойство 2. Если $\xi \in N(0,1)$, то $\eta = \xi^2 \in \Gamma_{1/2,1/2}$.

Доказательство. Найдём функцию распределения случайной величины $\eta = \xi^2$. Для положительных x ($x \geqslant 0$) запишем

$$F_\eta(x) = P(\xi^2 < x) = P(-\sqrt{x} < \xi < \sqrt{x}) = 2\int\limits_0^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

Сделаем замену переменных $z = t^2$. Тогда $t = \sqrt{z}$, $dt = dz/2\sqrt{z}$ и

$$F_\eta(x) = \int\limits_0^x \frac{1}{\sqrt{2\pi z}} e^{-z/2} dz = \int\limits_0^x \frac{(1/2)^{1/2}}{\Gamma(1/2)} z^{1/2-1} e^{-z/2} dz = F_{\Gamma_{1/2,1/2}}(x).$$

Свойство 3. Гамма-распределение устойчиво по суммированию, т.е. если ξ и η независимы и $\xi \in \Gamma_{\lambda,\alpha_1}$, $\eta \in \Gamma_{\lambda,\alpha_2}$, то

$$\zeta = \xi + \eta \in \Gamma_{\lambda,\alpha_1+\alpha_2}.$$

Согласно свойствам преобразования Фурье (теорема умножения), для характеристической функции запишем

$$\gamma_\zeta(t) = \gamma_\xi(t)\gamma_\eta(t) = \Big(1 - \frac{it}{\lambda}\Big)^{-\alpha_1}\Big(1 - \frac{it}{\lambda}\Big)^{-\alpha_2} = \Big(1 - \frac{it}{\lambda}\Big)^{-\alpha_1-\alpha_2},$$

а это и есть характеристическая функция распределения $\Gamma_{\lambda,\alpha_1+\alpha_2}$ (21.7).

21.3. Бета-распределение

♦ Случайная величина ξ называется *бета-распределённой*, если её плотность распределения вероятностей имеет вид (рис. 72)

$$f(x) = \begin{cases} \dfrac{1}{\mathrm{B}(a,b)} x^{a-1}(1-x)^{b-1} & x \in [0,1], \\ 0 & x \notin [0,1], \end{cases} \tag{21.8}$$

где $a > 0$, $b > 0$, $\mathrm{B}(a,b)$ — бета-функция (см. [2]).

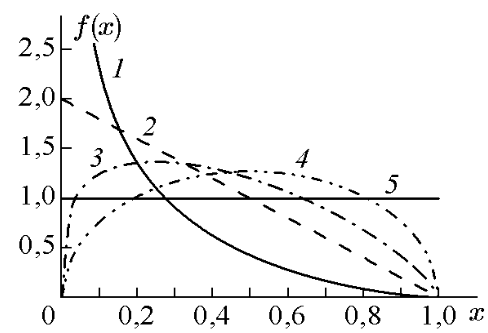


Рис. 72.

Очевидно, что

$$F_{\mathrm{B}}(x) = \begin{cases} 0, x \leqslant 0; \\ \dfrac{1}{\mathrm{B}(a,b)} \displaystyle\int\limits_0^x u^{a-1}(1-u)^{b-1} du, & 0 \leqslant x \leqslant 1; \\ 1, & x \geqslant 1. \end{cases}$$

Найдём числовые характеристики бета-распределения:

$$M(\xi) = \frac{1}{\mathrm{B}(a,b)} = \int\limits_0^1 x^a(1-x)^{b-1}dx = \frac{\mathrm{B}(a+1,b)}{\mathrm{B}(a,b)} = \frac{a}{a+b}$$

и

$$M(\xi^2) = \frac{1}{\mathrm{B}(a,b)} = \int\limits_0^1 x^{a+1}(1-x)^{b-1}dx = \frac{\mathrm{B}(a+2,b)}{\mathrm{B}(a,b)} = \frac{(a+1)a}{(a+b+1)(a+b)}.$$

Следовательно,

$$\begin{gathered} D(\xi) = M(\xi^2) - [M(\xi)]^2 = \frac{a}{a+b}\Big(\frac{a+1}{a+b+1} - \frac{a}{a+b}\Big) = \\ = \frac{a}{a+b}\frac{(a+1)(a+b) - a(a+b+1)}{(a+b+1)(a+b)} = \\ = \frac{a}{(a+b)^2}\frac{a(a+b)+(a+b)-a(a+b)-a}{a+b+1} = \frac{ab}{(a+b)^2(a+b+1)}. \end{gathered}$$

Аналогично

$$A(\xi) = \frac{2(-a+b)\sqrt{1+a+b}}{\sqrt{a}\sqrt{b}(2+a+b)};$$

$$E(\xi) = -3 + \frac{3(1+a+b)[(-6+a+b)+2(a+b)^2]}{ab(2+a+b)(3+a+b)}.$$

♦ Функция

$$\mathrm{B}_{ab}(x) = \int\limits_0^x \frac{1}{B(a,b)}x^{a-1}(x-1)^{b-1}dx \tag{21.9}$$

называется *неполной бета-функцией.*

21.4. χ-Распределение. χ^2-Распределение. Распределения Стьюдента и Фишера

Пусть ξ_j, $j=\overline{0,k}$, — независимые случайные величины, распределённые по нормальному закону с $\sigma=1$, $m=0$ ($\xi_j \in N(0,1)$).

♦ Случайная величина η

$$\eta = \sum_{j=1}^k \xi_j^2 \tag{21.10}$$

называется χ^2-*распределённой* с k степенями свободы.

♦ Случайная величина ζ

$$\zeta = \sqrt{\sum_{j=1}^k \xi_j^2} \tag{21.11}$$

называется χ-*распределённой* с k степенями свободы.

♦ χ-распределение при $k=3$ называется *распределением Максвелла.*

♦ Случайная величина τ

$$\tau = \frac{\xi_0}{\sqrt{\dfrac{1}{k}\sum\limits_{j=1}^{k}\xi_j^2}} \tag{21.12}$$

называется *распределённой по закону Стьюдента с k степенями свободы* (t-распределение Стьюдента) и обозначается $\tau \in S_k$. (Здесь ξ_0 — случайная величина, подчиняющаяся нормальному распределению с $m = 0$, $\sigma = 1$ и не зависящая от величин ξ_i.)

Теорема 21.1. *Плотность вероятности χ-распределения имеет вид*

$$f_\chi(x) = \begin{cases} \dfrac{x^{k-1}}{2^{k/2-1}\Gamma(k/2)}e^{-x^2/2} & x \geqslant 0; \\ 0 & x < 0. \end{cases} \tag{21.13}$$

Доказательство. Вероятность события $x < \zeta < x + dx$ можно получить из k-мерной плотности распределения нормального закона

$$f(\vec{x}) = \frac{1}{(2\pi)^{k/2}\sigma^k}e^{-\vec{x}^{\,2}/2\sigma^2}, \quad \vec{x} = (x_1, \ldots, x_k),$$

интегрируя её по k-мерному сферическому слою радиуса x и толщины dx. В результате, поскольку $(k-1)$-мерный объём $(k-1)$-мерной сферы x пропорционален x^{k-1}, получим

$$f_\chi(x) = C_k x^{k-1}e^{-x^2/2}.$$

Из условия нормировки

$$\int\limits_0^\infty f(x)dx = 1$$

найдём постоянную C_k:

$$C_k = \int\limits_0^\infty x^{k-1}e^{-x^2/2}dx = 2^{k/2-1}\Gamma(k/2)C_k = 1.$$

Таким образом, теорема доказана.

Теорема 21.2. *χ^2-Распределение является частным случаем гамма-распределения с $\alpha = k/2$, $\lambda = 1/2$, т.е. функция*

$$f_{\chi^2}(x) = \begin{cases} \dfrac{x^{k/2-1}}{2^{k/2}\Gamma(k/2)}e^{-x^2/2} & x \geqslant 0; \\ 0 & x < 0 \end{cases} \tag{21.14}$$

является плотностью вероятности для χ^2-распределения.

Доказательство. Используя связь (11.11) между плотностями

$$f_{\chi^2}(x) = \frac{1}{2\sqrt{x}}f_\chi(\sqrt{x}),$$

получим утверждение теоремы. Здесь мы воспользовались тем, что если положительная случайная величина ξ имеет плотность распределения $f(x)$, то случайная величина $\eta = \varphi(\xi)$ имеет плотность распределения $\Phi(x) = \varphi'(x)f\big(\varphi(x)\big)$. Таким образом, теорема доказана.

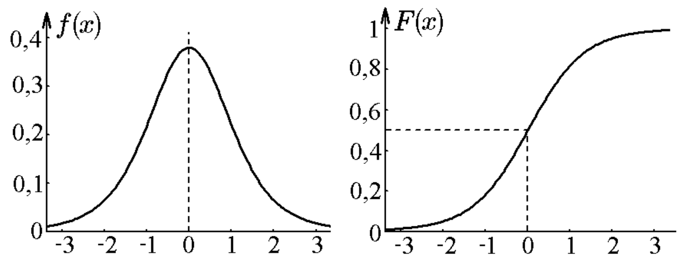


Рис. 73. Графики $f_S(x)$ и $F_S(x)$ t-распределения при $\nu = \omega = 10$

Теорема 21.3. *Плотность распределения случайной величины τ, распределённой по закону Стьюдента, имеет вид* (*см. рис. 73*)

$$f_{S_k}(x) = \frac{\Gamma((k+1)/2)}{\sqrt{k\pi}\,\Gamma(k/2)}\left(1 + \frac{x^2}{k}\right)^{-(k+1)/2}, \quad -\infty < x < \infty. \tag{21.15}$$

Доказательство аналогично доказательству предыдущей теоремы.

Числовые характеристики t-распределения:

$$M(\xi) = x_{\text{med}} = x_{\text{mod}} = 0, \quad D(\xi) = \frac{k}{k-2}, \quad A = 0, \quad E = \frac{6}{k-4}.$$

Можно показать, что при $k \to \infty$

$$\frac{\chi_k^2}{k} = \frac{1}{k}\sum_{i=1}^{k}\xi_i^2 \xrightarrow{p} M(\xi^2) = M(\chi_1^2) = 1,$$

поэтому $S_k \to N(0,1)$ при $k \to \infty$.

Таким образом, при больших степенях свободы ($k > 30$) t-распределение Стьюдента практически совпадает с нормальным распределением $N(0;\,1)$.

На рис. 74 приведены для сравнения графики плотностей распределения:
- нормального с $m = 0$, $\sigma = 1$ (штриховая линия) — $f(x)$ для $N(0,1)$;
- Стьюдента с $m = k$ (тонкая сплошная линия) — $f_{\text{St}}(x)$ для $\text{St}(5)$;
- Стьюдента с $k = 1$ (жирная сплошная линия) — $f_{\text{St}}(x)$ для $\text{St}(1)$.

Эти графики наглядно иллюстрируют, что с увеличением степени свободы k распределение Стьюдента стремится к нормальному:

$$\lim_{k\to\infty} f_S(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

$\Diamond$ В частном случае при $k = 1$ распределение Стьюдента переходит в распределение Коши.

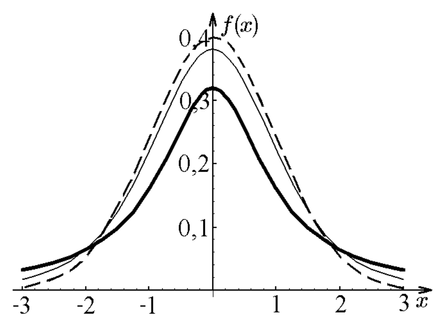


Рис. 74.

Плотность вероятности для распределения Коши имеет вид

$$f(x) = \frac{1}{\pi(1+x^2)}. \tag{21.16}$$

Найдём функцию распределения:

$$F(x) = \frac{1}{\pi}\int\limits_{-\infty}^{x} = \frac{dt}{1+t^2} = \frac{1}{\pi}\operatorname{arctg} x + \frac{1}{2}. \tag{21.17}$$

Характеристическая функция распределения Коши определится соотношением

$$\gamma(t) = \frac{1}{\pi}\int\limits_{-\infty}^{\infty}\frac{e^{-itx}}{1+x^2}.$$

Вычислив интеграл (см., например, [2]), получим

$$\gamma(t) = e^{-|t|}. \tag{21.18}$$

Пример 21.1. Случайные величины ξ и η независимы и имеют стандартное нормальное распределение. Найти закон распределения случайной величины

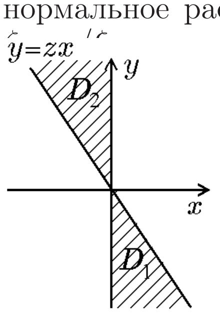


Рис. 75.

Решение. Так как случайные величины ξ и η имеют стандартное нормальное распределение, то их плотность распределения имеет вид

$$f_\xi(x) = f_\eta(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \quad -\infty < x < +\infty.$$

Так как эти величины независимы, то плотность совместного распределения этих величин равна произведению плотностей:

$$f_{\xi,\eta}(x,y) = f_\eta(x)f_\xi(y) = \frac{1}{2\pi}e^{-(x^2/2+y^2/2)}, \quad -\infty < x < \infty, \quad -\infty < y < \infty.$$

Найдём функцию распределения величины ζ. По определению,

$$F_\zeta(z) = P(\mu < z) = P(\eta/\xi < z) = \iint\limits_{D_1: y/x<z} f_{\xi,\eta}(x,y)dx\,dy =$$

$$= \iint\limits_{D_2: y<zx, x>0} f_{\xi,\eta}(x,y)dx\,dy + \iint\limits_{D_2: y>zx, x<0} f_{\xi,\eta}(x,y)dx\,dy.$$

Пусть $z < 0$ (см. рис. 75). Перейдём к полярным координатам:

$$F_\zeta(z) = \frac{1}{2\pi}\biggl(\iint\limits_{D_1: y<zx, x>0} e^{-(x^2+y^2)/2}dx\,dy + \iint\limits_{D_2: y>zx, x<0} e^{-(x^2+y^2)/2}dx\,dy\biggr) =$$

$$= \frac{1}{2\pi}\biggl(\int\limits_{-\pi/2}^{\operatorname{arctg} z} d\varphi\int\limits_0^\infty e^{-\rho^2/2}\rho\,d\rho + \int\limits_{\pi/2}^{\operatorname{arctg} z+\pi} d\varphi\int\limits_0^\infty e^{-\rho^2/2}\rho\,d\rho\biggr) =$$

$$= \frac{1}{2\pi}\Bigl(\operatorname{arctg} z + \frac{\pi}{2} + \operatorname{arctg} z + \pi - \frac{\pi}{2}\Bigr) = \frac{1}{\pi}\Bigl(\operatorname{arctg} z + \frac{\pi}{2}\Bigr).$$

Легко показать, что в случае $z > 0$ получим такой же результат. Таким образом, функция распределения определяется соотношением

$$F_\zeta(z) = \frac{1}{\pi}\int\limits_{-\infty}^{z}\frac{dt}{1+t^2} = \frac{1}{\pi}\Bigl(\operatorname{arctg} z + \frac{\pi}{2}\Bigr),$$

т.е. ζ имеет распределение Коши.

Найдём числовые характеристики рассматриваемых распределений.

1. Поскольку распределение χ^2 является частным случаем гамма-распределения с $\lambda = 1/2$, $\alpha = k/2$, то

$$M(\xi) = \frac{\alpha}{\lambda} = k, \quad D(\xi) = \frac{\alpha}{\lambda^2} = 2k, \quad \sigma = \sqrt{2\alpha}, \quad A = \sqrt{\frac{8}{k}}, \quad E = \frac{12}{k}.$$

2. Для χ-распределения получим

$$M(\xi^l) = \int\limits_0^\infty \frac{x^{k+l-1}}{2^{k/2-1}\Gamma(k/2)} e^{-x^2/2} dx.$$

Сделав замену переменных

$$\frac{x^2}{2} = u, \quad x = \sqrt{2u}, \quad dx = \frac{du}{\sqrt{2u}},$$

получим

$$M(\xi^l) = \frac{2^{l/2}}{\Gamma(k/2)} \int\limits_0^\infty u^{(k+l)/2-1} e^{-u^2} du = \frac{\Gamma((k+l)/2)}{\Gamma(k/2)} 2^{l/2}.$$

В частности,

$$M(\xi) = \sqrt{2}\frac{\Gamma((k+1)/2)}{\Gamma(k/2)}, \quad M(\xi^2) = 2\frac{\Gamma((k+2)/2)}{\Gamma(k/2)}$$

и

$$D(\xi) = \frac{2}{\Gamma(k/2)}\left[\Gamma\Big(\frac{k+2}{2}\Big) - \frac{\Gamma^2((k+1)/2)}{\Gamma(k/2)}\right].$$

3. Моменты $M(\xi^l)$ для распределения Стьюдента существуют только для $l < k$, при этом начальные моменты равны нулю, а для центральных моментов получим

$$\mu[\xi^{2l}] = k^l \frac{\Gamma(l/2 + 1/2)\Gamma(k/2 - l)}{\sqrt{\pi}\Gamma(k/2)}, \quad 2 \leqslant 2l < k.$$

Теорема 21.4. *Функция распределения $F_{S_k}(x)$ выражается через неполную бета-функцию:*

$$F_{S_k}(x) = 1 - \frac{1}{2\mathrm{B}(k/2, 1/2)} \int\limits_0^{k/(k+x^2)} u^{k/2-1}(1-u)^{k/2-1} du.$$

Доказательство. Действительно,

$$F_{S_k}(x) = \frac{\Gamma((k+1)/2)}{\sqrt{k}\sqrt{\pi}\Gamma(k/2)} \int\limits_{-\infty}^{x} \left(1 + \frac{x^2}{k}\right)^{-(k+1)/2} dx.$$

Сделаем замену переменных

$$u = \frac{k}{k+x^2}, \qquad x = \sqrt{\frac{k}{u} - k} = \sqrt{k}\sqrt{\frac{1-u}{u}},$$

тогда

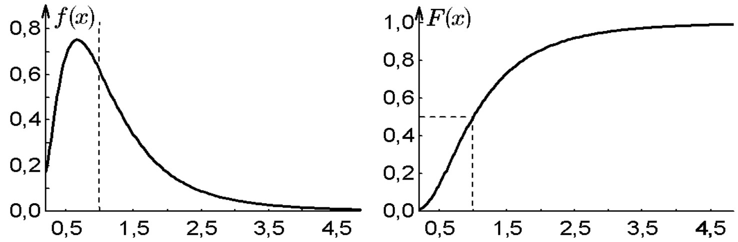


Рис. 76. Графики $f(x)$ и $F(x)$ закона F-распределения Фишера

$$dx = \frac{\sqrt{k}}{2}\frac{1}{\sqrt{1/u-1}}\frac{-du}{u^2} = -\frac{\sqrt{k}}{2}\sqrt{\frac{u}{1-u}}\frac{du}{u^2},$$

$x_1 = -\infty$, $u_1 = 0$, $x_2 = x$, $u_2 = k/(k+x^2)$. Таким образом,

$$F_{S_k}(x) = 1 - \frac{1}{\sqrt{k}\mathrm{B}(k/2,1/2)}\int\limits_0^{k/(k+x^2)} u^{(k+1)/2}(1-u)^{k/2-1}u^{k/2-2}\Big(-\frac{\sqrt{u}}{2}\Big)du =$$

$$= 1 - \frac{1}{2\mathrm{B}(k/2,1/2)}\int\limits_0^{k/(k+x^2)} u^{k/2-1}(1-u)^{k/2-1}du,$$

что и требовалось доказать.

◊ Распределение Стьюдента ($k = 1, 3, 5$) используется в экономике для описания флуктуаций стоимости ценных бумаг. Флуктуации, описываемые распределением Стьюдента, в силу степенного характера функции плотности распределения многократно превосходят флуктуации, описываемые нормальным распределением.

♦ Непрерывная случайная величина ξ имеет *F-распределение Фишера*, если она представима в виде отношения двух независимых случайных величин, распределённых по закону хи-квадрат со степенями свободы ν и ω, и плотность распределения вероятностей имеет вид (см. рис. 76)

$$f_F(x) = \frac{\Gamma((\nu+\omega)/2)}{\Gamma(\nu/2)\Gamma(\omega/2)}\Big(\frac{\nu}{\omega}\Big)^{\nu/2}x^{\nu/2-1}\Big(1+\frac{\nu}{\omega}x\Big)^{-(\nu+\omega)/2}, \quad 0 < x < \infty. \qquad (21.19)$$

Основные числовые характеристики F-распределения:

$$M(\xi) = \frac{\omega}{\omega-2}, \quad D(\xi) = \frac{2\omega^2(\nu+\omega-2)}{\nu(\omega-2)^2(\omega-4)},$$

$$A(\xi) = \frac{2\sqrt{2}\sqrt{-4+\omega}(-2+2\nu+\omega)}{\sqrt{\nu}(-6+\omega)\sqrt{-2+\nu+\omega}},$$

$$E(\xi) = \frac{12[(-4+\omega)(-2+\omega)^2+\nu(-2+\nu+\omega)(-22+5\omega)]}{\nu(-8+\omega)(-6+\omega)(-2+\nu+\omega)}.$$

Графики плотности $f_F(x)$ для F-распределения Фишера для различных параметров ν и ω представлены на рис. 77.

◊ Если случайная величина ξ распределена по закону Фишера с параметром λ ($\xi \in F_\lambda$) и $a > 0$, то случайная величина $\eta = a\xi$ также распределена по закону Фишера с параметром λ/a ($\eta = a\xi \in F_{\lambda/a}$).

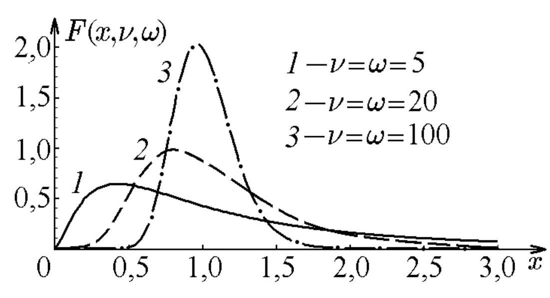


Рис. 77.

♦ Распределение Фишера–Снедекора при целых $\nu = m$ и $\omega = n$ называется *распределением F-отношения*

$$\bar{F} = \frac{\frac{1}{m}\chi^2(m)}{\frac{1}{n}\chi^2(n)},$$

где $\chi^2(m)$ и $\chi^2(n)$ — случайные величины, имеющие χ^2-распределение с m и n степенями свободы соответственно.

Эллипсоид рассеяния

Пусть случайная величина $\vec{\xi} = \{\xi_1, \xi_2, \ldots, \xi_n\}$ имеет нормальное распределение с вектором средних $\vec{\mu} = \mu_1, \mu_2, \ldots, \mu_n$ и матрицей ковариаций A.

Рассмотрим в пространстве переменных $\xi_1, \xi_2, \ldots, \xi_n$ область, определяемую неравенством

$$(\vec{\xi} - \vec{\mu})^{\intercal} A^{-1} (\vec{\xi} - \vec{\mu}) < \delta,$$

где δ — некоторое положительное число (вектор рассматриваем как матрицу-столбец). Левая часть неравенства есть положительно определённая квадратичная форма (при условии, что хотя бы один из коэффициентов корреляции $|\rho_{ij}| < 1$). Следовательно, данное неравенство определяет в пространстве $\mathbb{R}^n$ область, ограниченную эллипсоидом. Центр эллипсоида находится в точке $(\mu_1, \mu_2, \ldots, \mu_n)$, оси параллельны собственным векторам матрицы A, а полуоси пропорциональны $\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}$, где $\lambda_1, \lambda_2, \ldots, \lambda_n$ — собственные значения матрицы A.

С другой стороны, $\vec{\eta} = (\vec{\xi} - \vec{\mu})^{\intercal} A^{-1} (\vec{\xi} - \vec{\mu})$ есть некоторая одномерная случайная величина, являющаяся функцией величин $\xi_1, \xi_2, \ldots, \xi_n$. Найдём закон распределения $\vec{\eta}$. Как известно (см. свойство 3), многомерную нормальную случайную величину $\vec{\eta}$ с параметрами $\vec{\mu}$ и A можно получить путем линейного преобразования вектора $\vec{\xi}$, компоненты которого независимы и имеют стандартное нормальное распределение

$$\vec{\xi} = B\vec{\zeta} + \vec{\mu},$$

где матрица B связана с матрицей ковариаций A соотношением $BB^{\intercal} = A$. Но тогда

$$\eta = (B\vec{\zeta})^{\intercal} (BB^{\intercal})^{-1} (B\vec{\zeta}) = \vec{\zeta}^{\intercal}\vec{\zeta} = \sum_{i=1}^{n} \zeta_i^2,$$

т.е. η есть сумма квадратов независимых стандартных нормальных случайных величин. Следовательно, η имеет распределение χ^2 с n степенями свободы. Тогда по таблицам распределения χ^2 для заданной вероятности β можно определить значение δ, для которого $P(\eta < \delta) = \beta$, т.е. δ будет квантилью χ^2-распределения уровня β. Соответственно, с вероятностью β многомерная случайная величина $\vec{\xi}$ будет принадлежать области, ограниченной эллипсоидом

$$(\vec{x} - \vec{\mu})^{\intercal} A^{-1} (\vec{x} - \vec{\mu}) < \delta, \qquad \vec{x} = \{x_1, x_2, \ldots, x_n\}.$$

Такой эллипсоид называется *эллипсоидом рассеяния*.

Пример 21.2. Для системы $\{\xi, \eta\}$, имеющей нормальное распределение с параметрами

$$\vec{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

построить эллипс рассеяния, соответствующий вероятности 0,9.

Решение. Пусть $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, тогда

$$(\vec{x} - \vec{\mu})^{\intercal} A^{-1} (\vec{x} - \vec{\mu}) = 2x^2 - 2xy + y^2.$$

По таблицам распределения χ^2 для числа степеней свободы $\nu = 2$ найдём квантиль уровня 0,9: $\tau_{0,9} = 4{,}506$. Соответствующий эллипс рассеяния будет определяться уравнением

$$2x^2 - 2xy + y^2 = 4{,}506.$$

Заметим, что при числе степеней свободы $\nu = 2$ распределение χ^2 совпадает с показательным распределением с параметром $\lambda = 1/2$. Поэтому τ_β можно было бы найти как корень уравнения $F(\tau_\beta) = \beta$, где $F(x)$ — функция распределения показательного закона $F(x) = 1 - e^{-\lambda x}$. Положив $\lambda = \tau_\beta$ и приравняв $F(x)$ к β, получим $1 - e^{-\tau_\beta/2} = \beta$ или $\tau_\beta = -2\ln(1-\beta)$.

22. Распределения Леви, Парето и логистическое распределение

В заключение рассмотрим ряд распределений, широко распространённых в экономических и биологических приложениях.

22.1. Распределение Леви

В экономических моделях часто встречается распределение Леви. Так распределены флуктуации индексов и курсов акций. Главная особенность экономических флуктуаций в том, что в отличие от физических величин, у которых больших флуктуаций мало, поскольку эти флуктуации имеют в основном гауссов характер и в силу этого малы, распределения флуктуаций в экономике имеют так называемые «тяжёлые хвосты», т.е. значительное количество больших флуктуаций.

♦ Случайная величина ξ называется *распределённой по закону Леви*, если её характеристическая функция распределения имеет вид

$$\gamma_L(t) = \exp(-|t|^\alpha), \quad 0 < \alpha \leqslant 2. \tag{22.1}$$

Проведя обратное преобразование Фурье для функции плотности распределения, запишем

$$f_L(x) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{ixt} e^{-|t|^\alpha} dt.$$

С учётом чётности характеристической функции можно записать

$$f_L(x) = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} \cos(xt)\, e^{-|t|^\alpha} dt. \tag{22.2}$$

Соотношение (22.2) определяет так называемое двустороннее распределение Леви, когда $x \in \mathbb{R}$. Пример распределения Леви с $\alpha = 1/2; 1; 3/2; 2$ приведён на рис. 78.

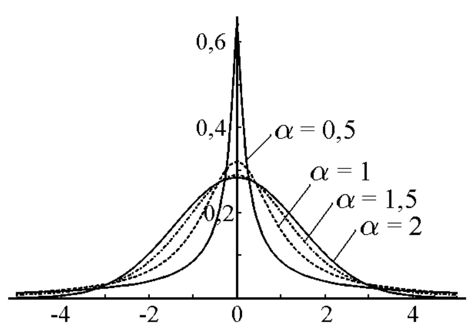


Рис. 78.

Если $x \in [0, \infty[$, получим так называемое одностороннее распределение Леви:

$$f_L = \frac{1}{\pi} \int\limits_0^\infty \cos(xt) e^{-t^\alpha} dt. \tag{22.3}$$

$\Diamond$ При $\alpha = 2$ формула (22.1) совпадает с (19.7), а распределение Леви совпадает с нормальным распределением (19.1); при $\alpha = 1$ соотношение (22.1) совпадает с (19.8), а распределение Леви — с распределением Коши (21.16).

22.2. Распределение Парето

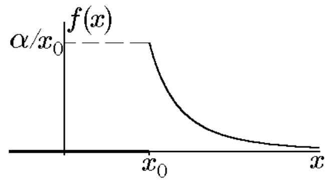


Рис. 79.

$\blacklozenge$ Говорят, что величина ξ имеет *распределение Парето* с параметрами (x_0, α), если (см. рис. 79)

$$f_\xi(x) = \begin{cases} \dfrac{\alpha}{x_0}\Big(\dfrac{x_0}{x}\Big)^{\alpha+1}, & x > x_0; \\ 0, & x \leqslant x_0. \end{cases}$$

$\Diamond$ Нетрудно заметить, что распределение Парето усечением на интервале $]x_0, \infty[$ степенного распределения с параметром α и плотностью вероятности

$$f(x) = \begin{cases} \alpha x^{-(\alpha+1)}, & x > 0; \\ 0, & x \leqslant 0. \end{cases}$$

Непосредственной проверкой можно убедиться, что у распределения Парето существуют моменты k-го порядка ($k < \alpha$):

$$\alpha_k = \frac{\alpha}{\alpha - k} x_0^k, \qquad D\xi = \begin{cases} \dfrac{\alpha x_0^2}{(\alpha-1)(\alpha-2)}, & \alpha > 2; \\ \infty, & \alpha \leqslant 2. \end{cases}$$

22.3. Логистическое распределение

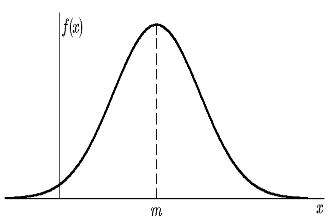


Рис. 80.

$\blacklozenge$ Говорят, что величина ξ имеет *логистическое распределение* с параметрами (m, σ^2), если (см. рис. 80)

$$f_\xi(x) = \frac{\pi \exp\Big[-\dfrac{\pi}{\sqrt{3}}\dfrac{(x-m)}{\sigma}\Big]}{\sigma\sqrt{3}\Big\{1 + \exp\Big[-\dfrac{\pi}{\sqrt{3}}\dfrac{(x-m)}{\sigma}\Big]\Big\}^2}, \quad x \in \mathbb{R}.$$

$\Diamond$ Функция плотности вероятности логистического распределения по форме мало отличается от функция плотности вероятности нормального распределения и наряду с последней широко используется в приложениях.

Логистическое распределение лежит в основе теории популяций (бактерий, человека и т.д.). Было обнаружено, что высота и вес растений и животных подчиняется логистическому закону.

Характеристическая функция логистического распределения имеет вид

$$\gamma_\xi(t) = e^{imt}\Gamma\Big(1 - i\frac{\sigma\sqrt{3}}{\pi}t\Big)\Gamma\Big(1 + i\frac{\sigma\sqrt{3}}{\pi}t\Big).$$

Основные числовые характеристики случайной величины ξ определяются соотношениями

$$M\xi = m, \quad D\xi = \sigma^2, \quad A = 0, \quad E = 1{,}2.$$

ГЛАВА 3

Предельные теоремы. Закон больших чисел

Выше мы уже неоднократно отмечали, что экспериментально замечена устойчивость относительной частоты m/n реальных событий (m — число появлений события A в n опытах).

Относительная частота — это простейший пример среднего случайных величин ξ_k $(k = \overline{1, n})$, каждая из которых есть число появлений события A в k-м опыте

$$\mu^*(A) = \frac{m}{n} = \frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n} \underset{n\to\infty}{\sim} P(A).$$

Связь относительной частоты и вероятности этого события устанавливается теоремой Бернулли. Исторически теорема Бернулли — первая из доказанных предельных теорем (1713 г.). После неё было доказано довольно большое число теорем такого рода, выясняющих поведение среднего различных случайных величин при $n \to \infty$.

Суть предельных теорем состоит в том, что среднее при неограниченном увеличении числа опытов теряет характер случайного, и поэтому его поведение можно предсказать.

Это положение, доказанное в ряде предельных теорем, называется *законом больших чисел*. Накопленный опыт отмечает хорошее согласие предельных теорем с реальной действительностью.

На предельных теоремах фактически основана вся математическая статистика, поскольку они позволяют приближённо вычислить характеристики случайной величины на основе экспериментальных данных.

Из всего множества предельных теорем рассмотрим теоремы Чебышева, Бернулли, а также центральную предельную теорему Ляпунова и её частный случай — теорему Муавра–Лапласа.

Здесь нами будут рассматриваться только одномерные случайные величины, потому с целью упрощения будем называть их просто случайными величинами.

23. Сходимость случайных последовательностей

Напомним, что случайная величина есть функция, определённая на пространстве элементарных событий Ω. Следовательно, сходимость последовательности случайных величин можно понимать в разных смыслах, аналогично тому, как для функциональных последовательностей существуют поточечная сходимость, равномерная, сходимость по норме и т.д.

♦ Говорят, что *последовательность случайных величин* $\{\xi_n(\omega)\}$ *сходится к величине* $\xi(\omega)$ *с вероятностью единица* или «почти наверное», и пишут $\xi_n(\omega) \to \xi(\omega)$, если $\lim\limits_{n\to\infty} \xi_n(\omega) = \xi(\omega)$ для любого $\omega \in \Omega$ кроме, может быть, $\omega \in A$, где множество A (событие A) имеет нулевую вероятность.

Заметим, что в общем случае, чтобы говорить о сходимости с вероятностью единица, необходимо знать структуру отображений $\omega \to \xi_n(\omega)$, в задачах же теории вероятностей известны, как правило, не сами случайные величины, а их распределения. Поэтому более распространёнными оказываются следующие определения сходимости.

♦ Говорят, что *последовательность случайных величин* $\{\xi_n\}_{n=1}^{\infty}$ *сходится к величине* ξ *по вероятности*, и пишут $\xi_n \xrightarrow[n\to\infty]{p} \xi$, если для любого $\varepsilon > 0$ справедливо

$$P(|\xi_n - \xi| > \varepsilon) \xrightarrow[n\to\infty]{} 0 \quad \text{или} \quad P(|\xi_n - \xi < \varepsilon) \xrightarrow[n\to\infty]{} 1.$$

Кроме сходимостей «почти наверное» и по вероятности, в теории вероятностей используется также сходимость по распределению.

♦ Говорят, что *последовательность случайных величин* $\{\xi_n\}_{n=1}^{\infty}$ *сходится к величине* ξ *по распределению*, и пишут $\xi_n \Rightarrow \xi$ или $F_{\xi_n}(x) \xrightarrow[n\to\infty]{} F_\xi(x)$, если функциональная последовательность $\{F_{\xi_n}(x)\}_{n=1}^{\infty}$ сходится поточечно к функции $F_\xi(x)$, т.е.

$$\lim_{n\to\infty} F_{\xi_n}(x) = F_\xi(x)$$

для любого x, принадлежащего к области непрерывности функции $F_\xi(x)$.

Свойства сходимости по вероятности

Свойство 1. Пусть $\{\xi_n\}_{n=1}^{\infty}$ и $\{\eta_n\}_{n=1}^{\infty}$ — последовательности случайных величин, определённых на вероятностном пространстве Ω. Если $\xi_n \xrightarrow[n\to\infty]{p} \xi$, $\eta_n \xrightarrow[n\to\infty]{p} \eta$, то предел по вероятности суммы (произведения) последовательностей равен сумме (произведению) их пределов, т.е.

$$\xi_n + \eta_n \xrightarrow[n\to\infty]{p} \xi + \eta, \qquad \xi_n\eta_n \xrightarrow[n\to\infty]{p} \xi\eta.$$

Свойство 2. Пусть $g(x)$ — непрерывная функция. Тогда случайная функция $g(\xi_n)$ непрерывна по вероятности, т.е. для любой случайной последовательности $\{\xi_n\}_{n=1}^{\infty}$, сходящейся по вероятности к случайной величине ξ ($\xi_n \xrightarrow[n\to\infty]{p} \xi$), справедливо $g(\xi_n) \xrightarrow[n\to\infty]{p} g(\xi)$.

Свойство 3. Из сходимости с вероятностью единица следует сходимость по вероятности, т.е. если $\{\xi_n\}_{n=1}^{\infty}$ — последовательность случайных величин, сходящаяся с вероятностью единица к случайной величине ξ ($\xi_n \to \xi$), то $\xi_n \xrightarrow{p} \xi$.

Свойство 4. Если последовательность случайных величин $\{\xi_n\}_{n=1}^{\infty}$ сходится по распределению к ξ ($\xi_n \xrightarrow[n\to\infty]{} \xi$), а последовательность $\{\eta_n\}_{n=1}^{\infty}$ сходится по вероятности к постоянной c ($\eta_n \xrightarrow[n\to\infty]{p} c$), то справедливо $\xi_n\eta_n \xrightarrow[n\to\infty]{} c\xi$, $\eta_n + \xi_n \xrightarrow[n\to\infty]{} c + \xi$.

Свойство 5. Если последовательность случайных величин $\{\xi_n\}_{n=1}^{\infty}$ сходится по распределению к постоянной ($\xi_n \xrightarrow[n\to\infty]{} c$), то последовательность $\{\eta_n\}_{n=1}^{\infty}$ сходится к постоянной c по вероятности, т.е. $\xi_n \xrightarrow[n\to\infty]{p} c$.

Свойство 6. Из сходимости по вероятности следует сходимость по распределению, т.е. если $\xi_n \xrightarrow{p} \xi$, то $\xi_n \xrightarrow[n\to\infty]{} \xi$.

Свойство 7. Пусть $\gamma_{\xi_n}(t)$, $\gamma_\xi(t)$ — характеристические функции величин ξ_n и ξ. Тогда если $F_{\xi_n}(x) \to F_\xi(x)$, то $\gamma_{\xi_n}(x) \to \gamma_\xi(x)$, и наоборот: если $\gamma_{\xi_n}(x) \to \gamma_\xi(x)$, то $F_{\xi_n}(x) \to F_\xi(x)$.

Свойство 8. Если последовательность $\{\xi_n\}_{n=1}^{\infty}$ сходится по распределению к ξ ($\xi_n \xrightarrow[n\to\infty]{} \xi$) и функция распределения $F_\xi(x)$ непрерывна в точках a и b, то $P(\xi_n \in [a,b]) \xrightarrow[n\to\infty]{} P(\xi \in [a,b])$.

Докажем первое свойство. Рассмотрим $P(|\xi_n + \eta_n - \xi - \eta| > \varepsilon)$. Заметим, что $|\xi_n + \eta_n - \xi - \eta| \leqslant |\xi_n - \xi| + |\eta_n - \eta|$. Тогда, если $|\xi_n + \eta_n - \xi - \eta| > \varepsilon$, то и $|\xi_n - \xi| + |\eta_n - \eta| > \varepsilon$, причем $P(|\xi_n + \eta_n - \xi - \eta| > \varepsilon) \leqslant P(|\xi_n - \xi| + |\eta_n - \eta| > \varepsilon)$, так как событие $\{\xi_n + \eta_n - \xi - \eta| > \varepsilon\}$ влечет событие $\{|\xi_n - \xi| + |\eta_n - \eta| > \varepsilon\}$. Следовательно,

$$P(|\xi_n + \eta_n - \xi - \eta| > \varepsilon) \leqslant P(|\xi_n - \xi| + |\eta_n - \eta| > \varepsilon) \leqslant$$
$$\leqslant P(|\xi_n + \eta_n - \xi - \eta| > \varepsilon) \leqslant P(|\xi_n - \xi| > \varepsilon/2 \text{ или } |\eta_n - \eta| > \varepsilon/2) \leqslant$$
$$\leqslant P(|\xi_n - \xi| > \varepsilon/2) + P(|\eta_n - \eta| > \varepsilon/2) \xrightarrow[n\to\infty]{} 0.$$

Остальные свойства доказываются аналогично.

Пример 23.1. Пусть $\{\xi_n\}_{n=1}^{\infty}$ — последовательность случайных величин, каждая из которых может принимать два значения: 0 и n с вероятностями $(n-1)/n$ и $1/n$ соответственно. Показать, что эта последовательность сходится по вероятности к нулю.

Решение. Действительно, для произвольного ε, начиная с некоторого $n > \varepsilon$, будет выполнено

$$P(|\xi_n - \xi| > \varepsilon) = P(|\xi_n| > \varepsilon) = P(\xi_n = n) = \frac{1}{n} \xrightarrow[n\to\infty]{} 0.$$

$\Diamond$ В отличие от сходимости по вероятности сходимость с вероятностью единица зависит от отображения $\omega \to \xi_n(\omega)$. Пусть, например, ω — точка, случайным образом выбранная на отрезке $[0,1]$, а $\xi_n(\omega) = 0$, если $\omega \in [0,(n-1)/n]$, и $\xi_n(\omega) = n$, если $\omega \in](n-1)/n, 1]$. Тогда для любого $\omega \in [0,1[$, начиная с некоторого n, $\omega \in [0,(n-1)/n]$ и, следовательно, $\xi_n(\omega) = 0$, т.е. последовательность ξ_n сходится к нулю с вероятностью единица. Пусть теперь ω — точка, случайным образом выбранная на окружности, и $\xi_n(\omega) = n$, если ω принадлежит интервалу углов $2\pi[S_{n-1}, S_{n-1} + 1/n]$, где $S_{n-1} = \sum\limits_{k=1}^{n-1} 1/k$ $(n > 1)$, и $\xi_n(\omega) = 0$ в противном случае. То есть интервал длиной $2\pi/n$, соответствующий значениям случайной величины $\xi_n(\omega) = n$, перемещается по окружности так, что конец предыдущего интервала является началом следующего. Поскольку гармонический ряд является расходящимся, то этот интервал бесконечное число раз накроет любую точку ω, и, следовательно, говорить о сходимости с вероятностью единица неправомерно. Заметим, что если бы вероятности $P(\xi(\omega) = n)$ стремились к нулю достаточно быстро (например, $P(\xi(\omega) = n) = 1/n^2$), то построенная аналогичным образом случайная величина уже сходилась бы к нулю с вероятностью единица.

Пример 23.2. Пусть $\xi_1, \xi_2 \ldots, \xi_n$ — независимые, равномерно распределённые на отрезке $[1,4]$ случайные величины. Доказать, что $\max\{\xi_1, \xi_2 \ldots, \xi_n\} \xrightarrow[n\to\infty]{p} 4$.

Решение. Обозначим $\max\{\xi_1, \xi_2 \ldots, \xi_n\} = \eta_n$. Найдём функцию распределения $F_{\eta_n}(x)$. По определению,

$$F_{\eta_n}(x) = P(\eta_n < x) = P(\text{все } \xi_i < x,\ x = \overline{1,n}) =$$
$$= P(\xi_1 < x)P(\xi_2 < x)\cdots P(\xi_n < x) = P^n(\xi_1 < x) = F^n_{\xi_1}(x).$$

Так как ξ_1 имеет равномерное на $[1,4]$ распределение, то

$$F_{\xi_1} = \begin{cases} 0, & x \leqslant 1; \\ \dfrac{x-1}{3}, & 1 < x \leqslant 4; \\ 1, & x > 4. \end{cases}$$

Тогда

$$F_{\eta_n} = \begin{cases} 0, & x \leqslant 1; \\ \left(\dfrac{x-1}{3}\right)^n, & 1 < x \leqslant 4; \\ 1, & x > 4. \end{cases}$$

В задаче требуется доказать, что $P(|\eta_n - 4| > \varepsilon) \to 0$ для всех $\varepsilon > 0$. Найдём $P(|\eta_n - 4| > \varepsilon)$. Так как $\eta_n \leqslant 4$, то

$$P(|\eta_n - 4| > \varepsilon) = P(4 - \eta_n > \varepsilon) = P(\eta_n < 4 - \varepsilon) =$$
$$= F_{\eta_n}(4 - \varepsilon) = \left(\frac{3-\varepsilon}{3}\right)^n = \left(1 - \frac{\varepsilon}{3}\right)^n.$$

Поскольку $\varepsilon > 0$, то при $n \to \infty$

$$P(|\eta_n - 4| > \varepsilon) = \left(1 - \frac{\varepsilon}{3}\right)^n \to 0.$$

Следовательно, $\eta_n \xrightarrow[n\to\infty]{p} 4$, что и требовалось доказать.

Можно было бы решить задачу и иначе. Найдём предел при $n \to \infty$ функции распределения $F_{\eta_n}(x)$:

$$\lim_{n\to\infty} F_{\eta_n}(x) = \begin{cases} 0, & x < 4; \\ 1, & x \geqslant 4. \end{cases}$$

Значения этого предела в точках непрерывности совпадают со значениями функции распределения постоянной величины, равной 4, т.е. $F_{\eta_n}(x) \xrightarrow[n\to\infty]{} 4$. Следовательно, согласно свойствам сходимости по распределению, $\eta_n \xrightarrow[n\to\infty]{p} 4$.

24. Неравенство Чебышева

Теорема 24.1 (первое неравенство Чебышева). *Если неотрицательная случайная величина ξ имеет конечное математическое ожидание $M(\xi)$, то для любого $\varepsilon > 0$ справедливо неравенство*

$$P(\xi \geqslant \varepsilon) \leqslant \frac{M(\xi)}{\varepsilon} \quad \text{или} \quad P(\xi < \varepsilon) \geqslant 1 - \frac{M(\xi)}{\varepsilon}. \tag{24.1}$$

Доказательство. Доказательство проведём для непрерывной случайной величины. По определению, для положительной случайной величины

$$M(\xi) = \int\limits_0^\infty x\, f(x)dx \geqslant \int\limits_\varepsilon^\infty x\, f(x)dx \geqslant \varepsilon \int\limits_\varepsilon^\infty f(x)dx = \varepsilon P(\xi \geqslant \varepsilon).$$

Следовательно,

$$P(\xi \geqslant \varepsilon) \leqslant \frac{M(\xi)}{\varepsilon},$$

что и требовалось доказать.

Теорема 24.2 (второе неравенство Чебышева). *Для любой случайной величины ξ, имеющей конечную дисперсию $D(\xi)$, при любом $\varepsilon > 0$ имеет место неравенство*

$$P\big\{|\xi - M(\xi)| \geqslant \varepsilon\big\} \leqslant \frac{D(\xi)}{\varepsilon^2}. \tag{24.2}$$

Это неравенство Чебышева дает оценку вероятности того, что отклонение любой случайной величины ξ от центра распределения $m = M(\xi)$ превзойдёт заданное положительное число ε.

Доказательство. *1 способ.* Положим $\eta = [\xi - M(\xi)]^2$. Заметим, что $M(\eta) = D(\xi)$. В соответствии с первым неравенством Чебышева для величины η имеем

$$P(\eta \geqslant \varepsilon^2) \leqslant \frac{M(\eta)}{\varepsilon^2} = \frac{D(\xi)}{\varepsilon^2}.$$

Учтя, что событие $(\eta \geqslant \varepsilon^2) = \{[\xi - M(\xi)]^2 \geqslant \varepsilon^2\}$ эквивалентно событию $|\xi - M(\xi)| \geqslant \varepsilon$, получим

$$P(|\xi - M(\xi)| \geqslant \varepsilon) \leqslant \frac{D(\xi)}{\varepsilon^2}.$$

2 способ. Пусть величина ξ непрерывна. Изобразим некоторые её частные значения и математическое ожидание в виде точек на числовой оси Ox (рис. 81).

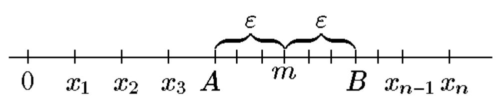


Рис. 81.

Зададим некоторое значение $\varepsilon > 0$ и вычислим вероятность того, что случайная величина ξ отклонится от своего математического ожидания $m = M(\xi)$ больше, чем на ε:

$$P\big(|\xi - M(\xi)| \geqslant \varepsilon\big) = P\big(|\xi - m| \geqslant \varepsilon\big). \tag{24.3}$$

Для этого отложим от точки m вправо и влево по отрезку длиной ε; получим отрезок AB. Вероятность (24.3) есть не что иное, как вероятность того, что при случайном бросании точка ξ попадёт не внутрь отрезка AB, а вовне его:

$$P\big(|\xi - m| \geqslant \varepsilon\big) = P(\xi \notin [A, B])$$

Нужно просуммировать вероятности для всех x, которые лежат вне отрезка $[A, B]$.

Пусть $f(x)$ есть плотность вероятности случайной величины ξ:

$$P\big(|\xi - m| \geqslant \varepsilon\big) = \int\limits_{|x-m|>\varepsilon} f(x)dx.$$

Для дисперсии имеем

$$D(\xi) = \int\limits_{-\infty}^{\infty} (x-m)^2 f(x)dx = \int\limits_{-\infty}^{\infty} |x-m|^2 f(x)dx \geqslant \int\limits_{|x-m|>\varepsilon} |x-m|^2 f(x)dx,$$

где выражение $|x - m| > \varepsilon$ означает, что интегрирование распространяется на внешнюю часть отрезка $[A, B]$.

Заменив $|x-m|$ под знаком интеграла на ε, получим

$$D(\xi) \geqslant \varepsilon^2 \int\limits_{|x-m|>\varepsilon} f(x)dx = \varepsilon^2 P\big(|\xi-m| \geqslant \varepsilon\big),$$

откуда и вытекает неравенство Чебышева для непрерывных величин (24.2).

$\lozenge$ Очевидно, что утверждение этой теоремы равносильно утверждению

$$P\Big\{|\xi - M(\xi)| < \varepsilon\Big\} \geqslant 1 - \frac{D(\xi)}{\varepsilon^2}, \tag{24.4}$$

на том основании, что события $|\xi-m| < \varepsilon$ и $|\xi-m| \geqslant \varepsilon$ противоположны.

$\lozenge$ Неравенство Чебышева дает оценку вероятности отклонения случайной величины от её математического ожидания, по существу, при любом заданном законе распределения.

Причём речь идёт только о верхней границе вероятности заданного отклонения, которую последняя не может превзойти ни при каком законе распределения. Для большинства законов данная оценка является очень грубой. Так, мы знаем, что отклонение значений нормально распределённой случайной величины от математического ожидания на величину, большую чем 3σ, возможно с вероятностью $P \approx 0{,}0027$. Если оценить эту же вероятность, используя неравенство Чебышева, то получим

$$P(|\xi - M(\xi)| \geqslant 3\sigma) \leqslant \frac{\sigma^2}{(3\sigma)^2} = \frac{1}{9} \approx 0{,}1111,$$

т.е. в 49 раз (!!!) больше.

Пример 24.1. Число пасмурных дней в году является случайной величиной ξ со средним значением 100 дней. Оценить вероятность события: число пасмурных дней в году больше двухсот ($\xi \geqslant 200$), если а) значение $D(\xi)$ неизвестно; б) известно, что $\sigma_\xi = \sqrt{D(\xi)} = 20$ дням.

Решение. а) Заданная случайная величина ξ является положительной случайной величиной с $M(\xi) = 100$. Следовательно, в соответствии с (24.1), $P(\xi \geqslant 200) \leqslant 100/200 = 0{,}5$ для $\varepsilon = 200$.

б) Так как событие $\{\xi \geqslant 200\}$ эквивалентно событию $\{\xi - M(\xi) \geqslant 100\}$, которое, в свою очередь, является подмножеством события $\big\{|\xi - M(\xi)| \geqslant 100\big\}$, то

$$P(\xi \geqslant 200) \leqslant P(|\xi - M(\xi)| \geqslant 100) \leqslant \frac{20^2}{100^2} = 0{,}04.$$

Пример 24.2. Среднее значение длины детали 50 см, а дисперсия — 0,1. Пользуясь неравенством Чебышева, оценить вероятность того, что случайно взятая деталь окажется по длине не меньше 49,5 см и не больше 50,5 см.

Решение. Пусть ξ — длина случайно взятой детали. Из условия задачи следует, что $M(\xi) = 50$, а $D(\xi) = 0{,}1$. Требуется оценить величину $P(49{,}5 < \xi < 50{,}5)$. Так как неравенство $49{,}5 < \xi < 50{,}5$ равносильно неравенству $|\xi - 50| < 0{,}5$, то требуется оценить вероятность $P(|\xi - 50| < 0{,}5)$. Согласно неравенству Чебышева (24.4), при $\varepsilon = 0{,}5$ получим

$$P(|\xi - 50| < 0{,}5) \geqslant 1 - \frac{D_\xi}{\varepsilon^2} = 1 - \frac{0{,}1}{0{,}5^2} = 0{,}6.$$

25. Теорема Чебышева

Пусть имеется последовательность $\{\xi_n\}_{n=1}^{\infty}$ независимых и одинаково распределённых случайных величин, имеющих конечные математические ожидания и дисперсии. Так как все величины одинаково распределены, то математические ожидания и дисперсии у них также одинаковы и равны, например, $M(\xi_1)$, $D(\xi_1)$. Составим случайную величину

$$\eta_n = \frac{1}{n}\sum_{i=1}^{n}\xi_i.$$

Очевидно, что η_n — случайная величина с числовыми характеристиками

$$M(\eta_n) = M\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right) = \frac{1}{n}\sum_{i=1}^{n}M(\xi_i) = \frac{1}{n}\sum_{i=1}^{n}M(\xi_1) = M(\xi_1),$$
$$D(\eta_n) = D\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}D(\xi_i) = \frac{1}{n^2}\sum_{i=1}^{n}D(\xi_1) = \frac{D(\xi_1)}{n}.$$

Таким образом, среднее арифметическое η_n есть случайная величина с тем же математическим ожиданием, что и у величин ξ_i, и с дисперсией, которая неограниченно убывает с ростом n, то есть при больших n ведёт себя, почти как неслучайная величина. Применим к величине η_n неравенство Чебышева:

$$P(|\eta_n - M(\eta_n)| < \varepsilon) \geqslant 1 - \frac{D(\eta_n)}{\varepsilon^2} = 1 - \frac{D(\xi_1)}{n\varepsilon^2}.$$

Заметим, что при $n \to \infty$ эта вероятность стремится к 1. Действительно,

$$\lim_{n\to\infty} P(|\eta_n - M(\eta_n)| < \varepsilon) = \lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - M(\xi_i)\right| < \varepsilon\right) \geqslant \lim_{n\to\infty}\left[1 - \frac{D(\xi_1)}{n\varepsilon^2}\right] = 1.$$

Таким образом, если $\{\xi_n\}_{n=1}^{\infty}$ — последовательность независимых и одинаково распределённых случайных величин, имеющих конечное математическое ожидание $M(\xi_1)$ и дисперсию $D(\xi_1)$, то среднее арифметическое этих величин сходится по вероятности к их математическому ожиданию.

Это утверждение можно обобщить и на случай, когда случайные величины имеют различные математические ожидания и дисперсии.

Теорема 25.1 (закон больших чисел в форме Чебышева). *Если $\{\xi_n\}_{n=1}^{\infty}$ — последовательность попарно независимых случайных величин, имеющих конечные дисперсии, ограниченные одной и той же постоянной C, т.е.*

$$D(\xi_n) \leqslant C, \qquad n = \overline{1,\infty},$$

то для любых постоянных $\varepsilon > 0$ имеет место соотношение

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}M(\xi_k)\right| < \varepsilon\right\} = 1, \tag{25.1}$$

где

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k$$

— среднее арифметическое n величин $\xi_1, \xi_2, \ldots, \xi_n$.

Доказательство. Так как ξ_k — взаимно независимые величины, то

$$D\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right) = \frac{1}{n^2}\sum_{k=1}^{n}D(\xi_k).$$

Следовательно,

$$D\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right) \leqslant \frac{C}{n},$$

так как $D(\xi_k) \leqslant C$. Согласно неравенству Чебышева (24.4),

$$P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}M(\xi_k)\right| < \varepsilon\right\} \geqslant 1 - \frac{1}{\varepsilon^2}D\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right) \geqslant 1 - \frac{C}{n\varepsilon^2}.$$

Переходя к пределу при $n \to \infty$, получим

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}M(\xi_k)\right| < \varepsilon\right\} \geqslant 1,$$

а так как вероятность не может быть больше единицы, то утверждение этой теоремы справедливо.

$\Diamond$ Если случайные величины ξ_k имеют одинаковые математические ожидания ($M(\xi_k) = M(\xi_1)$) и дисперсии ($D(\xi_k) = D(\xi_1)$), то из приведённого доказательства следует, в частности,

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - M(\xi_1)\right| < \varepsilon\right) \geqslant 1 - \frac{D[\xi_1]}{n\varepsilon^2},$$

или

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - M(\xi_1)\right| \geqslant \varepsilon\right) \leqslant \frac{D(\xi_1)}{n\varepsilon^2} \qquad (25.2)$$

Пример 25.1. Результат измерения некоторой физической величины является случайной величиной со значением среднеквадратичного отклонения $\sigma = 0{,}02$. Проводится серия независимых измерений и вычисляется среднее этих измерений, которое используется в качестве оценки математического ожидания результатов измерения. Требуется: а) определить с вероятностью, не меньшей 0,95, верхнюю границу абсолютной погрешности среднего пяти измерений; б) число измерений, которое необходимо произвести, чтобы с вероятностью, не меньшей 0,95, получить абсолютную погрешность среднего, меньшую чем 0,01.

Решение. а) Необходимо определить ε, для которого

$$P\left(\left|\frac{1}{5}\sum_{i=1}^{5}\xi_i - M(\xi_1)\right| < \varepsilon\right) \geqslant 0{,}95.$$

Так как результаты измерений есть одинаково распределённые случайные величины, то в соответствии с (25.2)

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - M(\xi_1)\right| < \varepsilon\right) \geqslant 1 - \frac{D(\xi_1)}{n\varepsilon^2}.$$

Следовательно,

$$1 - \frac{D(\xi_1)}{n\varepsilon^2} = 0{,}95,$$

откуда следует, что

$$\varepsilon^2 = \frac{\sigma^2_{\xi_1}}{0{,}05n} = \frac{4 \cdot 10^{-4}}{0{,}05 \cdot 5} = 1{,}6 \cdot 10^{-3} \quad \text{и} \quad \varepsilon = 4 \cdot 10^{-2}.$$

б) С помощью этого же соотношения (25.2) для $\varepsilon = 0{,}01$ и $P = 0{,}95$ получим

$$n \geqslant \frac{D(\xi_1)}{0{,}05 \cdot \varepsilon^2} = \frac{4 \cdot 10^{-4}}{0{,}05 \cdot 10^{-4}} = 80.$$

Отметим, что поскольку погрешность характеризуется значением среднеквадратичного отклонения σ, а для среднего арифметического

$$\sigma\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right) = \sqrt{\frac{D(\xi_1)}{n}} = \frac{\sigma_{\xi_1}}{\sqrt{n}},$$

то погрешность среднего как оценки математического ожидания убывает пропорционально $\sqrt{n}$. Поэтому, чтобы сделать погрешность меньшей в четыре раза, потребовалось бы в 16 раз больше измерений.

♦ Говорят, что последовательность случайных величин $\{\xi_n\}_{n=1}^{\infty}$ *удовлетворяет закону больших чисел*, если

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}M(\xi_i) \xrightarrow[n\to\infty]{p} 0.$$

Фактически это означает, что среднее арифметическое последовательности случайных величин сходится по вероятности к среднему арифметическому их математических ожиданий. Под средним арифметическим математических ожиданий здесь понимается предел при $n \to \infty$ величины

$$\frac{1}{n}\sum_{i=1}^{n}M(\xi_i).$$

◊ Предположение об ограниченности моментов второго порядка в законе больших чисел связано исключительно с приведённым способом доказательства. Закон больших чисел остается верным и в случае существования только первых моментов последовательности случайных величин. Более того, справедлива следующая теорема.

Теорема 25.2 (закон больших чисел в форме Хинчина). *Если $\{\xi_n\}_{n=1}^{\infty}$ — последовательность независимых и одинаково распределённых случайных величин, имеющих конечное математическое ожидание $M(\xi_1)$, то среднее арифметическое этих величин сходится с вероятностью единица к их математическому ожиданию:*

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i \to M(\xi_1). \tag{25.3}$$

Доказательство см., например, в [22].

Пример 25.2. Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — независимые случайные величины. При любом $k \geqslant 1$ величина ξ_{2k-1} имеет распределение Пуассона с параметром $\alpha = 2$, а величина $\xi_{2k} \in N_{0,1}$. Найти предел η по вероятности последовательности $\{\eta_k\}_{n=1}^{\infty}$, где $\eta_n = (\xi_1 + \xi_2 + \ldots + \xi_n)/n$.

Решение. Поскольку дисперсии величин ξ_n ограничены ($D(\xi_{2k-1}) = 2$, $D(\xi_{2k}) = 1$), то в соответствии с законом больших чисел среднее арифметическое этих величин сходится по вероятности к среднему арифметическому их математических ожиданий, т.е.

$$\eta_n \xrightarrow[n\to\infty]{p} \eta = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} M(\xi_i).$$

Найдём $\lim\limits_{n\to\infty} \sum\limits_{i=1}^{n} M(\xi_i)$. Для четных $n = 2k$:

$$\frac{1}{2k}\sum_{i=1}^{2k} M(\xi_i) = \frac{2k + 0\cdot k}{2k} = 1,$$

и, соответственно,

$$\lim_{k\to\infty} \frac{1}{2k}\sum_{i=1}^{2k} M(\xi_i) = 1.$$

Для нечетных $n = 2k - 1$:

$$\frac{1}{2k-1}\sum_{i=1}^{2k-1} M(\xi_i) = \frac{2k + (k-1)\cdot 0}{2k-1} = \frac{2k}{2k-1},$$

и, соответственно,

$$\lim_{k\to\infty} \frac{1}{2k-1}\sum_{i=1}^{2k-1} M(\xi_i) = 1.$$

Таким образом, $\eta_n \xrightarrow[n\to\infty]{p} 1$.

Пример 25.3. Дана последовательность независимых случайных величин $\{\xi_n\}_{n=0}^{\infty}$, каждая из которых равномерно распределена на отрезке $[a_n, b_n]$, причем $b_0 - a_0 = 0$ и

$$b_n - a_n = b_{n-1} - a_{n-1} + \frac{1}{n}$$

для $n \geqslant 1$. Удовлетворяет ли эта последовательность закону больших чисел?

Решение. Последовательность случайных величин $\{\eta_n\}_{n=1}^{\infty}$ удовлетворяет закону больших чисел, если

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n} M(\xi_i) \xrightarrow[n\to\infty]{p} 0.$$

Согласно неравенству Чебышева, если

$$D\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right) \to 0,$$

то последовательность удовлетворяет закону больших чисел.

Найдём

$$D\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right).$$

Для равномерно распределённой на отрезке $[a,b]$ величины дисперсия определяется выражением

$$D = \frac{(b-a)^2}{12}.$$

Тогда

$$\begin{aligned}
&D(\xi_1) = \frac{1^2}{12};\\
&D(\xi_2) = \frac{(1+1/2)^2}{12};\\
&D(\xi_3) = \frac{(1+1/2+1/3)^2}{12};\\
&\dots\dots\dots\dots\dots\dots;\\
&D(\xi_n) = \frac{1}{12}\left(\sum_{i=1}^{n}\frac{1}{n}\right)^2.
\end{aligned}$$

Поскольку ξ_i независимы, то

$$D\left(\frac{1}{n}\sum_{l=1}^{n}\xi_l\right) = \frac{1}{n^2}\sum_{l=1}^{n}D(\xi_l) = \frac{1}{n^2}\left[\sum_{l=1}^{n}\frac{1}{12}\left(\sum_{k=1}^{l}\frac{1}{k}\right)^2\right] = \frac{1}{12n^2}\sum_{l=1}^{n}\left(\sum_{k=1}^{l}\frac{1}{k}\right)^2.$$

Оценим сумму, стоящую в скобках:

$$\sum_{k=1}^{l}\frac{1}{k} < 1 + \int\limits_{1}^{l}\frac{dx}{x} = 1 + \ln l.$$

Следовательно,

$$D\left(\frac{1}{n}\sum_{l=1}^{n}\xi_i\right) < \frac{1}{12n^2}\sum_{l=1}^{n}(1+\ln l)^2 < \frac{1}{12n^2}n(1+\ln n)^2 = \frac{(1+\ln n)^2}{12n}.$$

Тогда

$$\lim_{n\to\infty} D\left(\frac{1}{n}\sum_{l=1}^{n}\xi_l\right) \leqslant \lim_{n\to\infty}\frac{(1+\ln n)^2}{12n} = 0,$$

т.е. последовательность удовлетворяет закону больших чисел.

26. Теорема Бернулли

Одним из важных следствий закона больших чисел является так называемая теорема Бернулли, объясняющая факт устойчивости относительной частоты события в серии независимых испытаний и дающая теоретическое обоснование статистическому определению вероятности события.

Теорема 26.1. *Пусть μ — число наступлений события A в n независимых испытаниях и p — вероятность наступления события A в каждом из испытаний. Тогда, каково бы не было $\varepsilon > 0$,*

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = 1,$$

т.е. каким бы малым положительным числом ε ни было, вероятность события

$$\left|\frac{\mu}{n} - p\right| < \varepsilon$$

стремится к единице.

Доказательство. Рассмотрим случайную величину $\xi = \mu/n$. В разделе «Схема Бернулли» мы убедились, что для распределения Бернулли $M(\mu) = np$ и $D(\mu) = npq$. Тогда

$$M\left(\frac{\mu}{n}\right) = \frac{1}{n}M(\mu) = \frac{np}{n} = p; \tag{26.1}$$

$$D\left(\frac{\mu}{n}\right) = \frac{1}{n^2}D(\mu) = \frac{npq}{n^2} = \frac{pq}{n}. \tag{26.2}$$

Запишем теорему Чебышева в форме (24.4). Для нашего случая с учётом (26.1) и (26.2) получим

$$P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} \geqslant 1 - \frac{pq}{n\varepsilon^2}. \tag{26.3}$$

Переходя к пределу при $n \to \infty$, найдём

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = 1. \tag{26.4}$$

Таким образом, вероятность того, что отклонение относительной частоты μ/n от вероятности p события A превзойдёт заданное число $\varepsilon > 0$, стремится к нулю при неограниченном увеличении числа опытов ($n \to \infty$). Иными словами, можно записать следующую теорему.

Теорема 26.2 (Бернулли). *Большие отклонения относительной частоты μ/n от вероятности события при любом сколь угодно большом n маловероятны.*

◊ Теорема Бернулли в символическом виде записывается следующим образом:

$$\frac{\mu}{n} \xrightarrow[n\to\infty]{p} p$$

и читается так: относительная частота μ/n стремится по вероятности к числу p при $n \to \infty$.

Пример 26.1. Монету подбрасывают 10 000 раз. Оценить вероятность того, что отклонение относительной частоты от вероятности выпадения орла составит менее 0,01.

Решение. Имеем $n = 10000$, $\varepsilon = 0{,}01$, $p = q = 0{,}5$. Тогда в соответствии с (26.3)

$$P(|\mu - p| < \varepsilon) > 1 - \frac{pq}{n\varepsilon^2} = 1 - \frac{0{,}25}{10000\cdot(0{,}01)^2} = 0{,}75.$$

27. Центральная предельная теорема Ляпунова

Иногда приходится иметь дело со случайной величиной, которая является суммой большого числа малых и независимых между собой случайных величин, законы распределения которых неизвестны. Такой случайной величиной будет, например, ошибка при измерении, являющаяся суммой большого числа малых и независимых между собой ошибок, порождаемых различными факторами (состояние измерительного прибора, зрения, влажности, температуры, атмосферного давления, ошибка наблюдения и т.д.). Наблюдатель имеет дело лишь с суммарным действием отдельных факторов, и его интересует закон распределения «суммарной ошибки».

Ответ на вопрос о поведении функции распределения суммы таких величин даёт так называемая центральная предельная теорема Ляпунова.

Рассмотрим систему n взаимно независимых случайных величин $\{\xi_k\}_{k=1}^n$:

$$\sum_{k=1}^{n} \xi_k.$$

Центрируем и нормируем её. Так как

$$M\left(\sum_{k=1}^{n} \xi_k\right) = \sum_{k=1}^{n} m_k;$$

$$D\left(\sum_{k=1}^{n} \xi_k\right) = \sum_{k=1}^{n} D_k,$$

где обозначено $m_k = M(\xi_k)$, $D_k = D(\xi_k)$, то величины

$$\eta_n = \frac{\sum\limits_{k=1}^{n} \xi_k - M\left(\sum\limits_{k=1}^{n} \xi_k\right)}{\sigma\left(\sum\limits_{k=1}^{n} \xi_k\right)} = \frac{\sum\limits_{k=1}^{n} \xi_k - \sum\limits_{k=1}^{n} m_k}{\sqrt{\sum\limits_{k=1}^{n} D_k}}$$

являются нормированными. Справедливо следующее утверждение.

Теорема 27.1 (Ляпунова). *Пусть простые случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ взаимно независимы, имеют конечные математические ожидания $m_1, m_2, \ldots$, конечные дисперсии $D_1, D_2, \ldots$ и конечные третьи центральные моменты $\beta_1, \beta_2, \ldots$ Тогда функция распределения нормированной и центрированной случайной величины*

$$\eta_n = \frac{\sum\limits_{k=1}^{n} \xi_k - \sum\limits_{k=1}^{n} m_k}{\sqrt{\sum\limits_{k=1}^{n} D_k}} \tag{27.1}$$

стремится при $n \to \infty$ к функции распределения нормальной случайной величины с параметрами $m = 0$ и $\sigma = 1$ равномерно относительно x:

$$F_{\eta_n}(x) = P(\eta_n < x) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2} dt = \frac{1}{2} + \Phi(x), \tag{27.2}$$

если выполнено условие Ляпунова

$$\frac{\sum\limits_{k=1}^{n}\beta_k}{\left(\sum\limits_{k=1}^{n}D_k\right)^{3/2}} \xrightarrow[n\to\infty]{} 0. \tag{27.3}$$

Здесь $\Phi(x)$ *— функция Лапласа* (19.6).

Доказательство. Доказательство проведём при упрощающем предположении, что случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ имеют одинаковые математические ожидания $M(\xi_1) = m$ и дисперсии $D(\xi_1) = \sigma^2$. Тогда случайная величина (27.1) запишется в виде

$$\eta_n = \frac{\sum\limits_{k=1}^{n}\xi_k - \sum\limits_{k=1}^{n}M(\xi_1)}{\sqrt{\sum\limits_{i=1}^{n}D(\xi_1)}} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\xi_i - M(\xi_1)}{\sqrt{D(\xi_1)}}. \tag{27.4}$$

Поскольку $[\xi_i - M(\xi_i)]/\sqrt{D(\xi_1)}$ — центрированные и нормированные величины, то, не нарушая общности, будем считать, что $\xi_1, \xi_2, \ldots, \xi_n$ — стандартные случайные величины с математическим ожиданием, равным нулю, и дисперсией, равной единице, и рассмотрим величину

$$\eta_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i.$$

Заметим, что $M(\eta_n) = 0$, $D(\eta_n) = 1$. Пусть $\gamma_\xi(t)$ — характеристическая функция величин ξ_i. Согласно свойствам (12.18) и (12.22) характеристической функции, запишем $\gamma_{\eta_n}(t) = \gamma_\xi^n(t/\sqrt{n})$. Разложим $\gamma_\xi(t/\sqrt{n})$ в ряд Маклорена, учитывая, что, согласно (12.17), $g(0) = 1$, а согласно (12.19), $\gamma^{(k)}(t)|_{t=0} = i^k m_k$:

$$\gamma_{\eta_n}(t) = \left(1 + im_1\frac{t}{\sqrt{n}} + \frac{i^2 m_2 t^2}{2!n} + \ldots\right)^n = \left[1 - \frac{t^2}{2n} + o(t^2/n)\right]^n$$

или

$$\ln[\gamma_{\eta_n}(t)] \sim n\ln\left[1 - \frac{t^2}{2n} + o\Big(\frac{t^2}{n}\Big)\right] = n\Big[-\frac{t^2}{2n} + o\Big(\frac{t^2}{n}\Big)\Big] = -\frac{t^2}{2} + n\cdot o\Big(\frac{t^2}{n}\Big).$$

Переходя к пределу при $n \to \infty$, получим

$$\lim_{n\to\infty}\ln[\gamma_{\eta_n}(t)] = \lim_{n\to\infty}\Big[-\frac{t^2}{2} + n\cdot o\Big(\frac{t^2}{n}\Big)\Big] = -\frac{t^2}{2}$$

или $\lim\limits_{n\to\infty}\gamma_{\eta_n}(t) = e^{-t^2/2}$, то есть $\gamma_\eta(t)$ стремится к характеристической функции (12.18) стандартного нормального закона $N(0,1)$. Следовательно, по свойству сходимости по распределению функция распределения величины

$$\eta = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i$$

стремится к функции распределения $F_{0,1}(x)$ стандартной нормальной величины (19.15), что и требовалось доказать.

Иначе говоря, при выполнении условий теоремы нормальный закон

$$F(x) = \frac{1}{2} + \Phi(x)$$

является предельным для закона распределения центрированной и нормированной суммы (или, что то же самое, среднего) случайных величин (27.1).

Итак, при выполнении условий теоремы и большом n

$$F_{\eta_n}(x) \approx \frac{1}{2} + \Phi(x). \tag{27.5}$$

Тогда закон распределения $\sum\limits_{k=1}^{n} \xi_k$ является приближенно нормальным с математическим ожиданием $\sum\limits_{k=1}^{n} m_k$ и стандартным отклонением $\sqrt{\sum\limits_{k=1}^{n} D_k}$, т.е.

$$P_k\left(\sum_{k=1}^{n} \xi_k < x\right) \approx \frac{1}{2} + \Phi\left(\frac{x - \sum\limits_{k=1}^{n} m_k}{\sqrt{\sum\limits_{k=1}^{n} D_k}}\right) = \frac{1}{2} + \Phi\left(\frac{x-m}{\sigma}\right). \tag{27.6}$$

Следствие 27.1.1. Если случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ независимы, одинаково распределены и имеют конечное математическое ожидание $M(\xi_1)$ и дисперсию $D(\xi_1)$, то при $n \to \infty$ для любого $x \in \mathbb{R}$ справедливо

$$P\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\xi_i - M(\xi_1)}{\sqrt{D(\xi_1)}} < x\right) \xrightarrow[n\to\infty]{} F_{0,1}(x). \tag{27.7}$$

Пример 27.1. Случайная величина ξ является средним арифметическим независимых и одинаково распределённых случайных величин, дисперсия каждой из которых равна 5. Сколько нужно взять таких величин, чтобы ξ с вероятностью, не меньшей 0,9973, имела отклонение от своего математического ожидания, не превосходящее 0,1. Решить задачу, используя: а) неравенство Чебышева; б) центральную предельную теорему.

Решение. Согласно определениям и условию задачи,

$$\xi = \frac{1}{n}\sum_{i=1}^{n} \xi_i, \quad M(\xi) = M(\xi_1), \quad D(\xi) = \frac{D(\xi_1)}{n} = \frac{5}{n}.$$

а) Согласно неравенству Чебышева,

$$P(|\xi - M(\xi)| < \varepsilon) \geqslant 1 - \frac{D(\xi)}{\varepsilon^2}$$

или

$$P(|\xi - M(\xi)| < 0{,}1) \geqslant 1 - \frac{5}{0{,}1^2 n},$$

откуда получим $n \approx 185185$.

б) Согласно центральной предельной теореме, распределение ξ является приближенно нормальным с параметрами

$$M(\xi) = M(\xi_1), \qquad D(\xi) = \frac{D(\xi_1)}{n}.$$

Тогда

$$P(|\xi - M(\xi)| < \varepsilon) \approx 2\Phi\Big(\frac{\varepsilon}{\sqrt{D(\xi)}}\Big) = 2\Phi\Big(\frac{0{,}1}{\sqrt{5/n}}\Big).$$

По условию $P(|\xi - M(\xi)| < \varepsilon) = 0{,}9973$, т.е.

$$2\Phi\Big(\frac{0{,}1}{\sqrt{5/n}}\Big) = 0{,}9973.$$

По таблицам функции Лапласа найдём x, для которого $\Phi(x) = 0{,}9973/2$: $x = 3$. Тогда

$$\frac{0{,}1}{\sqrt{5/n}} = 3, \qquad n = 150.$$

Оцените разницу!!! Таким образом, оценка, полученная с помощью неравенства Чебышева, является более грубой, чем аналогичная оценка, полученная с помощью центральной предельной теоремы.

28. Теоремы Муавра–Лапласа и Пуассона

Теорема 28.1 (Муавра–Лапласа). *Пусть ξ — случайная величина, распределённая по биномиальному закону с параметрами n и p, тогда при $n \to \infty$ случайная величина*

$$\eta_n = \frac{\xi - np}{\sqrt{npq}} \tag{28.1}$$

является стандартной нормальной случайной величиной ($\eta_\infty \in N(0,1)$).

Доказательство. Пусть ξ — число успехов в n испытаниях по схеме Бернулли. Представим ξ как $\xi = \sum\limits_{i=1}^{n} \xi_i$, где ξ_i — число успехов в i-том испытании. Очевидно, что каждая из величин ξ_i может принимать два значения: $\xi_i = 1$ с вероятностью p и $\xi_i = 0$ с вероятностью $q = 1 - p$. Тогда $M(\xi_i) = 1 \cdot p + 0 \cdot q = p$, $D(\xi_i) = (1-p)^2 p + (0-p)^2 q = pq$.

Следовательно, при $n \to \infty$ для случайной величины $\xi = \sum\limits_{i=1}^{n} \xi_i$ справедлива центральная предельная теорема:

$$\frac{\sum\limits_{i=1}^{n} \xi_i - nM[\xi_1]}{\sqrt{nD[\xi_1]}} = \frac{\xi - np}{\sqrt{npq}} = \eta_n, \qquad \eta_\infty \in N(0,1),$$

что и требовалось доказать.

Как следствие теоремы Муавра–Лапласа сразу же получаем интегральную формулу Лапласа.

Теорема 28.2 (интегральная теорема Муавра–Лапласа). *Пусть ξ — число появлений события A в n независимых испытаниях, в каждом из которых вероятность появления события A равна p $(0 < p < 1)$ (схема Бернулли). Тогда*

$$P(\alpha < \xi < \beta) \xrightarrow[n\to\infty]{} \Phi\left(\frac{\beta - np}{\sqrt{npq}}\right) - \Phi\left(\frac{\alpha - np}{\sqrt{npq}}\right), \tag{28.2}$$

где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-t^2/2} dt$$

— функция Лапласа.

Эта теорема находит применение в приближённых вычислениях биномиальных вероятностей и их сумм.

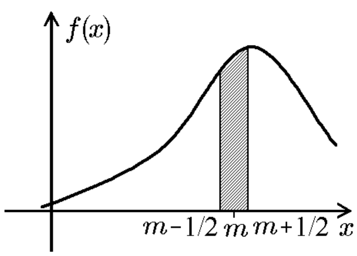


Рис. 82.

Вероятность того, что биномиальная случайная величина примет значение m, можно считать приближённо равной вероятности попадания нормальной случайной величины в малый интервал $(m - 1/2, m + 1/2)$ (рис. 82). Поскольку в пределах этого интервала плотность вероятности можно считать постоянной и равной $f(m)$, то эта вероятность приближённо равна площади прямоугольника:

$$P(\xi = m) \approx f(m) \cdot 1 = f(m) = \frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-(m-np)^2/(2npq)}.$$

Таким образом, мы приходим к локальной формуле Муавра–Лапласа.

Теорема 28.3 (локальная формула Муавра–Лапласа). *Пусть ξ — случайная величина, распределённая по биномиальному закону с параметрами n и p, тогда при $n \to \infty$*

$$P(\xi = m) \to \frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-(m-np)^2/(2npq)} = \frac{1}{\sqrt{npq}} \varphi\left(\frac{m - np}{\sqrt{npq}}\right). \tag{28.3}$$

Пример 28.1. Произведено 700 независимых испытаний, в каждом из которых вероятность наступления события A равна 0,7. Найти вероятность того, что частота появления события A окажется заключённой между 460 и 600.

Решение. Применим интегральную теорему (формулу) Муавра–Лапласа (28.2), где $\alpha = 460$; $\beta = 600$; $n = 700$; $np = 700 \cdot 0{,}7 = 490$; $npq = 490 \cdot 0{,}3 = 147$; $\sqrt{npq} \approx 12{,}124$ и

$$\frac{\alpha - np}{\sqrt{npq}} = \frac{460 - 490}{12{,}124} = -2{,}47;$$
$$\frac{\beta - np}{\sqrt{npq}} = \frac{600 - 490}{12{,}124} = 9{,}07.$$

Тогда

$$P(460 < \xi < 600) = \Phi(9{,}09) - \Phi(-2{,}47) = \Phi(9{,}09) + \Phi(2{,}47) =$$
$$= \frac{1}{2} + 0{,}49324 = 0{,}9932.$$

Пример 28.2. Найти вероятность того, что в результате 1000 бросаний монет число выпаданий герба будет заключено в интервале]475,525[.

Решение. Воспользуемся формулой (27.5), где $\alpha = 475$; $\beta = 525$; $n = 1000$; $p = q = 0{,}5$; $np = 500$; $npq = 250$; $\sqrt{npq} \approx 15{,}8114$ и

$$\frac{\alpha - np}{\sqrt{npq}} = \frac{475 - 500}{15{,}8114} = -1{,}58;$$
$$\frac{\beta - np}{\sqrt{npq}} = \frac{525 - 500}{15{,}8114} = 1{,}58.$$

Тогда

$$P(475 < \xi < 525) = \Phi(1{,}58) - \Phi(-1{,}58) = 2\Phi(1{,}58) = 0{,}8858.$$

Теорема 28.4 (Пуассона). *Пусть ξ — случайная величина, распределённая по закону Пуассона с параметром λ, тогда при $\lambda \to \infty$ случайная величина $(\xi - \lambda)/\sqrt{\lambda}$ сходится по распределению к нормальному закону $N(0, 1)$.*

Доказательство. Для доказательства воспользуемся свойствами характеристических функций. Рассмотрим величину $\eta = (\xi - \lambda)/\sqrt{\lambda}$. Так как $M(\xi) = D(\xi) = \lambda$, то $M(\eta) = 0$, $D(\eta) = 1$. Поскольку характеристическая функция величины ξ определяется соотношением (12.14): $\gamma_\xi(t) = e^{\lambda(e^{it}-1)}$, то, согласно свойству характеристической функции $\gamma_{a\xi+b}(t) = \gamma_\xi(at)e^{ibt}$, для величины η получим

$$\gamma_\eta(t) = \gamma_\xi\Big(\frac{t}{\sqrt{\lambda}}\Big)e^{-it\sqrt{\lambda}} = e^{\lambda(e^{it/\sqrt{\lambda}}-1)}e^{-it\sqrt{\lambda}} = e^{\lambda(e^{it/\sqrt{\lambda}}-1)-it\sqrt{\lambda}}.$$

Разложим экспоненту в ряд Маклорена по степеням $t/\sqrt{\lambda}$ с точностью до слагаемых второго порядка малости и получим

$$\lambda(e^{it/\sqrt{\lambda}} - 1) - it\sqrt{\lambda} = \lambda\Big[1 + \frac{it}{\sqrt{\lambda}} + \frac{(it)^2}{\lambda 2!} + o\Big(\frac{t^2}{\lambda}\Big) - 1\Big] - it\sqrt{\lambda} = -\frac{t^2}{2} + \lambda o\Big(\frac{t^2}{\lambda}\Big).$$

Следовательно,

$$\lim_{\lambda\to\infty}\gamma_\eta(t) = \lim_{\lambda\to\infty} e^{-t^2/2+\lambda o(t^2/\lambda)} = e^{-t^2/2},$$

что совпадает с характеристической функцией стандартного нормального закона. Таким образом, закон распределения η при $\lambda \to \infty$ стремится к закону $N(0, 1)$, что и требовалось доказать.

ЧАСТЬ II

Элементы математической статистики

В противоположность теории вероятностей математическая статистика — раздел прикладной математики. Для неё характерно индуктивное построение, поскольку в этом случае мы идём от события к гипотезе.

Современная математическая статистика разрабатывает способы определения числа необходимых испытаний до начала исследования (планирование эксперимента), в ходе исследования (последовательный анализ) и решает многие другие задачи. Современную математическую статистику определяют как науку о принятии решений в условиях неопределённости.

Первая задача математической статистики — указать способы сбора и группировки статистических сведений, полученных в результате наблюдений или в результате специально поставленных экспериментов.

Вторая задача математической статистики — разработать методы анализа статистических данных в зависимости от целей исследования. Сюда относятся:

а) оценка неизвестной вероятности события; оценка неизвестной функции распределения; оценка параметров распределения, вид которого известен; оценка зависимости случайной величины от одной или нескольких случайных величин и др.

б) проверка статистических гипотез о виде неизвестного распределения или о величине параметров распределения, вид которого известен.

Итак, задача математической статистики состоит в создании методов сбора и обработки статистических данных для получения научных и практических выводов.

ГЛАВА 1

Выборочный метод

29. Статистический ряд и его характеристики

29.1. Генеральная совокупность и выборка

Пусть требуется изучить совокупность объектов, однородных относительно некоторого *качественного* или *количественного* признака, характеризующего эти объекты. Например, если имеется партия деталей, то качественным признаком может служить стандартность детали, а количественным — контролируемый её размер. Совокупность таких объектов будем обозначать через Ω.

Очевидно, что значения параметра, характеризующего объекты данной совокупности, можно рассматривать как значения некоторой случайной величины ξ, определенной на Ω.

На практике мы располагаем, как правило, ограниченными данными, полученными из этой совокупности (например, отдельными результатами измерения физических или экономических величин; результатами контроля отдельных образцов изделий; результатами социологического опроса некоторой группы людей и т.д.).

♦ Множество всех возможных значений случайной величины ξ, распределенной по закону $\mathcal{F}$, называется *генеральной совокупностью* $\mathcal{F}$.

♦ Множество $\{x_1, x_2, \ldots, x_n\}$ отдельных значений случайной величины ξ, полученных в серии из n независимых экспериментов (наблюдений), называется *выборочной совокупностью* или *выборкой* объема n из генеральной совокупности.

Итак, с одной стороны, выборка — это конкретный набор значений случайной величины. Однако, если мы повторим серию из n экспериментов, мы получим другой набор значений случайной величины $\{y_1, y_2, \ldots, y_n\}$, т.е. любое выборочное значение само является случайной величиной, распределенной по тому же закону $\mathcal{F}$. Эту величину будем обозначать прописной (заглавной) буквой латинского алфавита.

Таким образом, в математической модели выборка — совокупность независимых и одинаково распределенных случайных величин $\{X_1, X_2, \ldots, X_n\} = \{X_j\}_{j=1}^n$.

Наблюдаемые значения x_i случайной величины ξ называют *вариантами*, а последовательность вариант, записанных в порядке возрастания, — *вариационным рядом.*

♦ Вектор $\{X_{(j)}\}_{j=1}^n$, состоящий из упорядоченных по возрастанию: $X_{(1)} \leqslant X_{(2)} \leqslant \ldots \leqslant X_{(n)}$ элементов выборки $\{X_{(j)}\}_{j=1}^n$, называют *вариационным рядом.* Величину $X_{(k)}$, $k = \overline{1, n}$, называют при этом k-ой *порядковой статистикой.*

Если ξ — дискретная случайная величина, число возможных значений которой невелико, и соответственно с этим выборка содержит много повторяющихся значений, то поступают следующим образом.

Выписывают все неповторяющиеся значения в вариационном ряде ξ_i, $(i = \overline{1, m}$, $m \leqslant n)$. Подсчитывают частоты n_i — количество повторов каждого из значений x_i в выборке и определяют относительные частоты $\omega_i = n_i/n$. Очевидно, что

$$\sum_{i=1}^{m} n_i = n, \qquad \sum_{i=1}^{m} \omega_i = 1.$$

Далее результаты эксперимента записывают в виде статистического ряда.

♦ Совокупность пар чисел (x_i, n_i) называют *статистическим рядом абсолютных частот*, а совокупность пар чисел (x_i, ω_i) называют *статистическим рядом относительных частот.* Статистические ряды удобно представлять в виде таблиц.

Таблица 5

Статистический ряд абсолютных (x_i, n_i) и относительных (x_i, ω_i) частот

x_i	x_1	x_2	$\ldots$	x_m
n_i	n_1	n_2	$\ldots$	n_m
ω_i	ω_1	ω_2	$\ldots$	ω_m

Пример 29.1. Дана выборка:

$$\{1,\ 3,\ 0,\ 2,\ 4,\ 4,\ 2,\ 1,\ 3,\ 1,\ 1,\ 4,\ 2,\ 1,\ 2,\ 0,\ 1,\ 2,\ 3,\ 1\}.$$

Записать статистический ряд.

Решение. Объем выборки $n = 20$. Записываем вариационный ряд:

$$\{0,\ 0,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 2,\ 2,\ 2,\ 2,\ 2,\ 3,\ 3,\ 3,\ 4,\ 4,\ 4,\ 4\}.$$

Подсчитываем частоты и представляем выборочные данные в виде статистического ряда (табл. 6):

Таблица 6

Числовой статистический ряд абсолютных и относительных частот

x_i	0	1	2	3	4
n_i	2	6	5	3	4
ω_i	0,1	0,3	0,25	0,15	0,2

Если же величина ξ — непрерывная или число возможных значений ξ велико, то в этом случае проводят группировку данных. Для этого интервал, в котором содержатся все элементы выборки, делится на m равных (иногда неравных) последовательных, непересекающихся интервалов $[\tilde{x}_0; \tilde{x}_1[$, $[\tilde{x}_1; \tilde{x}_2[$, …, $[\tilde{x}_{m-1}; \tilde{x}_m[$ и подсчитывают частоты n_i — число элементов выборки, попавших в i-й интервал. При этом элемент, совпавший с границей интервала, относят к верхнему интервалу.

$\Diamond$ Для определения оптимальной величины интервала применяют формулу Стэрджеса
$$h = \frac{x_{\max} - x_{\min}}{1 + [\log_2 n]}. \tag{29.1}$$
Здесь квадратные скобки обозначают целую часть числа; $x_{\min}$, $x_{\max}$ — минимальный и максимальный варианты. Число интервалов группирования по формуле Стэрджеса принимают равным
$$m = 1 + [\log_2 n] \approx 1 + [3{,}322 \lg n].$$
Кроме того, разбиение на интервалы должно удовлетворять условию, чтобы частоты n_i для каждого из интервалов имели один порядок величины. В противном случае следует объединять соседние интервалы, добиваясь относительно равномерного распределения частот по интервалам. Далее подсчитываются относительные частоты $\omega_i = n_i/n$ для каждого из интервалов и плотности частот $f_i = \omega_i/\Delta_i$, где $\Delta_i = \tilde{x}_i - \tilde{x}_{i-1}$ — длины соответствующих интервалов группирования. В результате получаем следующий статистический ряд:

Таблица 7

Сгруппированный статистический ряд частот

$[\tilde{x}_{i-1}; \tilde{x}_i[$	$[\tilde{x}_0; \tilde{x}_1[$	$[\tilde{x}_1; \tilde{x}_2[$	$[\ldots[$	$[\tilde{x}_{m-1}; \tilde{x}_m[$
$x_i = (\tilde{x}_{i-1} + \tilde{x}_i)/2$	x_1	x_2	…	x_m
n_i	n_1	n_2	…	n_m
$\omega_i = n_i/n$	ω_1	ω_2	…	ω_m
$f_i = \omega_i/\Delta_i$	f_1	f_2	…	f_m

Пример 29.2. Дана выборка объемом 20 из некоторой генеральной совокупности:
$$\{0,70; -0,28; 1,24; 2,28; 2,20; 2,73; -1,18; 0,77; 2,10; -0,09; 0,31; -0,69; -0,85; 0,02; 0,23; -1,12; 0,43; 0,60; 1,13; 0,63\}.$$
Представить выборку в виде сгруппированного статистического ряда.

Решение. Записываем вариационный ряд:
$$\{-1,18; -1,12; -0,85; -0,69; -0,28; -0,09; 0,02; 0,23; 0,31; 0,43; 0,60; 0,63; 0,70; 0,77; 1,13; 1,24; 2,10; 2,20; 2,28; 2,73\}.$$
Определяем число интервалов группирования по формуле Стерджесса: $m = 1 + [\log_2 20] = 5$. Выберем в качестве нижней границы $\tilde{x}_0 = -1,2$, в качестве верхней $\tilde{x}_0 = 2{,}8$. Тогда длина каждого интервала (при условии равенстве длин интервалов): $\Delta = [2{,}8 - (-1{,}2)]/5 = 0{,}8$. Разбиваем на интервалы и формируем статистический ряд:

Таблица 8

Числовой группированный статистический ряд частот

$[\tilde{x}_{i-1}; \tilde{x}_i[$	$[-1{,}2; -0{,}4[$	$[-0{,}4; 0{,}4[$	[0,4;1,2[	[1,2; 2,0[	]2,0; 2,8[
x_i	$-0{,}8$	0,0	0,8	1,6	2,4
n_i	4	5	6	1	4
ω_i	0,2	0,25	0,3	0,05	0,2
f_i	0,25	0,3125	0,375	0,0625	0,25

29.2. Полигон и гистограмма

В качестве графического представления статистических и экспериментальных данных используют полигон и гистограмму.

♦ Полигоном частот называют ломаную линию, отрезки которой соединяют точки (x_i, ω_i) либо (x_i, f_i) (рис. 83).

Пример 29.3. Построить полигон относительных частот следующего распределения:

x_i	1,5	3,5	5,5	7,5
ω_i	0,1	0,2	0,4	0,3

Решение. См. рис. 83.

Для дискретной случайной величины полигон частот является оценкой многоугольника распределения, для непрерывной случайной величины полигон частот есть оценка кривой плотности распределения.

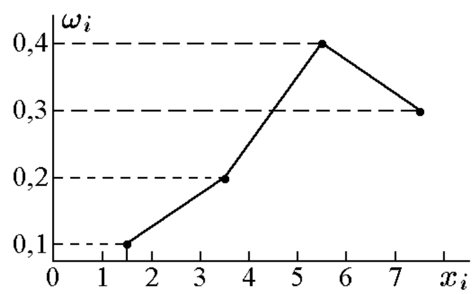


Рис. 83.

♦ Гистограмма частот — ступенчатая фигура, состоящая из m прямоугольников, опирающихся на частичные интервалы. Высота i-го прямоугольника полагается равной плотности частоты f_i. Соответственно площадь каждого прямоугольника равна $f_i\Delta_i = \omega_i$ — относительной частоте.

◊ Гистограмма частот является непараметрической оценкой плотности распределения (рис. 84).

Гистограмма исторически является первым способом оценивания функции плотности распределения, хотя и имеет невысокую точность.

29.3. Эмпирическая функция распределения

♦ *Эмпирической функцией распределения*, полученной по заданной выборке $\{x_1, x_2, \ldots, x_n\}$, называется функция, при каждом $x \in \mathbb{R}$ равная

$$F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} I(x_i < x), \tag{29.2}$$

где

$$I(x_i < x) = \begin{cases} 1, & X_i < x; \\ 0, & x_i \geqslant x. \end{cases}$$

Очевидно, что $F_n^*(x)$ — ступенчатая функция (рис. 85), имеющая разрывы в точках, соответствующих наблюдаемым выборочным значениям. Величина скачка

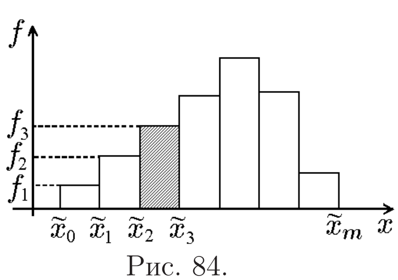


Рис. 84.

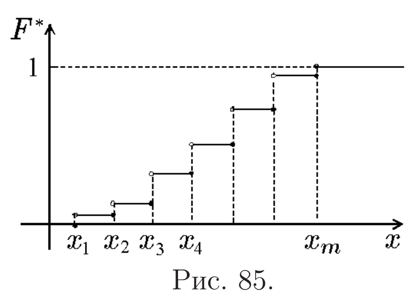


Рис. 85.

в точке x_i равна относительной частоте ω_i значения x_i. Эмпирическая функция распределения является оценкой функции распределения.

Для любого $x \in \mathbb{R}$ эмпирическая функция распределения является случайной величиной как функция случайных переменных $X_1, X_2, \ldots, X_n$.

29.4. Числовые характеристики выборки

Функцию $F^*(x)$ (29.2) можно рассматривать как функцию распределения дискретной случайной величины. Тогда в качестве числовых характеристик выборки естественно использовать следующие.

1. *Выборочным средним* называют величин

$$\bar{m} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{29.3}$$

2. *Выборочная дисперсия* определяется соотношением

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2. \tag{29.4}$$

3. *Несмещенной выборочной дисперсией* называют величину

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2. \tag{29.5}$$

4. *Выборочными начальными и центральными моментами* называют величины

$$\bar{\alpha}_k = \overline{X^k} = \frac{1}{n}\sum_{i=1}^{n} X_i^k, \qquad \bar{\mu}_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k. \tag{29.6}$$

5. *Выборочными коэффициентом асимметрии и эксцессом* называют величины

$$\bar{A} = \frac{\bar{\mu}_3}{s^3}, \qquad \bar{E} = \frac{\bar{\mu}_4}{s^4} - 3. \tag{29.7}$$

По статистическому ряду значения этих величин могут быть найдены по формулам

$$\bar{m} = \sum_{i=1}^{n} X_i\omega_i, \quad \bar{D} = \sum_{i=1}^{n}(X_i - \bar{m})^2\omega_i, \quad s^2 = \frac{n}{n-1}\sum_{i=1}^{n}(x_i - \bar{m})^2\omega_i,$$

$$\bar{\alpha}_k = \sum_{i=1}^{n} X_i^k\omega_i, \quad \bar{\mu}_k = \sum_{i=1}^{n}(X_i - \bar{m})^k\omega_i. \tag{29.8}$$

Для сгруппированных данных формулы (29.8) дают приближенные значения выборочных характеристик. Выборочные характеристики, очевидно, есть числовые характеристики дискретной случайной величины, ряд распределения которой совпадает со статистическим рядом. Выборочные характеристики являются приближенными значениями соответствующих числовых характеристик случайной величины ξ. Выборочные характеристики являются случайными величинами, так как являются функциями случайной выборки.

30. Свойства законов распределения

Рассмотрим свойства законов распределения и числовых характеристик выборок.

Теорема 30.1. *Пусть $X_1, X_2, \ldots, X_n$ — выборка из генеральной совокупности $\mathcal{F}$ с функцией распределения $F(x)$. Тогда для всех $x \in \mathbb{R}$ функция $F_n^*(x)$ при $n \to \infty$ стремится по вероятности к $F(x)$, т.е.*

$$F_n^*(x) \xrightarrow[n\to\infty]{p} F(x). \tag{30.1}$$

Доказательство. Напомним, что для среднего n независимых и одинаково распределенных случайных величин $X_1, X_2, \ldots, X_n$ справедливо:

$$M\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = M(X_1), \quad D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{D(X_1)}{n}, \quad \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} M(X_1).$$

По определению,

$$F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i < x).$$

Так как $X_1, X_2, \ldots, X_n$ независимы и одинаково распределены, то и случайные величины $I(X_1 < x), I(X_2 < x), \ldots, I(X_n < x)$ также независимы и одинаково распределены, причем каждая из них есть число успехов в одном испытании при вероятности успеха $p = P(X_1 < x) = F(x)$. Следовательно, $M(I(X_1 < x)) = p = F(x)$, $D(I(X_1 < x)) = pq = F(x)(1 - F(x))$. Тогда

$$M\left(\frac{1}{n}\sum_{i=1}^{n} I(X_i < x)\right) = M(I(X_1 < x)) = F(x),$$

$$D\left(\frac{1}{n}\sum_{i=1}^{n} I(X_i < x)\right) = \frac{D(I(X_1 < x))}{n} = \frac{F(x)(1 - F(x))}{n}.$$

Так как случайные величины имеют конечные математическое ожидание и дисперсию ($0 \leqslant F(x) \leqslant 1$), то в соответствии с законом больших чисел при $n \to \infty$

$$\frac{1}{n}\sum_{i=1}^{n} I(X_i < x) \xrightarrow[n\to\infty]{p} M(I(X_1 < x)) \quad \text{или} \quad F_n^*(x) \xrightarrow[x\to\infty]{p} F(x),$$

что и требовалось доказать.

Теорема 30.2 (Гливенко–Кантелли). *Пусть $X_1, X_2, \ldots, X_n$ — выборка из генеральной совокупности $\mathcal{F}$ с функцией распределения $F(x)$. Тогда $F_n^*(x)$ при $n \to \infty$ сходится к $F(x)$ равномерно, то есть*

$$\sup_{x\in\mathbb{R}} |F_n^*(x) - F(x)| \xrightarrow[n\to\infty]{} 0. \tag{30.2}$$

Доказательство см., например, в [4].

Теорема 30.3 (Колмогорова). *Пусть $X_1, X_2, \ldots, X_n$ — выборка из генеральной совокупности $\mathcal{F}$ с непрерывной функцией распределения $F(x)$. Тогда при $n \to \infty$ случайная величина*

$$\eta = \sqrt{n} \sup |F_n^*(x) - F(x)| \tag{30.3}$$

распределена по закону Колмогорова ($\eta \in K$).

Доказательство см., например, в [4].

Свойство 1. Эмпирическая функция распределения $F_n^*(x)$ (29.2) при $n \to \infty$ асимптотически нормальна, т.е.

$$\sqrt{n}[F_n^*(x) - F(x)] \xrightarrow[n\to\infty]{} N_{0,F(x)[1-F(x)]}. \tag{30.4}$$

Доказательство. Согласно центральной предельной теореме, если случайные величины $X_1, X_2, \ldots, X_n$ независимы и имеют одинаковые распределения с математическим ожиданием $M(X_1)$ и дисперсией $D(X_1)$, то случайная величина

$$\zeta = \left(\frac{1}{n}\sum_{i=1}^{n} X_i - M(X_1)\right) \Big/ \sqrt{D(X_1)/n}$$

сходится по распределению к нормальному распределению $N_{0,D(X_1)}$, т.е.

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - M(X_1)\right) \Rightarrow N_{0,D(X_1)}.$$

По определению (29.2),

$$F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i < x),$$

где $I(X_1 < x), I(X_2 < x), \ldots, I(X_n < x)$ независимы и одинаково распределены, причем $M(I(X_1 < x)) = F(x)$ и $D(I(X_1 < x)) = F(x)[1 - F(x)]$. Тогда

$$\sqrt{n}[F_n^*(x) - F(x)] = \sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n} I(X_i < x) - M(I(X_1 < x))\right] \Rightarrow$$
$$\Rightarrow N_{0,D(I(X_1<x))} = N_{0,F(x)[1-F(x)]},$$

что и требовалось доказать.

Теорема 30.4. *Пусть $f(x)$ — плотность распределения случайной величины ξ, а ω_i — относительная частота для i-го частичного интервала, $i = \overline{1,m}$, тогда при $n \to \infty$ и постоянном m*

$$\omega_i \xrightarrow{p} P(\tilde{x}_{i-1} < \xi < \tilde{x}_i) = \int\limits_{\tilde{x}_{i-1}}^{\tilde{x}_i} f(x)dx \tag{30.5}$$

для всех i, т.е. площадь столбца гистограммы при $n \to \infty$ стремится к площади под графиком функции плотности распределения над тем же интервалом.

Доказательство. Действительно,

$$\omega_i = F_n^*(\tilde{x}_i) - F_n^*(\tilde{x}_{i-1}) \xrightarrow[n\to\infty]{p} F(\tilde{x}_i) - F(\tilde{x}_{i-1}) = P(\tilde{x}_{i-1} < \xi < \tilde{x}_i),$$

что и требовалось доказать.

Теорема 30.5. *Пусть выборочные характеристики получены по выборке из генеральной совокупности $\mathcal{F}$ случайной величины ξ, имеющей математическое ожидание $M(\xi)$ и дисперсию $D(\xi)$. Тогда*

1. $M(\bar{X}) = M(\xi);$ (30.6)
2. $\bar{\xi} \xrightarrow{p} M(\xi);$ (30.7)
3. $M(\bar{D}) = \dfrac{n-1}{n}D(\xi), \quad M(s^2) = D(\xi);$ (30.8)
4. $\bar{D} \xrightarrow{p} D(\xi), \quad s^2 \xrightarrow{p} D(\xi);$ (30.9)
5. $M(\overline{X^k}) = M(\xi^k); \quad \overline{X^k} \xrightarrow{p} M(X^k) \quad$ (при $|M(X^k)| < \infty$). (30.10)

Доказательство. 1. По определению

$$M(\bar{X}) = M\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = M(\xi).$$

2. В соответствии с законом больших чисел справедливо

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} M(\xi).$$

3. По определению

$$M(\bar{D}) = M\left(\frac{1}{n}\sum_{i=1}^{n} \xi_i^2 - \bar{X}^2\right) = M\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) - M(\bar{X}^2),$$

где

$$M\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) = M(\xi^2)$$

как математическое ожидание среднего арифметического независимых и одинаково распределенных величин $X_1^2, X_2^2, \ldots, X_n^2$, а из формулы для дисперсии $D(\bar{X}) = M(\bar{X}^2) - M^2(\bar{X})$ имеем

$$M(\bar{X}^2) = D(\bar{X}) + M^2(\bar{X}) = \frac{D(\xi)}{n} + M^2(\xi).$$

Следовательно,

$$M(\bar{D}) = M(\xi^2) - M^2(\xi) - \frac{D(\xi)}{n} = D(\xi) - \frac{D(\xi)}{n} = \frac{n-1}{n}D(\xi).$$

Таким образом, выборочная дисперсия как оценка дисперсии генеральной совокупности является смещенной оценкой, то есть содержит систематическую ошибку, хотя асимптотически $\bar{D}$ является несмещенной оценкой дисперсии, так как $M(\bar{D}) \to D$ при $n \to \infty$. В отличие от $\bar{D}$, s^2 является несмещенной оценкой дисперсии, так как

$$M(s^2) = M\Big(\frac{n}{n-1}\bar{D}\Big) = M\Big(\frac{n}{n-1}\bar{D}\Big) = \frac{n}{n-1}\frac{n-1}{n}D(\xi) = D(\xi).$$

4. По определению

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2.$$

В соответствии с законом больших чисел для независимых и одинаково распределенных случайных величин

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p} M(\xi^2), \qquad \bar{X} \xrightarrow{p} M(\xi).$$

Следовательно,

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 \xrightarrow[n\to\infty]{p} M(\xi^2) - M^2(\xi) = D(\xi).$$

Аналогично $s^2 \xrightarrow[n\to\infty]{p} D(\xi)$, так как

$$\lim_{n\to\infty}\frac{n}{n-1} = 1.$$

5. Доказательство аналогично доказательству свойств 1, 2.

Пример 30.1. По выборке объема $n = 500$ построен сгруппированный статистический ряд распределения (таблица 9):

Таблица 9

Числовой сгруппированный статистический ряд частот по выборке объема $n = 500$

x_i	$-2{,}25$	$-1{,}75$	$-1{,}25$	$-0{,}75$	$-0{,}25$	0,25	0,75	1,25	1,75	2,25
ω_i	0,02	0,04	0,11	0,18	0,27	0,16	0,10	0,07	0 03	0,02

Построить гистограмму, полигон и эмпирическую функцию распределения, а также найти выборочные математическое ожидание, дисперсию, среднеквадратичное отклонение, асимметрию и эксцесс.

Решение. Прямой проверкой убеждаемся в справедливости условия $\sum\limits_{i=1}^{10}\omega_i = 1$.

По формулам (29.8) находим выборочные числовые характеристики:

$$\bar{m} = \sum_{i=1}^{10} x_i\omega_i = (-2{,}25)\cdot 0{,}02 + (-1{,}75)\cdot 0{,}04 + \ldots + 2{,}25\cdot 0{,}02 = -0{,}155;$$

$$\bar{D} = \sum_{i=1}^{10}(x_i - \bar{m})^2\omega_i = 0{,}858475; \quad s^2 = \frac{500}{499}\sum_{i=1}^{10}(x_i - \bar{m})^2\omega_i = 0{,}860195;$$

$$\bar{\mu}_3 = \sum_{i=1}^{10}(x_i - \bar{m})^3\omega_i = 0{,}235727; \quad \bar{\mu}_4 = \sum_{i=1}^{10}(x_i - \bar{m})^4\omega_i = 2{,}23323;$$

$$s = 0{,}927467; \quad \bar{A} = \frac{\bar{\mu}_3}{s^3} = 0{,}29636; \quad \bar{E} = \frac{\bar{\mu}_4}{s^4} - 3 = 0{,}03025.$$

Построим гистограмму плотности частот $f_i = \omega_i/\Delta_i = \omega_i/0{,}5 = 2\omega_i$ (рис. 86) — статистический аналог кривой плотности распределения. Здесь и ниже точками выделен массив $\{x_i, f_i\}$.

Построим полигон частот (рис. 87) — оценку графика плотности распределения — ломаную, отрезки которой соединяют точки $\{x_i, f_i\}$.

Если через точки массива $\{x_i, f_i\}$ провести плавную кривую (рис. 88), то мы получим приближение кривой плотности распределения $f(x)$.

В связи с этим представляет интерес сопоставить полученное эмпирическое распределение с теоретическим распределением по нормальному закону

$$f(x) = \frac{1}{s\sqrt{2\pi}} \exp\Big[-\frac{(x-m)^2}{2s^2}\Big],$$

график плотности которого для $m = -0{,}155$ и $s = 0{,}927467$ изображен на рис. 88 тонкой линией.

Близость полученного эмпирического распределения и соответствующего теоретического нормального распределения подтверждается не только графически на рис. 88, но и сравнением числовых характеристик:

Таблица 10

Числовые характеристики эмпирического и соответствующего теоретического нормального распределения

Числовые характеристики	Распределения	
	Эмпирическое	Нормальное
$\bar{\mu} \mid m$	$-0{,}155$	$-0{,}155$
$s \mid \sigma$	0,927467	0,927467
$\bar{A} \mid A$	0,29636	0
$\bar{E} \mid E$	0,03025	0

Причем по сравнению с кривой плотности нормального распределения (асимметрия $A = 0$, эксцесс $E = 0$), график эмпирической плотности распределения деформирован влево (выборочная асимметрия $\bar{A} = 0{,}29636 > 0$) и вытянут вверх (выборочный эксцесс $\bar{E} = 0{,}03025 > 0$).

Эмпирическую функцию распределения строим по формуле (29.2):

Таблица 11

Эмпирическая функция распределения

x_i	$-2{,}25$	$-1{,}75$	$-1{,}25$	$-0{,}75$	$-0{,}25$	0,25	0,75	1,25	1,75	2,25
$F^*(x_i)$	0,02	0,06	0,17	0,35	0,62	0,78	0,88	0,95	0,98	1,00

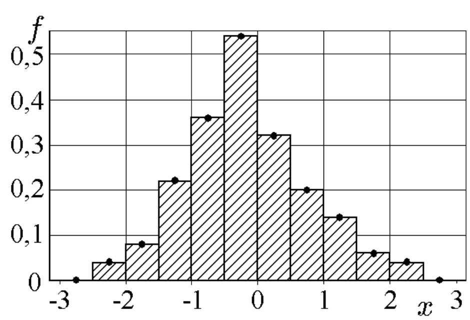


Рис. 86.

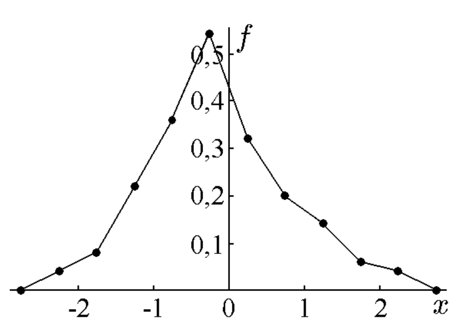


Рис. 87.

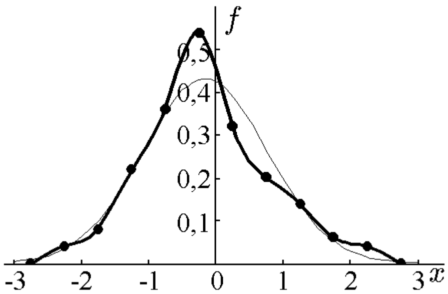


Рис. 88.

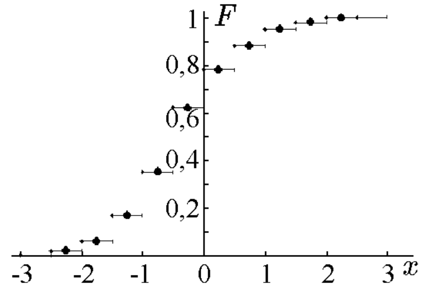


Рис. 89.

Так как эмпирическая функция распределения представлена сгруппированным рядом, то будем считать, что $F^*(x)$ меняется скачкообразно на величину ω_i при переходе через левую границу i-го интервала (рис. 89).

На рис. 89 жирными точками выделен массив $\{x_i, F^*(x_i)\}$.

Проведя через точки массива $\{x_i, F^*(x_i)\}$ плавную кривую, получим сглаженную эмпирическую функцию распределения — аналог кривой функции распределения (рис. 90).

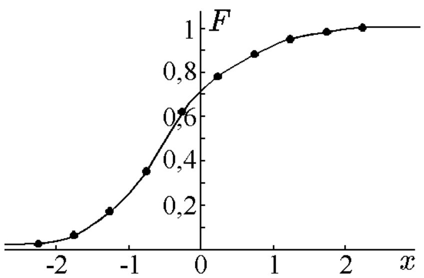


Рис. 90.

ГЛАВА 2

Точечные оценки

31. Основные понятия и определения

В статистике при проведении эксперимента зачастую тип распределения известен заранее, но неизвестны параметры или часть параметров этого распределения. Например, ошибки измерения предполагаются распределенными по нормальному закону с математическим ожиданием равным нулю (если систематическая ошибка равна нулю) и неизвестной дисперсией, число покупателей в магазине в течении часа имеет распределение Пуассона с неизвестным параметром λ и т.д. В результате мы имеем задачу статистического оценивания параметров этого распределения на основе выборочных данных.

♦ Класс распределений, целиком определяющийся значением параметра $\theta \in \Theta$ $(\theta = (\theta_1, \theta_2, \ldots, \theta_l)$, будем называть *параметрическим семейством распределений* $\mathcal{F}_\theta$.

Например, $\mathcal{F}_\theta = \Pi_\lambda$ — семейство распределений Пуассона с параметром $\lambda > 0$. Здесь $\theta = \lambda$, $\Theta = \mathbb{R}_+ =]0, \infty[$. Если $\mathcal{F}_0$ — семейство нормальных распределений, то $\theta = (m, \sigma^2)$, а $\Theta = \mathbb{R} \times \mathbb{R}_+$.

♦ Всякую однозначно определённую функцию $\theta^* = \theta^*_n = \theta^*(X_1, \ldots, X_n)$ результатов наблюдений называют *статистикой*. Статистику θ^*, с помощью которой судят о значении параметра θ, называют *оценкой* этого параметра.

Обозначим выборочную оценку параметра θ через θ^*_n. Следовательно, можно записать

$$\tilde{\theta}_n = f(X_1, X_2, \ldots, X_n).$$

Поскольку $(X_1, X_2, \ldots, X_n)$ — случайные величины, то и $\tilde{\theta}_n$ является случайной величиной.

Например, статистика

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

— выборочное среднее — может быть использована в качестве оценки математического ожидания, статистика

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

— для оценки дисперсии и т.д.

Выбор оценки, позволяющей получить «хорошее» приближение оцениваемого параметра, — основная задача теории оценивания. Основными свойствами оценок являются несмещенность, эффективность и состоятельность.

♦ Оценка θ^*_n параметра θ называется *несмещенной*, если её математическое ожидание равно оцениваемому параметру, т.е.

$$M(\theta^*_n) = \theta \tag{31.1}$$

для всех $\theta \in \Theta$. Если это равенство не выполняется, то оценка θ^*_n может либо завышать значение θ, либо занижать. В обоих случаях это приводит к систематическим ошибкам (ошибки одного знака) в оценке параметра θ.

Свойство несмещенности означает отсутствие ошибки в среднем, при систематическом использовании оценки. Так выборочное среднее является несмещенной оценкой математического ожидания (30.6), выборочная несмещенная дисперсия s^2 является несмещенной оценкой дисперсии (30.8). Выборочная дисперсия $\bar{D}$ является смещенной оценкой дисперсии (30.8).

♦ *Смещенной* называют оценку, математическое ожидание которой не равно оцениваемому параметру.

♦ Оценка параметра $\theta^* = \theta_n^*$ называется *асимптотически несмещенной оценкой параметра* θ, если

$$\lim_{n\to\infty} M(\theta_n^*) = \theta \tag{31.2}$$

для всех $\theta \in \Theta$. Так, выборочная дисперсия $\bar{D}$ является асимптотически несмещенной оценкой дисперсии.

Однако было бы ошибочным считать, что несмещенная оценка параметра всегда дает его хорошее приближение. Действительно, возможные значения θ_n^* могут быть сильно рассеяны около своего среднего значения, т.е. дисперсия $D(\theta_n^*)$ может быть значительной. В этом случае найденная по данным одной выборки оценка, например θ_n^{1*}, может оказаться весьма удалённой от среднего значения $\bar{\theta}_n^*$, а значит, и от самого оцениваемого параметра θ. Приняв θ_n^{1*} в качестве приближенного значения θ, мы допустили бы большую ошибку. Если же потребовать, чтобы дисперсия θ_n^* была малой, то возможность допустить большую ошибку будет исключена. По этой причине к статистической оценке предъявляется требование эффективности.

♦ Несмещенная оценка θ_n^*, которая имеет наименьшую дисперсию среди всех возможных несмещенных оценок параметра θ, вычисленных по выборкам одного и того же объёма, называется *эффективной*.

♦ Статистика θ_n^* называется *состоятельной оценкой параметра* θ, если при $n \to \infty$ она сходится по вероятности к оцениваемому параметру ($\theta_n^* \xrightarrow[n\to\infty]{p} \theta$), т.е. для любого $\varepsilon > 0$ выполняется следующее равенство:

$$\lim_{n\to\infty} P\{|\theta_n^* - \theta| < \varepsilon\} = 1. \tag{31.3}$$

Таким образом, состоятельная оценка должна стремиться к параметру с ростом объема выборки. Выборочное среднее и дисперсии в соответствии с (30.7), (30.9) являются состоятельными оценками математического ожидания и дисперсии.

Требования несмещенности и состоятельности являются основными требованиями, предъявляемыми к «качественным» оценкам.

Выборочное среднее $\bar{\xi}$ и выборочная дисперсия s^2 как оценки математического ожидания и дисперсии произвольного распределения не были выведены нами ни из каких математических формул, они просто разумны с практической точки зрения как аналоги характеристик генеральной совокупности, и, как видим, обладают неплохими свойствами с точки зрения требований, предъявляемых к оценке. Однако в общем случае необходим систематический и обоснованный подход к методам получения оценок.

32. Метод моментов. Метод Пирсона

Согласно *методу моментов*, в качестве оценок моментов генеральной совокупности принимаются их выборочные аналоги:

$$\tilde{\alpha}_k = \overline{m_k} = \overline{X^k}, \qquad \tilde{\mu}_k = \overline{\mu_k} = \overline{(X-m)^k}. \tag{32.1}$$

Для оценки других параметров распределения по методу моментов необходимо их выразить через начальные моменты случайной величины (в общем случае для этого необходимо знание закона распределения).

Пусть $X_1, X_2, \ldots, X_n$ — выборка из семейства распределений $\mathcal{F}_\theta$, где $\theta \in \Theta$. Пусть $M(X^k) = \alpha_k = h(\theta)$, где функция $h(x)$ имеет обратную. Тогда оценкой θ^* параметра θ будет служить решение уравнения $\overline{X^k} = h(\theta^*)$. Другими словами:

♦ Если параметр $\theta = h^{-1}(\alpha_k)$, то

$$\theta^* = h^{-1}(\alpha_k) \tag{32.2}$$

называется *оценкой метода моментов* параметра θ по k-му моменту. Если $\theta = (\theta_1, \theta_2, \ldots, \theta_l)$ — векторный параметр размерности l, то, согласно методу моментов, необходимо получить l уравнений для l различных моментов и разрешить их относительно координат вектора.

◊ Как видно, могут быть различные оценки (32.2) для одного и того же параметра (так как параметр может быть выражен через различные моменты).

◊ Как правило, априорно значения параметра считаются принадлежащими некоторому множеству, т.е. $\theta \in \Theta$, а значение полученной оценки θ^* может и не попасть в Θ. Тогда в качестве оценки θ^* берут ближайший к $h^{-1}(\alpha_k)$ элемент из множества Θ.

Например, если плотность распределения $f(x, \theta_1, \ldots, \theta_l)$ зависит от l параметров, считаем l каких-либо выборочных моментов (например, начальных) и для определения $\theta_1, \ldots, \theta_l$ получим систему уравнений

$$\begin{aligned}
&\int\limits_{-\infty}^{\infty} x f(x, \theta_1, \ldots, \theta_l) dx = \varphi_1(\theta_1, \ldots, \theta_l) = \alpha_1; \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\int\limits_{-\infty}^{\infty} x^l f(x, \theta_1, \ldots, \theta_l) dx = \varphi_l(\theta_1, \ldots, \theta_l) = \alpha_l.
\end{aligned}$$

◊ Вообще говоря, в методе моментов можно использовать любое уравнение вида $M(\varphi(\xi)) = h(\theta)$, тогда оценка $\tilde{\theta} = h^{-1}(\overline{\varphi(\xi)})$, где

$$\overline{\varphi(\xi)} = \frac{1}{n} \sum_{i=1}^{n} \varphi(\xi_i).$$

Функция $\overline{\varphi(\xi)}$ в этом случае называется *пробной функцией.*

Согласно методу моментов, в качестве оценок математического ожидания и дисперсии при любом законе распределения можно взять выборочное среднее $\bar{\xi}$ и дисперсию $\bar{D}$. (Действительно, так как $M(\xi) = \alpha_1$, то $\widetilde{M}(\xi) = \overline{\alpha_1} = \bar{\xi}$, а так как $D(\xi) = \alpha_2 - \alpha_1^2$, то $\tilde{D}(\xi) = \overline{\alpha_2} - \overline{\alpha_1}^2 = \bar{D}$).

Теорема 32.1 (состоятельность метода моментов). *Пусть $\tilde{\theta} = h^{-1}(\bar{\alpha}_k)$ — оценка параметра θ, полученная методом моментов, причем функция $h^{-1}(x)$ непрерывна, тогда $\tilde{\theta}$ состоятельна.*

Доказательство. Так как в соответствии с законом больших чисел средние $\bar{\alpha}_k$ сходятся по вероятности к α_k ($\bar{\alpha}_k \xrightarrow{p} \alpha_k$), а функция $h^{-1}(x)$ непрерывна, то $h^{-1}(\bar{\alpha}_k) \xrightarrow{p} h^{-1}(\alpha_k) = \theta$, что и требовалось доказать.

Пример 32.1. Пусть $X_1, X_2, \ldots, X_n$ — выборка из генеральной совокупности, равномерно распределенной на интервале $[0, \theta]$, где θ — неизвестный параметр. Требуется оценить θ по методу моментов.

Решение. В данном случае параметр θ может быть выражен через моменты любого порядка:

$$\alpha_k = \frac{1}{\theta}\int\limits_0^\theta x^k dx = \frac{1}{\theta}\frac{x^{k+1}}{k+1}\Big|_0^\theta = \frac{\theta^k}{k+1},$$

откуда $\theta = \sqrt[k]{(k+1)\alpha_k}$. Следовательно, $\theta^* = \sqrt[k]{(k+1)\bar{\alpha}_k}$.

Соответственно имеем различные оценки для θ: $\theta_1^* = 2\bar{\alpha}_1$, $\theta_2^* = \sqrt{3\bar{\alpha}_2}$ и т.д. Проверим несмещенность оценок: $M(\theta_1^*) = 2M(\bar{\alpha}_1) = 2\alpha_1 = \theta$, т.е. θ_1^* — несмещенная оценка, $M(\theta_2^*) = \sqrt{3}M(\sqrt{\bar{\alpha}_2})$. Чтобы оценка θ_2^* была несмещенной, очевидно, необходимо, чтобы выполнялось равенство $M(\sqrt{\bar{\alpha}_2}) = \sqrt{\alpha_2}$. Так как $M(\bar{\alpha}_2) = \alpha_2$, то это равносильно выполнению условия $M(\sqrt{\bar{\alpha}_2}) = \sqrt{M(\bar{\alpha}_2)}$, но, согласно неравенству Йенсена, $M(\varphi(\xi)) \neq \varphi(M(\xi))$, если $\varphi(x)$ не является линейной функцией. Следовательно, оценка θ_2^* — смещенная. Аналогично смещенными являются и другие оценки $\tilde{\theta}_k$ при $k > 2$.

Таким образом, несмещенной является только оценка θ_1^*, все остальные оценки являются смещенными. Заметим, что в любом случае оценка не может быть меньше, чем $x_{\max}$ — наибольшее из выборочных значений $X_1, \ldots, X_n$. Следовательно, в качестве оценки θ следует выбрать наибольшее из чисел θ_1^* и $x_{\max}$, т.е. $\theta^* = \max(\theta_1^*, \xi_{\max})$.

Пример 32.2. Пусть $X_1, X_2, \ldots, X_n$ — выборка из генеральной совокупности, распределенной по закону Коши с плотностью

$$f_\xi(x) = \frac{\theta}{\pi}\frac{1}{\theta^2 + x^2}, \qquad x \in \mathbb{R},$$

где $\theta > 0$ — неизвестный параметр. Найти оценку параметра θ по методу моментов.

Решение. Распределение Коши не имеет моментов выше первого порядка, т.е. для $k \geqslant 1$. Поэтому используем для построения оценок пробные функций вида $h^{-1}(x) = |x|^\kappa$, где $0 < \kappa < 1$.

Вычислим математическое ожидание $|x|^\kappa$:

$$M(|\xi|^\kappa) = \frac{\theta}{\pi}\int\limits_{-\infty}^{\infty} |x|^\kappa f_\xi(x)dx = \frac{\theta^\kappa}{\cos(\pi\kappa/2)} = \frac{\sqrt[m]{\theta}}{\cos(\pi/2m)}, \qquad m = \frac{1}{\kappa},$$

откуда получим

$$\theta = \left[M(|\xi|^\kappa)\cos\left(\frac{\pi\kappa}{2}\right)\right]^{1/\kappa} = \left[M|\xi|^{1/m}\cos\left(\frac{\pi}{2m}\right)\right]^m,$$

и соответственно оценки θ^* примут вид

$$\theta_\kappa^* = \left[\overline{|\xi|^\kappa}\cos\left(\frac{\pi\kappa}{2}\right)\right]^{1/\kappa} = \left[\overline{|\xi|^{1/m}}\cos\left(\frac{\pi}{2m}\right)\right]^m, \quad \overline{|\xi|^\kappa} = \frac{1}{n}\sum_{i=1}^n |\xi|^\kappa,$$

где $0 < \kappa < 1$, $m = 1/\kappa$.

33. Метод максимального правдоподобия. Метод Фишера

Пусть ξ – случайная величина, полученная в результате n испытаний, принимает значения $X_1, X_2, \ldots, X_n$. Пусть задан закон распределения величины ξ, но неизвестен параметр распределения $\theta = (\theta_1, \ldots, \theta_l)$.

При использовании *метода максимального правдоподобия* в качестве оценок параметров $\theta_1, \theta_2, \ldots$ закона распределения величины ξ берутся такие значения $\theta_1^*, \theta_2^*, \ldots$, при которых вероятность получить данную выборку $\xi = \{x_1, x_2, \ldots, x_n\}$ из генеральной совокупности величины ξ максимальна.

Выясним предварительно, что значит «вероятность получить данную выборку». Очевидно, что для дискретной случайной величины

$$P(\xi = \{x_1, x_2, \ldots, x_n\}) = P(\xi = x_1)P(\xi = x_2)\cdots P(\xi = x_n).$$

При этом мы предполагаем, что выборка получена в серии независимых испытаний. Вероятность того, что непрерывная случайная величина примет заданное значение, равна нулю, однако можно считать, что вероятность получить значение $\xi = x$ есть вероятность попасть в некоторый малый интервал Δx, содержащий x, т.е. $P(\xi = x) = f_\xi(x)\Delta x$. Тогда для непрерывной случайной величины

$$P(\xi = \{x_1, x_2, \ldots, x_n\}) = f_\xi(x_1)f_\xi(x_2)\cdots f_\xi(x_n)(\Delta x)^n.$$

♦ Функция

$$\Psi(x_1, \ldots, x_n, \theta) = \prod_{i=1}^{n} P(\xi = x_i) \tag{33.1}$$

называется *функцией правдоподобия* дискретной случайной величины.

♦ Функция

$$\Psi(x_1, \ldots, x_n, \theta) = \prod_{i=1}^{n} f(x_i) \tag{33.2}$$

называется *функцией правдоподобия* непрерывной случайной величины.

♦ Функция

$$L(x_1, \ldots, x_n, \theta) = \ln \Psi(x, \theta) \tag{33.3}$$

называется *логарифмической функцией правдоподобия.*

Как видим, функция правдоподобия — это и есть вероятность того, что случайная величина примет значения $x_1, x_2, \ldots, x_n$ (в случае непрерывной случайной величины с точностью до постоянного множителя Δx^n), причем эта вероятность зависит от неизвестного параметра распределения θ.

♦ *Оценкой параметра θ по методу максимального правдоподобия* называется значение θ^*, при котором функция максимального правдоподобия $\Psi(x_1, \ldots, x_n, \theta)$ достигает наибольшего значения как функция переменной θ в области Θ.

♢ Так как функция $\ln x$ монотонна, то максимумы функции $\Psi(x_1, \ldots, x_n, \theta)$ совпадают с максимумами функции $L(x_1, \ldots, x_n, \theta)$, которая более проста для исследования на экстремумы.

Пример 33.1. Пусть $X_1, X_2, \ldots, X_n$ — выборка из показательного распределения с неизвестным параметром λ. Найти оценку параметра λ методом максимального правдоподобия.

Решение. Для показательного распределения

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

составим функцию правдоподобия

$$\Psi(X_1,\ldots,X_n,\lambda)=\prod_{i=1}^{n}\lambda e^{\lambda X_i}=\lambda^n\exp\Big\{-\lambda\sum_{i=1}^{n}X_i\Big\}.$$

Тогда

$$\ln\Psi(X_1,\ldots,X_n,\lambda)=n\ln\lambda-\lambda\sum_{i=1}^{n}X_i.$$

Из условия

$$\frac{\partial\ln g}{\partial\lambda}\Big|_{\lambda=\lambda^*}=\frac{n}{\lambda}-\sum_{i=1}^{n}X_i\Big|_{\lambda=\lambda^*}=0$$

найдём

$$\frac{n}{\lambda^*}=\sum_{i=1}^{n}X_i.$$

Следовательно,

$$\frac{1}{\lambda^*}=\frac{1}{n}\sum_{i=1}^{n}X_i=\bar{X}.$$

Для второй производной запишем

$$\frac{\partial^2(\ln\Psi)}{\partial\lambda^2}=-\frac{n}{\lambda^2}<0,$$

следовательно, точка $1/\lambda^*=1/\bar{X}$ — точка максимума.

Пример 33.2. Пусть $X_1,X_2,\ldots,X_n$ — выборка из распределения Пуассона с неизвестным параметром λ. Требуется найти оценку λ методом максимального правдоподобия.

Решение. Для случайной величины, распределенной по закону Пуассона,

$$P(\xi=m)=\frac{\lambda^m e^{-\lambda}}{m!},\quad \lambda>0, m=\overline{0,\infty}.$$

Следовательно,

$$L(X_1,\ldots,X_n,\lambda)=\sum_{i=1}^{n}\ln\left(\frac{\lambda^{X_i}e^{-\lambda}}{X_i!}\right)=\sum_{i=1}^{n}[X_i\ln\lambda-\lambda-\ln(X_i!)]=$$
$$=\ln\lambda\sum_{i=1}^{n}X_i-\lambda n-\sum_{i=1}^{n}\ln(X_i!).$$

Найдем максимум функции $L(X_1,\ldots,X_n,\lambda)$. Вычислим

$$\frac{\partial L}{\partial\lambda}=\frac{1}{\lambda}\sum_{i=1}^{n}X_i-n$$

и из условия $\partial L/\partial\lambda=0$ получим оценку для λ:

$$\frac{1}{\lambda}\sum_{i=1}^{n}X_i-n=0.$$

Следовательно,

$$\lambda^* = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}.$$

Легко убедиться, что в точке λ^* функция правдоподобия достигает максимума.

Пример 33.3. Пусть $X_1, X_2, \ldots, X_n$ — выборка из генеральной совокупности, имеющей непрерывное распределение с плотностью $f_\xi(x)$, равной $e^{\theta - x}$ при $x \geqslant \theta$ и нулю при $x < \theta$, где θ — неизвестный параметр ($\theta > 0$). Найти оценку параметра θ по методу максимального правдоподобия.

Решение. Запишем функцию максимального правдоподобия

$$L(X_1, X_2, \ldots, X_n, \theta) = \sum_{i=1}^{n}(\ln e^{\theta - X_i}) = \sum_{i=1}^{n}(\theta - X_i) = n\theta - \sum_{i=1}^{n} X_i.$$

Данная функция как функция параметра θ является монотонно возрастающей. Следовательно, наибольшего она достигает при наибольшем из возможных значений θ. Но θ не может превосходить любое выборочное значение X_i. Таким образом, оценка по методу максимального правдоподобия равна наименьшему выборочному значению, т.е. $\theta^* = X_{(1)}$.

34. Сравнение оценок параметров распределения

Используя методы максимального правдоподобия и метод моментов, мы получили довольно много различных оценок для оценивания одних и тех же параметров. В результате возникает вопрос — какую из них предпочесть?

Как уже отмечалось выше, предпочтение следует отдать наиболее эффективной оценке, т.е. той, для которой разброс значений относительно параметра минимален.

♦ Величина $\sqrt{M([\theta^* - \theta]^2)}$ называется *среднеквадратичной ошибкой* оценки θ^* параметра θ ($M(\theta^*) = \theta$).

♦ Говорят что оценка θ_1^* лучше оценки θ_2^* в смысле среднеквадратичной ошибки или среднеквадратична, если для всех $\theta \in \Theta$ среднеквадратичная ошибка θ_1^* не больше среднеквадратичной ошибки θ_2^*, т.е.

$$M([\theta_1^* - \theta]^2) \leqslant M([\theta_2^* - \theta]^2) \qquad (34.1)$$

и хотя бы при одном θ неравенство (34.1) выполняется строго.

♢ Оценки могут быть и несравнимы: например, при некоторых θ оценка θ_1^*, согласно (34.1), лучше оценки θ_2^*, при других наоборот.

Пример 34.1. Пусть $X_1, X_2, \ldots, X_n$ — выборка из генеральной совокупности случайной величины ξ, равномерно распределенной на интервале $[0, \theta]$, где θ — неизвестный параметр. Требуется сравнить в среднеквадратичном оценку $\theta_1^* = 2\bar{X}$, полученную методом моментов по первому моменту, и оценку $\theta_2^* = X_{(n)} = \max\{X_1, X_2, \ldots, X_n\}$, найденную по методу максимального правдоподобия.

Решение. Так как математическое ожидание оценки $\theta_1^* = 2\bar{X}$ есть θ, то

$$M([\theta_1^* - \theta]^2) = D(\theta_1^*) = D(2\bar{X}) = 4D(\bar{X}) = 4\frac{D(\xi)}{n} = \frac{4\theta^2}{12n} = \frac{\theta^2}{3n}.$$

Оценка $\theta_2^* = X_{(n)}$ — смещенная, для нее имеем

$$M([\theta_2^* - \theta]^2) = M([\theta_2^*]^2 - 2\theta_2^*\theta + \theta^2) = M([\theta_2^*]^2) - 2\theta M(\theta_2^*) + \theta^2.$$

Чтобы найти первый и второй моменты случайной величины $\theta_2^* = X_{(n)}$, запишем ее функцию распределения и плотность распределения:

$$F_{X_{(n)}}(x) = P(X_{(n)} < x) = P(\text{все } X_i < x) = P(X_1 < x)\cdots P(X_n < x) = F_\xi^n(x).$$

Следовательно,

$$F_{X_{(n)}}(x) = \begin{cases} 0, & x \leqslant 0; \\ \dfrac{x^n}{\theta^n}, & 0 < x \leqslant \theta; \\ 1, & x > \theta; \end{cases} \quad f_{X_{(n)}}(x) = \begin{cases} 0, & x \leqslant 0; \\ \dfrac{nx^{n-1}}{\theta^n}, & 0 < x \leqslant \theta; \\ 0, & x > \theta. \end{cases}$$

Таким образом,

$$M(\theta_2^*) = M(X_{(n)}) = \int_0^\theta x\frac{nx^{n-1}}{\theta^n}dx = \frac{nx^{n+1}}{(n+1)\theta^n}\Big|_0^\theta = \frac{n}{n+1}\theta,$$

$$M([\theta_2^*]^2) = M(X_{(n)}^2) = \int_0^\theta x^2\frac{nx^{n-1}}{\theta^n}dx = \frac{nx^{n+2}}{(n+2)\theta^n}\Big|_0^\theta = \frac{n}{n+2}\theta^2.$$

Тогда

$$M([\theta_2^* - \theta]^2) = \frac{n}{n+2}\theta^2 - 2\frac{n}{n+1}\theta^2 + \theta^2 = \frac{2\theta^2}{(n+1)(n+2)}.$$

Можно заметить, что при $n = 1, 2$ квадратичные отклонения равны, а при $n > 2$ $M([\theta_1^* - \theta]^2) > M([\theta_2^* - \theta]^2)$, т.е. оценка $\theta_2^* = X_{(n)}$ лучше оценки $\theta_1^* = 2\bar{X}$ в смысле среднеквадратичной ошибки. Причем $(\theta_2^* - \theta)^2 \to 0$ на порядок быстрее, чем $(\theta_1^* - \theta)^2$.

Если оценка параметра есть функция вида $\varphi(\overline{X^k})$, то второй момент такой величины вряд ли удастся вычислить и, следовательно, для сравнения таких оценок среднеквадратичный подход оказывается трудно реализуемым. В этом случае, если оценка является асимптотически нормальной, то можно применить асимптотический подход.

♦ Оценка θ^* называется *асимптотически нормальной* оценкой (АНО) параметра θ с коэффициентом $\sigma^2(\theta)$, если с увеличением объема выборки ее распределение (при соответствующей нормировке) стремится к нормальному:

$$\sqrt{n}(\theta^* - \theta) \Rightarrow N_{0,\sigma^2(\theta)}. \tag{34.2}$$

◊ Для любой случайной величины ξ, имеющей конечные математическое ожидание и дисперсию, выборочное среднее $\bar{X}$, согласно центральной предельной теореме, является асимптотически нормальной оценкой математического ожидания с коэффициентом $D(\xi)$. Действительно,

$$\sqrt{n}[\bar{X} - M(\xi)] = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i - M(\xi)\right) = \frac{\sum\limits_{i=1}^n X_i - nM(\xi)}{\sqrt{n}} \Rightarrow N_{0,D(\xi)}.$$

Теорема 34.1. *Если θ^* — асимптотически нормальная оценка параметра θ, то она состоятельна.*

Доказательство. Так как θ^* — асимптотически нормальная оценка параметра θ, то $\sqrt{n}(\theta^* - \theta) \Rightarrow N_{0,\sigma^2(\theta)}$. Следовательно, величина $\eta = (\theta^* - \theta)$ распределена приближенно нормально (и тем более нормально, чем больше n) с математическим ожиданием, равным нулю, и дисперсией $D(\eta) = \sigma^2(\theta)/n$. Но тогда $(\tilde{\theta} - \theta) \Rightarrow N_{0,0} = 0$ при $n \to \infty$, что и требовалось доказать.

Таким образом, свойство асимптотической нормальности оценки показывает, что оценка сходится к параметру, причем разность ведет себя, как нормальная величина с математическим ожиданием, равным нулю, и среднеквадратическим отклонением, равным $\sigma(\theta)/\sqrt{n}$, т.е. скорость сходимости имеет порядок $1/\sqrt{n}$. Очевидно также, что при больших n разброс оценки относительно параметра тем больше, чем больше коэффициент асимптотической нормальности $\sigma^2(\theta)$, отсюда следует естественный способ сравнения асимптотически нормальных оценок.

Пусть θ_1^* — асимптотически нормальная оценка с коэффициентом $\sigma_1^2(\theta)$, а θ_2^* — такая же оценка с коэффициентом $\sigma_2^2(\theta)$.

♦ Оценка θ_1^* параметра θ *лучше* оценки θ_2^* этого параметра в смысле асимптотически нормальной ошибки, если для всех $\theta \in \Theta$ справедливо неравенство

$$\sigma_1^2(\theta) \leqslant \sigma_2^2(\theta), \tag{34.3}$$

и хотя бы при одном θ это неравенство выполняется строго.

Следующая теорема устанавливает асимптотическую нормальность оценок, полученных по методу моментов, что позволяет сравнивать такие оценки асимптотически.

Теорема 34.2. *Пусть параметр θ — некоторая функция моментов распределения: $\theta = \varphi(\alpha_k)$, а $\theta^* = \varphi(\overline{X^k})$ — оценка параметра θ по методу моментов. Тогда, если функция $\varphi(x)$ непрерывно дифференцируема в точке $x = \alpha_k$, причем $\varphi'(\alpha_k) \neq 0$, а величина X^k имеет конечную, отличную от нуля дисперсию $D(X^k)$, то оценка θ^* является асимптотически нормальной оценкой для параметра θ с коэффициентом асимптотической нормальности $\sigma^2(\theta) = [\varphi'(\alpha_k)]^2 D(X^k)$.*

Доказательство. Разложим $\varphi(x)$ в ряд Тейлора в окрестности точки α_k:

$$\varphi(x) = \varphi(\alpha_k) + \varphi'(\alpha_k)(x - \alpha_k) + r(x - \alpha_k),$$

где $r \to 0$ при $x \to \alpha_k$.

Следовательно,

$$\sqrt{n}[\varphi(x) - \varphi(\alpha_k)] = \varphi'(\alpha_k)\sqrt{n}(x - \alpha_k) + r\sqrt{n}(x - \alpha_k).$$

Положим $x = \overline{X^k}$ и получим

$$\sqrt{n}[\varphi(\overline{X^k}) - \varphi(\alpha_k)] = \varphi'(\alpha_k)\sqrt{n}(\overline{X^k} - \alpha_k) + r\sqrt{n}(\overline{X^k} - \alpha_k)$$

или $\sqrt{n}(\theta^* - \theta) = \varphi'(\alpha_k)\sqrt{n}(\overline{X^k} - \alpha_k) + r\sqrt{n}(\overline{X^k} - \alpha_k)$.

Заметим, что $r \xrightarrow{p} 0$ при $n \to \infty$, так как при $n \to \infty$ в соответствии с законом больших чисел $\overline{X^k} \xrightarrow{p} M(X^k) = \alpha_k$, а в соответствии с центральной предельной теоремой $\sqrt{n}(\overline{X^k} - \alpha_k) \Rightarrow N_{0,D(X^k)}$. Следовательно, $r\sqrt{n}(\overline{X^k} - \alpha_k) \xrightarrow{p} 0$, и

$$\sqrt{n}(\theta^* - \theta) \Rightarrow \varphi'(\alpha_k)\sqrt{n}(\overline{X^k} - \alpha_k) \Rightarrow N_{0,(\varphi'(\alpha_k))^2 D(X^k)},$$

что и требовалось доказать.

Пример 34.2. Пусть $X_1, X_2, \ldots, X_n$ — выборка из генеральной совокупности случайной величины ξ, равномерно распределенной на интервале $[0, \theta]$, где θ — неизвестный параметр. Сравнить асимптотически оценки $\theta_k^* = \sqrt[k]{(k+1)\overline{X^k}}$, полученные по методу моментов для разных k.

Решение. Чтобы найти коэффициенты асимптотической нормальности, вычислим статистические математическое ожидание $M(X^k)$ и дисперсию $D(X^k)$:

$$M(X^k) = \frac{1}{\theta}\int\limits_0^\theta x^k dx = \frac{\theta^k}{k+1},$$

$$D(X^k) = M(X^{2k}) - M^2(X^k) = \frac{\theta^{2k}}{2k+1} - \frac{\theta^{2k}}{(k+1)^2} = \frac{k^2\theta^{2k}}{(k+1)^2(2k+1)}.$$

Функция $\varphi(x) = \sqrt[k]{(k+1)x}$ дифференцируема для всех $x > 0$:

$$\varphi'(x) = \frac{\sqrt[k]{k+1}}{k}x^{1/k-1},$$

и в точке $x = M(X^k) = \alpha_k$

$$\varphi'(\alpha_k) = \frac{\sqrt[k]{k+1}}{k}\Big(\frac{\theta^k}{k+1}\Big)^{1/k-1} = \frac{k+1}{k}\theta^{1-k}.$$

Следовательно, оценки

$$\theta_k^* = \sqrt[k]{\frac{1}{n}\sum_{i=1}^n (k+1)X_i^k} = \sqrt[k]{(k+1)\overline{X^k}}$$

асимптотически нормальны с коэффициентами

$$\sigma_k^2(\theta) = [\varphi'(m_k)]^2 D(X^k) = \frac{(k+1)^2\theta^{2-2k}}{k^2}\frac{k^2\theta^{2k}}{(k+1)^2(2k+1)} = \frac{\theta^2}{2k+1}.$$

Так как коэффициент $\sigma_k^2(\theta)$ тем меньше, чем больше k, то в последовательности оценок θ_k каждая последующая лучше предыдущей. Наилучшей в смысле асимптотического подхода была бы оценка θ_∞^*, если бы она являлась асимптотически нормальной. Определим, к чему стремится θ_k^*, если $k \to \infty$:

$$\theta_k^* = \sqrt[k]{(k+1)\overline{X^k}} = \sqrt[k]{(k+1)\frac{1}{n}\sum_{i=1}^n X_i^k} \xrightarrow[k\to\infty]{} \sqrt[k]{\sum_{i=1}^n X_i^k} = \sqrt[k]{X_{(1)}^k + \ldots + X_{(n)}^k} =$$

$$= X_{(n)}\sqrt[k]{\Big(\frac{X_{(1)}}{\xi_{(n)}}\Big)^k + \Big(\frac{X_{(2)}}{\xi_{(n)}}\Big)^k + \ldots + 1} \xrightarrow[k\to\infty]{} X_{(n)}.$$

Однако $X_{(n)}$ не является асимптотически нормальной, так как $P(\sqrt{n}(X_{(n)} - \theta) > 0) = 0$ при любых n, а для асимптотически нормальной оценки при $n \to \infty$ должно быть $P(\sqrt{n}(\theta^* - \theta) > 0) = 0{,}5$. То есть, оценки $\theta_k^* = \sqrt[k]{(k+1)\overline{X^k}}$ и $X_{(n)}$ асимптотически не сравнимы.

Пример 34.3. Пусть $X_1, X_2, \ldots, X_n$ — выборка из генеральной совокупности, распределенной по нормальному закону N_{m,σ^2} с известным параметром m и неизвестным параметром σ^2. Сравнить при помощи асимптотического подхода оценки параметра σ^2:

$$s_0^2 = \frac{\pi}{2}\Big(\frac{1}{n}\sum_{i=1}^{n}|X_i - m|\Big)^2, \qquad s_1^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2.$$

Решение. Для статистики

$$s_1^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2$$

имеем

$$M((X_1 - m)^2) = D(X_1) = \sigma^2, \qquad D((X_1 - m)^2) = \sigma^4 D\Big(\frac{(X_1 - m)^2}{\sigma^2}\Big) = 2\sigma^4$$

(случайная величина $[(X_1 - m)/\sigma]^2 \in \chi_1^2$, $D(\chi_1^2) = 2$). Тогда в соответствии с центральной предельной теоремой

$$\sqrt{n}\Big[\frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2 - M((X_1 - m)^2)\Big] \Rightarrow N_{0,D((X_1 - m)^2)}$$

или $\sqrt{n}(s_1^2 - \sigma^2) \Rightarrow N_{0,2\sigma^4}$, т.е. оценка s_1^2 является асимптотически нормальной оценкой параметра σ^2 с коэффициентом асимптотической нормальности $2\sigma^4$.

Попробуем доказать асимптотическую нормальность статистики

$$s_0^2 = \frac{\pi}{2}\left(\frac{1}{n}\sum_{i=1}^{n}|X_i - m|\right)^2.$$

Покажем предварительно, что если статистика θ^* есть асимптотически нормальная оценка параметра θ, то статистика $(\theta^*)^2$ есть асимптотически нормальная оценка параметра θ^2. Пусть $\sqrt{n}(\theta^* - \theta) \Rightarrow N_{0,\sigma^2(\theta)}$, тогда $\sqrt{n}(\theta^{*2} - \theta^2) = \sqrt{n}(\theta^* - \theta)(\theta^* + \theta) \Rightarrow 2\theta\sqrt{n}(\theta^* - \theta) \Rightarrow N_{0,4\theta^2\sigma^2(\theta)}$ при условии $\theta \neq 0$. То есть статистика θ^{*2} есть асимптотически нормальная оценка параметра θ^2 с коэффициентом $4\theta^2\sigma^2(\theta)$.

Таким образом, остается доказать, что статистика

$$s_0 = \sqrt{\frac{\pi}{2}}\frac{1}{n}\sum_{i=1}^{n}|X_i - m|$$

является асимптотически нормальной оценкой параметра σ, и найти ее коэффициент асимптотической нормальности. Для этого найдем математическое ожидание и дисперсию статистики s_0:

$$M\Big(\sqrt{\frac{\pi}{2}}|X_1 - m|\Big) = \sqrt{\frac{\pi}{2}}\frac{1}{\sqrt{2\pi}\sigma}\int\limits_{-\infty}^{\infty}|x - m|e^{-(x-m)^2/2\sigma^2}dx = \sigma,$$

$$D\Big(\sqrt{\frac{\pi}{2}}|\xi_1 - a|\Big) = \frac{1}{\sqrt{2\pi}\sigma}\int\limits_{-\infty}^{\infty}\Big(\sqrt{\frac{\pi}{2}}|x - m| - \sigma\Big)^2 e^{-(x-m)^2/2\sigma^2}dx = \frac{1}{2}(\pi - 2)\sigma^2.$$

Следовательно, в соответствии с центральной предельной теоремой, статистика

$$s_0 = \sqrt{\frac{\pi}{2}} \frac{1}{n} \sum_{i=1}^{n} |X_i - m|$$

является асимптотически нормальной оценкой параметра σ с коэффициентом асимптотической нормальности $(\pi - 2)\sigma^2/2$. Тогда статистика

$$s_0^2 = \frac{\pi}{2} \left(\frac{1}{n} \sum_{i=1}^{n} |X_i - m| \right)^2$$

есть асимптотически нормальная оценка параметра σ^2 с коэффициентом

$$4\sigma^2 \frac{1}{2} (\pi - 2)\sigma^2 = 2(\pi - 2)\sigma^4 \approx 2{,}28\sigma^4.$$

Сравнив коэффициенты, делаем вывод, что статистика s_1^2 лучше статистики s_0^2 в смысле асимптотического подхода.

35. Эффективные оценки

Вернемся к сравнению оценок в смысле среднеквадратичного подхода. Сравнивать оценки попарно — далеко не лучший способ определения наилучшей оценки. Хотелось бы иметь метод, который бы позволил бы найти наилучшую оценку, или доказать, что данная оценка является наилучшей. Однако справедлива следующая теорема.

Теорема 35.1. *Среди всех возможных оценок наилучшей в смысле среднеквадратичного подхода не существует.*

Доказательство. Пусть $\theta \in \Theta$. Положим $\theta^* = \theta_0$, где θ_0 — некоторая постоянная, принадлежащая Θ. Очевидно, что для $\theta = \theta_0$ оценка θ^* будет обладать наименьшим среднеквадратичным отклонением, так как $M([\theta^* - \theta]^2) = 0$. Поскольку θ_0 произвольно, то, чтобы некоторая другая оценка θ_1^* могла быть наилучшей, необходимо, чтобы $M([\theta_1^* - \theta]^2) = 0$ для всех $\theta \in \Theta$, что возможно, лишь если эта оценка в точности совпадает с параметром, т.е. не является оценкой со статистической точки зрения.

Если среди всех оценок наилучшей не существует, то следует все оценки разбить на классы и определять лучшую в каждом классе. Обычно рассматривают классы оценок, имеющих одинаковое смещение $b(\theta) = M(\theta^*) - \theta$. Обозначим через K_b класс оценок, имеющих смещение $b(\theta)$.

♦ Оценка $\tilde{\theta} \in K_b$ называется эффективной в классе K_b, если она не хуже других оценок класса K_b в смысле среднеквадратического подхода. Эффективная оценка в классе K_0 (т.е. в классе с нулевым смещением) называется просто эффективной.

Заметим, что если $\theta^* \in K_0$, то $M([\theta^* - \theta]^2) = D(\theta^*)$, т.е. сравнение оценок в среднеквадратичном в классе K_0 — это сравнение их дисперсий. Если $\theta^* \in K_b$, то $M([\theta^* - \theta]^2) = D([\theta^*]) + M^2([\theta^* - \theta]) = D(\theta^*) + b^2(\theta)$, так что сравнение оценок с одинаковым смещением — это также сравнение их дисперсий.

Теорема 35.2 (единственность эффективной оценки). *Если θ_1^* и θ_2^* — две эффективные оценки в классе K_b, то $\theta_1^* = \theta_2^*$.*

Доказательство. Рассмотрим оценку $\theta^* = (\theta_1^* + \theta_2^*)/2$, которая также принадлежит K_b. Вычислим ее среднеквадратическое отклонение:

$$M([\theta^* - \theta]^2) = M\Big(\Big[\frac{\theta_1^* + \theta_2^*}{2} - \theta\Big]^2\Big) = M\Big(\Big[\frac{\theta_1^* - \theta + \theta_2^* - \theta}{2}\Big]^2\Big).$$

Заметим, что

$$\Big(\frac{a+b}{2}\Big)^2 + \Big(\frac{a-b}{2}\Big)^2 = \frac{a^2+b^2}{2},$$

поэтому

$$M\Big(\Big[\frac{\theta_1^* - \theta + \theta_2^* - \theta}{2}\Big]^2\Big) = \frac{M([\theta_1^* - \theta]^2) + M([\theta_2^* - \theta]^2)}{2} - M\Big(\Big[\frac{\theta_1^* - \theta_2^*}{2}\Big]^2\Big).$$

Так как $M([\theta_1^* - \theta]^2) = M([\theta_2^* - \theta]^2)$ (обе оценки эффективны), то получаем

$$M([\theta^* - \theta]^2) = M([\theta_1^* - \theta]^2) - \frac{1}{4}M([\theta_1^* - \theta_2^*]^2).$$

Но $M([\theta^* - \theta]^2) \geqslant M([\theta_1^* - \theta]^2)$, следовательно, $M([\theta_1^* - \theta_2^*]^2) = 0$ и соответственно $\theta_1^* = \theta_2^*$, что и требовалось доказать.

◊ Неравенство Рао–Крамера утверждает, что в классе K_b существует точная нижняя грань для среднеквадратичного отклонения $M([\theta^* - \theta]^2)$ любой оценки. Таким образом, если для данной оценки среднеквадратичное отклонение совпадает с этой гранью, то данная оценка эффективна. К сожалению, это неравенство верно лишь для так называемых регулярных семейств распределений.

♦ Пусть $\mathcal{F}_\theta$ — параметрическое семейство распределений с параметром $\theta \in \Theta$, имеющее плотность распределения $f(x, \theta)$ (для дискретных случайных величин $f(x, \theta) = P(\xi = x)$). Если для почти всех $x \in \mathbb{R}$ функция $\ln f(x, \theta)$ непрерывно дифференцируема по θ во всех точках $\theta \in \Theta$, *семейство $\mathcal{F}_\theta$ является регулярным.*

◊ Термин «для почти всех» означает «для всех, кроме, возможно, $x \in A$, где $P(\xi \in A) = 0$».

Пример 35.1. Привести примеры регулярного и нерегулярного семейства распределений.

Решение. 1. Рассмотрим показательное распределение с параметром $\lambda > 0$. Плотность этого распределения

$$f(x, \lambda) = \begin{cases} 0, & x \leqslant 0; \\ \lambda e^{-\lambda x}, & x > 0; \end{cases} \quad \ln f(x, \lambda) = \begin{cases} \nexists, & x \leqslant 0; \\ -\lambda x + \ln \lambda, & x > 0. \end{cases} \tag{35.1}$$

При любом $x > 0$ эта функция непрерывна по λ вместе со своей производной во всех точках $\lambda > 0$ и, следовательно, распределение (35.1) регулярно.

2. Рассмотрим равномерное распределение на отрезке $[0, \theta]$ с параметром $\theta > 0$. Плотность этого распределения имеет вид

$$f(x, \theta) = \begin{cases} \dfrac{1}{\theta}, & x \in [0, \theta]; \\ 0, & x \notin [0, \theta] \end{cases} \quad \text{или} \quad f(x, \theta) = \begin{cases} \dfrac{1}{\theta}, & \theta \geqslant x, \ x > 0; \\ 0, & \text{в противном случае}, \end{cases} \tag{35.2}$$

т.е. для любого $x > 0$ функция $f(x, \theta)$ не является даже непрерывной по θ и, следовательно, распределение (35.2) нерегулярно.

Теорема 35.3 (неравенство Рао–Крамера). *Пусть $\mathcal{F}_\theta$ — регулярное семейство распределений, имеющее плотность распределения $f(x,\theta)$, а θ^* — несмещенная оценка параметра θ, причем $D(\theta^*)$ ограничена на любом замкнутом промежутке $\Theta_0 \subset \Theta$. Тогда*

$$D(\theta^*) \geqslant \frac{1}{I_n(\theta)}, \tag{35.3}$$

где

$$I_n(\theta) = nM\Big(\Big[\frac{\partial}{\partial\theta}\ln f(\xi,\theta)\Big]^2\Big)$$

— информация Фишера, содержащаяся в выборке объема n относительно неизвестного параметра θ.

Доказательство приведено в [4].

Под информацией о неизвестном параметре θ, содержащейся в наблюдении случайной величины, понимают степень уменьшения неопределенности, касающейся неизвестного параметра θ, после наблюдения над случайной величиной. Очевидно, что чем сильнее зависимость выборочных значений от изменения значений параметра θ, тем больше информации о параметре содержит выборка.

Таким образом, если для несмещенной оценки регулярного семейства распределений будет выполнено

$$D(\theta^*) = \frac{1}{I_n(\theta)},$$

то оценка будет являться эффективной.

Пример 35.2. Проверить, является ли эффективной оценкой параметра λ распределения Пуассона выборочное среднее $\bar{X}$.

Решение. Для распределения Пуассона плотность распределения

$$f(x,\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad \lambda > 0,$$

удовлетворяет условиям регулярности при $x = \overline{0,\infty}$.

Найдем дисперсию оценки $\lambda^* = \bar{X}$:

$$D(\bar{X}) = \frac{D(\xi)}{n} = \frac{\lambda}{n}.$$

Найдем информацию Фишера относительно параметра λ:

$$I_n(\lambda) = nM\Big(\Big\{\frac{\partial}{\partial\lambda}\ln\Big[\frac{e^{-\lambda}\lambda^\xi}{\xi!}\Big]\Big\}^2\Big) = nM\Big(\Big\{\frac{\partial}{\partial\lambda}[-\lambda + \xi\ln\lambda - \ln\xi!]\Big\}^2\Big) =$$
$$= nM\Big(\Big[-1 + \frac{\xi}{\lambda}\Big]^2\Big) = \frac{n}{\lambda^2}M([\xi-\lambda]^2) = \frac{n}{\lambda^2}D(\xi) = \frac{n}{\lambda}.$$

Так как

$$D(\bar{X}) = \frac{1}{I_n(\theta)} = \frac{\lambda}{n},$$

то выборочное среднее $\bar{X}$ является эффективной оценкой параметра λ распределения Пуассона, т.е. обладает наименьшей дисперсией среди всех несмещенных оценок.

Для класса смещенных оценок со смещением $b(\theta)$ неравенство Рао–Крамера имеет вид

$$M([\theta^* - \theta]^2) \geqslant \frac{[1 + b'(\theta)]^2}{I_n(\theta)} + b^2(\theta)$$

или

$$D(\theta^*) \geqslant \frac{[1 + b'(\theta)]^2}{I_n(\theta)}. \tag{35.4}$$

$\Diamond$ Для нерегулярного семейства распределений среднеквадратичное отклонение оценки может быть меньше нижней грани, определяемой неравенствами (35.3), (35.4). Такая оценка называется *сверхэффективной*.

Заметим также, что невыполнение равенства не означает, что оценка не является эффективной.

ГЛАВА 3
Интервальное оценивание

36. Доверительный интервал и доверительная вероятность

В предыдущей главе были рассмотрены некоторые характеристики выборки (статистики), которые наилучшим образом в смысле несмещенности, эффективности и состоятельности оценивали параметры распределения генеральной совокупности.

Однако любая точечная оценка параметра распределения является приближенной величиной, поэтому, чтобы использовать ее, необходимо знать погрешность оценки, то есть границы a и b интервала, в котором находится истинное значение оцениваемого параметра. Поскольку эти границы могут быть определены только на основании случайных результатов опыта, то они также являются случайными величинами. Следовательно, необходимо не только указать интервал (a, b), но и указать надежность этого интервала, то есть вероятность того, что истинное значение параметра будет лежать в данном интервале. Следует заметить: чем больше уверенность, что параметр принадлежит интервалу, тем больше интервал. Так что искать интервал, которому принадлежит θ с вероятностью 1, бессмысленно — это вся область возможных значений параметра.

Интервальное оценивание особенно необходимо при малом числе наблюдений, когда точечная оценка в значительной мере случайна и, следовательно, мало надёжна.

♦ Интервал $]a, b[$, содержащий неизвестный параметр θ с заданной вероятностью β, называют *доверительным интервалом* (ДИ), соответствующим *доверительной вероятности* β. То есть, если $P(a < \theta < b) = \beta$, то $]a, b[$ — доверительный интервал, а β — доверительная вероятность.

♢ Так как случайными являются границы интервала, а не параметр θ, то обычно говорят «интервал $]a, b[$ накрывает параметр θ», а не «θ содержится (попадает) в интервале $]a, b[$».

♢ Для дискретных распределений точное равенство $P(a < \theta < b) = \beta$ возможно не для всех значений β. В этом случае под доверительным интервалом, соответствующим вероятности β, понимается интервал $]a, b[$, удовлетворяющий условию $P(a < \theta < b) \geqslant \beta$.

♦ Интервал $]a, b[$ называется *асимптотическим доверительным интервалом* (АДИ) для параметра θ, соответствующим *доверительной вероятности* β, если $P(a < \theta < b) \xrightarrow[n\to\infty]{} \beta$.

♦ Число $\alpha = 1 - \beta$ называют *уровнем значимости*, оно определяет вероятность того, что доверительный интервал не накроет оцениваемый параметр. Уровень значимости α отделяет события, практически невозможные, от возможных. Выбор конкретного значения α (или β) зависит от объема выборки и характера решаемой задачи. Обычно выбирают $\beta = 0{,}9;\ 0{,}95;\ 0{,}99$.

Пример 36.1. Пусть на двух предприятиях вероятность выпуска годных изделий $p = 1 - \alpha = 0{,}99$, т.е. вероятность выпуска бракованных изделий $\alpha = 0{,}01$. Можно ли в рамках математической теории, т.е. не интересуясь характером выпускаемых изделий, решать вопрос о том, мала или велика вероятность α.

Решение. Пусть одно из предприятий выпускает электролампы, а другое — парашюты. Если на 100 ламп встретится одна бракованная, то с этим мириться можно при условии, что выбросить 1% ламп дешевле, чем перестроить технологический процесс. Если же на 100 парашютов встретится один бракованный, что может повлечь за собой серьезные последствия, то с таким положением мириться нельзя. Следовательно, в первом случае вероятность брака α приемлема, а во втором нет, поэтому выбор доверительной вероятности следует производить, исходя из конкретных условий задачи.

Общий принцип построения доверительных интервалов таков:

1) Находим статистику $\eta(X_1, \ldots, X_n, \theta)$, зависящую от неизвестного параметра θ, закон распределения которой известен (и не зависит от θ). Причем необходимо, чтобы статистика $\eta(X_1, \ldots, X_n, \theta)$ была обратима относительно θ.

2) Находим квантили η_1 и η_2 распределения статистики $\eta(X_1, \ldots, X_n, \theta)$, такие что $P(\eta_1 < \eta(X_1, \ldots, X_n, \theta) < \eta_2) = \beta$. Заметим, что существует бесконечное множество пар чисел η_1, η_2, для которых $P(\eta_1 < \eta(\vec{\xi}, \theta) < \eta_2) = \beta$. Обычно в качестве η_1, η_2 выбирают квантили распределения статистики $\eta(X_1, \ldots, X_n, \theta)$ уровней $\alpha/2$ и $1 - \alpha/2$ соответственно. Напомним, что квантилью порядка β случайной величины ξ называется значение τ_β, для которого $P(\xi < \tau_\beta) = \beta$.

3) Разрешив неравенство $\eta_1 < \eta(X_1, \ldots, X_n, \theta) < \eta_2$ относительно θ, находим границы доверительного интервала.

Аналогично находится и асимптотический доверительный интервал, с той лишь разницей, что на первом этапе находим статистику $\eta(X_1, \ldots, X_n, \theta)$, закон распределения которой при $n \to \infty$ стремится к известному закону, не содержащему параметр θ.

37. Оценка параметров нормального распределения

37.1. Построение доверительного интервала для математического ожидания при известном σ

Применим общий принцип построения доверительных интервалов для математического ожидания нормальной величины при известном среднеквадратичном отклонении σ.

Пусть $\{X_1, \ldots, X_n\}$ — выборка, полученная из нормальной генеральной совокупности N_{m,σ^2} с известным среднеквадратичным отклонением σ. Требуется построить доверительный интервал для параметра m, соответствующий доверительной вероятности β.

Так как каждая из величин X_i распределена нормально, то и выборочное среднее

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

распределено нормально с параметрами $M(\bar{X}) = m$, $D(\bar{X}) = \sigma^2/n$. Тогда статистика $\eta = \bar{X} - m$ (отклонение выборочного среднего от математического ожидания) имеет также нормальное распределение с параметрами $M(\eta) = 0$, $D(\eta) = \sigma^2/n$. Поэтому вероятность любого отклонения $|\bar{X} - m|$ может быть вычислена по формуле

$$P(|\bar{X} - m| < \delta) = 2\Phi\Big(\frac{\delta}{\sigma/\sqrt{n}}\Big) = 2\Phi\Big(\sqrt{n}\frac{\delta}{\sigma}\Big) = 2\Phi(\tau),$$

где $\tau = \sqrt{n}\delta/\sigma$.

Задавая определенную доверительную вероятность $\beta = 1 - \alpha$, по таблицам можно определить значение τ_β, для которого $\Phi(\tau_\beta) = \beta/2$. Тогда, учитывая, что $\delta = \tau\sigma/\sqrt{n}$, получим

$$P\Big(|\bar{X} - m| < \tau_\beta\frac{\sigma}{\sqrt{n}}\Big) = \beta,$$

или

$$P\Big(-\tau_\beta\frac{\sigma}{\sqrt{n}} < \bar{X} - m < \tau_\beta\frac{\sigma}{\sqrt{n}}\Big) = \beta,$$

или

$$P\Big(\bar{X} - \tau_\beta\frac{\sigma}{\sqrt{n}} < m < \bar{X} + \tau_\beta\frac{\sigma}{\sqrt{n}}\Big) = \beta.$$

Таким образом, интервал

$$\Big]\bar{X} - \tau_\beta\frac{\sigma}{\sqrt{n}}; \bar{X} + \tau_\beta\frac{\sigma}{\sqrt{n}}\Big[, \tag{37.1}$$

где τ_β находится из условия $\Phi(\tau_\beta) = (1-\alpha)/2 = \beta/2$, является доверительным для параметра m, соответствующим доверительной вероятности $\beta = 1-\alpha$. Поскольку интервал симметричен относительно оценки $\bar{X}$, то можно утверждать, что $\delta = \tau_\beta(\sigma/\sqrt{n})$ есть точность точечной оценки $\bar{X}$ параметра m.

♦ Оценку

$$|\bar{X} - m| < \frac{\tau_\beta\sigma}{\sqrt{n}}$$

называют *классической*. Из формулы $\delta = \tau_\beta\sigma/\sqrt{n}$ следует, что

1) при возрастании n число δ уменьшается и, следовательно, точность оценки возрастает;

2) увеличение надёжности $1-\alpha = 2\Phi(\tau_\beta)$ приводит к увеличению τ_β и, следовательно, к возрастанию δ, т.е. увеличение надёжности классической оценки приводит к уменьшению точности.

Пример 37.1. Найти доверительный интервал для математического ожидания нормальной случайной величины ξ с надежностью $\beta = 0{,}9$, если $n = 16$, $\bar{X} = 20{,}9$, $\sigma = 2$.

Решение. По условию, ξ — нормальная случайная величина с известным σ. Требуется построить доверительный интервал для математического ожидания этой величины, то есть для параметра m. По таблицам функции Лапласа находим $\tau = 1{,}645$, для которого $\Phi(\tau) = 0{,}90/2 = 0{,}45$. На рис. 91 заштрихованной областью на графике плотности нормального распределения выделена площадь, численно равная β:

$$\beta = 2\Phi(\tau) = 2\int_0^\tau \varphi(x)dx = 2\frac{1}{\sqrt{2\pi}}\int_0^\tau \exp\Big(-\frac{x^2}{2}\Big)dx.$$

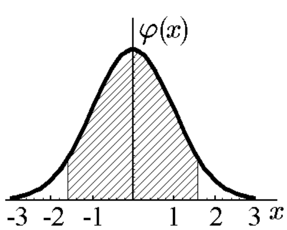


Рис. 91.

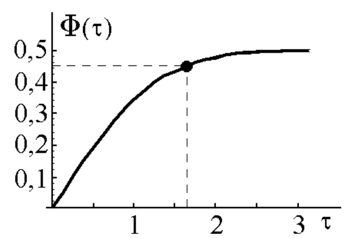


Рис. 92.

На рис. 92 на графике функции Лапласа выделена точка $(\tau, \Phi(\tau))$. Следовательно,

$$\delta = \frac{\tau\sigma}{\sqrt{n}} = \frac{1{,}645\cdot 2}{\sqrt{16}} = 0{,}8225.$$

Таким образом, с вероятностью 0,9 имеем $20{,}9 - 0{,}8225 < m < 20{,}9 + 0{,}8225$ или математическое ожидание m с вероятностью 0,9 лежит в интервале]20,0775, 21,7225[.

Если значение σ неизвестно, то с помощью статистики $\eta = \bar{X} - m$ невозможно построить точный доверительный интервал для параметра m нормальной случайной величины. Однако при больших n величину σ можно заменить состоятельной оценкой

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2} \quad \text{или} \quad \sqrt{\bar{D}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2},$$

построив статистику

$$\eta_1 = \frac{\bar{X} - m}{s/\sqrt{n}}.$$

Так как $s \xrightarrow{p} \sigma$, то

$$\eta_1 = \frac{\bar{X} - m}{s/\sqrt{n}} = \frac{\sigma}{s}\frac{\bar{X} - m}{\sigma/\sqrt{n}} \Rightarrow \eta_0 \in N_{0,1},$$

то есть статистику η_1 можно использовать для построения асимптотического доверительного интервала для параметра m. Тогда, если τ_β — корень уравнения $\Phi(\tau_\beta) = \beta/2$, то

$$P(-\tau_\beta < \eta_1 < \tau_\beta) \xrightarrow[n\to\infty]{} P(-\tau_\beta < \eta_0 < \tau_\beta) = \beta,$$

и искомый интервал имеет вид

$$\bar{X} - \frac{\tau_\beta s}{\sqrt{n}} < m < \bar{X} + \frac{\tau_\beta s}{\sqrt{n}}.$$

Статистика $\eta = (\bar{X} - m)/(\sqrt{\sigma^2/n})$ непригодна для построения доверительного интервала для σ нормальной случайной величины при известном параметре m, а тем более при неизвестном m. Действительно, разрешая неравенство относительно σ, мы получим

$$\sigma > \frac{\sqrt{n}|\bar{X} - m|}{\tau_\beta}$$

— бесконечный доверительный интервал.

37.2. Построение доверительного интервала для дисперсии σ^2 при известном математическом ожидании

Пусть случайная величина ξ распределена нормально, причём среднеквадратичное отклонение σ неизвестно. Требуется построить доверительный интервал для m.

Статистика

$$\eta = \frac{\bar{X} - m}{\sqrt{\sigma^2/n}}$$

непригодна для построения доверительного интервала для σ нормальной случайной величины при известном параметре m, а тем более при неизвестном m. Действительно, разрешая неравенство относительно σ, мы получим

$$\sigma > \frac{\sqrt{n}(\bar{X} - m)}{\eta_2}$$

(при условии $\eta_1 = -\eta_2$) — бесконечный доверительный интервал.

Рассмотрим другую статистику. Ранее мы отмечали, что распределение суммы квадратов независимых стандартных нормальных величин есть распределение хи-квадрат. Величины $(X_i - m)/\sigma$ распределены по нормальному закону $N_{0,1}$, поэтому статистика

$$\eta = \sum_{i=1}^{n} \left(\frac{X_i - m}{\sigma} \right)^2$$

будет иметь распределение хи-квадрат с n степенями свободы.

Если обозначить через

$$\bar{D}_0 = \frac{1}{n} \sum_{i=1}^{n} (X_i - m)^2$$

оценку дисперсии при известном m, то

$$\eta = \frac{n\bar{D}_0}{\sigma^2}.$$

Распределение χ^2 несимметрично, поэтому границы доверительного интервала выберем так, чтобы вероятности выхода параметра за их пределы была одинаковой и равной $\alpha/2$. Пусть $\tau_{\alpha/2}$, $\tau_{1-\alpha/2}$ — квантили распределения χ^2_n уровней $\alpha/2$ и $1-\alpha/2$ соответственно. Тогда

$$P(\tau_{\alpha/2} < \eta < \tau_{1-\alpha/2}) = 1 - \alpha = \beta$$

или

$$P\Big(\tau_{\alpha/2} < \frac{n\bar{D}_0}{\sigma^2} < \tau_{1-\alpha/2}\Big) = \beta,$$

откуда

$$P\Big(\frac{n\bar{D}_0}{\tau_{1-\alpha/2}} < \sigma^2 < \frac{n\bar{D}_0}{\tau_{\alpha/2}}\Big) = \beta.$$

Таким образом, искомый интервал с надежностью β имеет вид

$$\Big]\frac{n\bar{D}_0}{\tau_{1-\alpha/2}}; \frac{n\bar{D}_0}{\tau_{\alpha/2}}\Big[. \tag{37.2}$$

Итак, остается построить для нормального распределения следующие доверительные интервалы:

1. для параметра m при неизвестном σ;
2. для параметра σ^2 при неизвестном m.

37.3. Лемма Фишера

Чтобы построить доверительные интервалы для дисперсии σ^2 при неизвестном математическом ожидании, для параметров нормального распределения рассмотрим предварительно лемму Фишера.

Теорема 37.1 (об ортогональном преобразовании нормального вектора). *Пусть $\vec{\xi} = \{\xi_1, \xi_2, \ldots, \xi_n\}$ — случайный вектор, координаты которого независимы и имеют стандартное нормальное распределение, а $\vec{\eta} = C\vec{\xi}$, где C — ортогональная матрица порядка n (т.е. $CC^\mathsf{T} = E$). Тогда координаты $\eta_i = \sum\limits_{j=1}^{n} C_{ij}\xi_j$ вектора $\vec{\eta}$ независимы и имеют стандартное нормальное распределение.*

Доказательство. Запишем плотность распределения вектора $\vec{\xi} = \{\xi_1, \xi_2, \ldots, \xi_n\}$. Так как величины $\xi_1, \xi_2, \ldots, \xi_n$ независимы и имеют стандартное нормальное распределение, то

$$f_{\vec{\xi}}(\vec{x}) = f_{\vec{\xi}}(x_1, x_2, \ldots, x_n) = \prod_{j=1}^{n} f_{\xi_j}(x_j) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-x_j^2/2} =$$

$$= \frac{1}{(\sqrt{2\pi})^n} \exp\Big\{ -\frac{1}{2}\sum_{j=1}^{n} x_j^2 \Big\} = \frac{1}{(\sqrt{2\pi})^n} e^{-\|\vec{x}\|^2/2},$$

где

$$\|\vec{x}\|^2 = \vec{x}^{\mathsf{T}}\vec{x} = \sum_{j=1}^{n} x_j^2.$$

Чтобы записать плотность распределения вектора $\vec{\eta}$, воспользуемся формулой для плотности при линейном преобразовании вектора: если $\vec{\eta} = C\vec{\xi}$, то

$$f_{\vec{\eta}}(\vec{x}) = \frac{1}{|\det C|} f_{\vec{\xi}}(C^{-1}\vec{x}).$$

Тогда с учетом того, что $C^{-1} = C^{\mathsf{T}}$ и $\det C = 1$, получим

$$f_{\vec{\eta}}(\vec{x}) = f_{\vec{\xi}}(C^{\mathsf{T}}\vec{x}) = \frac{1}{(\sqrt{2\pi})^n} e^{-\|C^{\mathsf{T}}\vec{x}\|/2}.$$

Но умножение вектора на ортогональную матрицу не меняет нормы вектора: действительно, $\|C^{\mathsf{T}}\vec{x}\|^2 = (C^{\mathsf{T}}\vec{x})^{\mathsf{T}}(C^{\mathsf{T}}\vec{x}) = \vec{x}^{\mathsf{T}}CC^{\mathsf{T}}\vec{x} = \vec{x}^{\mathsf{T}}\vec{x} = \|\vec{x}\|^2$.

Следовательно,

$$f_{\vec{\eta}}(\vec{x}) = \frac{1}{(\sqrt{2\pi})^n} e^{-\|\vec{x}\|^2/2} = f_{\vec{\xi}}(\vec{x}),$$

т.е. величины η_j так же, как и величины ξ_j, независимы и имеют стандартное нормальное распределение.

Теорема 37.2 (лемма Фишера). *Пусть $\vec{\xi} = \{\xi_1, \xi_2, \ldots, \xi_n\}$ — выборка из $N_{0,1}$ и $\vec{\eta} = C\vec{\xi}$, где C — ортогональная матрица порядка n. Тогда для любого $k = \overline{1, n}$ статистика*

$$\eta(\vec{\xi}) = \sum_{i=1}^{n} \xi_i^2 - \eta_1^2 - \ldots - \eta_k^2$$

распределена по закону χ^2_{n-k} и не зависит от $\eta_1, \eta_2, \ldots, \eta_k$.

Доказательство. Так как $\vec{\eta} = C\vec{\xi}$, то

$$\sum_{i=1}^{n} \eta_j^2 = \sum_{i=1}^{n} \xi_j^2$$

(см. доказательство предыдущей теоремы). Тогда

$$\eta(\vec{\xi}) = \sum_{i=1}^{n} \xi_i^2 - \eta_1^2 - \ldots - \eta_k^2 = \sum_{i=1}^{n} \eta_i^2 - \eta_1^2 - \ldots - \eta_k^2 = \eta_{k+1}^2 + \ldots + \eta_n^2 \in \chi^2_{n-k},$$

что и требовалось доказать.

Следствие 37.2.1. Пусть $\xi_1, \xi_2, \ldots, \xi_n$ независимы и имеют распределение N_{a,σ^2},

$$\bar{\xi} = \frac{1}{n}\sum_{i=1}^{n}\xi_i, \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\xi_i - \bar{\xi})^2, \quad \bar{D}_0 = \frac{1}{n}\sum_{i=1}^{n}(\xi_i - a)^2.$$

Тогда

1) случайная величина

$$\eta = \frac{(n-1)s^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{\xi_i - \bar{\xi}}{\sigma}\right)^2 \in \chi^2_{n-1} \tag{37.3}$$

распределена по закону χ^2_{n-1}, т.е. $\eta \in \chi^2_{n-1}$;

2) случайные величины $\bar{\xi}$ и s^2 независимы;

3) случайная величина

$$\eta = \sqrt{n}\frac{\bar{\xi} - a}{\sqrt{s^2}} \in T_{n-1} \tag{37.4}$$

распределена по закону Стьюдента T_{n-1}, т.е. $\eta \in T_{n-1}$.

Доказательство. 1. Рассмотрим статистику

$$\eta(\vec{\xi}) = \frac{(n-1)s^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{\xi_i - \bar{\xi}}{\sigma}\right)^2.$$

Введем стандартные нормальные величины $z_i = (\xi_i - a)/\sigma$ и выразим $\eta(\vec{\xi})$ через z_i:

$$\eta(\vec{\xi}) = \sum_{i=1}^{n}\left(\frac{\xi_i - \bar{\xi}}{\sigma}\right)^2 = \sum_{i=1}^{n}\left(\frac{\xi_i - a}{\sigma} - \frac{\bar{\xi} - a}{\sigma}\right)^2 = \sum_{i=1}^{n}(z_i - \bar{z})^2,$$

где

$$\bar{z} = \frac{\bar{\xi} - a}{\sigma} = \frac{1}{n}\sum_{i=1}^{n} z_i,$$

то есть можно изначально считать, что величины ξ_i имеют стандартное нормальное распределение. Применим к $\eta(\vec{\xi})$ лемму Фишера. Представим $\eta(\vec{\xi})$ в виде

$$\eta(\vec{\xi}) = (n-1)s^2 = \sum_{i=1}^{n}(\xi_i - \bar{\xi})^2 = \sum_{i=1}^{n}(\xi_i)^2 - n\bar{\xi}^2 = \sum_{i=1}^{n}(\xi_i)^2 - \eta_1^2,$$

где $\eta_1 = \sqrt{n}\bar{\xi}$.

Покажем, что найдется ортогональная матрица C такая, что вектор $\vec{\eta} = C\vec{\xi}$ будет иметь координату $\eta_1 = \sqrt{n}\bar{\xi}$. Возьмем в качестве первой строки матрицы C строку

$$\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\right).$$

Тогда

$$\eta_1 = \frac{1}{\sqrt{n}}\xi_1 + \frac{1}{\sqrt{n}}\xi_2 + \ldots + \frac{1}{\sqrt{n}}\xi_n = \sqrt{n}\bar{\xi}.$$

Так как норма этой строки (длина вектора) равна 1, то эту строку всегда можно дополнить до ортогональной матрицы (строки и столбы ортогональной матрицы есть ортонормированные векторы). Тогда в соответствии с леммой Фишера статистика

$$\eta(\vec{\xi}) = (n-1)s^2 = \sum_{i=1}^{n}(\xi_i)^2 - \eta_1^2$$

имеет распределение хи-квадрат с $n-1$ степенью свободы.

2. В соответствии с леммой Фишера величина

$$\eta_1 = \frac{1}{\sqrt{n}}\bar{\xi}$$

и статистика

$$\eta(\vec{\xi}) = (n-1)s^2 = \sum_{i=1}^{n}(\xi_i)^2 - \eta_1^2$$

независимы, то есть s^2 и $\bar{\xi}$ независимы.

3. Преобразуем

$$\sqrt{n}\frac{\bar{\xi}-a}{\sqrt{s^2}} = \sqrt{n}\frac{\bar{\xi}-a}{\sigma}\frac{1}{\sqrt{\dfrac{(n-1)s^2}{\sigma^2}\Big/(n-1)}}.$$

Величина $\sqrt{n}(\bar{\xi}-a)/\sigma \in N_{0,1}$, а величина $[(n-1)s^2]/\sigma^2 \in \chi^2_{n-1}$, и по следствию 4 эти величины независимы. Следовательно,

$$\sqrt{n}\frac{\bar{\xi}-a}{\sqrt{s^2}} = \sqrt{n}\frac{\bar{\xi}-a}{\sigma}\frac{1}{\sqrt{\dfrac{(n-1)s^2}{\sigma^2}\Big/(n-1)}} \in T_{n-1}.$$

37.4. Доверительные интервалы для параметров нормального распределения

Перейдем к построению доверительных интервалов для параметров нормального распределения.

1. Для параметра m при известном σ (см. (37.1)).

С вероятностью $\beta = 1-\alpha$:

$$\bar{X} - \tau_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < m < \bar{X} + \tau_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},$$

где $\tau_{1-\alpha/2}$ — квантиль стандартного нормального распределения уровня $1-\alpha/2$.

2. Для параметра m при неизвестном σ.

Из следствия 3 из леммы Фишера, учитывая симметрию распределения Стьюдента, с вероятностью $\beta = 1-\alpha$ получим

$$\bar{X} - \tau_{1-\alpha/2}\frac{s}{\sqrt{n}} < m < \bar{X} + \tau_{1-\alpha/2}\frac{s}{\sqrt{n}}, \tag{37.5}$$

где $\tau_{1-\alpha/2}$ — квантиль распределения Стьюдента уровня $1-\alpha/2$.

◊ Квантиль уровня $1-\alpha/2$ распределения Стьюдента называется коэффициентом Стьюдента уровня β и обозначается $t_\beta = \tau_{1-\alpha/2}$.

3. Для параметра σ^2 при неизвестном m (см. (35.4)).

С вероятностью $\beta = 1 - \alpha$ из (37.2) получим

$$\frac{n\bar{D}_0}{\tau_{1-\alpha/2}} < \sigma^2 < \frac{n\bar{D}_0}{\tau_{\alpha/2}}, \tag{37.6}$$

где $\tau_{\alpha/2}$, $\tau_{1-\alpha/2}$ — квантили распределения хи-квадрат с n степенями свободы уровней $\alpha/2$ и $1 - \alpha/2$ соответственно.

4. Для параметра σ^2 при неизвестном m.

Из следствия 1 леммы Фишера с вероятностью $\beta = 1 - \alpha$ получим

$$\frac{(n-1)s^2}{\tau_{1-\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\tau_{\alpha/2}}, \tag{37.7}$$

где $\tau_{\alpha/2}$, $\tau_{1-\alpha/2}$ — квантили распределения хи-квадрат с $n-1$ степенью свободы уровней $\alpha/2$ и $1 - \alpha/2$ соответственно.

Пример 37.2. Найти доверительный интервал для математического ожидания нормальной случайной величины с надежностью $\beta = 0{,}95$, если $n = 50$, $\bar{X} = -0{,}155$, $s = 0{,}936$.

Решение. В данном случае при $n = 50$ и $\beta = 0{,}95$ по таблице распределения Стьюдента для $k = 49$ и $\alpha = 1 - \beta = 0{,}05$ находим $t_\beta = 2{,}009$ (см. также пример **??**). На рис. 93 заштрихованной областью на графике плотности распределения Стьюдента выделена площадь, численно равная β.

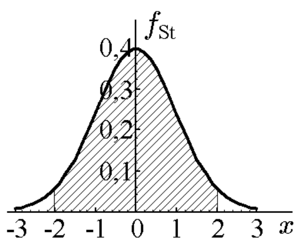


Рис. 93.

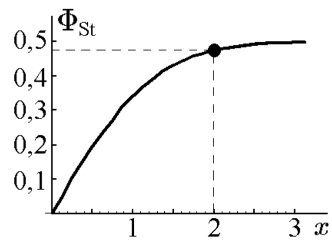


Рис. 94.

На рис. 94 на графике функции $\Phi_{\mathrm{St}}(t)$, аналогичной функции Лапласа, для распределения Стьюдента выделена точка $(t, \Phi_{\mathrm{St}}(t))$.

Вычисляем далее по формуле (31.2)

$$\frac{t_\beta s}{\sqrt{n}} = \frac{2{,}009 \cdot 0{,}936}{\sqrt{50}} \approx 0{,}266$$

и записываем доверительный интервал $]-0{,}155 - 0{,}266; -0{,}155 + 0{,}266[$. Таким образом, с вероятностью $\beta = 0{,}95$ справедливо неравенство

$$-0{,}421 < m < 0{,}111.$$

Пример 37.3. По данным выборки объема $n = 20$ была найдена выборочная несмещенная дисперсия $s^2 = 0{,}876$; найти с надежностью $\beta = 0{,}90$ доверительный интервал для параметра σ^2 при неизвестном m.

Решение. В данном случае $n = 20$ и $\alpha = 1 - \beta = 1 - 0{,}90 = 0{,}10$, следовательно $k = n - 1 = 19$, $\alpha_1 = \alpha/2 = 0{,}05$ и $\alpha_2 = 1 - \alpha/2 = 0{,}95$; с помощью таблиц χ^2-распределения по $k = 19$ и $\alpha_1 = 0{,}05$ находим $\tau_2 = 30{,}1$, а по $k = 19$ и $\alpha_2 = 0{,}95$ находим $\tau_1 = 10{,}1$ (см. также пример **??**). На рис. 95 заштрихованной областью на графике плотности χ^2-распределения выделена площадь, численно равная β, согласно формуле

$$\beta = F_\chi(\tau_2) - F_\chi(\tau_1) = \int\limits_{\tau_1}^{\tau_2} f_\chi(x)dx.$$

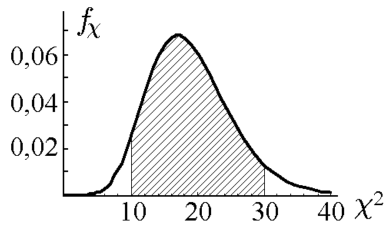


Рис. 95.

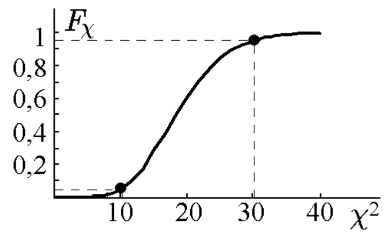


Рис. 96.

На рис. 96 на графике функции χ^2-распределения $F_\chi(\chi^2)$ выделены точки $(\tau_1,\, F_\chi(\tau_1))$ и $(\tau_2,\, F_\chi(\tau_2))$.

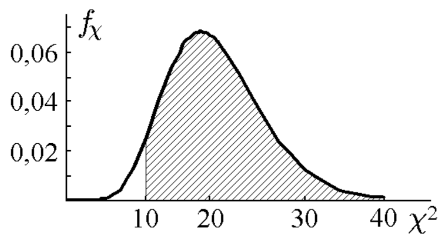


Рис. 97.

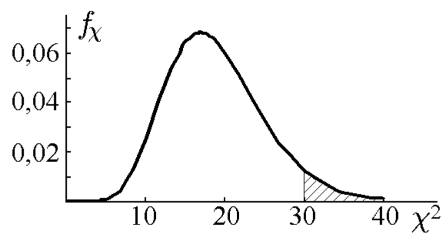


Рис. 98.

На рис. 97 и 98 заштрихованными областями на графике плотности χ^2-распределения выделены площади, численно равные $\alpha_1 = 0{,}95$ и $\alpha_2 = 0{,}05$ соответственно, согласно формулам

$$P(\chi^2 > \tau_2) = \frac{\alpha}{2} = \alpha_2, \quad P(\chi^2 > \tau_1) = 1 - \frac{\alpha}{2} = \alpha_1.$$

Наконец, по (37.7) определяем границы доверительного интервала:

$$\frac{(n-1)s^2}{\tau_2} = \frac{19 \cdot 0{,}876}{30{,}1} \approx 0{,}553; \quad \frac{(n-1)s^2}{\tau_1} = \frac{19 \cdot 0{,}876}{10{,}1} \approx 1{,}648.$$

38. Доверительный интервал для параметров показательного распределения и распределения Пуассона

Применим общий принцип построения доверительных интервалов в случае асимптотического доверительного интервала для параметра λ распределения Пуассона.

Пусть $\{X_1, X_2, \ldots, X_n\}$ — выборка, полученная из генеральной совокупности случайной величины ξ, распределенной по закону Пуассона Π_λ с неизвестным параметром λ. Требуется построить доверительный интервал для параметра λ, соответствующий доверительной вероятности β.

Рассмотрим статистику

$$\eta_1 = \frac{\bar{\xi} - M(\xi)}{\sigma(\xi)/\sqrt{n}} = \sqrt{n}\frac{\bar{\xi} - \lambda}{\sqrt{\lambda}}.$$

В соответствии с центральной предельной теоремой, $\eta_1 \Rightarrow \eta_0 \in N_{0,1}$ при $n \to \infty$. Пусть $\tau_{1-\alpha/2}$ — квантиль распределения $N_{0,1}$ уровня $1 - \alpha/2$ $(\alpha = 1 - \beta)$, тогда

$$P(-\tau_{1-\alpha/2} < \eta_1 < \tau_{1-\alpha/2}) = P\Big(-\tau_{1-\alpha/2} < \sqrt{n}\frac{\bar{\xi} - \lambda}{\sqrt{\lambda}} < \tau_{1-\alpha/2}\Big) \xrightarrow[n\to\infty]{} \xrightarrow[n\to\infty]{} P(-\tau_{1-\alpha/2} < \eta_0 < \tau_{1-\alpha/2}) = \beta.$$

Однако разрешить неравенство относительно λ не просто из-за корня в знаменателе. Попробуем заменить в знаменателе λ на состоятельную оценку этого параметра $\lambda^* = \bar{\xi}$, построив статистику $\eta_2 = \sqrt{n}(\bar{\xi} - \lambda/\sqrt{\bar{\xi}}$. Не изменится ли при этом характер сходимости? Вспомним свойство сходимости по распределению: если $\xi_n \Rightarrow \xi$, а $\eta_n \xrightarrow{p} c$, то $\eta_n \xi_n \Rightarrow c\xi$. Тогда

$$\eta_2 = \sqrt{n}\frac{\bar{\xi} - \lambda}{\sqrt{\bar{\xi}}} = \sqrt{\frac{\lambda}{\bar{\xi}}}\sqrt{n}\frac{\bar{\xi} - \lambda}{\sqrt{\lambda}} \Rightarrow \eta \in N_{0,1},$$

так как $\sqrt{\lambda/\bar{\xi}} \xrightarrow{p} 1$. Следовательно,

$$P\Big(-\tau_{1-\alpha/2} < \sqrt{n}\frac{\bar{\xi} - \lambda}{\sqrt{\bar{\xi}}} < \tau_{1-\alpha/2}\Big) \xrightarrow[n\to\infty]{} \beta,$$

или

$$P\Big(\bar{\xi} - \tau_{1-\alpha/2}\frac{\sqrt{\bar{\xi}}}{\sqrt{n}} < \lambda < \bar{\xi} + \frac{\sqrt{\bar{\xi}}}{\sqrt{n}}\tau_{1-\alpha/2}\Big) \xrightarrow[n\to\infty]{} \beta.$$

Таким образом, искомый асимптотический доверительный интервал уровня β имеет вид

$$\bar{\xi} - \tau_{1-\alpha/2}\frac{\sqrt{\bar{\xi}}}{\sqrt{n}} < \lambda < \bar{\xi} + \frac{\sqrt{\bar{\xi}}}{\sqrt{n}}\tau_{1-\alpha/2}. \tag{38.1}$$

Аналогичным образом применим общий принцип построения доверительных интервалов в случае асимптотического доверительного интервала для параметра α показательного распределения.

Пусть $\{X_1, X_2, \ldots, X_n\}$ — выборка, полученная из генеральной совокупности случайной величины ξ, распределенной по показательному закону E_α с неизвестным параметром α. Требуется построить доверительный интервал для параметра α, соответствующий доверительной вероятности β.

Рассмотрим статистику

$$\eta = \frac{\bar{X} - M(\xi)}{\sigma(\xi)/\sqrt{n}} = \sqrt{n}\frac{\bar{X} - 1/\alpha}{1/\alpha} = \sqrt{n}(\alpha\bar{X} - 1).$$

В соответствии с центральной предельной теоремой, $\eta \Rightarrow \eta_0 \in N_{0,1}$ при $n \to \infty$. Пусть $\tau_{1-\alpha/2}$ — квантиль распределения $N_{0,1}$ уровня $1 - \alpha/2$ $(\alpha = 1 - \beta)$, тогда

$$P(-\tau_{1-\alpha/2} < \eta < \tau_{1-\alpha/2}) = P(-\tau_{1-\alpha/2} < \sqrt{n}(\alpha\bar{X} - 1) < \tau_{1-\alpha/2}) \xrightarrow[n\to\infty]{} \xrightarrow[n\to\infty]{} P(-\tau_{1-\alpha/2} < \eta_0 < \tau_{1-\alpha/2}) = \beta$$

или

$$P\Big(\frac{1}{\bar{X}} - \frac{\tau_{1-\alpha/2}}{\sqrt{n}\bar{X}} < \alpha < \frac{1}{\bar{X}} + \frac{\tau_{1-\alpha/2}}{\sqrt{n}\bar{X}}\Big) \xrightarrow[n\to\infty]{} \beta.$$

Таким образом, искомый асимптотический доверительный интервал уровня β имеет вид

$$\frac{1}{\bar{X}} - \frac{\tau_{1-\alpha/2}}{\sqrt{n}\bar{X}} < \alpha < \frac{1}{\bar{X}} + \frac{\tau_{1-\alpha/2}}{\sqrt{n}\bar{X}}. \tag{38.2}$$

39. Построение доверительной области для вектора средних многомерной нормальной случайной величины

Пусть по результатам n наблюдений из генеральной совокупности k-мерной нормальной случайной величины ξ найден вектор выборочных средних $\bar{X}$. Требуется найти с надежностью β доверительную область для компонент вектора генеральных средних $\mu = \{\mu_1, \mu_2, \ldots, \mu_k\}$.

Пусть ковариационная матрица A случайной величины ξ известна.

Для одномерной нормальной случайной величины, для построения доверительного интервала для среднего m при известном σ использовалась статистика

$$\eta = \sqrt{n}\frac{\bar{X} - m}{\sigma}, \quad \eta \in N_{0,1}.$$

Если возвести η в квадрат, получим статистику, распределенную по закону хи-квадрат с одной степенью свободы:

$$\eta^2 = n\frac{(\bar{X} - m)}{\sigma^2} = n(\bar{X} - m)^\intercal(\sigma^2)^{-1}(\bar{X} - m), \quad \eta^2 \in \chi^2_1.$$

Обобщим на многомерный случай, построим статистику

$$\eta^2 = n(\bar{X} - \mu)^\intercal A^{-1}(\bar{X} - \mu), \tag{39.1}$$

где $\bar{X}$ — вектор выборочных средних многомерной случайной величины ξ; A — матрица ковариаций. Статистика η^2 в случае k-мерного нормального распределения генеральной совокупности будет иметь распределение хи-квадрат с k-степенями свободы. Действительно, случайную величину ξ можно получить путем линейного преобразования многомерной нормальной случайной величины ζ, компоненты которой независимы и имеют стандартное нормальное распределение: $\xi = B\zeta + \mu$, где невырожденная матрица B связана с матрицей ковариаций случайной величины ξ соотношением $BB^\intercal = A$. Но тогда $(\zeta - \mu)^\intercal A^{-1}(\zeta - \mu) = \zeta^\intercal\zeta$ есть сумма квадратов стандартных нормальных случайных величин. Учитывая, что вектор выборочных средних $\bar{X}$ имеет нормальное распределение с параметрами μ и A/n, получим, что статистика η^2 имеет распределение хи-квадрат с k степенями свободы.

Пусть χ^2_α — критическая точка распределения хи-квадрат с k степенями свободы уровня α ($\alpha = 1 - \beta$). Уравнение

$$n(\bar{X} - \mu)^\intercal A^{-1}(\bar{X} - \mu) = \chi^2_\alpha$$

определяет k-мерный эллипсоид в системе координат $\{\mu_1, \mu_2, \ldots, \mu_k\}$ с центром $\bar{X}$, поскольку левая часть уравнения есть положительно определенная квадратичная форма. Соответственно, доверительной областью уровня β будет область, ограниченная этим эллипсоидом, т.е. область, определяемая неравенством

$$n(\bar{X} - \mu)^\intercal A^{-1}(\bar{X} - \mu) \leqslant \chi^2_\alpha.$$

Пусть теперь ковариационная матрица A случайной величины ξ неизвестна. Для одномерной случайной величины, для построения доверительного интервала для среднего μ в этом случае использовалась статистика

$$t = \sqrt{n}\frac{\bar{X} - \mu}{s}, \quad t \in T_{n-1}.$$

Возведя в квадрат, получим статистику

$$t^2 = n(\bar{X} - \mu)^\intercal (s^2)^{-1} (\bar{X} - \mu).$$

Статистика t^2 будет иметь распределение Фишера с $\nu_1 = 1$, $\nu_2 = n-1$ степенями свободы ($t^2 \in F_{1,n-1}$). Заметим, что в этом случае распределение Фишера совпадает с распределением Хотеллинга, которое связано с распределением Фишера соотношением $T^2_{\nu_1,\nu_2} = \frac{\nu_1(\nu_1+\nu_2-1)}{\nu_2} F_{\nu_1,\nu_2}$, где $T^2_{\nu_1,\nu_2}$ — случайная величина, имеющая распределение Хотеллинга с ν_1, ν_2 степенями свободы, а F_{ν_1,ν_2} — случайная величина, имеющая распределение Фишера с ν_1, ν_2 степенями свободы.

Заменим s^2 на S, построим статистику

$$T^2 = n(\bar{X} - \mu)^\intercal S^{-1} (\bar{X} - \mu), \tag{39.2}$$

где $S = \frac{1}{n-1} X^\intercal X$ — несмещенная оценка матрицы ковариаций (X — матрица выборочных данных размера $n \times k$, а каждый столбец матрицы X есть выборочные значения соответствующей координаты вектора ξ). Статистика T^2 в случае k-мерного нормального распределения генеральной совокупности будет иметь распределение Хотеллинга с $\nu_1 = k$, $\nu_2 = n - k$ степенями свободы.

Пусть T^2_α — критическая точка распределения Хотеллинга с $\nu_1 = k$, $\nu_2 = n-k$ степенями свободы. Тогда доверительная область уровня β будет определяться неравенством

$$n(\bar{X} - \mu)^\intercal S^{-1} (\bar{X} - \mu) \leqslant T^2_\alpha, \quad T^2_\alpha = \frac{k(n-1)}{n-k} F_\alpha.$$

Если требуется построить доверительную область лишь для части компонент (для l компонент) вектора μ, можно использовать частичную статистику Хотеллинга. Пусть C — матрица размерами $k \times l$, причем строки матрицы C, соответствующие компонентам, для которых строится доверительная область, являются строками единичной матрицы порядка l, остальные строки нулевые. В соответствии с теоремами о линейном преобразовании нормального случайного вектора, l-мерный вектор $C^\intercal$ имеет нормальное распределение $N(C^\intercal \mu, C^\intercal A C)$. Тогда статистика Хотеллинга вида

$$T^2 = n(C^\intercal(\bar{X} - \mu))^\intercal (C^\intercal S C)^{-1} C^\intercal (\bar{X} - \mu) \tag{39.3}$$

будет иметь распределение Хотеллинга с $\nu_1 = l$, $\nu_2 = n - l$ степенями свободы.

Пример 39.1. По данным выборки объема $n = 10$ двумерной нормальной случайной величины ξ были найдены вектор выборочных средних $\bar{X} = \begin{pmatrix} 11{,}69 \\ 5{,}33 \end{pmatrix}$ и несмещенная выборочная матрица ковариаций $S = \begin{pmatrix} 1{,}008 & -0{,}072 \\ -0{,}072 & 0{,}049 \end{pmatrix}$. Найти с надежностью $\beta = 0{,}95$ доверительную область для вектора средних $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ случайной величины ξ.

Решение. Найдем матрицу

$$S^{-1} = \begin{pmatrix} 1{,}108 & 1{,}629 \\ 1{,}629 & 22{,}80 \end{pmatrix},$$

обратную к S, и выпишем квадратичную форму

$$(\bar{X}-\mu)^{\intercal}S^{-1}(\bar{X}-\mu) = 1{,}108\mu_1^2+3{,}25\mu_1\mu_2+22{,}8\mu_2^2-43{,}28\mu_1-281{,}14\mu_2+1002{,}19.$$

Из таблиц распределения Фишера для $\alpha = 0{,}05$, $\nu_1 = k = 2$ и $\nu_2 = n-k = 8$ найдем критическую точку $F_\alpha = 4{,}46$ и вычислим значение

$$\frac{1}{n}T_\alpha^2 = \frac{1}{n}\frac{k(n-1)}{n-k}F_\alpha = 1{,}003.$$

Следовательно, доверительная область будет определяться неравенством

$$1{,}108\mu_1^2+3{,}25\mu_1\mu_2+22{,}8\mu_2^2-43{,}28\mu_1-281{,}14\mu_2+1002{,}19 \leqslant 1{,}003$$

(рис. ???).

ГЛАВА 4

Проверка статистических гипотез

40. Статистические гипотезы и критерии

♦ Любое предположение H о распределении выборочных наблюдений называется *статистической гипотезой* H.

Процедура обоснованного сопоставления высказанной гипотезы с выборочными данными называется статистической проверкой гипотезы. Поскольку по выборке ограниченного объема невозможно сделать безошибочные выводы, всегда существует вероятность принятия неверной гипотезы.

Гипотеза называется *простой*, если она однозначно определяет распределение, в противном случае гипотеза называется *сложной*. Если гипотез всего две, то одна из них называется *основной* (обозначается H_0), гипотеза же, принимаемая при отклонении H_0, называется *альтернативной* (обозначается H_1).

Примеры «классических» статистических гипотез:

1. $H_0 : F = F_1$ (простая); $H_1 : F \neq F_1$ (сложная).

2. $H_1 : F = F_1$; $H_2 : F = F_2$; $\ldots$; $H_k : F = F_k$ (все гипотезы простые).

3. $H_0 : F = \{F_1\}$ (сложная — гипотеза о виде закона распределения); $H_1 : F \neq \{F_1\}$ (сложная).

4. $H_0 : F = F_\theta$, $\theta = \theta_0$ (простая); $H_1 : F = F_\theta$, $\theta \neq \theta_0$ (сложная).

5. Гипотеза однородности. Заданы несколько выборок: $\vec{X}_1$, $\vec{X}_2$, $\ldots$, $\vec{X}_k$. $H_0 : F_1 = F_2 = \ldots = F_k$ (сложная); $H_1 : H_0$ — неверна (сложная).

♦ Правило $\delta(\vec{X})$, по которому на основе выборочных значений принимается одна из гипотез $\{H_1, H_2, \ldots, H_3\}$, называется *статистическим критерием* $\delta(\vec{X})$.

Таким образом, статистический критерий $\delta(\vec{X})$ есть отображение

$$\delta : \{X_1, X_2, \ldots, X_n\} \to \{H_1, H_2, \ldots, H_k\}. \tag{40.1}$$

◊ Статистический критерий не отвечает на вопрос — верна или нет проверяемая гипотеза. Он лишь решает, противоречат или нет выдвинутой гипотезе выборочные данные.

◊ Если есть одна основная гипотеза, а все остальные — нежелательные отклонения от нее, то вывод «данные противоречат гипотезе» всегда весомее, чем «данные не противоречат гипотезе».

Пусть имеется всего две гипотезы: основная H_0 и альтернативная H_1. Каков бы не был критерий $\delta(\vec{X})$, он принимает не более двух значений, то есть все пространство $\mathbb{R}^n$ выборочных значений делится на две части S и $\mathbb{R}^n \setminus S$, таких, что

$$\delta(\vec{X}) = \begin{cases} H_0, & \text{если } \vec{X} \in \mathbb{R}^n \setminus S; \\ H_1, & \text{если } \vec{X} \in S. \end{cases}$$

♦ Область $\mathbb{R}^n \setminus S$ называется *областью принятия гипотезы*, а область S — *критической областью*.

♦ Для критерия $\delta : \vec{X} \to \{H_0, H_1\}$ говорят, что произошла *ошибка 1-го рода*, если гипотеза H_0 отвергнута критерием, в то время как она верна. Если же гипотеза H_0 принята критерием, в то время как была верной гипотеза H_1, то говорят, что произошла *ошибка 2-го рода*.

♦ Вероятность ошибки первого рода $\alpha(\delta)$ называется *уровнем значимости критерия*, а величина $1 - \beta$, где $\beta(\delta)$ — вероятность ошибки второго рода, — *мощностью критерия*.

Заметим, что вероятности ошибок первого и второго рода α и β вычисляются при разных предположениях о распределении генеральной совокупности, а следовательно, никакой прямой связи (вроде $\alpha = 1 - \beta$) между ними нет. Хотя, как правило, естественное стремление уменьшить одну из ошибок, построив соответствующий критерий, приводит к увеличению другой ошибки.

♦ Критерий $\delta : \vec{X} \to \{H_0, H_1\}$ для проверки гипотезы H_0 против альтернативы H_1 называется *состоятельным*, если при $n \to \infty$ $\beta(\delta) \to 0$.

♢ В случае сложной альтернативы H_1 предполагается, $\beta(\delta) \to 0$ для любой простой альтернативы H_1.

Заметим, что так же, как в случае статистического оценивания параметров распределения, в задаче проверки гипотез возникают следующие вопросы:

1) как построить на основе выборочных данных критерий?

2) если существует несколько критериев, какой из них предпочесть?

Ограничимся для простоты случаем двух гипотез.

Как правило, ошибки первого и второго рода неравноправны. Например, H_0 может представлять утверждение, которое при отсутствии явно противоречащих данных желательно бы считать справедливым. В этом случае при выборе критерия фиксируется уровень ошибки первого рода α и рассматриваются критерии, обладающие такой же или еще меньшей вероятностью ошибки первого рода. Очевидно, что среди этих критериев следует предпочесть тот, который обладает большей мощностью.

♦ Критерий δ_0 называют *наиболее мощным критерием* (НМК) уровня α, если $\beta(\delta_0) \leqslant \beta(\delta)$, для любого другого критерия уровня α.

Если ошибки первого и второго рода равноправны, можно использовать для сравнения критериев минимаксный подход.

♦ Говорят, что критерий δ_1 не хуже δ_2 в смысле минимаксного подхода, если $\max\{\alpha(\delta_1), \beta(\delta_1)\} \leqslant \max\{\alpha(\delta_2), \beta(\delta_2)\}$.

♦ Критерий δ_0 называют *минимаксным критерием*, если он не хуже других критериев в смысле минимаксного подхода.

Если априори известно, что с вероятностью p справедлива гипотеза H_0, а с вероятностью $q = 1 - p$ гипотеза H_1, то можно использовать байесовский подход, по которому выбирается критерий, обладающей наименьшей полной вероятностью ошибки.

♦ Говорят, что критерий δ_1 не хуже δ_2 в смысле байесовского подхода, если $p\alpha(\delta_1) + q\beta(\delta_1) \leqslant p\alpha(\delta_2) + q\beta(\delta_2)$.

♦ Критерий δ_0 называют *байесовским критерием*, если он не хуже других критериев в смысле байесовского подхода.

Рассмотрим далее принципы построения критериев на примере критериев максимального правдоподобия и согласия. Начнем с критерия максимального правдоподобия.

Вспомним, что функция правдоподобия

$$\Psi(\vec{\xi}) = \begin{cases} \prod\limits_{i=1}^{n} f(X_i), & \text{если } \xi \text{ — непрерывная случайная величина;} \\ \prod\limits_{i=1}^{n} P(X_i), & \text{если } \xi \text{ — дискретная случайная величина,} \end{cases}$$

есть фактически вероятность получить данную выборку (для непрерывной случайной величины с точностью до постоянного множителя). Если у нас имеются две гипотезы H_0 и H_1 о распределении случайной величины ξ, то в предположении истинности каждой из них можно получить две функции правдоподобия:

$\Psi_{H_0}(\vec{X})$ и $\Psi_{H_1}(\vec{X})$. Тогда можно, очевидно, построить следующий критерий:

$$\delta(\vec{X}) = \begin{cases} H_0, & \text{если } \Psi_{H_0}(\vec{X}) > \Psi_{H_1}(\vec{X}); \\ H_1, & \text{если } \Psi_{H_0}(\vec{X}) < \Psi_{H_1}(\vec{X}). \end{cases}$$

Однако такой критерий обладает очевидными недостатками:

1) он имеет фиксированный уровень ошибок первого и второго рода;

2) не решает задачу, если $\Psi_{H_0}(\vec{X}) = \Psi_{H_1}(\vec{X})$ на множестве, мера которого отлична от нуля.

Поэтому такой критерий требует некоторой модификации.

♦ Отношение

$$\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})}$$

называется *отношением правдоподобия.*

♦ *Критерием максимального правдоподобия* называется критерий

$$\delta(\vec{\xi}) = \begin{cases} H_0, \text{ если } \dfrac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} < c; \\ H_1, \text{ если } \dfrac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} > c; \\ H_1 \text{ с вероятностью } p,\ H_0 \text{ с } 1-p, \text{ если } \dfrac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} = c. \end{cases} \tag{40.2}$$

Уровень *значимости* критерия правдоподобия:

$$\alpha(c,p) = P_{H_0}(\delta(\vec{X}) = H_1) = P_{H_0}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} > c\Big) + pP_{H_0}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} = c\Big). \tag{40.3}$$

Вероятность ошибки второго рода:

$$\beta(c,p) = P_{H_1}(\delta(\vec{X}) = H_0) = P_{H_1}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} < c\Big) + (1-p)P_{H_1}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} = c\Big). \tag{40.4}$$

♢ Заметим, что построенный таким образом критерий отличается от критерия (40.1): покажем, что функция (40.1) является случайной функцией. Критерии подобного вида называются *рандомизированными.* (Строго говоря, рандомизированный критерий есть отображение $\delta : \vec{X} \to p$, где $p \in [0,1]$ — вероятность принять гипотезу H_1.)

♢ Если отношение правдоподобия

$$\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})}$$

имеет непрерывное распределение, то величина

$$P_{H_0}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} = c\Big) = 0,$$

и критерий правдоподобия принимает вид

$$\delta(\vec{X}) = \begin{cases} H_0, & \text{если } \dfrac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} < c; \\ H_1, & \text{если } \dfrac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} \geqslant c. \end{cases} \tag{40.5}$$

Лемма Неймана–Пирсона

1. Если постоянные c и p выбраны так, что $\alpha(c,p) = \beta(c,p)$, то критерий правдоподобия является минимаксным критерием.

2. Если постоянные c и p выбраны так, что $\alpha(c,p) = \alpha_0$, то критерий правдоподобия является наиболее мощным критерием уровня α_0.

Примем лемму без доказательства, покажем лишь, что для любого $\alpha_0 \in [0,1]$ уравнение (40.3) разрешимо относительно p и c.

Рассмотрим невозрастающую функцию

$$\varphi(c) = P_{H_0}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} > c\Big),$$

где $c \in [0,\infty)$. Заметим, что при $c \to \infty$:

$$\varphi(c) = P_{H_0}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} > c\Big) =$$

$$= \int\limits_{\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} > c} \Psi_{H_0}(\vec{x})d(\vec{x}) < \frac{1}{c} \int\limits_{\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} > c} \Psi_{H_1}(\vec{x})d(\vec{x}) = \frac{1}{c} P_{H_1}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} > c\Big) \xrightarrow[c\to\infty]{} 0.$$

Найдем

$$\varphi(0) = P_{H_0}\Big(\frac{\Psi_1(\vec{\xi})}{\Psi_0(\vec{\xi})} > 0\Big) = P_{H_0}(\Psi_1(\vec{\xi}) > 0).$$

Пусть $\varphi(0) \geqslant \alpha_0$ (если это не так, то полагаем $c = 0$). Имеем $\varphi(0) \geqslant \alpha_0$, $\varphi(c)$ не возрастает и стремится к нулю с ростом c. Следовательно, найдется такая точка c, что $\varphi(c) = \alpha_0$, либо такая, что $\varphi(c-0) > \alpha_0$, $\varphi(c) < \alpha_0$, так как $\varphi(c)$ непрерывна в каждой точке справа (сравните с функцией распределения, которая непрерывна в каждой точке слева). В первом случае полагаем $p = 0$, а во втором случае положим

$$p = \frac{\alpha_0 - \varphi(c)}{\varphi(c-0) - \varphi(c)}.$$

Так как, очевидно,

$$\varphi(c-0) - \varphi(c) = P_{H_0}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} = c\Big),$$

то получим

$$\alpha(p,c) = P_{H_0}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} > c\Big) + pP_{H_0}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} = c\Big) =$$

$$= \varphi(c) + \frac{\alpha_0 - \varphi(c)}{\varphi(c-0) - \varphi(c)}[\varphi(c-0) - \varphi(c)] = \alpha_0.$$

Пример 40.1. Имеется выборка $\{\xi_1\}$ объема $n = 1$ и две гипотезы $H_0 : X_1 \in U_{1,5}$, $H_1 : X_1 \in U_{0,2}$. Требуется построить наиболее мощный критерий уровня $\alpha_0 = 1/5$.

Решение. Запишем функции правдоподобия и отношение правдоподобия:

$$\Psi_{H_0}(X_1) = \begin{cases} 1/4, & X_1 \in [1,5]; \\ 0, & X_1 \notin [1,5]; \end{cases} \quad \Psi_{H_1}(X_1) = \begin{cases} 1/2, & X_1 \in [0,2]; \\ 0, & X_1 \notin [0,2]; \end{cases}$$

$$\frac{\Psi_{H_1}(X_1)}{\Psi_{H_0}(X_1)} = \begin{cases} \infty, & X_1 \in [0,1[; \\ 2, & X_1 \in [1,2[; \\ 0, & X_1 \in]2,5[. \end{cases}$$

Так как отношение правдоподобия постоянно на каждом из трех интервалов, для постоянной c существуют лишь несколько возможностей, при которых критерии будут различаться: $c = 0$, $0 < c < 2$, $c = 2$, $2 < c < \infty$, $c = \infty$. Исследуем поведение функции

$$\varphi(c) = P_{H_0}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} > c\Big),$$

чтобы определить необходимое значение c.

1. $0 \leqslant c < 2$. В этом случае $\Psi_{H_1}(X_1)/\Psi_{H_0}(X_1) > c$ при $X_1 \in [0,2]$ и

$$\varphi(c) = P_{H_0}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} > c\Big) = P_{H_0}(X_1 \in [0,2]) = P_{H_0}(X_1 \in [1,2]) = \frac{1}{4}.$$

2. $2 \leqslant c < \infty$. В этом случае $\Psi_{H_1}(X_1)/\Psi_{H_0}(X_1) > c$ при $X_1 \in [0,1[$ и

$$\phi(c) = P_{H_0}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} > c\Big) = P_{H_0}(X_1 \in [0,1[) = 0.$$

Следовательно, $c = 2$, и остается определить вероятность p так, чтобы $\alpha(c,p) = 1/5$. Так как

$$P_{H_0}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} = 2\Big) = P_{H_0}(X_1 \in [1,2]) = \frac{1}{4},$$

то из уравнения (40.3) получим

$$0 + p\frac{1}{4} = \frac{1}{5}$$

или $p = 4/5$. Таким образом, критерий заданного уровня $\alpha = 1/5$ будет иметь вид

$$\delta(X_1) = \begin{cases} H_0, & \text{если } X_1 \in]2,5]; \\ H_1, & \text{если } X_1 \in [0,1[; \\ H_0 \text{ с вероятностью } 1/5,\ H_1 \text{ с вероятностью } 4/5, & \text{если } X_1 \in [1,2]. \end{cases}$$

Вероятность ошибки второго рода при этом равна

$$\beta = P_{H_1}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} < 2\Big) + \frac{1}{5}P_{H_1}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} = c\Big) =$$
$$= P_{H_1}(X_1 \in]2,5]) + \frac{1}{5}P_{H_1}(X_1 \in [1,2]) = 0 + \frac{1}{5}\cdot\frac{1}{2} = \frac{1}{10}.$$

Соответственно мощность критерия $1 - \beta = 9/10$.

Пример 40.2. Имеется выборка $\{X_1, X_2, \ldots, X_n\}$ из нормальной совокупности с неизвестным параметром m и дисперсией, равной единице. Требуется построить наиболее мощный критерий уровня α_0 для проверки гипотезы $H_0 : m = m_0$ против альтернативы $H_1 : m = m_1$ $(m_0 < m_1)$.

Решение. Запишем функции правдоподобия и отношение правдоподобия:

$$\Psi_{H_0}(\vec{X}) = \Big(\frac{1}{\sqrt{2\pi}}\Big)^n \exp\Big(-\frac{1}{2}\sum_{i=1}^{n}(X_i - m_0)^2\Big),$$

$$\Psi_{H_1}(\vec{X}) = \Big(\frac{1}{\sqrt{2\pi}}\Big)^n \exp\Big(-\frac{1}{2}\sum_{i=1}^{n}(X_i - m_1)^2\Big),$$

$$\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} = \exp\Big(\frac{1}{2}\sum_{i=1}^{n}[2X_i(m_1 - m_0) + m_0^2 - m_1^2]\Big)$$

или

$$\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} = \exp(\mu\bar{X} + \nu),$$

где $\mu = n(m_1 - m_0) > 0$, $\nu = n(m_0^2 - m_1^2)/2$. Заметим, что условие

$$\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} > c$$

равносильно условию $\bar{X} > c_1$. А так как величина $\bar{X}$ имеет непрерывное распределение, то наиболее мощный критерий уровня α будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & \text{если } \bar{X} < c_1; \\ H_1, & \text{если } \bar{X} > c_1, \end{cases}$$

где c_1 есть корень уравнения $\alpha(c) = P_{H_0}(\bar{X} > c_1) = \alpha_0$.

При условии истинности H_0 $\bar{X} \in N_{a_0, 1/n}$,

$$\frac{\bar{X} - m_0}{1/\sqrt{n}} \in N_{0,1}.$$

Тогда с вероятностью α_0 справедливо

$$\frac{\bar{X} - m_0}{1/\sqrt{n}} > \tau_{1-\alpha_0}$$

или $\bar{X} > m_0 + \tau_{1-\alpha_0}/\sqrt{n}$, где $\tau_{1-\alpha_0}$ — квантиль распределения $N_{0,1}$ уровня $1 - \alpha_0$. Сравнивая, находим $c_1 = m_0 + \tau_{1-\alpha_0}/\sqrt{n}$.

Таким образом, наиболее мощный критерий будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & \text{если } \bar{X} < c_1; \\ H_1, & \text{если } \bar{X} > c_1, \end{cases}$$

где $c_1 = m_0 + \tau_{1-\alpha_0}/\sqrt{n}$; $\tau_{1-\alpha_0}$ — квантиль распределения $N_{0,1}$.

Вероятность ошибки второго рода при этом

$$\beta = P_{H_1}(\delta(\vec{X}) = H_0) = P_{H_1}\Big(\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} < c\Big) = P_{H_1}(\bar{X} < c_1) = F_{0,1}\Big(\frac{c_1 - m_1}{1/\sqrt{n}}\Big).$$

Пример 40.3. Имеется выборка $\{X_1, X_2, \ldots, X_n\}$ из нормальной совокупности с неизвестным параметром σ и средним, равным нулю. Требуется построить минимаксный критерий для проверки гипотезы $H_0 : \sigma = \sigma_0$ против альтернативы $H_1 : \sigma = \sigma_1$ $(\sigma_0 < \sigma_1)$.

Решение. Запишем функции правдоподобия и отношение правдоподобия:

$$\Psi_{H_0}(\vec{X}) = \Big(\frac{1}{\sqrt{2\pi}\sigma_0}\Big)^n \exp\Big\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^n (X_i)^2\Big\},$$

$$\Psi_{H_1}(\vec{X}) = \Big(\frac{1}{\sqrt{2\pi}\sigma_1}\Big)^n \exp\Big\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^n (X_i)^2\Big\},$$

$$\frac{\Psi_{H_1}(\vec{X})}{\Psi_{H_0}(\vec{X})} = \Big(\frac{\sigma_0}{\sigma_1}\Big)^n \exp\Big[\frac{1}{2}\Big(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\Big)\sum_{i=1}^n (X_i)^2\Big].$$

Заметим, что условие $\Psi_{H_1}(\vec{X})/\Psi_{H_0}(\vec{X}) > c$ равносильно условию

$$\frac{1}{n}\sum_{i=1}^n (X_i)^2 = \overline{X^2} > c_1.$$

Так как величина $\overline{X^2}$ имеет непрерывное распределение, то минимаксный критерий будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & \text{если } \overline{X^2} < c_1; \\ H_1, & \text{если } \overline{X^2} > c_1, \end{cases}$$

где c_1 есть корень уравнения $P_{H_0}(\overline{X^2} > c_1) = P_{H_1}(\overline{X^2} < c_1)$. При условии истинности H_0 величина $n\overline{X^2}/\sigma_0^2 \in \chi_n^2$, а при условии истинности H_1 величина $n\overline{X^2}/\sigma_1^2 \in \chi_n^2$ и, соответственно, условие $P_{H_0}(\overline{X^2} > c_1) = P_{H_1}(\overline{X^2} < c_1)$ можно переписать в виде

$$P_{H_0}(n\overline{X^2}/\sigma_0^2 > c_1 n/\sigma_0^2) = P_{H_1}(n\overline{X^2}/\sigma_1^2 < c_1 n/\sigma_1^2)$$

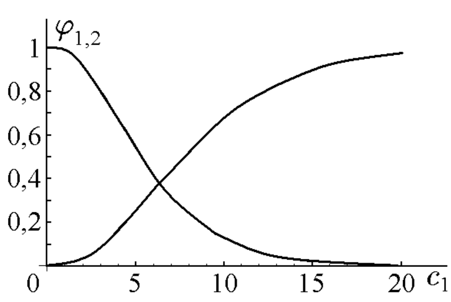


Рис. 99.

или $P(\chi_n^2 > c_1 n/\sigma_0^2) = P(\chi_n^2 < c_1 n/\sigma_1^2)$, где χ_n^2 — случайная величина, имеющая распределение хи-квадрат с n степенями свободы. Поскольку функция $\varphi_1(c_1) = P(\chi_n^2 > c_1 n/\sigma_0^2)$ — убывающая от 1 до 0, а функция $\varphi_2(c_1) = P(\chi_n^2 > c_1 n/\sigma_1^2)$ — возрастающая от 0 до 1 (см. рис. 99), то это уравнение имеет ровно один корень, который, к сожалению, можно найти только численно. Итак, окончательно имеем

$$\delta(\vec{X}) = \begin{cases} H_0, & \text{если } \overline{X^2} < c_1; \\ H_1, & \text{если } \overline{X^2} > c_1, \end{cases}$$

где c_1 есть корень уравнения $P(\chi_n^2 > c_1 n/\sigma_0^2) = P(\chi_n^2 < c_1 n/\sigma_1^2)$.

Рассмотрим далее принципы построения критериев на примере критерия согласия.

♦ Назовем *критерием согласия* (критерием значимости) критерий для проверки простой или сложной гипотезы H_0 против сложной альтернативы H_1, построенный следующим образом:

1) Задается статистика $\rho(\vec{X})$, называемая *статистикой критерия*, которая характеризует степень отклонения эмпирического распределения от теоретического, причем если H_0 верна, то $\rho(\vec{X}) \in F$ или $\rho(\vec{X}) \Rightarrow F$, где F — некоторое известное распределение. Если же H_0 не верна, то, как правило, $|\rho(\vec{X})| \xrightarrow{p} \infty$ либо $\rho(\vec{X}) \Rightarrow F_1$ ($F_1 \neq F$).

2) Для заданного уровня значимости α определяется множество значений $V_{\text{кр}}$ (критическая область) статистики $\rho(\vec{X})$, для которого $P_{H_0}(\rho \in V_{\text{кр}}) = \alpha$. Обычно критическая область $V_{\text{кр}}$ — это одно из множеств вида $[c; +\infty[$, $]-\infty; -c]$, $]-\infty; -c] \cup [c; +\infty[$, где c — некоторая положительная постоянная.

3) Строим критерий $\delta(\vec{X})$ следующим образом:

$$\delta(\vec{X}) = \begin{cases} H_0, & \text{если } \rho(\vec{X}) \notin V_{\text{кр}} \\ H_1, & \text{если } \rho(\vec{X}) \in V_{\text{кр}}. \end{cases}$$

Построенный таким образом критерий имеет уровень значимости α (или стремится к этому уровню при $n \to \infty$) и работает следующим образом: если для данной выборки функция отклонения велика (по абсолютной величине), то гипотеза отвергается, и наоборот. Так как при истинности H_1 $|\rho(\vec{X})| \xrightarrow{p} \infty$, то вероятность не попасть в критическую область при $n \to \infty$ стремится к нулю, то есть построенный критерий является состоятельным.

Заметим, что критерий согласия имеет вполне определенную ошибку первого рода, но ошибка второго рода может быть вычислена только, если известно конкретное распределение, соответствующее альтернативной гипотезе. Поэтому будем рассматривать ошибку второго рода как функцию от альтернативного распределения.

Положение критической области на множестве значений статистики $\rho(\vec{X})$ зависит от распределения статистики $\rho(\vec{X})$ и формулировки альтернативной гипотезы H_1. Если распределение статистики таково, что в предположении истинности гипотезы H_1 более вероятны отклонения одного знака, то выбирают одностороннюю критическую область (правостороннюю или левостороннюю). Если же отклонения разных знаков равновероятны, то выбирают двухстороннюю критическую область.

41. Гипотезы о законе распределения

Рассмотрим проверку гипотезы о законе распределения, используя в качестве критерия согласия критерий Колмогорова.

Пусть $\{X_1, X_2, \ldots, X_n\}$ — выборка из генеральной совокупности F с непрерывной функцией распределения. Проверяется гипотеза $H_0 : F = F_1$ против альтернативы $H_1 : F \neq F_1$.

Пусть $\rho(\vec{X}) = \sqrt{n} \sup |F_n^*(x) - F_1(x)|$. Покажем, что эту статистику можно использовать в качестве статистики критерия согласия для проверки истинности H_0.

1) $\rho(x)$ характеризует отклонение эмпирической функции распределения от теоретической.

2) Если H_0 верна, то по теореме Колмогорова $\rho(x) \Rightarrow K$, где K — распределение Колмогорова.

3) Если H_0 неверна, то по теореме Гливенко–Кантелли $F_n^*(x) \xrightarrow{p} F_2(x)$ для любого x. Поскольку $F_1 \neq F_2$, то найдется такое x, что $|F_2(x) - F_1(x)| > 0$. Тогда для этого x

$$|F_n^*(x) - F_1(x)| \xrightarrow{p} |F_2(x) - F_1(x)| > 0.$$

Поэтому $\rho(\vec{X}) = \sqrt{n} \sup |F_n^*(x) - F_1(x)| \xrightarrow{p} \infty$.

Пусть $\tau_{1-\alpha}$ — квантиль распределения Колмогорова уровня $1-\alpha$, тогда критическая область для статистики уровня α будет определяться неравенством $\rho(x) \geqslant \tau_{1-\alpha}$. Таким образом, критерий Колмогорова уровня α имеет вид

$$\delta(\vec{X}) = \begin{cases} H_0, & \rho(\vec{X}) < \tau_{1-\alpha}; \\ H_1, & \rho(\vec{X}) \geqslant \tau_{1-\alpha}, \end{cases} \quad \rho(\vec{X}) = \sqrt{n} \sup |F_n^*(x) - F_1(x)|, \tag{41.1}$$

где $\tau_{1-\alpha}$ — квантиль распределения Колмогорова.

Величина $\tilde{\alpha}(\vec{X}) = P_{H_0}(\rho \geqslant \rho(\vec{X})$ называется *достигнутым* (или наблюдаемым) уровнем значимости критерия. Эта величина позволяет оценить, на каком уровне значимости мы можем принять нулевую гипотезу. Используя $\tilde{\alpha}(\vec{X})$, критерий согласия Колмогорова можно было бы построить в виде

$$\delta(\vec{X}) = \begin{cases} H_0, & \tilde{\alpha}(\vec{X}) > \alpha, \\ H_1, & \tilde{\alpha}(\vec{X}) \leqslant \alpha. \end{cases}$$

Наряду с критерием Колмогорова применяется также его модификация: критерий Колмогорова–Смирнова, основанный на статистике $\rho(\vec{X}) = \sup |F_n^*(x) - F_1(x)|$. В случае критерия Колмогорова–Смирнова достигнутый уровень значимости $\tilde{\alpha}$ рассчитывался приближенно (для $0{,}01 < \alpha < 0{,}2$ и $n > 10$) по формуле

$$\tilde{\alpha} \approx 2 \exp \left[\sqrt{\frac{1}{2} + \left(1 + \frac{9}{2}n\right)^2 - 18n^2 \left(\rho(\vec{X}) + \frac{1}{6n}\right)^2} - \frac{9}{2}n - 1 \right]. \tag{41.2}$$

Рассмотрим также проверку гипотезы о законе распределения, используя в качестве критерия согласия критерий χ^2 (Пирсона).

Пусть $\{X_1, X_2, \ldots, X_n\}$ — выборка из генеральной совокупности F. Проверяется гипотеза $H_0 : F = F_1$ против альтернативы $H_1 : F \neq F_1$.

Представим выборку в виде сгруппированного ряда, разбив предполагаемую область значений случайной величины на k интервалов. Пусть n_i — число элементов выборки, попавших в i-ый интервал, а p_i — теоретическая вероятность попадания в этот интервал при условии истинности H_0. Составим статистику

$$\rho(\vec{X}) = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i},$$

которая характеризует сумму квадратов отклонения наблюдаемых значений n_i от ожидаемых np_i по всем интервалам группирования.

Теорема 41.1 (Пирсона). *Если H_0 верна, то при фиксированном k и $n \to \infty$*

$$\rho(\vec{X}) = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} \Rightarrow \chi^2_{k-1}. \tag{41.3}$$

Доказательство. Докажем теорему для $k = 2$. Тогда $n_2 = n - n_1$, $p_2 = 1 - p_1$ и

$$\rho(\vec{X}) = \frac{(n_1 - np_1)^2}{np_1} + \frac{(n_2 - np_2)^2}{np_2} = \frac{(n_1 - np_1)^2}{np_1} + \frac{[n - n_1 - n(1 - p_1)]^2}{n(1 - p_1)} =$$
$$= \frac{(n_1 - np_1)^2}{np_1} + \frac{(-n_1 + np_1)^2}{n(1 - p_1)} = \frac{(n_1 - np_1)^2(1 - p_1 + p_1)}{np_1(1 - p_1)} =$$
$$= \frac{(n_1 - np_1)^2}{np_1(1 - p_1)} = \Big(\frac{n_1 - np_1}{\sqrt{np_1q_1}}\Big)^2.$$

Но n_1 есть случайная величина, распределенная по биномиальному закону с параметрами n и p_1. Следовательно, $M(n_1) = np_1$, $D(n_1) = np_1q_1$, и в соответствии с теоремой Муавра–Лапласа случайная величина

$$\frac{n_1 - np_1}{\sqrt{np_1q_1}} \Rightarrow \xi \in N_{0,1}.$$

Но тогда

$$\Big(\frac{n_1 - np_1}{\sqrt{np_1q_1}}\Big)^2 \Rightarrow \xi^2 \in \chi^2_1.$$

Если H_1 такова, что хотя бы для одного $i \in \{1, 2, \ldots, k\}$, вероятность попадания в i-ый интервал $\tilde{p}_i \neq p_i$, то с учетом закона больших чисел

$$\frac{(n_i - np_i)^2}{np_i} = \frac{n(n_i/n - p_i)^2}{p_i} \xrightarrow{p} \frac{n(\tilde{p}_i - p)}{p_i} \to \infty.$$

Таким образом, статистика $\rho(\vec{X})$ подходит в качестве статистики критерия согласия для проверки гипотезы о виде закона распределения, который будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & \rho(\vec{X}) < \tau_{1-\alpha}; \\ H_1, & \rho(\vec{X}) \geqslant \tau_{1-\alpha}, \end{cases} \quad \rho(\vec{X}) = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}, \tag{41.4}$$

где $\tau_{1-\alpha}$ — квантиль распределения χ^2_{k-1}.

Данный критерий называется критерием χ^2 или критерием согласия Пирсона.

$\lozenge$ Критерий несостоятелен для альтернатив, для которых $\tilde{p}_i = p_i$ для всех $i \in \{1, 2, \ldots, k\}$. Поэтому следует стремиться к как можно большему числу интервалов группирования. Однако, с другой стороны, сходимость к χ^2 величины $(n_i - np_i)^2/(np_i)$ обеспечивается центральной предельной теоремой, а погрешность этого приближения $\sim 1/\sqrt{np_iq_i}$, то есть ожидаемое значение np_i для каждой ячейки не должно быть слишком мало. Поэтому обычно число интервалов выбирают таким образом, чтобы $np_i \geqslant 5$.

Рассмотрим критерий χ^2 Пирсона для сложной гипотезы.

Пусть $\{X_1, X_2, \ldots, X_n\}$ — выборка из генеральной совокупности F. Проверяется сложная гипотеза $H_0 : F = F_\theta$, где θ — неизвестный параметр распределения F (или вектор параметров), против альтернативы $H_1 : F \neq F_\theta$.

Пусть выборка по-прежнему представлена в виде группированного ряда и n_i — число элементов выборки, попавших в i-ый интервал, $i \in \{1, 2, \ldots, k\}$. Статистику (41.2) мы не можем в этом случае использовать для построения критерия Пирсона, так как не можем вычислить теоретические значения вероятностей p_i, которые зависят от неизвестного параметра θ. Пусть θ^* — оценка параметра θ по методу максимального правдоподобия, а $p_i^*(\theta^*)$ — соответствующие ей оценки вероятностей p_i. Составим статистику

$$\rho(\vec{X}) = \sum_{i=1}^{k} \frac{(n_i - np_i^*)^2}{np_i^*}.$$

Теорема 41.2 (Пирсона). *Если H_0 верна и l — число компонент вектора θ (число неизвестных параметров распределения), то при фиксированном k и $n \to \infty$*

$$\rho(\vec{X}) = \sum_{i=1}^{k} \frac{(n_i - np_i^*)^2}{np_i^*} \Rightarrow \chi^2_{k-l-1}. \tag{41.5}$$

(Без доказательства.)

Таким образом, критерий Пирсона для параметрической гипотезы будет иметь вид:

$$\delta(\vec{X}) = \begin{cases} H_0, & \rho(\vec{X}) < \tau_{1-\alpha}; \\ H_1, & \rho(\vec{X}) \geqslant \tau_{1-\alpha}, \end{cases} \quad \rho(\vec{X}) = \sum_{i=1}^{k} \frac{(n_i - np_i^*)^2}{np_i^*}, \tag{41.6}$$

где $\tau_{1-\alpha}$ — квантиль распределения χ^2_{k-l-1}.

Пример 41.1. По критерию Пирсона для параметрической гипотезы (41.5), (41.6) при уровне значимости $\alpha = 0{,}05$ проверить гипотезу о распределении случайной величины ξ по нормальному закону, если задано n_k попаданий выборочных значений случайной величины ξ в подинтервал $\Omega_i = (a_i, b_i)$ с серединой $\bar{x}_i = (a_i + b_i)/2$ (табл. 12).

Таблица 12

Группированный статистический ряд частот случайной величины ξ

Ω_i	$]12, 14]$	$]14, 16]$	$]16, 18]$	$]18, 20]$	$]20, 22]$
$\bar{x}_i$	13	15	17	19	21
n_i	6	8	12	16	13

Решение. Определим объем выборки:

$$\sum_{i=1}^{k} n_i = n = 55$$

и найдем по выборке оценки математического ожидания и дисперсии (таким образом, $l = 2$) предполагаемого нормального распределения случайной величины ξ,

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{5} \bar{x}_i n_i = 17{,}8;$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{5} (\bar{x}_i - \bar{x})^2 n_i \approx 6{,}75556; \quad s \approx 2{,}6.$$

Подсчитаем оценку вероятности p_i для предполагаемого нормального распределения случайной величины ξ по формуле

$$p_i^* = P(\xi \in \Omega_i) = \Phi\Big(\frac{b_i - \bar{x}}{s}\Big) - \Phi\Big(\frac{a_i - \bar{x}}{s}\Big), \quad k = 1, 2, 3, 4, 5,$$

где a_i и b_i — соответственно нижняя и верхняя границы подинтервалов Ω_i, причем крайние границы расширены до бесконечности ($a_1 = -\infty$, $b_5 = \infty$, а $\Phi(x)$ — значение функции Лапласа, вычисляемое по таблице. Добавим также строку плотностей частот

$$f_i = \frac{\omega_i}{\Delta_i},$$

где $\omega_i = n_i/n$ — относительные частоты, а $\Delta_i = b_i - a_i$ — длины соответствующих интервалов группирования. Таким образом, расширенную таблицу выборочного распределения можно представить в виде

Таблица 13

Расширенный группированный статистический ряд

Ω_i	$]-\infty, 14]$	$]14, 16[$	$[16, 18[$	$[18, 20[$	$[20, \infty[$
n_i	6	8	12	16	13
p_i^*	0,07	0,17	0,29	0,27	0,20
f_k	0,055	0,085	0,11	0,145	0,1175

Далее вычисляем статистическое значение критерия (41.6):

$$\rho = \sum_{i=1}^{5} \frac{(n_i - np_i^*)^2}{np_i^*} \approx 2{,}83.$$

Затем, при уровне значимости $\alpha = 0{,}05$, определяем $\rho_{\text{кр}}$ по таблице критических точек распределения χ^2 с $k - l - 1 = 5 - 2 - 1 = 2$ степенями свободы: $\rho_{\text{кр}}(\alpha, r - l - 1) = \rho_{\text{кр}}(0{,}05, 2) = 6{,}0$. Наконец, так как $\rho = 2{,}83 < 6{,}0 = \rho_{\text{кр}}$, то делаем вывод о том, что гипотеза о распределении случайной величины ξ по нормальному закону не отвергается. Отметим, что квантиль распределения $\tau_{1-\alpha}$ уровня $1 - \alpha$ есть критическая точка $\rho_{\text{кр}}$ этого распределения уровня α.

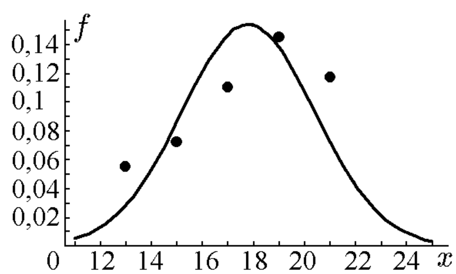


Рис. 100.

Изобразим на фоне точек массива $\{\bar{x}_i, f_i\}$ график кривой плотности нормального распределения $f(x)$, беря в качестве математического ожидания и дисперсии их подсчитанные оценки $\bar{x} = 17{,}8$; $s \approx 2{,}6$.

42. Гипотезы о параметрах распределения одномерной случайной величины

Рассмотрим проверку гипотезы о параметрах распределения, используя критерии, основанные на доверительных интервалах.

Пусть $\{X_1, X_2, \ldots, X_n\}$ — выборка из генеральной совокупности F_θ. Проверяется гипотеза $H_0 : \theta = \theta_0$ против альтернативы $H_1 : \theta \neq \theta_0$. Если имеется доверительный интервал для параметра θ (точный или асимптотический) уровня $1 - \alpha$, то можно легко построить критерий согласия уровня α для проверки H_0. Действительно, пусть (a, b) — доверительный интервал для θ уровня $1 - \alpha$, тогда

$$\delta(\vec{X}) = \begin{cases} H_0, & \theta_0 \in]a, b[; \\ H_1, & \theta_0 \notin]a, b[. \end{cases} \tag{42.1}$$

Так как условие $P(\theta_0 \in]a, b[) = 1 - \alpha$ равносильно условию $P(\rho_1 < \rho(\vec{\xi}, \theta_0) < \rho_2) = 1 - \alpha$, где ρ — статистика, используемая для построения доверительного интервала, то можно использовать при построении критерия непосредственно ее, не находя доверительный интервал. Так как длина правильно построенного доверительного интервала с ростом n стремится к нулю, то вероятность не попасть в него, если H_0 неверна, стремится к единице, то есть критерий, построенный на основе доверительного интервала, является состоятельным.

Рассмотрим далее гипотезы о значении параметров нормальной совокупности:

1) Пусть $\{X_1, X_2, \ldots, X_n\}$ — выборка из нормальной совокупности N_{m,σ^2} с неизвестным параметром m. Проверяется простая гипотеза $H_0 : m = m_0$ против сложной альтернативы $H_1 : m \neq m_0$. Возьмем в качестве статистики критерия статистику (41.1):

$$\rho(\vec{X}) = \sqrt{n}\frac{\bar{X} - m_0}{\sigma},$$

использованную нами для построения доверительного интервала для параметра m при известном σ и имеющую распределение $N_{0,1}$ в случае истинности H_0. Тогда критерий уровня α будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & |\rho(\vec{X})| < \tau_{1-\alpha/2}; \\ H_1, & |\rho(\vec{X})| \geqslant \tau_{1-\alpha/2}, \end{cases} \quad \rho(\vec{X}) = \sqrt{n}\frac{\bar{X} - m_0}{\sigma}, \tag{42.2}$$

где $\tau_{1-\alpha/2}$ — квантиль распределения $N_{0,1}$ уровня $1 - \alpha/2$.

Если альтернативная гипотеза имеет вид, например, $H_1 : m > m_0$, то при истинности H_1 более вероятны положительные отклонения от m_0 и, следовательно, разумнее выбрать одностороннюю критическую область и использовать критерий вида

$$\delta(\vec{X}) = \begin{cases} H_0, & \rho(\vec{X}) < \tau_{1-\alpha}; \\ H_1, & \rho(\vec{X}) \geqslant \tau_{1-\alpha}, \end{cases} \quad \rho(\vec{X}) = \sqrt{n}\frac{\bar{X} - m_0}{\sigma}. \tag{42.3}$$

2) Пусть $\{X_1, X_2, \ldots, X_n\}$ — выборка из нормальной совокупности N_{m,σ^2} с неизвестными параметрами m и σ^2. Проверяется сложная гипотеза $H_0 : m = m_0$ против сложной альтернативы $H_1 : m \neq m_0$. Возьмем в качестве статистики критерия статистику (41.1):

$$\rho(\vec{X}) = \sqrt{n}\frac{\bar{X} - m_0}{s},$$

использованную нами для построения доверительного интервала для параметра m при неизвестном σ и имеющую распределение Стьюдента с $n - 1$ степенью свободы в случае истинности H_0. Тогда критерий уровня α будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & |\rho(\vec{X})| < \tau_{1-\alpha/2}; \\ H_1, & |\rho(\vec{X})| \geqslant \tau_{1-\alpha/2}, \end{cases}, \quad \rho(\vec{X}) = \sqrt{n}\frac{\bar{X} - m_0}{s}, \tag{42.4}$$

где $\tau_{1-\alpha/2}$ — квантиль распределения Стьюдента с $n-1$ степенью свободы уровня $1 - \alpha/2$.

3) Пусть $\{X_1, X_2, \ldots, X_n\}$ — выборка из нормальной совокупности N_{m,σ^2} с неизвестным параметром σ^2. Проверяется простая гипотеза $H_0 : \sigma^2 = \sigma_0^2$ против сложной альтернативы $H_1 : \sigma^2 \neq \sigma_0^2$. Возьмем в качестве статистики критерия статистику (41.1):

$$\rho(\vec{X}) = \frac{n\bar{D}_0}{\sigma_0^2},$$

использованную нами для построения доверительного интервала для параметра σ^2 при известном m и имеющую распределение хи-квадрат с n степенями свободы в случае истинности H_0. Тогда критерий уровня α будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & \tau_{\alpha/2} < \rho(\vec{X}) < \tau_{1-\alpha/2}; \\ H_1, & \rho(\vec{X}) \leqslant \tau_{\alpha/2} \text{ или } \rho(\vec{X}) \geqslant \tau_{1-\alpha/2}; \end{cases}$$

$$\rho(\vec{X}) = \frac{n\bar{D}_0}{\sigma_0^2} = \sum_{i=1}^{n} \left(\frac{X_i - a}{\sigma_0}\right)^2, \tag{42.5}$$

где $\tau_{\alpha/2}$, $\tau_{1-\alpha/2}$ — квантили распределения χ^2 с n степенями свободы уровня $\alpha/2$ и $1-\alpha/2$ соответственно.

4) Пусть $\{X_1, X_2, \ldots, X_n\}$ — выборка из нормальной совокупности N_{m,σ^2} с неизвестными параметрами m и σ^2. Проверяется сложная гипотеза $H_0 : \sigma^2 = \sigma_0^2$ против сложной альтернативы $H_1 : \sigma^2 \neq \sigma_0^2$. Возьмем в качестве статистики критерия статистику (41.1):

$$\rho(\vec{X}) = \frac{(n-1)s^2}{\sigma_0^2},$$

использованную нами для построения доверительного интервала для параметра σ^2 при неизвестном m и имеющую распределение хи-квадрат с $n-1$ степенями свободы в случае истинности H_0. Тогда критерий будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & \tau_{\alpha/2} < \rho(\vec{X}) < \tau_{1-\alpha/2}; \\ H_1, & \rho(\vec{X}) \leqslant \tau_{\alpha/2} \text{ или } \rho(\vec{X}) \geqslant \tau_{1-\alpha/2}, \end{cases}$$

$$\rho(\vec{X}) = \frac{(n-1)s^2}{\sigma_0^2} = \sum_{i=1}^{n} \left(\frac{X_i - \bar{\xi}}{\sigma_0}\right)^2, \tag{42.6}$$

где $\tau_{\alpha/2}$, $\tau_{1-\alpha/2}$ — квантили распределения χ^2 с $n-1$ степенью свободы уровня $\alpha/2$ и $1-\alpha/2$ соответственно.

Рассмотрим *гипотезу о равенстве дисперсий* двух нормальных выборок.

Пусть имеются две независимые выборки $\{X_1, X_2, \ldots, X_{n_1}\}$ объема n_1 и $\{Y_1, Y_2, \ldots, Y_{n_2}\}$ объема n_2 из генеральных совокупностей, имеющих распределения $N(m_1, \sigma_1^2)$ и $N(m_2, \sigma_2^2)$ соответственно. Проверяется сложная гипотеза $H_0 : \sigma_1^2 = \sigma_2^2$ против альтернативы $H_1 : \sigma_1^2 \neq \sigma_2^2$.

1) Пусть математические ожидания m_1, m_2 обеих совокупностей известны. Согласно лемме Фишера, величины

$$\frac{n_1\bar{D}_1}{\sigma_1^2} \in \chi^2_{n_1}, \qquad \frac{n_2\bar{D}_2}{\sigma_2^2} \in \chi^2_{n_2}.$$

Здесь

$$\bar{D}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1}(X_i - m_1)^2, \quad \bar{D}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2}(Y_i - m_2)^2.$$

Тогда при условии истинности H_0: $\sigma_1^2 = \sigma_2^2$ статистика

$$\rho(\vec{X}, \vec{Y}) = \frac{n_1\bar{D}_1/\sigma_1^2 n_1}{n_2\bar{D}_2/\sigma_2^2 n_2} = \frac{\bar{D}_1}{\bar{D}_2} \in F_{n_1,n_2}.$$

Если гипотеза H_0 верна, т.е. $\rho(\vec{X},\vec{Y})$ имеет распределение Фишера с n_1, n_2 степенями свободы, то, согласно свойствам этого распределения, $\rho(\vec{X},\vec{Y}) \xrightarrow{p} 1$. Если же H_0 неверна, то

$$\rho(\vec{X},\vec{Y}) \xrightarrow{p} \frac{\sigma_1^2}{\sigma_2^2} \neq 1.$$

Это означает, что при $n \to \infty$ вероятность того, что значение статистики попадет в критическую область, если H_0 неверна, стремится к единице. Таким образом, эту статистику можно использовать для построения критерия сравнения дисперсий, который будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & \tau_{\alpha/2} < \rho(\vec{X}) < \tau_{1-\alpha/2}; \\ H_1, & \rho(\vec{X}) \leqslant \tau_{\alpha/2} \quad \text{или} \quad \rho(\vec{X}) \geqslant \tau_{1-\alpha/2}, \end{cases}$$

$$\rho(\vec{X},\vec{Y}) = \frac{\bar{D}_1}{\bar{D}_2} = \frac{n_2}{n_1}\frac{\sum\limits_{i=1}^{n_1}(X_i - a_1)^2}{\sum\limits_{i=1}^{n_2}(Y_i - a_2)^2}, \tag{42.7}$$

где $\tau_{\alpha/2}$, $\tau_{1-\alpha/2}$ — квантили распределения Фишера с n_1, n_2 степенями свободы уровня $\alpha/2$ и $1-\alpha/2$ соответственно.

Можно заметить, что если $\bar{D}_1 > \bar{D}_2$, то значение статистики $\rho(\vec{X},\vec{Y}) > 1$ и для проверки гипотезы можно использовать только один квантиль $\tau_{1-\alpha}$.

2) Пусть теперь m_1, m_2 неизвестны. Согласно лемме Фишера, величины

$$\frac{(n_1-1)s_1^2}{\sigma_1^2} \in \chi^2_{n_1-1}, \qquad \frac{(n_2-1)s_2^2}{\sigma_2^2} \in \chi^2_{n_2-1}.$$

Тогда при условии истинности H_0: $\sigma_1^2 = \sigma_2^2$ статистика

$$\rho(\vec{X},\vec{Y}) = \frac{s_1^2}{s_2^2} \in F_{n_1-1,n_2-1}$$

и, соответственно, критерий сравнения дисперсий будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & \tau_{\alpha/2} < \rho(\vec{X}) < \tau_{1-\alpha/2}; \\ H_1, & \rho(\vec{X}) \leqslant \tau_{\alpha/2} \quad \text{или} \quad \rho(\vec{X}) \geqslant \tau_{1-\alpha/2}, \end{cases}, \quad \rho(\vec{X},\vec{Y}) = \frac{s_1^2}{s_2^2}, \tag{42.8}$$

где $\tau_{\alpha/2}$, $\tau_{1-\alpha/2}$ — квантили распределения Фишера с n_1-1, n_2-1 степенями свободы уровня $\alpha/2$ и $1-\alpha/2$ соответственно.

Можно заметить, что если $\bar{D}_1 > \bar{D}_2$, то значение статистики $\rho(\vec{X},\vec{Y}) > 1$ и для проверки гипотезы можно использовать только одну квантиль $\tau_{1-\alpha/2}$.

Рассмотрим **гипотезы о равенстве средних** двух нормальных выборок.

Пусть имеются две независимые выборки: $\{X_1, X_2, \ldots, X_{n_1}\}$ объема n_1 и $\{Y_1, Y_2, \ldots, Y_{n_2}\}$ объема n_2 из генеральных совокупностей, имеющих распределения N_{m_1,σ_1^2} и N_{m_2,σ_2^2} соответственно. Проверяется сложная гипотеза $H_0 : m_1 = m_2$ против альтернативы $H_1 : m_1 \neq m_2$.

1. Пусть дисперсии σ_1^2, σ_2^2 обеих совокупностей известны. Так как величина $\bar{X}-m_1 \in N_{0,\sigma_1^2/n_1}$, а величина $\bar{Y}-m_2 \in N_{0,\sigma_2^2/n_2}$, то их разность $(\bar{X}-m_1)-(\bar{Y}-m_2)$ имеет также нормальное распределение с математическим ожиданием, равным нулю, и дисперсией, равной

$$\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

Тогда при условии истинности H_0: $m_1 = m_2$ статистика

$$\rho(\vec{X}, \vec{Y}) = \frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \in N_{0,1}.$$

Соответственно критерий сравнения средних будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & |\rho(\vec{X}, \vec{Y})| < \tau_{1-\alpha/2}; \\ H_1, & |\rho(\vec{X}, \vec{Y})| \geqslant \tau_{1-\alpha/2}, \end{cases} \quad \rho(\vec{X}, \vec{Y}) = \frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}, \tag{42.9}$$

где $\tau_{1-\alpha/2}$ — квантиль распределения $N_{0,1}$ уровня $1 - \alpha/2$.

2. Пусть дисперсии σ_1^2, σ_2^2 обеих совокупностей неизвестны, но совпадают (или принята гипотеза о равенстве дисперсий): $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

1) Тогда величина $\bar{X} - a_1 \in N_{0,\sigma^2/n_1}$, величина $\bar{Y} - m_2 \in N_{0,\sigma^2/n_2}$, и разность $(\bar{X} - m_1) - (\bar{Y} - m_2)$ имеет нормальное распределение

$$N\Big(0, \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}\Big) = N\Big(0, \sigma^2 \frac{n_1 + n_2}{n_1 n_2}\Big),$$

а величина

$$\xi = \frac{1}{\sigma}\sqrt{\frac{n_1 n_2}{n_1 + n_2}}[(\bar{X} - m_1) - (\bar{Y} - m_2)]$$

имеет стандартное нормальное распределение.

2) Согласно лемме Фишера, величина

$$\frac{(n_1 - 1)s_1^2}{\sigma^2}$$

имеет распределение χ^2 с $n_2 - 1$ степенью свободы, а величина

$$\frac{(n_2 - 1)s_2^2}{\sigma^2}$$

имеет распределение χ^2 с $n_1 - 1$ степенью свободы, тогда их сумма

$$\eta = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{\sigma^2}$$

имеет распределение χ^2 с $n_1 + n_2 - 2$ степенями свободы.

3) По определению, величина

$$\frac{\xi}{\sqrt{\eta/(n_1 + n_2 - 2)}} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \frac{(\bar{X} - m_1) - (\bar{Y} - m_2)}{\sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}}$$

имеет распределение Стьюдента с $n_1 + n_2 - 2$ степенями свободы. Тогда статистика

$$\rho(\vec{X}, \vec{Y}) = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \frac{\bar{X} - \bar{Y}}{\sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}}$$

при условии истинности H_0 будет иметь такое же распределение, и, соответственно, критерий сравнения средних будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & |\rho(\vec{X},\vec{Y})| < \tau_{1-\alpha/2}; \\ H_1, & |\rho(\vec{X},\vec{Y})| \geqslant \tau_{1-\alpha/2}, \end{cases}$$

$$\rho(\vec{X},\vec{Y}) = \sqrt{\frac{n_1 n_2}{n_1+n_2}} \frac{\bar{X}-\bar{Y}}{\sqrt{\dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}}, \qquad (42.10)$$

где $\tau_{1-\alpha/2}$ — квантиль распределения Стьюдента с n_1+n_2-2 степенями свободы уровня $1-\alpha/2$.

3. Если дисперсии σ_1^2, σ_2^2 обеих совокупностей неизвестны и не равны либо отклонена гипотеза о равенстве этих дисперсий, то приближенный критерий равенства средних в этом случае имеет вид

$$\delta(\vec{X}) = \begin{cases} H_0, & |\rho(\vec{X},\vec{Y})| < \tau_{1-\alpha/2}; \\ H_1, & |\rho(\vec{X},\vec{Y})| \geqslant \tau_{1-\alpha/2}, \end{cases} \quad \rho(\vec{X},\vec{Y}) = \sqrt{\frac{n_1 n_2}{n_1+n_2}} \frac{\bar{X}-\bar{Y}}{\sqrt{s^2}}, \qquad (42.11)$$

где

$$s^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2},$$

$\tau_{1-\alpha/2}$ — квантиль уровня $1-\alpha/2$ распределения Стьюдента с

$$k = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\dfrac{(s_1^2/n_1)^2}{n_1-1} + \dfrac{(s_2^2/n_2)^2}{n_2-1}}$$

степенями свободы.

Пример 42.1. По двум независимым выборкам объемов $n_1 = 26$ и $n_2 = 29$ нормальных распределений найдены исправленные выборочные дисперсии $s_1^2 = 0{,}56$ и $s_2^2 = 0{,}38$. При уровне значимости $\alpha = 0{,}10$ проверить нулевую гипотезу H_0: $D_1 = D_2$ о равенстве дисперсий при конкурирующей гипотезе H_1: $D_1 > D_2$.

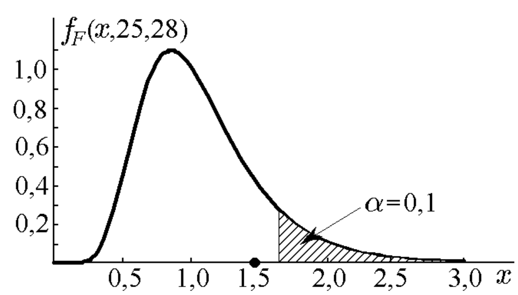


Рис. 101.

Решение. Найдем, согласно (42.8), отношение большей исправленной дисперсии к меньшей:

$$F_{\text{набл}} = \frac{s_{\text{Б}}^2}{s_{\text{М}}^2} = \frac{0{,}56}{0{,}38} \approx 1{,}47.$$

По условию, конкурирующая гипотеза имеет вид H_1: $D_1 > D_2$, поэтому критическая область — правосторонняя, и находим только одну квантиль $\tau_{1-\alpha}$. По таблице критических точек распределения Фишера по уровню значимости $\alpha = 0{,}10$ и числам степеней свободы $k_1 = n_1 - 1 = 26 - 1 = 25$ и $k_2 = n_2 - 1 = 29 - 1 = 28$ находим критическую точку

$$\tau_{1-\alpha} = F_{\text{кр}}(\alpha, k_1, k_2) = F_{\text{кр}}(0{,}10, 25, 28) = 1{,}65.$$

Так как $F_{\text{набл}} = 1{,}47 < 1{,}65 = F_{\text{кр}}$, то нет оснований отвергать H_0 о равенстве дисперсий. Другими словами, исправленные выборочные дисперсии различаются незначимо.

На рис. 101 под кривой плотности вероятностей F-распределения Фишера $f(x, 25, 28)$ выделена критическая область $\{x > 1{,}65 = F_{\text{кр}}\}$ площадью $\alpha = 0{,}10$ и вне ее жирной точкой значение $F_{\text{набл}} = 1{,}47$.

Пример 42.2. По двум независимым выборкам объемов $n_1 = 40$ и $n_2 = 50$ нормальных распределений найдены выборочные значения математических ожиданий $\bar{\xi} = 130$ и $\bar{\eta} = 135$. Дисперсии известны: $\sigma_1^2 = 80$ и $\sigma_2^2 = 100$. При уровне значимости $\alpha = 0{,}02$ проверить нулевую гипотезу H_0: $m_1 = m_2$ о равенстве математических ожиданий при конкурирующей гипотезе H_1: $m_1 \neq m_2$.

Решение. Найдем наблюдаемое значение критерия, согласно (42.9),

$$\rho_{\text{набл}} = \frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} = \frac{130 - 135}{\sqrt{80/40 + 100/50}} = -2{,}5.$$

По условию, конкурирующая гипотеза имеет вид H_1: $m_1 \neq m_2$, поэтому критическая область — двусторонняя. Найдем критическую точку $\rho_{\text{кр}} = \tau_{1-\alpha/2}$ из равенства

$$\Phi(\rho_{\text{кр}}) = (1 - \alpha)/2 = (1 - 0{,}02)/2 = 0{,}49.$$

По таблице функции Лапласа находим $\rho_{\text{кр}} = 2{,}33$.

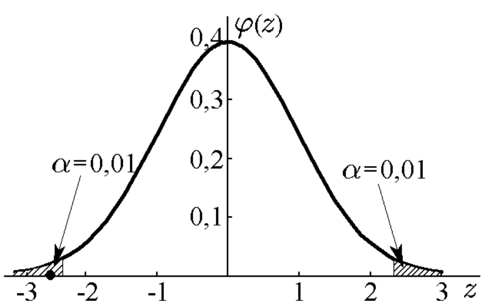


Рис. 102.

Так как $|\rho_{\text{набл}}| = |-2{,}5| = 2{,}5 > 2{,}33 = \rho_{\text{кр}}$, то нулевую гипотезу отвергают. Другими словами, математические ожидания различаются значимо.

На рис. 102 под кривой $\varphi(z)$ плотности вероятностей нормального распределения $N_{0,1}$ выделена двусторонняя критическая область $\{|\rho| > 2{,}33 = \rho_{\text{кр}}\}$ общей площадью $\alpha = 0{,}02$ и внутри нее жирной точкой значение $\rho_{\text{набл}} = -2{,}5$.

Пример 42.3. По двум малым независимым выборкам объемов $n_1 = 12$ и $n_2 = 18$ нормальных распределений найдены выборочные значения математических ожиданий $\bar{X} = 31{,}2$ и $\bar{Y} = 29{,}2$ и исправленные выборочные дисперсии $s_1^2 = 0{,}84$ и $s_2^2 = 0{,}40$. При уровне значимости $\alpha = 0{,}05$ проверить нулевую гипотезу H_0: $m_1 = m_2$ при конкурирующей гипотезе H_1: $m_1 \neq m_2$.

Решение. Исправленные выборочные дисперсии различны, поэтому проверим предварительно гипотезу о равенстве дисперсий, используя критерий Фишера.

Найдем отношение большей исправленной дисперсии к меньшей:

$$\rho_{\text{набл}} = \frac{s_{\text{Б}}^2}{s_{\text{М}}^2} = \frac{0{,}84}{0{,}4} = 2{,}1.$$

Дисперсия s_1^2 значительно больше дисперсии s_2^2, поэтому в качестве конкурирующей гипотезы примем гипотезу $H_1 : D_\xi > D_\eta$. В этом случае критическая область — правосторонняя. По таблице критических точек распределения Фишера, по уровню значимости $\alpha = 0{,}05$ и числам степеней свободы $k_1 = n_1 - 1 = 12 - 1 = 11$ и $k_2 = n_2 - 1 = 18 - 1 = 17$ находим (см. также пример **??**) критическую точку

$$\tau_{1-\alpha/2} = \rho_{\text{кр}}(\alpha, k_1, k_2) = \rho_{\text{кр}}(0{,}05, 11, 17) = 2{,}41.$$

Так как $\rho_{\text{набл}} = 2{,}1 < 2{,}41 = \rho_{\text{кр}}$, то нет оснований отвергать H_0 о равенстве дисперсий. Предположение о равенстве дисперсий не отвергается, поэтому далее проверим гипотезу о равенстве математических ожиданий по формуле (42.10), то есть вычислим наблюдаемое значение критерия Стьюдента

$$T_{\text{набл}} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \frac{\bar{\xi} - \bar{\eta}}{\sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}} \approx 7{,}1.$$

По условию, конкурирующая гипотеза имеет вид $H_1 : m_1 \neq m_2$, поэтому критическая область — двусторонняя. По таблице критических точек распределения Стьюдента, по уровню значимости $\alpha = 0{,}05$ и числу степеней свободы $k = n_1 + n_2 - 2 = 12 + 18 - 2 = 28$ находим двустороннюю критическую точку, т.е. такое $t_{\text{кр}}$, для которого $p(|T| \geqslant t_{\text{кр}}) = \alpha$):

$$t_{\text{кр}}(\alpha, k) = t_{\text{кр}}(0{,}05, 28) = 2{,}05.$$

Так как $|T_{\text{набл}}| = 7{,}1 > 2{,}05 = t_{\text{кр}}$, то нулевую гипотезу о равенстве математических ожиданий отвергаем. Другими словами, математические ожидания различаются значимо.

Пример 42.4. По двум малым независимым выборкам объемов $n_1 = 5$ и $n_2 = 5$ нормальных распределений найдены выборочные значения математических ожиданий $\bar{X} = 13{,}32$ и $\bar{Y} = 13{,}80$ и исправленные выборочные дисперсии $s_1^2 = 3{,}37$ и $s_2^2 = 0{,}46$. При уровне значимости $\alpha = 0{,}05$ проверить нулевую гипотезу H_0: $m_1 = m_2$ при конкурирующей гипотезе H_1: $m_1 \neq m_2$.

Решение. Исправленные выборочные дисперсии различны, поэтому проверим предварительно гипотезу о равенстве дисперсий, используя критерий Фишера. Найдем отношение большей исправленной дисперсии к меньшей:

$$F_{\text{набл}} = \frac{s_{\text{Б}}^2}{s_{\text{М}}^2} = \frac{3{,}37}{0{,}46} \approx 7{,}33.$$

Дисперсия s_1^2 значительно больше дисперсии s_2^2, поэтому в качестве конкурирующей гипотезы примем гипотезу H_1: $D_\xi > D_\eta$. В этом случае критическая область — правосторонняя. По таблице критических точек распределения Фишера, по уровню значимости $\alpha = 0{,}05$ и числам степеней свободы $k_1 = n_1 - 1 = 5 - 1 = 4$ и $k_2 = n_2 - 1 = 5 - 1 = 4$ находим критическую точку

$$F_{\text{кр}}(\alpha, k_1, k_2) = F_{\text{кр}}(0{,}05, 4, 4) = 6{,}39.$$

Так как $F_{\text{набл}} = 7{,}33 > F_{\text{кр}} = 6{,}39$, то гипотеза о равенстве дисперсий отклоняется. Поэтому вычислим наблюдаемое значение критерия Стьюдента по формуле (42.11):

$$T_{\text{набл}} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \frac{\bar{X} - \bar{Y}}{\sqrt{s^2}} \approx -0{,}55.$$

Число степеней свободы

$$k = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}} = \frac{(3{,}37/5 + 0{,}46/5)^2}{\dfrac{(3{,}37/5)^2}{5 - 1} + \dfrac{(0{,}46/5)^2}{5 - 1}} \approx 5.$$

По условию, конкурирующая гипотеза имеет вид H_1: $m_1 \neq m_2$, поэтому критическая область — двусторонняя. По таблице критических точек распределения Стьюдента, по уровню значимости $\alpha = 0{,}05$ и числу степеней свободы $k = 5$ находим двустороннюю критическую точку

$$t_{\text{кр}}(\alpha, k) = t_{\text{кр}}(0{,}05; 5) = 2{,}57.$$

Так как $|T_{\text{набл}}| = 0{,}55 < 2{,}57 = t_{\text{кр}}$, то нет оснований отвергать нулевую гипотезу о равенстве математических ожиданий. Другими словами, математические ожидания различаются незначимо.

43. Непараметрические критерии

Теорию статистических выводов называют непараметрической статистикой, если эти выводы не зависят от неизвестного теоретического распределения и, в частности, от его параметров.

Непараметрические методы позволяют обрабатывать данные «низкого качества» из выборок малого объёма с переменными, о распределении которых мало что или вообще ничего не известно. Непараметрические методы как раз и разработаны для тех ситуаций, достаточно часто возникающих на практике, когда исследователь ничего не знает о параметрах исследуемой популяции (отсюда и название методов — *непараметрические*). Говоря более специальным языком, непараметрические методы не основываются на оценке параметров (таких как среднее или стандартное отклонение) при описании выборочного распределения интересующей величины. Поэтому эти методы иногда также называются *свободными от параметров* или *свободно распределенными*.

По существу, для каждого параметрического критерия имеется, по крайней мере, одна непараметрическая (ранговая) альтернатива. Для проверки гипотезы о законе распределения используется одновыборочный критерий Колмогорова–Смирнова. Различия между независимыми выборками оцениваются с помощью двухвыборочного критерия Колмогорова–Смирнова, критерия серий Вальда–Вольфовица и U-теста Манна–Уитни. Различия между зависимыми выборками оцениваются с помощью критерия знаков и критерия Вилкоксона. Наконец, при оценке зависимости между двумя переменными непараметрическим аналогом коэффициенту корреляции Пирсона является коэффициент корреляции Спирмена.

Если вы имеете несколько независимых групп данных, то непараметрическими аналогами дисперсионного анализа являются ранговый дисперсионный анализ *Краскела–Уоллиса* и *медианный* тест. Если рассматривается более двух зависимых переменных, относящихся к одной и той же выборке, то альтернативным дисперсионному анализу с повторными измерениями непараметрическим методом является *ранговый дисперсионный анализ Фридмана* или Q-критерий Кохрена (последний применяется, например, если переменная измерена в номинальной шкале, то есть принимает конечное множество значений, например, Пол $= \{$Муж, Жен$\}$). Q-критерий Кохрена используется также для оценки изменений частот (долей). Обобщением R-статистики Спирмена на случай зависимости между несколькими переменными является так называемый *коэффициент согласия* (*конкордации*) Кендалла. Этот тест часто используется для оценки согласованности мнений независимых экспертов (судей), в частности баллов, выставленных одному и тому же субъекту.

43.1. Проверка гипотезы о законе распределения

Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — взаимно независимые и одинаково распределенные случайные величины, и пусть $\eta_1 \leqslant \eta_2 \leqslant \ldots \leqslant \eta_n$ — те же ξ_i, но расположенные в порядке возрастания их значений. *Эмпирическим* называют распределение дискретной случайной величины ξ^*, которая принимает значения $\eta_1, \eta_2, \ldots, \eta_n$ с одинаковыми вероятностями, равными $1/n$:

$$P\{\xi^* = \eta_i | \eta_1, \eta_2, \ldots, \eta_n\} = \frac{1}{n} \quad (i = \overline{1, n}).$$

Функция эмпирического распределения задается равенством (29.2):

$$F_n^*(x|\eta_1, \eta_2, \ldots, \eta_n) = P(\xi^* < x|\eta_1, \eta_2, \ldots, \eta_n) = \begin{cases} 0, & \text{если } x \leqslant \eta_1; \\ m/n, & \text{если } \eta_m < x \leqslant \eta_{m+1}; \\ 1, & \text{если } x > \eta_n, \end{cases}$$

$1 \leqslant m \leqslant n-1$, и при каждом действительном x является случайной величиной (функцией от $\eta_1, \eta_2, \ldots, \eta_n$). В дальнейшем функцию эмпирического распределения мы будем обозначать $F_n(x)$, не указывая явно зависимость от величин η_i.

Так как (см. доказательство теоремы 30.1)

$$M(F_n^*) \equiv F(x), \qquad D(F_n^*) \equiv \frac{1}{n} F(x)[1 - F(x)] \to 0 \quad (n \to \infty),$$

где $F(x)$ — функция распределения исходных величин ξ_i (ее называют *теоретической функцией распределения*, то $F_n(x)$ — несмещенная и состоятельная оценка для $F^*(x)$. Если функция теоретического распределения достоверно неизвестна и лишь высказывается гипотеза, согласно которой этой функцией является некоторая заданная функция непрерывного распределения $F(x)$, не содержащая неизвестных параметров, то, обозначив такую гипотезу символом H_0, мы условимся формально записывать ее в виде тождества::

$$H_0 : M(F_n^*(x)) \equiv F(x), \quad |x| < \infty.$$

Точно так же пусть неравенствами, приведенными ниже, выражаются гипотезы, конкурирующие с H_0:

$$H_1^+\{\psi[F(x)]\} : \sup_{|x|<\infty} \psi[F(x)](M(F_n^*(x)) - F(x)) > 0,$$
$$H_1^-\{\psi[F(x)]\} : \inf_{|x|<\infty} \psi[F(x)](M(F_n^*(x)) - F(x)) < 0,$$
$$H_1\{\psi[F(x)]\} : \sup_{|x|<\infty} \psi[F(x)](M(F_n^*(x)) - F(x)) > 0,$$

где $\psi(F)$ — заданная неотрицательная функция (ее часто называют весовой функцией).

Критерии Колмогорова и Смирнова предназначены для проверки гипотезы H_0 при конкурирующей гипотезе H_1 (критерий Колмогорова) и H_1^+ или H_1^- (критерий Смирнова). Статистики критериев задаются формулами

$$D_n = \sup_{|x|<\infty} |F_n^*(x) - F(x)|, \; D_n^+ = \sup_{|x|<\infty} (F_n^*(x) - F(x)), \; D_n^- = -\inf_{|x|<\infty} (F_n^*(x) - F(x)),$$

где в левых частях знаки $+$ и $-$, а также отсутствие знака указывают соответствующую конкурирующую гипотезу.

Для практических вычислений этих статистик полезны другие формулы, эквивалентные предыдущим:

$$D_n^+ = \max_{1\leqslant m\leqslant n}\Big(\frac{m}{n}-F(\eta_m)\Big),\quad D_n^- = \max_{1\leqslant m\leqslant n}\Big(F(\eta_m)-\frac{m-1}{n}\Big),\quad D_n=\max(D_n^+,D_n^-). \tag{43.1}$$

Если гипотеза H_0 верна, то статистики D_n^+ и D_n^- распределены одинаково, поэтому в дальнейшем мы будем рассматривать лишь критерий, основанный на статистике D_n^+.

Согласно асимптотической теории [**?**, 15], при $n\to\infty$ и $0<\varepsilon\leqslant x=O(n^{1/3})$

$$P\Big\{\frac{(6nD_n^+ +1)^2}{18n}<x\Big\}=(1-e^{-x})+e^{-x}\frac{2x^2-4x-1}{18n}+O\Big(\frac{1}{n\sqrt{n}}\Big),$$

$$P\Big\{\frac{(6nD_n+1)^2}{18n}<x\Big\}=K\Big(\sqrt{\frac{x}{2}}\Big)-\frac{1}{18n}\sum_{k=-\infty}^{\infty}e^{-k^2x}[P_k(x)-2k^4x-k^2]+O\Big(\frac{1}{n\sqrt{n}}\Big),$$

где

$$K(y)=\sum_{k=-\infty}^{\infty}(-1)^k e^{-2k^2y^2} \tag{43.2}$$

— функция распределения Колмогорова, которое является предельным распределением случайной величины $\sqrt{n}D_n$ при $n\to\infty$, и

$$P_k(x)=\Big[k^2-\frac{1-(-1)^k}{2}\Big](1-2k^2x)+2k^2x(k^2x-3).$$

Иными словами, при больших значениях n статистика $(6nD_n^+ +1)^2/9n$ приближенно распределена, как χ^2 с двумя степенями свободы, а статистика $(6nD_n+1)^2/18n$ приближенно распределена по закону с функцией распределения $K(\sqrt{x/2})$. Оба эти приближения действуют практически удовлетворительно при $n\geqslant 20$. С ростом n погрешности убывают как $1/n$.

Пусть Q — заданный уровень значимости, выраженный в процентах ($0<Q\leqslant 50\%$), и пусть $D_n^+(Q)$ и $D_n(Q)$ — критические значения статистик D_n^+ и D_n соответственно, определяемые как решения уравнений

$$P\{D_n^+\geqslant D_n^+(Q)\}=0{,}01Q,\qquad P\{D_n\geqslant D_n(Q)\}=0{,}01Q.$$

Если в результате эксперимента окажется, что $D_n\geqslant D_n(Q)$, то, согласно критерию Колмогорова с уровнем значимости Q, гипотеза H_0 должна быть отвергнута (аналогичный вывод по критерию Смирнова делается при $D_n^+\geqslant D_n^+(Q)$).

Если $Q\leqslant 20\%$, то с большой точностью $D_n(Q)\approx D_n^+(0{,}5Q)$. Погрешность этого приближенного равенства при $Q=20$ и 10% не превышает соответственно $5\cdot 10^{-4}$ и $5\cdot 10^{-5}$; с уменьшением Q погрешность быстро убывает.

При $n\geqslant 10$ для определения $D_n(Q)$ на отрезке $1\%\leqslant Q\leqslant 20\%$ и $D_n^+(Q)$ при $Q\geqslant 0{,}5\%$ можно воспользоваться приближенным выражением

$$\sqrt{\frac{1}{2n}\Big(y-\frac{2y^2-4y-1}{18n}\Big)}-\frac{1}{6n}\sim\sqrt{\frac{y}{2n}}-\frac{1}{6n},$$

где $y=-\ln(0{,}01Q)$, если вычисляется $D_n^+(Q)$, и $y=-\ln(0{,}005Q)$, если вычисляется $D_n(Q)$. Для приближенного определения $D_n(Q)$ при $20\%\leqslant Q\leqslant 30\%$ и

$10 \leqslant n \leqslant 50$ рекомендуется полагать y равным корню уравнения $K(\sqrt{y/2}) = 1 - 0{,}01Q$ и применять более точную формулу

$$D_n(Q) \approx \sqrt{\frac{1}{2n}\Big\{y - \frac{1}{18n}\Big[(2y^2 - 4y - 1) - \Big(\frac{Q}{100}\Big)^3\Big(3y^2 - y + \frac{1}{2}\Big)\Big]\Big\}} - \frac{1}{6n}.$$

При $n \geqslant 100$ указанные приближенные формулы позволяют надежно оценивать критические значения $D_n(Q)$ и $D_n^+(Q)$ на отрезке $0{,}1\% \leqslant Q \leqslant 50\%$ (оценки $D_n^+(Q)$ будут удовлетворительными при всех $Q \geqslant 0{,}01\%$).

43.2. Проверка гипотезы об однородности двух выборок

Пусть, помимо выборки $\xi_1, \xi_2, \ldots, \xi_n$, имеются также взаимно независимые случайные величины $\xi_1', \xi_2', \ldots, \xi_m'$, распределенные одинаково и непрерывно, и пусть $\eta_1' \leqslant \eta_2' \leqslant \ldots \leqslant \eta_m'$ — те же величины ξ_i', но расположенные в порядке возрастания их значений (объемы выборок m и n могут быть различными). Обозначим символом $G_m^*(x)$ функцию эмпирического распределения, соответствующую выборке $\xi_1', \xi_2', \ldots, \xi_m'$. Основная гипотеза H_0, подлежащая проверке, заключается в предположении, что обе выборки извлечены из одной и той же совокупности и, значит, функции распределения случайных величин ξ и ξ' одинаковы. Эту гипотезу можно выразить тождеством

$$H_0 : M(F_n^*(x)) \equiv M(G_m^*(x)),$$

где $F_n^*(x)$ — функция эмпирического распределения, построенного на выборке $\xi_1, \xi_2, \ldots, \xi_n$. Возможные конкурирующие гипотезы запишем в виде неравенств:

$$H_1^+ : \sup_{|x|<\infty} M(G_m^*(x) - F_n^*(x)) > 0, \quad H_1^- : \inf_{|x|<\infty} M(G_m^*(x) - F_n^*(x)) < 0,$$
$$H_1 : \sup_{|x|<\infty} M(G_m^*(x) - F_n^*(x)) > 0.$$

В случае конкурирующих гипотез H_1^+ и H_1^- для проверки гипотезы H_0 можно воспользоваться критериями, основанными на статистиках

$$D_{m,n}^+ = \sup_{|x|<\infty} (G_m^*(x) - F_n^*(x)), \qquad D_{m,n}^- = -\inf_{|x|<\infty} (G_m^*(x) - F_n^*(x)).$$

Если гипотеза H_0 верна, то случайные величины $D_{m,n}^-$, $D_{n,m}^-$, $D_{m,n}^+$ и $D_{n,m}^+$ распределены одинаково. Поэтому в дальнейшем мы будем рассматривать лишь статистику $D_{m,n}^+$, причем для определенности будет предполагаться, что $m \leqslant n$.

При конкурирующей гипотезе H_1 для проверки H_0 применяется критерий, статистика которого задается выражением

$$D_{m,n} = \sup_{|x|<\infty} |G_m^*(x) - F_n^*(x)|.$$

Так как $D_{m,n} = D_{n,m}$, то без ограничения общности можно рассматривать лишь статистику $D_{m,n}^+$, предполагая, что $m \leqslant n$.

Практически значения статистик рекомендуется вычислять по формулам, эквивалентным предыдущим:

$$D_{m,n}^+ = \max_{1\leqslant r\leqslant m} \Big(\frac{r}{m} - F_n^*(\eta_r')\Big) = \max_{1\leqslant s\leqslant m} \Big(G_m^*(\eta_s) - \frac{s-1}{n}\Big);$$
$$D_{m,n}^- = \max_{1\leqslant r\leqslant m} \Big(F_n^*(\eta_r') - \frac{r-1}{m}\Big) = \max_{1\leqslant s\leqslant m} \Big(\frac{s}{n} - G_m^*(\eta_s)\Big),$$

$$D_{m,n} = \max(D^+_{m,n}, D^-_{m,n}).$$

Если гипотеза H_0 верна и объемы выборок неограниченно увеличиваются, то

$$\begin{aligned} &\lim_{m\to\infty} P\Big\{\sqrt{\frac{mn}{m+n}}D^+_{m,n} < y\Big\} = 1 - e^{-2y^2}, \\ &\lim_{m\to\infty} P\Big\{\sqrt{\frac{mn}{m+n}}D_{m,n} < y\Big\} = K(y), \end{aligned} \qquad y > 0,$$

где $K(y)$ — функция распределения Колмогорова, заданная формулой (43.2). Таким образом, при больших выборок нормированные остатки $D^+_{m,n}$ и $D_{m,n}$ подчиняются тем же распределениям, что и статистики одностороннего и двустороннего критериев, заданные формулами (43.1).

При умеренных значениях m и n для приближенной оценки функций распределения статистик $D^+_{m,n}$ и $D_{m,n}$ могут оказаться полезными асимптотические формулы.

Если $d = d(m,n)$ — наибольший общий делитель чисел m и n и если

$$b(m,n) = -\frac{d}{N}\sum_{j=2}^{m/d}\frac{\lambda_j}{1-\lambda_j}, \quad N = m + n,$$

где $\lambda_2, \lambda_3, \ldots, \lambda_{m/d}$ — корни уравнения $m\lambda^{N/d} - N\lambda^{m/d} + n = 0$, по модулю меньше единицы, то при $d \to \infty$ для функций распределения случайных величин

$$\begin{aligned} \Delta^+_{m,n} &= \frac{2N}{mn}\Big[\frac{mn}{N}D^+_{m,n} + \frac{n-m}{6N} + b(m,n)\Big]^2, \\ \Delta_{m,n} &= \frac{2N}{mn}\Big[\frac{mn}{N}D_{m,n} + \frac{n-m}{6N} + b(m,n)\Big]^2 \end{aligned}$$

справедливы асимптотические формулы

$$\begin{aligned} P\{\Delta^+_{m,n} < x\} = (1 - e^{-x}) - e^{-x}\Big[\Big(\frac{n-m}{N}\Big)^2(1-2x) + \\ +2\Big(1 - \frac{mn}{N^2}\Big)x(3-x)\Big]\frac{N}{18mn} + O\Big[\Big(\frac{N}{mn}\Big)^{3/2}\Big], \\ P\{\Delta_{m,n} < x\} = K\Big(\sqrt{\frac{x}{2}}\Big) + O\Big(\frac{N}{mn}\Big). \end{aligned} \tag{43.3}$$

Указанные оценки остатков равномерны относительно x, принимающего возможные значения статистик $\Delta^+_{m,n}$ и $\Delta_{m,n}$ в любом фиксированном конечном интервале. Формулы (43.3) имеют ограниченную область применения, так как они действуют хорошо при не слишком малых m. Поэтому для практических вычислений полезны другие формулы, подчиненные единственному условию $n \to \infty$ (m может быть любым целым положительными числом, не превосходящим n). Основой этих формул служит тот факт, что случайные величины

$$D^+_{m,n} - \frac{N}{6mn}\Big(1 - 6b(m,n) - \frac{n-m}{N}\Big), \quad D_{m,n} - \frac{N}{6mn}\Big(1 - 6b(m,n) - \frac{n-m}{N}\Big)$$

при больших n распределены приближенно, как D^+_ν и D_ν соответственно, где $\nu = mn/N$ (D^+_ν и D_ν — указанные выше случайные величины, определенные при целых ν формулами (43.1); их функции распределения имеют смысл и при дробных ν). Более точно, если x — какое-либо из возможных значений $D^+_{m,n}$ или $D_{m,n}$, то при $N \to \infty$ и $x = O(1/\sqrt{\nu})$

$$P\{D^+_{m,n} < x\} = P\Big\{D^+_\nu < x - \frac{1}{6\nu}\Big(1 - 6b(m,n) - \frac{n-m}{N}\Big)\Big\} + O\Big(\frac{1}{N}\Big). \qquad (43.4)$$

Аналогичная формула для $D_{m,n}$ получается из формулы (43.4) после отбрасывания знаков $+$.

Критические значения $\Delta^+_{m,n}(Q)$ и $\Delta_{m,n}(Q)$ статистик $\Delta^+_{m,n}$ и $\Delta_{m,n}$, соответствующие уровню значимости Q ($0 < Q \leqslant 50\%$), представляют собой решения неравенств

$$P\{D_{m,n} \geqslant D_{m,n}(Q)\} \leqslant 0{,}01Q, \qquad P\{D_{m,n} < D_{m,n}(Q)\} > 1 - 0{,}01Q \qquad (43.5)$$

на множестве возможных значений указанных статистик (для $\Delta^+_{m,n}(Q)$ неравенства записываются аналогично (43.5)). С помощью наименьшего общего кратного $k = k(m,n)$ эти решения можно выразить в виде соотношений

$$\Delta^+_{m,n}(Q) = \frac{r^+(Q)}{k}, \qquad \Delta_{m,n}(Q) = \frac{r(Q)}{k}, \qquad (43.6)$$

где r^+ и r — целые числа. Из формул (43.3) следует, что для оценки $r^+(Q)$ и $r(Q)$ можно использовать приближенные равенства

$$\begin{aligned} r^+(Q) &\approx 1 + \Big[k(m,n)\Big\{\sqrt{-\frac{N}{2mn}\ln(0{,}01Q)} - \frac{N}{mn}\Big(\frac{n-m}{6N} + b(m,n)\Big)\Big\}\Big], \\ r(Q) &\approx 1 + \Big[k(m,n)\Big\{\sqrt{\frac{N}{mn}K^{-1}(1 - 0{,}01Q)} - \frac{N}{mn}\Big(\frac{n-m}{6N} + b(m,n)\Big)\Big\}\Big], \end{aligned} \qquad (43.7)$$

где $[y]$ — целая часть числа, а $K^{-1}(P)$ — квантиль распределения Колмогорова. Если $Q \leqslant 10\%$, то $r(Q) = r^+(0{,}5Q)$ и, значит, $D_{m,n}(Q) = D^+_{m,n}(0{,}5Q)$.

В общем случае вычисления непосредственно по (43.7) затруднены тем, что функция $b(m,n)$ определяется весьма сложно. Для вычисления приближенных значений функции $b(m,n)$ можно воспользоваться равенством

$$b(m,n) \approx \frac{1}{2}\frac{m - d(m,n)}{N + d(m,n)}.$$

Рассмотрим **гипотезу о среднем**, применяя непараметрический критерий, основанный на медиане.

Пусть $\{X_1, X_2, \ldots, X_n\}$ — выборка из генеральной совокупности случайной величины ξ, распределение которой симметрично относительно математического ожидания. Проверяется сложная гипотеза H_0: $M(\xi) = m_0$ против сложной альтернативы H_1: $M(\xi) \neq m_0$. Если распределение симметрично относительно $M(\xi)$, то математическое ожидание одновременно является медианой распределения, следовательно, каждое значение X_i с вероятностью $1/2$ может быть больше либо меньше медианы. Если обозначить через m число значений в выборке больших m_0, то при условии истинности H_0 случайная величина m будет иметь биномиальное распределение с параметрами $1/2$ и n. Но тогда, согласно теореме Муавра–Лапласа, величина

$$\rho(\vec{X}) = \frac{m - 0{,}5n}{\sqrt{0{,}5n(1 - 0{,}5)}} = \frac{2m - n}{\sqrt{n}} \Rightarrow N_{0,1}.$$

Соответственно асимптотический критерий истинности H_0 уровня α будет иметь вид

$$\delta(\vec{X}) = \begin{cases} H_0, & |\rho(\vec{X})| < \tau_{1-\alpha/2}; \\ H_1, & |\rho(\vec{X})| \geqslant \tau_{1-\alpha/2}, \end{cases} \quad \rho(\vec{X}) = \frac{2m-n}{\sqrt{n}}, \tag{43.8}$$

где $\tau_{1-\alpha/2}$ — квантиль распределения $N_{0,1}$ уровня $1-\alpha/2$.

Критерий называется *непараметрическим*, поскольку проверяет гипотезу о числовых характеристиках распределения вне зависимости от вида закона распределения и связи его числовых характеристик с параметрами распределения.

Рассмотрим далее в качестве примера непараметрического критерия U-критерий Манна–Уитни для проверки гипотезы H_0 об однородности двух выборок, представляющий непараметрическую альтернативу t-критерию Стьюдента для независимых выборок. U-критерий Манна–Уитни предполагает, что все значения по обеим выборкам случайных величин ξ и η объемов n и m соответственно ранжируются, то есть записываются в один ряд в порядке возрастания. После этого каждый элемент выборок характеризуется рангом — порядковым номером каждого элемента выборок в общем ранжированном ряду из обеих выборок. Наблюдаемое значение критерия U вычисляется по формуле

$$U = W - \frac{1}{2}m(m+1) = \sum_{i=1}^{n}\sum_{j=1}^{m}\delta_{ij},$$

где W — значение критерия Уилкоксона, численно равное сумме рангов элементов второй выборки (объема m) в общем ранжированном ряду;

$$\delta_{ij} = \begin{cases} 1, & \text{если } X_i < Y_j; \\ 0 & \text{в противном случае.} \end{cases}$$

Таким образом, критерий U считает общее число тех случаев, в которых элементы второй выборки превосходят элементы первой выборки.

Распределение случайной величины U асимптотически нормально с параметрами $M[U] = nm/2$ и $D[U] = nm(n+m+1)/12$, чем и пользуются на практике, если $\min\{n, m\} > 25$, для определения критического значения $U_{\text{кр}}(\alpha, n, m)$, соответствующего заданному уровню значимости α. Для случаев, когда n и $m < 25$, пользуются специальными таблицами (Каждан и др., 1990; Большев и др., 1983).

Проверка гипотезы о равенстве средних, определенных по двум выборкам объемов n_1 и n_2, с помощью ξ-критерия Ван-дер-Вардена начинается с того, что все значения по обеим выборкам ранжируются, то есть записываются в один ряд в порядке возрастания. ξ-критерий представляет собой величину

$$\xi = \sum_{i=1}^{n_2}\psi\Big(\frac{i}{n_1+n_2+1}\Big),$$

где i — порядковый номер каждого значения второй выборки в общем ряду; ψ — функция, обратная функции нормального распределения, вычисляется, например, с помощью калькулятора распределения вероятности по нормальному закону.

Вычисленное значение критерия ξ сравнивается с $X_{\text{кр}}$, определенным по специальным таблицам для заданного уровня значимости и объемов выборок (Каждан и др., 1990; Большев и др., 1983). Если $|\xi| > X_{\text{кр}}$, то гипотеза о равенстве выборочных средних отвергается.

При этом следует учитывать особенности применения непараметрических критериев, например, ранговый ξ-критерий Ван-дер-Вардена (Каждан и др., 1990) рекомендуется применять, если предполагается, что наблюдения близко следуют нормальному закону; статистическим U-критерием Манна–Уитни для проверки гипотезы об однородности двух выборок ξ и η объемов n_ξ и n_η следует пользоваться на практике, если только $\min\{n_\xi, n_\eta\} > 25$ [**?**]; критерии серий Вальда–Вольфовица предполагает, что рассматриваемые переменные являются непрерывными и измерены в порядковой шкале (Боровиков В.П., 2003). Заметим, что двухвыборочный критерий Колмогорова–Смирнова (уровень значимости α_{2K-S}), основанный на сравнении эмпирических функций распределения двух выборок и проверяющий гипотезу однородности двух выборок (Боровиков В.П., 2003), является чувствительным как к различию в положении двух выборок, так и к различию общих форм распределений двух выборок (в частности, различия в рассеянии, асимметрии и т. д.).

43.3. Оценка зависимости между переменными

Если выборки малы или распределения существенно отличаются от нормального закона, то для проверки гипотезы о наличии корреляционной связи можно использовать непараметрический аналог коэффициента корреляции r Пирсона — ранговый коэффициент корреляции R Спирмена, вычисляемый аналогично r заменой наблюдаемых значений случайных величин их рангами (порядковыми номерами наблюдаемых значений в объединенной выборке, записанной в порядке возрастания). Значимость коэффициента корреляции R Спирмена проверяется аналогично значимости коэффициента корреляции r Пирсона.

Парные коэффициенты корреляции можно построить из обобщенного коэффициента корреляции Γ:

$$\Gamma = \frac{\sum a_{ik} b_{ik}}{\sqrt{\sum a_{ik}^2 \sum b_{ik}^2}}, \tag{43.9}$$

где ${}_{ik} = a(X_i X_k)$, $b_{ik} = b(Y_i, Y_k)$ — некоторые функции пар наблюдений X и Y соответственно, суммирование ведется по всем парам i, j.

Заметим, что при ${}_{ik} = X_k - {}_i$, $b_{ik} = Y_k - Y_i$ из (43.9) получаем обычный коэффициент корреляции Пирсона r.

Если переменные ранжированы, то удобно работать с их рангами. Упорядочим значения X_i по возрастанию, то есть построим вариационный ряд этих величин. Номер величины ${}_i$ в этом ряде называется её рангом и обозначается R_i. Затем упорядочим значения Y_i в порядке возрастания. Номер величины Y_i в этом ряде называется ее рангом и обозначается S_i. Коэффициент ранговой корреляции Спирмена R вычисляется как обобщенный коэффициент парной корреляции (43.9) с заменой наблюдений их рангами. Формально для обобщенного коэффициента корреляции (43.9) нужно положить $a_{ik} = R_k - R_i$, $b_{ik} = S_k - S_i$. В этом случае можно привести выражение коэффициента ранговой корреляции Спирмена R к виду [?]

$$R = 1 - 6 \sum_{i=1}^{n} \frac{(R_i - S_i)^2}{n^3 - n} \tag{43.10}$$

или, что равносильно,

$$R = \frac{12}{n^3 - n} \sum_{i=1}^{n} \Big(R_i - \frac{n+1}{2}\Big)\Big(S_i - \frac{n+1}{2}\Big).$$

Значение R меняется от -1 до $+1$, причем $R = +1$, когда последовательности рангов полностью совпадают, т.е. $R_i = S_i$, $i = 1, \ldots, n$, и $R = -1$, когда последовательности рангов полностью противоположны, т.е. $R_i = n + 1 - S_i$, $i = 1, \ldots, n$. Коэффициент ранговой корреляции Спирмена R применяется для проверки гипотезы о независимости двух признаков. По величине отклонения R от нуля можно сделать вывод о зависимости или независимости признаков. Для построения соответствующего критерия вычисляется R по формуле (43.10) и при $4 \leqslant n \leqslant 10$ используют таблицы точного распределения [**?**, **?**] для вычисления критического значения, а при $n > 10$ можно воспользоваться, например, тем, что случайная величина $\sqrt{n-1}R$ при $n \to \infty$ распределена асимптотически нормально с параметрами $(0, 1)$. В последнем случае гипотеза о независимости отвергается, если $|R| > u_{1-\alpha/2}/\sqrt{n-1}$, где $u_{1-\alpha/2}$ есть корень уравнения $\Phi(u) = 1 - \alpha/2$ ($\Phi(u)$ — функция стандартного нормального распределения).

Если в формуле (43.9) для обобщенного коэффициента Γ положить $_{ij} = 1$ при $R_i < R_j$ и $_{ij} = -1$ при $R_i > R_j$, то вычисляется так называемый коэффициент корреляции Кендалла τ. Величины b_{ij} задаются аналогичными соотношениями с заменой рангов R_j на ранги S_i наблюдений Y. В этом случае можно привести выражение коэффициента ранговой корреляции Кендалла τк виду [**?**]

$$\tau = -1 + \frac{4N}{n^2 - n}, \tag{43.11}$$

где N — число элементов выборки, для которых одновременно $k > i$ и $S_k > S_i$ (S_i — ранг Y, принадлежащего той паре (X, Y), для которой ранг , т.е. R_i, равен i.

Значение τ меняется от -1 до $+1$. Коэффициент ранговой корреляции Кендалла применяется для проверки гипотезы независимости двух случайных величин. При небольшом объеме выборки ($4 \leqslant n \leqslant 10$) проверка статистической гипотезы независимости производится с помощью специальных таблиц [**?**]. При $n > 10$ пользуются асимптотически нормальным приближением для распределения τ: если $|\tau| > u_{\alpha/2}\sqrt{2(2n+5)/9n(n-1)}$, то гипотеза о независимости отвергается.

Итак, мы ясно видим, что идея всех корреляций возникает из одного и того же источника. Относительно различия этих корреляций можно заметить следующее. По сравнению с параметрической корреляцией (r Пирсона) ранговая корреляция R Спирмена, то есть рассмотрение рангов, а не самих наблюдений, в действительности улучшает оценку зависимости между переменными, так как «подавляет» случайную изменчивость и уменьшает воздействия выбросов. При этом, если статистику Спирмена R можно представить себе как вычисленную по рангам корреляцию Пирсона, т.е. в терминах доли изменения одной величины, связанной с изменением другой, то статистика Кендалла скорее оценивает вероятность, точнее разность между вероятностью того, что наблюдаемые значения переменных имеют один и тот же порядок, и вероятностью того, что порядок различен.

Для измерения степени тесноты статистической связи между более чем двумя m ($m > 2$) порядковыми переменными обычно используется множественный вариант ранговой корреляции — так называемый коэффициент конкордации (согласованности) Кендалла $W(m)$, выборочная версия которого определяется по формуле

$$\bar{W}(m) = \frac{12}{m^2(n^3 - n)} \sum_{i=1}^{n} \left[\sum_{j=1}^{m} R_i(j) - \frac{m(n+1)}{2} \right]^2,$$

где $R_i(j)$ — ранг i-го наблюдения j-й случайной переменной.

Этот тест часто используется для оценки согласованности мнений независимых экспертов (судей). Если все мнения совпадают, то $W = 1$. Если различия между ними очень велики, коэффициент W составляет малую величину. Увеличение W от 0 до 1 означает все большую согласованность рангов. Значимость наблюдаемого значения $\bar{W}(m)$ можно оценить в следующих двух приближениях: в общем случае при любых значениях m и n можно прибегнуть к аппроксимации, основанной на использовании известного в статистике z-распределения Фишера:

$$z = \frac{1}{2}\ln\frac{(m-1)W}{1-W}$$

со степенями свободы $\nu_1 = n - 1 - 2/m$ и $\nu_2 = (m-1)\nu_1$, а в частном случае, если $n > 7$, допускается также аппроксимация, основанная на использовании известного в статистике χ^2-распределения:

$$\chi^2 = m(n-1)W$$

с $\nu = n - 1$ степенями свободы.

Задания для самоконтроля

Теоретические вопросы

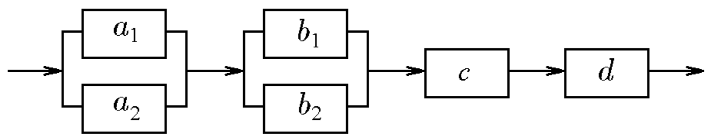
1. Случайное событие. Правила действия над событиями. Элементарное, достоверное, невозможное, противоположное события. Пространство событий, полная группа событий.
2. Частота. Свойства частоты.
3. Аксиомы теории вероятностей.
4. Свойства вероятности. Комбинаторные формулы.
5. Классическое определение вероятности. Геометрические вероятности.
6. Условная вероятность. Попарная независимость событий и независимость в совокупности.
7. Формула полной вероятности. Формула Байеса.
8. Схема Бернулли. Наивероятнейшее число успехов в схеме Бернулли.
9. Предельные теоремы Пуассона и Муавра–Лапласа.
10. Случайная величина. Функция и закон распределения случайной величины. Дискретные и непрерывные случайные величины.
11. Числовые характеристики дискретной случайной величины и их свойства (математическое ожидание, дисперсия, среднеквадратичное отклонение, моменты и их производящая функция).
12. Числовые характеристики непрерывной случайной величины и их свойства (математическое ожидание, дисперсия, среднеквадратичное отклонение, моменты и их производящая функция).
13. Дискретная случайная величина. Биномиальное распределение. Математическое ожидание, дисперсия и среднеквадратичное отклонение.
14. Дискретная случайная величина. Распределение Пуассона. Математическое ожидание, дисперсия и среднеквадратичное отклонение.
15. Непрерывная случайная величина. Равномерное и показательное распределения. Математическое ожидание, дисперсия и среднеквадратичное отклонение.
16. Непрерывная случайная величина. Нормальное распределение (закон Гаусса). Математическое ожидание, дисперсия и среднеквадратичное отклонение.
17. Непрерывная случайная величина. Гамма-распределение. Математическое ожидание, дисперсия и среднеквадратичное отклонение, производящая функция.
18. Непрерывная случайная величина. Бета-распределение. Математическое ожидание, дисперсия и среднеквадратичное отклонение.
19. Понятие многомерной случайной величины. Функция распределения и плотность распределения вероятностей двумерной случайной величины. Свойства.
20. Числовые характеристики двумерной случайной величины и их свойства. Коэффициент корреляции.
21. Предельные теоремы. Закон больших чисел. Сущность и назначение предельных теорем.
22. Неравенство Чебышева.
23. Теоремы Чебышева и Бернулли.
24. Центральная предельная теорема Ляпунова (формулировка). Центрированная и нормированная случайная величина.

Индивидуальные задания

Случайные события и их вероятность

Вариант № 1

1.1 Система S состоит из четырех независимых подсистем S_a, S_b, S_c и S_d. Неисправность хотя бы одной подсистемы ведет к неисправности всей системы (подсистемы соединены последовательно). Подсистемы S_a и S_b состоят из двух независимых дублирующих блоков a_k и b_k $(k = 1, 2)$ (схема параллельного соединения блоков в подсистемах).



Используя определения суммы и произведения событий, записать событие, состоящее в том, что а) система исправна, б) система неисправна. Для контроля использовать свойство противоположного события.

1.2. В условиях предыдущей задачи найти надежность системы — вероятность того, что система будет исправна в течении некоторого времени, если известны надежности блоков $P(a_k) = 0{,}8$; $P(b_k) = 0{,}9$; $P(c) = 0{,}99$; $P(d) = 0{,}95$. Для контроля использовать свойство связи вероятности события с вероятностью противоположного события.

1.3. Доказать тождество: $\overline{(A - B) + (A - C)} = \bar{A} + BC$.

1.4. Колода карт (36 листов) делится случайным образом на две равные части. Найти вероятность того, что в каждой пачке будет по два туза.

1.5. На одной полке наудачу расставляется 8 книг. Найти вероятность того, что определённые 3 книги окажутся поставленными рядом.

1.6. Среди 10 лотерейных билетов 6 выигрышных. Наудачу взяли 4 билета. Определить вероятность того, что среди них хотя бы 2 выигрышных.

1.7. В лифт 6-этажного дома вошли 4 пассажира. Каждый независимо от других с одинаковой вероятностью может выйти на любом (начиная со второго) этаже. Определить вероятность того, что: а) все вышли на разных этажах; б) по крайней мере, трое вышли на одном этаже.

1.8. В отрезке единичной длины наудачу выбираются две точки. Определить вероятность того, что расстояние между точками не превосходит $1/4$.

1.9. Моменты начала двух событий наудачу распределены в промежутке времени длиной 200 минут. Одно из событий длится 10 мин., другое — 5 мин. Определить вероятность того, что: а) события «перекрываются» по времени; б) «не перекрываются».

1.10. В сфере радиуса 2 случайно и независимо друг от друга разбросаны 10 точек. Найти вероятность того, что расстояние от центра до ближайшей точки не меньше 1.

1.11. Вероятность попадания в цель при одном выстреле равна 0,7. Произведено 3 выстрела. Какова вероятность, что будет: а) три попадания; б) один промах; в) хотя бы одно попадание?

1.12. Урна содержит 12 шаров с номерами от 1 до 12. Шары извлекаются по одному без возвращения. Рассматриваются следующие события: A — номера шаров в порядке поступления образуют последовательность 1,2,...,12; B — хотя бы один раз совпадает номер шара и порядковый номер извлечения; C — нет ни одного совпадения номера шара и порядкового номера извлечения. Определить вероятности событий A, B, C. Найти предельные значения вероятностей, если число шаров в урне стремится к бесконечности.

1.13. Бросают три монеты. Определить, зависимы или нет события $A = \{$выпал орёл на первой монете$\}$ и $B = \{$выпала хотя бы одна решка$\}$.

1.14. Мышь может выбрать наугад один из 5 лабиринтов. Известно, что вероятности её выхода из различных лабиринтов за три минуты равны 0,5; 0,6; 0,2; 0,1;

0,1. Пусть оказалось, что мышь выбралась из лабиринта через три минуты. Какова вероятность того, что она выбрала первый лабиринт? Второй лабиринт?

1.15. В первом ящике из 6 шаров 4 красных и 2 чёрных, во втором ящике из 7 шаров 2 красных и 5 чёрных. Из первого ящика во второй переложили один шар, затем из второго в первый переложили один шар. Найти вероятность того, что шар, извлечённый после этого из первого ящика, чёрный.

1.16. Для проверки геодезических работ назначена группа экспертов, состоящая из трёх подгрупп. В первой подгруппе 1 человек, во второй 4 и в третьей 5. Эксперты первой подгруппы принимают верное решение с вероятностью 0,8; эксперты второй подгруппы — с вероятностью 0,6; эксперты третьей подгруппы — с вероятностью 0,5. Наудачу вызванный эксперт принимает 3 независимых решения. Найти вероятность того, что: а) ровно 3 решения приняты верно; б) принимал решения эксперт из первой подгруппы, если 3 решения приняты верно.

1.17. Вероятность выигрыша в лотерею на один билет равна 0,3. Куплены 10 билетов. Найти наивероятнейшее число выигравших билетов и соответствующую вероятность.

1.18. Монету бросают до тех пор, пока орёл не выпадет 3 раза. Определить вероятность того, что при этом решка выпадет 2 раза.

1.19. Вероятность «сбоя» в работе телефонной станции при каждом вызове равна 0,003. Поступило 500 вызовов. Определить вероятность того, что будет более 2 «сбоев».

1.20. Вероятность наступления некоторого события в каждом из 100 независимых испытаний равна 0,8. Определить вероятность того, что число m наступлений события удовлетворяет неравенству $80 \leqslant m \leqslant 90$.

1.21. Из 100 изделий, среди которых имеется 6 нестандартных, выбраны случайным образом 6 изделий для проверки их качества. Используя классическое определение вероятности, формулу Бернулли, формулу Пуассона и локальную теорему Лапласа, определить вероятность того, что среди выбранных 6 изделий окажется ровно 1 нестандартное изделие.

Случайные величины

Вариант № 1

1.1. Имеются 5 ключей, из которых только один подходит к замку. Случайная величина ξ — число проб при открывании замка (испробованный ключ в последующих пробах не участвует). Найти 1) ряд распределения случайной величины ξ; 2) функцию распределения; 3) математическое ожидание, дисперсию, среднеквадратичное отклонение, коэффициент асимметрии и эксцесс распределения.

1.2. Задана плотность распределения непрерывной случайной величины ξ: $f_\xi(x) = A\cos(x)$ при $x \in [-\pi/2; \pi/2]$; $f_\xi(x) = 0$ при $x \notin [-\pi/2; \pi/2]$. Найти коэффициент A и функцию распределения $F_\xi(x)$; построить графики $f_\xi(x)$ и $F_\xi(x)$; найти $M(\xi)$, $D(\xi)$, $\sigma(\xi)$, коэффициент асимметрии $A(\xi)$ и эксцесс распределения $E(\xi)$; найти вероятность попадания случайной величины в интервал $]-3; \pi/4[$.

1.3. Задана функция распределения непрерывной случайной величины ξ:

$$F_\xi(x) = \begin{cases} 0 & \text{при } x \leqslant -2, \\ Ax + B & \text{при } -2 < x \leqslant 4, \\ 1 & \text{при } x \geqslant 4. \end{cases}$$

Найти: а) постоянные A и B; б) плотность вероятности $f_\xi(x)$; в) вероятность попадания случайной величины в интервал $[-3; 1]$, г) $M(\xi)$, $D(\xi)$. Построить графики $f_\xi(x)$ и $F_\xi(x)$.

1.4. Случайная величина ξ может принимать два значения: 2 и -2 с равной вероятностью. Найти характеристическую функцию случайной величины $\gamma(t)$ и, используя ее, вычислить математическое ожидание и дисперсию случайной величины ξ.

1.5. При записи программы на неисправном накопителе появляется в среднем 4 ошибки (поток ошибок предполагается простейшим). Какова вероятность безошибочной записи? Сколько раз в среднем надо записывать программу, чтобы получить безошибочную запись?

1.6. Вероятность выиграть хотя бы на один билет из 100 в лотерею равна 0,8. Сколько в среднем из 100 билетов выигрышных? Каково наивероятнейшее число выигрышных билетов? Предполагается, что вероятность выигрыша на каждый билет одинакова.

1.7. Время работы элемента до отказа подчинено показательному закону распределения с параметром $\alpha = 2 \cdot 10^{-5}$ ч$^{-1}$. Найти среднее время между появлением двух смежных отказов и вероятность безотказной работы к моменту среднего времени после включения технического устройства.

1.8. Коробки с шоколадом упаковываются автоматически: их масса есть нормальная случайная величина со средним 1.06 кг. Найти среднеквадратичное отклонение случайной величины — массы коробок, если известно, что 5% коробок имеют массу меньше 1 кг.

1.9. Случайная величина ξ распределена равномерно на интервале $]-\pi/2, \pi/2[$. Найти плотность распределения случайной величины $\eta = \operatorname{tg} \xi$.

1.10. Случайная величина ξ распределена равномерно на отрезке $[-2; 1]$. Найти плотность распределения случайной величины $\eta = 1/\xi^2$.

Системы случайных величин

Вариант № 1

1.1. Двумерная случайная величина $\{\xi, \eta\}$ распределена равномерно в области D, ограниченной снизу осью Ox, а сверху кривой $\eta = e^{-x^2}$. Найти совместную плотность распределения $f_{\xi,\eta}(x, y)$, плотности распределения $f_\xi(x)$ и $f_\eta(y)$, условные плотности распределения $f_\xi(x/y)$ и $f_\eta(y/x)$, основные числовые характеристики величин ξ и η, коэффициент корреляции между ξ и η.

1.2. Пусть ξ и η — независимые случайные величины, причём ξ имеет равномерное на отрезке $[-1, 1]$ распределение, а η имеет биномиальное распределение с параметрами 2 и 1/2. Найти функцию и плотность распределения суммы $\xi + \eta$.

1.3. Пусть ξ и η — независимые случайные величины, имеющие показательные распределения с параметрами λ_1 и λ_2 соответственно. Доказать, что случайные величины $\xi - \eta$ и $\min\{\xi, \eta\}$ независимы.

1.4. Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — независимые случайные величины. При любом $k \geqslant 1$ величина ξ_{2k-1} имеет распределение Пуассона с параметром $\lambda = 3$, а величина $\xi_{2k} \in N_{0,1}$. Найти предел по вероятности последовательности $(\xi_1 + \xi_2 + \ldots + \xi_n)/n$.

1.5. Дана последовательность независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_n, \ldots$. Случайная величина ξ_n $(n = 0, 1, 2, \ldots)$ может принимать два значения: $\pm \ln^2 n$ с вероятностями, равными 1/2. Удовлетворяет ли эта последовательность закону больших чисел Чебышева?

1.6. Складывается 10^4 чисел, каждое из которых округлено с точностью до 10^{-m}. Предполагается, что ошибки от округления независимы и равномерно распределены в интервале $]-0{,}5 \cdot 10^{-m}, 0{,}5 \cdot 10^{-m}[$. Используя центральную предельную теорему, найти пределы, в которых с вероятностью, не меньшей 0,99, будет суммарная ошибка.

1.7. Случайная величина ξ является средней арифметической независимых и одинаково распределённых случайных величин, среднеквадратичное отклонение каждой из которых равно 2. Сколько нужно взять таких величин, чтобы случайная величина ξ с вероятностью, не меньшей 0,92, имела отклонение от своего математического ожидания, не превосходящее 0,05. Решить задачу, используя а) неравенство Чебышева; б) центральную предельную теорему.

Список литературы

1. Айвазян С.А. *Прикладная статистика; Основы эконометрики*: Учебник: В 2-х т. — М.: ЮНИТИ-ДАНА, 2001 — Т. 2: Основы эконометрики. — 2001. — 432 с.
2. Багров В.Г., Белов В.В., Задорожный В.Н., Трифонов А.Ю. *Методы математической физики.* Т. II, ч. 1: *Специальные функции.* — Томск: Изд-во НТЛ, 2002. — 352 с.
3. Задорожный В.Н., Зальмеж В.Ф., Трифонов А.Ю., Шаповалов А.В. *Высшая математика для технических университетов.* IV. *Ряды*: Учебное пособие. — Изд. 2. — Томск: Изд-во Томского политехнического университета, 2009. — 343 с.
4. Боровков А.А. *Математическая статистика.* — М.: Наука, 1984.
5. Вентцель Е.С. *Теория вероятностей.* — М.: Физматгиз, 1962
6. Вентцель Е.С., Овчаров Л.А. *Прикладные задачи теории вероятностей.* — М.: Радио и связь, 1983.
7. Гмурман В.Е. *Теория вероятностей и математическая статистика.* — М.: Высшая школа, 1998.
8. Гмурман В.Е. *Руководство к решению задач по теории вероятностей и математической статистики.* — М. Высшая школа, 1999.
9. Гнеденко Б.В. *Курс теории вероятностей.* — М.: Наука, 1969.
10. Горелова Г.В., Кацко И.А. Теория вероятностей и математическая статистика в примерах и задачах с применением Excel. — Ростов н/Д: Феникс, 2006.
11. Данко П.Е., Попов А.Г., Кожевникова Т.Я. *Высшая математика в упражнениях и задачах.* — М.: Высшая школа, 1980.
12. Дженкинс Г., Ваттс Д. *Спектральный анализ и его приложения.* — М,: Мир, 1971.
13. Дороговцев А.Я., Сильвестров Д.С., Скороход А.В., Ядренко М.И. *Теория вероятностей. Сборник задач.* — Киев: Вища школа, 1980.
14. Емельянов Г.В., Скитович В.П. *Задачник по теории вероятностей и математической статистике.* — Л.: Изд-во Ленинград. ун-та, 1967.
15. Колмогоров А.Н., и др. *Введение в теорию вероятностей.* — М.: Наука, 1982.
16. Кремер Н.Ш. *Теория вероятностей и математическая статистика*: Учебник для вузов. — М.: ЮНИТИ-ДАНА, 2003. — 543 с.
17. Магазинников Л.И. *Курс лекций по теории вероятностей.* — Томск: Изд-во ТГУ, 1989.
18. Маленво Э. *Статистические методы в эконометрии.* — Вып. 1, 2. — М.: Статистика, 1976.
19. *Справочник по прикладной статистике.* В 2-х т. Т. 2 / Под ред. Э. Ллойда и др. — М.: Финансы и статистика, 1990.
20. Романовский М.Ю., Романовский Ю.М. *Введение в эконофизику. Статистические и динамические модели.* — М.–Ижевск: Институт компьютерных исследований; НИЦ «Регулярная и хаотическая динамика», 2007.
21. *Справочник по теории вероятностей и математической статистике* / В.С. Королюс и др. — М.: Наука, 1985.
22. Феллер В. *Введение в теорию вероятностей и её приложения* (в 2-х томах) — М.: Мир, т. 1 1964, т. 2, 1966.
23. Чистяков В.П. *Курс теории вероятностей.* — М.: Наука, 1982.
24. *Вероятность и математическая статистика*: Энциклопедия. — М.: БРЭ, 1999. — 910 с.

Крицкий Олег Леонидович
Михальчук Александр Александрович
Трифонов Андрей Юрьевич
Шинкев Михаил Леонидович

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА
для технических университетов. I

Теория вероятности

Учебное пособие

Технический редактор *В.Н. Романенко*

Набор и верстка выполнены на компьютерной технике
в издательской системе TEX – LATEX
с использованием семейства шрифтов Computer Modern

Подписано к печати . .2009.
Формат 60×84/8. Бумага офсетная.
Печать RISO. Усл. печ. л. – . Уч.-изд. л. – .
Тираж экз. Заказ № . Цена свободная.
Издательство ТПУ. 634050, г. Томск, проспект Ленина, 30