ВСЯ ВЫСШАЯ МАТЕМАТИКА

М. Л. Краснов
А. И. Киселев
Г. И. Макаренко
Е. В. Шикин
В. И. Заляпин
С. К. Соболев

5

Рекомендовано
Министерством образования
Российской Федерации
в качестве учебника для студентов
высших технических учебных заведений

Эдиториал УРСС • Москва • 2001

ББК 22.1я73

Краснов Михаил Леонтьевич, Киселев Александр Иванович, Макаренко Григорий Иванович, Шикин Евгений Викторович, Заляпин Владимир Ильич, Соболев Сергей Константинович

Вся высшая математика: Учебник. Т. 5. — М.: Эдиториал УРСС, 2001. — 296 с.

ISBN 5-8360-0306-8

Предлагаемый учебник впервые вышел в свет в виде двухтомника сначала на английском и испанском языках в 1990 году, а затем на французском. Он пользуется большим спросом за рубежом.

В 1999 году книга стала лауреатом конкурса по созданию новых учебников Министерства образования России.

Этот учебник адресован студентам высших учебных заведений (в первую очередь будущим инженерам и экономистам) и охватывает практически все разделы математики, но при этом представляет собой не набор разрозненных глав, а единое целое.

Пятый том включает в себя материал по теории вероятностей, математической статистике и теории игр.

Директор — Доминго Марин Рикой
Заместители директора — Наталья Финогенова, Ирина Макеева
Главный редактор — Елена Кудряшова
Компьютерный дизайн — Виктор Романов, Василий Подобед
Верстка — Михаил Кириллов, Ксения Пулькина, Анна Чикунова, Андрей Стулов, Анар Мелик-Еганов
Редакционно-корректурные работы — Василий Подобед, Александр Савченко
Указатель — Василий Подобед, Андрей Стулов
Обработка графики — Василий Подобед, Наталья Аринчева, Елена Колокольчикова
Дизайн обложки — Ирина Макеева
Техническая поддержка — Наталья Аринчева
Менеджер по продажам — Алексей Петяев

Издательство «Эдиториал УРСС». 113208, г. Москва, ул. Чертановская, д. 2/11, к. п.
Лицензия ИД № 03216 от 10.11.2000 г. Гигиенический сертификат на выпуск книжной продукции № 77.ФЦ.8.953.П.270.3.99 от 30.03.99 г. Подписано к печати 20.07.2001 г.
Формат 70×100/16. Тираж 2500 экз. Печ. л. 18,5. Зак. № 244.

Отпечатано в АООТ «Политех-4». 129110, г. Москва, ул. Б. Переяславская, 46.

9 785836 003067 >

ISBN 5-8360-0150-2 (Полное произведение)
ISBN 5-8360-0306-8 (Том 5)

© Эдиториал УРСС, 2001

Все права защищены. Никакая часть настоящей книги не может быть воспроизведена или передана в какой бы то ни было форме и какими бы то ни было средствами, будь то электронные или механические, включая фотокопирование и запись на магнитный носитель, если на то нет письменного разрешения Издательства.

Издательство УРСС
научная и учебная литература

Тел./факс: 7(095)135-44-23
Тел./факс: 7(095)135-42-46
E-mail: urss@urss.ru
Каталог изданий в *Internet*: http://urss.ru

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

При изучении различных явлений в природе и обществе исследователь сталкивается с двумя видами экспериментов — теми, результаты которых однозначно прогнозируемы в данных условиях, и теми, результаты которых в условиях, контролируемых исследователем, однозначно спрогнозировать нельзя, а можно лишь высказать предположение о спектре возможных результатов. В первом случае говорят о детерминированных явлениях, во втором — о явлениях, носящих случайный характер. При этом имеют в виду, что a priori (заранее, до проведения эксперимента или завершения наблюдения за явлением) в первом случае мы в состоянии предсказать результат, а во втором — нет. Для дальнейшего несущественно, чем вызвана подобная непредсказуемость — законами природы, лежащими в основе изучаемого явления или неполнотой информации о процессах, обуславливающих это явление. Важным обстоятельством является наличие самого факта непредсказуемости.

Теория вероятностей, изложению основ которой посвящен этот раздел, призвана дать исследователю возможность описывать подобного рода эксперименты и явления и предоставляет ему надежный инструмент для изучения реальности в ситуациях, когда детерминистическое описание невозможно.

Глава XXXVII

ЭЛЕМЕНТАРНЫЕ СООБРАЖЕНИЯ

Прежде чем переходить к строгим определениям основных понятий теории вероятностей, мы рассмотрим несколько простых и в то же время типичных ситуаций, призванных проиллюстрировать идейную сторону дальнейшего изложения.

§ 1. Схема случаев

Потребность в теоретико-вероятностных методах, как правило, возникает в ситуации, когда исход изучаемого явления по тем или иным причинам не может быть однозначно спрогнозирован. Одной из простейших моделей такого сорта ситуаций может служить *схема случаев*. Под *схемой случаев* мы понимаем такую ситуацию, когда, во-первых, множество возможных исходов рассматриваемого эксперимента образует конечную совокупность, а во-вторых, каждый из исходов имеет такие же шансы на осуществление, как и любой другой. При этом предполагается, что исследуемое явление может наблюдаться в идентичных условиях[1] неограниченное число раз — эксперимент обладает свойством повторяемости. Очевидно, что в этом случае шансы на осуществление того или иного исхода в каждом отдельно взятом эксперименте тем меньше, чем больше самих исходов, а шансы на осуществление какой-либо группы исходов пропорциональны количеству исходов в рассматриваемой группе.

Пример 1. Рассмотрим эксперимент, состоящий в подбрасывании монеты. Пренебрегая «нештатными» возможностями — монета стала на ребро, закатилась в щель, прилипла к потолку и т. п. — будем считать, что возможные исходы этого эксперимента — это выпадение герба или решетки (решки). Если предположить дополнительно, что монета является физически симметричной и что эксперимент производится «честно», то мы получим простейший пример схемы случаев с двумя равновозможными исходами. Заметим, что если монета несимметрична, то рассмотренный эксперимент схемой случаев в нашем понимании описан быть не может, так как исходы уже не будут обладать свойством равновозможности.

Пример 2. Пусть в урне лежит N физически идентичных шаров, пронумерованных от 1 до N. Эксперимент состоит в извлечении шара из урны, возможный исход однократного извлечения — любой номер от 1 до N. Если шары тщательно перемешаны и извлекаются из урны наугад, то мы имеем пример схемы случаев с N равновозможными исходами.

1.1. Вероятность исхода. Событие. Вероятность события

Для описания возможности осуществления того или иного исхода в схеме случаев введем количественную характеристику указанной возможности — ***вероятность исхода***, которую определим как величину, обратно пропорциональную общему количеству N равновозможных при однократном проведении эксперимента исходов:

$$P = \frac{k}{N}, \tag{1}$$

где k — некоторый коэффициент пропорциональности.

---

[1] По крайней мере в принципе.

Для того чтобы понять, каким он должен быть, введем понятие *события*, которое может (или не может) осуществляться в эксперименте:

Событием в эксперименте, описываемом схемой случаев, назовем любую совокупность исходов рассматриваемого эксперимента.

Мы будем говорить, что событие осуществилось, если в результате однократного проведения эксперимента реализовался один из составляющих это событие исходов.

Пример 3. Рассмотрим эксперимент, состоящий в извлечении ровно одной карты из тщательно перетасованной колоды, содержащей 36 карт. Этот эксперимент может быть описан схемой случаев с 36 равновозможными исходами.

Примерами событий в рассматриваемой ситуации могут служить следующие:

— $\spadesuit = \{$извлеченная карта имеет масть пик$\}$. Это событие состоит из всех (9) пиковых карт. Оно осуществляется (происходит) в эксперименте, если мы извлекли из колоды *любую* пиковую карту.

— Десятка $= \{$извлеченная карта — десятка$\}$. Это событие состоит из всех (4) карт с изображением десятки. Оно происходит, если мы извлекли из колоды *любую* десятку.

— Картинка $= \{$извлеченная карта — валет, дама, король или туз$\}$.

Если у нас есть пара событий A и B, то можно сконструировать из них новые события, пользуясь следующими простыми правилами действий:

Суммой двух событий A и B назовем событие, происходящее, если происходит либо событие A, либо событие B, либо оба эти события одновременно[2]. Легко понять, что сумма событий составляется из всех исходов входящих либо в A, либо в B, при этом общие исходы (т. е. входящие одновременно и в A, и в B) входят в сумму однократно.

Обозначение для суммы: $A \cup B$.

Пример 4. В описанном выше эксперименте с извлечением одной карты из 36-листовой колоды суммой событий $A = \{$извлеченная карта масти пик$\} = \{6\spadesuit, \ldots, T\spadesuit\}$ и $B = \{$извлеченная карта — король$\} = \{K\clubsuit, K\diamondsuit, K\heartsuit, K\spadesuit\}$ будет событие $C = \{6\spadesuit, \ldots, T\spadesuit, K\clubsuit, K\diamondsuit, K\heartsuit\}$, состоящее в извлечении карты масти пик или любого короля.

Совмещением (произведением) двух событий A и B назовем событие, происходящее, если события A и B осуществляются одновременно.

Совмещение событий состоит из всех общих для событий A и B исходов.

Обозначение для совмещения: $A \cap B$. Чаще всего знак совмещения $\cap$ опускают, обозначая совмещение событий $A \cdot B = AB$.

Если у событий отсутствуют общие исходы, то такие события вместе не происходят. Они называются *несовместными.* При сложении несовместных событий обычно используется значок «$+$» вместо значка объединения $\cup$ — сумма обозначается в этом случае как $A + B$.

Пример 5. В описанном выше эксперименте совмещением событий A и B будет событие, состоящее в извлечении короля пик.

Здесь же события $\spadesuit$ и $\heartsuit$ — несовместны.

Отрицанием события A или событием, *противоположным* событию A, назовем событие, происходящее, когда событие A не происходит. Отрицание (противоположное событие) состоит из всех тех исходов эксперимента, которые не входят в A.

---

[2] В соответствии с определением события, оно происходит в эксперименте, если осуществляется любой из благоприятствующих ему исходов. Одновременное осуществление событий понимается как реализация в эксперименте их общего исхода.

Обозначение для отрицания (противоположного события): $\overline{A}$.

Пример 6. В описанном выше эксперименте отрицанием события B будет событие, состоящее в извлечении любой карты, не являющейся королем.

Рассмотрим теперь событие, которое в дальнейшем будем обозначать буквой Ω, составленное из *всех* исходов рассматриваемого эксперимента. Очевидно, что при *любой* реализации эксперимента какой-нибудь исход обязательно осуществится, а следовательно, осуществится событие Ω, что дает основание назвать это событие *достоверным*. Оно происходит всегда, когда проводится эксперимент. Пополним множество *возможных* в рассматриваемом эксперименте событий событием *невозможным*, которое в данном эксперименте не происходит. Невозможное событие будем обозначать символом $\oslash$.

Легко понять, что если мы имеем дело со схемой случаев с N равновозможными исходами, то общее количество всех событий в рассматриваемом эксперименте равно 2^N.

◀ Действительно, событий, состоящих ровно из одного исхода, будет N, из двух исходов — C_N^2, и вообще, событий, состоящих из S исходов, будет C_N^S. Таким образом, число возможных событий равно

$$N + C_N^1 + C_N^2 + \ldots + C_N^S + \ldots + C_N^N;$$

если добавить теперь к этому количеству еще одно — невозможное — событие, получим искомый результат, так как известно[3], что

$$1 + N + C_N^1 + C_N^2 + \ldots + C_N^S + \ldots + C_N^N = (1+1)^N = 2^N,$$

откуда и следует искомое. ▶

Отметим несколько очевидных соотношений:

$$\forall A \quad A \cup \overline{A} = \Omega, \quad A\overline{A} = \oslash, \quad \overline{\overline{A}} = A,$$

$$\overline{\Omega} = \oslash, \quad \overline{\oslash} = \Omega.$$

Естественно под вероятностью *события* понимать величину, пропорциональную количеству входящих в него исходов — если некоторое событие составлено S исходами, то его вероятность положим равной

$$S \cdot P = S \cdot \frac{k}{N}. \tag{2}$$

При этом ясно, что чем больше исходов входят в событие (говорят — благоприятствуют осуществлению события), тем больше шансы на его осуществление, как следствие — тем больше его вероятность. Все события, осуществляющиеся в эксперименте, с точки зрения шансов на осуществление естественно располагаются между невозможным и достоверным событиями.

Вероятность невозможного события положим равной нулю, отмечая тем самым, что шансов на осуществление невозможного события нет.

Вероятность достоверного события может быть принята равной любому положительному числу — никаких запретов или ограничений на это значение нет. Так как достоверное событие включает все возможные исходы, то его вероятность больше вероятности любого другого события в этом эксперименте и равна, в силу (2), сумме вероятностей всех исходов, т. е. $P(\Omega) = N \cdot k/N = k$. Таким образом, коэффициент пропорциональности в соотношении (1) равен вероятности достоверного события.

[3] Здесь использована формула Ньютона для разложения бинома.

Поскольку выбор значения вероятности достоверного события не влияет на содержательную сторону описания возможности осуществления того или иного исхода в схеме случаев, а меняет только масштаб шкалы измерения вероятностей, положим

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

и, тем самым, завершим определение вероятностей исхода и события в схеме случаев.

Суммируя вышеизложенное, еще раз отметим, что все события, происходящие в эксперименте, могут быть естественным образом ранжированы в соответствии с их шансами на осуществление при однократной реализации эксперимента. В этой ранжировке они располагаются между невозможным событием, которое не происходит никогда, и достоверным, которое реализуется всегда, когда реализуется эксперимент.

Мерой осуществимости любого события A выступает его ***вероятность***, определяемая как отношение количества S благоприятствующих осуществлению события исходов к общему числу N всех возможных исходов:

$$\forall A \quad P(A) = \frac{S}{N}.$$

Отметим некоторые свойства вероятности события в схеме случаев.

1. Вероятность любого события, происходящего в рассматриваемом эксперименте, задается положительным числом, заключенным в пределах между 0 и 1

$$0 = P(\oslash) < P(A) < P(\Omega) = 1.$$

2. Если события A и B — несовместны, то вероятность суммы равна сумме вероятностей

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

в частности, справедлив так называемый *принцип дополнительности*

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

Рассмотрим несколько примеров, иллюстрирующих технику нахождения вероятностей событий в схеме случаев.

Пример 7. Пусть в урне лежит $m+n$ физически идентичных шаров, окрашенных соответственно в белый (n шаров) и черный (m шаров) цвета. Эксперимент состоит в извлечении из урны одного шара. Найдем вероятность извлечения шара белого цвета

◀ Поскольку всего возможных исходов $N = m + n$, а благоприятствующих извлечению белого шара — n, то искомая вероятность дается отношением $n/(m+n)$. ▶

Пример 8. В условиях предыдущего примера производится извлечение двух шаров. Какова вероятность того, что оба извлеченных шара — белые? Извлечены разноцветные шары?

◀ Ответ на первый вопрос задачи может быть получен, например, с помощью следующих рассуждений.

Всего различных пар шаров из урны, содержащей $m + n$ физически идентичных шаров, можно составить C^2_{m+n}, так что можно считать, что всего различных исходов в рассматриваемом эксперименте C^2_{m+n}. В то же время количество пар белых шаров дается числом C^2_n, откуда искомая вероятность равна

$$P(\text{б}, \text{б}) = \frac{C^2_n}{C^2_{m+n}} = \frac{n(n-1)}{(m+n)(m+n-1)}.$$

Аналогичные соображения для второго вопроса дают.

$$N = C^2_{m+n}, \quad S = mn, \quad P(\text{б}, \text{ч}) = \frac{mn}{C^2_{m+n}} = \frac{2mn}{(m+n)(m+n-1)}.$$ ▶

Заметим, что рассматриваемый эксперимент — извлечение пары шаров — эквивалентен двукратному последовательному извлечению шаров из урны.

Пример 9. В условиях предыдущего примера производится последовательное извлечение двух шаров *с возвращением каждого извлеченного шара обратно в урну*. Какова вероятность того, что оба извлеченных шара — белые? Извлечены разноцветные шары?

◀ В отличие от предыдущей ситуации, общее количество возможных исходов в рассматриваемом эксперименте будет уже равно $N = (m+n)^2$, а количество S исходов, благоприятствующих интересующему нас событию (оба шара — белые) — n^2, и, следовательно, искомая вероятность равна

$$P(\text{б}, \text{б}) = \left(\frac{n}{m+n}\right)^2.$$

Аналогичные рассуждения для второго вопроса дают:

$$N = (m+n)^2, \quad S = 2mn, \quad P(\text{б}, \text{ч}) = \frac{2mn}{(m+n)^2}. \ ▶$$

Пример 10. Ребенок, играя с четырьмя карточками разрезной азбуки, на которых изображены буквы **А, А, М, М**, случайным образом выкладывает их в ряд. Какова вероятность того, что у него получится слово **МАМА**?

◀ Предполагая, что все возможные расстановки четырех карточек в ряд равновозможны, получаем схему случаев с общим количеством исходов $N = 4! = 24$. Количество исходов, благоприятствующих интересующему нас событию, равно $S = 4$, откуда искомая вероятность равна $4/24 = 1/6$. ▶

Пример 11. Студент подготовил к экзамену 40 из 50 вопросов, охватывающих программу изученного курса. На экзамене ему предлагается дать ответ на два случайным образом выбранных из общего списка вопроса. Какова вероятность того, что студент знает ответ на оба предложенных ему вопроса?

◀ Легко видеть, что всего различных вариантов выбора пары различных вопросов из общего списка, содержащего 50 вопросов, будет $N = C_{50}^2$. Интересующее нас событие состоит из таких пар вопросов, оба из которых известны студенту. Их количество $S = C_{40}^2$. Таким образом, искомая вероятность дается отношением

$$P = \frac{C_{40}^2}{C_{50}^2} \approx 0{,}6. \ ▶$$

Пример 12. В урне лежит всего 10 черных и белых шаров. Из урны извлекают без возвращения пару шаров. Известно, что вероятности извлечения одноцветных шаров относятся как 2 : 5. Можно ли по этим данным установить состав шаров в урне?

◀ Пусть в урне лежит n белых (черных) (соответственно, $10 - n$ черных (белых)) шаров. Тогда вероятность извлечения из урны пары белых (черных) шаров будет равна

$$P(\text{б}, \text{б}) = \frac{n(n-1)}{10 \cdot 9},$$

аналогично, пары черных (белых) шаров

$$P(\text{ч}, \text{ч}) = \frac{(10-n)(9-n)}{10 \cdot 9}.$$

Из условия задачи следует

$$\frac{P(\text{б}, \text{б})}{P(\text{ч}, \text{ч})} = \frac{n(n-1)}{(10-n)(9-n)} = \frac{2}{5},$$

откуда для n получаем уравнение

$$n^2 + 11n - 60 = 0,$$

единственное целое положительное решение ($n = 4$) которого дает ответ — в урне возможно наличие 4 белых и 6 черных, либо 4 черных и 6 белых шаров. ▶

Пример 13. В урне лежит некоторое количество белых и черных шаров, так что вероятность извлечения пары белых шаров равна 0,5. Какое минимально возможное количество шаров находится в урне? Каков при этом состав шаров в урне?

◀ Пусть в урне лежит всего N шаров, из которых $n < N$ — белые. Вероятность извлечения пары белых шаров (см. предыдущий пример) дается соотношением

$$P(\text{б}, \text{б}) = \frac{n(n-1)}{N(N-1)},$$

и условие задачи приводит к уравнению, связывающему N и n

$$\frac{n(n-1)}{N(N-1)} = \frac{1}{2}.$$

Учитывая, что величины n и N — целые положительные числа, удовлетворяющие условию $n < N$, выразим из этого соотношения n через N

$$n = \frac{1 + \sqrt{1 + 2N(N-1)}}{2}$$

и, придавая величине N последовательно значения $1, 2, \ldots$, найдем, что наименьшее значение N, при котором n — целое положительное, равняется 4. Значение n при этом равно 3. ▶

Пример 14. Среди выпущенных N лотерейных билетов n — выигрышных. Некто приобрел $r < N - n$ лотерейных билетов. Какова вероятность того, что среди них по крайней мере один выигрышный?

◀ Заметим, что событие $A = \{$среди приобретенных r билетов по крайней мере один выигрышный$\}$ противоположно событию $\overline{A} = \{$среди приобретенных r билетов выигрышных нет$\}$. Для решения задачи воспользуемся принципом дополнительности: $P(A) = 1 - P(\overline{A})$.

Найдем вероятность $P(\overline{A})$. Всего возможных исходов (т. е. различных наборов из r лотерейных билетов) C_N^r, среди них таких, которые не содержат ни одного выигрышного билета — C_{N-n}^r. Для вероятности $P(\overline{A})$ получаем

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

откуда искомая вероятность дается соотношением

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

Пример 15. Среди выпущенных N лотерейных билетов n — выигрышных. Некто приобрел $r < \min\{n, N - n\}$ лотерейных билетов. Какова вероятность того, что среди них ровно k выигрышных?

◀ Как и выше, всего возможных исходов (т. е. различных наборов из r лотерейных билетов) будет C_N^r. Набор, содержащий ровно k выигрышных билетов, образуется в результате объединения любых k выигрышных билетов с любыми $r - k$ невыигрышными. Количество различных наборов из k выигрышных билетов равно C_n^k, количество различных наборов из $r - k$ невыигрышных билетов равно C_{N-n}^{r-k}. Следовательно, количество различных комбинаций из выигрышных и невыигрышных билетов дается числом $C_n^k \cdot C_{N-n}^{r-k}$, а искомая вероятность равна

$$P = \frac{C_n^k \cdot C_{N-n}^{r-k}}{C_N^r}. \blacktriangleright$$

§2. Геометрические вероятности

Другая схема описания экспериментов с неоднозначно прогнозируемыми исходами, которая позволяет довольно просто ввести количественную характеристику осуществимости того или иного события — это схема *геометрических вероятностей*, которая, как и рассмотренная выше схема случаев, эксплуатирует идею о *равновозможности* исходов эксперимента.

Аналогично тому, как это было проделано в схеме случаев, количественная характеристика осуществимости события — его вероятность — определяется как нормированная некоторым образом величина, пропорциональная запасу исходов, благоприятствующих осуществлению события.

Пусть множество исходов исследуемого эксперимента может быть описано как множество Ω точек некоторого «геометрического континуума» — каждому исходу соответствует некоторая точка и каждой точке отвечает некоторый исход. В качестве «геометрического континуума» Ω может выступать отрезок на прямой, дуга спрямляемой кривой на плоскости или в пространстве, квадрируемое множество на плоскости (треугольник, прямоугольник, круг, эллипс и т. п.) или часть квадрируемой поверхности, некоторый объем в пространстве (многогранник — призма, пирамида, шар, эллипсоид и т. п.)

Событием назовем любое квадрируемое подмножество множества Ω.

Как и в схеме случаев, событие состоит из точек-исходов, однако уже не любая совокупность исходов образует событие, а только такая, меру которой (длину, площадь, объем) мы можем измерить.

Предполагая равновозможность исходов, назовем *вероятностью события A* число, пропорциональное мере подмножества A множества Ω:

$$P(A) = k \cdot \mu(A).$$

Если $\oslash$ — событие, невозможное в данном эксперименте, а Ω — достоверное, то положим $P(\oslash) = 0$, $P(\Omega) = 1$. Вероятность любого события A будет при этом заключена между нулем — вероятностью события невозможного, и единицей — вероятностью события достоверного[4]. Условие нормировки позволяет найти константу k — коэффициент пропорциональности, задающий вероятность. Он оказывается равен $1/\mu(\Omega)$.

Таким образом, в схеме геометрических вероятностей *вероятность любого события определяется как отношение меры подмножества A, описывающего событие, к мере множества Ω, описывающего эксперимент в целом*:

$$P(A) = \frac{\mu(A)}{\mu(\Omega)}. \tag{3}$$

Отметим некоторые свойства так определенной вероятности:

1. $\forall A \subset \Omega \quad 0 \leqslant P(A) \leqslant 1.$

2. Если $A \subseteq B$, то $P(A) \leqslant P(B)$.

◀ Свойство очевидно следует из того обстоятельства, что множество, содержащееся внутри другого, не может быть больше последнего. ▶

Как и в схеме случаев, события в схеме геометрических вероятностей можно объединять, совмещать и строить на их основе противоположные — при этом будут получаться, вообще говоря, отличные от исходных события. Следующее свойство весьма важно.

3. Если события A и B — несовместны, то $P(A \cup B) = P(A) + P(B)$, в частности, справедлив *принцип дополнительности*: $P(\overline{A}) = 1 - P(A)$.

◀ Это свойство, называемое обычно *правилом сложения вероятностей*, очевидно следует из аддитивности меры[5]. ▶

В заключение отметим, что вероятность осуществления любого исхода в схеме геометрических вероятностей всегда равна нулю, равно как равна нулю вероятность любого события, описываемого «тощим» множеством точек, т. е. множеством, мера которого (соответственно — длина, площадь, объем) равна нулю.

Рассмотрим несколько примеров, иллюстрирующих вычисление вероятностей в схеме геометрических вероятностей.

Пример 1. Эксперимент состоит в случайном выборе точки из отрезка $[a, b]$. Найти вероятность того, что выбрана точка, лежащая в левой половине рассматриваемого отрезка.

◀ По определению, вероятность выбора точки из любого множества на отрезке $[a, b]$ пропорциональна длине этого множества. Следовательно, искомая вероятность равна 0,5:

$$P\left(x \in \left[a, \frac{a+b}{2}\right]\right) = \frac{l([a, (a+b)/2])}{l(a, b)} = \frac{(b-a)/2}{b-a} = \frac{1}{2}. \text{ ▶}$$

---

[4] Однако, в отличие от схемы случаев, здесь уже *возможны* события, имеющие нулевую вероятность, и могут *не быть достоверными* события с вероятностью единица.

[5] Мера (длина, площадь, объем) суммы непересекающихся множеств равна сумме их мер.

Пример 2. Эксперимент состоит в случайном выборе точки $M(\xi, \eta)$ из квадрата $K = \{-1 \leqslant \xi, \ \eta \leqslant 1\}$. Какова вероятность того, что уравнение

$$x^2 + 2\xi \cdot x + \xi \cdot \eta = 0$$

имеет действительные корни? Равные корни?

◀ Хорошо известно, что у квадратного уравнения корни действительны, если его дискриминант неотрицателен. В рассматриваемом случае дискриминант D дается соотношением

$$D = \xi^2 - \xi \cdot \eta$$

и будет неотрицателен, если (ξ, η) удовлетворяют условию

$$A\colon \ \begin{cases} \xi \geqslant 0, \\ \xi \geqslant \eta \end{cases} \bigcup \begin{cases} \xi \leqslant 0, \\ \xi \leqslant \eta, \end{cases}$$

т е если точка $M(\xi, \eta)$ будет выбрана из множества A, являющегося пересечением квадрата K и множества точек, описываемого вышеприведенными условиями (рис. 1).

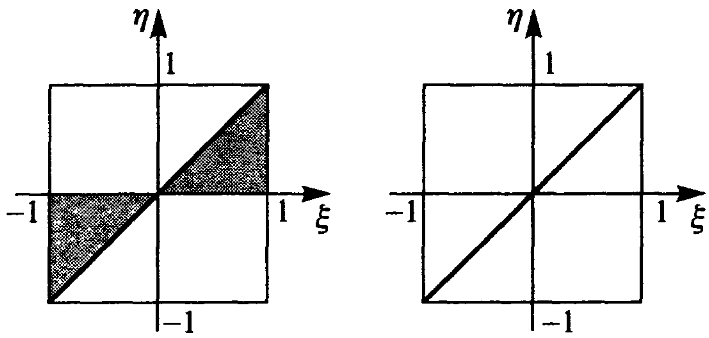


Рис. 1. Случайный выбор точки из квадрата K

Следовательно, для искомой вероятности получаем:

$$P(\text{корни действительны}) = \frac{S(A)}{S(K)} = \frac{1}{4}.$$

Далее, корни квадратного уравнения совпадают, если $D = 0$. Этому значению дискриминанта отвечает отрезок оси $O\xi$ от -1 до $+1$ и отрезок биссектрисы первого и третьего координатного угла, лежащий внутри квадрата K (рис. 1). Легко понять, что площадь этого множества точек равна нулю и, следовательно, вероятность совпадения корней рассматриваемого уравнения равна нулю. ▶

Следующий пример является классическим[6] и призван проиллюстрировать то простое соображение, что понятие «случайности» не является очевидным и одинаково понимаемым всеми, а потому должно быть, вообще говоря, аккуратно формализовано, иначе использование вероятностных соображений может привести к недоразумениям.

Пример 3. В круге радиуса R случайным образом выбрана хорда. Какова вероятность того, что длина этой хорды больше радиуса?

◀ В первую очередь следует понять, что значит *хорда выбрана случайно*.

1. Поскольку длина хорды однозначно определяется расстоянием этой хорды от центра круга, то одна из возможных интерпретаций случайного выбора может выглядеть так:

Случайный выбор хорды эквивалентен случайному выбору точки на диаметре круга.

Длина хорды, находящейся на расстоянии d от центра, равна $2\sqrt{R^2 - d^2}$, и для того чтобы длина хорды превышала длину радиуса круга, нужно, чтобы выбранная точка была расположена от центра круга на расстоянии, не преаышающем $d = R\sqrt{3}/2$ (рис. 2)

Поэтому искомая вероятность раана

$$P = \frac{R\sqrt{3}}{2R} = \frac{\sqrt{3}}{2}.$$

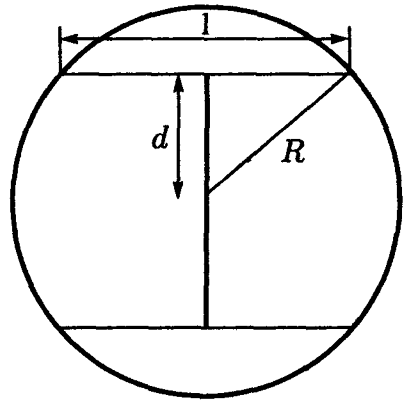


Рис. 2. Задача Бертрана — первый способ

[6] Вероятно впервые эта задача обсуждалась в книге Ж. Л. Бертрана «Исчисление вероятностей», вышедшей в 1889 году. Заметим, что помимо рассмотренных ниже, существует еще по крайней мере три различных возможности формализации данной задачи.

2 Всякая хорда может быть задана парой точек на окружности, являющихся ее концами. Поэтому другая интерпретация случайного выбора хорды может быть сформулирована так.

Случайный выбор хорды эквивалентен случайному выбору пары точек на дуге окружности.

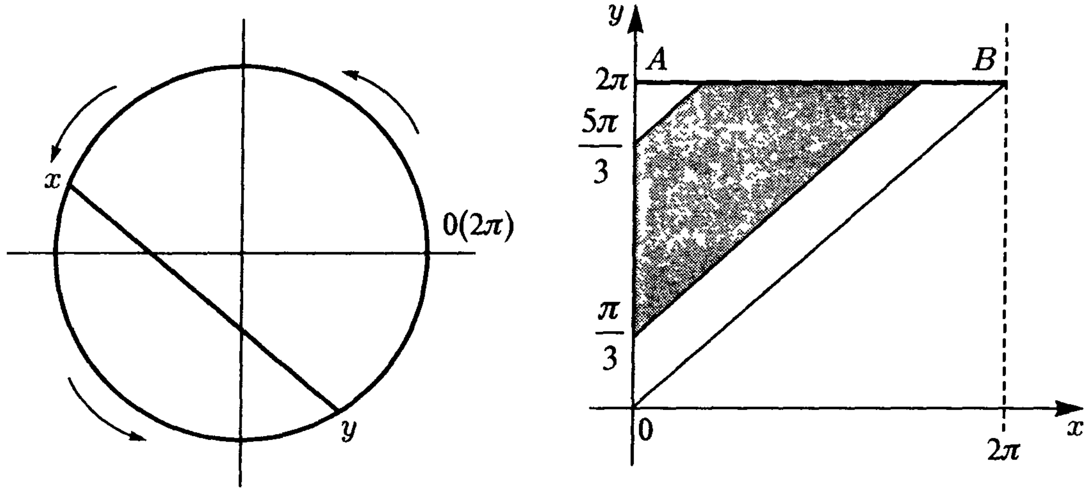


Рис. 3. Задача Бертрана — второй способ

Выбирая на окружности начало отсчета и задавая направление обхода (например, против часовой стрелки), пометим положение любой точки на окружности ее координатой, меняющейся в пределах от 0 до 2π. Множество хорд может быть описано множеством упорядоченных пар чисел (x, y), $0 \leqslant x \leqslant y \leqslant 2\pi$ — координат начала и конца каждой из хорд (рис. 3, слева). Это множество на плоскости координат (x, y) изображается треугольником OAB (рис. 3, справа). Понятно, что длина хорды будет больше радиуса, если координаты начала и конца хорды удовлетворяют условиям

$$y - x \geqslant \frac{\pi}{3} \cup 2\pi - y + x \geqslant \frac{\pi}{3}.$$

Последние могут быть записаны одним двойным неравенством

$$x + \frac{\pi}{3} \leqslant y \leqslant x + \frac{5\pi}{3}.$$

Множество точек (x, y), удовлетворяющих этому условию, заштриховано на рис. 3 (справа). Теперь легко находим

$$S_{\text{штрих}} = \frac{4\pi^2}{3}, \quad S_{OAB} = 2\pi^2,$$

и искомая вероятность равна

$$P = \frac{S_{\text{штрих}}}{S_{OAB}} = \frac{2}{3},$$

что отличается от результата, полученного выше. ▶

§ 3. Условные вероятности. Взаимное влияние и независимость

Информация о реализации некоторого события в эксперименте может менять наши представления о шансах на осуществление других событий.

Пример 1. Пусть в эксперименте с бросанием симметричной монеты рассматриваются события Г — выпадение герба и Р — выпадение решки. Очевидно, что если нам известно: выпал герб, т. е. осуществилось событие Г, то осуществление события Р — выпадение решки в этом эксперименте невозможно.

Пример 2. Если в эксперименте с извлечением карты извлечена карта $K♠$, то очевидно, что одновременно осуществились события ♠ $= \{$извлечена карта масти пик$\}$ и $K = \{$извлечен король$\}$. Другими словами, осуществление события $K♠$ влечет за собой осуществление и этих событий.

Но, конечно, может оказаться и так, что осуществление одного из событий в эксперименте ничего не говорит нам об осуществлении или неосуществлении другого, точнее, не меняет наших представлений о шансах на его осуществление.

Пример 3. Рассмотрим эксперимент, состоящий в двукратном извлечении шаров из урны с последующим возвращением извлеченного шара обратно в урну. Пусть в урне лежит $N = m + n$ соответственно черных (m) и белых (n) шаров. Рассмотрим события: А — шар, извлеченный первым, белый, В — шар, извлеченный вторым, белый. Поскольку после каждого извлечения шар возвращается в урну, то ясно, что зависимости между этими событиями нет.

Из общих соображений понятно, что при условии осуществления одного из событий шансы на осуществление другого должны быть пропорциональны запасу их общих исходов[7] — чем значительнее общая часть рассматриваемых событий, тем выше должны быть шансы на осуществление одного из них, в предположении, что другое произошло.

Введем соответствующее формальное понятие.

Условной вероятностью осуществления события A относительно события B назовем число

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

где $S(A \cap B)$, $S(B)$ — запас исходов эксперимента, благоприятствующих осуществлению соответственно событий $A \cap B$ и B.

Пусть событие B фиксировано и таково, что $P(B) > 0$. Тогда условная вероятность обладает следующими очевидными свойствами:

1. $\forall A \ \ 0 \leqslant P(A|B) \leqslant 1$.

2. Если события A и B — несовместны, то $P(A|B) = 0$, если же события A и B таковы, что B составляет часть A[8], то $P(A|B) = 1$, в частности $P(B|B) = 1$.

3. Для условных вероятностей справедливо правило сложения

$$P(A_1 \cup A_2 | B) = P(A_1|B) + P(A_2|B),$$

если только события A_1 и A_2 несовместны.

Таким образом, условная вероятность обладает всеми свойствами вероятности и описывает шансы на осуществление события A при уже происшедшем событии B. Очевидно, что, вообще говоря, $P(A|B) \neq P(B|A)$.

Сразу же заметим, что условная вероятность может быть вычислена как отношение вероятности совместного осуществления событий A и B к вероятности события-условия B:

$$P(A|B) = \frac{S(A \cap B)}{S(B)} = \frac{S(A \cap B)}{S(B)} \cdot \frac{S(\Omega)}{S(\Omega)} = \frac{\dfrac{S(A \cap B)}{S(\Omega)}}{\dfrac{S(B)}{S(\Omega)}} = \frac{P(A \cap B)}{P(B)}.$$

Из последнего соотношения следует *правило умножения вероятностей*

$$P(A \cap B) = P(AB) = P(A|B)P(B) = P(B|A)P(A), \tag{4}$$

справедливое для событий с положительной вероятностью.

Рассмотрим несколько примеров, иллюстрирующих введенное понятие.

Пример 4. Из урны, содержащей n белых и m черных шаров, извлекают без возвращения пару шаров. Какова вероятность извлечь вторым черный шар, если известно, что первым был извлечен черный?

◀ Очевидно, что если первым был извлечен черный шар, то в урне осталось всего $n+m-1$ шаров, среди которых черных $m-1$. Поэтому искомая вероятность равна $\dfrac{m-1}{n+m-1}$. ▶

---

[7] Под *запасом исходов* мы понимаем здесь *количество исходов* в схеме случаев либо *меру (длину, площадь, объем) соответствующего множества исходов* в схеме геометрических вероятностей.

[8] В этом случае говорят, что событие A *подчинено* событию B и записывают это так: $A \supseteq B$.

Пример 5. Эксперимент состоит в случайном выборе точки из квадрата $K = \{(x, y)\colon\ 0 \leqslant x, y \leqslant 1\}$. Найти вероятность того, что первая координата точки не превышает 0,5, если известно, что выбрана точка, лежащая выше биссектрисы первого и третьего координатных углов (рис. 4).

◀ Из рисунка легко усмотреть, что вероятность события $B = \{$выбрана точка, лежащая выше биссектрисы$\}$ равна 0,5, а вероятность совместного осуществления событий B и $A = \{x \leqslant 0{,}5\}$ — $3/8$. Отсюда для искомой вероятности получаем: $P(A \mid B) = 3/4$. ▶

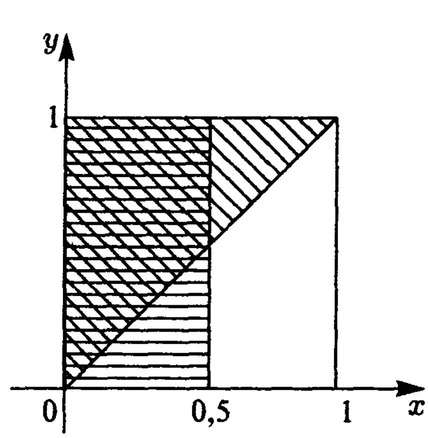


Рис. 4. Случайный выбор точки из квадрата

Понятие условной вероятности позволяет ввести также количественную меру, характеризующую степень влияния одного из событий на другое.

Будем говорить, что событие A не зависит от события B, если осуществление события A не меняет вероятности осуществления события B, т. е. если условная вероятность $P(A|B)$ совпадает с безусловной $P(A)$:

$$P(A|B) = P(A), \qquad P(B) \neq 0.$$

В противном случае будем говорить, что событие A зависит от B.

Сразу же отметим, что понятия зависимости-независимости, несмотря на явную несимметричность определения, носят взаимный характер — если событие A зависит (не зависит) от события B, то и событие B зависит (не зависит) от события A.

◀ Действительно, пусть событие A не зависит от события B. Рассмотрим

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{P(A|B)P(B)}{P(A)},$$

но, в силу независимости, $P(A|B) = P(A)$, откуда и следует независимость B от A. ▶

В случае независимых событий правило умножения вероятностей принимает особенно простой вид: *вероятность совместного осуществления двух событий равна произведению их вероятностей:*

$$P(AB) = P(A)P(B). \tag{5}$$

Соотношение (5) может быть принято в качестве определения независимости.

Нижеследующие примеры иллюстрируют использование правила умножения при вычислении вероятностей событий.

Пример 6. Из урны, содержащей n белых и m черных шаров, извлекают три шара. Какова вероятность того, что среди извлеченных есть хотя бы один белый шар?

◀ Заметим, что интересующее нас событие A противоположно событию — все извлеченные шары черные. В соответствии с принципом дополнительности $P(A) = 1 - P(\overline{A})$. Вероятность события $\overline{A}$ найдем, воспользовавшись тем, что $\overline{A} = 1_\text{ч} \cap 2_\text{ч} \cap 3_\text{ч}$, где события $i_\text{ч}$ означают, что шар, извлеченный i-м — черный. В соответствии с правилом умножения получаем

$$P(\overline{A}) = P(3_\text{ч} \mid 1_\text{ч} \cap 2_\text{ч})P(1_\text{ч} \cap 2_\text{ч}) = P(1_\text{ч})P(2_\text{ч} \mid 1_\text{ч})P(3_\text{ч} \mid 1_\text{ч} \cap 2_\text{ч}).$$

Для вероятностей, участвующих в этом соотношении, легко получаем

$$P(1_\text{ч}) = \frac{m}{m+n}, \quad P(2_\text{ч} \mid 1_\text{ч}) = \frac{m-1}{m+n-1}, \quad P(3_\text{ч} \mid 1_\text{ч} \cap 2_\text{ч}) = \frac{m-2}{m+n-2},$$

откуда ответ

$$P(A) = 1 - \frac{m}{m+n} \cdot \frac{m-1}{m+n-1} \cdot \frac{m-2}{m+n-2}. \ ▶$$

Пример 7. В круге радиуса R случайным образом независимо друг от друга выбрано N точек. Найти вероятность того, что расстояние от центра круга до ближайшей из них будет не менее r.

◀ Ясно, что если ближайшая из точек находится от центра на расстоянии не меньшем r, то и все прочие будут находиться от центра на не меньшем расстоянии.

Вероятность того, что случайная в круге точка находится от центра на расстоянии не меньшем r, дается отношением

$$P(d \geqslant r) = \frac{\pi R^2 - \pi r^2}{\pi R^2} = 1 - \frac{r^2}{R^2}.$$

В соответствии с правилом умножения (5), искомая вероятность равна

$$P(d_1 \geqslant r \cap d_2 \geqslant r \cap \ldots d_N \geqslant r) = P(d_1 \geqslant r) \ldots P(d_N \geqslant r) = \left(1 - \frac{r^2}{R^2}\right)^N. ►$$

Пример 8. Некто забыл последнюю цифру номера телефона и набирает ее наугад. Какова вероятность того, что он дозвонится до нужного абонента не более чем за три попытки?

◄ Воспользуемся принципом дополнительности — противоположным рассматриваемому событию A будет событие $\overline{A}$, состоящее в том, что первые три попытки дозвониться до нужного абонента оказались безуспешными. Последовательные попытки дозвониться до нужного абонента $\overline{A}_1, \overline{A}_2, \overline{A}_3$ — зависимые события, так как однажды набранная и не принесшая успеха цифра в дальнейшем уже не набирается. Применим праило умножения (4):

$$P(\overline{A}) = P(\overline{A}_1 \cdot \overline{A}_2 \cdot \overline{A}_3) = P(\overline{A}_3 | \overline{A}_1 \cdot \overline{A}_2) P(\overline{A}_1 \cdot \overline{A}_2) = P(\overline{A}_3 | \overline{A}_1 \cdot \overline{A}_2) P(\overline{A}_2 | \overline{A}_1) P(\overline{A}_1).$$

Для сомножителей очевидно имеем

$$P(\overline{A}_1) = \frac{9}{10}, \quad P(\overline{A}_2 | \overline{A}_1) = \frac{8}{9}, \quad P(\overline{A}_3 | \overline{A}_1 \cdot \overline{A}_2) = \frac{7}{8}.$$

Отсюда

$$P(\overline{A}) = \frac{9}{10} \cdot \frac{8}{9} \cdot \frac{7}{8},$$

и для искомой вероятности получаем $P(A) = 1 - P(\overline{A}) = 0{,}3$. ►

Пример 9[9]. Исследовать связь между темным цветом глаз у отца (событие T_o) и сына (событие T_c) на основании следующих данных, полученных при переписи населения Англии и Уэльса в 1891 году.

Темноглазые отцы и темноглазые сыновья (T_oT_c) составляли 5% среди всех обследованных, темноглазые отцы и светлоглазые сыновья ($T_o\overline{T}_c$) — 7,9%, светлоглазые отцы и темноглазые сыновья ($\overline{T}_oT_c$) — 8,9%, светлоглазые отцы и светлоглазые сыновья ($\overline{T}_o\overline{T}_c$) — 78,2%.

◄ Для оценки исследуемой связи найдем условные вероятности $P(T_c | T_o)$ и $P(T_c | \overline{T}_o)$ и сравним их с соответствующей безусловной $P(T_c)$.

По определению имеем

$$P(T_c | T_o) = \frac{P(T_oT_c)}{P(T_o)}.$$

Условия задачи дают основания для следующей оценки вероятностей[10]

$$P(T_oT_c) = 0{,}05, \quad P(T_o\overline{T}_c) = 0{,}079, \quad P(\overline{T}_oT_c) = 0{,}089, \quad P(\overline{T}_o\overline{T}_c) = 0{,}782.$$

Поскольку очевидно, что $T_oT_c + T_o\overline{T}_c = T_o$, постольку $P(T_o) = P(T_oT_c) + P(T_o\overline{T}_c) = 0{,}129$. Отсюда

$$P(T_c | T_o) = \frac{P(T_oT_c)}{P(T_o)} = \frac{0{,}05}{0{,}129} = 0{,}3876.$$

В то же время $T_oT_c + \overline{T}_oT_c = T_c$ и $P(T_c) = P(T_oT_c) + P(\overline{T}_oT_c) = 0{,}139$. Сравнивая значения условной и безусловной вероятностей, делаем заключение о наличии связи между темным цветом глаз у отца и темным цветом глаз у сына — у темноглазых отцов темноглазые сыновья встречаются почти втрое чаще, чем вообще среди обследованных. Заметим, между прочим, что светлоглазые сыновья у темноглазых отцов астречаются примерно а 6 случаях из 10 ($P(\overline{T}_c | T_o) = 1 - P(T_c | T_o) = 0{,}6124$).

Подсчитаем теперь вероятность $P(T_c | \overline{T}_o)$. Рассуждения, аналогичные вышеприведенным, дают:

$$P(T_c | \overline{T}_o) = \frac{P(\overline{T}_oT_c)}{P(\overline{T}_o)} = 0{,}1022.$$

Заключаем, что светлоглазые отцы, вообще говоря, могут иметь темноглазых сыновей, однако значительно реже, чем светлоглазых — примерно в одном случае из 10 у светлоглазых отцов темноглазые сыновья и, соответственно, в 9 случаях из 10 — светлоглазые. ►

---

[9] Задача заимствована из задачника Л. Д. Мешалкина.

[10] Как обычно, полагаем вероятность пропорциональной запасу случаев.

§ 4. Формула полной вероятности

Введенные в предыдущем разделе понятия условной вероятности и независимости событий позволяют получить простое соотношение, облегчающее вычисление вероятностей в многоальтернативных ситуациях — когда событие, вероятность которого отыскивается, может происходить совместно с другими событиями, относительно которых подсчет вероятностей интересующего нас события оказывается по каким-то причинам проще. Это соотношение носит название *формулы полной вероятности.*

Пусть события $H_1, H_2, \ldots, H_N$ имеют ненулевую вероятность, несовместны и вместе исчерпывают все возможные исходы эксперимента:

$$H_i \cap H_j = \varnothing, \quad i \neq j, \quad H_1 + H_2 + \ldots + H_N = \Omega.$$

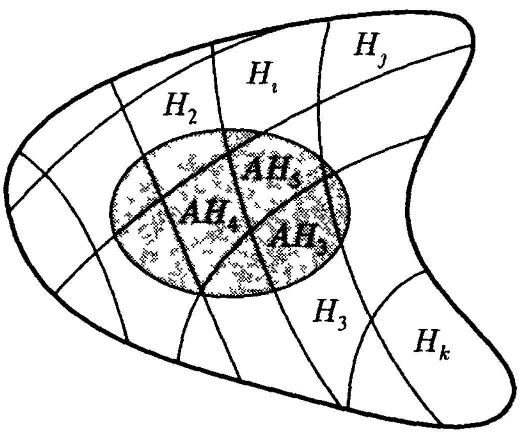


Рис 5 Разбиение полной системой несовместных событий

Совокупность событий, обладающих перечисленными свойствами, задает альтернативное разбиение множества всех исходов эксперимента и обычно называется *полной группой несовместных событий.*

Если A — некоторое событие, то очевидно, что разбиение эксперимента, задаваемое полной группой событий H_i, $i = 1, \ldots, N$, задает и разбиение множества исходов, образующих событие A (рис. 5):

$$A = AH_1 + AH_2 + \ldots + AH_N,$$

(при этом, конечно, некоторые слагаемые в приведенной сумме могут оказаться невозможными событиями).

Поскольку события H_i, $i = 1, \ldots, N$ — несовместны, то и события AH_i, $i = 1, \ldots, N$ — также несовместны, и по правилу сложения заключаем, что

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

Правило же умножения (4) позволяет вычислить каждое из слагаемых

$$P(AH_i) = P(A|H_i)P(H_i), \quad i = 1, 2, \ldots, N.$$

Объединяя два последних соотношения, получаем искомую формулу

$$P(A) = P(A|H_1)P(H_1) + P(A|H_2)P(H_2) + \ldots + P(A|H_N)P(H_N). \tag{6}$$

Пример 1. На книжной полке стоит два десятка книг, из которых 4 уже прочитаны хозяином, а оставшиеся еще нет Хозяин выбирает случайным образом книгу и читает (или перечитывает) ее, после чего ставит обратно на полку После этого он выбирает наугад очередную книгу. Какова вероятность того, что вновь выбранная книга еще не была прочитана?

◄ Рассуждения выглядят следующим образом пусть H_1 — событие, состоящее в том, что первая книга читанная, H_2 — нечитанная, A — вновь выбранная книга еще не была прочитана

Легко установить, что имеют место следующие соотношения

$$P(A|H_1) = \frac{16}{20} = \frac{4}{5}, \qquad P(A|H_2) = \frac{15}{20} = \frac{3}{4},$$
$$P(H_1) = \frac{4}{20} = \frac{1}{5}, \qquad P(H_2) = \frac{16}{20} = \frac{4}{5},$$

и формула (6) для искомой вероятности дает

$$P(A) = \frac{4}{5} \cdot \frac{1}{5} + \frac{3}{4} \cdot \frac{4}{5} = \frac{19}{25}. \blacktriangleright$$

Пример 2. На перегоне между двумя остановками в автобусе едут 3 пассажира, каждый из которых, независимо от прочих, с вероятностью 0,1 покидает автобус на ближайшей остановке. На этой остановке ожидают транспорт 3 пассажира, каждый из которых, независимо от прочих ожидающих, садится в подошедший автобус с вероятностью 0,3. Какова вероятность того, что после отправления с этой остановки количество пассажиров в салоне автобуса не изменится?

◀ Очевидно, что количество пассажиров в салоне автобуса останется неизменным, если количество вышедших будет равно количеству вошедших.

Пусть события H_0, H_1, H_2, H_3 состоят соответственно в том, что в автобус не сел никто, сел ровно один, два или три пассажира, событие A — количество пассажиров в салоне автобуса не изменилось. По формуле полной вероятности получаем

$$P(A) = P(A\,|\,H_0)P(H_0) + P(A\,|\,H_1)P(H_1) + P(A\,|\,H_2)P(H_2) + P(A\,|\,H_3)P(H_3).$$

Учитывая независимость посадки пассажиров в автобус, легко находим

$$P(H_0) = (1-0{,}3)^3 = (0{,}7)^3 = 0{,}343, \qquad P(H_1) = 3(1-0{,}3)^2 0{,}3 = 0{,}441,$$

$$P(H_2) = 3(1-0{,}3)0{,}3^2 = 0{,}189, \qquad P(H_3) = 0{,}3^3 = 0{,}027.$$

Заметим, что условные вероятности $P(A\,|\,H_i)$ есть вероятности того, что из автобуса вышло на остановке ровно i пассажиров.

$$P(A\,|\,H_0) = (1-0{,}1)^3 = 0{,}9^3 = 0{,}729, \qquad P(A\,|\,H_1) = 3(1-0{,}1)^2 0{,}1 = 0{,}243,$$

$$P(A\,|\,H_2) = 3(1-0{,}1)0{,}1^2 = 0{,}027, \qquad P(A\,|\,H_3) = 0{,}1^3 = 0{,}001.$$

Окончательно

$$P(A) = 0{,}729\cdot 0{,}343 + 0{,}243\cdot 0{,}441 + 0{,}027\cdot 0{,}189 + 0{,}001\cdot 0{,}027 \approx 0{,}36.$$ ▶

Формула полной вероятности вместе с формулой условной вероятности позволяет выносить некие суждения о «правдоподобности» гипотез:

Если событие A может происходить в эксперименте совместно с одним из альтернативных событий H_i, $i = 1, 2, \ldots, n$, и в результате эксперимента это событие осуществилось, то можно попытаться ответить на вопрос — с каким именно из событий H_i оно произошло вместе. Для этого оценим условную вероятность события (гипотезы) H_i при условии, что событие A реализовалось:

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

В соответствии с правилом умножения, вероятность, стоящая в числителе, дается выражением $P(A \cap H_i) = P(A\,|\,H_i)P(H_i)$, а стоящая в знаменателе может быть подсчитана с помощью формулы полной вероятности (6), что дает

$$P(H_i\,|\,A) = \frac{P(A\,|\,H_i)P(H_i)}{P(A)} = \frac{P(A\,|\,H_i)P(H_i)}{\sum P(A\,|\,H_j)P(H_j)}. \qquad (7)$$

Сравнивая вероятности $P(H_i\,|\,A)$ для различных значений i, будем считать ту из гипотез наиболее вероятной, для которой эта вероятность наибольшая.

Формула (7) называется формулой Бейеса, вероятности $P(H_j)$ — *априорными*, т. е. доопытными вероятностями, вероятности $P(H_j\,|\,A)$ — *апостериорными*, т. е. послеопытными вероятностями.

Пример 3. Из урны, в которой находится 4 белых и 6 черных физически идентичных шаров, извлекли наугад один шар и положили его в урну, содержащую 5 белых и 4 черных шаров. Случайно извлеченный из второй урны шар оказался черным. Что более вероятно — из первой урны был извлечен черный шар или белый?

◀ Пусть событие $H_б$ состоит в том, что из первой урны извлечен белый шар, $H_ч$ — из первой урны извлечен черный шар, Ч — из второй урны извлечен черный шар.

Очевидно, что

$$P(Ч) = P(Ч|H_б)P(H_б) + P(Ч\,|\,H_ч)P(H_ч) = \frac{4}{10}\cdot\frac{4}{10} + \frac{5}{10}\cdot\frac{6}{10} = \frac{46}{100}.$$

Для оценки апостериорной вероятности $P(H_6 \,|\, Ч)$ воспользуемся формулой (7):

$$P(H_6 \,|\, Ч) = \frac{P(Ч \,|\, H_6)\,P(H_6)}{P(Ч)} = \frac{16}{46}.$$

Полученный результат позволяет считать гипотезу о том, что из первой урны был извлечен белый шар менее предпочтительной в сравнении с гипотезой, предполагающей извлечение из первой урны черного шара. ▶

Для содержательного заключения о правдоподобности той или иной гипотезы важно, чтобы рассматриваемые события были действительно *случайными* в контексте рассматриваемых проблем. В противном случае выводы могут оказаться неадекватными реальному положению дел.

Пример 4. В одном из телевизионных шоу ведущий предлагает игроку выбрать один из стоящих перед ним ларцов, предупреждая, что только в одном из ларцов заключен ценный приз (скажем, ключи от автомобиля). После того как игрок произвел выбор, ведущий открывает один из оставшихся ларцов и, демонстрируя, что в нем ничего нет, предлагает игроку еще раз подумать и, если захочется, изменить свое решение — выбрать оставшийся ларец, вместо того, который был выбран первоначально. Имеет ли смысл игроку менять свое решение?

◀ Ответ на вопрос задачи зависит от того, случаен или нет выбор ведущим одного из ларцов для демонстрации его содержимого.

1. Пусть выбор ведущего случаен. Обозначим через H_1 событие, состоящее в том, что ларец, выбранный игроком, содержит приз, через H_2 — что он приза не содержит, через A — что ларец, выбранный ведущим, пуст. Тогда по формуле полной вероятности заключаем, что

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

А формула Бейеса (7) дает следующую оценку, например, для вероятности $P(H_1|A)$:

$$P(H_1 \,|\, A) = \frac{P(A \,|\, H_1)\,P(H_1)}{P(A)} = \frac{1/3}{2/3} = \frac{1}{2},$$

и, следовательно, в этом случае игроку нет нужды менять свой выбор, так как шансы его на получение приза одинаковы, остановится ли он на своем первоначальном выборе или сменит его.

2. Пусть выбор ведущего не случаен, и он, зная где лежит приз, всегда открывает для всеобщего обозрения ларец, приза не содержащий, т. е. событие A не является случайным и происходит с вероятностью 1 В этом случае по формуле условной вероятности заключаем, что

$$P(H_1 \,|\, A) = \frac{P(AH_1)}{P(A)} = P(H_1) = \frac{1}{3},$$

и в двух случаях из трех игроку выгоднее изменить свой выбор, чем настаивать на первоначальном[11]. ▶

---

[11] Впрочем, формальное использование формулы Бейеса дает тот же результат:

$$P(H_1 \,|\, A) = \frac{P(A \,|\, H_1)P(H_1)}{P(A)} = \frac{1/3}{1} = \frac{1}{3}.$$

Глава XXXVIII

СЛУЧАЙНЫЕ СОБЫТИЯ. ВЕРОЯТНОСТЬ

§1. Эксперимент. Исходы эксперимента. События

1.1. Эксперимент. Элементарные исходы эксперимента

Всякое изучаемое в дальнейшем явление будем называть *экспериментом*. Для дальнейшего не важна его содержательная сторона эксперимента (предметом рассмотрения может быть физическое, химическое, социальное или иное явление), а важно то обстоятельство, что комплекс условий, обуславливающий рассматриваемое явление, может быть (по крайней мере, в принципе) воспроизведен сколько угодно раз. Важно также, что a priori возможно перечисление всех событий, которые могут наступить в случае осуществления указанного выше комплекса условий. В отдельных случаях эти события могут наступать или не наступать в разных комбинациях. Во множество исходов эксперимента включаются все возможные варианты появлений или непоявлений рассматриваемых событий. Таким образом, с каждым экспериментом мы связываем список всех заведомо возможных в этом эксперименте событий, что дает основание для следующего определения.

Определение. *Пространством элементарных исходов эксперимента* будем называть произвольное множество Ω:

$$\Omega = \{\omega\}.$$

Элементы ω множества Ω будем называть *элементарными случайными событиями* или *элементарными исходами*, а само Ω — *пространством элементарных случайных событий*.

При построении пространства элементарных событий, описывающего конкретный эксперимент, в список элементарных исходов Ω включают элементарные исходы ω, удовлетворяющие требованиям *полноты* и *несовместности*. Первое из требований означает, что исходы, включенные в список, заведомо исчерпывают все возможные при однократном проведении эксперимента события; второе — что при однократном проведении эксперимента может произойти один и только один из исходов, включенных в список. Конечно, при составлении списка Ω прибегают к разумной степени идеализации реального эксперимента.

Пример 1. Пусть рассматриваемый эксперимент состоит в подбрасывании монеты. Пренебрегая возможностью того, что в итоге монета закатится в щель, встанет на ребро и т. п., выделим следующие два возможных исхода: Р — «монета упала вверх решкой», Г — «монета упала вверх гербом». Пространство элементарных исходов рассматриваемого эксперимента состоит из двух элементов,

$$\Omega = \{\Gamma;\ \mathrm{P}\}.$$

Пример 2. В урне находится m черных и n белых шаров. Эксперимент состоит в извлечении наудачу одного шара из урны. Один из вариантов пространства Ω для этого эксперимента может быть следующим: в список Ω включены два события Б — извлеченный шар белый и Ч — извлеченный шар черный,

$$\Omega = \{\text{Б}, \text{Ч}\}.$$

Пример 3. Из колоды игральных карт (52 листа четырех мастей: $\spadesuit$ — пики, $\heartsuit$ — черви, $\diamondsuit$ — бубны и $\clubsuit$ — трефы, в каждой масти 13 листов — 1, 2, ..., Валет, Дама, Король; карта единичного достоинства носит название Туз) извлекают одну карту. Возможные списки Ω при описании этого эксперимента могут быть следующими:

$$\Omega_1 = \{\spadesuit, \heartsuit, \diamondsuit, \clubsuit\};$$

$$\Omega_2 = \{1, 2, 3, \ldots, 10, \text{В}, \text{Д}, \text{К}\};$$

$$\Omega_3 = \{\text{Кт}, \text{Нкт}\};$$

$$\Omega_4 = \{\text{Кр}, \text{Чр}\};$$

$$\Omega_5 = \{1\spadesuit, \ldots, \text{К}\spadesuit, 1\heartsuit, \ldots, \text{К}\heartsuit, 1\diamondsuit, \ldots, \text{К}\diamondsuit, 1\clubsuit, \ldots, \text{К}\clubsuit\}.$$

Здесь приняты следующие обозначения:

$\spadesuit$ — извлечена карта масти пик; аналогично $\heartsuit$, $\diamondsuit$, $\clubsuit$;

1, 2, ..., 10, В, Д, К — извлечена карта соответствующего достоинства; (скажем, 3 — извлечена тройка (пик, червей, бубей или треф));

Кт — извлечена картинка (к картинкам относятся карты Туз, Валет, Дама, Король);

Нкт — извлечена некартинка (карта достоинством от 2 до 10);

Кр — извлечена карта красной масти (черви или бубны);

Чр — извлечена карта черной масти (пики или трефы).

Пример 4. Монету подбрасывают в одних и тех же условиях до тех пор, пока не появится герб. Пространство элементарных исходов этого эксперимента может быть задано следующим (бесконечным) списком:

$$\Omega = \{\text{Г}; \text{РГ}; \text{РРГ}; \ldots; \underbrace{\text{РР}\ldots\text{Р}}_{n-1}\text{Г}; \ldots\};$$

здесь исход, обозначенный как

$$\underbrace{\text{РР}\ldots\text{Р}}_{n-1}\text{Г},$$

описывает ситуацию, когда впервые герб появился при n-м бросании монеты.

Пример 5. Эксперимент состоит в наблюдении за временем безотказной работы некоторого агрегата. Идеализируя ситуацию, предположим, что мы умеем измерять любые, сколь угодно малые промежутки времени (в действительности, конечно, мы в состоянии измерять длительности только с точностью до некоторого «кванта» времени Δt). Тогда пространство элементарных исходов, отвечающее рассматриваемому эксперименту, может быть, например, задано следующим (континуальным!) списком:

$$\Omega = \{\omega\colon\ 0 \leqslant \omega < +\infty\}.$$

Здесь исход $\omega \in [0, +\infty)$ описывает ситуацию, когда агрегат вышел из строя в момент $t_0 + \omega$, где t_0 — время начала наблюдения.

Как показывают приведенные примеры, в зависимости от потребностей исследователя один и тот же эксперимент может быть описан различными пространствами элементарных исходов.

Принятая выше схема описания реального эксперимента позволяет отождествлять однократное проведение эксперимента с выбором одного из элементов списка Ω.

1.2. События

Следующее понятие, которое мы рассмотрим, это понятие *события*.

Событием будем называть любое подмножество A множества элементарных исходов Ω.

Мы говорим, что в результате однократного проведения эксперимента событие A осуществилось, если был выбран элемент $\omega \in A$. Элементарный исход $\omega \in A$ будем называть *благоприятствующим* осуществлению события A. *Достоверное* событие происходит при каждом осуществлении эксперимента, а *невозможное* событие не происходит ни при каком осуществлении эксперимента. Все остальные события с этой точки зрения занимают промежуточное положение между невозможным и достоверным. Невозможное событие будем обозначать символом $\varnothing$.

Пример 6. Опишем эксперимент Ω, состоящий в извлечении одной карты из полной (52 листа) колоды игральных карт списком

$$\Omega_5 = \{1\spadesuit, \dots, \text{К}\spadesuit, 1\heartsuit, \dots, \text{К}\heartsuit, 1\diamondsuit, \dots, \text{К}\diamondsuit, 1\clubsuit, \dots, \text{К}\clubsuit\}$$

(см. пример 3). Следующие подмножества множества Ω_5 образуют события в рассматриваемом эксперименте:

$$\text{Кт} = \{\text{«извлечена карта-картинка»}\} =$$
$$= \{\text{В}\spadesuit, \text{В}\heartsuit, \text{В}\diamondsuit, \text{В}\clubsuit, \text{Д}\spadesuit, \text{Д}\heartsuit, \text{Д}\diamondsuit, \text{Д}\clubsuit, \text{К}\spadesuit, \text{К}\heartsuit, \text{К}\diamondsuit, \text{К}\clubsuit, \text{Т}\spadesuit, \text{Т}\heartsuit, \text{Т}\diamondsuit, \text{Т}\clubsuit\};$$

$$\text{Кр} = \{\text{«извлечена карта красной масти»}\} = \{1\heartsuit, \dots, \text{К}\heartsuit, 1\diamondsuit, \dots, \text{К}\diamondsuit\};$$

$$\spadesuit = \{\text{«извлечена карта масти пик»}\} = \{1\spadesuit, \dots, \text{К}\spadesuit\}.$$

Пример 7. Пусть эксперимент Ω состоит в трехкратном подбрасывании монеты. Опишем его следующим списком

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

Примерами событий в этом эксперименте могут быть следующие:

$$A_{\geqslant 2} = \{\text{«герб выпал не менее 2 раз»}\} = \{\text{ГГГ, ГГР, РГГ, ГРГ}\};$$

$$A_{=2} = \{\text{«герб выпал ровно 2 раза»}\} = \{\text{ГГР, РГГ, ГРГ}\};$$

$$A_{=0} = \{\text{«герб не выпал ни разу»}\} = \{\text{РРР}\}.$$

Легко сообразить, что

в случае, когда множество исходов эксперимента Ω конечно и содержит N элементов, множество всех возможных событий также конечно и содержит 2^N элементов.

◀ Действительно, событий, состоящих ровно из k «благоприятствующих» элементарных исходов будет C_N^k, $k = 0, 1, 2, \dots, N$. Поэтому общее число всех возможных событий дается равенством

$$C_N^0 + C_N^1 + \dots + C_N^N = (1+1)^N = 2^N. \blacktriangleright$$

Пример 8. Монету бросают до первого появления герба. Пространство элементарных событий опишем списком (см. пример 4)

$$\Omega = \{\text{Г, РГ}, \dots, \text{РР} \dots \text{РГ}, \dots\}.$$

В рассматриваемом эксперименте можно наблюдать, например, следующие события:

$$A_{\leqslant 10} = \{\text{«герб появился до одиннадцатого броска»}\} = \{\text{Г, РГ, РРГ}, \dots, \underbrace{\text{РР} \dots \text{Р}}_{9}\text{Г}\};$$

$$A_{=10} = \{\text{«герб появился при десятом испытании»}\} = \{\underbrace{\text{РР} \dots \text{Р}}_{9}\text{Г}\};$$

$$A_{>10} = \{\text{«герб не появился ранее одиннадцатого испытания»}\} = \{\underbrace{\text{РР} \dots \text{Р}}_{10}\text{Г}, \text{РР} \dots \text{Р}_{11}\text{Г}, \dots\};$$

$${}_{10\geqslant}A_{\geqslant 4} = \{\text{«герб появился между четвертым и десятым испытаниями»}\} =$$
$$= \{\text{РРРГ, РРРРГ}, \dots, \underbrace{\text{РР} \dots \text{Р}}_{9}\text{Г}\}.$$

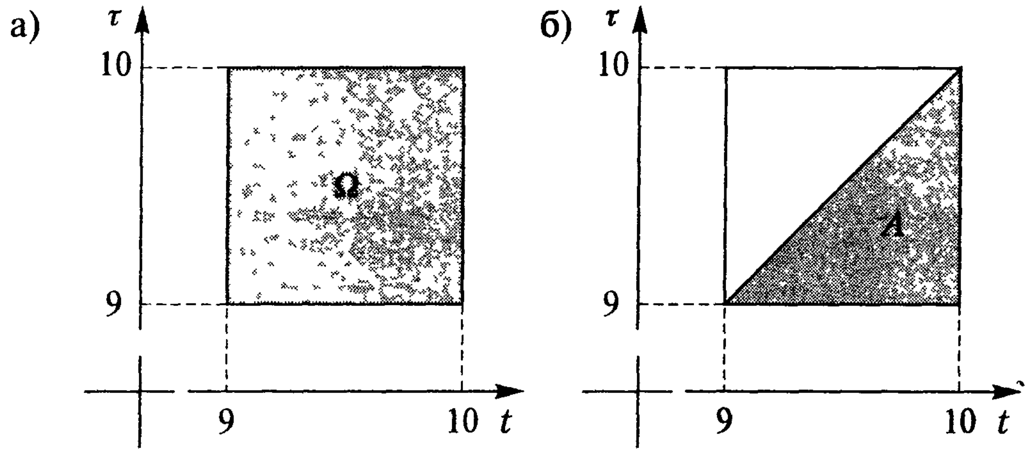


Рис. 1

Пример 9. Известно, что автобус приходит к остановке в некоторый момент времени t, заключенный между 9 и 10 часами утра Студент выходит на остановку в течение этого же промежутка времени. Эксперимент состоит в наблюдении за временем прихода к остановке автобуса и студента. Опишем множество элементарных исходов следующим списком

$$\Omega = \{(t, \tau)\colon\ 9 \leqslant t \leqslant 10;\ 9 \leqslant \tau \leqslant 10\}.$$

Здесь t — момент прибытия автобуса к остановке, τ — момент прибытия студента. Множество Ω представляет совокупность всех упорядоченных пар чисел (t, τ), заключенных в промежутке от 9 до 10 каждое Если откладывать t по оси абсцисс, а τ — по оси ординат, то Ω будет соответствовать множеству точек квадрата на плоскости (t, τ) (рис. 1 а). Рассмотрим событие

$$A = \{\text{«Студент пришел на остановку не позднее, чем туда прибыл автобус»}\} = \{(t, \tau)\colon\ \tau \leqslant t\}.$$

Это событие может быть описано теми точками квадрата, в которых $\tau < t$ (рис. 1 б).

Пример 10. Пусть множество ω — время безотказной работы агрегата из примера 5. Множество

$$\Omega = \{\omega\colon\ 0 \leqslant \omega < +\infty\}$$

может быть отождествлено с множеством точек числовой полупрямой. Событие

$$A = \{\text{«агрегат вышел из строя в промежутке времени от } \tau_0 \text{ до } \tau_1\text{»}\} = \{\omega\colon\ \tau_0 \leqslant \omega \leqslant \tau_1\}$$

изобразится отрезком этой полупрямой

Для дальнейшего важным является то обстоятельство, что уже в случае континуальных пространств элементарных событий запас возможных событий становится необозримым. Причем среди них могут быть и весьма экзотические.

Пример 11. Канторово множество. Приведем пример экзотического множества, которое обладает на первый взгляд противоречивыми свойствами Рассмотрим отрезок $[0, 1]$ Разобьем его на три равные части и удалим средний интервал, т е интервал $(1/3, 2/3)$. У нас останется два отрезка — $[0, 1/3]$ и $[2/3, 1]$ С каждым из этих отрезков поступим аналогично: разделим на три равные части и удалим средний интервал После этого останется четыре отрезка С каждым из них поступим так же — получится восемь отрезков, и так далее до бесконечности. То, что останется после бесконечного числа выбрасываний, и есть канторово множество (на рис. 2 показана схема его построения). В него входят концы выбрасываемых интервалов, но интересно, что входит и еще много других точек Множество концов выбрасываемых интервалов счетно, в то время как канторово множество имеет мощность континуума.

Действительно, запишем числа из интервала $[0, 1]$ в троичной системе счисления, т. е в виде

$$0, \alpha_1\alpha_2 \ldots \alpha_n \ldots, \tag{1}$$

где цифра α_i, стоящая в i-й позиции после запятой, может быть нулем, единицей или двойкой. Выбрасываемые интервалы состоят из точек, соответствующих числам вида (1), в записи которых встречается хотя бы одна 1 Соответственно, канторово множество состоит из всех точек, описываемых числами вида (1), где каждое α_i равно или 0, или 2. Следовательно, если мы сопоставим точке $0, \alpha_1\alpha_2 \ldots \alpha_n \ldots$ в троичной системе точку $0, \beta_1\beta_2 \ldots \beta_n \ldots$ в двоичной системе, где $\beta_i = \alpha_i/2$, то получим взаимно однозначное отображение канторова множества на весь отрезок $[0, 1]$, и, значит, мощность канторова множества оказывается равной континууму Тем самым, в смысле мощности канторово множество «большое» Однако оно «маленькое» в смысле «размера» — его мера равна 0 Действительно, под-

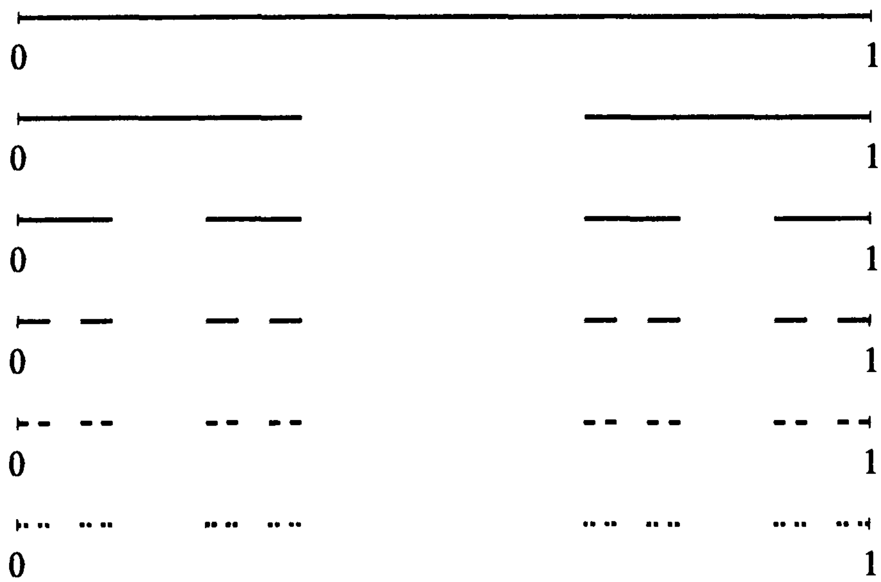


Рис. 2

считаем сумму длин выброшенных интервалов; она оказывается равной единице, т. е. длине всего отрезка $[0, 1]$:

$$\frac{1}{3}+\frac{2}{9}+\frac{4}{27}+\ldots+\frac{2^{n-1}}{3^n}+\ldots=\frac{1/3}{1-2/3}=1.$$

Назовем числовое множество *нигде не плотным*, если в любом интервале можно указать отрезок, не содержащий ни одной точки данного множества. Интуитивно свойство «нигде-не-плотности» означает, что множество очень «разрежено» (таким будет, например, множество, состоящее из изолированных точек). Нигде не плотным будет также множество, содержащее только конечное число неизолированных точек. Назовем множество *совершенным*, если у него вообще нет изолированных точек. Такое множество должно быть «густым», а не «разреженным»; например, отрезок есть совершенное множество. Кажется невероятным, чтобы некоторое множество было бы нигде не плотным и совершенным одновременно, и тем не менее канторово множество именно таково: оно нигде не плотно, т. к. в любом интервале найдется точка, в троичной записи которой есть 1, а значит, и целый выбрасываемый интервал, и оно совершенно, т. к. в любой окрестности точки, в троичной записи которой есть только 0 и 2, найдутся другие точки с тем же свойством, т. е. изолированных точек в канторовом множестве нет. Перечислим еще раз рассмотренные свойства канторова множества:

1. Мощность канторова множества равна континууму.
2. Мера канторова множества равна нулю.
3. Канторово множество нигде не плотно.
4. Канторово множество совершенно.

Интуитивно кажется, что как первые два свойства, так и последние два противоречат друг другу.

§ 2. Алгебра событий

2.1. Диаграммы Эйлера—Венна

Пусть Ω — пространство элементарных исходов некоторого эксперимента, A, B, C, ... — события, среди которых Ω — достоверное событие и $\varnothing$ — невозможное событие. Выше мы уже отмечали, что каждое событие является подмножеством множества Ω, а осуществление события A в эксперименте мы понимаем как принадлежность реализовавшегося в данном испытании элементарного исхода ω множеству A:

$$\boxed{\Omega \to \omega: \begin{cases} \omega \in A \text{ — событие } A \text{ произошло,} \\ \omega \notin A \text{ — событие } A \text{ не произошло.} \end{cases}}$$

В дальнейшем это соглашение будем иллюстрировать следующим образом: пространство элементарных исходов Ω (независимо от природы составляющих его элементов и мощности) будем изображать прямоугольником; элементарные случайные события — точками прямоугольника; события — подмножествами прямоугольника (рис. 3). В этой модели достоверное событие — множество всех точек прямоугольника; невозможное событие — пустое подмножество; испытание — выбор одной (ровно одной!) точки из прямоугольника, представляющей реализовавшееся в данном испытании элементарное случайное событие Такие иллюстрации носят название *диаграмм Эйлера—Венна*.

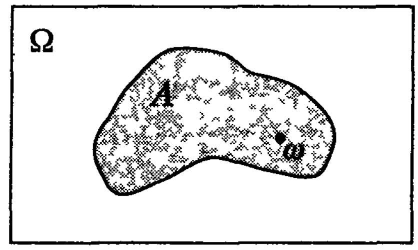


Рис 3

2.2. Совмещение событий

Определение. *Совмещением* двух событий $A \in \Omega$ и $B \in \Omega$ назовем событие, состоящее из тех и только из тех исходов, которые благоприятствуют событиям A и B одновременно.

Обозначение: $A \cap B = AB$ (A «и» B).

Ясно, что совмещение событий A и B имеет место (происходит) только тогда, когда события A и B происходят одновременно, т. е. элементарный исход $\omega \in \Omega$, реализовавшийся в данном испытании, принадлежит одновременно и событию A и событию B. Если у событий A и B нет общих элементарных исходов, они называются *несовместными*. Это обстоятельство будем обозначать так:

$$\boxed{AB = A \cap B = \varnothing\ .}$$

Совмещение любого числа событий можно рассматривать аналогично.

Пример 1. Ω — эксперимент, состоящий в извлечении карты из полной колоды (52 листа) игральных карт

$$\Omega = \{1\spadesuit, \ldots, K\spadesuit, 1\heartsuit, \ldots, K\heartsuit, 1\diamondsuit, \ldots, K\diamondsuit, 1\clubsuit, \ldots, K\clubsuit\};$$
$$A = \{\text{«извлечен туз»}\} = \{1\spadesuit, 1\heartsuit, 1\diamondsuit, 1\clubsuit\},$$
$$B = \{\text{«извлечена карта черной масти»}\} = \{1\spadesuit, \ldots, K\spadesuit, 1\clubsuit, \ldots, K\clubsuit\};$$
$$AB = \{\text{«извлечен туз черной масти»}\} = \{1\spadesuit, 1\clubsuit\}.$$

Пример 2. Проверяется партия телевизоров, содержащая 10 образцов Событие

$$A_i, \quad i = 1, \ldots, 10$$

состоит в том, что i-й телевизор дефектен Событие

$$\bigcap_{i=1}^{10} A_i = A_1 A_2 \ldots A_{10}$$

— в том, что все проверенные образцы дефектны

Пример 3. Релейная схема составлена из агрегатов A_1, A_2, A_3, A_4 и устроена так, как показано на рис. 4 Схема выходит из строя за время T при одновременном выходе из строя за это время всех агрегатов Пусть A_i, $i = 1, 2, 3, 4$, — событие, состоящее в том, что i-й агрегат вышел из строя Тогда событие A — «схема вышла из строя» — является совмещением событий A_i

$$A = A_1 A_2 A_3 A_4.$$

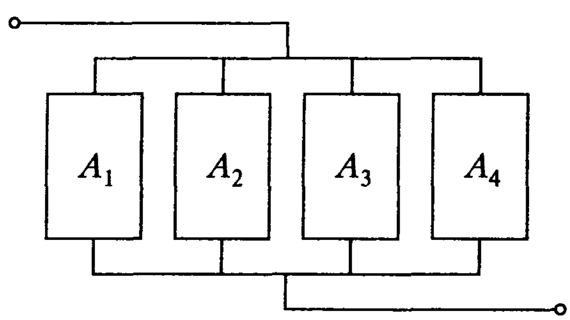


Рис. 4

Пример 4. Релейная схема составлена из агрегатов B_1, B_2, B_3, B_4 и устроена, как показано на рис. 5. Схема работает безотказно в течение времени T, если все агрегаты в течение этого времени работают безотказно. Пусть B_j, $j = 1, 2, 3, 4$, — событие, состоящее в том, что j-й агрегат в течение времени T не вышел из строя. Тогда совмещением событий B_j будет событие B — «схема работает безотказно в течение времени T».

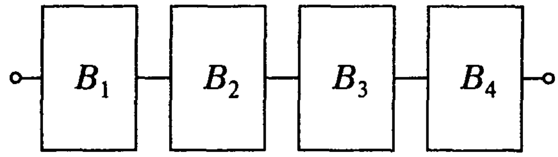


Рис. 5

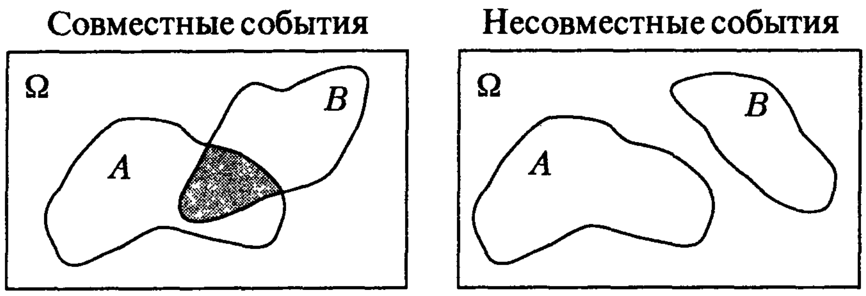


Рис. 6

С использованием диаграмм Эйлера—Венна операцию совмещения событий можно изобразить так, как показано на рис. 6. Отметим два очевидных свойства операции совмещения:

1. $A \cap \varnothing = \varnothing$;
2. $A \cap \Omega = A$.

2.3. Подчиненность событий

Будем говорить, что *событие B подчинено событию A* и записывать это отношение как

$$\boxed{A \subseteq B,}$$

если осуществление события A *влечет за собой* осуществление события B, т. е. все исходы A являются исходами B, но у B возможны исходы, не являющиеся исходами A (рис. 7). По определению положим также, что $\varnothing \subseteq A$ для любого A.

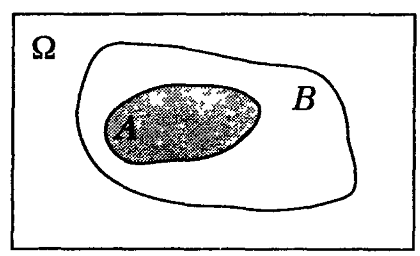


Рис 7

Отметим, что события, происходящие в эксперименте Ω, упорядочены лишь частично — отношение подчиненности определено не для любой пары событий.

Имеют место следующие соотношения:

1. $\forall A$: $A \subseteq \Omega$.
2. $A \cap B \subseteq A$, $A \cap B \subseteq B$.
3. $C \subseteq AB \Leftrightarrow C \subseteq A$, $C \subseteq B$.
4. $A \subseteq B$, $B \subseteq C \Rightarrow A \subseteq C$.

Свойства **1** и **2** очевидным образом следуют из определений. Свойства **3** и **4** легко усмотреть из диаграмм Эйлера—Венна, приведенных на рис. 8.

Пример 5. В эксперименте с извлечением карты из полной колоды игральных карт (52 листа) событие $B = \{$«извлечена карта-картинка»$\}$ подчинено событию $A = \{$«извлечен король»$\}$. $A \subseteq B$.

Пример 6. В этом же эксперименте события $B = \{$«извлечена карта-картинка»$\}$ и $C = \{$«извлечена карта черной масти»$\}$ не состоят друг с другом в отношении подчиненности.

$C \subseteq AB \Leftrightarrow C \subseteq A$ и $C \subseteq B$ $\quad$ $(A \subseteq B) \cap (B \subseteq C) \Rightarrow A \subseteq C$

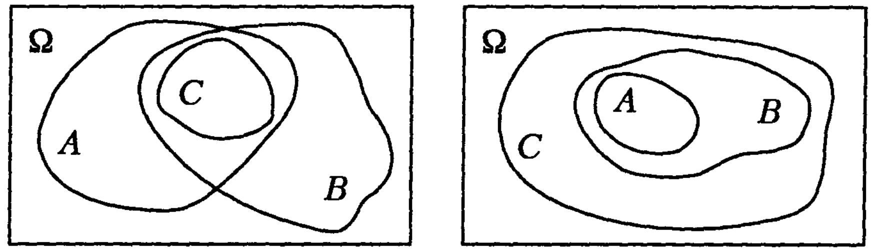


Рис. 8

2.4. Сумма событий

Определение. *Суммой двух событий* $A \in \Omega$ и $B \in \Omega$ назовем событие, состоящее из элементарных исходов, входящих в A или в B.

Обозначение: $A \cup B$ (A «или» B).

Событие, являющееся суммой событий A и B, происходит, если произошло по крайней мере одно из этих событий — или произошло A, или произошло B, или оба этих события произошли вместе. Диаграммы, иллюстрирующие эту операцию над событиями, приведены на рис. 9.

Сумма произвольного числа событий определяется аналогично. В случае несовместности событий A и B вместо знака объединения «$\cup$» часто употребляют знак суммы «$+$» (рис. 10).

Отметим следующие соотношения:

1. $A \cup \varnothing = A$.
2. $A \cup \Omega = \Omega$.
3. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
4. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Последние два из них — свойства дистрибутивности умножения относительно сложения и сложения относительно умножения — нуждаются в доказательстве. Покажем, как это делается, на примере соотношения **4.**

◀ Событие $A \cup (B \cap C)$ состоит из тех элементарных исходов, которые входят или в A, или в $B \cap C$, или в оба этих события одновременно (рис. 11). Следовательно, событие $A \cup (B \cap C)$ происходит, когда происходит либо событие A, либо события B и C одновременно. Но если произошло A, то произошло и $A \cup B$ и $A \cup C$, а, значит, они оба, т. е. $(A \cup B) \cap (A \cup C) \supseteq A$. Если же осуществилось $B \cap C$, то осуществились $A \cup B$ и $A \cup C$ вместе: $(B \cap C) \subseteq (A \cup B) \cap (A \cup C)$. Следовательно,

$$A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C).$$

С другой стороны (рис. 12), события $(A \cup B)$ и $(A \cup C)$ происходят одновременно, если происходит событие A, либо, если A не произошло, то одновременно должны произойти B и C. Следовательно,

$$(A \cup B) \cap (A \cup C) \subseteq A \cup (B \cap C).$$

Вместе с полученным выше отношением подчиненности это дает искомое утверждение. ▶

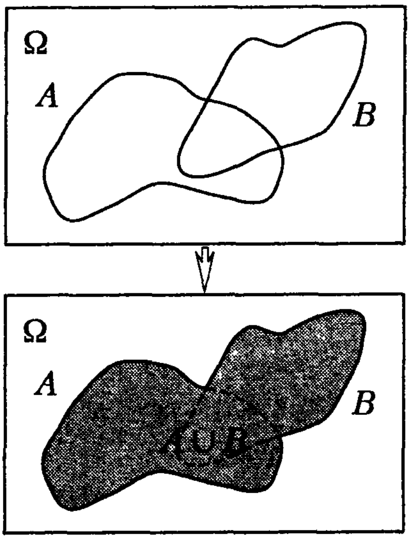


Рис. 9

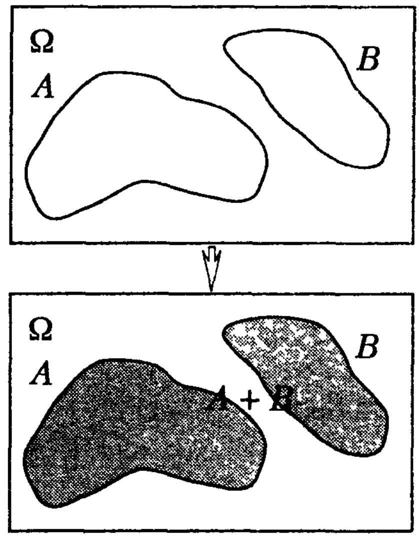


Рис. 10

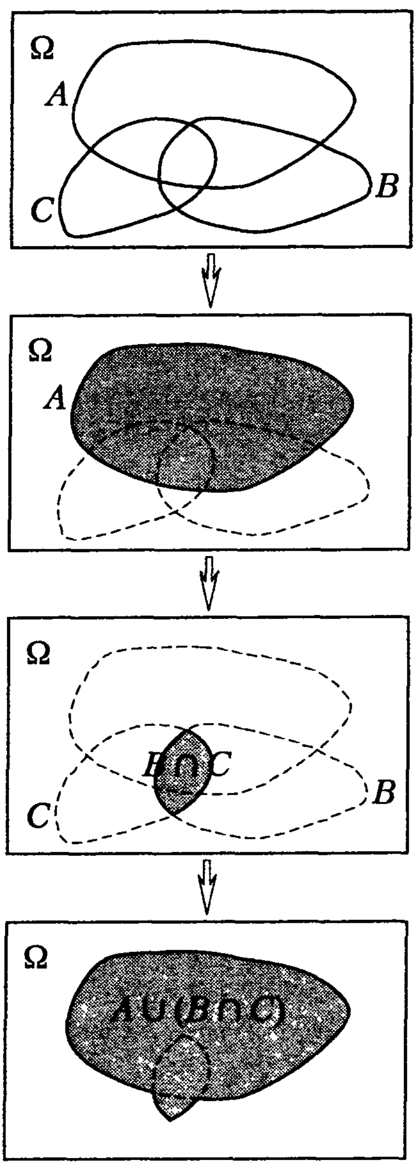


Рис 11

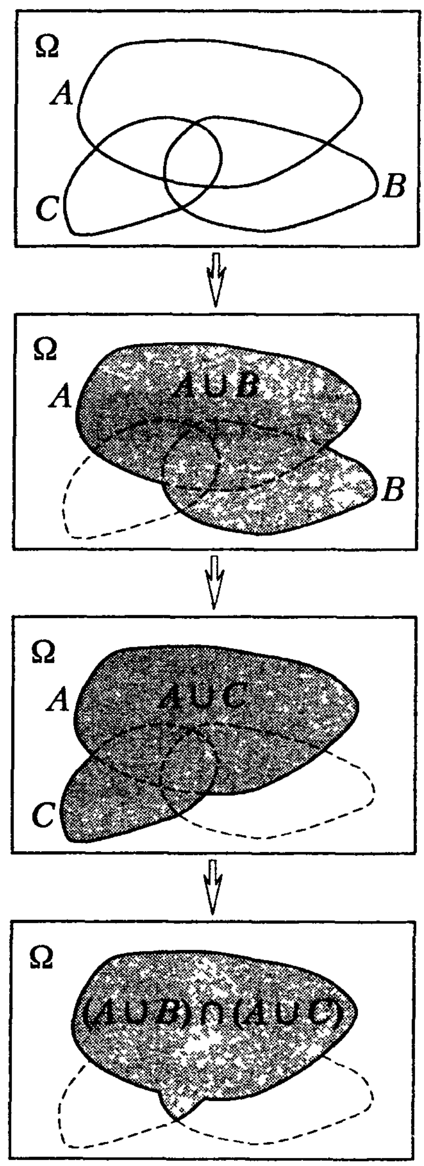


Рис 12

Пример 7. Пусть в эксперименте с извлечением карты из полной колоды (52 листа) игральных карт событие $A = \{$«извлечена карта черной масти»$\}$, а событие $B = \{$«извлечена карта-картинка»$\}$, тогда их сумма $A \cup B$ — событие, состоящее в том, что извлечена карта черной масти или красная картинка

Пример 8. Релейная схема составлена из агрегатов A_1, A_2, A_3, B_1, B_2, как показано на рис 13 Пусть события A_j, B_i, $j = 1, 2, 3$, $i = 1, 2$, состоят в том, что одноименные агрегаты работают безотказно в течение времени T, а U — событие, состоящее в том, что схема работает безотказно в течение времени T Тогда

$$U = B_1 B_2 (A_1 \cup A_2 \cup A_3).$$

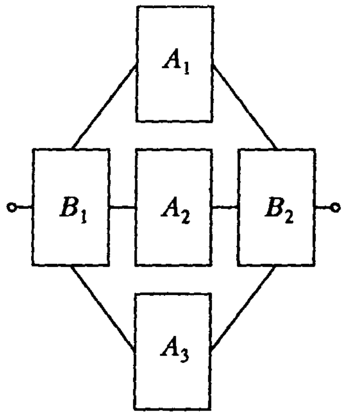


Рис. 13

2.5. Противоположные события. Тождества де Моргана

Определение. Событием, *противоположным* событию A, назовем событие, состоящее из тех исходов Ω, которые не входят в A.

Обозначение: $\overline{A}$ («не A»).

Событие «не A» происходит всегда, когда не происходит A (рис. 14).

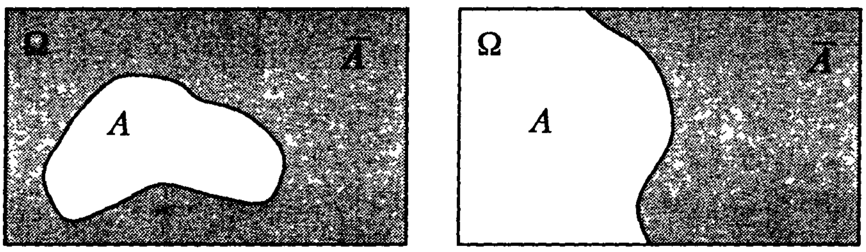


Рис 14

Ясно, что

$$\boxed{\overline{\Omega} = \varnothing, \quad \overline{\varnothing} = \Omega, \quad \overline{\overline{A}} = A.}$$

Кроме того,

$$\boxed{\forall A \in \Omega\colon\ \overline{A} \cap A = \varnothing, \quad A \cup \overline{A} = \Omega.}$$

Имеют место следующие тождества (*тождества де Моргана*):

$$\boxed{\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B},}$$

которые легко обобщаются на случай конечного числа слагаемых (сомножителей)

$$\boxed{\overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \overline{A_i}, \quad \overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \overline{A_i}.}$$

Доказательство тождеств де Моргана легко усмотреть из рис. 15 и 16 соответственно.

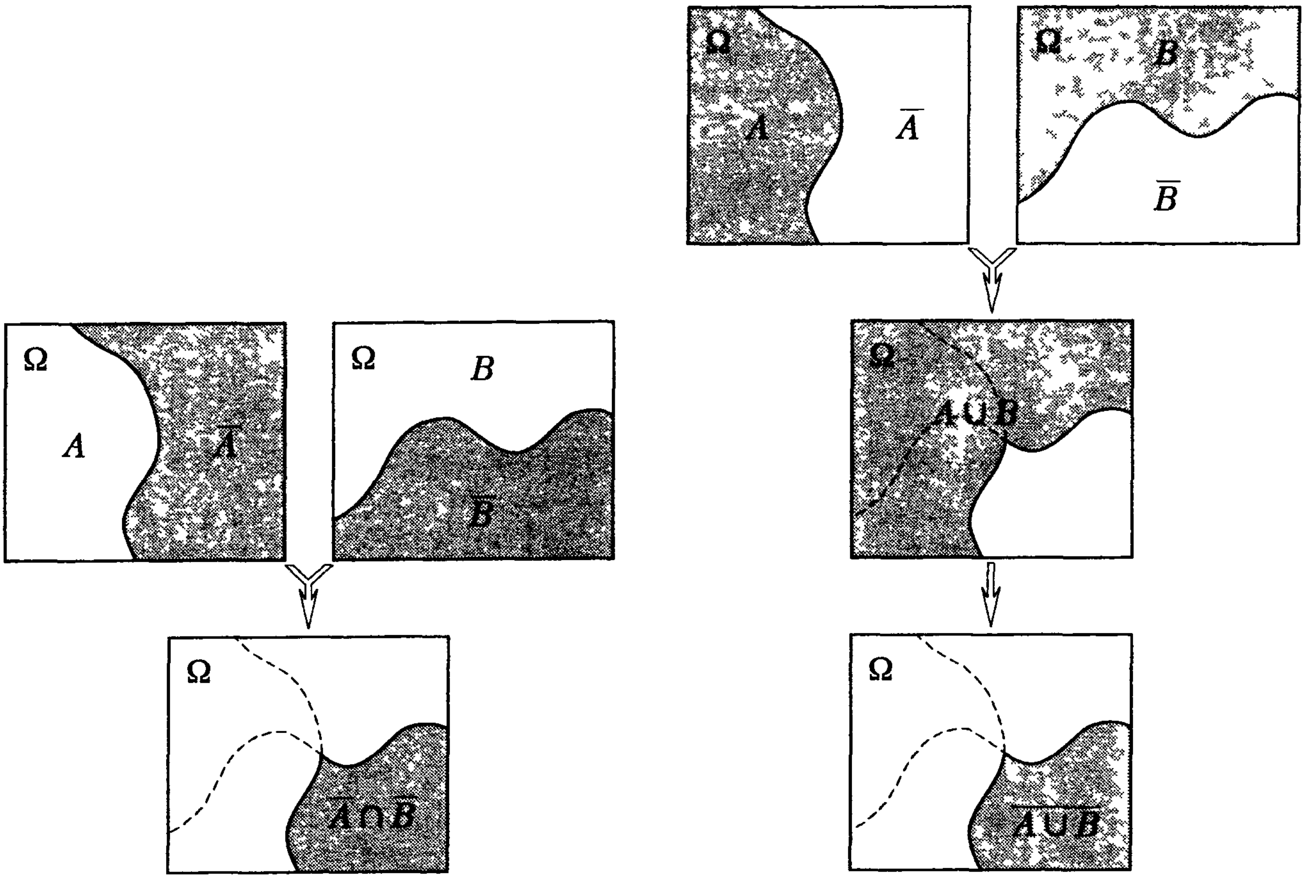

Ω
A
$\overline{A}$
Ω
B
$\overline{B}$
Ω
$A \cup B$
Ω
$\overline{A \cup B}$
Ω
A
$\overline{A}$
Ω
B
$\overline{B}$
Ω
$\overline{A} \cap \overline{B}$

Рис 15

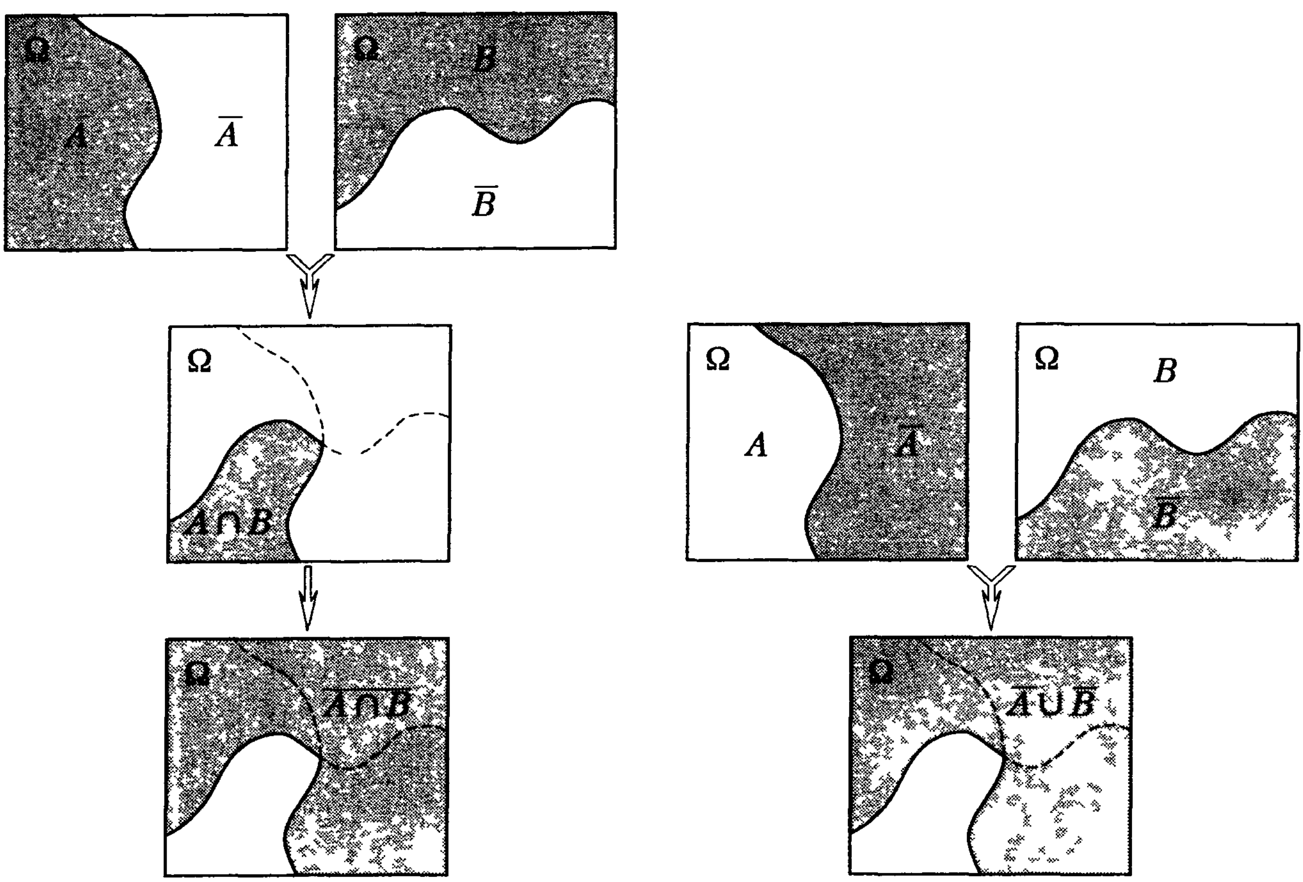

Ω
A
$\overline{A}$
Ω
B
$\overline{B}$
Ω
$A \cap B$
Ω
$\overline{A \cap B}$
Ω
A
$\overline{A}$
Ω
B
$\overline{B}$
Ω
$\overline{A} \cup \overline{B}$

Рис 16

Пример 9. Релейная схема составлена из элементов $A_1, A_2, \ldots, A_n$, как показано на рис. 17. Схема выходит из строя за время T, если за это время выйдет из строя по крайней мере один из элементов A_j, $j = 1, 2, \ldots, n$.

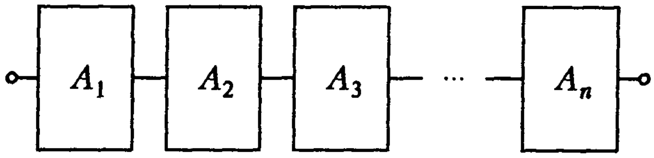


Рис. 17

Обозначим теми же буквами A_j, $j = 1, 2, \ldots, n$, события, состоящие в том, что j-й элемент схемы за время T из строя не вышел. Тогда событие U, состоящее в том, что схема не выйдет из строя за время T, описывается соотношением

$$U = A_1 A_2 \ldots A_n,$$

а событие $\overline{U}$ — «схема выйдет из строя за время T» — будет выглядеть так

$$\overline{U} = \overline{A_1} \cup \overline{A_2} \cup \ldots \cup \overline{A_n} = \overline{A_1 A_2 \ldots A_n}.$$

Пример 10. Релейная схема составлена из элементов $A_1, A_2, \ldots, A_n$ так, как показано на рис. 18. Схема выходит из строя за время T, если за это время выйдут из строя все элементы A_j, $j = 1, 2, \ldots, n$. Пусть события U и A_j — те же, как и в примере 9. Тогда имеем

$$U = A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n,$$
$$\overline{U} = \overline{A_1}\,\overline{A_2} \ldots \overline{A_n} = \overline{A_1 \cup A_2 \cup \ldots \cup A_n}.$$

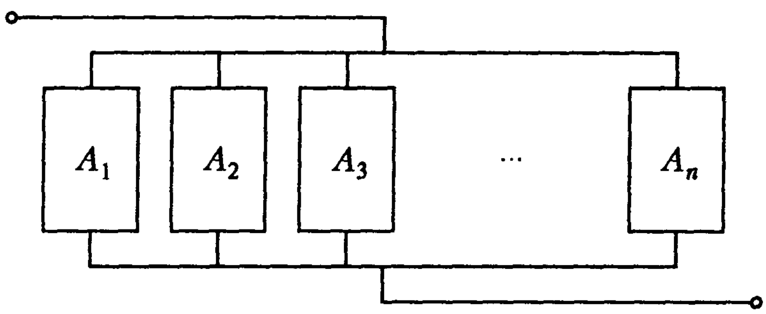


Рис. 18

2.6. Разность событий

Определение. *Разностью* событий $A \in \Omega$ и $B \in \Omega$ будем называть событие, состоящее из всех тех элементарных исходов, которые входят в A, но не входят в B.

Обозначение: $A \setminus B$ (A «без» B).

Это событие происходит только тогда, когда происходит A, но не происходит B, т. е. $A \setminus B = A \cap \overline{B}$ (рис. 19). Ясно, что (рис. 20)

$$\boxed{\begin{gathered} \Omega \setminus A = \overline{A}, \quad A \setminus \Omega = \varnothing, \\ A = (A \setminus B) + A \cap B, \\ (A \setminus B) \cap AB = \varnothing, \\ A \cup B = (A \setminus B) + (B \setminus A) + AB. \end{gathered}}$$

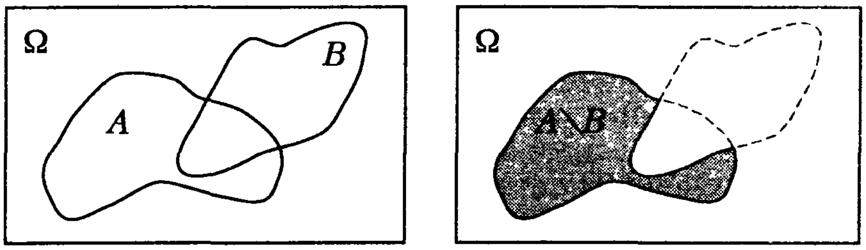


Рис. 19

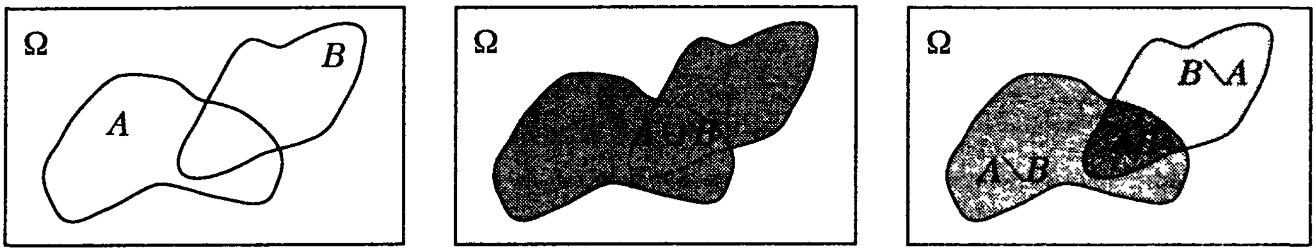


Рис. 20

В заключение этого краткого обзора правил действий над событиями отметим, что если

$$A_1 \supseteq A_2 \supseteq \ldots \supseteq A_n \supseteq \ldots$$

— бесконечная последовательность подчиненных событий, то существует событие $A_{\min}$, которому подчинены все A_j

$$A_{\min} = \bigcap_{j=1}^{\infty} A_j.$$

Это событие происходит тогда, когда одновременно происходят все события A_j. Этот факт будем записывать так

$$\boxed{A_n \downarrow A \quad \Longleftrightarrow \quad A_{\min} = \lim_{n\to\infty} A_n = \bigcap_{j=1}^{\infty} A_j.}$$

Если же

$$A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subseteq \ldots,$$

то существует событие $A_{\max}$, которое подчинено всем A_j

$$A_{\max} = \bigcup_{j=1}^{\infty} A_j.$$

В этом случае будем использовать также и символ lim

$$\boxed{A_n \uparrow A \quad \Longleftrightarrow \quad A_{\max} = \lim_{n\to\infty} A_n = \bigcup_{j=1}^{\infty} A_j.}$$

§ 3. Вероятность. Случайные события

Пусть Ω — пространство элементарных исходов некоторого эксперимента n и A — событие. В результате однократного испытания событие A происходит или нет в зависимости от того, принадлежит событию A реализовавшийся в этом испытании исход ω

или не принадлежит. Наша ближайшая задача состоит в указании способа числовой оценки правдоподобности осуществления или неосуществления события A в эксперименте Ω.

Пусть $P\{A\}$ — вещественная числовая функция, приписывающая каждому событию A число — степень его осуществимости в эксперименте Ω. Опишем предварительно некоторые интуитивно ясные свойства такой функции. Как говорилось выше, любое событие A (с точки зрения шансов на осуществимость) занимает промежуточное положение между невозможным событием $\varnothing$, которое не происходит никогда, и достоверным событием Ω, которое происходит всегда. Разумно считать в связи с этим, что

$$\boxed{P\{\varnothing\} \leqslant P\{A\} \leqslant P\{\Omega\}.}$$

Припишем невозможному событию оценочное нулевое значение

$$\boxed{P\{\varnothing\} = 0,}$$

а достоверному событию — значение единица

$$\boxed{P\{\Omega\} = 1.}$$

Столь же естественно считать, что если событие A влечет за собой событие B ($A \subseteq B$), то $P\{A\} \leqslant P\{B\}$, так как осуществление события A всегда вызывает осуществление события B; в то же время возможно осуществление B без осуществления A.

Отметим, наконец, еще одно естественное свойство: если A и B — пара несовместных событий, то степень достоверности осуществления события $A + B$ должна быть равна сумме соответствующих оценок для A и для B

$$P\{A + B\} = P\{A\} + P\{B\}.$$

Если бы такая функция P была бы построена, то она давала бы нам некоторую информацию об осуществимости событий: априорное осуществление события A тем более вероятно, чем больше оценочное значение $P\{A\}$ ему приписано; из двух событий заведомо скорее следует ожидать осуществления в эксперименте того, для которого значение P больше.

Ниже мы увидим, что если функцию с указанными свойствами удается построить, то она обязательно имеет очень простой (частотный) смысл (см. Закон больших чисел в форме Бернулли).

3.1. Вероятность события

Пусть Ω — эксперимент и A — событие, которое может происходить в этом эксперименте. *Вероятностью события* A будем называть числовую функцию $P\{A\}$, обладающую следующими свойствами:

1. $P\{A\} \geqslant 0$, $P\{\Omega\} = 1$ (свойство нормировки); (1)
2. $P\left\{\bigcup\limits_{i=1}^{\infty} A_n\right\} = \sum\limits_{n=1}^{\infty} P\{A_n\}$, если только $A_i \cap A_j = \varnothing$ (свойство счетной аддитивности). (2)

Из определения немедленно следует, что если функция со свойствами (1) и (2) может быть построена на множестве событий, происходящих в эксперименте Ω, то она обладает дополнительно следующими свойствами.

1. Конечная аддитивность

$$\boxed{\text{Если } A \cap B = \varnothing, \text{ то } P(A+B) = P(A) + P(B).}$$

2. Правило вычисления вероятности противоположного события

$$\boxed{P(\overline{A}) = 1 - P(A).} \tag{3}$$

◀ Поскольку $\Omega = A + \overline{A}$, то

$$P\{\Omega\} = P\{A\} + P\{\overline{A}\} = 1. \ ▶$$

3. Монотонность вероятности

$$\boxed{\text{Если } A \subseteq B, \text{ то } P\{A\} \leqslant P\{B\}.} \tag{4}$$

◀ Действительно, $B = (B \setminus A) + A$ (рис. 21), при этом $(B \setminus A) \cap A = \varnothing$. По правилу сложения (2) имеем

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

Из того, что $P(B \setminus A) \geqslant 0$, следует искомое. ▶

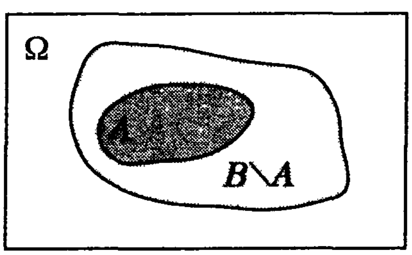


Рис. 21

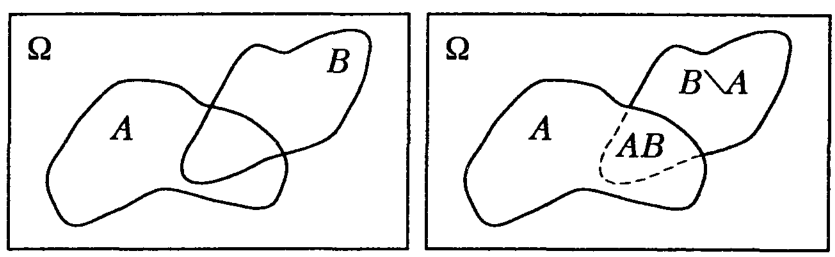


Рис. 22

4. Правило вычитания вероятностей

$$\boxed{\text{Если } A \subseteq B, \text{ то } P\{B \setminus A\} = P\{B\} - P\{A\}.} \tag{5}$$

5. Правило сложения вероятностей

$$\boxed{P\{A \cup B\} = P\{A\} + P\{B\} - P\{AB\}.} \tag{6}$$

◀ Доказательство следует из представления

$$A \cup B = A + (B \setminus AB)$$

(рис. 22), где $A \cap (B \setminus AB) = \varnothing$, $AB \subseteq B$, к которому нужно только применить свойство конечной аддитивности 1 и правило вычитания вероятностей 4. ▶

6. Непрерывность вероятности. Пусть

$$A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subseteq \ldots$$

и

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n = \varnothing .$$

Тогда

$$\boxed{\lim_{n\to\infty} P\{A_n\} = 0.} \tag{7}$$

◀ Положим $A_0 = \Omega$ и рассмотрим очевидное тождество

$$(A_0 \setminus A_1) + (A_1 \setminus A_2) + \ldots + (A_{n-1} \setminus A_n) + \ldots = \Omega.$$

Поскольку, как легко проверить,

$$(A_{n-1} \setminus A_n) \cap (A_{s-1} \setminus A_s) = \varnothing,$$

то, применив правило сложения (2) к указанному тождеству, получим

$$1 = P\{\Omega\} = P\{\Omega \setminus A_1\} + \ldots + P\{A_{n-1} \setminus A_n)\} + \ldots = \sum_{i=0}^{n-1} P\{A_i \setminus A_{i+1}\} + R_n,$$

где

$$R_n = P\left\{\bigcap_{i=n}^{\infty}(A_i \setminus A_{i+1})\right\}.$$

В силу сходимости приведенного выше ряда

$$\lim_{n\to\infty} R_n = 0.$$

Но

$$\bigcap_{i=n}^{\infty}(A_i \setminus A_{i+1}) = A_n,$$

откуда

$$\lim_{n\to\infty} P(A_n) = 0. \blacktriangleright$$

Остается установить, можно ли такую функцию определить для *всех* возможных в эксперименте Ω событий.

3.2. Вероятность в экспериментах с дискретной структурой

Будем называть эксперимент Ω *экспериментом с дискретной структурой* или просто *дискретным экспериментом*, если множество его исходов не более чем счетно,

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

Для дискретных экспериментов рассмотренная выше схема определения вероятности, как будет сейчас показано, всегда осуществима. Действительно, пусть F — совокупность всех возможных в эксперименте Ω событий, в которую включены, конечно, и невозможное событие $\varnothing$, и достоверное Ω,

$$F = \Big\{\varnothing, \Omega, \{\omega_1\}, \{\omega_2\}, \ldots, \{\omega_n\}, \ldots\Big\}.$$

Пусть $p_1, p_2, \ldots, p_n, \ldots$ — произвольный счетный набор действительных положительных чисел, таких, что ряд

$$p_1 + p_2 + \ldots + p_n + \ldots$$

сходится и его сумма равна 1. Положим по определению

$$P\{\omega_j\} = p_j, \quad j = 1, 2, \ldots, n, \ldots \tag{8}$$

Всякое событие A из совокупности F состоит из не более чем счетного числа элементарных исходов ω_j^A, $j = 1, 2, \ldots, m, \ldots$:

$$A = \Big\{\omega_1^A, \omega_2^A, \ldots, \omega_m^A, \ldots\Big\}.$$

Определим его вероятность $P\{A\}$ следующей формулой

$$\boxed{P\{A\} = \sum_{\omega_j \in A} P(\omega_j) = \sum_{j=1}^{\infty} P\{\omega_j^A\}.} \tag{9}$$

Легко видеть, что при таком способе задания вероятности *каждому* событию будет приписана вероятность (при помощи равенства (9)), удовлетворяющая требованиям (1) и (2).

Отметим, что теория вероятностей как математическая дисциплина начиналась с задач, описываемых конечными пространствами элементарных исходов

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

и одним из основных предположений, положенных в основание теории, было предположение о равновозможности всех элементарных исходов:

$$P(\omega_j) = p \quad \forall\, j = 1, 2, \ldots, N.$$

Учитывая требования (1) и (2), предъявляемые к вероятности, равновозможность элементарных исходов и конечность эксперимента, т. е. $P\{\Omega\} = \sum P(\omega_j) = N \cdot p = 1$, приходим к следующему правилу вычисления вероятностей:

$$\boxed{P(\omega_j) = p = \frac{1}{N},}$$

$$\boxed{A = \{\omega_1^A, \ldots, \omega_k^A\}, \quad P\{A\} = \frac{1}{N} + \frac{1}{N} + \ldots + \frac{1}{N} = \frac{k}{N}.}$$

Вероятность любого события A равна отношению количества исходов эксперимента Ω, благоприятствующих осуществлению события A, к общему числу всех элементарных исходов Ω.

Это так называемое *классическое определение вероятности*. Однако еще раз подчеркнем, что приведенное выше правило вычисления вероятностей справедливо лишь в экспериментах с *конечным* числом *равновозможных* исходов.

В приложениях важнейшим вопросом является вопрос об источнике «начальных» вероятностей p_j, $j = 1, 2, \ldots$. Теория никаких указаний по этому поводу не дает — «правильным» с точки зрения теории является любой способ задания вероятностей p_j, лишь бы выполнялись условия положительности $p_j > 0$ и нормировки $\sum\limits_{j=1}^{\infty} p_j = 1$. Другое дело, насколько хорошо тот или иной способ задания вероятностей соответствует содержанию конкретной задачи. Рассмотрим нескольких примеров.

Пример 1. Пусть

$$\Omega = \{\Gamma, \mathrm{P}\}$$

— пространство элементарных исходов эксперимента, связанного с бросанием монеты. Принятое выше определение позволяет приписать исходам Г и Р любые значения $p = P\{\Gamma\}$ и $q = P\{\mathrm{P}\}$, лишь бы $p + q = 1$, $p \geqslant 0$, $q \geqslant 0$. Однако, когда речь идет о конкретном эксперименте с конкретной монетой, то принятое распределение вероятностей должно соответствовать индивидуальным характеристикам именно этой монеты и этого эксперимента. Например, если монета физически симметрична, то следует ожидать, что в эксперименте шансы на появление исхода Г и исхода Р должны быть одинаковы, т. е.

$$P\{\Gamma\} = P\{\mathrm{P}\},$$

а это немедленно влечет за собой

$$p = q = \frac{1}{2},$$

т. к. $p + q = 1$. Если же монета несимметрична (например, смещен центр тяжести), то следует ожидать, что один из двух возможных исходов эксперимента будет появляться чаще другого (та сторона монеты, к которой смещен центр тяжести, будет выпадать реже). В этом случае $p \neq q$, и следует так задать p и q, чтобы соотношение шансов на появление исхода Г и исхода Р соответствовало обнаруженной асимметрии. В простейших ситуациях вневероятностные соображения (физическая симметрия эксперимента, равновозможность и т. п.) обычно дают указание на то, как именно следует задавать вероятностные распределения p_j. Эти соображения, конечно, не дают ответа на вопрос о соответствии выбранных значений p_j конкретной ситуации, а являются лишь наводящими. Для установления адекватности принятой вероятностной модели реальному эксперименту следует прибегнуть к специальным статистическим процедурам, о которых будет сказано ниже.

Пример 2. В урне лежит $m + n$ шаров, m — черных, n — белых. Эксперимент состоит в извлечении одного шара из урны. Найти вероятность извлечения черного (белого) шара.

◀ **1.** Предположим, что все шары физически идентичны, тщательно перемешаны и извлекаются «случайным образом» (последнее предположение является интуитивным предположением о честности экспериментатора и отсутствии закономерностей, по крайней мере осознанных, в очередности извлечения черных и белых шаров). Тогда резонно предполагать, что вероятность извлечения шара фиксированного цвета должна быть пропорциональна количеству шаров данного цвета и

$$\frac{P\{\text{«извлечен белый шар»}\}}{P\{\text{«извлечен черный шар»}\}} = \frac{n}{m}.$$

Поскольку события

$$\text{Б} = \{\text{«извлечен белый шар»}\}$$

и

$$\text{Ч} = \{\text{«извлечен черный шар»}\}$$

противоположны ($\overline{\text{Б}} = \text{Ч}$) и несовместны ($\text{Б} \cap \text{Ч} = \varnothing$), имеем

$$\text{Б} + \text{Ч} = \Omega$$

и

$$P\{\text{Б}\} + P\{\text{Ч}\} = 1.$$

Вместе с соотношением, приведенным выше, это дает для искомых вероятностей следующие значения

$$P\{\text{Б}\} = \frac{n}{m+n}, \quad P\{\text{Ч}\} = \frac{m}{m+n}.$$

2. Предположим, что все шары физически идентичны, тщательно перемешаны и извлекаются «случайным образом». Мысленно перенумеруем все шары и опишем рассматриваемый эксперимент следующим списком элементарных исходов

$$\Omega = \{\omega_1^{\text{Б}}, \ldots, \omega_n^{\text{Б}}, \omega_1^{\text{Ч}}, \ldots \omega_m^{\text{Ч}}\},$$

где $\omega_j^{\text{Б}}$, $\omega_i^{\text{Ч}}$ — события, состоящие в извлечении j-го белого или i-го черного шара, соответственно. Предположения об идентичности шаров и условиях их извлечения делают разумным предположение о равновозможности всех элементарных исходов в эксперименте Ω,

$$P\{\omega_j^{\text{Б}}\} = P\{\omega_i^{\text{Ч}}\}, \quad \forall\, j = 1, 2, \ldots, n, \quad i = 1, 2, \ldots, m.$$

Из условия нормировки

$$\sum P\{\omega_j^{\text{Б}}\} + \sum P\{\omega_i^{\text{Ч}}\} = 1$$

немедленно следует, что

$$P(\omega_j^{\text{Б}}) = P(\omega_i^{\text{Ч}}) = \frac{1}{m+n}.$$

Событие $\text{Б} = \{\text{«извлечен белый шар»}\}$ описывается следующим списком благоприятствующих ему элементарных исходов

$$\text{Б} = \left\{\omega_1^{\text{Б}}, \omega_2^{\text{Б}}, \ldots, \omega_n^{\text{Б}}\right\}.$$

В соответствии с соотношением (9)

$$P\{\text{Б}\} = \sum_{j=1}^{n} \omega_j^{\text{Б}} = \frac{n}{m+n}, \quad P\{\text{Ч}\} = 1 - P\{\text{Б}\} = \frac{m}{m+n}. \ ▶$$

Пример 3. Эксперимент состоит в одновременном бросании n симметричных монет. Найти вероятность того, что ровно на $k \leqslant n$ монетах выпадет герб.

◀ Опишем пространство элементарных исходов рассматриваемого эксперимента списком Ω, составленным из всех упорядоченных наборов символов Г и Р длины n:

$$\Omega = \{\underbrace{\text{ГГ}\ldots\text{Г}}_{n}, \underbrace{\text{РГ}\ldots\text{Г}}_{n}, \underbrace{\text{РРГГ}\ldots\text{Г}}_{n}, \ldots, \underbrace{\text{РГГРР}\ldots\text{Г}}_{n}, \ldots \underbrace{\text{РР}\ldots\text{Р}}_{n}\}.$$

Каждый элементарный исход описывает результат одновременного испытания n монет, причем на j-м месте стоит Г или Р (в зависимости от того, гербом или решкой завершилось испытание j-й монеты). Предполагая, что испытания различных монет не оказывают влияния друг на друга, и учитывая симметричность монет, будем считать все исходы равновозможными. Поскольку их всего $N = 2^n$, то из правила сложения и условия нормировки вероятности получаем

$$P(\omega) = \frac{1}{2^n} \quad \forall\, \omega \in \Omega.$$

Событие

$$A_{=k} = \{\text{«ровно на } k \text{ монетах выпал герб»}\}$$

описывается списком, составленным из тех элементарных исходов, которые содержат ровно k символов Г (и, соответственно, $n-k$ символов Р). Например, в эксперименте с 4 монетами

$$A_{=2} = \{\text{ГГРР, ГРГР, ГРРГ, РГРГ, РРГГ, РГГР}\}.$$

Число способов, которыми можно расставить k символов по n позициям, $C_n^k = C_n^{n-k}$, и есть количество элементарных исходов, содержащих ровно k символов Г, поэтому в соответствии с соотношением (9)

$$P\{A_{=k}\} = C_n^k \cdot \frac{1}{2^n}. \ ▶$$

Пример 4. В условиях предыдущего эксперимента найти вероятность того, что не более чем на k монетах выпадет герб.

◀ Пусть искомое событие $A_{\leqslant k}$. Тогда очевидно, что

$$A_{\leqslant k} = A_{=0} + A_{=1} + \ldots + A_{=k},$$

и по правилу сложения

$$P\{A_{\leqslant k}\} = \sum_{j=0}^{k} P\{A_{=j}\} = \sum_{j=0}^{k} C_n^j \cdot \frac{1}{2^n}. \ ▶$$

3.3. Вероятность в экспериментах с недискретной структурой

В случае экспериментов, множество исходов которых более чем счетно, реализовать рассмотренную схему построения вероятностей в том виде, как она была описана в п. 3.1, уже нельзя. Дело в том, что требование (2) — аддитивности вероятности — делает невозможным приписывание вероятностей абсолютно всем событиям, которые могут происходить в эксперименте.

Пример 1. Эксперимент состоит в случайном выборе точки из отрезка $[a, b] \subset \mathbb{R}$. Случайность здесь понимается как равновозможность выбора точки ω из любого подотрезка $[\alpha, \beta]$, независимо от того, в каком месте отрезка $[a, b]$ этот подотрезок расположен. В этой ситуации естественно считать, что вероятность выбора точки из любого подотрезка $[\alpha, \beta]$ пропорциональна его длине

$$P\{\omega \in [\alpha, \beta]\} = \frac{\beta - \alpha}{a - b}.$$

Оказывается, однако, что в рассматриваемом случае существуют события (т. е. подмножества множества $[a, b]$), которым не удается разумным образом приписать вероятность так, чтобы она была согласована с определением, принятым выше. Возьмем для определенности $a = -1$, $b = 2$, т. е. рассмотрим отрезок $[-1, 2]$. Для любого $x \in [0, 1]$ через A_x обозначим множество таких $y \in [0, 1]$, что $x - y \in \mathbb{Q}$, т. е. $x - y$ — рациональное число. Получаем некоторую совокупность множеств из $[0, 1]$. Ясно, что любые два множества A_x и A_y или не имеют общих точек, или совпадают. Действительно, если $z \in A_x$ и $z \in A_y$, то $x - z \in \mathbb{Q}$ и $y - z \in \mathbb{Q}$, отсюда следует, что $x - y \in \mathbb{Q}$, или $y \in A_x$. Следовательно, $A_y = A_x$. Значит, отрезок $[0, 1]$ разбит на непересекающиеся подмножества. Выберем из каждого подмножества разбиения по одному элементу и обозначим через A совокупность всех выбранных элементов. Пусть событие A состоит в выборе точки из множества $A \in [-1, 2]$ и $P\{A\}$ — вероятность этого события. Пусть $r_1, r_2, \ldots, r_n, \ldots$ — рациональные числа из отрезка $[0, 1]$. Заметим, что если $i \neq j$, то $\{A + r_i\} \cap \{A + r_j\} = \varnothing$. По свойству (2), которым должна обладать вероятность, имеем

$$P\left\{\bigcup_{j=1}^{\infty} \{A + r_j\}\right\} = \sum_{j=1}^{\infty} P\{A + r_j\} = \sum_{j=1}^{\infty} P\{A\} = \lim_{n\to\infty} nP\{A\}.$$

Но этот предел не существует, если $P\{A\} > 0$, и равен нулю, если $P\{A\} = 0$. Последнее, однако, невозможно, так как

$$[0, 1] \subset \bigcup_{j=1}^{\infty}\{A + r_j\} \subset [-1, 2],$$

и в силу монотонности вероятности

$$\frac{1}{3} = P\{[0, 1]\} \leqslant P\{\cup_{j=1}^{\infty}\{A + r_j\}\} \leqslant P\{[-1, 2]\} = 1. \blacktriangleright$$

Отсюда следует, что приписать событию A вероятность разумным образом невозможно. Поэтому, либо следует отказаться от требования аддитивности, предъявляемого к вероятностной функции, либо сузить класс событий, подлежащих рассмотрению из числа происходящих в эксперименте, но сохранить свойство (2).

В современной теории вероятностей принят последний подход. Помимо чисто математических соображений, это вызвано тем обстоятельством, что в реальной ситуации всякий эксперимент допускает лишь дискретное описание. Недискретные модели — удобная математическая абстракция, идеализация ситуации, но хотелось бы, чтобы идеализированная модель вероятностного описания эксперимента включала в себя важнейшие черты и характеристики, присущие реальному эксперименту, а поскольку дискретные вероятностные модели, обладающие свойством аддитивности, на практике себя зарекомендовали как весьма полезные и эффективные при прогнозировании явлений, природа которых случайна, то резонно свойство аддитивности вероятностной функции сохранить и в моделях, структура которых более сложна, чем дискретная.

«Подправленная» в соответствии с приведенными выше соображениями схема вероятностного описания экспериментов выглядит так.

Пусть $\Omega = \{\omega\}$ — пространство элементарных исходов эксперимента. Выделим среди *всех* возможных в эксперименте Ω событий некоторую совокупность событий $\mathcal{U}$, обладающих следующими свойствами:

1. $\Omega \in \mathcal{U}$;

2. $\forall\, A \in \mathcal{U} \Rightarrow \overline{A} \in \mathcal{U}$; в частности, $\varnothing \in \mathcal{U}$;

3. $\forall\, A_1, A_2, \dots, A_n, \dots,\ A_i \in \mathcal{U} \Rightarrow \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{U}$.

Заметим, что из указанных свойств совокупности $\mathcal{U}$ вытекает, например, что $\forall\, A, B \in \mathcal{U} \Rightarrow A \cap B \in \mathcal{U}$ (т. к. из равенств де Моргана следует $A \cap B = \overline{\overline{A} \cup \overline{B}}$), а также $A \setminus B \in \mathcal{U}$ (т. к. $A \setminus B = A \cap \overline{B}$).

Совокупность $\mathcal{U}$ событий, обладающую свойствами **1–3**, называют *σ-алгеброй (сигма-алгеброй) событий.*

Вероятность, удовлетворяющую требованиям неотрицательности, нормировки и счетной аддитивности, в соответствии со схемой п. 3.1 определим теперь не для всех событий в эксперименте, а для входящих в сигма-алгебру $\mathcal{U}$. Эти события называются *случайными событиями*, происходящими в эксперименте Ω.

Практически построение σ-алгебры случайных событий осуществляется для конкретного эксперимента следующим образом:

в эксперименте выделяется некоторая совокупность событий, важная с точки зрения исследователя и называемая в дальнейшем совокупностью «наблюдаемых» событий;

множество «наблюдаемых» событий расширяется за счет добавления в него всех отрицаний и не более чем конечных объединений событий из числа «наблюдаемых»;

для всех событий из полученной совокупности определяют вероятность, удовлетворяющую требованиям неотрицательности, нормировки (1) и счетной аддитивности (2);

строится $\mathscr{F}$ — наименьшая σ-алгебра, включающая в себя все указанные выше события (такая всегда существует) и определяются вероятности всех событий, входящих в $\mathscr{F}$, с сохранением определения вероятности, уже принятого для порождающей $\mathscr{F}$ совокупности наблюдаемых событий (это также всегда возможно, причем единственным способом).

Реализация этой программы для недискретных экспериментов произвольной природы достаточно сложна и громоздка как с идейной, так и с технической точки зрения, чтобы быть приведенной в этой книге. Мы остановимся здесь лишь на случае экспериментов, множество элементарных исходов которых может быть описано точками некоторого подмножества $\mathbb{R}^n$.

3.4. Числовая прямая

Допустим, что множество исходов эксперимента Ω может быть описано множеством точек некоторого интервала (полуинтервала, отрезка) $\langle a, b\rangle$ на числовой прямой $\mathbb{R}$ — каждому исходу ω поставлена в соответствие точка x на $\langle a, b\rangle$ и, наоборот, каждой точке $x \in \langle a, b\rangle$ отвечает некоторый элементарный исход $\omega \in \Omega$ (случаи $a = -\infty$, $b = +\infty$ не исключаются).

Однократное проведение эксперимента Ω будем интерпретировать как выбор точки из промежутка $\langle a, b\rangle$. «Наблюдаемыми» в этом эксперименте событиями будем считать события, состоящие в выборе точки из интервала $(\alpha, \beta) \subset \langle a, b\rangle$. Расширим множество «наблюдаемых» событий (так, как было указано выше), добавив к нему события, получающиеся из «наблюдаемых» в результате допустимых действий над последними: построения противоположных событий, а также всевозможных не более чем счетных объединений и пересечений.

Пусть теперь $\Phi(x)$ — произвольная монотонно неубывающая ограниченная функция, определенная на всей числовой прямой и такая, что для любого «наблюдаемого» события $A = \{\omega\colon\ \alpha \leqslant \omega < \beta\}$

$$P\{A\} = P\{\alpha \leqslant \omega < \beta\} = \Phi(\beta) - \Phi(\alpha). \tag{10}$$

Для выполнения условия нормировки вероятности следует потребовать, чтобы $\Phi(x)$ дополнительно удовлетворяла условию

$$\lim_{x\to\alpha+} \Phi(x) = 0, \qquad \lim_{x\to\beta-} \Phi(x) = 1. \tag{11}$$

В силу монотонности и ограниченности $\Phi(x)$ пределы (11) существуют.

Далее, если

$$B = A_1 + A_2,$$

где $A_1 = \{\omega\colon\ \alpha_1 \leqslant \omega < \beta_1\}$, $A_2 = \{\omega\colon\ \alpha_2 \leqslant \omega < \beta_2\}$, $\alpha_1 < \beta_1 \leqslant \alpha_2 < \beta_2$, то

$$P\{B\} = P\{A_1\} + P\{A_2\}, \tag{12}$$

где $P\{A_i\}$ определены соотношением (10).

Таким образом, вероятность определена на конечном объединении промежутков в $\langle a, b\rangle$. Можно показать, что она будет удовлетворять требованию σ-аддитивности. Отсюда следует, что вероятность определена на наименьшей σ-алгебре, содержащей промежутки.

Определение. Наименьшая σ-алгебра, содержащая все интервалы из $\langle a, b\rangle$, называется *борелевской σ-алгеброй* промежутка $\langle a, b\rangle$, а множества, входящие в борелевскую σ-алгебру — *борелевскими*.

Итак, с помощью ограниченной неубывающей функции $\Phi(x)$ мы определили вероятность на борелевской σ-алгебре $\langle a, b\rangle$. Этот способ задания вероятности является универсальным, так как если на борелевской σ-алгебре промежутка $\langle a, b\rangle$ задана произвольная вероятность, то она определяется по функции

$$\boxed{\Phi(x) = P\{\omega\colon\ a \leqslant \omega < x\}}$$

вышеуказанным способом.

Пример 1. Дискретные вероятностные пространства. Пусть

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

— конечное или счетное множество точек на прямой и

$$p_1, p_2, \ldots, p_n, \ldots$$

— последовательность чисел таких, что

$$0 \leqslant p_i \leqslant 1 \quad \text{и} \quad \sum_{i=1}^{\infty} p_i = 1.$$

Определим вероятность произвольного множества $A \subset \mathbb{R}$ так

$$P\{A\} = \sum_{x_i \in A} p_i$$

(складываются числа p_i с теми же номерами, с которыми члены последовательности $\{x_n\}$ входят в A). Ясно, что определенная таким образом вероятность задается с помощью функции

$$\Phi(x) = \sum_{x_i < x} p_i.$$

Если все точки $x_1, x_2, \ldots, x_n, \ldots$ — изолированные, то $\Phi(x)$ — ступенчатая функция с разрывами в точках $\{x_i\}$, причем

$$\Phi(x_i + 0) - \Phi(x_i - 0) = p_i.$$

(рис. 23). Если же точки последовательности $\{x_n\}$ не изолированы, то $\Phi(x)$ может иметь довольно сложную структуру и ее график так просто изобразить невозможно.

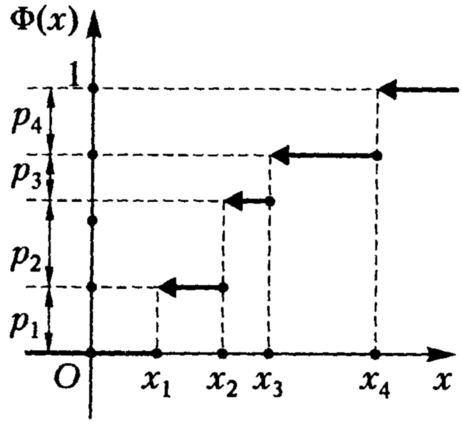


Рис. 23

Пример 2. Равномерное распределение на отрезке. Если функция $\Phi(x)$ линейна на отрезке $[a, b]$, $\Phi(x) = 0$ при $x < a$ и $\Phi(x) = 1$ при $x > b$, то такой способ определения вероятности называется *равномерным распределением* вероятности на $[a, b]$, т. к. вероятность отрезка $[\alpha, \beta]$ пропорциональна его длине и не зависит от его расположения внутри $[a, b]$ (рис. 24).

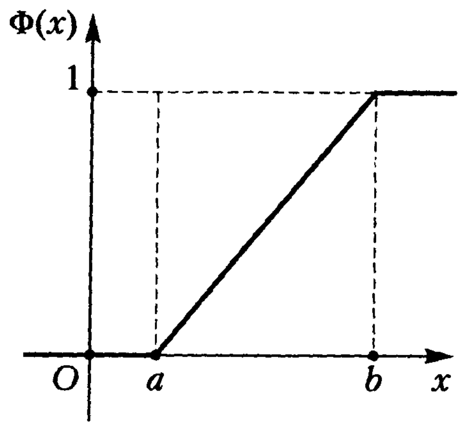


Рис. 24

Пример 3. Пусть существует неотрицательная интегрируемая функция $f(x)$ такая, что

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Для любого $x \in \mathbb{R}$ положим

$$\Phi(x) = \int_{-\infty}^{x} f(t)\,dt.$$

Определенное таким образом распределение вероятности называется *абсолютно непрерывным распределением*, а функция $f(x)$ — *плотностью* этого распределения. Легко видеть, что рассмотренное в примере 2 равномерное распределение является абсолютно непрерывным с плотностью

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{при } x \in (a, b), \\ 0 & \text{при } x \notin (a, b). \end{cases}$$

Дискретное распределение (см. пример 1) не является абсолютно непрерывным, т. к. вероятность одноточечного множества при абсолютно непрерывном распределении равна 0 ($\Phi(x)$ — непрерывная функция), а при дискретном положительна.

3.5. Координатная плоскость в $\mathbb{R}^2$

Предположим, что множество исходов эксперимента Ω может быть описано множеством $D \subset \mathbb{R}^2$, т. е. существует взаимно однозначное соответствие между точками множества D и исходами эксперимента Ω. «Наблюдаемыми» событиями будем считать события, состоящие в выборе точки из прямоугольника, принадлежащего D, с включенными или нет отрезками границы. Определив вероятность для «наблюдаемых» событий, распространим ее вначале на конечные объединения и пересечения прямоугольников, а потом и на наименьшую σ-алгебру, содержащую прямоугольники и называемую, как и в одномерном случае, борелевской.

Аналогично одномерному случаю вероятность можно задавать с помощью функции

$$\Phi(x, y) = P\big\{(u, v) \in D\colon\ u < x,\ v < y\big\};$$

однако здесь (тем более при переходе к $\mathbb{R}^n$ для $n > 2$) эта схема весьма громоздка и редко применяется на практике.

Пример 1. Дискретные распределения. Пусть $x_1, x_2, \dots, x_n, \dots$ — конечное или счетное множество точек плоскости $\mathbb{R}^2$ (или $D \subset \mathbb{R}^2$) и $p_1, p_2, \dots, p_n, \dots$ — последовательность неотрицательных чисел таких, что $\sum\limits_{i=1}^{\infty} p_i = 1$. Вероятность множества $A \subset \mathbb{R}^2$ определим так

$$p\{A\} = \sum_{x_i \in A} p_i.$$

Это и есть дискретное распределение вероятностей (полная аналогия с одномерным случаем).

Пример 2. Абсолютно непрерывные распределения. Пусть $f(x, y)$ — неотрицательная интегрируемая функция на $\mathbb{R}^2$ (или в D) и

$$\iint f(x, y)\, dx\, dy = 1.$$

Так же, как и в одномерном случае при абсолютно непрерывном распределении, вероятностью множества A называем число

$$P\{A\} = \iint\limits_A f(x, y)\, dx\, dy.$$

Аналогичным образом определяются вероятности в пространствах элементарных исходов, равномощных некоторым кубируемым подмножествам в $\mathbb{R}^n$. «Наблюдаемыми» будем считать события, состоящие в выборе случайной точки из прямоугольного параллелепипеда (с включенными или нет гранями — это вопрос соглашения), и действуя далее точно так же, как в $\mathbb{R}^2$.

3.6. Геометрические вероятности

Важным для приложений частным случаем изложенных выше схем являются так называемые *геометрические вероятности.*

Пусть Ω — кубируемое подмножество $\mathbb{R}^n$, $\mu(\Omega)$ — его мера (при $n = 1$ — длина, при $n = 2$ — площадь, при $n = 3$ — объем и т. п.). События — борелевские подмножества Ω, т. е. наименьшая σ-алгебра, порожденная параллелепипедами

$$\Pi_n = \{\omega\colon\ a_j \leqslant \omega_j < b_j,\ j = \overline{1, n}\}.$$

По определению, вероятность пропорциональна мере

$$P\{\Pi_n\} = \frac{\mu(\Pi_n)}{\mu(\Omega)}. \tag{13}$$

В частности:

если Ω — кривая на плоскости, начинающаяся в точке M_1 и заканчивающаяся в точке M_2, и случайное событие — выбор точки из дуги A (рис. 25), то вероятность этого события

$$P\{A\} = \frac{l(A)}{l(\Omega)}, \tag{14}$$

где $l(A)$ — длина дуги A;

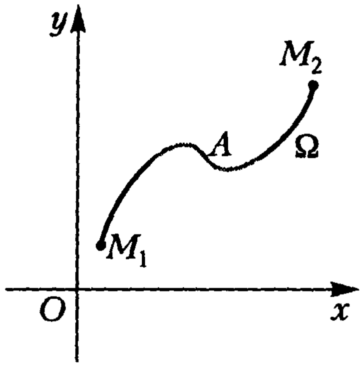


Рис. 25

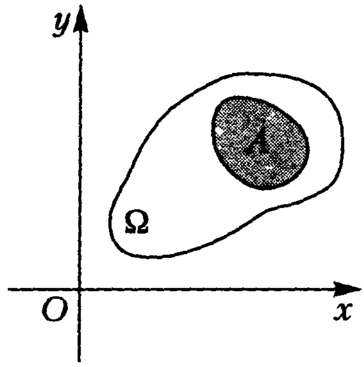


Рис. 26

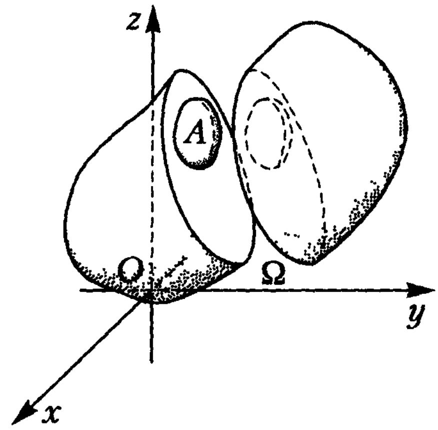


Рис. 27

если Ω — область на плоскости (рис. 26), то

$$P\{A\} = \frac{S(A)}{S(\Omega)}, \tag{15}$$

где $S(A)$ — площадь множества A;

если Ω — тело в пространстве (рис. 27), то

$$P\{A\} = \frac{V(A)}{V(\Omega)}, \tag{16}$$

где $V(A)$ — объем множества A.

Как правило, геометрические вероятности возникают в ситуации, когда множество всех возможных исходов эксперимента можно интерпретировать как «геометрическое», а сам эксперимент обладает физической (или геометрической) симметрией, обеспечивающей «равновозможность» всех исходов; при этом вероятность равномерно распределена по всему множеству.

Рассмотрим несколько примеров.

Пример 1. Точка случайным образом брошена на окружность радиуса 1. Какова вероятность того, что длина проекции радиус-вектора этой точки на выбранное направление не превышает 0,5?

◀ Пусть выбранное направление совпадает с направлением оси Ox, а центр окружности лежит в точке $x = 0$, $y = 0$ (рис. 28). Случайность бросания позволяет считать, что вероятность попадания точки на любую дугу пропорциональна длине этой дуги и не зависит от ее положения на окружности. Тогда для любого события $A = \{\omega\colon\ \omega \in L\}$ в соответствии с (14) имеем

$$P\{A\} = \frac{l(L)}{2\pi}.$$

Для того, чтобы длина проекции не превышала 0,5, нужно, чтобы точка попала либо на дугу M_0M_1, либо на дугу M_2M_3. Поэтому $B = \{\omega\colon\ \omega \in \smile M_0M_1 \cup \smile M_2M_3\}$ и

$$P\{B\} = \frac{l(M_0M_1)}{\pi} = \frac{1}{3}. ▶$$

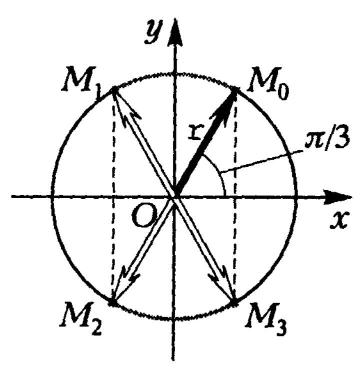


Рис. 28

Пример 2. Пусть (ξ, η) — координаты точки, выбранной случайным образом из единичного квадрата $0 \leqslant x, y \leqslant 1$. Найти вероятность того, что корни уравнения

$$x^2 + \xi x + \eta = 0$$

а) действительны; б) оба положительны

◀ а) Корни уравнения действительны, если выполняется соотношение

$$\xi^2 - 4\eta \geqslant 0.$$

Множество $\{(\xi, \eta)\colon\ \eta \leqslant \xi^2/4;\ 0 \leqslant \xi \leqslant 1;\ 0 \leqslant \eta \leqslant 1\}$ на рис. 29 заштриховано. Получаем

$$P\{\text{«корни действительны»}\} = \int_0^1 \frac{\xi^2}{4}\, d\xi = \frac{1}{12}.$$

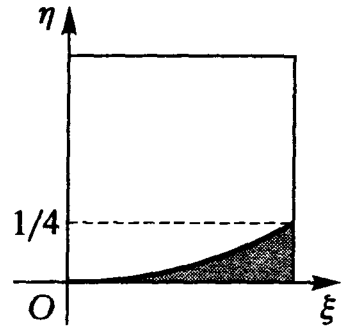


Рис. 29

б) Легко видеть, что искомая вероятность равна нулю, так как положительность корней требует отрицательности ξ (по теореме Виета $x_1 + x_2 = -\xi > 0$), что невозможно. ▶

Пример 3 (задача Бюффона). Плоскость разграфлена параллельными прямыми, отстоящими одна от другой на расстоянии L. Найти вероятность того, что наудачу брошенная игла длины $l < L$ пересечет какую-нибудь из этих прямых.

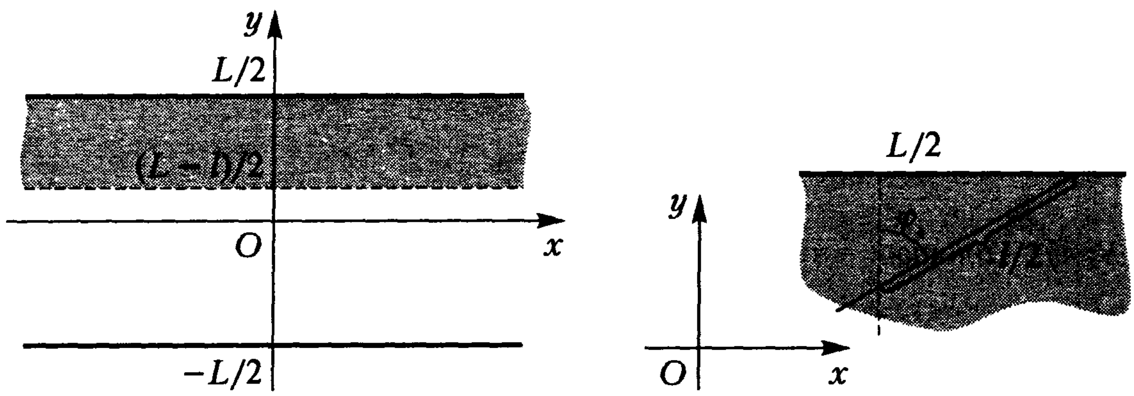


Рис. 30

◀ Игла, брошенная наудачу на плоскость, разграфленную параллельными прямыми, однозначно фиксируется следующими координатами: положением ее середины и углом поворота относительно некоторого фиксированного направления. В силу периодичности картины достаточно рассмотреть одну полосу ширины L. Систему координат введем так: начальную точку O возьмем на середине полосы, ось абсцисс проведем параллельно прямым, ограничивающим полосу, а ось ординат — перпендикулярно. Пусть y — ордината середины иглы, брошенной на полосу. Считая, что положение иглы случайно, полагаем вероятность попадания середины в любой промежуток внутри $[-L/2;\ L/2]$ пропорциональной длине промежутка. Игла пересекает границу полосы, если выполнены следующие условия:

$$\frac{L}{2} > |y| > \frac{L-l}{2}$$

и угол φ, составленный иглой с положительным направлением оси ординат, не меньше предельного угла для этой ординаты,

$$\cos\varphi \geqslant \frac{L/2 - |y|}{l/2} = \cos\varphi_0$$

(рис. 30). Возможные положения иглы на плоскости (φ, y) описываются точками прямоугольника

$$\Pi = \left\{(\varphi, y)\colon\ -\tfrac{\pi}{2} \leqslant \varphi \leqslant \tfrac{\pi}{2};\ -\tfrac{L}{2} \leqslant y \leqslant \tfrac{L}{2}\right\}$$

а множество M положений иглы, при которых она пересекает прямую, описывается так

$$M = \left\{(\varphi, y)\colon\ \cos\varphi \geqslant \tfrac{L-2|y|}{l};\ \tfrac{L-l}{2} < |y| < \tfrac{L}{2}\right\}.$$

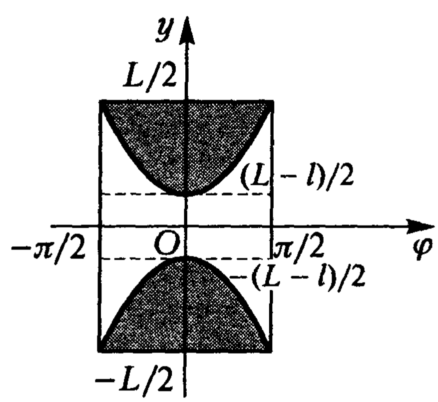


Рис. 31

Для искомой вероятности получаем

$$P = \frac{S(M)}{S(\Pi)} = \frac{2l}{L\pi}.\ ▶$$

§ 4. Условные вероятности

Пусть Ω — эксперимент, $\mathscr{F}$ — совокупность случайных событий, $P\{A\}$ — правило вычисления вероятностей для любого случайного события $A \in \mathscr{F}$. Часто возникает потребность изучения совместного осуществления нескольких случайных событий в следующей постановке: известно, что в результате однократного осуществления эксперимента Ω реализовалось случайное событие B. Что можно сказать о степени достоверности осуществления события A? Ясно, что если события A и B несовместны, то осуществление события B означает невозможность осуществления события A в том же испытании. Если же событие A подчинено событию B (B влечет за собой A, $B \subseteq A$), то осуществление B влечет за собой осуществление A (рис. 32). Однако эти две ситуации в известном смысле крайние, и для любого другого события A достоверность его осуществления вместе с B должна быть тем выше, чем большую долю в B составляют элементарные исходы A (рис. 33).

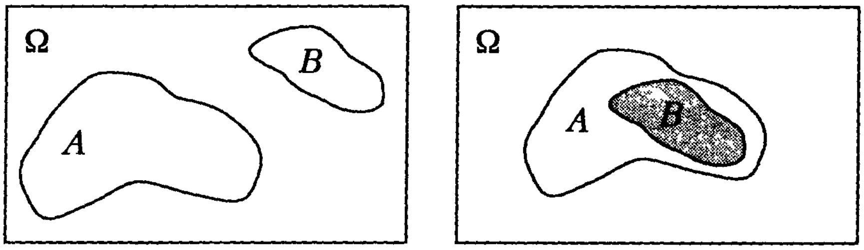


Рис 32

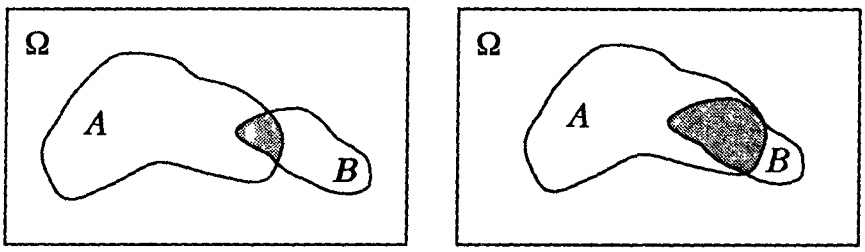


Рис. 33

Вышеизложенное дает основание для следующего определения.

Определение. *Условной вероятностью* случайного события A относительно случайного события B, $P\{B\} > 0$, называется число

$$P\{A|B\} = \frac{P\{A \cap B\}}{P\{B\}}. \tag{1}$$

Отметим очевидные свойства условной вероятности:

1. A: $0 \leqslant P\{A|B\} \leqslant 1$.

◀ $P\{AB\} \leqslant P\{B\}$, поскольку $AB \subseteq B$. ▶

2. $P\{B|B\} = 1$.

3. $P\{\varnothing \,|B\} = 0$, $P\{\Omega|B\} = 1$.

4. Если $A_1 \cap A_2 = \varnothing$, то $P\{A_1 \cup A_2|B\} = P\{A_1|B\} + P\{A_2|B\}$.

◀ Из свойства дистрибутивности сложения событий

$$\left.\begin{matrix} (A_1 \cup A_2)B = A_1B \cup A_2B, \\ A_1B \cap A_2B = \varnothing \end{matrix}\right\} \Rightarrow P\{A_1B \cup A_2B\} = P\{A_1B\} + P\{A_2B\}.$$

Отсюда следует искомое. ▶

5. $P\{A|\Omega\} = P\{A\}$.

Условные вероятности позволяют ввести в рассмотрение одну важную теоретико-вероятностную конструкцию, связанную с осуществившимся в эксперименте Ω случайным событием B (ненулевой вероятности) — конструкцию *сужения*, или *ограничения*, эксперимента Ω на событие B.

Рассмотрим эксперимент Ω_B такой, что

$$\Omega_B = B, \quad \mathcal{F}_B = \{\Omega_B, \varnothing_B, \ldots, A_B, \ldots\},$$

$$\forall\, A \subseteq \Omega \;\; A_B = A \cap B, \quad P\{A_B\} = \frac{P\{AB\}}{P\{B\}} = P\{A \mid B\}$$

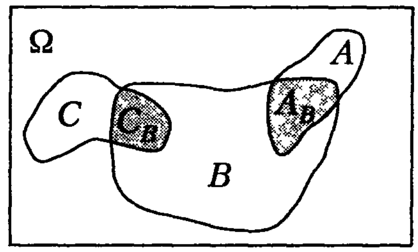


Рис. 34

(см. рис. 34). (Заметим, что $\mathcal{F}_B$ — является сигма-алгеброй.)

Определение эксперимента Ω_B корректно в том смысле, что любое событие из σ-алгебры случайных событий $\mathcal{F}_B$ обладает вероятностью. Эксперимент Ω_B называется *сужением* эксперимента Ω.

Пример 1. Эксперимент Ω состоит в одновременном подбрасывании двух игральных кубиков [1)] с последующей фиксацией значений, оказавшихся на верхних гранях кубиков. Найти вероятность того, что на кубиках выпали разные числа, если известно, что сумма выпавших очков оказалась равна 8.

◄ Опишем пространство элементарных исходов рассматриваемого эксперимента списком Ω, элементами которого являются упорядоченные наборы пар чисел (i, j), где i — число, выпавшее на первом кубике, j — на втором. Имеем

$$\Omega = \{(i, j), \quad i = 1, 2, \ldots, 6; \; j = 1, 2, \ldots, 6\}.$$

Число всех возможных исходов $|\Omega| = 36$. В силу физической симметрии эксперимента положим

$$P\{(i, j)\} = \frac{1}{|\Omega|} = \frac{1}{36}, \quad \forall\, i, j.$$

Событие

$$A = \{(i, j)\colon \; i \neq j\}$$

состоит из 30 равновозможных элементарных исходов, поэтому

$$P\{A\} = \frac{30}{36} = \frac{5}{6};$$

событие

$$B = \{(i, j)\colon \; i + j = 8\}$$

из 5 равновозможных исходов —

$$P\{B\} = \frac{5}{36}.$$

Совмещение событий A и B имеет следующую структуру:

$$A \cap B = \{(i, j)\colon \; i \neq j, \; i + j = 8\} = \{(2, 6); (3, 5); (5, 3); (6, 2)\}$$

Формула условной вероятности дает

$$P\{A \mid B\} = \frac{P\{AB\}}{P\{B\}} = \frac{4/36}{5/36} = \frac{4}{5}.$$

Тот же результат можно получить, рассмотрев сужение эксперимента Ω, индуцированное событием B. В этом конкретном эксперименте имеем

$$\Omega_B = \{(i, j)\colon \; i + j = 8\} = \{(2, 6); (3, 5); (5, 3); (6, 2)\},$$

$$A_B = \{(2, 6); (3, 5); (5, 3); (6, 2)\}, \quad P\{A_B\} = \frac{4}{5}.$$

[1)] Игральный кубик — симметричный однородный куб, на каждую грань которого нанесены символы от 1 до 6 с соблюдением следующего правила: сумма значений на противоположных гранях равна 7. Таких кубиков существует ровно два, но для дальнейшего не важно, о каком из них идет речь.

Проиллюстрируем приведенные выше рассуждения таблицей

(1, 6)	$\boxed{(2, 6)}$	(3, 6)	(4, 6)	(5, 6)	(6, 6)
(1, 5)	(2, 5)	$\boxed{(3, 5)}$	(4, 5)	(5, 5)	(6, 5)
(1, 4)	(2, 4)	(3, 4)	$\boxed{(4, 4)}$	(5, 4)	(6, 4)
(1, 3)	(2, 3)	(3, 3)	(4, 3)	$\boxed{(5, 3)}$	(6, 3)
(1, 2)	(2, 2)	(3, 2)	(4, 2)	(5, 2)	$\boxed{(6, 2)}$
(1, 1)	(2, 1)	(3, 1)	(4, 1)	(5, 1)	(6, 1)

▶

Отметим еще несколько полезных соотношений, связанных с понятием условной вероятности.

4.1. Правило умножения вероятностей

$$\boxed{P\{AB\} = P\{A\,|\,B\}P\{B\}, \quad P\{B\} > 0.} \tag{2}$$

Это соотношение очевидным образом следует из определения (1) и может быть легко обобщено на случай n совмещаемых событий

$$\boxed{P\{A_1A_2\ldots A_n\} = P\{A_1\,|\,A_2A_3\ldots A_n\}P\{A_2\,|\,A_3A_4\ldots A_n\}\ldots P\{A_{n-1}\,|\,A_n\}P\{A_n\},} \tag{3}$$

если только $P\{A_2A_2\ldots A_n\} > 0$.

Пример 1. Список экзаменационных вопросов содержит 25 вопросов, из которых студент подготовил только 20. Экзаменатор задает ему два вопроса, выбирая их из списка случайным образом. Найти вероятность того, что студент знает ответ на оба предложенных ему вопроса.

◀ 1. Один из способов получения ответа на поставленный вопрос следующий: пространство элементарных исходов рассматриваемого эксперимента Ω состоит из всевозможных (вообще говоря, неупорядоченных) пар номеров $\{(i, j),\ \ i \neq j,\ \ i, j = \overline{1,25}\}$ вопросов, которые может предложить преподаватель. Всего их $|\Omega| = C_{25}^2$. Событие A, состоящее в том, что студент знает ответ на оба предложенных вопроса, состоит из тех пар (i, j) вопросов, которые он подготовил. Их C_{20}^2. Учитывая случайность выбора вопросов преподавателем можем считать все исходы в Ω равновозможными, а потому

$$P\{A\} = \frac{C_{20}^2}{C_{25}^2} = \frac{20\cdot 19}{25\cdot 24} = \frac{19}{30}.$$

2. Другой способ — использование понятия условной вероятности: пусть событие B — «первый из предложенных студенту вопросов ему известен», событие C — «второй из предложенных студенту вопросов ему известен». Тогда событие A — «студент знает ответ на оба предложенные ему вопроса» — есть совмещение событий B и C и по правилу умножения (2)

$$P\{A\} = P\{BC\} = P\{C\,|\,B\}P\{B\}. \tag{4}$$

Но $P\{B\} = 20/25$ — вероятность извлечь из совокупности 25 равновозможных вопросов любой из 20 известных студенту, тогда вероятность $P\{C\,|\,B\}$ извлечь из оставшихся 24 вопросов любой из оставшихся 19 известных студенту дается соотношением

$$P\{C\,|\,B\} = \frac{19}{24},$$

и теперь осталось применить формулу умножения (4). ▶

4.2. Формула полной вероятности

Пусть случайные события $H_1, H_2, \ldots H_n$ попарно несовместны, $P\{H_j\} > 0$ и $H_1 + H_2 + \ldots + H_n = \Omega$. Совокупность событий $\{H_j\}$ будем называть *полной группой несовместных событий* в эксперименте Ω. Тогда для любого случайного события A справедливо равенство

$$\boxed{P\{A\} = P\{A\,|\,H_1\}P\{H_1\} + \ldots + P\{A\,|\,H_n\}P\{H_n\}.} \tag{5}$$

◀ Действительно, из

$$\bigcup_{j=1}^{n} H_j = \Omega$$

следует

$$\bigcup_{i=1}^{n} AH_i = A$$

и при этом

$$AH_i \cap AH_j = \varnothing\ .$$

По правилу сложения вероятностей отсюда получаем

$$P\{A\} = P\{AH_1\} + P\{AH_2\} + \ldots + P\{AH_n\}.$$

Правило же умножения (3) дает

$$P\{AH_j\} = P\{A\,|\,H_j\}P\{H_j\},$$

откуда и следует требуемое. ▶

Заметим, что в рамках принятой теории соотношение (5) формализует естественный перебор альтернатив: если $H_1, H_2, \ldots, H_n$ — альтернативные события, то A происходит либо вместе с H_1, либо с $H_2, \ldots$, либо с H_n. Следующий ниже пример иллюстрирует эту мысль.

Пример. Известно, что 5% мужчин и 0,25% женщин — дальтоники. Найти вероятность того, что наугад выбранный человек — дальтоник (предполагается, что число мужчин $N_м$ и число женщин $N_ж$ одинаковы, $N_м = N_ж = N$).

◀ Вероятность того, что наугад выбранный человек — дальтоник, должна быть, очевидно, пропорциональна общему количеству дальтоников $N_д$, которое в свою очередь является суммой количества дальтоников-мужчин $N_{дм}$ (5% от количества мужчин) и количества дальтоников-женщин $N_{дж}$ (0,25% от количества женщин):

$$N_д = 0{,}05N_м + 0{,}0025N_ж = N(0{,}05 + 0{,}0025) = N(0{,}05 + 0{,}0025) = 0{,}0525 \cdot N.$$

Доля дальтоников дается соотношением

$$P\{\text{«человек — дальтоник»}\} = \frac{0{,}0525 \cdot N}{2N} = 0{,}02625.$$

Проведенные наводящие рассуждения формализуются следующим образом: пусть Д — событие, состоящее в том, что наугад выбранный человек — дальтоник, М — что он мужчина, Ж — женщина. Заметим, что события М и Ж в рассматриваемом эксперименте (случайный выбор индивидуума из содержащей одинаковое число мужчин и женщин совокупности) образуют полную группу несовместных событий. В силу $N_м = N_ж$ можно считать что

$$P\{М\} = P\{Ж\} = \frac{1}{2}.$$

Для вероятности события Д из (5) получаем

$$P\{Д\} = P\{Д\,|\,М\}P\{М\} + P\{Д\,|\,Ж\}P\{Ж\}.$$

В качестве $P\{Д\,|\,М\}$ разумно, исходя из условия задачи, взять долю дальтоников среди мужчин — $P\{Д\,|\,М\} = 0{,}05$. Аналогично и для $P\{Д\,|\,Ж\} = 0{,}0025$. Поэтому

$$P\{Д\} = 0{,}05 \cdot \frac{1}{2} + 0{,}0025 \cdot \frac{1}{2} = 0{,}02625.\ ▶$$

4.3. Формула Бейеса

Формула условной вероятности (1) может быть прочитана и следующим образом: пусть A и H — два события ненулевой вероятности. Формальными выкладками получаем,что поскольку $P\{AH\} = P\{HA\} = P\{A\,|\,H\}P\{H\}$, то

$$\boxed{P\{H\,|\,A\} = \frac{P\{AH\}}{P\{A\}} = \frac{P\{A\,|\,H\}P\{H\}}{P\{A\}}.} \qquad (6)$$

Последняя формула интерпретируется так: пусть a priori (до проведения испытания) мы оценили вероятности событий A и H. Проведем эксперимент Ω и, если событие A осуществилось, пересчитаем вероятность события H в соответствии с (6). Получим $P\{H\,|\,A\}$ — *апостериорную* (послеопытную) вероятность H, которая уточняет наше знание о возможностях осуществления события H в последующих испытаниях. Эта интерпретация становится особо интересной, если применить формулу (6) к полной системе несовместных событий $H_1, H_2, \ldots, H_n$:

$$\boxed{\begin{array}{l} P\{H_j\,|\,A\} = \dfrac{P\{A\,|\,H_j\}P\{H_j\}}{P\{A\}} = \dfrac{P\{A\,|\,H_j\}P\{H_j\}}{\sum\limits_{i=1}^{n} P\{A\,|\,H_i\}P\{H_i\}} \\ \forall\, j = 1, 2, \ldots, n. \end{array}} \tag{7}$$

Соотношение (7) позволяет при известных априорных вероятностях событий H_j,

$$P\{H_j\}, \quad \sum_{j=1}^{n} P\{H_j\} = 1,$$

и сведении о том, какое событие A осуществилось в данном испытании, найти апостериорное распределение вероятностей

$$P\{H_j\,|\,A\}, \quad \sum_{j=1}^{n} P\{H_j\,|\,A\} = 1.$$

События H_j в этой трактовке носят имя «гипотез», а само соотношение (7) служит основой весьма плодотворного подхода в современной статистике, носящем имя «бейесовского» подхода. Формула (7) называется *формулой Бейеса* или *формулой гипотез.*

Пример. Оператор на мониторе локатора может зафиксировать сигнал, являющийся следом либо объекта, либо его имитации. При этом в одном случае из пяти оператор ошибается. Известно, что в поле наблюдения оператора объекты появляются вдвое реже, чем их имитации. Оператор идентифицировал сигнал на экране монитора как сигнал от объекта. Какова вероятность того, что в поле наблюдения находится действительно объект, а не его имитация?

◀ Пусть H_1 (H_2) — гипотеза, состоящая в том, что в поле наблюдения находится объект (имитация). Событие O — «сигнал на экране монитора идентифицирован как сигнал от объекта». Задача состоит в нахождении вероятности $P\{H_1\,|\,O\}$. По формуле (7) получаем

$$P\{H_1\,|\,O\} = \frac{P\{O\,|\,H_1\}P\{H_1\}}{P\{O\}}.$$

Условия задачи дают основания для следующей оценки априорных вероятностей:

$$P\{H_1\} = \frac{1}{3}; \quad P\{H_2\} = \frac{2}{3}$$

$\left(\text{так как } \frac{P\{H_1\}}{P\{H_2\}} = \frac{1}{2},\ P\{H_1\} + P\{H_2\} = 1\right)$;

$$P\{O\,|\,H_1\} = \frac{4}{5}; \quad P\{O\,|\,H_2\} = \frac{1}{5}.$$

Окончательно получаем

$$P\{H_1\,|\,O\} = \frac{4/5 \cdot 1/3}{4/5 \cdot 1/3 + 1/5 \cdot 2/3} = \frac{2}{3}. \blacktriangleright$$

4.4. Независимость случайных событий

Формула условной вероятности позволяет ввести в рассмотрение еще одно важное понятие — понятие зависимости и независимости случайных событий. Если трактовать независимость случайных событий в некотором эксперименте как отсутствие влияния осуществления или неосуществления одного или нескольких событий на осуществление или неосуществление другого или других событий, то разумно оценивать степень такой зависимости изменением вероятности осуществления событий второй группы при наличии информации об осуществлении событий первой группы. Приведенное соображение дает основание для следующего определения.

Определение. Будем говорить, что событие A *не зависит от события* B, если

$$P\{A\,|\,B\} = P\{A\}, \;\; P\{B\} > 0. \tag{8}$$

Сразу же отметим, что несмотря на внешнюю асимметрию данного определения, понятие независимости взаимно: *если A не зависит от B, то и B не зависит от A.*

◀ Действительно, если событие A не зависит от события B, то правило умножения вероятностей (2) принимает вид

$$P\{AB\} = P\{A\,|\,B\}P\{B\} = P\{A\}P\{B\}. \tag{9}$$

Отсюда легко получаем

$$P\{B\,|\,A\} = \frac{P\{AB\}}{P\{A\}} = \frac{P\{A\}P\{B\}}{P\{A\}} = P\{B\},$$

если только $P\{A\} > 0$. Последнее соотношение и означает независимость B от A в смысле определения (8). ▶

Для событий положительной вероятности соотношение (9) и определение (8) равносильны, и любое из них может быть принято в качестве определения независимости пары случайных событий.

Если события A и B независимы, то также независимы следующие пары событий:

$$A \text{ и } \overline{B}, \quad \overline{A} \text{ и } B, \quad \overline{A} \text{ и } \overline{B},$$

что легко проверить, воспользовавшись определением (8) и тем обстоятельством, что независимы пары событий (A, Ω) и (B, Ω).

Пример 1. В эксперименте с бросанием монеты события Г (выпадение герба) и Р (выпадение решки), очевидно, зависимы, так как осуществление одного из них влечет за собой невозможность осуществления другого. Формально

$$P\{\text{Г}\} = \frac{1}{2}, \quad P\{\text{Г}\,|\,\text{Р}\} = 0 \;\Rightarrow\; P\{\text{Г}\,|\,\text{Р}\} \neq P\{\text{Г}\}.$$

•

Пример 2. По ведомостям о расходе запасных частей было установлено, что при ремонте автомобильных двигателей деталь A была заменена в среднем в 36 % случаев, деталь B — в 42 % случаев, а обе эти детали одновременно заменялись в среднем в 30 % случаев. Связаны ли между собой выход из строя детали A и детали B?

◀ Пусть A — событие, состоящее в том, что деталь A вышла из строя и подлежит замене; B — вышла из строя и подлежит замене деталь B. Условие задачи дает основание для следующей оценки вероятностей событий A, B и $A \cap B$:

$$P\{A\} = 0{,}36, \quad P\{B\} = 0{,}42, \quad P\{A \cap B\} = 0{,}3.$$

По формуле условной вероятности получаем

$$P\{A\,|\,B\} = \frac{0{,}3}{0{,}42} \approx 0{,}71 \neq P\{A\} = 0{,}36.$$

Это означает, что замена детали A при условии, что заменяется деталь B, происходит почти вдвое чаще, чем безусловная замена этой детали, что, конечно, свидетельствует о наличии связи между выходом из строя деталей A и B.

Определение. Будем говорить, что события $A_1, A_2, \ldots, A_n$ *независимы в совокупности*, если для любого r, $1 \leqslant r \leqslant n$, любых k, s, $k + s = r$, и любых $i_1, i_2, \ldots, i_k$, $j_1, j_2, \ldots, j_s$, $i_m \neq i_l$, $j_m \neq j_l$, выполняются соотношения

$$P\{A_{i_1}A_{i_2}\ldots A_{i_k} \,|\, A_{j_1}A_{j_2}\ldots A_{j_s}\} = P\{A_{i_1}A_{i_2}\ldots A_{i_k}\}. \tag{10}$$

Например, независимость в совокупности событий A_1, A_2, A_3 по определению (10) означает, что они независимы попарно, т. е.

$$P\{A_1 \,|\, A_2\} = P\{A_1\}, \quad P\{A_1 \,|\, A_3\} = P\{A_1\}, \quad P\{A_2 \,|\, A_3\} = P\{A_2\},$$

но кроме того, должны иметь место еще и соотношения

$$P\{A_1 \,|\, A_2A_3\} = P\{A_1\}, \quad P\{A_2 \,|\, A_1A_3\} = P\{A_2\}, \quad P\{A_3 \,|\, A_1A_2\} = P\{A_3\},$$

или им эквивалентные, например

$$P\{A_2A_3 \,|\, A_1\} = P\{A_2A_3\}, \quad P\{A_1A_3 \,|\, A_2\} = P\{A_1A_3\}, \quad P\{A_1A_2 \,|\, A_3\} = P\{A_1A_2\}.$$

Приводимый ниже пример иллюстрирует недостаточность попарной независимости для обеспечения независимости событий в совокупности.

Пример 3. В урне лежат четыре геометрически идентичных шара, три из которых окрашены соответственно в красный, синий и зеленый цвета, а окраска четвертого шара содержит все эти три цвета. Эксперимент состоит в извлечении шара из урны.

Обозначим события следующим образом:

К — «извлеченный шар содержит красный цвет»,

С — «извлеченный шар содержит синий цвет»,

З — «извлеченный шар содержит зеленый цвет».

Пространство элементарных исходов Ω опишем списком равновозможных исходов

$$\{\omega_{\text{к}}, \omega_{\text{с}}, \omega_{\text{з}}, \omega_{\text{ксз}}\},$$

означающих соответственно, что извлечен красный, синий, зеленый или трехцветный шар. При этом

$$\text{К} = \{\omega_{\text{к}}, \omega_{\text{ксз}}\}, \quad \text{С} = \{\omega_{\text{с}}, \omega_{\text{ксз}}\}, \quad \text{З} = \{\omega_{\text{з}}, \omega_{\text{ксз}}\}.$$

Получаем

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

Заметим, что

$$\text{К} \cap \text{С} = \text{К} \cap \text{З} = \text{С} \cap \text{З} = \{\omega_{\text{ксз}}\},$$

поэтому

$$P\{\text{К} \cap \text{С}\} = P\{\text{К} \cap \text{З}\} = P\{\text{С} \cap \text{З}\} = \frac{1}{4},$$

откуда следует попарная независимость событий К, С и З. В то же время $\text{К} \cap \text{С} \cap \text{З} = \{\omega_{\text{ксз}}\}$ и $P\{\text{КСЗ}\} = 1/4$. Отсюда, например,

$$P\{\text{КС} \,|\, \text{З}\} = \frac{P\{\text{КСЗ}\}}{P\{\text{З}\}} = \frac{0{,}25}{0{,}5} = \frac{1}{2},$$

в то время как

$$P\{\text{КС}\} = \frac{1}{4} \neq \frac{1}{2} = P\{\text{КС} \,|\, \text{З}\}.$$

Из определения (10) следует, что

события $A_1, A_2, \ldots, A_n$ независимы в совокупности тогда и только тогда, когда для любого r, $1 \leqslant r \leqslant n$, и любых i_1, i_2, $1 \leqslant i_1, \leqslant i_2 \leqslant \ldots \leqslant i_r \leqslant n$, выполняется равенство

$$P\left\{\bigcap_{s=1}^{r} A_{i_s}\right\} = \prod_{s=1}^{r} P\{A_{i_s}\}. \tag{11}$$

Замечание. Отметим, что наличие «длинной» цепочки соотношений (11) не гарантирует справедливость «более коротких» цепочек.

Пример 4. В урне лежит 8 физически идентичных шаров — 2 красных, 2 зеленых и по одному синему, красно-синему, сине-зеленому и трехцветному. Эксперимент состоит в извлечении шара из урны.

Пространство элементарных исходов опишем следующим списком:

$$\Omega = \left\{\omega_{\text{к}}^1, \omega_{\text{к}}^2, \omega_{\text{с}}, \omega_{\text{кс}}, \omega_{\text{сз}}, \omega_{\text{ксз}}, \omega_{\text{з}}^1, \omega_{\text{з}}^2\right\}.$$

Пусть события К, С и З такие же, как и в предыдущем примере. Тогда

$$\text{К} = \{\omega_{\text{к}}^1, \omega_{\text{к}}^2, \omega_{\text{кс}}, \omega_{\text{ксз}}\} \quad \text{и} \quad P\{\text{К}\} = \frac{4}{8} = \frac{1}{2},$$

$$\text{С} = \{\omega_{\text{с}}, \omega_{\text{кс}}, \omega_{\text{сз}}, \omega_{\text{ксз}}\} \quad \text{и} \quad P\{\text{С}\} = \frac{4}{8} = \frac{1}{2},$$

$$\text{З} = \{\omega_{\text{з}}^1, \omega_{\text{з}}^2, \omega_{\text{сз}}, \omega_{\text{ксз}}\} \quad \text{и} \quad P\{\text{З}\} = \frac{4}{8} = \frac{1}{2},$$

$$\text{КСЗ} = \{\omega_{\text{ксз}}\} \quad \text{и} \quad P\{\text{КСЗ}\} = \frac{1}{8} = P\{\text{К}\}P\{\text{С}\}P\{\text{З}\}.$$

В то же время

$$\text{КЗ} = \{\omega_{\text{ксз}}\}, \ P\{\text{КЗ}\} = \frac{1}{8} \neq P\{\text{К}\}P\{\text{З}\} = \frac{1}{4}. \blacktriangleright$$

§ 5. Совмещение экспериментов

Часто возникает потребность в совместном рассмотрении событий из разных экспериментов. Формальная теория, изложенная выше, вообще говоря, не позволяет этого сделать, так как все определения и утверждения предыдущих разделов относятся к единому для рассматриваемых в них событий пространству элементарных исходов. Ниже будет предложен способ построения общего пространства элементарных исходов для конечной совокупности различных экспериментов.

Пусть $\{\Omega_1, F_1, P_1\}$ и $\{\Omega_2, F_2, P_2\}$ — два вероятностных пространства, описывающих два различных эксперимента.

Отметим, что совместное рассмотрение любых событий из этих экспериментов должно давать, во-первых, возможность рассмотрения в рамках общей модели любых событий каждого из экспериментов в отдельности, а во-вторых, в предположении отсутствия какого бы то ни было взаимодействия между экспериментами, события из разных экспериментов в рамках общей модели должны быть независимыми. Этим требованиям удовлетворяет конструкция *совмещения экспериментов*.

Рассмотрим эксперимент, состоящий в совместном осуществлении экспериментов, описываемых вероятностными пространствами $\{\Omega_1, F_1, P_1\}$ и $\{\Omega_2, F_2, P_2\}$, при условии отсутствия взаимодействия между ними. Совокупность элементарных исходов этого эксперимента опишем множеством всех возможных упорядоченных пар $(\omega^1; \omega^2)$, где $\omega^1 \in \Omega_1$, $\omega^2 \in \Omega_2$.

$$\boxed{\Omega = \left\{(\omega^1; \omega^2), \ \omega^i \in \Omega_i, \ i = 1, 2\right\},} \tag{1}$$

или

$$\Omega = \Omega_1 \times \Omega_2.$$

Случайными событиями в этом эксперименте назовем любые подмножества множества (1) вида

$$\boxed{A = (A_1, A_2) = \left\{(\omega^1; \omega^2): \ \omega^1 \in A_1 \in F_1, \ \omega^2 \in A_2 \in F_2\right\},} \tag{2}$$

т. е. любые упорядоченные пары случайных событий из рассматриваемых экспериментов, дополнив их всевозможными не более чем счетными объединениями, пересечениями и отрицаниями и т.д. (т.е. построим σ-алгебру случайных событий — наименьшую σ-алгебру, включающую в себя события вида (2)).

Вероятность на множестве случайных событий определим соотношением

$$\boxed{P\{A\} = P\{(A_1, A_2)\} = P_1\{A_1\}P_2\{A_2\}.} \tag{3}$$

Определение. Вероятностное пространство $\{\Omega, F, P\}$, задаваемое соотношениями (1), (2), (3), будем называть *совмещением экспериментов* $\{\Omega_1, F_1, P_1\}$ и $\{\Omega_2, F_2, P_2\}$.

Такой способ определения вероятностного пространства корректен в том смысле, что любому событию (не только вида (2)) приписана вероятность, удовлетворяющая всем предъявляемым к ней требованиям.

Любое событие A_1 из первого эксперимента может быть отождествлено с событием $\widehat{A}_1 = (A_1, \Omega_2)$ из совмещения экспериментов и при этом

$$P\{\widehat{A}_1\} = P_1\{A_1\}P_2\{\Omega_2\} = P_1\{A_1\}, \tag{4}$$

аналогично $\widehat{A}_2 = (\Omega_1, A_2)$ отождествляется с A_2 и

$$P\{\widehat{A}_2\} = P_1\{\Omega_1\}P_2\{A_2\} = P_2\{A_2\}. \tag{5}$$

Отметим, что определение (3), мотивированное отсутствием взаимодействия между экспериментами, постулирует независимость в рамках совмещения экспериментов любых событий из различных экспериментов. Действительно

$$P\{\widehat{A}_1\widehat{A}_2\} = P\{(A_1, \Omega_2) \cap (\Omega_1, A_2)\} = P\{(A_1, A_2)\} = P_1\{A_1\}P_2\{A_2\} = \\ = P\{\widehat{A}_1\}P\{\widehat{A}_2\}.$$

Эта схема очевидным образом распространяется на случай любого конечного числа совмещаемых экспериментов.

Рассмотрим несколько примеров, иллюстрирующих рассмотренную конструкцию.

Пример 1. Пусть $\{\Omega_1, F_1, P_1\}$ — вероятностное пространство, описывающее эксперимент, связанный с бросанием симметричного игрального кубика:

$$\Omega_1 = \{\omega_1^1, \omega_2^1, \omega_3^1, \omega_4^1, \omega_5^1, \omega_6^1\}, \quad p(\omega_j^1) = \frac{1}{6}, \quad j = 1, 2, \ldots, 6;$$

$$F_1 = \{\varnothing, \Omega_1, A_1\}, \quad A_1 = \{\omega_i^1\}, \quad P\{A\} = \sum P_1\{\omega_i^1\}.$$

Вероятностное пространство $\{\Omega_2, F_2, P_2\}$ описывает аналогичный предыдущему и независимый от него эксперимент. Тогда $\Omega = \Omega_1 \times \Omega_2$ — эксперимент, описывающий двукратное бросание игрального кубика (или, что то же, одновременное бросание двух физически идентичных симметричных игральных кубиков) — описывается так (рис. 35):

$$\Omega = \{(\omega_i^1, \omega_j^2), i = 1, 2, \ldots, 6; j = 1, 2, \ldots, 6\},$$

$$P\{(\omega_i^1, \omega_j^2)\} = P_1\{\omega_i^1\}P_2\{\omega_j^2\} = \frac{1}{36},$$

$$A = \{(\omega_i^1, \omega_j^2)\colon\ \omega_i^1 \in A_1;\ \omega_j^2 \in A_2\}, \quad P\{A\} = \sum P\{(\omega_i^1, \omega_j^2)\}.$$

Например, событие «на первом кубике выпала двойка» описывается всеми точками Ω вида

$$(\omega_2^1, \Omega_2) = \{(\omega_2^1, \omega_1^2), (\omega_2^1, \omega_2^2), (\omega_2^1, \omega_3^2), \ldots, (\omega_2^1, \omega_6^2)\},$$

событие «на обоих кубиках выпали одинаковые числа» — точками Ω вида

$$\{(\omega_i^1, \omega_i^2),\ \omega_i^1 = \omega_i^2,\ i = 1, 2, \ldots, 6\}$$

и т. д.

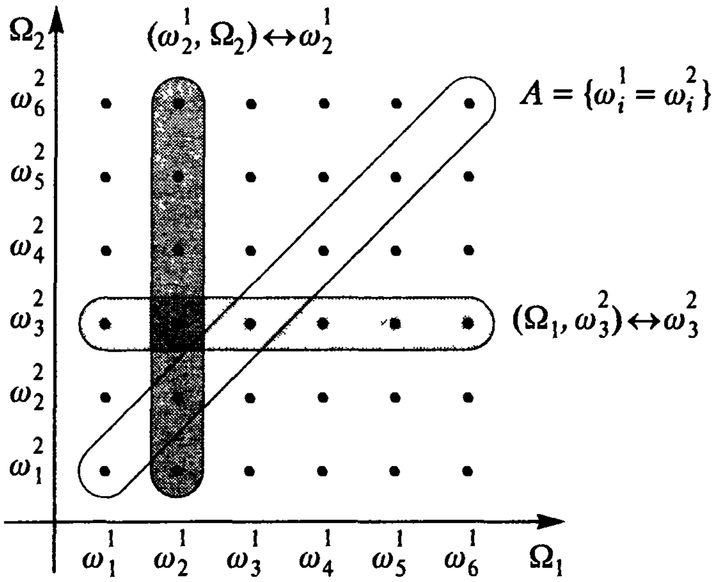


Рис. 35

Пример 2. Биномиальный эксперимент. В эксперименте, описанном вероятностным пространством $\{\Omega, F, P\}$, рассмотрим некоторое событие Y, которое в дальнейшем будем называть «успех», и для которого $P\{Y\} = p > 0$. Событие $\overline{Y}$ будет расцениваться как «неуспех»: $P\{\overline{Y}\} + P\{Y\} = 1$.

Свяжем с описанным выше экспериментом дискретный эксперимент с двумя элементарными исходами

$$\Omega^B = \{Y, \overline{Y}\}, \quad F = \{\varnothing, \Omega, Y, \overline{Y}\}, \quad P\{Y\} = p; \quad P\{\overline{Y}\} = q = 1 - p,$$

который будем называть *экспериментом Бернулли* с *параметром* p.

Последовательность n независимых испытаний, в каждом из которых событие Y происходит с неизменной вероятностью p, является n-кратным совмещением эксперимента Ω^B.

Элементарными исходами, составляющими пространство элементарных событий

$$\Omega = \underbrace{\Omega^B \times \ldots \times \Omega^B}_{n} = (\Omega^B)^n,$$

будут всевозможные упорядоченные наборы длины n, составленные из символов Y или $\overline{Y}$. Легко подсчитать, что общее число таких наборов (элементарных исходов) $|\Omega| = 2^n$. Каждому элементарному исходу в соответствии с правилом (3) припишем вероятность $p^k q^{n-k}$, если только этот исход содержит ровно k символов Y (описывает такой результат эксперимента, когда в серии из n независимых испытаний наблюдалось k успехов и, соответственно, $n - k$ неудач.) Введем в рассмотрение следующие события:

$A_{=k} = \{$«ровно k успехов в серии из n испытаний»$\}$;
$A_{\geqslant k} = \{$«не менее k успехов в серии из n испытаний»$\}$;
$A_{\leqslant k} = \{$«не более k успехов в серии из n испытаний»$\}$;
${}_{r\leqslant}A_{\leqslant k} = \{$«количество успехов в серии из n испытаний заключена в пределах от r до k»$\}$.

Сразу заметим, что поскольку $A_{=k} \cap A_{=r} = \varnothing$, если только $k \neq r$, то

$$A_{\leqslant k} = A_{=0} + A_{=1} + \ldots + A_{=k}$$

и

$$P\{A_{\leqslant k}\} = \sum_{j=0}^{k} P\{A_{=j}\}.$$

Аналогично

$$P\{A_{\geqslant k}\} = \sum_{j=k}^{n} P\{A_j\}, \quad P\{{}_{r\leqslant}A_{\leqslant k}\} = \sum_{j=r}^{k} P\{A_{=j}\}.$$

Для событий $A_{=k}$ имеем

$$A_{=k} = \{\overline{Y}Y\overline{Y} \ldots Y, Y\overline{Y}\overline{Y} \ldots Y, \ldots, YY \ldots \overline{Y}Y\overline{Y}\},$$

где каждый элементарный исход содержит ровно k символов Y. Поскольку всего таких исходов C_n^k, то

$$P\{A_{=k}\} = C_n^k p^k (1-p)^{n-k}. \qquad (6)$$

Эксперимент, являющийся n-кратным совмещением бернуллиевского эксперимента с параметром p, будем называть *биномиальным экспериментом с параметрами* (n, p).

Схема биномиального эксперимента находит широкое применение в различных прикладных задачах. Рассмотрим два конкретных примера.

Пример 1. Телевизионный завод выпускает в среднем 2 % некачественных кинескопов. Сборочный цех получил 10 кинескопов. Какова вероятность того, что среди полученных не более 2 некачественных?

◀ Событие Y — «соответствующий кинескоп качественный», $\overline{Y}$ — «соответствующий кинескоп некачественный». Количество испытаний $n = 10$. В качестве вероятности успеха условие задачи дает основание принять величину 0,98. Следует найти $P\{A_{\geqslant 8}\}$. Имеем

$$P\{A_{\geqslant 8}\} = \sum_{j=8}^{10} P\{A_{=j}\} = C_{10}^{8}(0{,}98)^3 \cdot (0{,}02)^2 + +C_{10}^{9}(0{,}98)^9 \cdot (0{,}02) + C_{10}^{10}(0{,}98)^{10} \approx 0{,}999. ▶$$

Пример 2. Урновая схема с возвращением. В урне лежит $N = m + l$ геометрически идентичных шаров, соответственно m черного и l белого цвета. Эксперимент состоит в n-кратном извлечении шара из урны с последующим возвращением извлеченного шара (после фиксации его цвета). Какова вероятность того, что среди извлеченных шаров число шаров белого цвета окажется не менее r?

◀ Пусть Y — событие, состоящее в том, что извлеченный шар белый, $\overline{Y}$ — черный. Тогда $p\{Y\} = \frac{l}{m+l}$; количество испытаний — n. Следует найти $P\{A_{\geqslant r}\}$:

$$P\{A_{\geqslant r}\} = \sum_{j=r}^{n} P\{A_{=j}\} = \sum_{j=r}^{n} C_n^j \left(\frac{l}{m+l}\right)^j \cdot \left(\frac{m}{m+l}\right)^{n-j}. ▶$$

Существенным обстоятельством, позволившим применить конструкцию совмещения экспериментов (в рассмотренном примере — совмещения однократных извлечений шара из урны) является *неизменность* вероятности извлечения белого шара — следствие процедуры выбора с возвращением.

Пример 3. Полиномиальный эксперимент. Пусть $H_1, H_2, \dots H_k$ — полная система несовместных событий, $P\{H_j\} = p_j$, $\sum\limits_{j=1}^{k} p_j = 1$. Свяжем с рассматриваемым экспериментом дискретный эксперимент

$$\Omega^B = \{H_1, H_2, \dots H_k\}, \quad P\{H_j\} = p_j$$

и рассмотрим последовательность n независимых испытаний, в каждом из которых может происходить любое из событий H_j с вероятностью p_j — n-кратное совмещение эксперимента Ω^B. Его элементарные исходы — упорядоченные наборы длины n, составленные из символов $\{H_j\}_{j=1}^k$ — на любом месте набора может стоять любой из символов H_j. Заметим, что $|\Omega| = |(\Omega^B)^n| = k^n$. В соответствии с правилом (2) каждому исходу, составленному из r_1 символов H_1, r_2 символов $H_2, \dots, r_k$ символов H_k, $r_1 + r_2 + \dots + r_k = n$, припишем вероятность

$$p = p_1^{r_1} \cdot p_2^{r_2} \dots p_k^{r_k}.$$

Введем в рассмотрение событие

$$A_{=r_1 r_2 \dots r_k} = \{\text{«ровно } r_1 \text{ осуществлений } H_1, \dots, \text{ровно } r_k \text{ осуществлений } H_k\text{»}\},$$

где $r_1 + r_2 + \dots + r_k = r$, $0 \leqslant r_j$. Все исходы, составляющие событие $A_{=r_1 r_2 \dots r_k}$, имеют одну и ту же вероятность $p_1^{r_1}, \dots, p_k^{r_k}$, и всего их $\frac{n!}{r_1! r_2! \dots r_k!}$ (см. Приложение). Тем самым,

$$P\{A_{=r_1 r_2 \dots r_k}\} = \frac{n!}{r_1! r_2! \dots r_k!} p_1^{r_1} p_2^{r_2} \dots p_k^{r_k}. \tag{7}$$

Этот эксперимент будем называть *полиномиальным экспериментом с параметрами* $(n;\ p_1, p_2, \dots, p_k)$.

Рассмотрим конкретный пример.

Пример. Двое друзей договорились встретиться в определенном месте в 10 часов утра. Известно, что каждый из них приходит к месту встречи случайным образом в любой момент времени от 9.45 до 10.15 независимо от момента прихода к месту встречи другого. Найти вероятность того, что одному из них придется ожидать другого не более 10 минут.

Пусть Ω_1 — пространство элементарных исходов, описывающее моменты прихода к месту встречи одного из друзей, Ω_2 — второго:

$$\Omega_1 = \left\{t\colon\ 9\frac{3}{4} \leqslant t \leqslant 10\frac{1}{4}\right\},$$

$$\Omega_2 = \left\{\tau\colon\ 9\frac{3}{4} \leqslant \tau \leqslant 10\frac{1}{4}\right\}.$$

В каждом из указанных экспериментов естественными наблюдаемыми событиями будут события, состоящие в приходе к месту встречи в некотором временнóм промежутке $(\alpha, \beta) \subset \left\langle 9\frac{3}{4}, 10\frac{1}{4}\right\rangle$. Единицей измерения времени выбран час.

Вероятность события

$$A_1 = \{t\colon\ t \in \langle \alpha, \beta \rangle\}$$

в силу случайности момента прихода к месту встречи первого из друзей резонно считать пропорциональной длине промежутка $\langle \alpha, \beta \rangle$:

$$P\{A_1\} = k(\beta - \alpha).$$

Коэффициент пропорциональности k определим из условия нормировки

$$P\{\Omega_1\} = k\left(10\frac{1}{4} - 9\frac{3}{4}\right) = \frac{1}{2}k = 1, \quad \text{откуда } k = 2.$$

Для Ω_2 поступаем аналогично. Заметим, что случайными событиями в рассматриваемых экспериментах будут элементы борелевской σ-алгебры на промежутке $\left\langle 9\frac{3}{4}, 10\frac{1}{4}\right\rangle$.

Совмещением экспериментов Ω_1 и Ω_2 будет эксперимент, пространство элементарных исходов которого можно описать точками квадрата на плоскости (t, τ):

$$\Omega = \Omega_1 \times \Omega_2 = \left\{(t, \tau)\colon\ 9\frac{3}{4} \leqslant t, \tau \leqslant 10\frac{1}{4}\right\}.$$

В соответстаии с рассмотренной выше конструкцией вероятности любого случайного события из совмещения

$$(A_1, A_2) = \{(t, \tau) : \alpha_1 \leqslant t \leqslant \beta_1;\ \alpha_2 \leqslant \tau \leqslant \beta_2\}$$

будут определяться соотношением

$$P\{(A_1, A_2)\} = 2(\beta_1 - \alpha_1) \cdot 2(\beta_2 - \alpha_2) = 4(\beta_1 - \alpha_1)(\beta_2 - \alpha_2),$$

(вероятность попадания в прямоугольник пропорциональна его площади). Отсюда, для любого события, описываемого квадрируемым подмножеством множества точек квадрата Ω, вероятность пропорциональна площади

$$P\{A\} = \frac{S(A)}{S(\Omega)} = 4S(A). \tag{8}$$

Интересующее нас событие — «время ожидания одним из друзей другого не превышает 10 минут» — описывается множеством тех точек Ω, для которых $|t - \tau| \leqslant 1/6$ (напомним, что время измеряется в часах: 10 минут = 1/6 часа)

$$A = \{(t, \tau)\colon\ |t - \tau| \leqslant \frac{1}{6};\ (t, \tau) \in \Omega\}.$$

Поэтому, в соответствии с (8) имеем

$$P\{A\} = 4S(A) = 4 \cdot \frac{5}{36} = \frac{5}{9}.$$

Собственно вычисление площадей проиллюстрировано на рис. 36.

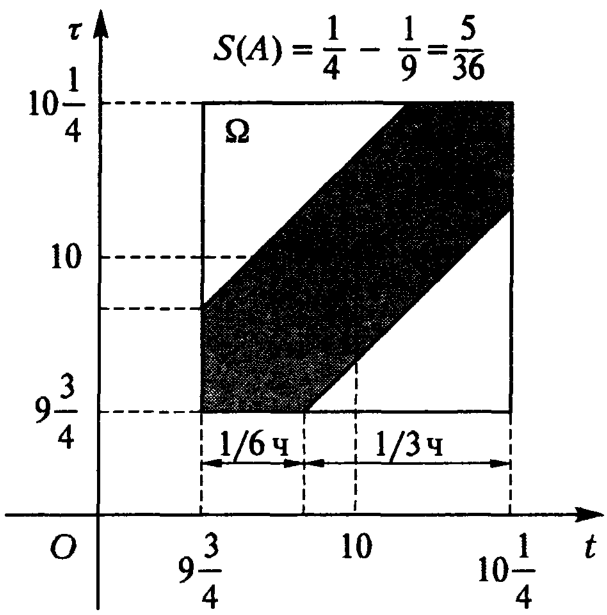


Рис. 36

§ 6. Счетные совмещения

В предыдущем параграфе построена конструкция, позволяющая совместно рассматривать события из конечного числа разных экспериментов. Однако во многих случаях возникает потребность рассматривать события из бесконечного множества различных экспериментов. Например, в теории случайных процессов изучается связь между случайными событиями, происходящими в разные моменты времени, и предполагается, что для любого момента времени задан свой эксперимент. В этой задаче множество экспериментов континуально.

Здесь мы рассмотрим случай счетного множества экспериментов. Пусть $\{\Omega_i, F_i, P_i\}$, $i = 1, 2, \dots$, — счетное множество вероятностных пространств. По аналогии с конечным совмещением экспериментов совокупность элементарных исходов счетного совмещения естественно описать как множество всех последовательностей $(\omega^1, \omega^2, \dots, \omega^n, \dots)$, где $\omega^i \in \Omega_i$, $i = 1, 2, \dots$,

$$\boxed{\Omega = \left\{(\omega^1, \omega^2, \dots, \omega^n),\ \omega^i \in \Omega_i,\ i = 1, 2, \dots\right\}} \tag{1}$$

или

$$\boxed{\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n \times \dots = \prod_{i=1}^{\infty} \Omega_i.}$$

Однако если мы захотим продолжить аналогию дальше и называть случайными событиями множества вида

$$A = (A_1, A_2, \dots) = \left\{(\omega^1, \omega^2, \dots),\ \omega^i \in A_i \in F_i\right\},$$

а вероятность определять по аналогии как произведение, то почти всегда

$$P\{A\} = P_1\{A_1\}P_2\{A_2\} \dots = 0$$

(здесь рассматривается бесконечное произведение). Поэтому (а также потому, что рассмотренная ниже конструкция применяется во многих задачах) рассмотрим следующие построения.

Определение. Множество $A \subset \Omega$ назовем *цилиндрическим*, если существует конечное число номеров $i_1, i_2, \dots, i_n$ и событий $A_{i_1} \in F_{i_1}, A_{i_2} \in F_{i_2}, \dots, A_{i_n} \in F_{i_n}$ таких, что

$$A = \left\{(\omega^1, \omega^2, \dots):\ \omega^{i_1} \in A_{i_1}, \dots, \omega^{i_n} \in A_{i_n}\right\},$$

т. е. A состоит из таких последовательностей $(\omega^1, \omega^2, \dots) \in \Omega$, что только на местах $i_1, i_2, \dots, i_k$ элементы ω^{i_k} содержатся в соответствующем A_{i_k}, а остальные ω^i произвольны.

Как легко видеть, вероятностные конечные объединения цилиндрических множеств образуют алгебру событий. Определим вероятность на алгебраическом множестве соотношением

$$\boxed{P\{A\} = P_{i_1}\{A_{i_1}\}P_{i_2}\{A_{i_2}\} \dots P_{i_n}\{A_{i_n}\},} \tag{2}$$

а на конечных объединениях непересекающихся цилиндрических множеств — как соответствующую сумму. Можно доказать (хотя это довольно сложно), что определенная в (2) вероятность обладает свойством σ-аддитивности, и, следовательно, ее можно продолжить на наименьшую σ-алгебру, содержащую цилиндрические множества. Вероятностное пространство $\{\Omega, F, P\}$, задаваемое соотношениями (1) и (2), где F — наименьшая σ-алгебра, содержащая цилиндрические множества, будем называть совмещением экспериментов $\{\Omega_i, F_i, P_i\}$, $i = 1, 2, \dots$.

Случайному событию из i-го эксперимента $A_i \in F_i$ так же, как и раньше, сопоставляется случайное событие $\widehat{A}_i \in F$,

$$\widehat{A}_i = \left\{(\omega^1, \omega^2, \dots),\ \omega^i \in A_i\right\}.$$

Из (2) следует, что вероятности A_i и $\widehat{A}_i$ равны: $P\{A_i\} = P\{\widehat{A}_i\}$. В счетном совмещении экспериментов события из разных экспериментов независимы в совокупности, что сразу следует из (2). Итак, мы видим, что наша конструкция, как и в конечном случае, обладает всеми нужными свойствами: в новом эксперименте $\left(\Omega = \prod_{i=1}^{\infty} \Omega_i\right)$

вероятности всех событий не изменились и события из разных экспериментов остались независимыми.

Пример 1. В примере 2 предыдущего параграфа рассмотрен эксперимент Бернулли с параметром p — $\Omega^B = \{1, 0\}$ (здесь мы обозначили событие Y через 1, а событие $\overline{Y}$ через 0). На основе Ω^B был определен биномиальный эксперимент с параметрами (n, p) как n-кратное совмещение бернуллиевского эксперимента. Рассмотрим теперь счетное совмещение бернуллиевского эксперимента с параметром p.

В соответствии с (1) элементами

$$\Omega = \Omega^B \times \Omega^B \times \ldots = \Omega^{B*\infty}$$

являются всевозможные последовательности из нулей и единиц. Поставив в соответствие такой последовательности действительное число, для которого онв является разложением в двоичную дробь, получим $\Omega = [0, 1]$. Итак, мы видим, что Ω континуально, даже несмотря на то, что в Ω^B всего два элемента. Далее, в Ω обязательно существует событие, не имеющее вероятности, так как в отличие от биномиального эксперимента $\Omega^{B*\infty}$ дискретным экспериментом не является. Аналогия Ω и $[0, 1]$ идет и дальше: оказывается, что при указанном отождествлении Ω и $[0, 1]$ σ-алгебра F, содержащая цилиндрические множества в Ω, соответствует σ-алгебре борелевских множеств $[0, 1]$. Так происходит потому, что цилиндрические множества соответствуют отрезкам с концами в двоично-рациональных точках, т. е. σ-алгебры, содержащие цилиндрические множества в Ω и двоично-рациональные отрезки в $[0, 1]$, совпадают. Очевидно, что σ-алгебра на $[0, 1]$, содержащая все двоично-рациональные отрезки, совпадает с σ-алгеброй борелевских множеств $[0, 1]$. Итак, совпадают не только множества $\Omega^{B*\infty}$ и $[0, 1]$, но и σ-алгебры случайных множеств.

Обратимся к вероятности в $\Omega^{B*\infty}$ и в $[0, 1]$ при $p = 1/2$ (в остальных случаях этот вопрос более сложен). Оказывается, что в этом случае вероятность соответствует длине; точнее, если событию $A \subset \Omega^{B*\infty}$ соответствует промежуток $\langle\alpha; \beta\rangle \subset [0, 1]$, то $P\{A\} = \beta - \alpha$. Итак, при $p = 1/2$ счетное совмещение $\Omega^{B*\infty}$ полностью совпадает с отрезком $[0, 1]$ — и как множество, и в смысле совпадения случайных событий, а вероятность в $\Omega^{B*\infty}$ совпадает с длиной в $[0, 1]$.

Пример 2. Бросание монеты до первого появления герба. Пусть вероятность выпадения герба при одном бросании равна p. Эксперимент состоит в том, что мы подбрасываем монету до тех пор, пока не появится герб. Интуитивно ясно, как построить пространство элементарных событий и как задать на нем вероятность: если обозначить через 1 выпадение герба, а через 0 — решки, то все исходы эксперимента перечисляются так —

$$1, 01, 001, \ldots, \underbrace{00\ldots0}_{n}1, \ldots, \omega_0.$$

(ω_0 — герб не выпал никогда). Случайными событиями являются все подмножества, и

$$P\{\underbrace{00\ldots0}_{n}1\} = pq^n, \quad P\{\omega_0\} = 0,$$

что соответствует независимости бросаний. Однако с помощью нашей конструкции счетного совмещения событий мы можем изучить данную задачу более строго. Рассмотрим счетное совмещение бернуллиевских экспериментов $\Omega^{B*\infty}$ с параметром p. Событие A_n, состоящее в том, что n раз выпала решка, а в $(n+1)$-й раз — герб, соответствует цилиндрическому множеству

$$A_n = \{(\omega^1, \omega^2, \ldots)\colon\ \omega^1 = 0,\ \omega^2 = 0,\ \ldots,\ \omega^n = 0,\ \omega^{n+1} = 1\}.$$

Следовательно, $p\{A_n\} = pq^n$. Теперь мы также можем доказать, что $P\{\omega_0\} = 0$. В самом деле,

$$\omega_0 = \{(\omega^1, \omega^2, \ldots)\colon\ \omega^i = 0, i = 1, 2, \ldots\}.$$

Отсюда видно, что ω_0 является счетным пересечением убывающих цилиндрических множеств. Значит, по свойству непрерывности вероятности, $p\{\omega\} = 0$. Итак, за пространство элементарных исходов при бросании монеты до первого появления герба возьмем совокупность цилиндрических множеств A_n, $n = 0, 1, 2, \ldots$, и ω_0 из σ-алгебры F в $\Omega^{B*\infty}$ вместе с предписанными им в $\Omega^{B*\infty}$ вероятностями. Мы получаем дискретное распределение вероятностей, так как

$$\sum_{n=0}^{\infty} pq^n = 1.$$

Вероятность произвольного события в нашем эксперименте находится, как обычно в дискретном распределении суммированием вероятностей элементарных исходов, благоприятствующих данному событию. Нвйдем для примера вероятность того, что для первого появления герба надо сделать не более трех подбрасываний. Если через A обозначить интересующее нас событие, то

$$A = \{A_0, A_1, A_2\} \text{ и}$$

$$P\{A\} = P\{A_0\} + P\{A_1\} + P\{A_2\} = p + pq + pq^2 = p(1 + q + q^2) = 1 - q^3.$$

Глава XXXIX

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ И ЗАКОНЫ РАСПРЕДЕЛЕНИЯ

Центральным понятием теории вероятностей является понятие *случайной величины*, формализующее в рамках принятых выше соглашений возможность численной фиксации результатов эксперимента. Как уже отмечалось, пространство элементарных событий, описывающее эксперимент, отождествляется с совокупностью всех возможных исходов, случайная же величина представляется как совокупность чисел, которые экспериментатор с тем или иным исходом связывает. Каждому исходу эксперимента ставится в соответствие число (или несколько чисел), являющееся результатом измерения одного показателя (или нескольких), описывающего течение эксперимента. При этом однократное проведение эксперимента (испытания) дает возможность получить одно измерение каждого из показателей. Отметим, что значения измеряемых показателей являются «случайными» в том смысле, что зависят от исхода, реализовавшегося в данном испытании. Есть смысл наряду с множеством возможных значений «случайной» величины характеризовать измеряемый параметр вероятностью, с которой может быть получено то или иное значение в ряду всех возможных. Любая такая характеристика носит название *закона распределения вероятностей на множестве значений случайной величины*. Обсуждению этих понятий и изучению свойств связанных с ними конструкций посвящена настоящая глава.

§ 1. Случайная величина

Пусть $\{\Omega, F, P\}$ — вероятностное пространство некоторого эксперимента, Ω — список элементарных исходов, F — список случайных событий, P — правило вычисления вероятностей для событий из списка F. Пусть каждому элементарному исходу $\omega \in \Omega$ поставлено в соответствие действительное число $\xi(\omega)$.

Определение. Будем называть $\xi(\omega)$ *случайной величиной*, если для любого полуинтервала $[a, b) \subset \mathbb{R}$ множество тех ω, для которых $\xi(\omega) \in [a, b)$, является случайным событием.

Пример 1. Игрок бросает монету — при выпадении герба он выигрывает 1 рубль, решки — проигрывает 1 рубль. Случайная величина ξ — выигрыш игрока будет принимать значения $+1$ или -1 в зависимости от того, чем закончился эксперимент — гербом или решкой.

Пример 2. Эксперимент — одновременное бросание двух игральных кубиков, случайная величина — сумма выпавших очков, может принимать все целые значения от 2 до 12, в зависимости от выпавшей комбинации.

Пример 3. Эксперимент — n-кратное повторение эксперимента с бросанием монеты, случайная величина — количество выпавших гербов — может принимать все целые значения от 0 до n.

Пример 4. Эксперимент — извлечение шара из урны, содержащей равное количество белых и черных шаров, с возвращением шара в урну после каждого извлечения, случайная величина — количество извлечений до первого появления белого шара. Эта случайная величина может принимать все целые положительные значения: $1, 2, 3, \ldots, n, \ldots$.

Пример 5. Эксперимент — случайный выбор точки из отрезка $[0, 1]$. Случайная величина — координата точки. Эта случайная величина может принимать любые значения от 0 до 1.

Пример 6. Эксперимент — наблюдение за временем безотказной работы некоторого устройства: от момента включения до первого выхода из строя. Случайная величина — время безотказной работы — может принимать все действительные значения от 0 до $+\infty$.

Подчеркнем, что если обозначить через $\xi^{-1}([a, b))$ полный прообраз полуинтервала $[a, b)$ при отображении ξ, то $\xi(\omega)$ будет случайной величиной в том случае, когда $\xi^{-1}([a, b)) \in \boldsymbol{F}$ и, следовательно, можно говорить о вероятности этого случайного события. Для любого полуинтервала $[a, b)$ числовой прямой в этом случае будет определена функция

$$\boxed{\Phi_\xi([a, b)) = P\{\xi^{-1}([a, b))\},}$$

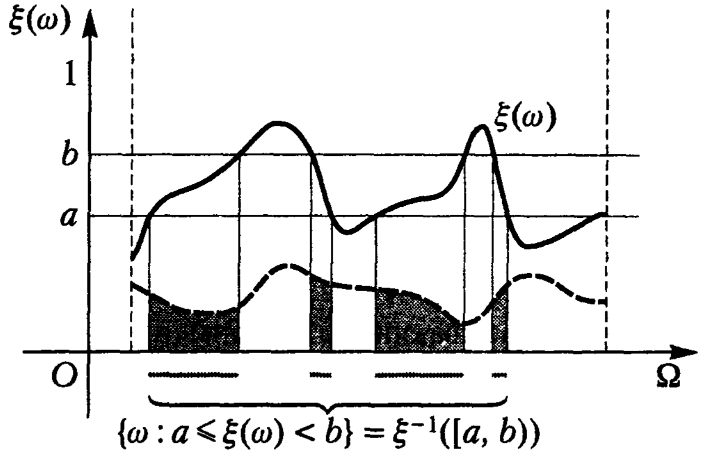


Рис. 1

которую в дальнейшем будем называть *вероятностной функцией случайной величины* ξ. Рис. 1 иллюстрирует данные выше определения; жирная пунктирная кривая изображает плотность вероятности, речь о которой пойдет ниже, в п. 2.1; горизонтальная разрывная линия условно показывает прообраз $\xi^{-1}([a, b))$; вероятность $P\{\xi^{-1}([a, b))\}$ равна площади закрашенной серым фигуры.

Во всех указанных примерах легко установить, что события $\xi^{-1}([a, b))$ являются случайными событиями, т. е. элементами списка $\boldsymbol{F}$.

Заметим, что если $\xi(\omega)$ — случайная величина, то $\xi^{-1}\{(-\infty, x)\}$, $\xi^{-1}\{(x, +\infty)\}$, $\xi^{-1}\{(a, b)\}$, $\xi^{-1}\{[a, b]\}$, $\xi^{-1}\{(a, b]\}$, $\xi^{-1}(\alpha)$ являются случайными событиями и, как следствие, прообраз любого множества на числовой прямой, полученного из интервалов не более чем счетным числом операций объединения, пересечения и дополнения (борелевское множество), будет случайным событием. Доказательство немедленно следует из очевидных соотношений

$$\{\omega\colon\ \xi(\omega) < x\} = \{\omega\colon\ -\infty < \xi(\omega) < x\},$$
$$\{\omega\colon\ \xi(\omega) \geqslant x\} = \{\omega\colon\ -\infty < \xi(\omega) < x\},$$
$$\{\omega\colon\ \xi(\omega) \in [a, b]\} = \bigcap_{k=1}^{\infty} \left\{\omega\colon\ \xi(\omega) \in \left[a, b + \frac{1}{k}\right)\right\},$$
$$\{\omega\colon\ \xi(\omega) = \alpha\} = \bigcap_{k=1}^{\infty} \left\{\omega\colon\ \xi(\omega) \in \left[\alpha, \alpha + \frac{1}{k}\right)\right\}$$

и т. д.

Это позволяет говорить о вероятности попадания случайной величины в полупрямую $(-\infty, x)$, интервал (a, b), отрезок $[a, b]$, точку α и т. п. Вообще, если M — борелевское множество на прямой, то осмысленным является высказывание: «вероятность попадания случайной величины ξ в множество M равна p», так как множество тех ω, для которых $\xi(\omega) \in M$, является случайным событием, и, следовательно, можно говорить о вероятности этого случайного события

$$P\{\xi(\omega) \in M\} = P\{\omega\colon\ \xi(\omega) \in M\}. \tag{1}$$

Законом распределения вероятностей на множестве значений случайной величины ξ (коротко — *законом распределения*) будем называть любое правило, позволяющее вычислять вероятности (1).

Закон распределения может быть задан в произвольной форме, однако наиболее употребительными являются следующие:

— функция распределения вероятностей;

— плотность распределения вероятностей;

— ряд распределения вероятностей.

1.1. Функция распределения

Функцией распределения вероятностей на множестве значений случайной величины ξ (коротко: *функцией распределения случайной величины* ξ) будем называть функцию $F_\xi(x)$ такую, что

$$\forall\, x \in \mathbb{R}\colon\ F_\xi(x) = P\{\xi < x\}. \tag{2}$$

Отметим, что определение (2) корректно и $F_\xi(x)$ действительно определена во всех точках числовой прямой, так как множество $\{\omega\colon\ \xi(\omega) < x\}$ является случайным событием и ему можно приписать вероятность.

Свойства функции распределения

1. Функция $F_\xi(x)$ ограничена на всей числовой прямой

$$\boxed{0 \leqslant F_\xi(x) \leqslant 1.} \tag{3}$$

2.

$$\boxed{P\{a \leqslant \xi < b\} = F_\xi(b) - F_\xi(a).}$$

◄ Если $a < b$, то события $\{\omega\colon\ \xi(\omega) < a\}$ и $\{\omega\colon\ a \leqslant \xi(\omega) < b\}$ несовместны и

$$\{\omega\colon\ \xi(\omega) < a\} + \{\omega\colon\ a \leqslant \xi(\omega) < b\} = \{\omega\colon\ \xi(\omega) < b\}.$$

Поэтому

$$P\{\xi < a\} + P\{a \leqslant \xi < b\} = P\{\xi < b\}.$$

С учетом определения (2) последнее равенство можно переписать так

$$F_\xi(a) + P\{a \leqslant \xi < b\} = F_\xi(b), \tag{4}$$

или

$$P\{a \leqslant \xi < b\} = F_\xi(b) - F_\xi(a). \ ► \tag{5}$$

3. Функция $F_\xi(x)$ монотонно не убывает на всей числовой прямой

$$\boxed{\forall\, a < b \quad F_\xi(a) \leqslant F_\xi(b).} \tag{6}$$

◄ С учетом неотрицательности вероятности $P\{a \leqslant \xi < b\}$ следует из соотношения (4). ►

4.

$$\boxed{\lim_{x\to-\infty} F_\xi(x) = 0, \quad \lim_{x\to+\infty} F_\xi(x) = 1.} \tag{7}$$

◀ В силу монотонности и ограниченности функции $F_\xi(x)$ на $\mathbb{R}$ у нее существуют пределы при $x \to \pm\infty$. Пусть

$$\lim_{x\to-\infty} F_\xi(x) = m_- > 0, \qquad \lim_{x\to+\infty} F_\xi(x) = m^+ < 1.$$

Заметим, что

$$P\{\xi \in \mathbb{R}\} = P\{\omega\colon\ -\infty < \xi < \infty\} = P\{\Omega\} = 1.$$

В силу непрерывности вероятности

$$1 = P\{-\infty < \xi < +\infty\} = \lim_{\substack{b\to+\infty\\ a\to-\infty}} P\{a \leqslant \xi < b\} =$$

$$= \lim_{b\to+\infty} F_\xi(b) - \lim_{a\to-\infty} F_\xi(a) = m^+ - m_-,$$

что возможно лишь если $m^+ = 1$ и $m_- = 0$. ▶

5. Функция $F_\xi(x)$ имеет не более чем счетное число точек разрыва, при этом в каждой точке существуют односторонние пределы

$$\boxed{F_\xi(x_0+) = \lim_{x\to x_0+} F_\xi(x), \qquad F_\xi(x_0-) = \lim_{x\to x_0-} F_\xi(x).}$$

Звмечвние. Функция распределения непрерывна слева в каждой точке числовой прямой

$$\lim_{x\to x_0-} F_\xi(x) = F_\xi(x_0).$$

◀ Как и выше, в силу монотонности это утверждение легко получить из непрерывности вероятности

$$F_\xi(x_0) = P\{\xi < x_0\} = P\left\{\bigcup_{n=1}^{\infty}\left(\xi < x_0 - \frac{1}{n}\right)\right\} =$$

$$= \lim_{n\to\infty} P\left\{\xi < x_0 - \frac{1}{n}\right\} = \lim_{x\to x_0-} P\{\xi < x\}. \ ▶$$

6. Если x_0 — точка непрерывности функции $F_\xi(x)$, то

$$\boxed{P\{\xi = x_0\} = 0,}$$

если же x_0 — точка разрыва, то

$$\boxed{P\{\xi = x_0\} = F_\xi(x_0+) - F_\xi(x_0)}$$

(рис. 2).

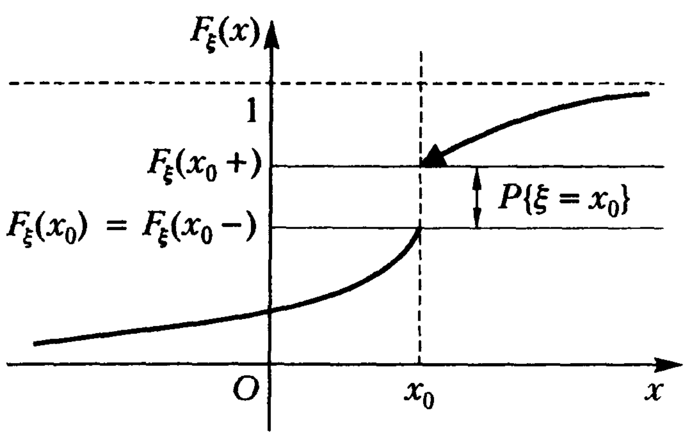


Рис. 2

Для дальнейшего важным является то обстоятельство, что функция распределения есть *универсальная* форма задания закона распределения — у любой случайной величины существует функция распределения. Не менее важно и то, что функция распределения *не дает исчерпывающей информации* о случайной величине — разные случайные величины могут обладать одинаковыми законами (в частности, функциями) распределения.

Действительно, пусть ξ — случайная величина, описывающая выигрыш при игре в орлянку (см. пример 1). Легко установить, что

$$F_\xi(x) = \begin{cases} 0, & x \leqslant -1, \\ \dfrac{1}{2}, & -1 < x \leqslant 1, \\ 1, & x > 1. \end{cases}$$

Пусть η — случайная величина, определенная в эксперименте со случайным выбором точки из отрезка $[0, 1]$ так, что

$$\eta(\omega) = -1, \quad \omega \in \left[0, \frac{1}{2}\right), \qquad \eta(\omega) = 1, \quad \omega \in \left[\frac{1}{2}, 1\right].$$

Функция распределения случайной величины η тождественна функции $F_\xi(x)$:

$$F_\eta(x) = \begin{cases} 0, & x \leqslant -1, \\ \dfrac{1}{2}, & -1 < x \leqslant 1, \\ 1, & x > 1. \end{cases}$$

У этой случайной величины ζ, определенной в таком же эксперименте, что и η, следующим образом

$$\zeta(\omega) = -1, \quad \omega \in \left[0, \frac{1}{4}\right) \cup \left[\frac{3}{4}, 1\right] \quad \text{и} \quad \zeta(\omega) = 1, \quad \omega \in \left[\frac{1}{4}, \frac{3}{4}\right),$$

та же функция распределения:

$$F_\zeta(x) = \begin{cases} 0, & x \leqslant -1, \\ \dfrac{1}{2}, & -1 < x \leqslant 1, \\ 1, & x > 1. \end{cases}$$

Рассмотренные примеры показывают, что каждая функция распределения описывает, вообще говоря, много различных случайных величин. Правда, эти случайные величины, если так можно выразиться, «не очень разные» — в реальном эксперименте у нас нет никакой возможности различить случайные величины ξ, η и ζ, потому что реальный эксперимент, как правило, дает нам возможность произвести измерения *значений* случайной величины, но не дает возможности установить, как эти значения связаны с элементарными *исходами*, которые в подавляющем большинстве реальных экспериментов остаются «вещью в себе». Поэтому на практике экспериментатор имеет дело с *функцией распределения вероятностей на множестве значений случайной величины*, а не с функцией от элементарных исходов на пространстве элементарных исходов.

Интересно поэтому выяснить, какие функции могут быть функциями распределения и как, зная функцию распределения, указать по крайней мере одну из случайных величин, описываемых этой функцией. Оказывается, что любая монотонно-неубывающая на $\mathbb{R}$ функция $F(x)$, непрерывная слева в каждой точке и такая, что

$$\lim_{x\to-\infty} F(x) = 0, \qquad \lim_{x\to+\infty} F(x) = 1,$$

может служить функцией распределения. Точнее, имеет место

Теорема. *Пусть $F(x)$ — функция, определенная на всей числовой прямой и обладающая указанными выше свойствами. Тогда существует вероятностное пространство $\{\Omega, F, P\}$*

и случайная величина $\xi(\omega)$: $\Omega \to \mathbb{R}$ такие, что $F(x)$ является функцией распределения случайной величины ξ,

$$F_\xi(x) = F(x).$$

◄ Укажем одну из возможных конструкций требуемого пространства. Пусть эксперимент состоит в случайном выборе точки из $\mathbb{R}$ так, что $\Omega = \{\omega\colon\ \omega \in \mathbb{R}\}$. В качестве наблюдаемых случайных событий рассмотрим наименьшую σ-алгебру, порожденную полуинтервалами $[a, b)$ (борелевскую σ-алгебру на прямой), т. е. совокупность всех множеств, полученных из полуинтервалов с помощью не более чем счетного количества объединений, пересечений и дополнений. Определим вероятности на множестве случайных событий следующим образом:

1. Если $A = \{\omega\colon\ a \leqslant \omega < b\}$, то

$$P\{A\} = F(b) - F(a), \tag{8}$$

2. Если $A = \bigcup\limits_{i=1}^{n} A_i$, $A_i = \{\omega\colon\ a_i \leqslant \omega < b_i\}$, $a_1 < b_1 \leqslant a_2 < b_2 < \ldots \leqslant a_n < b_n$, то

$$P\{A\} = \sum_{i=1}^{n} P\{A_i\} = \sum_{i=1}^{n} \bigl(F(b_i) - F(a_i)\bigr). \tag{9}$$

Можно показать, что определенное соотношениями (8) и (9) правило вычисления вероятностей удовлетворяет всем требованиям, предъявляемым к вероятности: условию неотрицательности ($P \geqslant 0$), нормировки ($P\{\Omega\} = 1$) и счетной аддитивности.

В качестве вероятностного пространства возьмем пространство $\{\Omega, F, P\}$, где $\Omega = \{\omega\colon\ \omega \in \mathbb{R}\}$, F – σ-алгебра борелевских множеств, P — правило вычисления вероятностей, порожденное соотношениями (8) и (9).

Положим теперь на Ω

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

Легко видеть, что

$$F_\xi(x) = P\{\omega\colon\ \xi(\omega) < x\} = F(x),$$

чем и завершается доказательство. ►

Рассмотрим несколько примеров.

1. **Равномерное распределение на отрезке** $[a, b]$. Случайную величину ξ будем называть *равномерно распределенной на отрезке* $[a, b]$, если вероятность попадания значений этой случайной величины в любой подпромежуток $\langle\alpha, \beta\rangle$ отрезка $[a, b]$ не зависит от его положения на $[a, b]$ и пропорциональна длине этого промежутка:

$$P\{\alpha \leqslant \xi < \beta\} = P\{\alpha < \xi < \beta\} = P\{\alpha < \xi \leqslant \beta\} = k \cdot (\beta - \alpha). \tag{10}$$

Условие нормировки $P\{a \leqslant \alpha \leqslant b\} = 1$ позволяет легко определить значение k —

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

Функция распределения дается соотношением

$$\boxed{F_\xi(x) = \begin{cases} 0, & x \leqslant a, \\ \dfrac{x-a}{b-a}, & a < x \leqslant b, \\ 1, & x > b. \end{cases}} \tag{11}$$

Равномерную на отрезке $[a, b]$ случайную величину будем обозначать $\xi = R[a, b]$. Числа a и b называются *параметрами равномерного распределения*.

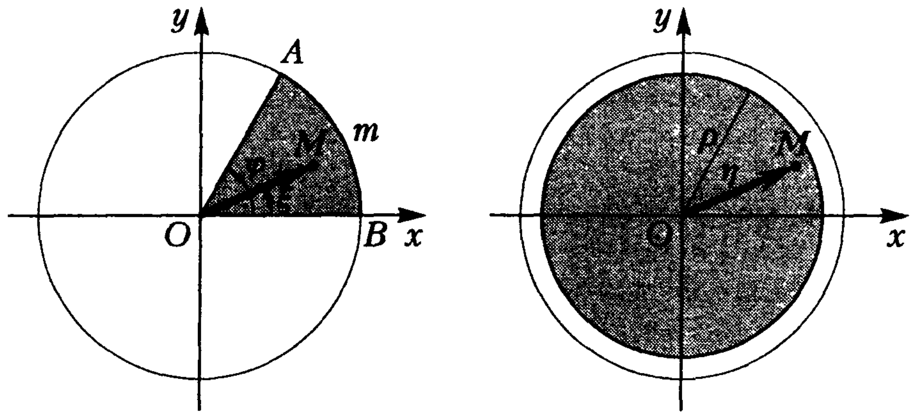


Рис. 3

2. Распределение аргумента и длины радиус-вектора случайной точки в круге радиуса r. Пусть эксперимент состоит в извлечении случайной точки из круга радиуса r с центром в начале координат, так, что вероятность точке быть извлеченной из некоторого борелевского множества D пропорциональна площади этого множества

$$P\{M \in D\} = \frac{S(D)}{\pi r^2}. \tag{12}$$

Пусть ξ — угол, образованный радиус-вектором точки M с положительным направлением оси Ox, η — длина радиус-вектора точки M. Отметим, что ξ и η в рассматриваемом эксперименте являются случайными величинами, и для их функций распределения получаем (см. рис. 3)

$$\forall\ 0 < \varphi < 2\pi \quad P\{\xi < \varphi\} = P\{M \in AmBO\} = \frac{S(\text{сектора})}{\pi r^2} = \frac{\varphi}{2\pi},$$

$$\forall\ 0 \leqslant p \leqslant r \quad P\{\eta < p\} = P\{M \in K_{x^2+y^2<p^2}\} = \frac{\pi p^2}{\pi r^2} = \frac{p^2}{r^2},$$

так, что

$$\boxed{\begin{gathered} F_\xi(\varphi) = \begin{cases} 0, & \varphi \leqslant 0, \\ \dfrac{\varphi}{2\pi}, & 0 < \varphi \leqslant 2\pi, \\ 1, & \varphi > 2\pi, \end{cases} \\ F_\eta(p) = \begin{cases} 0, & p \leqslant 0, \\ \dfrac{p^2}{r^2}, & 0 < p \leqslant R, \\ 1, & p > R. \end{cases} \end{gathered}} \tag{13}$$

Заметим, что ξ — равномерная на $[0, 2\pi]$ случайная величина.

3. Распределение Коши. Пусть L — прямая и A — точка, расположенная на расстоянии $a > 0$ от прямой L (рис. 4) Выберем начало отсчета в точке O, являющейся основанием перпендикуляра, опущенного из A на L. Пусть φ, $-\pi/2 \leqslant \varphi \leqslant \pi/2$ — угол, отсчитываемый от AO вправо или влево и составленный прямой, проходящей через точку A, и отрезком AO.

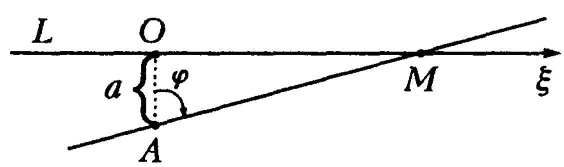


Рис. 4

Эксперимент состоит в следующем: случайным образом выбирается направление прямой, проходящей через точку A, и фиксируется координата точки M пересечения прямых AM и L.

Случайная величина ξ — координата точки M на прямой L — называется *случайной величиной, распределенной по Коши*, и обозначается следующим образом: $\xi = K[a]$.

По определению функции распределения

$$F_\xi(x) = P\{\xi < x\} = P\{a \operatorname{tg} \varphi < x\} = P\left\{ -\frac{\pi}{2} \leqslant \varphi < \operatorname{arctg} \frac{x}{a} \right\}.$$

Предположение о случайности выбора направления прямой AM дает основание для заключения о равномерности распределения угла φ на промежутке $\left[-\pi/2, \pi/2\right]$, при этом последняя вероятность будет равна

$$P\left\{-\frac{\pi}{2} \leqslant \varphi < \operatorname{arctg}\frac{x}{a}\right\} = \frac{\operatorname{arctg} x/a + \pi/2}{\pi} = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}\frac{x}{a},$$

так что

$$\boxed{F_\xi(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}\frac{x}{a}, \quad x \in \mathbb{R}.} \tag{14}$$

Число a называется *параметром распределения Коши*.

4. Экспоненциальное распределение. Эксперимент состоит в наблюдении за временем безотказной работы некоторого агрегата. Известно, что вероятность выхода агрегата из строя в любой промежуток времени длины t одна и та же и не зависит от расположения этого промежутка на временной оси. Это свойство носит название *свойства отсутствия последействия* и означает, что только что включенный прибор и этот же прибор, уже проработавший некоторое время T, имеют одинаковую вероятность выйти из строя в течение времени от 0 до t и от T до $T+t$ соответственно.

Пусть ξ — время безотказной работы агрегата. Тогда свойство отсутствия последействия означает, что

$$P\{\xi < T+t\colon\ \xi \geqslant T\} = P\{\xi < t\}, \qquad T \geqslant 0,\ t \geqslant 0.$$

Предполагая, что $P\{\xi \geqslant T\} > 0$, и используя формулу условной вероятности, получаем

$$\frac{P\{T \leqslant \xi < T+t\}}{P\{\xi \geqslant T\}} = P\{\xi < t\}. \tag{15}$$

Если ввести функцию распределения $F_\xi(t)$ времени безотказной работы агрегата, то соотношение (15) примет вид

$$F_\xi(T+t) - F_\xi(T) = [1 - F_\xi(T)]\,F_\xi(t). \tag{16}$$

Это уравнение относительно функции распределения $F_\xi(t)$, которая дополнительно к (16) обладает тем свойством, что $F_\xi(0) = 0$.

Предположим, что $F_\xi(t)$ — непрерывная функция. Тогда уравнение (16) имеет своим решением функцию

$$\boxed{F_\xi(t) = \begin{cases} 0, & t \leqslant 0, \\ 1 - e^{-\mu t}, & t > 0. \end{cases}} \tag{17}$$

Случайная величина ξ, имеющая распределение (17), называется *экспоненциальной случайной величиной с параметром* $\mu > 0$ и будет в дальнейшем обозначаться так — $\xi = \exp[\mu]$.

◀ Для доказательства формулы (17), заметим, что любая непрерывная на $t \geqslant 0$ функция $F_\xi(t)$, удовлетворяющая уравнению (16), является дифференцируемой. Действительно, пусть $\varphi(T)$ — произвольная дифференцируемая функция, обращающаяся в нуль вне некоторого промежутка $\langle\alpha, \beta\rangle$, $\beta > \alpha \geqslant 0$. Умножая обе части соотношения (16) на $\varphi(T)$ и интегрируя на полупрямой от 0 до $+\infty$, получаем

$$\int_0^\infty \varphi(T)F_\xi(T+t)\,dT = \int_0^\infty \varphi(T)F_\xi(T)\,dT + \int_0^\infty \varphi(T)F_\xi(t)\,dT - \int_0^\infty \varphi(T)F_\xi(T)F_\xi(t)\,dT,$$

или, полагая в первом из интегралов $T + t = u$, —

$$\int_t^\infty \varphi(u-t)F_\xi(u)\,du = A + BF_\xi(t), \tag{18}$$

где

$$A = \int_0^\infty \varphi(T)F_\xi(T)\,dT, \qquad B = \int_0^\infty \varphi(T)\,[1 - F_\xi(T)]\,dT$$

— постоянные. В силу дифференцируемости φ и непрерывности F_ξ интеграл в соотношении (18) — дифференцируемая по переменной t функция. Следовательно, $F_\xi(t)$ — дифференцируемая функция.

Продифференцируем теперь (16), например, по переменной T. Положив $T = 0$, получим

$$F'_\xi(t) = F'_\xi(0)(1 - F_\xi(t))$$

— обыкновенное дифференциальное уравнение первого порядка относительно $F_\xi(t)$ с начальным условием $F_\xi(0) = 0$. Для завершения доказательства формулы (17) заметим, что $F'(0) = \mu > 0$ в силу неубывания функции распределения $F_\xi(t)$. ▶

1.2. Плотность распределения

Плотностью распределения вероятностей на множестве значений случайной величины ξ (коротко: *плотностью распределения случайной величины*) будем называть функцию $f_\xi(x) \geqslant 0$ такую, что

$$\forall\, [a, b) \in \mathbb{R} \quad P\{a \leqslant \xi < b\} = \int\limits_a^b f_\xi(x)\, dx. \tag{19}$$

Не всякую случайную величину можно описать плотностью — существуют случайные величины, для которых функция $f_\xi(x)$, обладающая свойством (19), не может быть определена. Тем не менее, если плотность у случайной величины существует, то она обладает следующими свойствами:

1.

$$\boxed{F_\xi(x) = \int\limits_{-\infty}^{x} f_\xi(x)\, dx.} \tag{20}$$

◀ Легко следует из определения (19). ▶

Замечание. Отметим, что в тех точках $x \in \mathbb{R}$, где $f_\xi(x)$ — непрерывна, имеет место соотношение

$$\boxed{\frac{dF_\xi}{dx} = f_\xi(x).} \tag{21}$$

Вообще же, функция распределения $F_\xi(x)$ непрерывна на всей числовой прямой, что немедленно следует из (20).

2. Если x_0 — точка непрерывности $f_\xi(x)$, то

$$\boxed{P\{x_0 \leqslant \xi < x_0 + \Delta x\} = f_\xi(x_0)\Delta x + 0(\Delta x),} \tag{22}$$

или

$$\boxed{f_\xi(x_0) = \lim_{\Delta x \to 0} \frac{P\{x_0 \leqslant \xi < x_0 + \Delta x\}}{\Delta x}.} \tag{23}$$

Замечание. Последнее соотношение означает, что «плотность» распределения вероятностей действительно плотность· «количество вероятности», приходящееся на единицу длины.

◀ Оба соотношения (22) и (23) следуют из определения (19). ▶

3. Условие нормировки плотности

$$\boxed{\int\limits_{-\infty}^{\infty} f_\xi(x)\, dx = 1.} \tag{24}$$

◀ Следует из (20) и соответствующего свойства функции распределения. ▶

Так же, как и функция распределения (и любая другая форма, в которой может быть задан закон распределения), плотность не определяет однозначно случайную величину — разные случайные величины могут иметь одну и ту же плотность.

Любая неотрицательная интегрируемая на $\mathbb{R}$ функция, удовлетворяющая условию нормировки (20), может быть плотностью некоторой случайной величины. Примером такой случайной величины служит случайная величина ξ, функция распределения которой дается соотношением (20),

$$F_\xi(x) = \int_{-\infty}^{x} f_\xi(x)\,dx.$$

◀ Действительно, $F_\xi(x)$ монотонно не убывает (в силу неотрицательности $f(x)$), непрерывна на всей числовой прямой и принимает на концах значения 0 и 1 соответственно. Поэтому $F_\xi(x)$, как это следует из теоремы предыдущего пункта, является функцией распределения некоторой случайной величины, для которой $f_\xi(x)$ будет плотностью. ▶

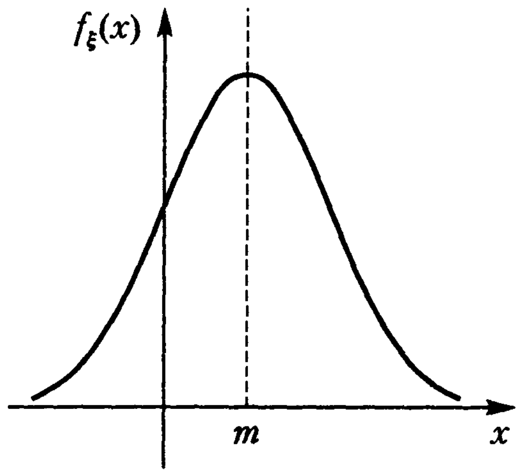


Рис. 5

Пример 1. Нормальная (гауссова) случайная величина. Случайная величина ξ называется *нормальной*, или *гауссовой случайной величиной*, если ее плотность распределения задается соотношением ($\sigma > 0$, m — произвольное действительное число)

$$\boxed{f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R}.} \tag{25}$$

◀ Отметим, что формула (25) действительно задает плотность распределения при любых значениях параметра m и любых положительных значениях параметра σ, так как

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} dx = \left| \begin{array}{c} \dfrac{x-m}{\sigma} = y \\ dx = \sigma dy \end{array} \right| = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2}\right\} dy = 1. \text{ ▶}$$

График функции $f_\xi(x)$, описываемой соотношением (25), приведен на рис. 5. При любых значениях параметров m и σ это симметричная (относительно $x = m$) колоколообразная кривая, быстро убывающая при $x \to \infty$. Исходя из вероятностного смысла плотности, случайная величина, подчиненная нормальному закону, с большей вероятностью принимает значения, лежащие около $x = m$, и с меньшей вероятностью — значения, заметно отличающиеся от $x = m$. Параметр m показывает расположение оси симметрии распределения

$$P\{\xi < m\} = P\{\xi \geqslant m\} = \frac{1}{2},$$

а параметр σ — степень «размазанности» вероятности по числовой прямой (рис. 6).

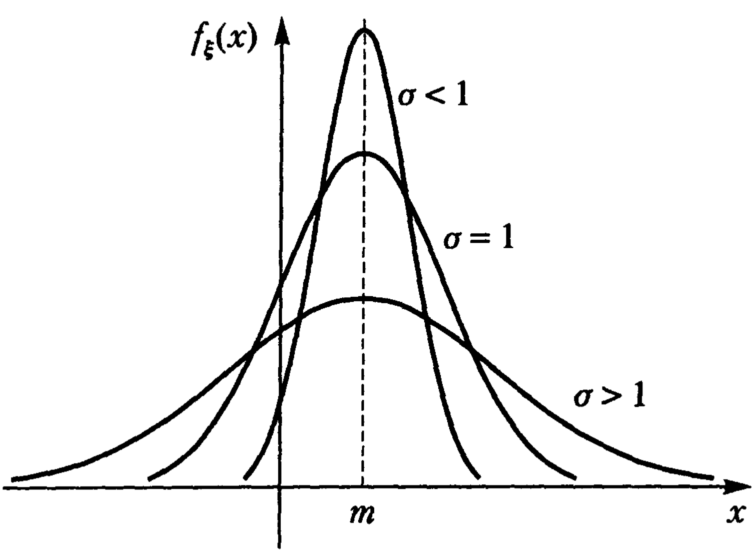


Рис. 6

Нормальное распределение характеризует следующую тенденцию в поведении случайной величины: в эксперименте с высокой вероятностью реализуются значения, близкие к m, причем чем меньше значение σ, тем меньше могут отклоняться наблюдаемые в эксперименте значения ξ от m. Этому качественному утверждению можно придать количественную форму.

Правило 3σ. *В подавляющем большинстве случаев значения, принимаемые случайной величиной, имеющей нормальное распределение с параметрами m и σ, отличаются от m не более чем на 3σ. Точнее*

$$P\{|\xi - m| < 3\sigma\} = 0{,}9973. \qquad (26)$$

◄ Рассмотрим

$$P\{|\xi - m| < 3\sigma\} = P\{m - 3\sigma < \xi < 3\sigma + m\} = F_\xi(3\sigma) - F_\xi(-3\sigma) =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{m-3\sigma}^{m+3\sigma} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} dx = \left| \begin{array}{c} \dfrac{x-m}{\sigma} = y, \ dx = \sigma dy \\ y_{\text{н}} = -3, \ y_{\text{в}} = 3 \end{array} \right| =$$

$$= \frac{1}{\sqrt{2\pi}} \int\limits_{-3}^{3} \exp\left\{-\frac{y^2}{2}\right\} dy = F(3) - F(-3),$$

где

$$F(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} \exp\left\{-\frac{y^2}{2}\right\} dy$$

— функция Лапласа, таблица значений которой приведена на с. 69. В силу симметрии $f_\xi(x)$

$$F(-x) = 1 - F(x),$$

и, так как $F(3) = 0{,}99865$, то для искомой вероятности получаем

$$P\{|\xi - m| < 3\sigma\} = 2F(3) - 1 = 0{,}9973. \ ►$$

Ниже будет показано (см. Центральную предельную теорему), что при достаточно широких предположениях случайные величины, являющиеся суммой большого числа независимых слагаемых, будут иметь распределение, близкое к нормальному.

Нормальную случвйную величину с пврвметрами m, σ будем обозначать так: $\xi = N[m, \sigma]$.

В заключение отметим, что рассмотренные выше равномерная на промежутке $[a, b]$ случайная величина, экспоненциальная случайная величина и величина, распределенная по Коши, обладают плотностями, графики которых изображены на рис. 7.

Распределение, обладающее плотностью, будем называть *абсолютно непрерывным распределением*, а соответствующую случайную величину — *непрерывной случайной величиной*.

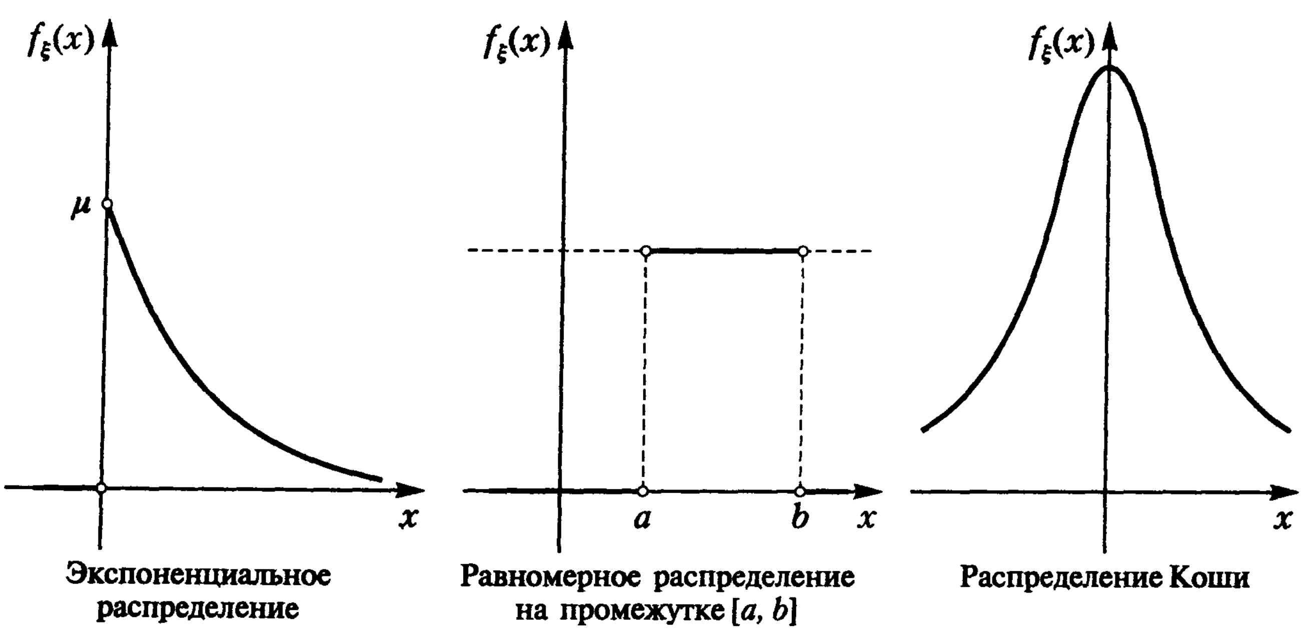


Рис. 7

Если случайная величина непрерывна, то каждое свое значение она принимает с нулевой вероятностью,

$$\forall\, x \in \mathbb{R} \quad P\{\xi = x\} = 0.$$

Это, очевидно, следует из соотношения (19). Не следует, однако, считать, что подобное свойство присуще лишь непрерывным случайным величинам — существуют (см. п. 2.1.4, пример 1) случайные величины, не являющиеся непрерывными, однако также обладающие этим свойством.

Таблица значений функции $F(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} \exp\left\{-\dfrac{y^2}{2}\right\} dy$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,5000	5040	5080	5120	5159	5199	5239	5279	5319	5359
0,1	5398	5438	5478	5517	5557	5596	5636	5675	5714	5753
0,2	5793	5832	5871	5909	5948	5987	6026	6103	6064	6141
0,3	6179	6217	6255	6293	6331	6368	6406	6443	6480	6517
0,4	6554	6591	6628	6664	6700	6736	6772	6808	6844	6879
0,5	6915	6950	6985	7019	7054	7088	7123	7157	7190	7224
0,6	7257	7291	7324	7356	7389	7421	7454	7486	7517	7549
0,7	7580	7611	7642	7673	7703	7734	7764	7793	7823	7852
0,8	7881	7910	7939	7967	7995	8023	8051	8078	8106	8133
0,9	8159	8186	8212	8238	8264	8289	8315	8340	8365	8389
1,0	8413	8437	8461	8485	8508	8531	8554	8577	8599	8621
1,1	8643	8665	8686	8708	8728	8749	8770	8790	8810	8830
1,2	8849	8869	8888	8906	8925	8943	8962	8980	8997	9015
1,3	9032	9049	9066	9082	9099	9115	9131	9147	9162	9177
1,4	9192	9207	9222	9236	9251	9265	9279	9292	9306	9319
1,5	9332	9345	9357	9370	9382	9394	9406	9418	9429	9441
1,6	9452	9463	9474	9484	9495	9505	9515	9525	9535	9545
1,7	9554	9564	9573	9582	9591	9599	9608	9616	9625	9633
1,8	9641	9648	9656	9664	9671	9678	9686	9692	9699	9706
1,9	9713	9719	9726	9732	9738	9744	9750	9756	9761	9767
2,0	9772	9778	9783	9788	9793	9798	9803	9808	9812	9817
2,1	9821	9826	9830	9834	9838	9842	9846	9850	9854	9857
2,2	9861	9864	9868	9871	9874	9878	9881	9884	9887	9890
2,3	9893	9896	9998	9901	9904	9906	9909	9911	9913	9916
2,4	9918	9920	9922	9924	9927	9929	9930	9932	9934	9936
2,5	9938	9940	9941	9943	9945	9946	9948	9949	9951	9952
2,6	9953	9955	9956	9957	9958	9960	9961	9962	9963	9964
2,7	9965	9966	9967	9968	9969	9970	9971	9972	9973	9974
2,8	9974	9975	9976	9977	9977	9978	9979	9979	9980	9981
2,9	9981	9982	9982	9983	9984	9984	9985	9985	9986	9986
3,0	0,9986		3,1	9990	3,2	99931	3,3	99952	3,4	99966
3,5	9998		3,6	9998	3,7	99989	3,8	99993	3,9	99995
4,0	999968									
4,5	999997									
5,0	99999997									

1.3. Ряд распределения

Определение. *Рядом распределения случайной величины* ξ называется такая форма задания закона распределения, когда перечисляются все возможные значения случайной величины с указанием положительных вероятностей, с которыми случайная величина эти значения принимает.

Ряд распределения может быть задан таблицей вида

ξ	a_1	a_2	...	a_n	...
$P\{\xi\}$	p_1	p_2	...	p_n	...

(*)

где все a_i различны, $a_i \neq a_j$, $i \neq j$, и

$$\sum_{i=1}^{\infty} p_i = 1. \tag{27}$$

При помощи подобного ряда можно задать закон распределения только такой случайной величины, которая принимает с положительными вероятностями не более чем счетное множество значений. Такие случайные величины называются *дискретными*.

Отметим следующие очевидные соотношения:

1. $$\boxed{\forall\,[a,b)\in\mathbb{R}\quad P\{\xi\in[a,b)\} = \sum_{i\ a_i\in[a,b)} p\{\xi = a_i\},} \tag{28}$$

2. $$\boxed{\forall\,x\in\mathbb{R}\quad F_\xi(x) = P\{\xi < x\} = \sum_{i\ a_i<x} p\{\xi = a_i\},} \tag{29}$$

первое из которых показывает, что ряд (*) является законом распределения, а второе дает выражение для функции распределения дискретной случайной величины.

Пример 1. Распределение Бернулли. Случайная величина, принимающая значение 1 с вероятностью p и значение 0 с вероятностью $q = 1 - p$, называется *бернуллиевой случайной величиной*. Ее ряд распределения дается таблицей

ξ	0	1
P	q	p

Функция распределения бернуллиевой случайной величины имеет вид

$$\boxed{F_\xi(x) = \begin{cases} 0, & x \leqslant 0, \\ q, & 0 < x \leqslant 1, \\ 1, & x > 1 \end{cases}}$$

(рис 8).

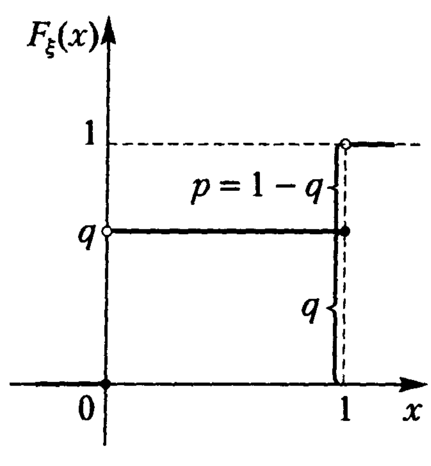


Рис 8

Бернуллиева случайная величина обычно появляется в приложениях как индикатор (характеристическая случайная величина) некоторого события если в эксперименте Ω выделить событие Y, интересующее исследователя, а все прочие события объединить в $\overline{Y}$, так, что $\overline{Y} + Y = \Omega$, то случайная величина $\xi(\omega) = 1$, $\omega \in Y$, $\xi(\omega) = 0$, $\omega \in \overline{Y}$, будет *индикатором* события Y

Пример 2. Биномиальное распределение.

Определение. Случайная величина ξ называется *биномиальной случайной величиной с параметрами* (n, p), если она описывает количество успехов в серии из n независимых испытаний, в каждом из которых вероятность успеха одна и та же и равна p.

Обозначение: $\xi = B[n; p]$.

Ряд распределения биномиальной случайной величины (см. биномиальный эксперимент) может быть задан соотношением

$$\forall k=0,1,\ldots,n \quad P\{\xi=k\}=C_n^k p^k(1-p)^{n-k}. \qquad (30)$$

Функция распределения биномиальной случайной величины имеет вид

$$\boxed{F_\xi(x)=\sum_{k=0}^{k<x} C_n^k p^k(1-p)^{n-k}.} \qquad (31)$$

Отметим, что если вероятность $P\{\xi = k\}$ изображать вертикальным отрезком с абсциссой в точке $\xi = k$, то распределение биномиальных вероятностей будет изображаться графической схемой, приведенной на рис. 9.

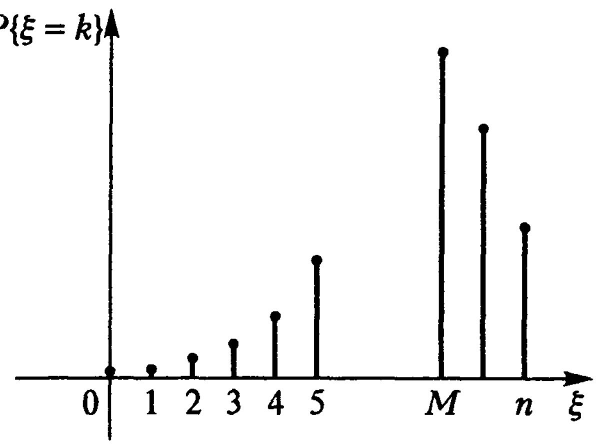


Рис. 9

Если $M = \underset{0\leqslant k\leqslant n}{\operatorname{argmax}} P\{\xi = k\}$, то легко видеть, что

$$P\{\xi = M-1\} \leqslant P\{\xi = M\} \quad \text{и} \quad P\{\xi = M+1\} \leqslant P\{\xi = M\},$$

поэтому

$$p(n+1)-1 \leqslant M \leqslant p(n+1). \qquad (32)$$

Отсюда ясно, что вероятности $P\{\xi = k\}$ монотонно растут от $k = 0$ до $k = M$, а затем монотонно убывают от $k = M$ до $k = n$. Значение $\xi = M$ называется *наиболее вероятным*. Заметим, что наиболее вероятных значений у биномиальной случайной величины не более двух и по крайней мере одно всегда есть.

Пример 3. Распределение Пуассона. Рассматривается последовательность n независимых экспериментов Ω с постоянной вероятностью успеха p в каждом испытании. Будем предполагать, что вероятность успеха p крайне мала, так что успех — событие в отдельном испытании редкое, однако количество испытаний n настолько велико, что $np = \lambda$ — постоянная величина. Тогда получим

$$C_n^k p^k(1-p)^{n-k} = \frac{n(n-1)\ldots(n-k+1)}{k!}\,p^k(1-p)^{n-k} =$$
$$= \frac{(np)^k}{k!}\cdot\frac{n}{n}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\cdot\ldots\cdot\frac{n-k}{n}\cdot(1-p)^{-k}\cdot\left(1-\frac{\lambda}{n}\right)^n = \frac{\lambda^k}{k!}\,e^{-\lambda}(1+o(1)).$$

Последнее дает основание для следующего определения:

Определение. *Пуассоновой случайной величиной с параметром* λ будем называть дискретную случайную величину, ряд распределения которой дается соотношениями

$$P\{\xi=k\}=e^{-\lambda}\frac{\lambda^k}{k!}, \quad k=0,1,2,\ldots,k,\ldots. \qquad (33)$$

Обозначение: $\xi = \Pi[\lambda]$.

Определение корректно, так как

$$\sum_{k=0}^{\infty} P\{\xi = k\} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}\cdot e^{\lambda} = 1.$$

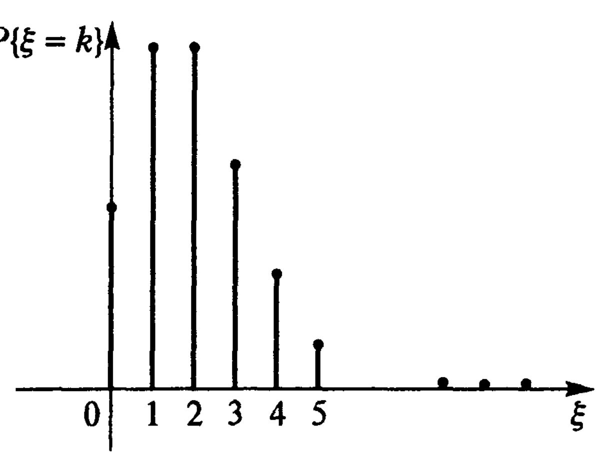


Рис. 10

Таким образом, пуассонова случайная величина описывает количество успехов в серии из неограниченно большого количества испытаний в предположении, что вероятность успеха в каждом отдельном испытании исчезающе мала. Пуассоновым распределением хорошо описываются такие случайные величины, как количество вызовов, поступающих на АТС в единицу времени, количество звезд, наблюдающихся в заданном объеме Галактики;

количество деревьев, приходящихся на единицу площади леса; количество радиоактивных частиц, регистрируемых счетчиком Гейгера в единицу времени, и другие. Даже такая экзотическая величина, как количество лиц, убитых ударом копыта в прусском армейском корпусе в течение года, — и та подчиняется пуассоновому распределению.

Графически распределение пуассоновых вероятностей на множестве целых неотрицательных чисел изображено на рис. 10 ($\lambda = 2$).

1.4. Теорема Лебега. Обобщенные плотности

Пусть ξ — случайная величина, $F_\xi(x)$ — ее функция распределения (являющаяся, напомним, универсальной формой задания закона распределения вероятностей на множестве значений случайной величины). В некоторых случаях, наряду с функцией распределения, можно указать и другие формы задания закона распределения: плотность распределения или ряд распределения. В первом из указанных случаев функция распределения $F_\xi(x)$ должна быть абсолютно непрерывной, т. е.

$$\exists\ f_\xi(x):\ F_\xi(x) = \int\limits_{-\infty}^{x} f_\xi(x)\,dx,$$

во втором функция распределения $F_\xi(x)$ должна быть кусочно постоянной. Ясно, что совокупность функций $F_\xi(x)$, которые могут служить функциями распределения этими двумя классами не исчерпывается (см. теорему раздела 2.1.1). Например, кусочно абсолютно непрерывные функции $F_\xi(x)$ описывают случайные величины, являющиеся смесью дискретных и непрерывных. Однако помимо довольно широкого класса подобных случайных величин (о них ниже) существуют еще случайные величины, не являющиеся ни дискретными, ни непрерывными, ни смешанными.

Пример 1 (Канторова лестница). Пусть C — канторово множество на отрезке $[0, 1]$. Определим функцию $F_\xi(x)$ следующим образом: $F_\xi(x) = 0$, $x \leqslant 0$, и $F_\xi(x) = 1$, $x > 1$. На отрезке $\left[\frac{1}{3}, \frac{2}{3}\right]$ положим $F_\xi(x) = \frac{1}{2}$, на отрезках $\left[\frac{1}{9}, \frac{2}{9}\right]$ и $\left[\frac{7}{9}, \frac{8}{9}\right]$ положим соответственно $F_\xi(x) = \frac{1}{4}$ и $F_\xi(x) = \frac{3}{4}$. На отрезках $\left[\frac{1}{27}, \frac{2}{27}\right]$, $\left[\frac{7}{27}, \frac{8}{27}\right]$, $\left[\frac{19}{27}, \frac{20}{27}\right]$ и $\left[\frac{25}{27}, \frac{26}{27}\right]$ припишем $F_\xi(x)$ значения $\frac{1}{8}$, $\frac{3}{8}$, $\frac{5}{8}$ и $\frac{7}{8}$ соответственно, и так далее, до бесконечности (рис. 11).

Отметим следующие свойства определенной таким образом функции $F_\xi(x)$:

1. $F_\xi(x)$ монотонно не убывает,
2. $F_\xi(x)$ принимает все значения из промежутка $[0, 1]$,
3. $F_\xi(x)$ непрерывна на $\mathbb{R}$,
4. $F_\xi(x)$ постоянна в окрестности любой точки, не принадлежащей множеству C.

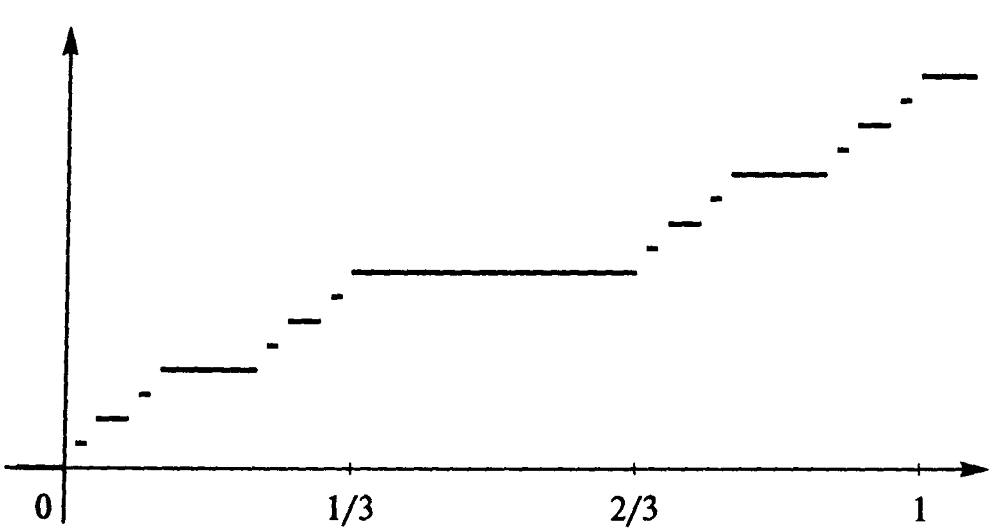


Рис. 11

Пусть теперь ξ — случайная величина, для которой $F_\xi(x)$ — функция распределения. ξ не является дискретной, так как $F_\xi(x)$ — непрерывна, и следовательно, $P\{\xi = x\} = 0 \ \forall\, x \in \mathbb{R}$. ξ также не является непрерывной, так как в этом случае функция $F_\xi(x)$ должна была бы восстанавливаться по своей производной, но по свойству 4 производная почти всюду (во всех точках отрезка $[0, 1]$ за исключением точек канторова множества) равна нулю.

Этот пример является в некотором смысле исчерпывающим: оказывается любая случайная величина — это «смесь» дискретной, непрерывной и сингулярной (подобной рассмотренной в примере 1 случайных величин. Точнее, имеет место следующая теорема.

Теорема Лебега. *Если ξ — случайная величина с функцией распределения $F_\xi(x)$, то существуют функции $F_{\text{н}}(x)$, $F_{\text{д}}(x)$ и $F_{\text{с}}(x)$ такие, что $F_\xi(x)$ единственным образом представима в виде*

$$F_\xi(x) = F_{\text{н}}(x) + F_{\text{д}}(x) + F_{\text{с}}(x) \tag{34}$$

и при этом

1.
$$\exists\, f(x) \geqslant 0,\ F_{\text{н}}(x) = \int_{-\infty}^{x} f(x)\,dx, \quad \int_{-\infty}^{\infty} f(x)\,dx = q < 1, \tag{35}$$

2.
$$\exists\, \{a_i\}_{i=1}^{\infty},\ \{p_i\}_{i=1}^{\infty},\ p_i > 0, \quad \sum_{i=1}^{\infty} p_i = p = 1 - q, \quad F_{\text{д}}(x) = \sum_{i:\ a_i < x} p_i, \tag{36}$$

3. $F_{\text{с}}(x)$ — *непрерывна на $\mathbb{R}$, монотонно не убывает и постоянна почти всюду, имеет изменение, равное $0 < r < 1$, так что $p + q + r = 1$.*

Если сингулярная компонента отсутствует, то случайная величина ξ является «смесью» дискретной и непрерывной компонент

$$\boxed{F_\xi(x) = F_{\text{н}}(x) + F_{\text{д}}(x).} \tag{37}$$

В дальнейшем в этом случае нам будет удобно использовать следующее соглашение: если через $\delta(x)$ обозначить дельта-функцию Дирака, обладающую тем свойством, что

$$\boxed{\begin{array}{l} \mathbf{1.}\ \displaystyle\int_{-\infty}^{\infty} \delta(x)\,dx = 1, \\ \mathbf{2.}\ \displaystyle\int_{-\infty}^{\infty} \delta(x) f(x)\,dx = f(0), \end{array}}$$

то *обобщенной плотностью* распределения (37) будем называть формальное выражение вида

$$\boxed{g_\xi(x) = f(x) + \sum_{i=1}^{\infty} \delta(x - a_i) p_i,} \tag{38}$$

где

$$\int_{-\infty}^{\infty} f(x)\,dx = q; \quad \sum_{i=1}^{\infty} p_i = p, \quad p + q = 1, \tag{39}$$

так что

$$P\{\xi \in [a, b)\} = \int_a^b g_\xi(x)\, dx$$

и

$$\int_{-\infty}^{x} g_\xi(x)\, dx = \int_{-\infty}^{x} f(x)\, dx + \sum_{i:\ a_i < x} p_i = F_{\text{н}}(x) + F_{\text{д}}(x) = F_\xi(x). \tag{40}$$

Последнее соотношение является формальным аналогом соотношения (20), дающего выражение функции распределения через плотность. В дальнейшем мы будем понимать обобщенную плотность $g_\xi(x)$ как условную форму записи интегральных соотношений (40).

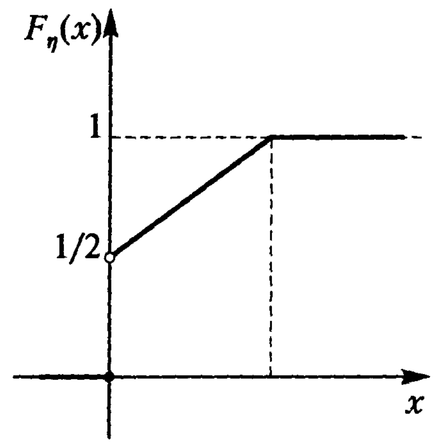


Рис. 12

Пример 2. Пусть $\xi = R[-1, 1]$, $\eta = \max(\xi, 0)$. Найдем закон распределения случайной величины η.

$$F_\xi(x) = P\{\eta < x\} = P\{\max(\xi, 0) < x\}.$$

По определению ξ не принимает значений, больших 1. Поэтому

$$P\{\eta < x\} = 1, \qquad x \geqslant 1;$$

в то же время $\max(\xi, 0) \geqslant 0$, откуда

$$P\{\eta < x\} = 0, \qquad x \leqslant 0.$$

Для $x \in (0, 1)$ получим

$$F_\xi(x) = P\{\max(\xi, 0) < x\} = P\{\xi < x\} = \frac{x+1}{2}$$

(рис. 12). Это распределение является распределением описанного выше типа. Обобщенная плотность может быть представлена здесь в виде

$$g_\xi(x) = \begin{cases} 0, & x < 0, \\ \dfrac{1}{2} + \dfrac{1}{2}\,\delta(0), & 0 \leqslant x < 1, \\ 0, & x > 1. \end{cases}$$

§ 2. Системы случайных величин — случайные векторы

Пусть $\{\Omega, F, P\}$ — вероятностное пространство и каждому элементарному исходу $\omega \in \Omega$ поставлена в соответствие упорядоченная совокупность n действительных чисел $\boldsymbol{\xi}(\omega) = \{\xi_1(\omega), \ldots, \xi_n(\omega)\}$.

Определение. Будем называть $\boldsymbol{\xi}(\omega)$ *случайным вектором*, или *системой n случайных величин*, если для любого параллелепипеда Π^n

$$\Pi^n = \left\{\mathrm{x} \in \mathbb{R}^n\colon\ a_i \leqslant x_i < b_i,\ i = 1, 2, \ldots, n\right\}$$

множество тех $\omega \in \Omega$, для которых $\boldsymbol{\xi}(\omega) \in \Pi^n$, является случайным событием.

В этом случае легко показать, что случайными событиями будут и все события вида $\{\omega\colon\ \boldsymbol{\xi}(\omega) = A\}$, где A — произвольное борелевское множество в $\mathbb{R}^n$ и тем самым можно определить *вероятностную функцию*

$$\boxed{\Phi_{\boldsymbol{\xi}}(A) = P\{\boldsymbol{\xi}^{-1}(A)\} = P\{\omega\colon\ \boldsymbol{\xi}(\omega) = A\},}$$

определенную на σ-алгебре борелевских множеств в $\mathbb{R}^n$.

Понятие случайного вектора формализует в теории вероятностей возможность в одном эксперименте зафиксировать несколько различных числовых показателей.

Пример 1. Эксперимент состоит в случайном выборе точки M из круга $K = \{(x, y)\colon x^2 + y^2 \leqslant 1\}$ на плоскости $\mathbb{R}^2$. Пусть ξ_1 — абсцисса, а ξ_2 — ордината выбранной точки. Тогда $\boldsymbol{\xi} = \{\xi_1, \xi_2\}$ — случайный вектор.

◀ В силу случайности выбора точки из круга K можно считать, что для любого борелевского множества $A \subseteq K$ вероятность выбора точки M пропорциональна площади этого множества

$$P\{M \in A\} = \frac{S(A)}{\pi}. \tag{1}$$

Вектор $\boldsymbol{\xi}$ — радиус-вектор точки M. Для любого борелевского $B \subset \mathbb{R}^2$ получаем

$$P\{\boldsymbol{\xi} \in B\} = \frac{S(B \cap K)}{\pi}$$

(рис. 13), и тем самым для произвольного борелевского множества $B \in \mathbb{R}^2$ определена вероятностная функция $\Phi_{\xi}(B) = P\{\boldsymbol{\xi} \in B\}$. ▶

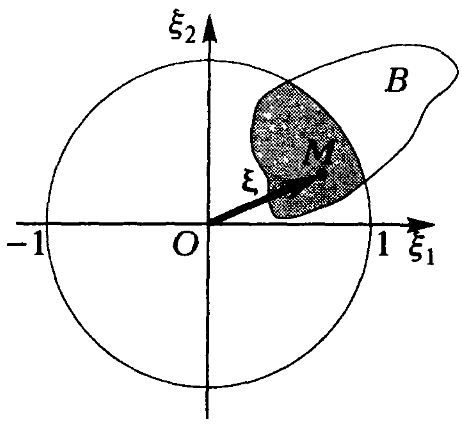


Рис. 13

Легко установить, что компоненты случайного вектора — случайные величины. Несколько менее очевидно, но столь же легко устанавливается, что если $\xi_1, \xi_2, \ldots, \xi_n$ — случайные величины в некотором эксперименте, и A — борелевское множество из $\mathbb{R}^n$, то событие $\{\omega\colon\ \boldsymbol{\xi}(\omega) \in A\}$ — является случайным событием в этом эксперименте и, следовательно, $\boldsymbol{\xi} = \{\xi_1, \ldots, \xi_n\}$ — случайный вектор.

С компонентами случайного вектора можно производить различные операции, и в подавляющем большинстве случаев результат этих операций является осмысленным с рассматриваемой — вероятностной — точки зрения, т. е. случайной величиной. Точнее, имеет место следующая

Теорема. *Пусть $\boldsymbol{\xi}$ — случайный вектор, а $\varphi(\mathrm{x}) = \varphi(x_1, \ldots, x_n)$ — борелевская функция в $\mathbb{R}^n$. Тогда $\eta = \varphi(\boldsymbol{\xi})$ является случайной величиной.*

◀ То обстоятельство, что φ — борелевская функция, означает, что прообраз полуинтервала $\varphi^{-1}([a, b))$ является борелевским множеством в $\mathbb{R}^n$, а это, в свою очередь, означает, что $\varphi(\boldsymbol{\xi})$ — случайная величина, так как

$$\{\omega\colon\ \varphi(\boldsymbol{\xi}) \in [a, b)\} = \left\{\omega\colon\ \varphi^{-1}([a, b)) \ni \boldsymbol{\xi}\right\}. ▶$$

Особо отметим, что сумма, разность, произведение, частное (если знаменатель «почти всегда» не нуль) пары случайных величин будут случайными величинами. Случайными величинами будут также результаты применения любых непрерывных, кусочно непрерывных или монотонных функций к случайным величинам.

2.1. Законы распределения случайных векторов

Пусть $\boldsymbol{\xi}\{\xi_1, \ldots, \xi_n\}$ — случайный вектор, и A — произвольное борелевское множество в $\mathbb{R}^n$. *Законом распределения* $\boldsymbol{\xi}$ будем называть любое правило, позволяющее вычислять вероятности

$$\Phi_{\boldsymbol{\xi}}(A) = P\{\omega\colon\ \boldsymbol{\xi}(\omega) \in A\}. \tag{2}$$

Как и в случае одномерных случайных величин, закон распределения случайного вектора может быть задан в произвольной форме. Универсальной формой задания закона распределения случайного вектора является *функция распределения*, определяемая соотношением

$$F_{\xi}(\mathrm{x}) = F_{\xi_1,\dots,\xi_n}(x_1, \dots , x_n) = P\{\xi_1 < x_1; \xi_2 < x_2; \dots ; \xi_n < x_n\}. \tag{3}$$

Как и выше (см. п. 2.1.1), легко устанавливаются следующие свойства функции (3).

1. Ограниченность

$$\boxed{0 \leqslant F_{\xi}(\mathrm{x}) \leqslant 1.}$$

2. $F_{\xi}(\mathrm{x})$ — закон распределения.

Чтобы не загромождать изложение выкладками, приведем в качестве примера формулу нахождения вероятностей $P\{\xi \in \Pi^n\}$ через значения $F_{\xi}(\mathrm{x})$ для $n = 2$ (рис. 14):

$$\begin{aligned} P\{\xi \in \Pi^2\} &= P\{a_1 \leqslant \xi_1 < b_1; a_2 \leqslant \xi_2 < b_2\} = \\ &= F_{\xi_1\xi_2}(b_1, b_2) - F_{\xi_1\xi_2}(a_1, b_2) - \\ &\quad - F_{\xi_1\xi_2}(b_1, a_2) + F_{\xi_1\xi_2}(a_1, a_2). \end{aligned} \tag{4}$$

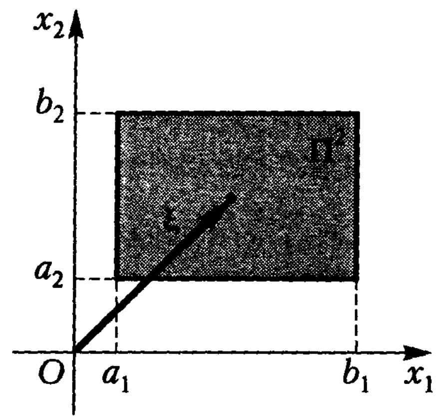


Рис. 14

3. «Монотонность» $\forall$ $\mathrm{x}, \mathrm{y} \in \mathbb{R}^n$ таких, что $\mathrm{x} \leqslant \mathrm{y}$ (т. е. $x_i \leqslant y_i$, $i = 1, 2, \dots , n$)

$$\boxed{F_{\xi}(\mathrm{x}) \leqslant F_{\xi}(\mathrm{y}).}$$

4. Существование пределов

$$\boxed{\lim_{|\mathrm{x}|\to-\infty} F_{\xi}(\mathrm{x}) = 0, \qquad \lim_{\min\limits_i x_i\to\infty} F_{\xi}(\mathrm{x}) = 1.}$$

5. Непрерывность слева

$$\boxed{\forall\, \mathrm{x}_n \to \mathrm{x}, \quad \mathrm{x}_n \leqslant \mathrm{x}, \quad n = 1, 2, \dots , \quad \forall \lim_{\mathrm{x}_n\to\mathrm{x}} F_{\xi}(\mathrm{x}).}$$

Аналог теоремы п. 2.1.1 справедлив и в этой ситуации.

Отметим, что знание функции распределения случайного вектора позволяет определить функцию распределения любого подвектора, в том числе и функции распределения отдельных компонент. Действительно, пусть $\{\xi_{i_1}, \xi_{i_2}, \dots , \xi_{i_k}\} = \xi_{\mathrm{i}}$ — подвектор случайного вектора $\xi = \{\xi_1, \xi_2, \dots , \xi_n\}$. Тогда

$$F_{\xi_{\mathrm{i}}}(x_{i_1}, x_{i_2}, \dots , x_{i_k}) = \lim_{\{x_j\to\infty,\, j\neq i_s,\, s=1,2,\dots,k\}} F_{\xi}(x_1, x_2, \dots , x_n), \tag{5}$$

в частности,

$$F_{\xi_i}(x_i) = \lim_{\{x_j\to\infty,\, j\neq i\}} F_{\xi}(x_1, x_2, \dots , x_n).$$

Знание законов распределения компонент вектора ξ не позволяет, вообще говоря, восстановить функцию распределения вектора. Впрочем, это понятно — индивидуальные законы распределения компонент не учитывают информацию о возможном их взаимодействии. Подробнее этот вопрос мы обсудим ниже.

Определение. *Плотностью распределения* случайного вектора ξ назовем такую функцию $f_{\xi}(\mathrm{x}) \geqslant 0$, $\mathrm{x} \in \mathbb{R}^n$, что для любого борелевского $A \subset \mathbb{R}^n$

$$P\{\xi \in A\} = \int\limits_A f_{\xi}(\mathrm{x})\, d\mathrm{x} = \iint\limits_A\dots\int f_{\xi}(x_1, \dots , x_n)\, dx_1 \dots dx_n. \tag{6}$$

Отсюда же ясно, что плотность $f_\xi(\mathrm{x})$ чаще всего действительно имеет механический смысл «плотности» — это «количество вероятности», приходящееся на единицу меры (длины, площади, объема, n-мерного объема). В самом деле, в каждой точке непрерывности функции $f_\xi(\mathrm{x})$ получаем, что существует предел

$$\lim_{\mu(A)\to 0} \frac{P\{\xi \in A\}}{\mu(A)} = f_\xi(\mathrm{x}). \tag{7}$$

Ясно, что

$$\int\limits_{\mathbb{R}^n} f_\xi(\mathrm{x})\, d\mathrm{x} = 1$$

и

$$F_\xi(\mathrm{x}) = \int\limits_{-\infty}^{\mathrm{x}} f_\xi(\mathrm{x})\, d\mathrm{x} = \int\limits_{-\infty}^{x_1}\ldots\int\limits_{-\infty}^{x_n} f_\xi(x_1, x_2, \ldots, x_n)\, dx_1 \ldots dx_n. \tag{8}$$

Во всех точках непрерывности плотности имеем дифференциальный аналог соотношения (7):

$$\frac{\partial^n F_\xi(\mathrm{x})}{\partial \mathrm{x}} = \frac{\partial^n F_\xi(\mathrm{x})}{\partial x_1 \partial x_2 \ldots \partial x_n} = f_\xi(x_1, \ldots, x_n). \tag{9}$$

Случайные векторы, обладающие плотностью распределения, называются непрерывными. Всякий подвектор непрерывного случайного вектора является непрерывным случайным вектором, плотность распределения которого можно найти из соотношения

$$f_{\xi_{i_1},\ldots,\xi_{i_k}}(x_{i_1}, \ldots, x_{i_k}) = \int\limits_{-\infty}^{\infty}\ldots\int\limits_{-\infty}^{\infty} f_\xi(x_1, \ldots, x_n)\, dx_{j_1} \ldots dx_{j_{n-k}}. \tag{10}$$

В частности, плотность i-ой компоненты дается формулой

$$f_{\xi_i}(x_i) = \int\limits_{-\infty}^{\infty}\ldots\int\limits_{-\infty}^{\infty} f_\xi(\mathrm{x})\, dx_1 \ldots dx_{i-1} dx_{i+1} \ldots dx_n.$$

Обратное неверно, т. е. непрерывность компонент не гарантирует непрерывности случайного вектора.

Как и в случае с функцией распределения, индивидуальные плотности не позволяют находить совместную плотность распределения случайного вектора, так как в общем случае не содержат информации о возможном взаимодействии компонент.

Если случайный вектор принимает не более чем счетное множество значений, то его закон распределения может быть задан в форме *ряда распределения*, в котором перечислены возможные значения вектора с вероятностями, соответствующими этим значениям

ξ	a_1	a_2	$\ldots$	a_n	$\ldots$
$P\{\xi = \mathrm{a}_i\}$	p_1	p_2	$\ldots$	p_n	$\ldots$

Все точки $\mathrm{a}_j \in \mathbb{R}^n$ в этой таблице различны и $\sum\limits_{j=1}^{\infty} p_j = 1$.

Очевидным образом имеем:

1.
$$\boxed{P\{\xi \in A\} = \sum_{\mathrm{a}_i \in A} P\{\xi = \mathrm{a}_i\}} \qquad (11)$$

для любого борелевского множества $A \subset \mathbb{R}^n$,

2.
$$\boxed{F_\xi(\mathrm{x}) = \sum_{\mathrm{a}_i < \mathrm{x}} P\{\xi = \mathrm{a}_i\}.} \qquad (12)$$

Неравенство $\mathrm{a}_i < \mathrm{x}$ в последнем соотношении понимается как покомнонентное

$$\mathrm{a}_i < \mathrm{x} \Leftrightarrow a_j^i < x_j, \quad j = 1, 2, \ldots, n.$$

Такие случайные векторы называются *дискретными*. Всякий подвектор дискретного вектора также является дискретным случайным вектором. Его ряд распределения задается таблицей

ξ_i	b_1	b_2	$\ldots$	b_s	$\ldots$
$P\{\xi_i = \mathrm{b}_i\}$	q_1	q_2	$\ldots$	q_s	$\ldots$

где $\mathrm{b}_s = \left\{\mathrm{a}_{i_1}^s, \mathrm{a}_{i_2}^s, \ldots, \mathrm{a}_{i_k}^s\right\}$ — подвектор вектора $\mathrm{a}_r = \{a_1, a_2, \ldots, a_n\}$, а вероятности

$$q_s = P\{\xi_p = \mathrm{b}_s\} = \sum_r P\{\xi = \mathrm{a}_r\},$$

причем суммирование в последней сумме ведется по всем значениям a_r вектора ξ, содержащим в качестве подвектора вектор $\mathrm{b}_s = \left\{a_{i_1}^s, a_{i_2}^s, \ldots, a_{i_k}^s\right\}$.

Ряд распределения i-й компоненты представлен в таблице

ξ_i	a_i^1	a_i^2	$\ldots$	a_i^s	$\ldots$
$P\left\{\xi_i = a_i^s\right\}$	q_1	q_2	$\ldots$	q_s	$\ldots$

Здесь a_i^s — значения i-х компонент векторов-значений случайного вектора ξ, т. е. a_s, $q_s = \sum\limits_r P\{\bar{\xi} = \mathrm{a}_r\}$, где суммирование ведется по всем r таким, что векторы a_r имеют одинаковую i-ю компоненту a_i^s.

Этими двумя классами случайных векторов (т. е. непрерывными и дискретными) многообразие возможных систем случайных величин, конечно, не исчерпывается. Отметим, в частности, еще два важных для приложений класса — это, во-первых, «смеси» дискретных и непрерывных случайных векторов, в том числе случайные векторы, у которых часть компонент дискретна, а часть — непрерывна; а во-вторых, абсолютно непрерывные распределения, сосредоточенные на множествах, размерность которых меньше размерности объемлющего пространства. Существуют и другие случайные векторы, на особенностях распределения которых мы останавливаться не будем, ввиду их «экзотичности». В реальной ситуации они, как правило, не встречаются.

Что касается перечисленных, то для единообразного описания их законов распределения воспользуемся, как и выше (п. 2.1.4), символической формой представления обобщенных плотностей, выразив последние через дельта-функцию множества.

Пусть γ — некоторое множество точек $\mathbb{R}^n$. *Дельта-функцией множества* γ назовем функцию $\delta_\gamma(\mathrm{x})$, определенную в $\mathbb{R}^n$ всюду, кроме точек множества γ, обращающуюся в нуль во всех точках $x \notin \gamma$ и обладающую тем свойством, что для любой непрерывной функции $f(x)$

$$\int_{\mathbb{R}^n} f(x)\delta_\gamma(x)\,dx = \int_\gamma f(x_\gamma)\,dx_\gamma, \quad x \in \mathbb{R}^n,\ x \in \gamma. \tag{13}$$

Определение. *Обобщенной плотностью* распределения случайного вектора $\xi = \{\xi_1, \xi_2, \ldots, \xi_n\}$ будем называть формальное выражение вида

$$g_\xi x = f_\xi(x) + \sum_{i=1}^{\infty} p_i\delta_{a_i}(x) + \sum_{j=1}^{\infty} \varphi_{\gamma_j}(x)\delta_{\gamma_j}(x), \tag{14}$$

где

1. функция $f_\xi(x)$ неотрицательна, $f_\xi(x) \geqslant 0$, интегрируема и

$$\int_{\mathbb{R}^n} f_\xi(x)\,dx = q, \quad 0 \leqslant q \leqslant 1,$$

2. $\sum\limits_{i=1}^{\infty} p_i = p$, $0 \leqslant p \leqslant 1$, где $\delta_{a_i}(x)$ — дельта-функция точки a_i,

3. для любого j функции $\varphi_{\gamma_i}(x)$ неотрицательны, интегрируемы на соответствующем множестве $\gamma_j \subset \mathbb{R}^n$ и

$$\int_{\gamma_j} \varphi_{\gamma_j}(x_j)\,dx_j = r_j, \qquad \sum_{j=1}^{\infty} r_j = r, \quad 0 \leqslant r \leqslant 1, \qquad p + q + r = 1.$$

При этом для любого борелевского множества $A \subset \mathbb{R}^n$ получаем, что

$$\boxed{P\{\xi \in A\} = \int_A q_\xi(x)\,dx = \int_A f_\xi(x)\,dx + \sum_{i:\ a_i \in A} p_i + \sum \int_{A\cap\gamma_j} \varphi_{\gamma_j}(x_j)\,dx_j.} \tag{15}$$

Именно в смысле равенства (15) мы и будем в дальнейшем понимать формальное выражение (14) для «обобщенной плотности».

Рассмотрим несколько примеров, иллюстрирующих введенные понятия.

Пример 1 (полиномиальное распределение). Пусть в эксперименте Ω выделена полная система несовместных событий N_j, $j = 1, 2, \ldots, k$, таких, что $P\{n_j\} = p_j$. Случайный вектор $\xi = \{m_1, m_2, \ldots, m_k\}$, j-я компонента которого — количество осуществлений события n_j в серии из n независимых испытаний, носит название полиномиального случайного вектора. Легко убедиться в том, что полиномиальный случайный вектор дискретен. Его ряд распределения дается соотношением

$$P\{\xi_1 = m_1, \xi_2 = m_2, \ldots, \xi_k = m_k\} = \frac{n!}{m_1!m_2!\ldots m_k!}\, p_1^{m_1}p_2^{m_2}\ldots p_k^{m_k}, \tag{16}$$

или, в векторной записи,

$$P\{\xi = m\} = \frac{n!}{m!}\,P^m, \quad \sum p_i = 1, \ \sum m_i = n.$$

Пример 2 (равномерное распределение в круге). Пусть (ξ_1, ξ_2) — координаты случайной точки в круге радиуса 1 с центром в начале координат. «Случайность» точки в круге понимается так, что вероятность ее попадания в любую область внутри круга пропорциональна площади этой области

$$\forall A \in K \quad P\{\xi \in A\} = \frac{S(A)}{\pi}.$$

Заметим, что в каждой точке $\mathbf{x} \in K$ существует предел (см. соотношение (6))

$$\lim_{S(A_{\mathbf{x}})\to 0} \frac{P\{\boldsymbol{\xi} \in A_{\mathbf{x}}\}}{S(A_{\mathbf{x}})} = \frac{1}{\pi},$$

где $A_{\mathbf{x}}$ — окрестность точки $\mathbf{x}$. Поэтому определена функция

$$f(\mathbf{x}) = \begin{cases} \dfrac{1}{\pi}, & \mathbf{x} \in K, \\ 0, & \mathbf{x} \notin K, \end{cases}$$

являющаяся плотностью распределения случайного вектора $\boldsymbol{\xi}$, так что для любого борелевского $B \in \mathbb{R}^2$

$$P\{\boldsymbol{\xi} \in B\} = \int\limits_B f(\mathbf{x})\, d\mathbf{x}.$$

Тем самым, случайный вектор $\boldsymbol{\xi}$ — непрерывен.

Пример 3. Пусть $\boldsymbol{\xi} = \{\xi_1, \xi_2\}$ — случайный вектор единичной длины. В данном случае случайность понимается так: вероятность того, что напрввление $\boldsymbol{\xi}$ на плоскости задается диапазоном углов от φ_0 до φ_1, пропорциональна этому диапазону:

$$P\{\varphi_0 \leqslant \varphi \leqslant \varphi_1\} = \frac{\varphi_1 - \varphi_0}{2\pi}; \quad 0 < \varphi_0 < \varphi_1 < 2\pi.$$

Множество значений вектора $\boldsymbol{\xi}$ совпадает с множеством точек единичной окружности $x^2 + y^2 = 1$. Следовательно $\boldsymbol{\xi}$ не является дискретным. Не является он и непрерывным, так как легко установить, что для всех точек, лежащих на единичной окружности, предел

$$\lim_{s\to 0} \frac{P\{\boldsymbol{\xi} \in A\}}{S(A)}$$

не существует, а во всех прочих точках плоскости равен нулю.

Эта случайная величина может быть описана обобщенной плотностью $g_\xi(x, y)$, задаваемой соотношением

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

где $\delta_{x^2+y^2=1}(x, y)$ — δ-функция окружности.

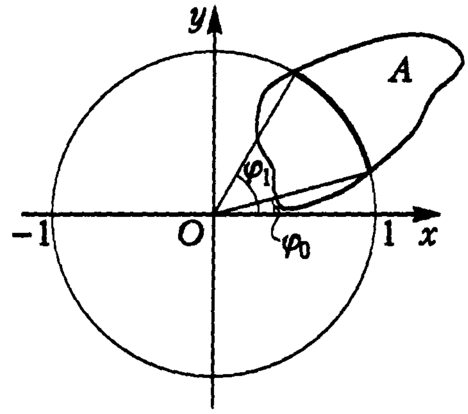


Рис. 15

Действительно, для любого борелевского $A \subset \mathbb{R}^2$ получаем

$$P\{\boldsymbol{\xi} \in A\} = \int\limits_A g_\xi(x, y)\, dx\, dy = \int\limits_{A\cap(x^2+y^2=1)} \frac{1}{2\pi}\, d\varphi = \frac{\varphi_1 - \varphi_0}{2\pi}$$

(см. рис. 15).

2.2. Независимость случайных величин. Условные законы распределения

Как уже было отмечено выше, знание индивидуальных законов распределения компонент случайного вектора не дает полной информации о совместном их законе распределения. Дело в том, что индивидуальные законы не позволяют учесть возможное взаимодействие компонент.

Пример 1. Пусть ξ_1 — равномерная на $[0, 1]$ случайная величина. Тогда $\xi_2 = \xi_1$ — также равномерна на $[0, 1]$, а совместное распределение (ξ_1, ξ_2) сосредоточено на отрезке

$$S = \{(x, y)\colon\ x = y;\ 0 \leqslant x \leqslant 1\}$$

и может быть описано обобщенной плотностью вида

$$g_{\xi_1\xi_2}(x, y) = \begin{cases} 0, & (x, y) \notin S, \\ \dfrac{1}{\sqrt{2}}\, \delta_s(x, y), & (x, y) \in S. \end{cases} \tag{17}$$

В то же время двумерная случайная величина $\boldsymbol{\xi} = \{\xi_1, \xi_2\}$, равномерно распределенная на квадрате $K = \{(x, y)\colon\ 0 \leqslant x,\ y \leqslant 1\}$ описывается плотностью

$$g_{\xi_1\xi_2}(x, y) = \begin{cases} 0, & (x, y) \notin K, \\ 1, & (x, y) \in K, \end{cases} \tag{18}$$

и ее компоненты, как следует из соотношения (10)

$$g_{\xi_1}(x) = \int_{-\infty}^{\infty} g_{\xi_1\xi_2}(x, y)\,dy = \int_{-\infty}^{0} g_{\xi_1\xi_2}(x, y)\,dy +$$

$$+ \int_{0}^{1} g_{\xi_1\xi_2}(x, y)\,dy + \int_{1}^{\infty} g_{\xi_1\xi_2}(x, y)\,dy = \begin{cases} 0, & x < 0 \cup x > 1 \\ 1, & 0 \leqslant x \leqslant 1, \end{cases}$$

также равномерны на $[0, 1]$. Таким образом, при одном и том же характере распределения компонент совместные распределения (17) и (18) совершенно различны.

Одно из центральных понятий теории вероятностей — независимой случайной величины — проливает свет на описанную выше ситуацию и позволяет внести ясность в рассматриваемый вопрос.

Определение. Случайные величины ξ_1 и ξ_2 называются ***независимыми***, если для любых борелевских множеств B_1 и B_2 на прямой случайные события $A_1 = \{\omega\colon\ \xi_1 \in B_1\}$ и $A_2 = \{\omega\colon\ \xi_2 \in B_2\}$ независимы.

Как мы знаем (см. гл. XXXVIII), для независимости событий A_1 и A_2 необходимо и достаточно, чтобы выполнялось соотношение

$$P\{\xi_1 \in B_1;\ \xi_2 \in B_2\} = P\{\xi_1 \in B_1\}P\{\xi_2 \in B_2\} \tag{19}$$

для любых B_1, B_2, а это, в свою очередь, имеет место тогда и только тогда, когда для любых полуинтервалов $[\alpha_1, \beta_1)$ и $[\alpha_2, \beta_2)$ справедлив аналог соотношения (19),

$$P\{\alpha_1 \leqslant \xi_1 < \beta_1;\ \alpha_2 \leqslant \xi_2 < \beta_2\} = P\{\alpha_1 \leqslant \xi_1 < \beta_1\}P\{\alpha_2 \leqslant \xi_2 < \beta_2\}. \tag{20}$$

Соотношение (19) показывает, что в случае независимости двух случайных величин совместный закон их распределения полностью и однозначно восстанавливается по индивидуальным законам распределения компонент. В частности, легко устанавливается следующая теорема.

Теорема. *Случайные величины ξ_1 и ξ_2 независимы тогда и только тогда, когда имеет место равенство*

$$F_{\xi_1\xi_2}(x, y) = F_{\xi_1}(x)F_{\xi_2}(y), \tag{21}$$

т. е. совместная функция распределения равна произведению индивидуальных.

◀ Необходимость немедленно следует из (19), если только $B_1 = \{-\infty, x\}$ и $B_2 = \{-\infty, y\}$.

Для доказательства достаточности воспользуемся формулой (4). Пусть равенство (21) имеет место. Тогда

$$P\{\alpha_1 \leqslant \xi_1 < \beta_1;\ \alpha_2 \leqslant \xi_2 < \beta_2\} =$$

$$= F_{\xi_1\xi_2}(\beta_1, \beta_2) - F_{\xi_1\xi_2}(\alpha_1, \beta_2) - F_{\xi_1\xi_2}(\beta_1\alpha_2) + F_{\xi_1\xi_2}(\alpha_1, \alpha_2) =$$

$$F_{\xi_1}(\beta_1)F_{\xi_2}(\beta_2) - F_{\xi_1}(\alpha_1)F_{\xi_2}(\beta_2) - F_{\xi_1}(\beta_1)F_{\xi_2}(\alpha_2) + F_{\xi_1}(\alpha_1)F_{\xi_2}(\alpha_2) =$$

$$= \left[F_{\xi_1}(\beta_1) - F_{\xi_1}(\alpha_1)\right]\ \left[F_{\xi_2}(\beta_2) - F_{\xi_2}(\alpha_2)\right] = P\{\alpha_1 \leqslant \xi_1 < \beta_1\} \cdot P\{\alpha_2 \leqslant \xi_2 < \beta_2\},$$

откуда в силу сделанного выше замечания о равенствах (19) и (20) следует искомое. ▶

Аналогичные утверждения можно установить и для специальных классов распределений: дискретных, описываемых рядами распределения, непрерывных, описываемых плотностями, а также более сложных, описываемых обобщенными плотностями. А именно, имеют место следующие утверждения.

Утверждение 1. *Если ξ_1 и ξ_2 — дискретные независимые случайные величины с рядами распределения соответственно $P\{\xi_1 = a_i\} = p_i$, $i = 1, \ldots, n, \ldots$ и $P\{\xi_2 = b_j\} = q_j$, $j = 1, 2, \ldots, m, \ldots$, то двумерный случайный вектор $\boldsymbol{\xi} = \{\xi_1, \xi_2\}$ дискретен и его ряд распределения дается соотношением*

$$P\{\xi_1 = a_i;\ \xi_2 = b_j\} = p_i q_j.$$

Верно и обратное: если компоненты двумерного дискретного вектора таковы, что

$$P\{\xi_1 = a_i;\ \xi_2 = b_j\} = P\{\xi_1 = a_i\}P\{\xi_2 = a_j\},$$

то они независимы.

◀ Доказательство немедленно следует из определения независимости и соотношения (20). ▶

Утверждение 2. ***Если случайные величины ξ_1 и ξ_2 непрерывны и независимы, то двумерный случайный вектор $\bar{\xi} = \{\xi_1, \xi_2\}$ непрерывен и его плотность равна произведению индивидуальных плотностей***

$$f_{\xi_1\xi_2}(x_1, x_2) = f_{\xi_1}(x_1)f_{\xi_2}(x_2). \tag{22}$$

Верно и обратное: если $\forall\ x_1, x_2$ совместная плотность распределения представима в виде (22), то компоненты вектора $\bar{\xi}$ — независимые случайные величины.

Замечание. Если компоненты случайного вектора непрерывны, то это, как показывает рассмотренный выше пример, не гарантирует непрерывности собственно вектора как двумерной случайной величины. Поэтому утверждение о непрерывности вектора $\{\xi_1, \xi_2\}$ является не менее важным, чем формула (22). Оба же эти утверждения оказываются справедливыми в силу независимости случайных величин ξ_1 и ξ_2.

◀ Пусть $f_{\xi_1}(x_1)$ — плотность распределения случайной величины ξ_1, а $f_{\xi_2}(x_2)$ — случайной величины ξ_2. Из независимости ξ_1 и ξ_2 заключаем, что $\forall\ a_1 \leqslant b_1,\ a_2 \leqslant b_2$

$$P\{a_1 \leqslant \xi_1 < b_1;\ a_2 \leqslant \xi_2 < b_2\} = P\{a_1 \leqslant \xi_1 < b_1\} \cdot P\{a_2 \leqslant \xi_2 < b_2\} =$$

$$= \int_{a_1}^{b_1} f_{\xi_1}(x_1)\,dx_1 \cdot \int_{a_2}^{b_2} f_{\xi_2}(x_2)\,dx_2 = \int_{a_1}^{b_1}\int_{a_2}^{b_2} f_{\xi_1}(x_1)\,f_{\xi_2}(x_2)\,dx_1\,dx_2.$$

Значит, для любого прямоугольника $\Pi = \{a_1 \leqslant x_1 < b_1;\ a_2 \leqslant x_2 < b_2\}$ функция $f_{\xi_1\xi_2}(x_1, x_2) = f_{\xi_1}(x_1)f_{\xi_2}(x_2)$ такова, что

$$P\{\boldsymbol{\xi} \in \Pi\} = \iint_{\Pi} f_{\xi_1\xi_2}(x_1, x_2)\,dx_1\,dx_2.$$

Отсюда следует справедливость аналогичного соотношения для любого борелевского A, что и завершает доказательство необходимости соотношения (22).

Так же просто убеждаемся в обратном

$$P\{\boldsymbol{\xi} \in \Pi\} = \iint_{\Pi} f_{\xi_1\xi_2}(x_1, x_2)\,dx_1\,dx_2 = \iint_{\Pi} f_{\xi_1}(x_1)f_{\xi_2}(x_2)\,dx_1 = \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} f_{\xi_1}(x_1)f_{\xi_2}(x_2)\,dx_2 =$$

$$= \int_{a_1}^{b_1} f_{\xi_1}(x_1)\,dx_1 \int_{a_2}^{b_2} f_{\xi_2}(x_2)\,dx_2 = P\{a_1 \leqslant \xi_1 < b_1\}P\{a_2 \leqslant \xi_2 < b_2\}. \blacktriangleright$$

Утверждение 3. *Если случайные величины ξ_1 и ξ_2 — независимы и обладают обобщенными плотностями $g_{\xi_1}(x_1)$ и $g_{\xi_2}(x_2)$ (см. соотношения (18)), то двумерный случайный вектор $\boldsymbol{\xi} = \{\xi_1, \xi_2\}$ обладает обобщенной плотностью (14) и имеет место равенство*

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

◀ Доказательство повторяет формальные выкладки из предыдущего утверждения с учетом следующих технических замечаний:

$$\delta(x_1 - a_i)\delta(x_2 - b_j) = \delta_{(a_i b_j)}(x_1, x_2),$$
$$\delta(x_1 - a_i)f_{\xi_2}(x_2) = f_{\xi_2}(x_2)\delta_{x_1=a_i}(x_1, x_2),$$
$$\delta(x_2 - b_j)f_{\xi_1}(x_1) = f_{\xi_1}(x_1)\delta_{x_2=b_j}(x_1, x_2). \blacktriangleright$$

Суммируя вышеизложенное, отметим, что независимость двух случайных величин содержательно эквивалентна отсутствию влияния одной из случайных величин на другую: изменение одной из них не влияет на закон распределения другой. Если же это не так, т. е. информация о значениях, принятых одной из случайных величин, меняет наши представления о законе распределения другой случайной величины, то мы говорим, что случайные величины зависимы. В этом случае разумно описывать влияние случайных величин друг на друга системой условных вероятностей, образующих так называемые *условные распределения*, к рассмотрению которых мы и перейдем.

Рассмотрим пару случайных величин (ξ_1, ξ_2), совместный закон распределения которых

$$P\{\xi_1 \in A_1,\ \xi_2 \in A_2\} = \Phi_{\xi_1\xi_2}(A_1, A_2).$$

Пусть $P\{\xi_2 \in A_2\} \neq 0$.

Определение. *Условным законом распределения* $\Phi_{\xi_1|\xi_2}(A_1 \,|\, A_2)$ *компоненты (случайной величины)* ξ_1 *относительно события* $\xi_2 \in A_2$ назовем любое правило, позволяющее вычислять условные вероятности $P\{\xi_1 \in A_1 \,|\, \xi_2 \in A_2\}$ для произвольных борелевских A_1.

По определению условной вероятности получаем, что

$$\boxed{\begin{aligned}\Phi_{\xi_1|\xi_2}(A_1 \,|\, A_2) &= P\{\xi_1 \in A_1 \,|\, \xi_2 \in A_2\} = \\ &= \frac{P\{\xi_1 \in A_1;\ \xi_2 \in A_2\}}{P\{\xi_2 \in A_2\}} = \frac{\Phi_{\xi_1\xi_2}(A_1, A_2)}{\Phi_{\xi_2}(A_2)}.\end{aligned}} \qquad (23)$$

В частности, если $P\{\xi_2 < x_2^\circ\} > 0$ для некоторого x_2°, то можно определить *условную функцию распределения* ξ_1 относительно ξ_2:

$$\boxed{F_{\xi_1|\xi_2}(x_1 \,|\, x_2) = \frac{P\{\xi_1 < x_1;\ \xi_2 < x_2\}}{P\{\xi_2 < x_2\}} = \frac{F_{\xi_1\xi_2}(x_1, x_2)}{F_{\xi_2}(x_2)};} \qquad (24)$$

она будет определена при всех $x_1 \in \mathbb{R}$ и $x_2 > x_2^\circ$.

Если ξ_1 и ξ_2 — независимы, то

$$\boxed{F_{\xi_1|\xi_2}(x_1 \,|\, x_2) = F_{\xi_1}(x_1).}$$

Из соотношения (23) получаем аналог «правила умножения» (21) функций распределения для зависимых случайных величин

$$\boxed{F_{\xi_1\xi_2}(x_1, x_2) = F_{\xi_1|\xi_2}(x_1 \,|\, x_2)F_{\xi_2}(x_2).} \qquad (25)$$

Аналогично определяются условные законы для ξ_2 относительно ξ_1.

Пример 2. Пусть ξ_1 — равномерная на $[-1, 1]$ случайная величина и $\xi_2 = \xi_1$. Индивидуальные функции распределения компонент даются в этом случае соотношениями:

$$F_{\xi_1}(x_1) = \begin{cases} 0, & x_1 < -1, \\ \dfrac{x_1+1}{2}, & -1 \leqslant x_1 < 1, \\ 1, & x_1 \geqslant 1; \end{cases}$$

$$F_{\xi_2}(x_2) = \begin{cases} 0, & x_2 < -1, \\ \dfrac{x_2+1}{2}, & -1 \leqslant x_2 < 1, \\ 1, & x_2 \geqslant 1. \end{cases}$$

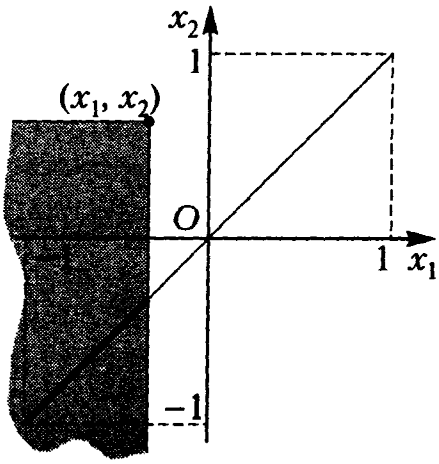


Рис. 16

Совместный закон распределения может быть описан обобщенной плотностью $g_{\xi_1\xi_2}(x_1, x_2) = (1/4)\sqrt{2} \cdot \delta_\gamma(x_1, x_2)$, где $\gamma = \{(x_1, x_2)\colon\ -1 \leqslant x_1 = x_2 \leqslant 1\}$. Распределение вектора $\boldsymbol{\xi}^\circ$ с компонентами (ξ_1, ξ_2) сосредоточено на отрезке прямой $x_1 = x_2$ (рис. 16). Функция распределения вектора $\boldsymbol{\xi}$

$$F_{\xi_1\xi_2}(x_1, x_2) = \begin{cases} 0, & (x_1 < -1) \cup (x_2 < -1), \\ \dfrac{\min(x_1, x_2)+1}{2}, & (|x_1| \leqslant 1) \cup (|x_2| \leqslant 1), \\ 1, & (x_1 > 1) \cap (x_2 > 1). \end{cases}$$

Условные законы имеют, очевидно, следующий вид (рис. 17):

$$F_{\xi_2|\xi_1}(x_2|x_1) = \begin{cases} 0, & x_2 < -1, \\ \dfrac{x_2+1}{x_1+1}, & x_2 < x_1, \\ 1, & x_2 > x_1; \end{cases}$$

$$F_{\xi_1|\xi_2}(x_1|x_2) = \begin{cases} 0, & x_1 < -1, \\ \dfrac{x_1+1}{x_2+1}, & x_1 < x_2, \\ 1, & x_1 > x_2. \end{cases}$$

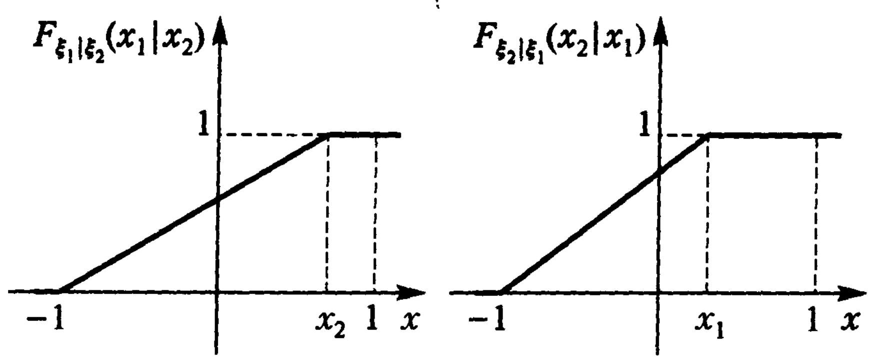


Рис. 17

Отметим, что несмотря на непрерывность компонент ξ_1 и ξ_2, вектор $\boldsymbol{\xi} = \{\xi_1, \xi_2\}$ непрерывным не является, так как непрерывность компонент случайного вектора — условие, лишь необходимое для непрерывности вектора; оно становится и достаточным только при условии независимости компонент.

Предположим теперь, что случайный вектор (ξ_1, ξ_2) непрерывен, т. е. обладает плотностью распределения $f_{\xi_1\xi_2}(x_1, x_2)$. В этом случае, как мы знаем (см. (10)), компоненты ξ_1 и ξ_2 также непрерывны с плотностями

$$f_{\xi_i}(x_i) = \int\limits_{-\infty}^{\infty} f_{\xi_1\xi_2}(x_1, x_2)\, dx_j, \quad i \neq j, \quad i, j = 1, 2.$$

Если в некоторой окрестности $S_{x_2}^{\Delta x_2}$ точки x_2 выполняется неравенство $f_{\xi_2}(x_2) > 0$, то можно определить условный закон распределения случайной величины ξ_1 относительно события $\{\xi_2 = x_2\}$:

$$\Phi_{\xi_1|\xi_2}(A_1|x_2) = P\{\xi_1 \in A_1|\xi_2 = x_2\} = \lim_{\Delta x_2 \to 0} P\{\xi_1 \in A_1|\xi_2 \in S_{x_2}^{\Delta x_2}\} =$$

$$= \lim_{\Delta x_2 \to 0} \frac{P\{\xi_1 \in A_1,\ \xi_2 \in S_{x_2}^{\Delta x_2}\}}{P\{\xi_2 \in S_{x_2}^{\Delta x_2}\}} = \lim_{\Delta x_2 \to 0} \frac{\int\limits_{A_1} dx_1 \int\limits_{x_2-\Delta x}^{x_2+\Delta x} f_{\xi_1\xi_2}(x_1, x_2)\, dx_2}{\int\limits_{x_2-\Delta x}^{x_2+\Delta x} f_{\xi_2}(x_2)\, dx_2} =$$

$$= \frac{\int\limits_{A_1} f_{\xi_1\xi_2}(x_1, x_2)\, dx_1}{f_{\xi_2}(x_2)} = \int\limits_{A_1} \frac{f_{\xi_1\xi_2}(x_1, x_2)}{f_{\xi_2}(x_2)}\, dx_1. \tag{26}$$

Последнее соотношение дает основание для следующего определения.

Определение. *Условной плотностью распределения вероятностей случайной величины* ξ_1 *относительно события* $\{\xi_2 = x_2\}$ называется функция $f_{\xi_1|\xi_2}(x_1|x_2)$ такая, что для любого борелевского A_1 выполняется соотношение

$$P\{\xi_1 \in A_1|\xi_2 \in x_2\} = \int\limits_{A_1} f_{\xi_1|\xi_2}(x_1|x_2)\, dx_1. \tag{27}$$

Сравнивая (26) и (27), заключаем, что у каждой компоненты двумерного непрерывного вектора существует условная плотность относительно значений другой компоненты, принимаемой с положительной вероятностью (например, относительно тех значений $\xi_2 = x_2$, для которых $P\{\xi_2 \in S_{x_2}^{\Delta x_2}\} > 0$). Она дается равенством

$$\boxed{f_{\xi_1|\xi_2}(x_1|x_2) = \frac{f_{\xi_1\xi_2}(x_1, x_2)}{f_{\xi_2}(x_2)}.} \tag{28}$$

Отсюда получаем правило умножения плотностей

$$\boxed{f_{\xi_1\xi_2}(x_1, x_2) = f_{\xi_1|\xi_2}(x_1|x_2) f_{\xi_2}(x_2) = f_{\xi_2|\xi_1}(x_2|x_1) f_{\xi_1}(x_1).} \tag{29}$$

В случае независимости компонент ξ_1 и ξ_2 из (28) следует

$$\boxed{f_{\xi_1|\xi_2}(x_1|x_2) = f_{\xi_1}(x_1),}$$

т. е. информация о значениях, принятых компонентой ξ_2, не меняет закона распределения компоненты ξ_1.

Определение (27) и соотношение (28) позволяют в рассматриваемой ситуации получить континуальные аналоги формул полной вероятности и Бейеса (см. гл. XXXVIII).

Теорема (формула полной вероятности). *Пусть* (ξ_1, ξ_2) *— непрерывный двумерный случайный вектор. Тогда*

$$f_{\xi_2}(x_2) = \int\limits_{-\infty}^{\infty} f_{\xi_2|\xi_1}(x_2|x_1) f_{\xi_1}(x_1)\, dx_1. \tag{30}$$

◀ Действительно, соотношение (10) дает

$$f_{\xi_2}(x_2) = \int\limits_{-\infty}^{\infty} f_{\xi_1\xi_2}(x_1, x_2)\, dx_1.$$

Остается только воспользоваться второй частью правила умножения (29). ▶

Теорема (формула Бейеса). *Пусть (ξ_1, ξ_2) — непрерывный двумерный случайный вектор. Тогда*

$$f_{\xi_1|\xi_2}(x_1|x_2) = \frac{f_{\xi_2|\xi_1}(x_2|x_1) f_{\xi_1}(x_1)}{\int\limits_{-\infty}^{\infty} f_{\xi_2\xi_1}(x_2, x_1) f_{\xi_1}(x_1)\, dx_1}. \tag{31}$$

◀ Доказательство следует из (28) с использованием правила умножения (29) и формулы полной вероятности (30). ▶

Рассмотрим пример, иллюстрирующий введенные понятия.

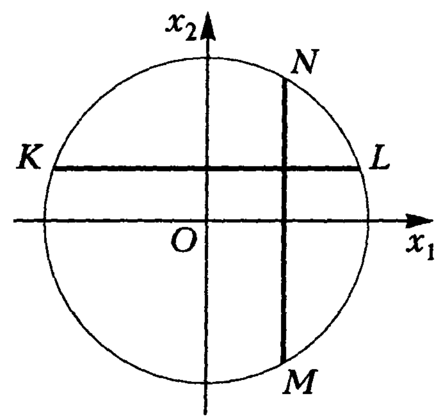


Рис. 18

Пример 3. Пусть $\boldsymbol{\xi} = \{\xi_1, \xi_2\}$ — двумерный вектор, равномерно распределенный в круге $K = \left\{(x_1, x_2)\colon\ x_1^2 + x_2^2 \leqslant 1\right\}$. Для любого борелевского $B \subset \mathbb{R}^2$

$$P\{\boldsymbol{\xi} \in B\} = \frac{S(B)}{S(K)} = \frac{S(B)}{\pi}.$$

Следовательно, плотность распределения случайной величины $\boldsymbol{\xi}$ дается соотношением:

$$f_{\xi_1\xi_2}(x_1, x_2) = \begin{cases} 0, & (x_1, x_2) \notin K, \\ \dfrac{1}{\pi}, & (x_1, x_2) \subset K. \end{cases}$$

Плотности компонент соответственно равны (рис. 18)

$$f_{\xi_1}(x_1) = \int\limits_{-\infty}^{\infty} f_{\xi_1\xi_2}(x_1, x_2)\, dx_2 = \int\limits_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \frac{1}{\pi}\, dx_2 = \begin{cases} \dfrac{2}{\pi}\sqrt{1-x_1^2}, & |x_1| < 1, \\ 0, & |x_1| > 1, \end{cases}$$

$$f_{\xi_2}(x_2) = \int\limits_{-\infty}^{\infty} f_{\xi_1\xi_2}(x_1, x_2)\, dx_1 = \int\limits_{-\sqrt{1-x_2^2}}^{\sqrt{1-x_2^2}} \frac{1}{\pi}\, dx_2 = \begin{cases} \dfrac{2}{\pi}\sqrt{1-x_2^2}, & |x_2| < 1, \\ 0, & |x_2| > 1. \end{cases}$$

Для условных плотностей из (28) получаем

$$f_{\xi_2|\xi_1}(x_2|x_1) = \begin{cases} \dfrac{1/\pi}{(2/\pi)\sqrt{1-x_1^2}} = \dfrac{1}{l_{x_1}}, & |x_1| < 1, \\ 0, & |x_1| > 1, \end{cases}$$

где $l_{x_1} = 2\sqrt{1-x_1^2}$ — длина хорды MN,

$$f_{\xi_1|\xi_2}(x_1|x_2) = \begin{cases} \dfrac{1/\pi}{(2/\pi)\sqrt{1-x_2^2}} = \dfrac{1}{l_{x_2}}, & |x_2| < 1, \\ 0, & |x_2| > 1, \end{cases}$$

где $l_{x_2} = 2\sqrt{1-x_2^2}$ — длина хорды KL.

Заметим, что при фиксированном значении x_1 компонента ξ_2 равномерно распределена на хорде MN; аналогично ξ_1 равномерно распределена на KL при фиксированном x_2. В то же время, частные распределения равномерными не являются.

Очевидна зависимость компонент ξ_1 и ξ_2 — изменение значений одной из компонент вызывает изменение диапазона возможных значений другой и, следовательно, ее закона распределения: если ξ_1 принимает значение, скажем, 1, то ξ_2 может принять только нулевое значение, если же $\xi_1 = 0$, то ξ_2 может принять любое значение в диапазоне от -1 до $+1$.

Наконец, если $\boldsymbol{\xi} = \{\xi_1, \xi_2\}$ — дискретный вектор с рядом распределения

$\boldsymbol{\xi}$	a_1	a_2	$\ldots$	a_k	$\ldots$
$P\{\boldsymbol{\xi} = \mathrm{a}_i\}$	p_1	p_2	$\ldots$	p_k	

и $\mathrm{a}_k = \left\{a_k^1, a_k^2\right\}$, то компоненты ξ_i также дискретны с рядом распределения

ξ_i	α_1^i	α_2^i	$\ldots$	α_k^i	$\ldots$
$P\left\{\xi \neq \alpha_j^i\right\}$	q_1^i	q_2^i	$\ldots$	q_k^i	

и $i = 1, 2$; α_j^1, α_j^2 — возможные значения соответствующей компоненты (т. е. просто различные первые — a_k^1 — или вторые — a_k^2 — координаты векторов a_k). При этом

$$q_k^1 = \sum P_s; \quad q_k^2 = \sum P_l,$$

где суммирование ведется по всем таким точкам $\mathrm{a}_s(\mathrm{a}_l)$, первые (вторые) координаты которых совпадают с $\alpha_k^1\left(\alpha_k^2\right)$.

Если $P\left\{\xi_2 = \alpha_k^2\right\} > 0$, то условный ряд распределения компоненты ξ_1 относительно события $\left\{\xi_2 = \alpha_k^2\right\}$ можно определить так

$$P\left\{\xi_1 = \alpha_s^1 / \xi_2 = \alpha_k^2\right\} = \frac{P\left\{\xi_1 = \alpha_s^1;\ \xi_2 = \alpha_k^2\right\}}{P\left\{\xi_2 = \alpha_k^2\right\}}. \tag{32}$$

Как и выше, из (32) можно получить правило умножения рядов распределения для зависимых компонент, а также дискретные аналоги соотношений (30) и (31), являющиеся обобщением формул полной вероятности и Бейеса на случай не более чем счетного числа гипотез.

Глава XL

ФУНКЦИИ СЛУЧАЙНЫХ ВЕЛИЧИН

Задача установления закона распределения функции от случайных величин по заданному закону распределения аргументов является основной.

Общая схема рассуждений здесь следующая. Пусть $\eta = \varphi(\bar{\xi})$, и $\Phi_{\bar{\xi}}(B)$ — закон распределения $\bar{\xi}$. Тогда очевидно имеем

$$\Phi_\eta\{[a, b)\} = P\{\eta \in [a, b)\} = P\{\varphi(\bar{\xi}) \in [a, b)\} = P\left\{\bar{\xi} \in \varphi^{-1}([a, b))\right\}, \qquad (*)$$

где $\varphi^{-1}\{[a, b)\}$ — полный прообраз полуинтервала $[a, b)$ в $\mathbb{R}^n$, т. е. совокупность тех значений вектора $\bar{\xi}$ из $\mathbb{R}^n$, для которых $a \leqslant \varphi(\bar{\xi}) < b$. Последняя вероятность легко может быть найдена, так как закон распределения случайных величин $\bar{\xi}$ известен

$$P\left\{\bar{\xi} \in \varphi^{-1}([a, b))\right\} = \Phi_{\xi}\left\{\varphi^{-1}([a, b))\right\}. \qquad (**)$$

Аналогично, в принципе, может быть найден закон распределения и векторной функции случайных аргументов.

Сложность реализации схемы (*)–(**) зависит только от конкретного вида функции φ и закона распределения аргументов.

Настоящая глава посвящена реализации схемы (*)–(**) в конкретных, важных для приложений, ситуациях.

§ 1. Функции одного переменного

Пусть ξ — случайная величина, закон распределения которой задан функцией распределения $F_\xi(x)$, $\eta = \varphi(\xi)$. Если $F_\eta(y)$ функция распределения случайной величины η, то приведенные выше соображения дают

$$F_\eta(y) = P\{\eta < y\} = P\{\varphi(\xi) < y\} = \\ = P\left\{\xi \in \varphi^{-1}(-\infty, y)\right\}, \qquad (1)$$

где через $\varphi^{-1}(-\infty, y)$ обозначен полный прообраз полупрямой $(-\infty, y)$. Соотношение (1) является очевидным следствием (*) и для рассматриваемого случая проиллюстрировано на рис. 1.

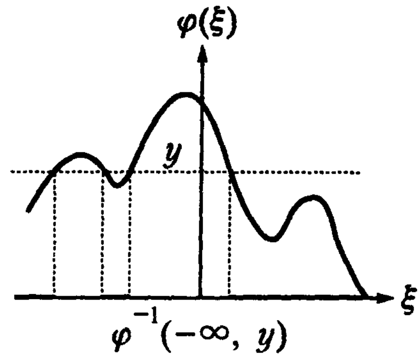


Рис. 1

Монотонное преобразование случайной величины

Пусть $\varphi(t)$ — непрерывная монотонная функция (для определенности — монотонно невозрастающая) и $\eta = \varphi(\xi)$. Для функции распределения $F_\eta(y)$ получаем

$$F_\eta(y) = P\{\eta < y\} = P\{\varphi(\xi) < y\} = P\left\{\xi > \varphi^{-1}(y)\right\} = 1 - F_\xi\left(\varphi^{-1}(y)\right) \qquad (2)$$

(здесь $\varphi^{-1}(y)$ — функция, обратная к $\varphi(t)$, существование которой обеспечивается монотонностью и непрерывность $\varphi(t)$). Для монотонно неубывающей $\varphi(t)$ аналогичные выкладки дают

$$F_\eta(y) = F_\xi\left(\varphi^{-1}(y)\right). \tag{3}$$

В частности, если φ — линейна, $\varphi(t) = at + b$, то при $a > 0$ (рис. 2)

$$F_\eta(y) = F_{a\xi+b}(y) = F_\xi\left(\frac{y-b}{a}\right), \tag{4}$$

а при $a < 0$

$$F_\eta(y) = 1 - F_\xi\left(\frac{y-b}{a}\right). \tag{5}$$

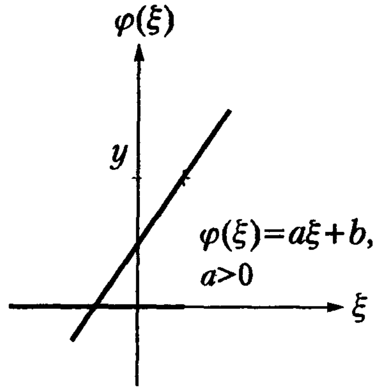


Рис. 2

Линейные преобразования не меняют характера распределения, а сказываются лишь на его параметрах.

Линейное преобразование равномерной на $[a, b]$ случайной величины

Пусть $\xi = R[a, b]$. Тогда $\eta = (\xi - a)/(b - a) = R[0, 1]$.

Линейное преобразование нормальной (m, σ) случайной величины

Пусть $\xi = N[m, \sigma]$. Тогда $\eta = (\xi - m)/\sigma = N[0, 1]$, и вообще, если $\eta = a\xi + b$, то $\eta = N[am + b, |a| \cdot \sigma]$.

◀ Пусть, например, $a > 0$. Из (4) заключаем, что

$$F_\eta(y) = P\{\eta < y\} = F_\xi\left(\frac{y-b}{a}\right) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{(y-b)/a} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} dx.$$

Положим в последнем интеграле $u = ax + b$. Эта замена дает

$$F_\eta(y) = \frac{1}{a\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{y} \exp\left\{-\frac{(u - am - b)^2}{2a^2\sigma^2}\right\} du = N[am + b, a\sigma]. ▶$$

Важное тождество, являющееся источником многих интересных приложений, может быть получено из соотношения (3) при $\varphi(t) = F_\xi(t)$.

Лемма. *Если ξ — случайная величина с непрерывной функцией распределения $F_\xi(x)$, то случайная величина $\eta = F_\xi(\xi)$ — равномерна на* $[0, 1]$.

◀ Имеем

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

$F_\xi(x)$ — монотонно не убывает и заключена в пределах от 0 до 1. Поэтому

$$P\{F_\xi(\xi) < y\} = \begin{cases} 0, & y < 0, \\ 1, & y \geqslant 1. \end{cases}$$

На промежутке же $0 \leqslant y < 1$ получаем

$$P\{\eta < y\} = P\{F_\xi(\xi) < y\} = P\left\{\xi < F_\xi^{-1}(y)\right\} = F_\xi\left(F_\xi^{-1}(y)\right) = y. ▶$$

Одним из возможных путей использования доказанной леммы является, например, процедура моделирования случайной величины с произвольным законом распределения $F_\xi(x)$. Как следует из леммы, для этого достаточно уметь получать значения равномерной на $[0, 1]$ случайной величины, тогда значения ξ могут быть получены из тождества

$$\xi = F_\xi^{-1}(\eta), \qquad \eta = R[0, 1]. \tag{6}$$

В заключение заметим, что если случайная величина ξ непрерывна и функция $\varphi(t)$ не только монотонна, но и дифференцируема, то η также непрерывна. При этом плотность случайной величины η легко может быть получена из (2) или (3):

$$f_\eta(y) = \frac{d}{dy}\, F_\eta(y) = \frac{d}{dy}\, F_\xi\left(\varphi^{-1}(y)\right) = f_\xi\left(\varphi^{-1}(y)\right)\left[\varphi^{-1}(y)\right]'. \tag{7}$$

Если $\varphi(t)$ свойством монотонности не обладает, то результат может быть получен скрупулезным следованием логике соотношения (1), как показывают приводимые ниже примеры.

Распределение квадрата равномерной на $[-1, 1]$ случайной величины

Пусть $\xi = R[-1, 1]$ и $\eta = \xi^2$. Рассмотрим (рис. 3) $0 \leqslant y \leqslant 1$:

$$F_\eta(y) = P\{\eta < y\} = P\left\{\xi^2 < y\right\} = P\{|\xi| < \sqrt{y}\} = \begin{cases} 0, & y < 0, \\ \sqrt{y}, & y \leqslant 1, \\ 1, & y > 1. \end{cases}$$

Отсюда для плотности $f_\eta(y)$ получим

$$f_\eta(y) = \begin{cases} 0, & y < 0 \cup y > 1, \\ \dfrac{1}{2\sqrt{y}}, & 0 \leqslant y \leqslant 1. \end{cases}$$

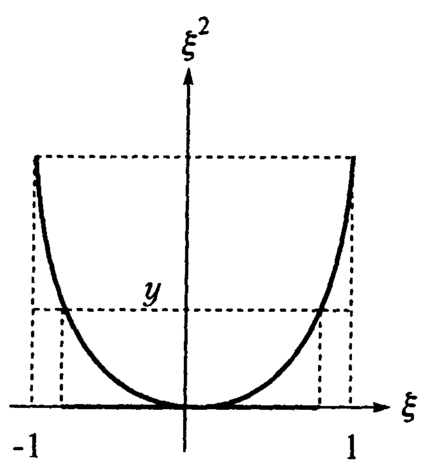


Рис. 3

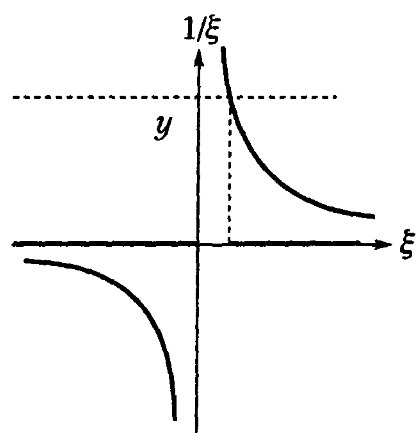


Рис. 4

Распределение случайной величины, обратной к случайной величине с распределением Коши

Пусть $\xi = k[1]$ — случайная величина, имеющая распределения Коши (см. п. 2.1.1) и $F_\xi(x) = 1/2 + 1/\pi \operatorname{arctg} x$, $x \in \mathbb{R}$. Положим $\eta = 1/\xi$ (рис. 4). Следуя (1), получаем:

$$F_\eta(y) = P\{\eta < y\} = P\left\{\frac{1}{\xi} < y\right\} = P\left\{\xi > \frac{1}{y} \cup \xi < 0\right\} =$$
$$= P\left\{\xi > \frac{1}{y}\right\} + P\{\xi < 0\} = 1 - F_\xi\left(\frac{1}{y}\right) + \frac{1}{2} = 1 - \frac{1}{2} - \frac{1}{\pi}\operatorname{arctg}\frac{1}{y} + \frac{1}{2} =$$
$$= 1 - \frac{1}{\pi}\operatorname{arctg}\frac{1}{y} = 1 - \frac{1}{\pi}\left(\frac{\pi}{2} - \operatorname{arctg} y\right) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} y.$$

Таким образом, если $\xi = k[1]$, то и $\frac{1}{\xi} = k[1]$.

§ 2. Функции двух переменных. Действия над случайными величинами

Пусть $\bar{\xi} = \{\xi_1, \xi_2\}$ — двумерный случайный вектор с законом распределения $\Phi_{\bar{\xi}}\{B\} = \Phi_{\xi_1\xi_2}\{B\} = P\{\bar{\xi} \in B\}$, $\varphi(A) = \varphi(t_1, t_2)$ — борелевская функция двух переменных, $\eta = \varphi(\bar{\xi}) = \varphi(\xi_1, \xi_2)$. Как и выше, задачу нахождения закона распределения случайной величины η решает схема (*)–(**). В частности, умение решать эту задачу дает возможность *производить арифметические действия* над случайными величинами: складывать, вычитать, умножать, делить и т. п. Особенно важным для дальнейшего является случай независимых компонент вектора $\bar{\xi}$, на котором мы будем всякий раз останавливаться подробнее.

Распределение максимума двух случайных величин

Пусть $\eta = \max\{\xi_1, \xi_2\}$, $F_{\bar{\xi}}(\mathbf{x})$ — функция распределения вектора $\bar{\xi}$.

$$F_\eta(y) = P\{\eta < y\} = P\{\max(\xi_1, \xi_2) < y\} = P\{\xi_1 < y;\ \xi_2 < y\} = F_{\bar{\xi}}(y, y). \tag{1}$$

Если вектор $\bar{\xi}$ непрерывен и существует плотность распределения $\bar{\xi}$ – $f_{\bar{\xi}}(\mathbf{x})$, то η также непрерывна и, очевидно, ее плотность дается соотношением

$$f_\eta(y) = \frac{dF_\eta}{dy} = \frac{\partial F_{\bar{\xi}}}{\partial x_1}(y, y) + \frac{\partial F_{\bar{\xi}}}{\partial x_2}(y, y) = \int\limits_{-\infty}^{y} f_{\bar{\xi}}(x_1, y)\,dx_1 + \int\limits_{-\infty}^{y} f_{\bar{\xi}}(y_1, x_2)\,dx_2. \tag{2}$$

Если дополнительно компоненты ξ_1 и ξ_2 — независимы, то

$$F_\eta(y) = F_{\xi_1}(y)F_{\xi_2}(y), \tag{3}$$

а в случае непрерывности

$$f_\eta(y) = f_{\xi_1}(y)F_{\xi_2}(y) + F_{\xi_1}(y)f_{\xi_2}(y). \tag{4}$$

Распределение минимума двух случайных величин

Пусть $\eta = \min\{\xi_1, \xi_2\}$, $F_{\bar{\xi}}(\mathbf{x})$ — функция распределения вектора $\bar{\xi}$.

$$F_\eta(y) = P\{\min(\xi_1, \xi_2) < y\} =$$
$$= 1 - P\{\min(\xi_1, \xi_2) \geqslant y\}.$$

Для последней вероятности получаем (рис. 5).

$$P\{\min(\xi_1, \xi_2) \geqslant y\} = P\{\xi_1 > y;\ \xi_2 > y\} =$$
$$= 1 - F_{\xi_1}(y) - F_{\xi_2}(y) + F_{\xi_1\xi_2}(y, y). \tag{5}$$

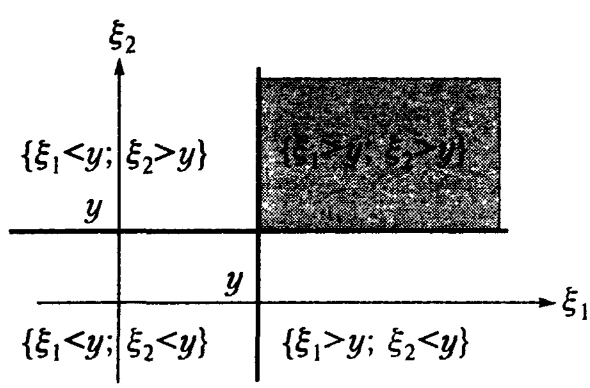


Рис. 5

Действительно, из

$$\{\xi_1 > y;\ \xi_2 > y\} + \{\xi_1 < y;\ \xi_2 > y\} + \{\xi_1 < y;\ \xi_2 < y\} + \{\xi_1 > y;\ \xi_2 < y\} = \{\bar{\xi} \in \mathbb{R}^2\}$$

следует

$$\{\xi_1 > y;\ \xi_2 > y\} + \{\xi_1 < y\} + \{\xi_2 < y\} - \{\xi_1 < y;\ \xi_2 < y\} = \{\bar{\xi} \in \mathbb{R}^2\},$$

а из последнего легко получается (5).

Окончательно имеем

$$F_\eta(y) = F_{\xi_1}(y) + F_{\xi_2}(y) - F_{\xi_1\xi_2}(y, y). \tag{6}$$

В случае непрерывности $\bar{\xi}$ с плотностью $f_{\bar{\xi}}(\mathbf{x})$ получаем, что η также непрерывна и

$$f_\eta(y) = f_{\xi_1}(y) + f_{\xi_2}(y) - \int\limits_{-\infty}^{y} f_\xi(x_1, y)\, dx_1 - \int\limits_{-\infty}^{y} f_\xi(y, x_2)\, dx_2. \tag{7}$$

Если ξ_1 и ξ_2 — независимы, то

$$F_\eta(y) = F_{\xi_1}(y) + F_{\xi_2}(y) - F_{\xi_1}(y)\cdot F_{\xi_2}(y) \tag{8}$$

и для непрерывных $\bar{\xi}$

$$f_\eta(y) = f_{\xi_1}(y) + f_{\xi_2}(y) - f_{\xi_2}(y)F_{\xi_1}(y) - f_{\xi_1}(y)F_{\xi_2}(y). \tag{9}$$

Отметим важную особенность экспоненциального распределения — если $\xi_1 = e^{\lambda_1}$, $\xi_2 = e^{\lambda_2}$ и они независимы, то $\min(\xi_1, \xi_2) = e^{\lambda_1+\lambda_2}$.

Действительно, (9) дает

$$f_\eta(y) = \lambda_1 e^{-\lambda_1 y} + \lambda_2 e^{-\lambda_2 y} - \lambda_2 e^{-\lambda_2 y}\left[1 - e^{-\lambda_1 y}\right] - \left[1 - e^{-\lambda_2 y}\right]\lambda_1 e^{-\lambda_1 y} =$$
$$= (\lambda_1 + \lambda_2)e^{-(\lambda_1+\lambda_2)y}.$$

Более того, как будет показано ниже, при достаточно широких предположениях относительно распределения независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_n$, величина $\eta = \min(\xi_1, \ldots, \xi_n)$ имеет распределение, близкое к экспоненциальному.

2.1. Сложение случайных величин. Свертка распределений

Пусть $\eta = \xi_1 + \xi_2$ и $F_{\bar{\xi}}(\mathbf{x})$ — функция распределения вектора $\bar{\xi}$. Для удобства изложения будем предполагать, что $\bar{\xi}$ обладает обобщенной плотностью $g_{\bar{\xi}}(\mathbf{x})$. Тогда получаем

$$F_\eta(y) = P\{\eta < y\} = P\{\xi_1 + \xi_2 < y\} = \iint\limits_{x_1+x_2<y} g_{\xi_1\xi_2}(x_1, x_2)\, dx_1\, dx_2 =$$
$$= \iint\limits_{x_1+x_2<y} f_{\bar{\xi}}(\mathbf{x})\, d\mathbf{x} + \sum_{i:\ a_i^1+a_i^2<y} p_i + \sum \int\limits_{\gamma_j\cap(x_1+x_2<y)} \varphi_{\gamma_j}(\mathbf{x})\, d\mathbf{x}_\gamma \tag{10}$$

(здесь $f_{\bar{\xi}}(\mathbf{x})$ — абсолютно непрерывная составляющая распределения $\bar{\xi}$, $p_i = P\{\bar{\xi} = \mathbf{a}_i\}$ дискретна, $\varphi_{\gamma_j}(\mathbf{x})$ — абсолютно непрерывная составляющая, сосредоточенная на гладкой линии γ_j в $\mathbb{R}^2$).

В частности, если вектор $\bar{\xi}$ непрерывен, то сумма η также непрерывна; соотношение (10) записывается в виде

$$F_\eta(y) = P\{\xi_1 + \xi_2 < y\} = \iint\limits_{x_1+x_2<y} f_{\xi_1\xi_2}(x_1, x_2)\, dx_1\, dx_2 = \int\limits_{-\infty}^{\infty} dx_1 \int\limits_{-\infty}^{y-x_1} f_{\xi_1\xi_2}(x_1, x_2)\, dx_2,$$

и плотность распределения η дается соотношением

$$f_\eta(y) = \int\limits_{-\infty}^{\infty} f_{\xi_1\xi_2}(x_1, y - x_1)\, dx_1. \tag{11}$$

Если $\bar{\xi}$ дискретна с рядом распределения $p_i = P\{\bar{\xi} = \mathbf{a}_i\}$, то сумма также дискретна и

$$P\{\xi_1 + \xi_2 = b_j\} = \sum_{i:\, a_i^1 + a_i^2 = b_j} p_i = q_j \tag{12}$$

— ее ряд распределения.

Рассмотрим теперь процедуру сложения *независимых* случайных величин. В этом случае соотношения (10) и (12) приобретают более компактный и завершенный вид.

Пусть $g_{\xi_1}(x_1)$ — плотность распределения случайной величины ξ_1, и $g_{\xi_2}(x_2)$ — плотность распределения случайной величины ξ_2

$$g_{\xi_1}(x_1) = f_{\xi_1}(x_1) + \sum p_i \delta(x_1 - a_i^1), \tag{13}$$

$$g_{\xi_2}(x_2) = f_{\xi_2}(x_2) + \sum q_j \delta(x_2 - a_j^2). \tag{14}$$

В случае независимости ξ_1 и ξ_2 получаем

$$g_{\xi_1\xi_2}(x_1, x_2) = g_{\xi_1}(x_1) \cdot g_{\xi_2}(x_2) = f_{\xi_1}(x_1) \cdot f_{\xi_2}(x_2) + \\ + \sum_j q_j \delta(x_2 - a_j^2) f_{\xi_1}(x_1) + \sum_i p_i \delta(x_1 - a_i^1) f_{\xi_2}(x_2) + \sum p_i q_j \delta_1 \delta_2,$$

и соотношение (10) можно записать в виде

$$F_\eta(y) = \int\limits_{-\infty}^{\infty} dx_1 \int\limits_{-\infty}^{y-x_1} f_{\xi_1}(x_1) f_{\xi_2}(x_2)\, dx_2 + \sum_{j=1}^{\infty} q_j \int\limits_{-\infty}^{y-a_j^2} f_{\xi_1}(x_1)\, dx_1 + \\ + \sum_{i=1}^{\infty} p_i \int\limits_{-\infty}^{y-a_i^1} f_{\xi_2}(x_2)\, dx_2 + \sum_{a_i^1 + a_j^2 < y} p_i q_j. \tag{15}$$

Обобщенная плотность суммы независимых случайных величин дается в этом случае соотношением

$$g_{\xi_1+\xi_2}(y) = \int\limits_{-\infty}^{\infty} f_{\xi_1}(x_1) f_{\xi_2}(y - x_1)\, dx_1 + \sum_j q_j \cdot f_{\xi_1}(y - a_j^2) + \\ + \sum p_i \cdot f_{\xi_2}(y - a_i^1) + \sum p_i q_j \delta(y - a_i^1 - a_j^2). \tag{16}$$

Отметим следующие, важные для приложений, частные случаи соотношений (15) и (16):

1. ξ_1 и ξ_2 непрерывны и независимы, тогда сумма непрерывна и

$$f_{\xi_1+\xi_2}(y) = \int\limits_{-\infty}^{\infty} f_{\xi_1}(t) f_{\xi_2}(y - t)\, dt. \tag{17}$$

2. ξ_1 и ξ_2 дискретны и независимы, тогда сумма дискретна и

$$P\{\xi_1 + \xi_2 = y\} = \sum_{\substack{i,\, j:\\ a_i^1 + a_j^2 = y}} P\{\xi_1 = a_i^1\} \cdot P\{\xi_2 = a_j^2\}. \tag{18}$$

3. ξ_1 непрерывна, ξ_2 любая

$$F_{\xi_1+\xi_2}(y) = \int\limits_{-\infty}^{\infty} f_{\xi_1}(t) F_{\xi_2}(y-t)\, dt. \tag{19}$$

4. ξ_1 дискретна, ξ_2 любая

$$F_{\xi_1+\xi_2}(y) = \sum_{i=1}^{\infty} P\{\xi_1 = a_i^1\} \cdot F_{\xi_2}\left(y - a_i^1\right). \tag{20}$$

5. ξ_1 непрерывна, ξ_2 дискретна, в этом случае сумма непрерывна и

$$f_{\xi_1+\xi_2}(y) = \sum_{j=1}^{\infty} P\{\xi_2 = a_j^2\} \cdot f_{\xi_1}\left(y - a_j^2\right). \tag{21}$$

Закон распределения суммы *независимых* случайных величин называется *сверткой* законов распределения слагаемых.

Например, соотношение (17) дает формулу свертки плотностей, (18) — свертки рядов распределения, (20) — свертки плотности с рядом распределения. Обычно свертка обозначается знаком «*». Это обозначение дает возможность символически представить функцию распределения суммы независимых слагаемых в виде

$$F_{\xi_1+\xi_2} = F_{\xi_1} * F_{\xi_2},$$

плотность распределения в виде

$$f_{\xi_1+\xi_2} = f_{\xi_1} * f_{\xi_2}$$

и т. д.

Как правило, при сложении независимых случайных величин характер распределения меняется, даже если складываются одинаково распределенные случайные величины.

Пример 1. ξ_1, ξ_2 — независимые, равномерные на $[0, 1]$ случайные величины. В соответствии с (15)

$$f_{\xi_1} * f_{\xi_2} = f_{\xi_1+\xi_2}(x) = \int\limits_{-\infty}^{\infty} f_{\xi_1}(t) f_{\xi_2}(x-t)\, dt =$$

$$= \begin{cases} \int\limits_0^x dt = x, & 0 \leqslant x \leqslant 1, \\ \int\limits_{x-1}^{1} dt = 2 - x, & 1 \leqslant x \leqslant 2, \end{cases} = \begin{cases} 0, & x < 0 \cup x > 2, \\ x, & 0 \leqslant x \leqslant 1, \\ 2 - x, & 1 \leqslant x \leqslant 2. \end{cases}$$

Таким образом, свертка двух равномерных на $[0, 1]$ случайных величин есть «треугольная» случайная величина (рис. 6).

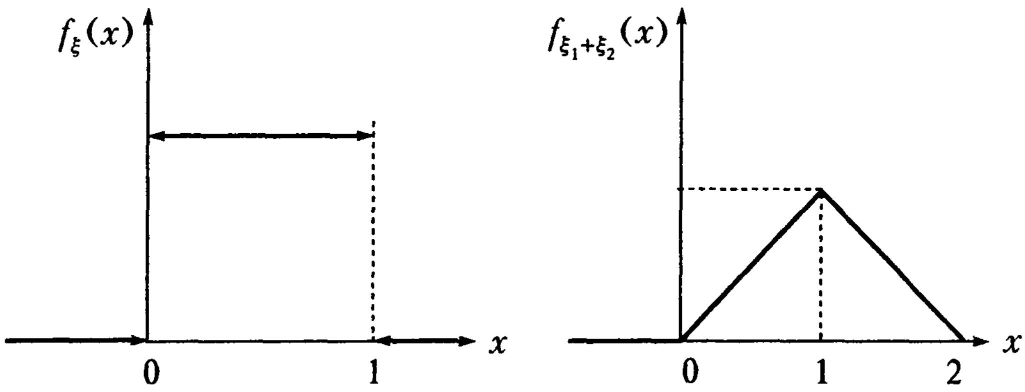


Рис. 6

Пример 2. $\xi_1 = \exp[\lambda_1]$; $\xi_2 = \exp[\lambda_2]$. Тогда

$$f_{\xi_1+\xi_2}(x) = \int\limits_{-\infty}^{\infty} f_{\xi_1}(t) f_{\xi_2}(x-t)\, dt =$$

$$= \begin{cases} \int\limits_0^x \lambda e^{-\lambda t} \cdot \lambda e^{-\lambda(x-t)}\, dt, & x \geqslant 0 \\ 0, & x < 0 \end{cases} = \begin{cases} \lambda^2 \int\limits_0^x e^{-\lambda x} dt & x \geqslant 0 \\ 0, & x < 0 \end{cases} = \begin{cases} \lambda^2 \cdot x e^{-\lambda x} & x \geqslant 0 \\ 0, & x < 0 \end{cases}.$$

Свертка двух экспоненциальных не является экспоненциальной случайной величиной (рис. 7).

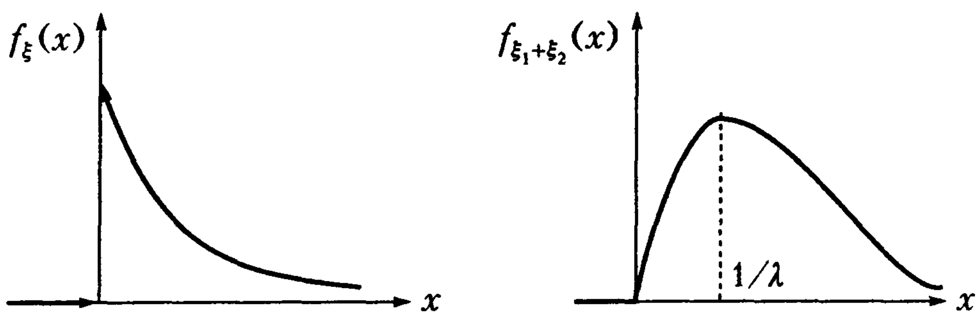


Рис. 7

Она является представителем семейства гамма-законов распределения (см. ниже).

Устойчивые относительно свертки распределения играют важную роль в теории и приложениях. Не касаясь вопроса о том, каким условиям должны удовлетворять и как описываются распределения, инвариантные относительно свертки, отметим инвариантность следующих часто встречающихся в приложениях распределений: нормального, пуассонова, гамма-распределения, распределения Коши и распределения Бернулли.

Сформулируем и докажем соответствующие утверждения для нормального распределения, распределения Пуассона и для гамма-распределения.

Теорема 1. *Если* $\xi_1 = N[m_1, \sigma_1]$ *и* $\xi_2 = N[m_2, \sigma_2]$, *то* $\xi_1 + \xi_2 = N\left[m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right]$:

$$N[m_1, \sigma_1] * N[m_2, \sigma_2] = N\left[m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right].$$

◀ По формуле (15) имеем

$$f_{\xi_1} * f_{\xi_2} = \int\limits_{-\infty}^{\infty} \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left\{-\frac{(t-m_1)^2}{2\sigma_1^2}\right\} \cdot \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left\{-\frac{(-x-t-m_2)^2}{2\sigma_2^2}\right\} dt =$$

$$= \left[\begin{array}{l} \dfrac{x-t-m_2}{\sigma_2} = \tau \\ t = x - m_2 - \sigma_2\tau \\ dt = \sigma_2 d\tau \\ M = m_1 + m_2 \end{array}\right] = \frac{1}{2\pi\sigma_1\sigma_2} \int\limits_{-\infty}^{\infty} \exp\left\{-\frac{(x-M-\sigma_2\tau)^2}{2\sigma_1^2} - \frac{\tau^2}{2}\right\} \cdot \sigma_2\, d\tau =$$

$$= \frac{1}{2\pi\sigma_1} \int\limits_{-\infty}^{\infty} \exp\left\{-\frac{A(\tau-\tau_0)^2 - B}{2}\right\} d\tau.$$

Элементарный (но несколько утомительный) подсчет дает

$$B = \frac{(x-m_1-m_2)^2}{2(\sigma_1^2+\sigma_2^2)}; \qquad A = \frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2}.$$

Поэтому

$$f_{\xi_1} * f_{\xi_2} = \frac{1}{2\pi\sigma_1} e^{-B} \int\limits_{-\infty}^{\infty} \exp\left\{\frac{A(\tau-\tau_0)^2}{2}\right\} d\tau = \left[\begin{array}{l} (\tau-\tau_0)\sqrt{A} = u \\ d\tau = \dfrac{du}{\sqrt{A}} \end{array}\right] =$$

$$= \frac{1 \cdot \cancel{\sigma_1}}{2\pi\, \cancel{\sigma_1} \sqrt{\sigma_1^2+\sigma_2^2}} \exp\left\{-\frac{(x-m_1-m_2)^2}{2\,(\sigma_1^2+\sigma_2^2)}\right\} \int\limits_{-\infty}^{\infty} \exp\left\{-\frac{u^2}{2}\right\} du =$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{\sigma_1^2+\sigma_2^2}} \exp\left\{\frac{(x-m_1-m_2)^2}{2\,(\sigma_1^2+\sigma_2^2)}\right\}. \blacktriangleright$$

Отметим, что нормальное распределение в некотором смысле «устойчиво» относительно свертки, а именно, если сумма двух независимых случайных величин имеет нормальное распределение, то оказывается слагаемые обязательно нормальны!

Теорема 2. *Если* $\xi_1 = \Pi[\lambda_1]$ *и* $\xi_2 = \Pi[\lambda_2]$, *то* $\xi_1 + \xi_2 = \Pi[\lambda_1 + \lambda_2]$:

$$\Pi[\lambda_1] * \Pi[\lambda_2] = \Pi[\lambda_1 + \lambda_2].$$

◀ $$P\{\xi_1 + \xi_2 = k\} = \sum_{r=0}^{k} P\{\xi_1 = r\} P\{\xi_2 = k - r\} = \sum_{r=0}^{k} e^{-\lambda_1} \frac{\lambda_1^r}{r!} \cdot e^{-\lambda_2} \frac{\lambda_2^{k-r}}{(k-r)!} =$$

$$= e^{-(\lambda_1+\lambda_2)} \sum_{r=0}^{k} \frac{1}{k!} C_k^r \lambda_1^r \lambda_2^{k-r} = e^{-(\lambda_1+\lambda_2)} \cdot \frac{(\lambda_1+\lambda_2)^k}{k!}. \blacktriangleright$$

Гамма-плотности

Будем говорить, что случайная величина ξ имеет гамма-распределение с параметрами (λ, ν), если ее плотность распределения задается соотношением ($\lambda > 0$; $\nu > 0$)

$$f_\xi(x) = \frac{1}{\Gamma(\nu)} \lambda^\nu x^{\nu-1} e^{-\lambda x}, \quad x \geqslant 0. \tag{22}$$

Обозначение: $\xi = \gamma[\lambda, \nu]$.

Здесь $\Gamma(\nu)$-гамма-функция Эйлера

$$\Gamma(\nu) = \int\limits_0^\infty x^{\nu-1} e^{-x}\, dx.$$

Отметим следующие, хорошо известные свойства гамма-функции:

1. $\Gamma(n+1) = n!$, n — натуральное.
2. $\nu \circ \Gamma(\nu) = \Gamma(\nu+1)$.

В справедливости этих свойств легко убедиться, интегрируя (22) по частям.

Определение плотности (22) корректно, так как для любых $\lambda > 0$ и $\nu > 0$ $f_\xi(x) \geqslant 0$ и выполнено условие нормировки

$$\int\limits_0^\infty f_\xi(x)\, dx = \int\limits_0^\infty \frac{1}{\Gamma(\nu)} \lambda^\nu x^{\nu-1} e^{-\lambda x}\, dx = \begin{bmatrix} \lambda x = t \\ dx = \dfrac{dt}{\lambda} \end{bmatrix} =$$

$$= \int\limits_0^\infty \frac{1}{\Gamma(\nu)} \lambda^\nu \cdot \frac{t^{\nu-1}}{\lambda^{\nu-1}} e^{-t} \cdot \frac{dt}{\lambda} = \frac{1}{\Gamma(\nu)} \int\limits_0^\infty t^{\nu-1} e^{-t}\, dt = \frac{1}{\Gamma(\nu)} \cdot \Gamma(\nu) = 1.$$

Имеет место теорема.

Теорема 3. *Если* $\xi_1 = \gamma[\lambda, \nu_1]$, $\xi_2 = \gamma[\lambda, \nu_2]$, *то* $\xi_1 + \xi_2 = \gamma[\lambda, \nu_1 + \nu_2]$:

$$\gamma[\lambda, \nu_1] * \gamma[\lambda, \nu_2] = \gamma[\lambda, \nu_1 + \nu_2].$$

$$\blacktriangleleft\ f_{\xi_1} * f_{\xi_2} = \int\limits_{-\infty}^\infty f_{\xi_1}(t) f_{\xi_2}(x-t)\, dt =$$

$$= \int\limits_0^x \frac{1}{\Gamma(\nu_1)} \lambda^{\nu_1} t^{\lambda_1 - 1} \cdot e^{-\lambda t} \cdot \frac{1}{\Gamma(\nu_2)} \lambda^{\nu_2} (x-t)^{\nu_2-1} e^{-\lambda(x-t)}\, dt =$$

$$= \frac{1}{\Gamma(\nu_1)\Gamma(\nu_2)} \lambda^{\nu_1+\nu_2} e^{-\lambda x} \int\limits_0^x t^{\nu_1-1} \cdot (x-t)^{\nu_2-1}\, dt.$$

В последнем интеграле положим $t = x\tau$. Тогда

$$\int\limits_0^x t^{\nu_1-1} \cdot (x-t)^{\nu_2-1}\, dt = \begin{bmatrix} t = x\tau \\ dt = x\, d\tau \end{bmatrix} = \int\limits_0^1 (x\tau)^{\nu_1-1}(x - \tau x)^{\nu_2-1} \cdot x\, d\tau =$$

$$= x^{\nu_1+\nu_2-1} \int\limits_0^1 t^{\nu_1-1}(1-\tau)^{\nu_2-1}\, d\tau,$$

откуда

$$f_{\xi_1} * f_{\xi_2} = \frac{1}{\Gamma(\nu_1)\Gamma(\nu_2)}\,\lambda^{\nu_1+\nu_2}\cdot x^{\nu_1+\nu_2-1}\cdot e^{-\lambda x}\cdot\int\limits_0^1 \tau^{\nu_1-1}(1-\tau)^{\nu_2-1}\,d\tau.$$

Поскольку $f_{\xi_1} * f_{\xi_2}$ — плотность, то для нее должно быть выполнено условие нормировки, поэтому

$$\int\limits_0^\infty f_{\xi_1} * f_{\xi_2}\,dx = \int\limits_0^\infty \frac{1}{\Gamma(\nu_1)\Gamma(\nu_2)}\lambda^{\nu_1+\nu_2}x^{\nu_1+\nu_2-1}e^{-\lambda x}\cdot\int\limits_0^1 \tau^{\nu_1-1}(1-\tau)^{\nu_2-1}\,d\tau\,dx =$$

$$= A(\nu_1,\nu_2)\cdot\int\limits_0^\infty \lambda^{\nu_1+\nu_2}\cdot x^{\nu_1+\nu_2-1}e^{-\lambda x}\,dx = A(\nu_1,\nu_2)\cdot\Gamma(\nu_1+\nu_2) = 1,$$

где

$$A(\nu_1,\nu_2) = \frac{1}{\Gamma(\nu_1)\Gamma(\nu_2)}\int\limits_0^1 \tau^{\nu_1-1}(1-\tau)^{\nu_2-1}\,d\tau = \frac{1}{\Gamma(\nu_1+\nu_2)}.$$

Отсюда окончательно заключаем, что

$$\gamma[\lambda,\nu_1]*\gamma[\lambda,\nu_2] = f_{\xi_1} * f_{\xi_2} = \frac{1}{\Gamma(\nu_1+\nu_2)}\,\lambda^{\nu_1+\nu_2}\cdot x^{\nu_1+\nu_2-1}e^{-\lambda x} =$$

$$= f_{\xi_1+\xi_2} = \gamma[\lambda,\nu_1+\nu_2]. \blacktriangleright$$

2.2. Другие действия над случайными величинами

Задача нахождения закона распределения результата других арифметических действий над случайными величинами решается аналогично. Отметим здесь основные соотношения для случая независимых операндов, следующие из (*)–(**).

Вычитание

$$\eta = \varphi(\xi_1,\xi_2) = \xi_1 - \xi_2, \qquad g_{\xi_1\xi_2}(x_1,x_2) = g_{\xi_1}(x_1)g_{\xi_2}(x_2)$$

(см. соотношения (13)–(14)).

$$F_\eta(y) = P\{\xi_1-\xi_2<y\} = \iint\limits_{x_1-x_2<y} g_{\xi_1\xi_2}(x_1,x_2)\,dx_1\,dx_2 =$$

$$= \iint\limits_{x_1-x_2<y} g_{\xi_1}(x_1)g_{\xi_2}(x_2)\,dx_1dx_2 = \int\limits_{-\infty}^{\infty}dx_1\int\limits_{-\infty}^{x_1-y} f_{\xi_1}(x_1)f_{\xi_2}(x_2)\,dx_2 +$$

$$+\sum_{j=1}^{\infty} q_j\int\limits_{a_j^2+y}^{\infty} f_{\xi_1}(x_1)\,dx_1 + \sum_{i=1}^{\infty} p_i\int\limits_{-\infty}^{a_i^1-y} f_{\xi_2}(x_2)\,dx_2 + \sum_{a_i^1-a_j^2<y} p_iq_j. \qquad (23)$$

Обобщенная плотность разности $g_\eta(y)$ при этом имеет вид

$$g_\eta(y) = \int\limits_{-\infty}^{\infty} f_{\xi_1}(x_1)f_{\xi_2}(x_1-y)\,dx_1 + \sum_j q_jf_{\xi_1}\left(a_j^2+y\right) +$$

$$+\sum_i p_if_{\xi_2}\left(a_i^1-y\right) + \sum p_iq_j\delta\left(y-a_i^1+a_j^2\right). \qquad (24)$$

Частные случаи (24), соответственно, для непрерывных

1. $f_{\xi_1\xi_2}(y) = \int\limits_{-\infty}^{\infty} f_{\xi_1}(x_1) f_{\xi_2}(x_1 - y)\, dx_1$

и дискретных случайных величин

2. $P\{\xi_1 - \xi_2 = y\} = \sum\limits_{a_i^1 - a_j^2 = y} P\{\xi_1 = a_i^1\} P\{\xi_2 = a_j^2\}.$

Аналоги соотношений (19), (20) и (21) очевидны.

Умножение

$$\eta = \xi_1 \cdot \xi_2, \quad g_{\xi_1\xi_2}(x_1, x_2) = g_{\xi_1}(x_1) g_{\xi_2}(x_2).$$

$$F_\eta(y) = P\{\eta < y\} = \iint\limits_{x_1 \cdot x_2 < y} g_{\xi_1\xi_2}(x_1, x_2)\, dx_1\, dx_2 = \iint\limits_{x_1 \cdot x_2 < y} g_{\xi_1}(x_1) g_{\xi_2}(x_2)\, dx_1\, dx_2 =$$

$$= \int\limits_{-\infty}^{0} dx_1 \int\limits_{y/x_1}^{\infty} g_{\xi_1}(x_1) g_{\xi_2}(x_2)\, dx_2 + \int\limits_{0}^{\infty} dx_1 \int\limits_{-\infty}^{y/x_1} g_{\xi_1}(x_1) g_{\xi_2}(x_2)\, dx_2 =$$

$$= \int\limits_{-\infty}^{0} g_{\xi_1}(x_1) \left[1 - F_{\xi_2}\left(\frac{y}{x_1}\right)\right] dx_1 + \int\limits_{0}^{\infty} g_{\xi_1}(x_1) F_{\xi_2}\left(\frac{y}{x_1}\right) dx_1. \qquad (25)$$

Выражение для обобщенной плотности произведения

$$g_\eta(y) = -\int\limits_{-\infty}^{0} g_{\xi_1}(x_1) g_{\xi_2}\left(\frac{y}{x_1}\right) dx_1 + \int\limits_{0}^{\infty} g_{\xi_1}(x_1) g_{\xi_2}\left(\frac{y}{x_1}\right) dx_1. \qquad (26)$$

Аналоги соотношений (17)–(21) очевидны.

Деление

$$\eta = \frac{\xi_1}{\xi_2}, \quad g_{\xi_1\xi_2}(x_1, x_2) = g_{\xi_1}(x_1) g_{\xi_2}(x_2).$$

Будем дополнительно предполагать, что $P\{\xi_2 = 0\} = 0$.

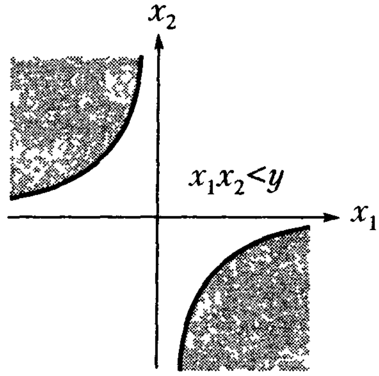
$$F_\eta(y) = P\{\eta < y\} = \iint\limits_{(x_1/x_2) < y} g_{\xi_1}(x_1) g_{\xi_2}(x_2)\, dx_1\, dx_2. \qquad (27)$$

Для $y < 0$ получаем (рис. 8)

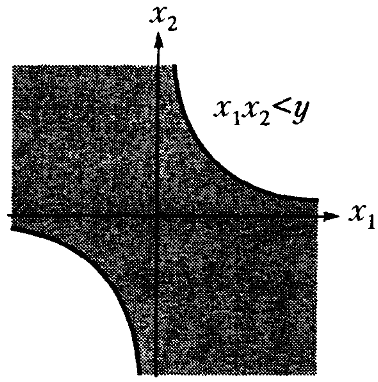
$$F_\eta(y) = \int\limits_{-\infty}^{0} dx_1 g_{\xi_1}(x_1)\, dx_1 \int\limits_{0}^{x_1/y} g_{\xi_2}(x_2)\, dx_2 + \int\limits_{0}^{\infty} g_{\xi_1}(x_1)\, dx_1 \int\limits_{x_1/y}^{0} g_{\xi_2}(x_2)\, dx_2, \quad y < 0.$$

Аналогично для $y > 0$ имеем

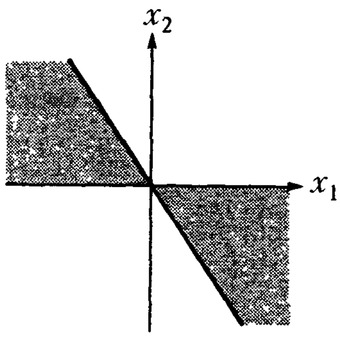
$$F_\eta(y) = 1 - \int\limits_{-\infty}^{0} g_{\xi_1}(x_1)\, dx_1 \int\limits_{x_1/y}^{0} g_{\xi_2}(x_2)\, dx_2 - \int\limits_{0}^{\infty} g_{\xi_1}(x_1)\, dx_1 - \int\limits_{0}^{x_1/y} g_{\xi_2}(x_2)\, dx_2, \quad y > 0.$$



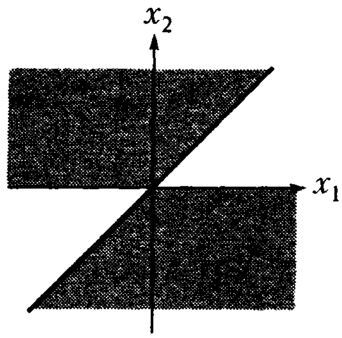
Область интегрирования $x_1x_2<y$ *для* $y<0$



Область интегрирования $x_1x_2<y$ *для* $y>0$



Область интегрирования $x_1/x_2<y$ *для* $y<0$



Область интегрирования $x_1/x_2<y$ *для* $y>0$

Рис. 8

Для дальнейшего нам понадобится выражение функции и плотности распределения частного в предположении, что знаменатель неотрицателен: $g_{\xi_2}(x_2) = 0,\ x_2 < 0$.

Учитывая вид области интегрирования (рис. 8, $x_2 > 0$), удобно в (27) расставить пределы интегрирования так, чтобы внешне интегрирование велось по x_2, а внутренне по x_1

$$F_\eta(y) = \int_0^\infty dx_2 \int_{-\infty}^{yx_2} g_{\xi_1}(x_1) g_{\xi_2}(x_2)\, dx_1 = \int_0^\infty F_{\xi_1}(yx_2) g_{\xi_2}(x_2)\, dx_2.$$

Обобщенная плотность при этом дается равенством

$$g_\eta(y) = \int_0^\infty g_{\xi_1}(yx_2) g_{\xi_2}(x_2) \cdot x_2\, dx_2. \tag{28}$$

§ 3. Функции нескольких переменных

В этом разделе мы остановимся на некоторых специфических функциях n переменных и их законах распределения, часто встречающихся в приложениях и играющих важную роль в статистике.

3.1. Экстремумы и порядковые статистики

Распределение максимума n случайных величин

Очевидное обобщение рассуждений предыдущего пункта (см. (1)) дает: если $\eta = \max\limits_{1\leqslant j\leqslant n} \{\xi_j\}$, $F_{\bar\xi}(\mathbf{x}) = P\{\bar\xi < \mathbf{x}\}$ — функция распределения $\bar\xi$, то

$$F_\eta(y) = P\left\{\max_j\{\xi_j\} < y\right\} = P\{\xi_1 < y;\ \xi_2 < y; \dots ; \xi_n < y\} = F_{\xi_1\xi_2\dots\xi_n}(y, y, \dots , y).$$

Если вектор $\bar\xi$ непрерывен с плотностью $f_{\bar\xi}(\mathbf{x})$, то плотность η дается соотношением

$$f_\eta(y) = \frac{dF_\eta}{dy} = \sum_{j=1}^{n} \frac{\partial F_{\bar\xi}}{\partial x_j}(y, y, \dots , y) = \sum_{j=1}^{n} \int\limits_{-\infty}^{y} f_{\bar\xi}(y_1, \dots , x_j, \dots , y)\, dx_j.$$

В случае независимости компонент вектора $\bar\xi$ в совокупности

$$F_\eta(y) = \prod_{j=1}^{n} F_{\xi_j}(y), \tag{1}$$

а в предположении непрерывности ξ_j

$$f_\eta(y) = \sum_{i=1}^{n} \prod_{j\neq i} f_{\xi_i}(y) F_{\xi_j}(y). \tag{2}$$

Распределение минимума n случайных величин

Обобщая соотношение (8) дословным повторением выкладок, получаем для вектора $\bar\xi$ с независимыми в совокупности компонентами

$$F_\eta(y) = P\left\{\min_j\{\xi_j\} < y\right\} = 1 - P\left\{\min_j\{\xi_j\} \geqslant y\right\} == 1 - \prod_{j=1}^{n}\bigl(1 - F_{\xi_j}(y)\bigr). \tag{3}$$

Для непрерывных $\bar\xi$ $\eta = \min\{\xi_j\}$ также непрерывна и

$$f_\eta(y) = \sum_{i=1}^{n} \prod_{j\neq i} f_{\xi_j}(y)\bigl(1 - F_{\xi_j}(y)\bigr). \tag{4}$$

Заметим, что здесь, как и для случая двухкомпонентного вектора $\bar\xi$, из условия экспоненциальности компонент следует экспоненциальность минимума.

Если n достаточно велико, то оказывается, что этот результат — экспоненциальность минимума — слабо зависит от характера распределения компонент. Точнее, имеет место следующее утверждение.

Теорема. *Пусть случайные величины ξ_j, $j = 1, 2, \dots , n$, — независимы в совокупности, непрерывны на $[0, +\infty)$ и одинаково распределены. Тогда при $n \to \infty$ распределение минимума ξ_j близко к экспоненциальному:*

$$F_{\min}(y) = \begin{cases} 0, & y < 0, \\ 1 - e^{-\mu_n \cdot y}, & y \geqslant 0. \end{cases}$$

Здесь $\mu_n = n\mu$, $\mu = f_{\xi_i}(0)$.

◀ Соотношение (3) в условиях теоремы дает

$$F_{\min}(y) = \begin{cases} 1 - [1 - F(y)]^n, & y > 0, \\ 0, & y < 0. \end{cases} \qquad (\star)$$

В силу непрерывности ξ_j в окрестности нуля (точнее, в правой полуокрестности) выполняется равенство

$$F(y) = f(0) \cdot y + o(y) = \mu \cdot y + o(y).$$

Из соотношения $(\star)$ ясно, что при $n \to \infty$ $F_{\min}(y) \to 1$, $y \neq 0$, и $F_{\min}(y) \to 0$, $y = 0$, т. е. вся информация о поведении минимума сосредоточена в окрестности нуля. Положим $t = n\mu y$. Тогда

$$F_{\min}(y) = 1 - [1 - F(y)]^n \approx 1 - \left[1 - \frac{t}{n}\right]^n \to 1 - e^{-t}. ▶$$

Указанное обстоятельство является теоретическим осмыслением т. н. «принципа слабого звена», широко используемого в теории надежности — надежность агрегата, функционирование которого необходимо зависит от надежности большого количества составляющих, определяется надежностью самого ненадежного из них и описывается экспоненциальным распределением.

Распределение порядковых статистик

Пусть $\bar{\xi} = \{\xi_1, \ldots, \xi_n\}$ — случайный вектор с законом распределения $F_{\bar{\xi}}(\mathbf{x})$. Вектор $\bar{\eta} = \{\eta_1, \eta_2, \ldots, \eta_n\}$ назовем вектором *порядковых статистик*, а его компоненты — *порядковыми статистиками*, если $\eta_1 = \min\limits_{1 \leqslant j \leqslant n} \{\xi_j\}$; $\eta_2 = \min\{\xi_j;\ \xi_j \neq \eta_1\}; \ldots ; \eta_n = \max\{\xi_j\}$. Компоненты вектора $\bar{\eta}$ расположены в порядке неубывания

$$\eta_1 \leqslant \eta_2 \leqslant \eta_3 \leqslant \ldots \leqslant \eta_n.$$

Найдем закон распределения m-й компоненты $\bar{\eta}$ в предположении независимости и одинаковой распределенности компонент $\bar{\xi}$, $1 < m < n$. Логику рассуждений рассмотрим на примере $n = 3$, $m = 2$.

Для того, чтобы вторая порядковая статистика приняла значение, меньшее y, нужно чтобы не менее двух из трех компонент вектора $\{\xi_1, \xi_2, \xi_3\}$ приняли значения, меньшие y. Это значит, что

$$P\{\eta_2 < y\} = P\{\xi_1 < y;\ \xi_2 < y;\ \xi_3 > y\} + \\ + P\{\xi_1 < y;\ \xi_2 > y;\ \xi_3 < y\} + P\{\xi_1 > y;\ \xi_2 < y;\ \xi_3 < y\} + \\ + P\{\xi_1 < y;\ \xi_2 < y;\ \xi_3 < y\} = \sum_{k=2}^{3} C_3^k F^k(y) [1 - F(y)]^{3-k}.$$

Аналогично для произвольных m и n

$$F_{\eta_m}(y) = \sum_{k=m}^{n} P\{\xi_{i_1} < y; \ldots ; \xi_{i_k} < y; \xi_{i_{k+1}} \geqslant y; \ldots ; \xi_{i_n} \geqslant y\} = \\ = \sum_{k=m}^{n} C_n^k F^k(y) (1 - F(y))^{n-k}. \qquad (5)$$

3.2. Кратные свертки. Некоторые специальные распределения

Столь же очевидно обобщается на случай произвольного конечного числа слагаемых понятие свертки случайных величин

$$F_{\xi_1} * F_{\xi_2} * \ldots * F_{\xi_n} = F_{\xi_1} * \left(F_{\xi_2} * F_{\xi_3} * \ldots * F_{\xi_n}\right) = F_{\xi_1+\xi_2+\ldots+\xi_n}.$$

Общие формулы при этом уже достаточно громоздки и необозримы, если только сворачиваемые распределения не являются устойчивыми относительно свертки — в последнем случае ситуация в техническом плане не сложнее, чем в случае двух переменных.

Особо отметим, что для нормальных, независимых в совокупности случайных величин из свойств линейного преобразования и теоремы 1 следует, что, если $\eta = \sum\limits_{i=1}^{n} \alpha_i\xi_i$, $\xi_i = N[m_i;\ \sigma_i]$, то $\eta = N\left[\sum \alpha_i m_i,\ \sum |\alpha_i|\sigma_i\right]$, т. е. линейная комбинация независимых нормальных случайных величин является нормальной случайной величиной.

Распределение χ^2 — Пирсона

Пусть $\chi^2 = \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2$, где $\xi_j = N[0,\ 1]$ независимы в совокупности.

Распределение суммы квадратов n независимых нормальных с параметрами $(0,\ 1)$ случайных величин называется *χ^2-распределением Пирсона с n степенями свободы.*

Читается — хи-квадрат. Обозначение: $\eta = \chi^2[n]$. Устойчивость χ^2-распределения относительно свертки усматривается непосредственно из определения

$$\chi^2[k] * \chi^2[m] = \chi^2[k+m]. \tag{6}$$

Для нахождения закона распределения случайной величины χ^2, заметим, что если $\xi = N[0,\ 1]$, то $\xi^2 = \gamma\left[1/2,\ 1/2\right]$.

◀ Действительно,

$$F_{\xi^2}(y) = P\left\{\xi^2 < y\right\} = P\{|\xi| < \sqrt{y}\} = \begin{cases} 0, & y < 0, \\ \dfrac{2}{\sqrt{2\pi}} \displaystyle\int\limits_0^{\sqrt{y}} e^{-x^2/2}\, dx, & y > 0, \end{cases}$$

откуда

$$f_{\xi^2}(y) = \frac{1}{\sqrt{2\pi}} \cdot y^{-1/2} \cdot e^{-y/2} = \gamma\left[\frac{1}{2},\ \frac{1}{2}\right]. ▶$$

В силу устойчивости гамма-распределения относительно свертки получаем

$$f_{\sum \xi_i^2} = f_{\xi_1^2} * f_{\xi_2^2} * \ldots * f_{\xi_n^2} = \gamma\left[\frac{1}{2},\ \frac{n}{2}\right] = \frac{y^{(n/2)-1}e^{-y/2}}{\Gamma(n/2)\cdot 2^{n/2}}. \tag{7}$$

t-распределение Стьюдента

Пусть $\eta, \xi_1, \xi_2, \ldots, \xi_n$ — независимые нормальные с параметрами $(0,\ 1)$ случайные величины и пусть

$$t = \frac{\eta\sqrt{n}}{\sqrt{\sum\limits_1^n \xi_i^2}}. \tag{8}$$

Распределение случайной величины t называется *распределением Стьюдента с n степенями свободы.*

Обозначение: $t = t[n]$.

Найдем выражение для закона распределения $t[n]$. Отметим, что числитель рассматриваемого отношения нормален с параметрами $(0, \sqrt{n})$, а знаменатель неотрицателен и его распределение дается выражением

$$P\left\{\sqrt{\sum_1^n \xi_1^2} < y\right\} = P\left\{\sum_1^n \xi_1^2 < y^2\right\} = P\{\chi^2[n] < y^2\}.$$

Из (7) заключаем, что

$$P\{\chi^2 < y^2\} = \frac{1}{\Gamma\left(\frac{n}{2}\right)2^{n/2}} \int_0^{y^2} z^{(n/2)-1} e^{-z/2}\, dz,$$

отсюда для плотности $\sqrt{\sum \xi_1^2}$ имеем

$$f_{\sqrt{\chi^2}}(y) = \frac{2y}{\Gamma\left(\frac{n}{2}\right)2^{n/2}} \cdot y^{n-2} e^{-y^2/2}.$$

Поэтому для частного $t[n]$, следуя (28), получаем

$$f_{t[n]}(y) = \int_0^\infty \frac{1}{\sqrt{2\pi n}} \exp\left\{-\frac{x^2}{2n}y^2\right\} \frac{2x}{\Gamma\left(\frac{n}{2}\right)2^{n/2}}\, x^{n-2}e^{-x^2/2}x\, dx =$$

$$= \frac{1}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)2^{(n-1)/2}} \int_0^\infty x^n \exp\left\{-\frac{x^2}{2}\left(1+\frac{y^2}{n}\right)\right\} dx.$$

Делая в последнем интеграле замену

$$u = \frac{x^2}{2}\left(1+\frac{y^2}{n}\right),$$

приходим к формуле

$$f_{t[n]}(y) = \frac{1}{\sqrt{\pi n}} \cdot \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(1+\frac{y^2}{n}\right)^{-(n+1)/2} \tag{9}$$

Заметим, что при $n = 1$ $f_{t[1]}(y) = K[1]$ — плотность распределения Коши.

$\mathcal{Z}$-распределение Фишера

Пусть $\xi_1, \xi_2, \ldots, \xi_n, \eta_1, \eta_2, \ldots, \eta_m$ — независимые в совокупности нормальные с параметрами $(0, 1)$ случайные величины. Положим

$$\mathcal{Z} = \frac{\frac{1}{n}\sum_1^n \xi_i^2}{\frac{1}{m}\sum_{i=1}^m \eta_i^2} = \frac{m\chi^2[n]}{n\chi^2[m]}. \tag{10}$$

Величина $\mathcal{Z}$ называется *случайной величиной Фишера-Снедекора.* Обозначение $\mathcal{Z} = \mathcal{Z}[n, m]$.

Закон распределения случайной величины $\mathcal{Z}$ найдем, используя (28) и (7). Имеем

$$f_{\mathcal{Z}\cdot(n/m)}(y) = \int\limits_0^\infty f_{\chi^2[n]}(ty) f_{\chi^2[m]}(t) t\, dt =$$

$$= \int\limits_0^\infty \frac{(ty)^{(n/2)-1} e^{-ty/2}}{\Gamma\left(\dfrac{n}{2}\right) 2^{n/2}} \cdot t \cdot \frac{t^{(m/2)-1} e^{-t/2}}{\Gamma\left(\dfrac{m}{2}\right) 2^{m/2}}\, dt =$$

$$= \frac{y^{(n/2)-1}}{2^{(m+n)/2}\Gamma\left(\dfrac{n}{2}\right)\Gamma\left(\dfrac{m}{2}\right)} \int\limits_0^\infty t^{((n+m)/2)-1} \exp\left\{-\frac{t}{2}(1+y)\right\} dt = \begin{bmatrix} \dfrac{t}{2}(1+y) = u \\ t = \dfrac{2u}{1+y} \\ dt = \dfrac{2du}{1+y} \end{bmatrix} =$$

$$= \frac{y^{(n/2)-1}}{2^{(m+n)/2}\Gamma\left(\dfrac{n}{2}\right)\Gamma\left(\dfrac{m}{2}\right)} \int\limits_0^\infty \frac{2^{((m+n)/2)-1} \cdot u^{((m+n)/2)-1}}{(1+y)^{((m+n)/2)-1}} \cdot e^{-u} \cdot 2\frac{du}{1+y} =$$

$$= \frac{\Gamma\left(\dfrac{m+n}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\Gamma\left(\dfrac{m}{2}\right)} y^{(n/2)-1} \frac{1}{(1+y)^{(m+n)/2}}. \tag{11}$$

Из соотношения (4) для линейного преобразования $\mathcal{Z} = \frac{m}{n} \cdot \left(\frac{n}{m}\mathcal{Z}\right)$ окончательно получаем

$$f_{\mathcal{Z}}(y) = \frac{n}{m} f_{\mathcal{Z}(n/m)}\left(\frac{ny}{m}\right). \tag{12}$$

3.3. Многомерное нормальное распределение

Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — независимые в совокупности нормальные с параметрами (0, 1) случайные величины, $\mathbf{A} = (a_j)$ — невырожденная матрица порядка n, $\mathbf{m} = (m_i)_{i=1}^n$ — столбец. Рассмотрим случайный вектор $\bar\eta = (\eta_i)_{i=1}^n$, задаваемый соотношением

$$\bar\eta = \mathbf{A}\bar\xi + \mathbf{m}. \tag{13}$$

Отметим, что каждая компонента η_j является линейной комбинацией нормальных с параметрами (0, 1) случайных величин

$$\eta_j = a_{j1}\xi_1 + a_{j2}\xi_2 + \ldots + a_{jn}\xi_n + m_j$$

и, в силу сделанного выше замечания, является нормальной случайной величиной.

Найдем закон распределения вектора (13). Пусть $B \subset \mathbb{R}^n$ — борелевское, тогда

$$P\{\bar\eta \in B\} = P\{\mathbf{A}\bar\xi + \mathbf{m} \in B\} = \int\limits_{\mathbf{x}\in\mathbb{R}^n:\ A\mathbf{x}+\mathbf{m}\in B} \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\sum x_i^2\right\} d\mathbf{x} =$$

$$= \frac{1}{(2\pi)^{n/2}} \int\limits_{\mathbf{x}\in\mathbb{R}^n:\ A\mathbf{x}+\mathbf{m}\in B} \exp\left\{-\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{x}\right\} d\mathbf{x}.$$

Сделаем в последнем интеграле замену переменных, положив

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{m}.$$

В силу невырожденности A эта замена невырождена и ее якобиан равен $|A|^{-1}$: $d\mathbf{x} = |A|^{-1} \cdot d\mathbf{y}$. Получим

$$\begin{aligned}
\mathbf{x} &= \mathbf{A}^{-1}(\mathbf{y} - \mathbf{m}),\\
\mathbf{x}^{\mathrm{T}} &= (\mathbf{y} - \mathbf{m})^{\mathrm{T}}(\mathbf{A}^{-1})^{\mathrm{T}},\\
\mathbf{x}^{\mathrm{T}} \cdot \mathbf{x} &= (\mathbf{y} - \mathbf{m})^{\mathrm{T}}(\mathbf{A}^{-1})^{\mathrm{T}}(\mathbf{A}^{-1})(\mathbf{y} - \mathbf{m}) = (\mathbf{y} - \mathbf{m})^{\mathrm{T}}(\mathbf{A}\mathbf{A}^{\mathrm{T}})^{-1}(\mathbf{y} - \mathbf{m}),
\end{aligned}$$

поэтому для любого борелевского $B \subset \mathbb{R}^n$ имеем

$$P\{\bar{\eta} \in B\} = \frac{1}{(2\pi)^{n/2}} \int\limits_B \exp\left\{-\frac{1}{2}(\mathbf{y} - \mathbf{m})^{\mathrm{T}}(\mathbf{A}\mathbf{A}^{\mathrm{T}})^{-1}(\mathbf{y} - \mathbf{m})\right\} \cdot \frac{1}{|\mathbf{A}|}\, d\mathbf{y}.$$

Положим $\mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{K}[\bar{\eta}]$, тогда $\det(\mathbf{A}\mathbf{A}^{\mathrm{T}}) = \det \mathbf{K} = |\mathbf{A}|^2$ и $|\mathbf{A}| = \sqrt{|\mathbf{K}|}$.

Для закона распределения вектора $\bar{\eta}$ получим

$$P\{\bar{\eta} \in B\} = \frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{K}|}} \int\limits_B \exp\left\{-\frac{1}{2}(\mathbf{y} - \mathbf{m})^{\mathrm{T}}\mathbf{K}^{-1}(\mathbf{y} - \mathbf{m})\right\} d\mathbf{y}, \tag{14}$$

отсюда следует, что $\bar{\eta}$ — непрерывный вектор, плотность которого дается равенством

$$f_{\bar{\eta}}(\mathbf{y}) = \frac{1}{(2\pi)^{1/2}\sqrt{|\mathbf{K}|}} \exp\left\{-\frac{1}{2}(\mathbf{y} - \mathbf{m})^{\mathrm{T}}\mathbf{K}^{-1}(\mathbf{y} - \mathbf{m})\right\}, \tag{15}$$

т. е. полностью определяется матрицей $\mathbf{K}$ и вектором $\mathbf{m}$.

Распределение (14)–(15) *называется невырожденным нормальным n-мерным распределением с параметрами* $(\mathbf{K}, \mathbf{m})$. Обозначение $\bar{\eta} = N[\mathbf{m}, \mathbf{K}]$.

Отметим, что здесь $\mathbf{K}$ — симметричная, положительно определенная матрица, $(\det \mathbf{A} \neq 0,\ \mathbf{K} = \mathbf{A}\mathbf{A}^{\mathrm{T}})$, $\mathbf{m}$ — произвольный вектор из $\mathbb{R}^n$.

Компоненты нормального случайного вектора — нормальные случайные величины. Однако нормальности компонент недостаточно для того, чтобы вектор был нормальным в смысле определения, данного выше.

Пример. Пусть $\xi_1 = N[0, 1]$, а $\xi_2 = \pm\xi_1$ с равными вероятностями

$$P\{\xi_2 = \xi_1\} = P\{\xi_2 = -\xi_1\} = 0.5.$$

◀ Компонента ξ_2 нормальна с параметрами $(0, 1)$, что немедленно следует из выкладок

$$\begin{aligned}
F_{\xi_2}(x_2) &= P\{\xi_2 < x_2\} = P\{\xi_2 < x_2/\xi_1 < x_2\}P\{\xi_1 < x_2\} + P\{\xi_2 < x_2/\xi_1 > -x_2\}P\{\xi_1 > -x_2\} =\\
&= \frac{1}{2}F_{\xi_1}(x_2) + \frac{1}{2}[1 - F_{\xi_1}(-x_2)] = \frac{1}{2}F_{\xi_1}(x_2) + \frac{1}{2}F_{\xi_1}(x_2) = F_{\xi_1}(x_2).
\end{aligned}$$

В то же время, совместное распределение $(\xi_1\xi_2)$ сосредоточено на паре прямых $x_1 = \pm x_2$, так что обобщенная плотность распределения этого вектора может быть представлена в виде

$$g_{\xi_1\xi_2}(x_1, x_2) = \frac{1}{2}(\delta(x_2 - x_1) + \delta(x_2 + x_1))f_{\xi_1}(x_1) \tag{16}$$

и не является плотностью совместного нормального распределения (15). ▶

Отметим еще одно важное свойство компонент нормального вектора: *они независимы тогда и только тогда, когда матрица* $\mathbf{K}$ *— диагональна.*

◀ Действительно, если $\xi_1, \xi_2, \ldots, \xi_n$ независимы и нормальны (m_i, σ_i), то для совместной плотности получаем

$$f_{\xi_1\xi_2\ldots\xi_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left\{-\frac{(x_i - m_i)^2}{2\sigma_1^2}\right\}. \tag{17}$$

Легко убедиться в том, что плотность (17) имеет вид (15) с $\mathbf{m} = [m_i]$ и $\mathbf{K} = (K_{ij})$, $K_{ii} = \sigma_i^2$, $k_{ij} = 0$.

Обратно, пусть $\mathbf{K} = \operatorname{diag}[\lambda_j]_{j=1}^n$, $\lambda_j > 0$. Тогда $\det \mathbf{K} = \lambda_1\lambda_2\ldots\lambda_n$, $\mathbf{K}^{-1} = \operatorname{diag}\left[\frac{1}{\lambda_j}\right]_{j=1}^n$ и плотность (15) принимает вид

$$f_{\bar\xi}(\mathbf{x}) = \frac{1}{(2\pi)^{1/2}\sqrt{\lambda_1\lambda_2\ldots\lambda_n}}\exp\left\{-\frac{1}{2}\sum_{i=1}^n (x_i - m_i)^2\cdot\frac{1}{\lambda_i}\right\} =$$

$$= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sqrt{\lambda_i}}\exp\left\{-\frac{(x_i - m_i)^2}{2\lambda_i}\right\},$$

т. е. представима в виде произведения нормальных плотностей, каждая из которых является индивидуальной плотностью распределения i-й компоненты, что и означает независимость компонент ξ_i. ▶

Пусть теперь $\xi_i = N[0, 1]$ — независимые, $\mathbf{A}$ — $m \times n$-матрица, ранг которой $\operatorname{rang}\mathbf{A} = r$. Можно показать, что вектор $\bar\eta = \mathbf{A}\bar\xi + \mathbf{m}$ имеет r-мерное нормальное распределение, которое в случае $r = m$ является невырожденным в $\mathbb{R}^m$ с параметрами $\left(\mathbf{K} = \mathbf{A}\mathbf{A}^{\mathrm{T}},\ \mathbf{m}\right)$, в случае же $r < m$ это распределение вырождено в $\mathbb{R}^m$ и сосредоточено на некотором подпространстве L, $\dim L = r$.

§ 4. Независимость функций независимых аргументов

В заключение этой главы рассмотрим одно важное свойство функций случайных аргументов.

Пусть $\bar\xi = \{\xi_i\}_{i=1}^n$, $\bar\eta = \{\eta_j\}_{j=1}^m$ — случайные векторы, законы распределения которых даются, соответственно, функциями $\Phi_{\bar\xi}(A)$ и $\Phi_{\bar\eta}(B)$, A — борелевское из $\mathbb{R}^n$, B — борелевское из $\mathbb{R}^m$. Пусть φ и Ψ — борелевские функции в $\mathbb{R}^n$ и $\mathbb{R}^m$ соответственно. Тогда имеет место

Теорема. *Если векторы $\bar\xi$ и $\bar\eta$ независимы, то случайные величины $\varphi(\bar\xi)$ и $\Psi(\bar\eta)$ — независимы.*

◀ Из независимости векторов $\bar\xi$ и $\bar\eta$ получаем

$$\Phi_{\bar\xi,\bar\eta}(A, B) = \Phi_{\bar\xi}(A)\cdot\Phi_{\bar\eta}(B), \quad \forall\, A \subset \mathbb{R}^n,\ B \subset \mathbb{R}^m. \tag{18}$$

Рассмотрим

$$P\left\{\varphi(\bar\xi) \in [a, b]\colon\ \Psi(\bar\eta) \in [c, d]\right\} = P\left\{\bar\xi \in \varphi^{-1}([a, b));\ \bar\eta \in \varphi^{-1}([c, d))\right\}.$$

В силу соотношения (18) последняя вероятность представима в виде

$$P\left\{\bar\xi \in \varphi^{-1}([a, b));\ \bar\eta \in \Psi^{-1}([c, d))\right\} =$$
$$= P\left\{\bar\xi \in \varphi^{-1}([a, b))\right\}\cdot P\left\{\bar\eta \in \Psi^{-1}([c, d))\right\} = P\left\{\varphi(\bar\xi) \in [a, b)\right\}\cdot P\left\{\Psi(\bar\eta) \in [c, d)\right\},$$

откуда и следует искомое. ▶

Глава XLI

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

Если ξ — случайная величина (может быть, и векторная), наблюдаемая в эксперименте Ω, то, как мы видели выше, исчерпывающую информацию о ней дает ее закон распределения $\Phi_\xi\{A\}$. Однако довольно часто можно, описывая случайную величину, ограничиться менее подробной, чем закон распределения, информацией, указав в каком-то смысле ее характерные значения и оценив, насколько наблюдаемая случайная величина может от этих значений уклоняться. Например, если ξ — результат измерения напряжения в бытовой электросети, то сказав, что это напряжение равно 220 В с возможными отклонениями плюс-минус 40 В, мы удовлетворим подавляющее большинство запросов пользователей этой сети.

Для некоторых случайных величин подобное описание может оказаться и неинформативным. Как правило, это происходит тогда, когда возможное рассеяние значений случайной величины оказывается значительным и соизмеримым со всем диапазоном принимаемых ею значений.

Мы рассмотрим здесь некоторые числовые показатели, называемые числовыми характеристиками случайной величины, которые отражают характерные для данной случайной величины аспекты ее поведения и позволяют описывать случайную величину в компактной форме. Важнейшими среди них являются *характеристики положения, характеристики рассеяния и характеристики связи.*

§ 1. Характеристики положения

Неслучайные числа, дающие представление о наиболее характерных значениях, принимаемых случайной величиной, будем называть *характеристиками положения.* К ним относятся математическое ожидание (центр распределения), мода (наиболее вероятное значение) и медиана (срединное значение). С точки зрения теории и ее приложений, важнейшей среди перечисленных характеристик является математическое ожидание, на изучении свойств которого мы и остановимся ниже.

1.1. Математическое ожидание случайной величины

Пусть $g_\xi(x)$ — обобщенная плотность распределения случайной величины ξ. *Математическим ожиданием*, или *средним*, случайной величины ξ назовем число, обозначаемое в дальнейшем $M[\xi]$, определяемое соотношением:

$$M[\xi] = \int_{-\infty}^{\infty} x g_\xi(x)\, dx. \tag{1}$$

Если интеграл (1) расходится, будем говорить, что у случайной величины ξ математическое ожидание отсутствует.

Для непрерывных случайных величин с плотностью $f_\xi(x)$ математическое ожидание дается равенством

$$M[\xi] = \int_{-\infty}^{\infty} x f_\xi(x)\, dx, \tag{2}$$

для дискретных, с рядом распределения $P\{\xi = a_i\} = p_i$, $i = 1, 2, \ldots, n, \ldots$ равенством

$$M[\xi] = \sum_{i=1}^{\infty} a_i p_i. \tag{3}$$

Если интерпретировать распределение вероятностей $g_\xi(x)$ как распределение единичной вероятностной массы по всей числовой прямой, то соотношения (1), (2) и (3) суть формулы для вычисления центра тяжести этой единичной массы. Именно это и имеют в виду, когда говорят, что математическое ожидание — это центр распределения.

Пример. Математическое ожидание экспоненциальной случайной величины.
Пусть $\xi = \exp[\mu]$. Тогда

$$f_\xi(x) = \begin{cases} 0, & x < 0, \\ \mu e^{-\mu x}, & x \geqslant 0, \end{cases}$$

и для $M[\xi]$ получаем

$$M[\xi] = \int_{-\infty}^{\infty} f_\xi(x) \cdot x\, dx = \int_{0}^{\infty} \mu \cdot x e^{\mu x}\, dx = \frac{1}{\mu}.$$

Пример. Математическое ожидание нормальной случайной величины.
Пусть $\xi = N[m, \sigma]$, т. е.

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}.$$

Для $M[\xi]$ имеем

$$M[\xi] = \int_{-\infty}^{\infty} \frac{x}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} dx = \begin{bmatrix} \dfrac{x-m}{\sigma} = u \\ x = \sigma u + m \\ dx = \sigma\, du \end{bmatrix} =$$

$$= \int_{-\infty}^{\infty} \frac{\sigma u + m}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} \cdot \sigma\, du = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u \exp\left\{-\frac{u^2}{2}\right\} du + \frac{m}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{u^2}{2}\right\} du = m.$$

Пример. Математическое ожидание бернуллиевой случайной величины.
Бернуллиевая случайная величина задается рядом распределения

$$\begin{array}{c||c|c} \xi & 0 & 1 \\ \hline p & q & p \end{array}, \quad p + q = 1; \quad p \geqslant 0; \quad q \geqslant 0.$$

Для нее соотношение (3) дает

$$M[\xi] = 0 \cdot q + 1 \cdot p = p.$$

Пример. Математическое ожидание случайной величины распределенной по Коши.
Плотность распределения Коши задается соотношением

$$f_\xi(x) = \frac{1}{\pi(1 + x^2)}, \quad x \in \mathbb{R}.$$

Из (2) получаем, что

$$M[\xi] = \int_{-\infty}^{\infty} x \frac{1}{\pi(1+x^2)}\, dx.$$

Легко установить, что последний интеграл расходится. У случайной величины, распределенной по Коши, математическое ожидание не существует.

1.2. Теорема о математическом ожидании функции от случайных величин. Свойства математического ожидания.

Важнейшие свойства математического ожидания являются следствиями следующей фундаментальной теоремы, которую мы примем без доказательства.

Теорема. *Пусть $\bar{\xi}$ — случайный вектор с законом распределения $g_{\bar{\xi}}(\mathbf{x}) = g_{\xi_1\xi_2\ldots\xi_n}(x_1x_2\ldots x_n)$ и функция $\varphi(\mathbf{t}) = \varphi(t_1, t_2, \ldots, t_n)$ такова, что интеграл*

$$\int\limits_{\mathbb{R}^n} |\varphi(\mathbf{t})| \cdot g_{\bar{\xi}}(\mathbf{t})\, d\mathbf{t}$$

сходится. Тогда

$$M\left[\varphi(\bar{\xi})\right] = \int\limits_{\mathbb{R}^n} \varphi(\mathbf{x}) g_{\bar{\xi}}(\mathbf{x})\, d\mathbf{x}. \tag{4}$$

Отметим следующие, важные для дальнейшего частные случаи формулы (4):

ξ — непрерывная скалярная случайная величина с плотностью $f_\xi(x)$

$$M[\varphi(\xi)] = \int_{-\infty}^{\infty} \varphi(x) f_\xi(x)\, dx; \tag{5}$$

ξ — дискретная скалярная случайная величина с рядом распределения $P\{\xi = a_i\} = p_i$, $i = 1, 2, \ldots$;

$$M[\varphi(\xi)] = \sum_{i=1}^{\infty} \varphi(a_i) p_i; \tag{6}$$

ξ — дискретно непрерывная: $g_\xi(x) = f(x) + \sum\limits_{i=1}^{\infty} \delta(x - a_i) p_i$

$$M[\varphi(\xi)] = \int_{-\infty}^{\infty} \varphi(x) f(x)\, dx + \sum_{i=1}^{\infty} p_i \varphi(a_i); \tag{7}$$

$\bar{\xi} = \{\xi_1, \xi_2\}$ — непрерывная с плотностью $f_{\xi_1\xi_2}(x_1, x_2)$

$$M\left[\varphi(\bar{\xi})\right] = \iint\limits_{\mathbb{R}^2} \varphi(x_1, x_2) f_{\xi_1\xi_2}(x_1, x_2)\, dx_1\, dx_2; \tag{8}$$

$\bar{\xi} = \{\xi_1, \xi_2\}$ — дискретно непрерывная с обобщенной плотностью $g_\xi(x_1, x_2)$

$$M[\varphi(\bar{\xi})] = \iint\limits_{\mathbb{R}^2} \varphi(x_1, x_2) g_{\xi_1\xi_2}(x_1, x_2)\, dx_1\, dx_2. \tag{9}$$

Из соотношения (4) и его вариантов (5)–(9) легко получаем следующие утверждения.

1. Математическое ожидание постоянной — постоянная

$$M[c] = c.$$

2. Если существуют $M[\xi_1]$, $M[\xi_2]$, а α и β — произвольные числа, то существует $M[\alpha\xi_1 + \beta\xi_2]$ и

$$M[\alpha\xi_1 + \beta\xi_2] = \alpha M[\xi_1] + \beta M[\xi_2].$$

3. Если существуют $M[\xi_1^2]$ и $M[\xi_2^2]$, а случайные величины ξ_1 и ξ_2 — независимы, то существует $M[\xi_1\xi_2]$ и

$$M[\xi_1\xi_2] = M[\xi_1] \cdot M[\xi_2].$$

4. Если существует $M[\xi]$, а ξ неотрицательна, то

$$M[\xi] \geqslant 0.$$

5. Если существуют $M[\xi_1]$, $M[\xi_2]$ и $\xi_1 \geqslant \xi_2$, то

$$M[\xi_1] \geqslant M[\xi_2].$$

6. Если существуют $M\left[\xi_1^2\right]$, $M\left[\xi_2^2\right]$, то существует $M[\xi_1\xi_2]$ и

$$M^2[\xi_1\xi_2] \leqslant M\left[\xi_1^2\right] M\left[\xi_2^2\right]$$

(неравенство Коши—Буняковского).

◀ Для доказательства последнего соотношения заметим, что в силу свойства 4 для любого $\lambda \in \mathbb{R}$

$$M[\lambda\xi_1 + \xi_2]^2 \geqslant 0.$$

Откуда для всех $\lambda \in \mathbb{R}$

$$\lambda^2 M\left[\xi_1^2\right] + 2\lambda M[\xi_1\xi_2] + M\left[\xi_2^2\right] \geqslant 0,$$

а это возможно лишь, если дискриминант квадратного трехчлена неположителен

$$D = M^2[\xi_1\xi_2] - M\left[\xi_1^2\right] M\left[\xi_2^2\right] \leqslant 0,$$

откуда следует искомое. ▶

Пример 1. Найдем математическое ожидание биномиальной случайной величины. Напомним, что $\xi = B[n; p]$ — количество успехов в серии из n независимых испытаний с неизменной вероятностью успеха p в каждом. Поэтому ξ представима в виде

$$\xi = \xi_1 + \xi_2 + \ldots + \xi_n,$$

где ξ_j — независимые случайные величины с рядом распределения $P\{\xi_j = 0\} = q = 1 - p$, $P\{\xi_j = 1\} = p$. Из свойства 2 для математического ожидания следует

$$M[\xi] = \sum_{j=1}^{\infty} M\xi_j = np.$$

Заметим, что мы одновременно доказали комбинаторное тождество

$$\sum_{k=0}^{n} C_n^k p^k q^{n-k} k = np.$$

Пример 2. Для нахождения математического ожидания случайной величины, имеющей распределение χ^2 с n степенями свободы, воспользуемся тем же свойством 2. Так как по определению

$$\chi^2[n] = \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2,$$

где ξ_j — независимые $(0,1)$-нормальные случайные величины, то

$$M[\chi^2] = \sum_{j=1}^{n} M\xi_j^2 = nM\left[\xi_j^2\right].$$

Из того, что $M\left[\xi_j^2\right] = 1$, получаем окончательно

$$M\left[\chi^2\right] = n.$$

1.3. Другие характеристики положения

Коротко остановимся еще на двух числовых характеристиках положения, описывающих «характерные» значения случайной величины.

Медианой случайной величины называется число, обозначаемое в дальнейшем $Me[\xi]$ и обладающее тем свойством, что

$$F_\xi(Me[\xi]) = 0{,}5. \tag{10}$$

Медиана может быть однозначно определена не для любой случайной величины. (Медиана распределения Коши равна нулю, в то время как ее математическое ожидание не существует; медиана нормального распределения и его математическое ожидание совпадают, аналогично для равномерного на $[a, b]$ и симметричного распределения.) Медиана может и не быть одним из возможных значений случайной величины.

Модой случайной величины называется число, обозначаемое в дальнейшем $Mo[\xi]$ и обладающее тем свойством, что это наиболее вероятное значение среди всех возможных значений случайной величины.

Для непрерывных случайных величин $Mo[\xi]$ — точка максимального значения плотности $f_\xi(x)$,

$$Mo[\xi] = \operatorname{argmax} f_\xi(x), \tag{11}$$

для дискретных — значение с наибольшей вероятностью,

$$Mo[\xi] = a_i, \quad P\{a_i\} \geqslant P\{a_j\} \quad \forall\, j.$$

Если таковых (наиболее вероятных значений) у распределения несколько, то оно называется *полимодальным*, если оно единственно, то *унимодальным*. Если наибольшее значение плотности достигается в концах промежутка множества значений случайной величины, а внутри этого промежутка имеется единственный минимум, то такое распределение называется *антимодальным*.

Отметим, что для симметричных унимодальных распределений мода, медиана и математическое ожидание (если последнее существует) совпадают. В общем случае между перечисленными характеристиками определенных соотношений нет.

Пример 3. Пусть $\xi = B[n; p]$ — биномиальная случайная величина с параметрами (n, p). Если $m = Mo[\xi]$, то по определению должны выполняться соотношения

$$P\{\xi = m-1\} \leqslant P\{\xi = m\} \quad \text{и} \quad P\{\xi = m+1\} \leqslant P\{\xi = m\}.$$

Поэтому

$$C_n^{m-1} p^{m-1} q^{n-m+1} \leqslant C_n^m p^m q^{n-m}, \qquad C_n^{m+1} p^{m+1} q^{n-m-1} \leqslant C_n^m p^m q^{n-m},$$

откуда

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

Вспоминая, что $M[\xi] = np$, получаем:

$$M[\xi] - q \leqslant Mo[\xi] \leqslant M[\xi] + p$$

или

$$p(n+1) - 1 \leqslant Mo[\xi] \leqslant p(n+1). \tag{12}$$

Из (12) следует, что мода у биномиальной случайной величины всегда есть и, при выполнении условия $p(n+1)$ — нецелое, она единственна; если же $p(n+1)$ — целое, то модой будут два соседних (равновероятных!) значения случайной величины ξ — $a = p(n+1) - 1$ и $b = p(n+1)$.

Пример 4. Аналогичные рассуждения для пуассоновой с параметром λ случайной величины $\xi = \Pi[\lambda]$ приведет к неравенству

$$\lambda - 1 \leqslant m \leqslant \lambda. \tag{13}$$

Выводы из соотношения (13) аналогичны приведенным в вышеизложенном примере.

§ 2. Характеристики рассеяния

Характеристики положения дают усредненное представление о характерных значениях, принимаемых случайными величинами. Информации в этих характеристиках тем больше, чем меньшие отклонения от них могут наблюдаться в реальном эксперименте.

Показатели, описывающие возможные отклонения значений случайной величины от «средних», называются *характеристиками рассеяния*. К ним относятся дисперсия, среднеквадратичное отклонение, срединное отклонение, коэффициент вариации и некоторые другие.

2.1. Дисперсия и ее свойства

Важнейшей из них является *дисперсия*.

Дисперсией случайной величины ξ (обозначение $D[\xi]$) называется математическое ожидание квадрата отклонения случайной величины ξ от своего среднего

$$D[\xi] = M[\xi - M\xi]^2. \tag{1}$$

Отметим некоторые свойства дисперсии.

1. $D[\xi] = M[\xi^2] - M^2[\xi]$.

◄ Из того, что $(\xi - M[\xi])^2 = \xi^2 - 2\xi M[\xi] + M^2[\xi]$, используя свойства математического ожидания, получаем

$$D[\xi] = M\left[\xi^2 - 2\xi M[\xi] + M^2[\xi]\right] = M[\xi^2] + 2M[\xi]M[\xi] + M^2[\xi] = M[\xi^2] - M^2[\xi]. \text{ ►}$$

2. $D[\xi] \geqslant 0$.

3. $D[c] = 0$.

4. $D[c\xi] = c^2 D[\xi]$.

5. $D[\xi_1 \pm \xi_2] = D[\xi_1] + D[\xi_2] \pm 2M\,[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)]$, в частности $D[\xi + c] = D[\xi]$.

$$\text{◄ } D[\xi_1 + \xi_2] = M\left[(\xi_1 + \xi_2 - M[\xi_1 + \xi_2])^2\right] = M\,[(\xi_1 - M\xi_1) + (\xi_2 - M\xi_2)]^2 =$$
$$= M[\xi_1 - M\xi_1]^2 + M[\xi_2 - M\xi_2]^2 \pm 2M\,[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)]. \text{ ►}$$

Отметим, что если случайные величины ξ_1 и ξ_2 — независимы, то из свойства 3 математического ожидания следует, что

$$M\left[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)\right] = M[\xi_1 - M\xi_1]M[\xi_2 - M\xi_2] = 0$$

и указанное свойство выглядит так:

$$D[\xi_1 \pm \xi_2] = D[\xi_1] + D[\xi_2].$$

6. Если $g_\xi(x)$ — обобщенная плотность распределения случайной величины ξ, то $D[\xi]$ может быть вычислена из соотношения

$$D[\xi] = \int_{-\infty}^{\infty} (x - M[\xi])^2 g_\xi(x)\,dx, \tag{2}$$

в частности, если ξ — непрерывная случайная величина с плотностью $f_\xi(x)$, то

$$D[\xi] = \int_{-\infty}^{\infty} (x - M[\xi])^2 f_\xi(x)\,dx = \int_{-\infty}^{\infty} x^2 f_\xi(x)\,dx - M^2[\xi], \tag{3}$$

если же ξ — дискретная случайная величина с рядом распределения $P\{\xi = a_i\} = p_i$, то

$$D[\xi] = \sum_{i=1}^{\infty} (a_i - M[\xi])^2\, p_i = \sum_{i=1}^{\infty} a_i^2 p_i - M^2[\xi]. \qquad (4)$$

Пример 1 (дисперсия бернуллиевой случайной величины).

Пусть ξ — бернуллиева случайная величина, $P\{\xi = 1\} = p$; $p\{\xi = 0\} = q = 1 - p$. В соответствие с соотношением (4), получаем ($M[\xi] = p$)

$$D[\xi] = 1^2 \cdot p + 0^2 \cdot q - p^2 = p(1-p) = pq.$$

Пример 2 (дисперсия биномиальной случайной величины).

Если ξ — биномиальная с параметрами (n, p), то, как было отмечено выше, ξ представима в виде

$$\xi = \xi_1 + \xi_2 + \ldots + \xi_n,$$

где ξ_j — независимые одинаково распределенные бернуллиевы с параметром p случайные величины. Поэтому (свойство дисперсии 5)

$$D[\xi] = D[\xi_1] + \ldots + D[\xi_n] = npq.$$

Одновременно доказано комбинаторное тождество

$$\sum_{n=0}^{n} k^2 \cdot C_n^k p^k (1-p)^{n-k} = npq + n^2p^2.$$

Пример 3 (дисперсия равномерной на $[a, b]$ случайной величины).

Пусть $\xi = R[a, b]$, $M\xi = \dfrac{a+b}{2}$. Имеем

$$D[\xi] = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \cdot \frac{1}{b-a}\, dx = \frac{(b-a)^2}{12}.$$

Пример 4 (дисперсия нормальной с параметрами (m, σ) случайной величины).

Пусть $\xi = N[m, \sigma]$, $M[\xi] = m$. Тогда

$$D[\xi] = \int_{-\infty}^{\infty} (x-m)^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} dx = \begin{bmatrix} \dfrac{x-m}{\sigma} = u \\ dx = \sigma\, du \end{bmatrix} =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma^2 u^2 \exp\left\{-\frac{u^2}{2}\right\} \sigma\, du = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 \exp\left\{-\frac{u^2}{2}\right\} du = \sigma^2.$$

Характеристикой рассеяния, тесно связанной с дисперсией, является *среднее квадратическое отклонение случайной величины*:

$$\sigma[\xi] = \sqrt{D\xi}. \qquad (5)$$

Обладая тем же качественным наполнением (содержа в себе ту же информацию), что и дисперсия, среднее квадратическое отклонение имеет то преимущество, что измеряется в тех же единицах, что и рассматриваемая случайная величина. Отметим, что из свойств дисперсии с очевидностью следует:

1. $\sigma[\xi] \geqslant 0$.
2. $\sigma[c] = 0$.
3. $\sigma[c\xi] = |c|\sigma[\xi]$.
4. $\sigma[\xi_1 \pm \xi_2] = \sqrt{\sigma^2[\xi_1] + \sigma^2[\xi_2]}$, если только ξ_1 и ξ_2 — независимы.

В заключение заметим, что если у случайной величины ξ существуют $M[\xi]$ и $D[\xi]$, то можно построить случайную величину $\overset{\circ}{\xi}$, обладающую теми же свойствами, что и ξ, но имеющую стандартные числовые характеристики: $M = 0$ и $D = 1$. Достаточно положить

$$\overset{\circ}{\xi} = \frac{\xi - m}{\sigma}. \qquad (6)$$

Переход от ξ к $\xi - m$ носит название *центрирование* случайной величины ξ, а переход от ξ к ξ/σ — *нормирование*. Таким образом, соотношение (6) описывает процедуру нормирования и центрирования случайной величины ξ. Очевидно, что центрирование ($\xi \to \xi - m$) не меняет дисперсии, в то время как нормирование, носящее характер масштабного преобразования, изменяет математическое ожидание в σ раз.

2.2. Неравенство Чебышёва

Из определения дисперсии (1) ясно, что она призвана качественно описывать рассеяние значений случайной величины относительно математического ожидания. Точный вероятностный смысл этого описания дается неравенством Чебышёва, которое мы здесь рассмотрим.

Теорема.

Пусть случайная величина ξ обладает математическим ожиданием $M[\xi] = m$ и дисперсией $D[\xi] = \sigma^2$. Тогда каково бы ни было $\varepsilon > 0$

$$P\{|\xi - m| \geqslant \varepsilon\} \leqslant \frac{\sigma^2}{\varepsilon^2}. \tag{7}$$

◀ Рассмотрим вспомогательную случайную величину η, заданную соотношением

$$\eta = \begin{cases} 0, & |\xi - m| < \varepsilon, \\ \varepsilon^2, & |\xi - m| \geqslant \varepsilon. \end{cases}$$

Заметим, что $\eta \leqslant |\xi - m|^2$, и потому

$$M[\eta] \leqslant M(\xi - m)^2 = \sigma^2.$$

По теореме о математическом ожидании функции от случайной величины получаем

$$M[\eta] = \varepsilon^2 P\{|\xi - m| \geqslant \varepsilon\},$$

откуда

$$\varepsilon^2 P\{|\xi - m| \geqslant \varepsilon\} \leqslant \sigma^2$$

или

$$P\{|\xi - m| \geqslant \varepsilon\} \leqslant \frac{\sigma^2}{\varepsilon^2},$$

чем и завершается доказательство. ▶

Отметим, что неравенство (7) часто используется в эквивалентной форме

$$P\{|\xi - m| < \varepsilon\} \geqslant 1 - \frac{\sigma^2}{\varepsilon^2}, \tag{8}$$

получающейся из (7) применением очевидного соотношения

$$P\{|\xi - m| < \varepsilon\} + P\{|\xi - m| \geqslant \varepsilon\} = 1.$$

Неравенство Чебышёва показывает, что чем меньше дисперсия, тем реже значения случайной величины ξ «сильно» (больше чем на ε) отклоняются от среднего m. При фиксированной дисперсии вероятности отклонений на величину, большую, чем ε, тем меньше, чем больше ε.

Неравенство (7) универсально. Оно не предъявляет никаких требований к характеру распределения случайной величины ξ — достаточно существования m и σ.

В силу своей универсальности оно малоинформативно количественно — для разумных значений ε оценки вероятностей $P\{|\xi - m| \geqslant \varepsilon\}$ и $P\{|\xi - m| < \varepsilon\}$ крайне грубы.

Пример. Для нормальной случайной величины с параметрами (0, 1) имеем

$$P\{|\xi| < 1\} = \frac{1}{\sqrt{2\pi}} \int\limits_{-1}^{1} \exp\left\{-\frac{x^2}{2}\right\} dx \approx 0{,}6827,$$

в то время как неравенство Чебышёва дает

$$P\{|\xi| < 1\} \geqslant 1 - \frac{1}{1} = 0,$$

что верно, но тривиально.

Для этой же случайной величины при $\varepsilon = 3$ точное значение вероятности $P\{|\xi| < 3\} = 0{,}9973$, а соотношение (8) приводит к оценке

$$P\{|\xi| < 3\} \geqslant 1 - \frac{1}{9} = \frac{8}{9} = 0{,}8889,$$

которая уже значительно лучше предыдущей.

Несмотря на достаточно грубый характер оценок (7)–(8), без дополнительных предположений о характере распределения случайной величины ξ, неравенство Чебышёва, как показывает следующий пример, улучшить нельзя — оно точное[1].

Пример. Пусть ξ — дискретная случайная величина, принимающая значения -1, 0 и 1 с вероятностями соответственно 1/4, 1/2, 1/4. Легко видеть, что $M[\xi] = 0$ и $D[\xi] = 1/2$. Положим $\varepsilon = 1$ и найдем значение вероятности $P\{|\xi| \geqslant 1\}$. Имеем

$$P\{|\xi| \geqslant 1\} = P\{\xi = -1\} + P\{\xi = 1\} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Неравенство (7) в этой ситуации дает оценку

$$P\{|\xi| \geqslant 1\} \leqslant \frac{D[\xi]}{\varepsilon^2} = \frac{1}{2} \cdot \frac{1}{1} = \frac{1}{2},$$

которая совпадает с точным значением оцениваемой вероятности.

2.3. Другие характеристики рассеяния

Из других характеристик рассеяния, часто используемых в приложениях, отметим *коэффициент вариации* и *срединное отклонение* (среднее арифметическое отклонение).

Пусть у случайной величины ξ существует $M[\xi] = m$ и $D[\xi] = \sigma^2$. *Коэффициентом вариации* случайной величины ξ называется величина

$$V[\xi] = \frac{\sigma^2}{m} \cdot 100\,\%. \tag{9}$$

Из (9) легко усмотреть, что $V[\xi]$ описывает рассеяние случайной величины ξ в долях по отношению к среднему. Как абсолютный показатель рассеяния коэффициент вариации не очень удобен, однако для совместно центрированных случайных величин (т. е. имеющих одинаковые математические ожидания) он позволяет эффективно сравнивать диапазоны изменения.

Пусть у случайной величины ξ существует $M[\xi] = m$.

Срединным отклонением случайной величины ξ называется величина U, задаваемая соотношением

$$U[\xi] = M|\xi - m|. \tag{10}$$

Срединное отклонение $U[\xi]$ качественно имеет тот же смысл, что и среднеквадратическое отклонение — чем больше срединное отклонение, тем больше рассеяние, чем меньше срединное отклонение — тем меньше рассеяние.

---

[1] В том смысле, что существует случайная величина ξ, для которой в неравенствах (7)–(8) при некотором ε достигается знак равенства.

Для конкретных классов распределений связь между этими показателями может быть установлена, однако в общем случае удобных для использования на практике соотношений между U и σ нет.

Пример 1. Пусть $\xi = N[m, \sigma]$ — нормально распределенная случайная величина. Тогда $U = \sigma\sqrt{\frac{2}{\pi}} < \sigma$.

◀ В этом случае

$$U = \int_{-\infty}^{\infty} |x-m| \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} dx = \left[\begin{array}{c} \frac{x-m}{\sigma} = u \\ dx = \sigma du \\ x - m = \sigma u \end{array}\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma|u| \exp\left\{-\frac{u^2}{2}\right\} du =$$

$$= \sigma\sqrt{\frac{2}{\pi}} \int_0^{\infty} u \exp\left\{-\frac{u^2}{2}\right\} du = \sigma\sqrt{\frac{2}{\pi}} \int_0^{\infty} \exp\left\{-\frac{u^2}{2}\right\} d\left(\frac{u^2}{2}\right) = \sigma\sqrt{\frac{2}{\pi}}. \blacktriangleright$$

Пример 2. Пусть $\xi = R[-a, a]$ — равномерно распределенная случайная величина. Тогда $U = a/2$.

◀ $M\xi = 0$.

$$U = \int_{-a}^{a} |x| \frac{1}{2a} dx = 2\int_0^a \frac{x}{2a} dx = \frac{a}{2}.$$

Отметим, что и в этом случае $\sigma = \frac{a\sqrt{3}}{3} > U$. ▶

Замеченное свойство $U < \sigma$ не случайно — оно имеет место для любых случайных величин (конечно, обладающих дисперсией).

Теорема. *Если у случайной величины ξ существует $D\xi = \sigma^2$, то*

$$U \leqslant \sigma.$$

◀ В неравенстве Коши—Буняковского (свойство 6 математического ожидания) положим $\xi_1 = |\xi - m|$, $\xi_2 \equiv 1$. Тогда

$$M^2[\xi_1\xi_2] = M^2[|\xi - m|] \leqslant M[\xi_1^2]M[\xi_2^2] = \sigma^2 \cdot 1,$$

откуда

$$U^2 \leqslant \sigma^2. \blacktriangleright$$

§ 3. Характеристики связи

Рассмотрим теперь двумерную случайную величину $\bar{\xi}$ с компонентами ξ и η: $\bar{\xi} = \{\xi, \eta\}$. Ясно, что для решения задачи компактного описания поведения этой пары случайных величин указания средних значений компонент и характеристик их рассеяния недостаточно — здесь важную роль играют, во-первых, факт наличия или отсутствия связи между компонентами, а во-вторых, характер связи.

Анализ свойств математического ожидания и дисперсии показывает, что описание взаимодействия компонент случайной величины $\bar{\xi}$ в какой-то степени может быть получено за счет изучения величины $M[(\xi - M\xi)(\eta - M\eta)]$. Действительно, вспомним, что дисперсия суммы двух случайных величин дается соотношением

$$D[\xi + \eta] = D[\xi] + D[\eta] + 2M[(\xi - M\xi)(\eta - M\eta)]$$

и в случае *независимости* ξ и η

$$M[(\xi - M\xi)(\eta - M\eta)] = 0.$$

Аналогичное положение и с математическим ожиданием произведения: в случае *независимости* ξ и η имеем

$$M[\xi\eta] = M[\xi]M[\eta],$$

в противном случае

$$M[\xi\eta] = M[\xi]M[\eta] + M[(\xi - M\xi)(\eta - M\eta)].$$

И вновь равенство $M[(\xi - M\xi)(\eta - M\eta)] = 0$ возникает как следствие независимости ξ и η. Эти соображения мотивируют более тщательное изучение указанной величины с целью установления степени ее «ответственности» за независимость компонент ξ и η.

3.1. Ковариация и корреляция

Ковариацией случайных величин ξ и η называется число

$$K[\xi, \eta] = K_{\xi\eta} = M[(\xi - M\xi)(\eta - M\eta)].$$

Свойства ковариации:

1. $K[\xi, \eta] = K[\eta, \xi]$.

2. Если $g_{\xi\eta}(x, y)$ — обобщенный закон распределения случайной величины $\bar{\xi} = \{\xi, \eta\}$, то

$$K[\xi, \eta] = \iint\limits_{\mathbb{R}^2} (x - M[\xi])(y - M[\eta])g_{\xi\eta}(x, y)\, dx\, dy.$$

3. Если случайные величины ξ и η независимы, то $K[\xi, \eta] = 0$.

4. Если $\sigma^2[\xi]$, $\sigma^2[\eta]$ — дисперсии случайных величин ξ и η соответственно, то

$$K^2[\xi, \eta] \leqslant \sigma^2[\xi]\sigma^2[\eta]. \tag{1}$$

◀ Последнее неравенство следует из неравенства Коши—Буняковского (свойство 6 математического ожидания). ▶

Нормированной ковариацией, или *коэффициентом корреляции*, случайных величин ξ и η называется число

$$\rho_{\xi\eta} = \rho[\xi, \eta] = \frac{K[\xi, \eta]}{\sigma[\xi]\sigma[\eta]}. \tag{2}$$

Наряду с очевидной переформулировкой отмеченных выше свойств ковариации отметим еще следующие:

5. $-1 \leqslant \rho[\xi, \eta] \leqslant 1$

◀ Очевидно следует из соотношения (1) и определения (2). ▶

6. $|\rho[\xi, \eta]| = 1$ тогда и только тогда, когда компоненты ξ и η линейно зависимы, при этом

$$\eta - M[\eta] = \rho \cdot \frac{\sigma[\eta]}{\sigma[\xi]}(\xi - M[\xi]), \tag{3}$$

где $|\rho| = 1$.

◀ Утверждение о линейной зависимости случайных величин ξ и η немедленно следует из того, что знак равенства в формуле (1) достигается в том и только в том случае, когда с вероятностью 1 случайные величины $\overset{\circ}{\xi} = \xi - M[\xi]$ и $\overset{\circ}{\eta} = \eta - M[\eta]$ линейно связаны, т. е. существуют постоянные α и β такие, что $P\{\overset{\circ}{\eta} = \alpha\overset{\circ}{\xi} + \beta\} = 1$.

Взяв математическое ожидание от обеих частей последнего равенства, заключаем, что $\beta = 0$. Для нахождения постоянной α перепишем упомянутое равенство в виде

$$\overset{\circ}{\eta} - \alpha\overset{\circ}{\xi} = 0$$

и найдем дисперсию от обеих частей этого соотношения. Имеем

$$D\left[\overset{\circ}{\eta} - \alpha\overset{\circ}{\xi}\right] = \sigma^2[\eta] + \alpha^2\sigma^2[\xi] - 2\alpha K[\xi, \eta] = 0.$$

Учитывая, что $\sigma^2[\eta] = \alpha^2\sigma^2[\xi]$ (это следует из равенства $\overset{\circ}{\eta} = \alpha\overset{\circ}{\xi}$), заключаем, что

$$\alpha\sigma^2[\xi] = K[\xi, \eta] \Rightarrow \alpha = \frac{K[\xi, \eta]}{\sigma^2[\xi]} = \frac{K[\xi, \eta]}{\sigma[\xi]\cdot\sigma[\eta]} \cdot \frac{\sigma[\eta]}{\sigma[\xi]} = \rho\frac{\sigma[\eta]}{\sigma[\xi]}. \blacktriangleright$$

Суммируя вышеизложенное, констатируем, что ковариация (коэффициент корреляции) пары случайных величин является надежным индикатором наличия между компонентами ξ и η *жесткой линейной связи* в случае $|\rho| = 1$. В то же время у нас нет оснований рассматривать ковариацию (коэффициент корреляции) как индикатор наличия или отсутствия какой-нибудь (не обязательно линейной) зависимости. Свойство 3 показывает, что ковариация обращается в нуль, если зависимость между ξ и η отсутствует. Как свидетельствуют приводимые ниже примеры, обратное утверждение, вообще говоря, неверно.

Пример 1. Пусть $(\xi, \eta) = R[x^2 + y^2 \leqslant 1]$ — равномерно распределенная в круге с центром в начале координат и радиусом 1 двумерная случайная величина. Легко сообразить, что компоненты ξ и η этой случайной величины зависимы. Действительно, закон распределения одной из компонент (например, η) зависит от значения, принятого другой: при $\xi = 0$ диапазон изменения η от -1 до 1, при $\xi = 1$ η принимает только нулевое значение. Ковариация же величин ξ и η, как показывают несложные выкладки, равна нулю: действительно, $M\xi = M\eta = 0$ и

$$K[\xi, \eta] = \iint\limits_{x^2+y^2\leqslant 1} x \cdot y \cdot \frac{1}{\pi}\, dx\, dy = 0.$$

Может показаться, что зависимость между ξ и η в рассмотренном примере не замечена ковариацией, потому что она (зависимость) в некотором смысле «ненастоящая». Однако следующий пример демонстрирует нам пару жестко связанных функциональной зависимостью случайных величин, ковариация которых также равна нулю.

Пример 2. Пусть $\xi = R[-1, 1]$, $\eta = \xi^2$. Имеем

$$M\xi = 0;$$

$$K[\xi, \eta] = M[\xi\cdot(\eta - M\eta)] = M[\xi\eta - \xi\cdot M\eta] = M[\xi\cdot\eta] = M[\xi^3],$$

но по теореме о математическом ожидании функции от случайной величины

$$M[\xi^3] = \int\limits_{-1}^{1} x^3 \cdot \frac{1}{2}\, dx = 0.$$

Чтобы понять роль ковариации (коэффициента корреляции) как показателя зависимости между случайными величинами ξ и η, рассмотрим следующую задачу: среди всех линейных функций от случайной величины ξ найти ту, которая наилучшим образом описывает связь между величинами ξ и η.

Другими словами, следует найти $l(\xi) = A\xi + B$ такую, что

$$\eta = A\xi + B + \zeta, \tag{4}$$

здесь ζ — остаток, обладающий тем свойством, что $M[\zeta] = 0$ и $D[\zeta] \to \min$.

Докажем следующее утверждение.

Теорема. *Если существуют* $D\xi = \sigma_\xi^2$, $D\eta = \sigma_\eta^2$, $M\xi = m_\xi$, $M\eta = m_\eta$, *то соотношение* (4) *принимает вид*

$$\eta - m_\eta = \rho\,\frac{\sigma_\eta}{\sigma_\xi}\,(\xi - m_\xi) + \zeta, \tag{5}$$

при этом

$$\min D[\zeta] = (1 - \rho^2)\,\sigma_\eta^2. \tag{6}$$

◀ В силу $M\zeta = 0$ имеем

$$m_\eta = Am_\xi + B \Rightarrow B = m_\eta - Am_\xi$$

и далее

$$D[\zeta] = D[\eta - A\xi - B] = D[\eta - A\xi] = \sigma_\eta^2 + A^2\sigma_\xi^2 - 2AK[\xi, \eta].$$

Этот квадратный (относительно A) трехчлен достигает своего наименьшего значения в точке

$$A_{\min} = \frac{K[\xi, \eta]}{\sigma_\xi^2} = \rho\,\frac{\sigma_\eta}{\sigma_\xi}.$$

Последнее соотношение вместе с ранее полученным выражением для B и доказывает зависимость (5). Равенство (6) немедленно получается подстановкой $A_{\min}$ в выражение для $D[\zeta]$. ▶

Вернемся к обсуждению роли коэффициента корреляции, как показателя зависимости между случайными величинами.

Соотношение (6) показывает, что если $\rho = 0$, то $D[\zeta] = D[\eta]$, а формула (5) принимает вид

$$\eta - m_\eta = \zeta,$$

т. е. линейной составляющей ξ в описании случайной величины η нет и вся изменчивость случайной величины η сосредоточена в «остатке» ζ.

Если же $\rho \neq 0$, то для дисперсии $D[\eta]$ получаем

$$D[\eta] = \sigma_\eta^2 = \rho^2\sigma_\eta^2 + (1 - \rho^2)\,\sigma_\eta^2, \tag{7}$$

где первое слагаемое отвечает линейной составляющей зависимости (5), а второе — «остатку» ζ. Чем ближе значение ρ к нулю, тем хуже линейная часть (5) описывает тенденции в изменении величины η (рис. 1).

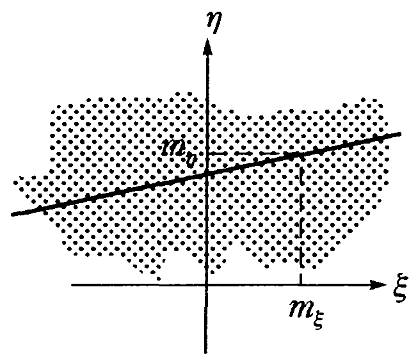
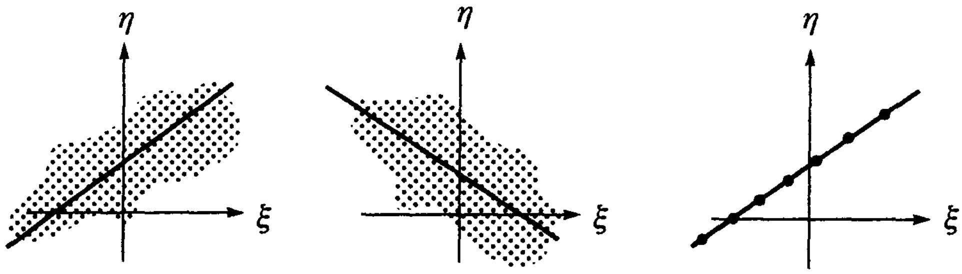


Рис. 1

Чем ближе значение $|\rho|$ к 1, тем бóльшая часть изменчивости случайной величины η описывается линейной составляющей соотношения (5); при $|\rho| = 1$ зависимость линейная, разброса у точек (ξ, η) относительно прямой (5) нет, величина ζ с вероятностью 1 равна нулю (рис. 2).

В случае $\rho > 0$ говорят о положительной коррелированности случайных величин ξ и η: с ростом одной из них почти линейно растет и другая, в случае $\rho < 0$ — об отрицательной (рис. 2 a, b).

Окончательно заключаем: *ковариация (коэффициент корреляции) является мерилом степени* **линейной** *зависимости между случайными величинами.* Равенство ковариации (коэффициента корреляции) нулю свидетельствует лишь об отсутствии линейной связи между ними. Отличие ковариации (коэффициента корреляции) от нуля свидетельствует о наличии связи между случайными величинами, причем, чем ближе коэффициент корреляции к единице, тем ближе эта связь к линейной.



положительная корреляция $\rho>0$
а)

отрицательная корреляция $\rho<0$
б)

линейная связь $\rho=1$
в)

Рис. 2

3.2. Корреляционное отношение

Для получения числовой характеристики, подобной коэффициенту корреляции, но охватывающей уже не только линейные, но и другие виды функциональных связей между случайными величинами, рассмотрим аналог задачи о представлении случайной величины η в виде (4): найдем функцию $r(\xi)$ такую, что

$$\eta = r(\xi) + \zeta, \tag{8}$$

где $M\zeta = 0,\ D[\zeta] \to \min$.

Можно показать, что решение этой задачи существует и дается так называемым *условным математическим ожиданием* η относительно ξ, если последнее определено. Для нас же важно, что представление (8) сопровождается аналогом дисперсионного разложения (7), а именно, дисперсия величины η раскладывается в сумму двух дисперсий $\sigma^2_{r(\xi)}$ и σ^2_ζ так, что величина каждой из них описывает вклад соответствующего слагаемого соотношения (8) в изменчивость η

$$D[\eta] = \sigma^2_\eta = \sigma^2_{r(\xi)} + \sigma^2_\zeta. \tag{9}$$

По аналогии с соотношением (6) введем в рассмотрение величину $\Theta > 0$ такую, что

$$\min\ D[\zeta] = \sigma^2_\zeta = (1 - \Theta^2)\sigma^2_\eta. \tag{10}$$

Отметим некоторые свойства показателя Θ.

1. $0 \leqslant \Theta^2 \leqslant 1$

◀ Следует из неотрицательности слагаемых в соотношении (9). ▶

2. $\sigma^2_{r(\xi)} = \Theta^2\sigma^2_\eta$.

3. Если $\Theta = 0$, то изменение случайной величины η не связано с изменением ξ. (Заметим, что это не означает стохастической независимости случайных величин η и ξ (см. пример 2), а говорит лишь об отсутствии тенденции к закономерному (описываемому функциональной зависимостью) изменению η с изменением ξ.

4. Если $\Theta = 1$, то между случайными величинами ξ и η существует жесткая функциональная связь, описываемая функцией $r(\xi)$.

Промежуточные значения показателя Θ (между нулем и единицей) дают представление о ширине корреляционного поля разброса точек (ξ, η) относительно линии $r(\xi)$ (рис. 3).

5. Если ρ — коэффициент корреляции пары (ξ, η), то

$$0 \leqslant |\rho| \leqslant \Theta \leqslant 1, \qquad |\rho| = \Theta \Leftrightarrow r(\xi) = l(\xi).$$

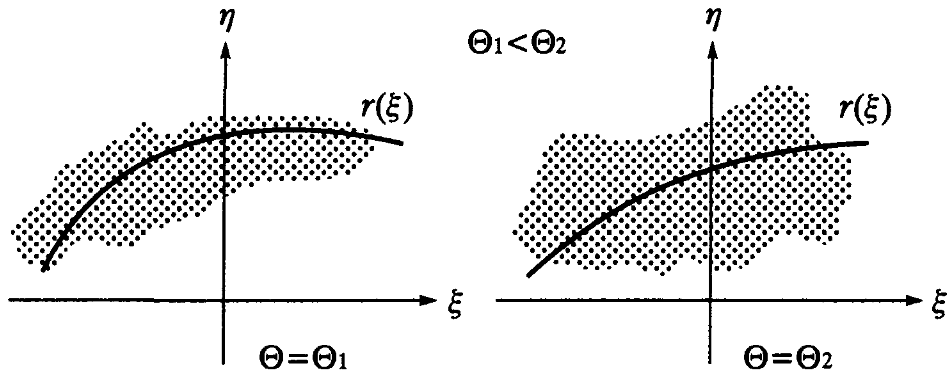


Рис. 3

◀ Для доказательства достаточно сравнить дисперсию остатка ζ для $r(\xi) = \operatorname{argmin} D[\zeta]$ и линейной функции $l(\xi)$. Если $r(\xi)$ линейна, то $r(\xi) \equiv l(\xi)$ и $\Theta^2 = \rho^2$ в силу (6). Если же $r(\xi)$ линейной функцией не является, то

$$D[\zeta] = (1 - \rho^2)\sigma_\eta^2 > (1 - \Theta^2)\sigma_\eta^2 = \min D[\zeta]$$

в силу оптимальности $r(\xi)$, откуда следует искомое. ▶

Величина Θ называется *корреляционным отношением*, а функция $r(\xi) = \operatorname{argmin} D[\zeta]$ *линией регрессии η на ξ.*

Заметим, что регрессия η на ξ и ξ на η — это, вообще говоря, различные линии! И корреляционное отношение Θ несимметрично относительно компонент ξ и η. Величина, которую мы определили выше, является корреляционным отношением величины η относительно величины ξ и, чтобы подчеркнуть это обстоятельство, обычно используют следующее обозначение

$$\Theta^2 = \Theta^2_{\eta/\xi} = 1 - \frac{\sigma^2_{\zeta/\xi}}{\sigma^2_\eta};$$

аналогично для корреляционного отношения относительно η

$$\Theta^2_{\xi/\eta} = 1 - \frac{\sigma^2_{\zeta/\eta}}{\sigma^2_\xi}.$$

Уже в случае линейной регрессии можно было заметить, что регрессия η на ξ, задаваемая соотношением

$$l_1(\xi) = \rho\,\frac{\sigma_\eta}{\sigma_\xi}\,(\xi - m_\xi),$$

и регрессия ξ на η, задаваемая аналогичным уравнением

$$l_2(\eta) = p\,\frac{\sigma_\xi}{\sigma_\eta}\,(\eta - m_\eta),$$

это различные прямые, совпадающие лишь в случае $|\rho| = 1$.

Определенных соотношений между $\Theta_{\eta/\xi}$ и $\Theta_{\xi/\eta}$ в общем случае нет. Как показывает рассмотренный выше пример 2, при $\Theta_{\xi/\eta} = 0$ возможно даже $\Theta_{\eta/\xi} = 1$.

§ 4. Старшие моменты. Асимметрия и эксцесс

Начальным моментом случайной величины ξ *порядка* s называется число

$$M\xi^s = m_s, \tag{1}$$

центральным моментом случайной величины ξ *порядка* s называется число

$$M(\xi - m_1)^s = \alpha_s. \tag{2}$$

Формулы (1)–(2) имеют смысл, если соответствующие интегралы сходятся. В противном случае говорят, что моменты не существуют.

Начальные и центральные моменты связаны между собой простыми соотношениями

$$\alpha_1 = 0; \ m_1 = M[\xi],$$

$$\alpha_2 = \sigma^2 = m_2 - m_1^2,$$

$$\alpha_3 = m_3 - 3m_2m_1 + 2m_1^2,$$

$$\alpha_4 = m_4 - 4m_3m_1 + 6m_2m_1^2 - 3m_1^4,$$

....

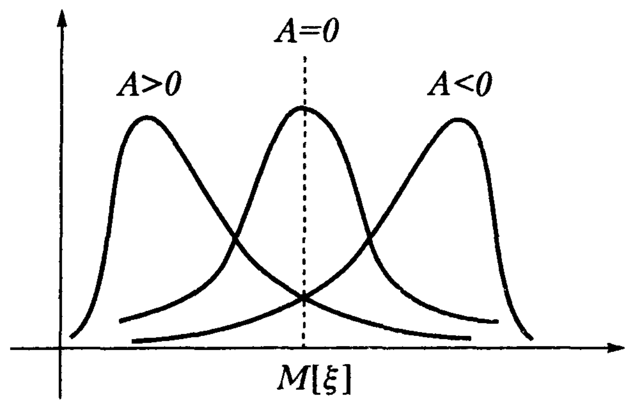


Рис. 4

Со старшими моментами (порядок которых выше второго) связаны две характеристики распределения, часто использующиеся в приложениях — асимметрия и эксцесс.

Асимметрией распределения случайной величины ξ называется число

$$A[\xi] = \frac{\alpha_3}{\sigma_\xi^3}. \tag{3}$$

Смысл асимметрии легко усматривается из формулы (3): если $A[\xi] < 0$ — распределение имеет правую асимметрию, если $A[\xi] > 0$ — левую, при $A[\xi] = 0$ — распределение симметрично относительно прямой $x = M[\xi]$ (рис. 4).

Четвертый центральный момент, нормированный четвертой степенью среднеквадратичного отклонения, называется *эксцессом* распределения

$$E[\xi] = \frac{\alpha_4}{\sigma_\xi^4} - 3. \tag{4}$$

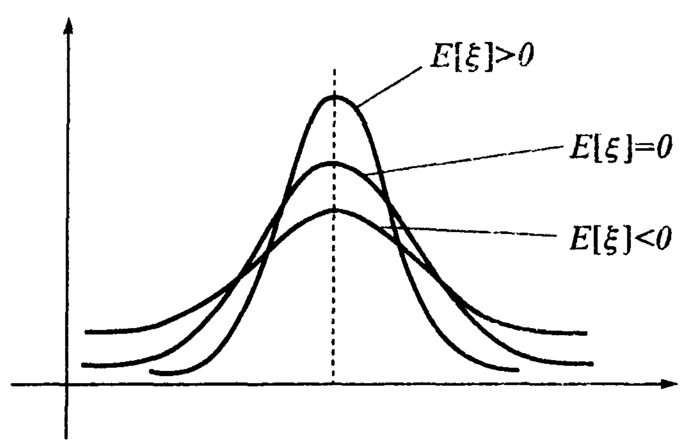


Рис. 5

Эксцесс характеризует сглаженность кривой плотности распределения в сравнении со стандартным нормальным распределением (для него $E[\xi] = 0$) (рис. 5).

Однако это утверждение справедливо лишь для распределений, похожих на нормальные.

Глава XLII

ЗАКОНЫ БОЛЬШИХ ЧИСЕЛ И ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

Важную роль в теории вероятностей и приложениях играют утверждения, устанавливающие характер поведения больших совокупностей случайных величин. Дело в том, что отдельно взятая величина плохо прогнозируема — значения, принимаемые ею в конкретном эксперименте, случайны, и, даже зная ее закон распределения, мы не в состоянии предугадать результат. Замечательным свойством больших совокупностей является то их свойство, что при весьма неограничительных условиях они, в отличие от единичной случайной величины, ведут себя почти детерминировано: *с увеличением числа рассматриваемых случайных факторов суммарное воздействие, обусловленное этими факторами, становится все менее случайным.*

Важнейшими из утверждений, формализующими подобные, давно подмеченные практиками закономерности, являются законы больших чисел и предельные теоремы. Их обсуждению мы и посвятим настоящую главу.

§ 1. Законы больших чисел

Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — попарно независимые случайные величины, обладающие математическими ожиданиями $M\xi_i = m_i$ и дисперсиями $D\xi_i = \sigma_i^2$.

Теорема (закон больших чисел в форме Чебышёва).
Если дисперсии σ_i^2 равномерно ограничены, т. е. $\exists\ c$: $\forall\ i\ |\sigma_i^2| < c$, то $\forall\ \varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\frac{\xi_1 + \ldots + \xi_n}{n} - \frac{m_1 + m_2 + \ldots + m_n}{n}\right| \geqslant \varepsilon\right\} = 0. \qquad (1)$$

◄ Для доказательства предельного соотношения (1) заметим, что случайная величина η, задаваемая соотношением

$$\eta = \frac{1}{n}\sum_{i=1}^{n} \xi_i,$$

обладает математическим ожиданием $M[\eta] = \frac{1}{n}\sum\limits_{i=1}^{n} m_i$, дисперсией $D[\eta] = \frac{1}{n^2}\sum\limits_{i=1}^{n} \sigma_i^2$ и, следовательно, удовлетворяет неравенству Чебышёва

$$P\{|\eta - M\eta| \geqslant \varepsilon\} \leqslant \frac{D[\eta]}{\varepsilon^2}.$$

Из условия теоремы следует, что

$$D[\eta] = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2 \leqslant \frac{c}{n},$$

поэтому для любого фиксированного значения $\varepsilon > 0$ имеем

$$P\{|\eta - M\eta| \geqslant \varepsilon\} \leqslant \frac{c}{n\varepsilon^2}.$$

Переходя в последнем соотношении к пределу при $n \to \infty$, получаем искомое. ▶

Доказанная теорема устанавливает, что с ростом количества случайных величин их среднее арифметическое сколь угодно мало отличается от детерминированного, неслучайного воздействия, обусловленного средним арифметическим математических ожиданий.

Было бы неверно понимать соотношение (1) буквально, как утверждение о близости среднего арифметического случайных величин и среднего арифметического их математических ожиданий. Какое бы (может быть, очень маленькое) число ε мы ни взяли, ***в подавляющем большинстве*** экспериментов разница между η и $M[\eta]$ будет не больше ε, причем доля таких случаев тем больше, чем больше n. В приведенной формулировке закон больших чисел оставляет возможность такого «плохого» эксперимента, когда разница между η и $M[\eta]$ будет большой (больше ε), однако доля таких, «плохих» экспериментов тем меньше, чем больше n.

В приложениях (особенно в статистической физике) важную роль играет частный случай приведенной выше теоремы, относящийся к одинаково распределенным независимым слагаемым.

Следствие 1 (закон больших чисел в форме Чебышёва для одинаково распределенных слагаемых).
Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — независимые, одинаково распределенные случайные величины с $M[\xi_i] = m$ и $D[\xi_i] = \sigma^2$. Тогда

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - m\right| \geqslant \varepsilon\right\} = 0. \tag{2}$$

Замечание. Условие существования дисперсий может быть опущено (закон больших чисел в форме Хинчина).

Другим важным следствием теоремы Чебышёва является теорема Я. Бернулли.

Следствие 2 (закон больших чисел в форме Бернулли).
Пусть S_n — число успехов в серии из n независимых испытаний с постоянной вероятностью успеха p в каждом испытании и $\nu_n = S_n/n$ — относительная частота числа успехов. Тогда с увеличением количества экспериментов n в подавляющем большинстве случаев частота ν_n будет мало отличаться от вероятности. Точнее

$$\lim_{n\to\infty} P\{|\nu_n - p| \geqslant \varepsilon\} = 0. \tag{3}$$

◀ Доказательство следует из (2), если взять ξ_i, $i = 1, 2, \ldots, n$, такие, что $P\{\xi_i = 1\} = p$, $P\{\xi_i = 0\} = q$. Тогда $M\xi_i = p$, дисперсия величины ξ_i существует, равна pq и применима теорема Чебышёва, при этом

$$\frac{1}{n}\sum \xi_i = \frac{S_n}{n} = \nu_n \quad \text{и} \quad m = M\xi = p. \blacktriangleright$$

Теорема Чебышёва и ее следствия позволяют делать по результатам наблюдений за экспериментом достаточно надежные заключения о поведении случайных величин: дело в том, что относительная частота ($\nu_n = S_n/n$) события — вещь, определяемая экспериментально, а утверждение (3) связывает ее с вероятностью события и, тем самым, дает путь определения последней по результатам наблюдений.

При использовании законов больших чисел полезно иметь в виду следующее важное обстоятельство: они устанавливают близость относительной частоты (ν_n) и вероятности (p), но они не утверждают, что среднее число успехов (pn) мало отличается от наблюдаемого числа успехов (S_n)! Поясним эту мысль следующим примером.

Пример 1. Пусть производится эксперимент с физически симметричной монетой, так что вероятности выпадения герба и решки можно считать одинаковыми, $p = q = 1/2$.

Можно ли утверждать, ссылаясь на закон больших чисел, что при достаточно длительном экспериментировании число выпавших гербов (Г) и решек (Р) будет примерно одинаковым, т. е. $|S_n(\text{Г}) - S_n(\text{Р})| \to 0$ при $n \to \infty$?

Пусть ε — фиксированное положительное число. Закон больших чисел позволяет заключить, что при некотором, может быть, достаточно большом n_ε выполняется (в подавляющем большинстве случаев) неравенство

$$\left|\frac{S_n(\text{Г})}{n_\varepsilon} - \frac{S_n(\text{Р})}{n_\varepsilon}\right| < 2\varepsilon. \tag{4}$$

Отсюда

$$|S_n(\text{Г}) - S_n(\text{Р})| < 2\varepsilon n_\varepsilon.$$

Но, как мы увидим ниже (п. 5.2.1), величина $\varepsilon n_\varepsilon$ ведет себя как $c\sqrt{n_\varepsilon}$, где c — постоянная, и следовательно, в достаточно длинной серии бросаний *разность* количества выпавших гербов и решек симметричной монеты *может* стать сколь угодно большой!

Звкон больших чисел в этой ситуации позволяет только утверждать, что *отношение* количеств выпавших гербов и решек близко к единице,

$$\frac{S_n(\text{Г})}{S_n(\text{Р})} \sim 1.$$

Пример 2. Пусть монету бросили 100 раз и 100 раз выпал герб. Какой результат следует ожидать при 101-м испытании?

Теоретически возможны следующие три способа рассуждения в этой ситуации:

1. Монета симметрична, однако *случайно* так получилось, что выпало 100 гербов подряд. В силу независимости испытаний в 101 раз следует ожидать с равными вероятностями герб или решку.

2. Частота выпадения герба оказалась равной 1. Поэтому (закон больших чисел) монета скорее всего несимметрична и $P\{\text{Г}\} \gg P\{\text{Р}\}$. В силу независимости испытаний в 101 раз скорее всего выпадет герб.

3. Монета симметрична, однако *случайно* так получилось, что выпало 100 гербов подряд. Поскольку закон больших чисел утверждает, что частоты должны быть близки к вероятностям, а вероятность выпадения решки, в силу предположения о симметричности, равна 0,5, то и частота должна быть близка к 0,5. Поэтому чем больше гербов выпало, тем вероятнее, что (для исправления искажения в частоте выпадения решек!) появится решка.

◀ Внимательный анализ закона больших чисел позволяет в рассматриваемой ситуации сделать следующие выводы:

— первое умозаключение теоретически безупречно, однако гипотеза о симметричности маловероятна — чтобы симметричная монета сто раз подряд в независимых испытаниях выпала гербом, необходимо чтобы осуществилось событие с вероятностью 2^{-100}. Поэтому скорее всего следует результат объяснить несимметричностью монеты;

— второе умозаключение теоретически безупречно и практически приемлемо;

— третье умозаключение просто неверно, так как неявно предполагает наличие у монеты «памяти» — вероятность выпадения решки связывается с количеством решек, выпавших до рассматриваемого испытания. ▶

Завершая обсуждение законов больших чисел, рассмотрим еще один пример, качественно иллюстрирующий логику их применения.

Пример 3 (правило среднего арифметического).

Пусть a — некоторая величина, определяемая в эксперименте путем измерений. Каждое измерение X_i, $i = 1, 2, \dots, n$, складывается из значения a измеряемой величины и погрешности измерения ξ_i

$$X_i = a + \xi_i, \quad i = 1, 2, \dots, n.$$

Практиками давно установлено, что для определения измеряемой величины a следует найти среднее арифметическое измерений

$$a \approx \bar{x} = \frac{1}{n} \sum_{i=1}^{n} X_i. \tag{5}$$

◀ Объяснение этому правилу дает закон больших чисел. Если предположить, что систематическая ошибка измерений отсутствует, т. е. $M\xi_i = 0$, то из соотношения (5) следует

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n}(a + \xi_i) = a + \frac{1}{n} \sum_{i=1}^{n} \xi_i, \tag{6}$$

и для последнего слагаемого из закона больших чисел получаем

$$P\left\{\left|\frac{1}{n} \sum \xi_i\right| > \varepsilon\right\} \to 0, \quad n \to \infty.$$

Тем самым, в подавляющем большинстве экспериментов ошибка измерений $\Delta = |\bar{x} - a| = \frac{1}{n}\left|\sum\limits_{i=1}^{n} \xi_i\right|$ может быть неограничено уменьшена за счет дублирования измерений.

Ниже (п. 5.2.2) мы уточним этот качественный результат количественно. ▶

Заметим в заключение, что законы больших чисел (в приведенных выше формулировках) утверждают устойчивость средних арифметических больших совокупностей случайных величин, в частности (в формулировке Бернулли), устойчивость частот. Это теоретические утверждения являются следствием постулатов, положенных в основание теории. Очевидно, что в реальной практической ситуации подобная устойчивость не обязана иметь место! Применимость же теории вероятностей к описанию различных явлений в природе и обществе базируется именно на наличии в этих явлениях статистической устойчивости.

Важно понимать, что подобная устойчивость должна быть установлена исследователем (или, по крайней мере, продекларирована) для того, чтобы теоретико-вероятностные выводы имели смысл. Наличие теорем типа законов больших чисел не может служить основанием для утверждений о наличии статистической устойчивости в той или иной конкретной ситуации.

§ 2. Предельные теоремы

Как следует из результатов предыдущего пункта, среднее арифметическое большой совокупности случайных величин при определенных условиях ведет себя почти детерминировано — мало отличается от среднего арифметического их математических ожиданий. Это утверждение носит качественный характер и в практической ситуации не всегда содержательно. Точный ответ на вопросы «При каких n...?» и «Насколько и как часто отличается совокупное *случайное* среднее от *неслучайного* среднего математических ожиданий?» в рамках законов больших чисел (в приведенных формулировках) получить нельзя.

Развитая выше теория (гл. XL, п. 3.2) говорит, что необходимо знание распределения суммы $\left(\sum\limits_{i=1}^{n} \xi_i\right)$ или, что то же, средней суммы $\left(\frac{1}{n} \sum\limits_{i=1}^{n} \xi_i\right)$. Однако, как уже отмечалось, задача нахождения закона распределения суммы случайных величин в общем случае — это довольно сложная в теоретическом плане и громоздкая в плане технической реализации задача, требующая знания законов распределения слагаемых. Так что для случайных величин, законы распределений которых «плохо» сворачиваются, получение распределений их сумм для достаточно больших значений n задача, скорее всего, практически нереализуемая. Впрочем, и для «хорошо» сворачивающихся распределений работа с распределением суммы большого числа слагаемых может оказаться технически затруднительной.

Спасти ситуацию может только чудо, и таким чудом в теории вероятностей как раз и являются предельные теоремы.

2.1. Теорема Муавра—Лапласа

Мы начнем изучение ситуации с довольно простого на первый взгляд, но важного для приложений случая, когда все ξ_i — независимые, одинаково распределенные случайные величины, принимающие значения 1 и 0 с вероятностями p и q соответственно, $p+q=1$. В этом случае $\sum\limits_{i=1}^{n}\xi_i = S_n$ — биномиальная случайная величина с параметрами (n, p), распределение которой нам хорошо известно

$$F(x) = P\{S_n < x\} = \sum_{k=0}^{k<x} C_n^k p^k q^{n-k}. \tag{1}$$

Вероятности, связанные с суммой S_n, в принципе легко могут быть найдены для

$$P\{a \leqslant S_n \leqslant b\} = \sum_{k=a}^{b} C_n^k p^k q^{n-k}. \tag{2}$$

Однако уже для небольших значений n воспользоваться соотношением (2) довольно трудно — вычисления оказываются очень громоздкими.

Пример 1. Сколько раз следует подбросить симметричную монету, чтобы с надежностью $\varkappa$, не худшей чем 0,99, частота появления герба отличалась от вероятности не больше, чем на 0,01?

◀ Для получения ответа на вопрос задачи следует решить неравенство

$$P\left\{\left|\frac{S_n}{n} - 0{,}5\right| < 0{,}01\right\} \geqslant 0{,}99$$

($p = 0{,}5$; $\varepsilon = 0{,}01$) относительно n. Использование соотношения (2) приводит к неравенству

$$\left(\sum_{0{,}49n}^{0{,}51n} C_n^k\right)(0{,}5)^n \geqslant 0{,}99, \tag{3}$$

для которого проверка, удовлетворяет ли ему, к примеру, число $n = 100$, — довольно утомительная вычислительная задача. ▶

К счастью, для больших значений n может быть указано сравнительно простое правило вычисления вероятностей (1)–(2).

Сначала заметим, что если $n \to \infty$, то выражение для индивидуальных вероятностей $P\{S_n = k\}$ может быть упрощено за счет замены факториалов приближенным выражением по известной из анализа формуле Стирлинга

$$n! = \sqrt{2\pi n}\cdot n^n e^{-n}\cdot(1+\Theta_n), \tag{4}$$

где $\Theta_n \to 0$ при $n \to \infty$. Действительно, легко убедиться в том, что

$$\begin{aligned} P\{S_n = k\} &= C_n^k p^k q^{n-k} = \frac{n!}{k!(n-k)!}\,p^k q^{n-k} = \\ &= \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{k(n-k)}}\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}(1+\Theta_{n,k}), \end{aligned} \tag{5}$$

где $\Theta_{n,k} \to 0$ при $n \to \infty$.

Ниже запись

$$A(x) \cong B(x)$$

будет обозначать, что $\lim\limits_{x \to x_0} A/B = 1$ или, что то же,

$$A(x) = B(x)(1 + \Theta),$$

где $\Theta \to 0$ при $x \to x_0$.

В силу закона больших чисел, при больших значениях n сумма S_n с заметной вероятностью принимает только те значения k, которые удовлетворяют условию $|k - np| \leqslant n\varepsilon$. Полагая

$$\Delta_k = k - np, \tag{6}$$

заключаем, что $n \to \infty$ влечет за собой $k \to \infty$ и $\Delta_k/n \to 0$. Подставим выражения для $np = k - \Delta_k$ и $nq = n - k + \Delta_k$ в формулу (5) и получим, что

$$P\{S_n = k\} \cong \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{k(n-k)}}\left(1 - \frac{\Delta_k}{k}\right)^k \left(1 + \frac{\Delta_k}{n-k}\right)^{n-k}. \tag{7}$$

Отсюда

$$\sqrt{\frac{n}{k(n-k)}} = \left(\frac{np}{k}\right)^{1/2}\left(\frac{nq}{n-k}\right)^{1/2}\frac{1}{\sqrt{npq}} \cong \frac{1}{\sqrt{npq}},$$

$$\left(1 - \frac{\Delta_k}{k}\right)^k = \exp\left\{k \ln\left(1 - \frac{\Delta_k}{k}\right)\right\} \cong \exp\left\{-\Delta_k - \frac{1}{2}\,\frac{\Delta_k^2}{k}\right\},$$

$$\left(1 + \frac{\Delta_k}{n-k}\right)^{n-k} = \exp\left\{(n-k)\ln\left(1 + \frac{\Delta_k}{n-k}\right)\right\} \cong \exp\left\{\Delta_k - \frac{1}{2}\,\frac{\Delta_k^2}{n-k}\right\}.$$

С учетом полученных соотношений равенство (7) примет вид

$$P\{S_n = k\} \cong \frac{1}{\sqrt{2\pi}}\,\frac{1}{\sqrt{npq}}\,\exp\left\{-\frac{1}{2}\,\Delta_k^2\,\frac{n}{k(n-k)}\right\}$$

или, после замены $\frac{n}{k(n-k)}$ на $\frac{1}{npq}$ и Δ_k на $k - np$,

$$P\{S_n = k\} \cong \frac{1}{\sqrt{2\pi}}\,\frac{1}{\sqrt{npq}}\,\exp\left\{-\frac{1}{2}(k - np)^2\frac{1}{npq}\right\}. \tag{8}$$

Последнее соотношение представляет собой приближенную формулу для подсчета индивидуальных биномиальных вероятностей и составляет содержание *локальной предельной теоремы Муавра—Лапласа.* Она пригодна в ситуациях, когда p и q не слишком малы и не слишком близки к единице (в этом случае, как мы знаем, можно использовать приближение Пуассона), а n достаточно велико. Практически же, если $np > 10$, $0{,}1 < p < 0{,}9$, то вычисление индивидуальных биномиальных вероятностей $P\{S_n = k\}$ по формуле (8) дает вполне приемлемые результаты.

Пример 1. Симметричную монету бросили 20 раз. Какова вероятность того, что герб появится ровно 7 раз?

◀ S_{20} — количество появлений герба в рассматриваемом эксперименте, $S_{20} = B[20; 0{,}5]$. Поэтому

$$P\{S_{20} = 7\} = C_{20}^7(0{,}5)^{20} = 0{,}074.$$

Формула (8) дает: $np = 20 \cdot 0{,}5 = 10$; $npq = 5$; $\sqrt{npq} = 2{,}236$

$$P\{S_{20} = 7\} \cong \frac{1}{\sqrt{2\pi}}\,\frac{1}{2{,}236}\,\exp\left\{-\frac{1}{2}(7 - 10)^2 \cdot \frac{1}{5}\right\} = 0{,}0727,$$

что практически неотличимо от точного результата. ▶

Для решения задачи приближенного вычисления **совокупных биномиальных вероятностей** (1)–(2) рассмотрим соотношение (2) с учетом выражения (8) для индивидуальных биномиальных вероятностей $P\{S_n = k\}$:

$$P\{a \leqslant S_n \leqslant b\} = \sum_{k=a}^{b} P\{S_n = k\} \cong \sum_{k=a}^{b} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{npq}} \exp\left\{-\frac{(k-np)^2}{2npq}\right\}. \tag{9}$$

Положим

$$x_k = \frac{k-np}{\sqrt{npq}}; \quad \Delta x_k = x_{k+1} - x_k = \frac{1}{\sqrt{npq}}; \quad x_a = \frac{a-np}{\sqrt{npq}}; \quad x_b = \frac{b-np}{\sqrt{npq}}.$$

Заметим, что при неограниченном увеличении n промежуток $[x_a, x_b]$ изменения переменной x передвигается вдоль числовой прямой неограниченно вправо, при этом его длина $l_n = x_b - x_a = \frac{b-a}{\sqrt{npq}}$ неограниченно уменьшается. Поэтому, вообще говоря, нет смысла пытаться вычислять сумму (9) при фиксированных a и b — с увеличением n эта сумма стремится к нулю.

Пусть x_a и x_b фиксированы. Тогда в принятых обозначениях сумма (9) запишется в виде

$$P\{a \leqslant S_n \leqslant b\} \cong \sum_{k=a}^{b} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x_k^2}{2}\right\} \Delta x_k.$$

При $n \to \infty$ выражение справа можно интерпретировать как интегральную сумму для интеграла

$$I = \int_{x_a}^{x_b} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx,$$

что приводит к следующему утверждению, впервые установленному Муавром для $p = q = 1/2$ и обобщенному Лапласом на случай произвольных p.

Теорема (интегральная теорема Муавра—Лапласа).
Пусть S_n — биномиальная случайная величина с параметрами n и p $(p \neq 0,\ p \neq 1)$. Тогда при $n \to \infty$ равномерно относительно $-\infty < x_a < x_b < +\infty$ справедливо соотношение

$$P\left\{x_a \leqslant \frac{S_n - np}{\sqrt{npq}} \leqslant x_b\right\} \to \frac{1}{\sqrt{2\pi}} \int_{x_a}^{x_b} \exp\left\{-\frac{x^2}{2}\right\} dx. \tag{10}$$

Предельное соотношение (10) является источником формул приближенного вычисления совокупных биномиальных вероятностей.

В самом деле, пусть n достаточно велико, так что (10) имеет место. Тогда

$$P\{a \leqslant S_n \leqslant b\} = P\left\{\frac{a-np}{\sqrt{npq}} \leqslant \frac{S_n - np}{\sqrt{npq}} \leqslant \frac{b-np}{\sqrt{npq}}\right\} \approx F\left(\frac{b-np}{\sqrt{npq}}\right) - F\left(\frac{a-np}{\sqrt{npq}}\right), \tag{11}$$

где $F(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} \exp\left\{-\frac{1}{2}\,x^2\right\} dx$ — функция Лапласа.

Уже при относительно небольших значениях n для средних (далеких от 0 и 1) значений p и q точность приближенного соотношения (11) достаточно высока. Приведенные ниже данные иллюстрируют порядок точности при замене совокупных биномиальных вероятностей по формуле (11) для различных значений n, p и k. В каждой ячейке таблицы в числителе приведено точное значение вероятности $P\{S_n \leqslant k\}$, а в знаменателе — ее приближенное значение для соответствующих n, p и k.

Теперь мы можем уточнить закон больших чисел в форме Бернулли и получить не только качественные, но и содержательные количественные заключения.

Поскольку
$$P\left\{\left|\frac{S_n}{n} - p\right| < \varepsilon\right\} = P\left\{\frac{|S_n - np|}{\sqrt{npq}} < \frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right\},$$
то, применяя соотношение (11) к последней вероятности, получаем
$$P\left\{\left|\frac{S_n}{n} - p\right| < \varepsilon\right\} \approx 2F\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) - 1. \tag{12}$$
Формула (12) позволяет получать конкретные ответы на конкретные вопросы «Как часто?», «С какой точностью?», «При каком n?».

n	k	p			n	k	p		
		0,2	0,4	0,7			0,2	0,4	0,7
20	10	0,804 0,711	0,126 0,085	0,000043 0	25	5	0,617 0,5	0,029 0,021	0 0,0005
	15	0,999 0,999	0,872 0,819	0,048 0,0255		10	0,994 0,994	0,586 0,5	0,0078 0,137
	20	1,0 1,0	0,999 0,999	0,762 0,687		20	1,0 1,0	0,999 0,999	0,909 0,862
50	20	0,999 0,999	0,561 0,5	0,0 0,0	100	40	0,999 1,0	0,543 0,5	0,0 0,0
	35	1,0 1,0	0,999 0,999	0,553 0,5		60	1,0 1,0	0,999 0,999	0,021 0,015
	40	1,0 1,0	1,0 1,0	0,960 0,939		80	1,0 1,0	1,0 1,0	0,991 0,986

Пример 1. Симметричную монету бросили 100 раз. **Как часто** число выпавших гербов будет отличаться от среднего не более чем на 10?

◀ Это прямой вопрос о величине вероятности
$$P\{|S_n - np| < 10\}$$
для $n = 100$ и $p = 0{,}5$. В соответствии с (12) имеем
$$P\{|S_n - np| < 10\} = 2F(2) - 1 \approx 0{,}9545$$
— в подавляющем большинстве случаев при 100 бросаниях симметричной монеты следует ожидать, что число выпавших гербов не меньше 40 и не больше 60. Отметим, что точное значение искомой вероятности равно 0,9540. ▶

Пример 2. Насколько большие отклонения частоты появления герба (в эксперименте с симметричной монетой) от 0,5 можно ожидать в подавляющем большинстве случаев, если монету бросили 50 раз? 100 раз?

◀ Вопрос задачи, это вопрос о величине ε, удовлетворяющей условию
$$P\left\{\left|\frac{S_n}{n} - p\right| < \varepsilon\right\} \geqslant \varkappa,$$
где $0 < \varkappa < 1$ — величина вероятности, отвечающая нашим представлениям о том, как понимать «в подавляющем большинстве случаев». Конечно, хотелось бы положить $\varkappa = 1$, однако в этом случае ответ на вопрос тривиален: ε — любое число, не меньшее 1. Для получения нетривиальной информации о точности поступимся надежностью. Возьмем $\varkappa$ близкое к единице настолько, чтобы событиями с вероятностью, меньшей $1 - \varkappa$, можно было пренебречь. Это, правда, даст нам возможность получить

оценку для ε, справедливую уже не для всех экспериментов, а только для $\varkappa$-доли из них. Но поскольку такие эксперименты достаточно часты ($\varkappa$ взята близкой к единице!), постольку полученная информация о величине ε оказывается информативной. Взяв, например, $\varkappa = 0{,}9$, получим для определения ε неравенство

$$P\left\{\left|\frac{S_n}{n} - 0{,}5\right| < \varepsilon\right\} \geqslant 0{,}9.$$

С учетом (12) это приводит к соотношению

$$2F\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) - 1 \geqslant 0{,}9$$

или

$$F\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) \geqslant 0{,}95.$$

Полагая $n = 50,\ 100,\ p = q = 0{,}5$ и учитывая монотонное неубывание функции Лапласа, получаем

$$\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \geqslant F^{-1}(0{,}95) = 1{,}6449,$$

откуда

$$\varepsilon \approx \frac{F^{-1}(0{,}95)\cdot\sqrt{pq}}{\sqrt{n}}. \tag{13}$$

Для $n = 50$ ε не меньше $0{,}126$, для $n = 100$ — $0{,}082$. Из (13) очевидно, что с ростом n ожидаемые отклонения (при фиксированной надежности $\varkappa$) убывают пропорционально $\sqrt{n}$.

Для заданной надежности при $n = 100$ точное решение поставленной задачи $\varepsilon \approx 0{,}080$. ▶

Пример 3. Сколько раз следует бросить монету, чтобы не менее чем в 99 случаях из 100 наблюдаемая частота выпадения герба отличалась от вероятности не более чем на $0{,}01$?

◀ Этот вопрос о количестве экспериментов, необходимых для оценивания *неизвестной* вероятности выпадения герба из неравенства

$$P\left\{\left|\frac{S_n}{n} - p\right| < \varepsilon\right\} \geqslant 0{,}99.$$

Соотношение (12) дает

$$2F\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) - 1 \geqslant 0{,}99 \Rightarrow F\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) \geqslant 0{,}995.$$

Заметим, что в силу неравенства

$$\frac{p+q}{2} \geqslant \sqrt{pq}$$

можно заключить, что

$$\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \geqslant 2\varepsilon\sqrt{n}.$$

В силу монотонности функции Лапласа для любых p и q

$$F\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) > F(2\varepsilon\sqrt{n}).$$

Поэтому, если n удовлетворяет условию

$$F(2\varepsilon\sqrt{n}) \geqslant 0{,}995,$$

то подавно

$$F\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) > F(2\varepsilon\sqrt{n}) \geqslant 0{,}995.$$

Определив n из последнего неравенства, мы заведомо решим поставленную задачу, может быть, и несколько завысив необходимое число экспериментов.

Несложные выкладки дают оценку

$$F(2\varepsilon\sqrt{n}) \geqslant 0{,}995 \Rightarrow 2\varepsilon\sqrt{n} \geqslant F^{-1}(0{,}995) = 2{,}5758,$$

откуда для n получаем

$$n \geqslant \left(\frac{2{,}5758}{2\varepsilon}\right)^2 = 16586.$$

Столь большое количество экспериментов объясняется высокими требованиями к надежности, заданными условиями задачи. Если требование к надежности несколько снизить, то и оценка для n уменьшится. Так для $\varkappa = 0{,}9$ $n \sim 6764$, а для $\varkappa = 0{,}8$ $n \sim 4106$.

В заключение отметим, что в реальных экспериментах по оценке симметричности монеты наблюдались следующие результаты:

эксперимент Бюффона	—	4040 бросаний,	$S_n = 2048$,	$\bar{p} = 0{,}5069$;
эксперимент Пирсона I	—	12000 бросаний,	$S_n = 6019$,	$\bar{p} = 0{,}5016$;
эксперимент Пирсона II	—	24000 бросаний,	$S_n = 12012$,	$\bar{p} = 0{,}5005$. ▶

2.2. Теорема Ляпунова

Теорема Муавра—Лапласа помимо возможности вычисления индивидуальных и совокупных биномиальных вероятностей предоставляет нам еще возможность по новому взглянуть на «взаимоотношение» различных классов случайных величин и их распределений. Действительно, соотношение (10) предыдущего пункта в форме

$$P\left\{\frac{S_n - np}{\sqrt{npq}} < x\right\} \approx F(x) \tag{14}$$

может быть прочитано как ***похожесть распределения центрированной и нормированной суммы независимых бернуллиевых случайных величин на нормальное распределение***, что достаточно удивительно.

Еще более удивительным и неожиданным оказывается тот факт, что если взять не бернуллиевы, а **любые другие** случайные величины, то при весьма необременительных требованиях центрированная и нормированная их сумма будет иметь распределение, близкое к **нормальному**!

Точные утверждения, формализующие это наблюдение, носят в теории вероятностей название **центральных предельных теорем**, формулировке одной из которых мы посвятим этот пункт.

Может быть доказана следующая

Теорема. *Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — независимые случайные величины с конечными математическими ожиданиями $M\xi_i = m_i$, дисперсиями $D\xi_i = \sigma_i^2$, третьими моментами $M|\xi_i - m_i|^3 = \tau_i^3$ и выполнено условие*

$$\lim_{n\to\infty} \frac{1}{\left(\sum\limits_{i=1}^{n} \sigma_i^2\right)^{3/2}} \sum_{i=1}^{n} \tau_i^3 = 0. \tag{15}$$

Тогда

$$P\left\{\frac{\sum \xi_i - \sum m_i}{\sqrt{\sum \sigma_i^2}} < x\right\} \to F(x) \tag{16}$$

равномерно относительно x.

Отметим, что теорема не предъявляет никаких требований к **законам распределения слагаемых**! Важное условие (15) может быть интерпретировано как требование «приблизительной одинаковости» слагаемых.

Понятно, что без этого (или другого подобного) условия теорема перестает быть верной, ибо доминирующее слагаемое может подавить прочие слагаемые в сумме и тем самым будет определять ее распределение.

Для практических приложений весьма важен случай одинаково распределенных величин. В этом случае условия теоремы упрощаются — можно доказать, что **если**

все ξ_i — независимые одинаково распределенные случайные величины с конечными математическим ожиданием m и дисперсией σ^2, то (16) справедливо без дополнительного условия (15).

Теорема Ляпунова указывает на ту особую роль, которую в теории вероятностей и ее приложениях играет закон нормального распределения — сумма большого числа примерно одинаковых случайных слагаемых имеет почти нормальное распределение. Поэтому на практике, когда приходится изучать воздействия, обусловленные большим числом факторов, принимают обычно гипотезу о нормальном характере суммарного воздействия[1].

Как и теорема Муавра—Лапласа, теорема Ляпунова может служить источником формул для приближенного вычисления вероятностей, связанных с большими суммами случайных величин.

Пример 1. Получим аналог соотношения (12) предыдущего пункта для закона больших чисел в форме Чебышёва.

◀ Имеем

$$P\left\{\left|\frac{1}{n}\sum \xi_i - \frac{1}{n}\sum m_i\right| < \varepsilon\right\} = P\left\{\frac{|\sum \xi_i - \sum m_i|}{\sqrt{\sum \sigma_i^2}} < \frac{\varepsilon n}{\sqrt{\sum \sigma_i^2}}\right\}.$$

В силу соотношения (16) для последней вероятности при достаточно больших n получим

$$P\left\{\frac{|\sum \xi_i - \sum m_i|}{\sqrt{\sum \sigma_i^2}} < \frac{\varepsilon n}{\sqrt{\sum \sigma_i^2}}\right\} \approx 2F\left(\frac{n\varepsilon}{\sqrt{\sum \sigma_i^2}}\right) - 1. \tag{17}$$

Это и есть искомое соотношение для вычисления вероятностей уклонений случайных средних от средних математических ожиданий. В случае одинаково распределенных слагаемых с конечными $M\xi_i = m$ и дисперсиями $D\xi_i = \sigma^2$ соотношение (17) принимает вид

$$P\left\{\left|\frac{1}{n}\sum \xi_i - m\right| < \varepsilon\right\} \approx 2F\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - 1. \blacktriangleright \tag{18}$$

Пример 2 (метод Монте-Карло вычисления интегралов).
Пусть требуется вычислить интеграл

$$I = \int_0^1 f(x)\,dx. \tag{19}$$

Рассмотрим случайную величину $\xi = R[0, 1]$ и заметим, что интеграл (17) может быть представлен как $Mf(\xi)$. Поэтому, если $X_1, X_2, \ldots, X_n$ — реализации случайной величины ξ, то, в силу закона больших чисел

$$I \approx \frac{1}{n}\sum_{i=1}^{n} f(X_i). \tag{20}$$

Для оценки точности приближенного соотношения (20) нужно уметь оценивать величину

$$\Delta_n = I - \frac{1}{n}\sum_{i=1}^{n} f(X_i),$$

а для этого, в свою очередь, следует уметь решать неравенство

$$P\{|\Delta_n| < \varepsilon\} \geqslant \varkappa \tag{21}$$

относительно $\varepsilon > 0$ при заданных значениях $\varkappa$ и n. Если n достаточно велико, то в основу (21) может быть положено соотношение (18).

Действительно, $f(X_i)$, $i = 1, 2, \ldots, n$, — независимые, одинаково распределенные случайные величины, для которых $Mf(X_i) = I$. Поэтому

$$P\{|\Delta_n| < \varepsilon\} \approx 2F\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - 1, \tag{22}$$

[1] Часто и тогда, когда для этого нет никаких оснований. Отметим, что как и выше, в случае с устойчивостью частот, теоретический факт — в данном случае теорема Ляпунова — может послужить только наводящим соображением о возможной нормальности, но не может служить доказательством наличия нормальности.

где $\sigma^2 = Df(\xi)$.

Для оценки величины σ^2 заметим, что в силу интегрируемости функция $f(x)$ ограничена — $|f(x)| < \alpha$ — и поэтому $\sigma^2 = Df(\xi) = M[f(X_i) - I]^2 \leqslant 4\alpha^2$. Отсюда, как и в примере 3 предыдущего пункта,

$$F\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) \geqslant F\left(\frac{\varepsilon\sqrt{n}}{2\alpha}\right)$$

и из $2F\left(\frac{\varepsilon\sqrt{n}}{2\alpha}\right) - 1 \geqslant \varkappa$ следует, что заведомо и

$$2F\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - 1 \geqslant \varkappa.$$

При фиксированном (и близком к единице) $\varkappa$ для ε имеем

$$\varepsilon \sim F^{-1}\left(\frac{1+\varkappa}{2}\right) \cdot \frac{2\alpha}{\sqrt{n}}.$$

Таким образом, увеличивая n, мы в подавляющем большинстве случаев гарантируем (с надежностью $\varkappa$) точность ε в вычислении интеграла (19) по формуле (20). ▶

Задачи к разделу

1. Построить пространство элементарных событий для эксперимента: одновременное бросание трех игральных костей. Описать в терминах элементарных исходов следующие случайные события:

$A = \{$сумма очков равна 9$\}$.

$B = \{$выпали одинаковые цифры на всех костях$\}$.

$C = \{$хотя бы на одной кости выпала единица$\}$.

Совместны ли события A и B? A и C? B и C? A, B и C? Зависимы ли события A и B? A и C? B и C? Найти вероятность того, что по крайней мере на одной кости выпала 3, если известно, что сумма выпавших очков равна 9.

2. Вероятность выбора точки из плоской области, лежащей внутри круга радиуса 1 с центром в начале координат, принимается пропорциональной кубу длины проекции этой области на ось абсцисс. Согласуется ли такое правило подсчета вероятностей с требованиями, предъявляемыми к вероятности?

3. Из множества точек, расположенных на отрезке $[0, 1]$, случайным образом выбирают одну, причем полагают вероятность выбора из подотрезка $[\alpha, \beta]$ пропорциональной $\sin\beta - \sin\alpha$. Согласуется ли такой способ подсчета вероятностей с требованиями, предъявляемыми к вероятности?

4. Вероятность выбора точки из дуги полуокружности радиуса 1, с центром в начале координат, расположенной в верхней полуплоскости, принимается пропорциональной длине проекции этой дуги на ось OY. Согласуется ли такое правило подсчета вероятностей с требованиями, предъявляемыми к вероятности?

5. Из множества точек, расположенных на отрезке $[0, 1]$, случайным образом выбирают одну, причем полагают вероятность выбора из подотрезка $[\alpha, \beta]$ пропорциональной $\operatorname{arctg}\beta - \operatorname{arctg}\alpha$. Согласуется ли такой способ подсчета вероятностей с требованиями, предъявляемыми к вероятности?

6. Вероятность выбора точки из дуги полуокружности радиуса 1, с центром в начале координат, расположенной в верхней полуплоскости, принимается пропорциональной синусу длины проекции этой дуги на ось OX. Согласуется ли такое правило подсчета вероятностей с требованиями, предъявляемыми к вероятности?

7. Вероятность выбора точки из дуги полуокружности радиуса 1, с центром в начале координат, расположенной в верхней полуплоскости, принимается пропорциональной синусу длины проекции этой дуги на ось OX. Согласуется ли такое правило подсчета вероятностей с требованиями, предъявляемыми к вероятности?

8. Указать ошибку в рассуждениях: агрегат выходит из строя, если выходит из строя по крайней мере один из его узлов A или B. Узел A выходит из строя с вероятностью p_1, B — с вероятностью p_2. Найдем вероятность выхода из строя агрегата. Событие $\mathcal{U} = \{$ агрегат вышел из строя$\}$ является объединением событий $A = \{$ узел A вышел из строя$\}$ и $B = \{$ узел B вышел из строя$\}$, $\mathcal{U} = A \cup B$. Поэтому: $P(\mathcal{U}) = P(A) + P(B) = p_1 + p_2$. Пусть, к примеру, $P(A) = p_1 = 0{,}4$; $P(B) = p_2 = 0{,}6$, тогда $P(\mathcal{U}) = 1$, т.е. агрегат наверняка не работает.

9. Вероятность попадания точки $M(x, y)$ в область D, лежащую в квадрате $K = \{(x, y) : |x| \leqslant 1,\ |y| \leqslant 1\}$, пропорциональна площади этой области: $P(M \in D) = \frac{1}{4}S(D)$. Найти $P\{|x - y| < 1/2\}$.

10. Вероятность попадания точки $M(x, y)$ в область D, лежащую в квадрате $K = \{|x| \leqslant 1, |y| \leqslant 1\}$, пропорциональна площади этой области $P\{M \in D\} = \frac{1}{4}S(D)$. Найти P $\{\min(x, y) < 1/2\}$.

11. Вероятность попадания точки $M(x, y)$ в область D, лежащую в квадрате $K = \{|x| \leqslant 1, |y| \leqslant 1\}$, пропорциональна площади этой области $P\{M \in D\} = \frac{1}{4}S(D)$. Найти $P\{xy < 0{,}5\}$.

12. Вероятность попадания точки $M(x, y)$ в область D, лежащую в квадрате $K = \{(x, y)\colon |x| \leqslant 1, |y| \leqslant 1\}$, пропорциональна площади этой области $P\{D \ni M\} = \frac{1}{4}S(D)$. Найти P $\{\max(x, y) < 1/2\}$.

13. Пусть (a, b) — координаты случайной точки в квадрате $K = \{(a, b)\colon |a| \leqslant 1, |b| \leqslant 1\}$. Найти вероятность того, что корни уравнения $x^2 + ax + b = 0$ действительны.

14. Пусть a, b координаты случайной точки в квадрате $K = \{(a, b)\colon |a| \leqslant 1, |b| \leqslant 1\}$. Найти вероятность того, что корни уравнения $x^2 + ax + b = 0$ положительны.

15. Вероятность попадания точки на дугу окружности $x^2 + y^2 = 1$ пропорциональна длине дуги D: $P\{M \in D\} = 1/(2\pi)L(D)$. Пусть (x, y) — координаты точки. Найти $P\{|x + y| \leqslant 1\}$.

16. В урне находится некоторое количество шаров, причем известно, что белых вдвое больше, чем черных. Из урны извлекают пару шаров. При этом $7P$ {извлечены одноцветные шары} $= 8P$ {извлечены разноцветные шары}. Можно ли по этим данным установить количество шаров, лежащих в урне?

17. Известно, что при двукратном бросании монеты вероятность наблюдения одноименных сторон больше вероятности наблюдения разноименных сторон этой монеты. Установить, не противоречат ли эти данные предположению о симметричности монеты.

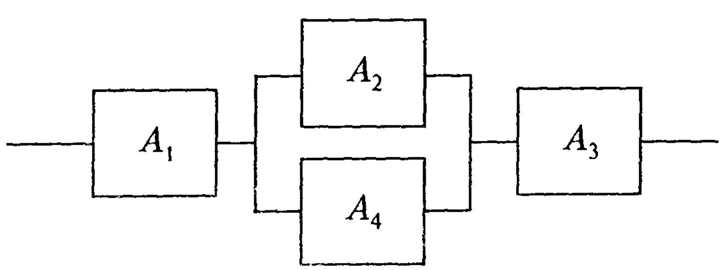
18. В урну, в которой было одинаковое количество белых и черных шаров, добавили один шар. При этом вероятность извлечения из урны пары разноцветных шаров изменилась на 1/48. Выяснить, уменьшилась или увеличилась эта вероятность, и установить, какое количество шаров было в урне.

19. В урне лежит 4 шара. Известно, что при извлечении из урны двух шаров (без возвращения) вероятность извлечения двух одноцветных шаров вдвое меньше вероятности извлечения разноцветной пары. Можно ли по этим данным установить состав шаров в урне?

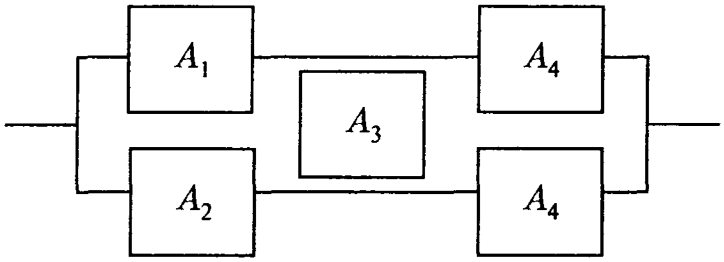
20. В урну, в которой было одинаковое количество белых и черных шаров, добавили один шар. При этом вероятность извлечения из урны пары разноцветных шаров изменилась на 1/15. Выяснить, уменьшилась или увеличилась эта вероятность, и установить, какое количество шаров было в урне.

21. Из урны, содержащей всего 6 белых и черных шаров, извлекают одновременно два шара. Известно, что вероятности извлечения одноцветных и разноцветных шаров относятся как 7 : 8. Можно ли по этим данным установить состав шаров в урне?

22. Надежность (т. е. вероятность безотказной работы в течение некоторого времени T) каждого из элементов A_i, $i = 1, 2, 3, 4$, объединенных в релейную схему (см. рис.), равна p_i. Найти надежность схемы.



23. Надежность (т. е. вероятность безотказной работы в течение некоторого времени T) каждого из элементов A_i, $i = 1, 2, 3, 4, 5$, объединенных в релейную схему (см. рис.), равна p_i. Найти надежность схемы.



24. Случайная величина ξ имеет пуассоновское распределение и известно, что ее математическое ожидание m и дисперсия D связаны соотношением $D^2 - m = 2$. Найти вероятность $P(\xi > 3)$.

25. Случайная величина ξ имеет пуассоновское распределение и известно, что $P\{\xi = 0\} = e^{-2}$. Найти математическое ожидание случайной величины $\xi^3 + \xi$.

26. Случайная величина ξ имеет пуассоновское распределение и известно, что $P(\xi > 0) = 0{,}8$. Найти математическое ожидание случайной величины $2\xi^2 - 5\xi^3$.

27. Случайная величина ξ имеет экспоненциальное распределение с математическим ожиданием m и дисперсией D. Известно, что $2D - 3m = 2$. Найти вероятность того, что $-2 \leqslant \xi \leqslant 2$.

28. Случайная величина ξ экспоненциально распределена и известно, что $P(\xi > 1) = 0{,}3679$. Найти математическое ожидание случайной величины $4\xi^2 + 4\xi + 4$.

29. Случайная величина ξ равномерно распределена на промежутке с центром в -3 и известно, что $P(\xi > -2) = 0{,}25$. Найти математическое ожидание случайной величины $\eta = 2 - 4\xi - 15\xi^2$.

30. Случайная величина ξ равномерно распределена на промежутке и известно, что ее математическое ожидание m и среднеквадратичное отклонение σ связаны соотношением

$$\begin{cases} 3m + 2\sqrt{3}\sigma = 4, \\ m = \sqrt{3}\sigma - 2. \end{cases}$$

Найти вероятность $P\{\xi \geqslant 1\}$.

31. Функция распределения случайной величины ξ непрерывна в $\mathbf{R}$ и дана соотношением

$$F_\xi(x) = \begin{cases} 0, & x \leqslant -2, \\ \alpha x^2 + \beta x + \dfrac{\beta^2}{4\alpha}, & -2 \leqslant x \leqslant 1, \\ 1, & x \geqslant 1. \end{cases}$$

Найти параметры α, β, плотность случайной величины ξ, математическое ожидание случайной величины $Y = 2\xi + 1$, дисперсию случайной величины $Y_1 = -3\xi - 4$ и вероятность того, что $\xi > 0{,}5$.

32. Плотность распределения случайной величины ξ дана соотношением

$$f_\xi(t) = \exp\{\alpha t^2 + \beta t + \gamma\}$$

и известно, что $M\xi = -4$, $D\xi = 9$. Найти параметры α, β и γ, а также вероятность того, что $|3\xi + 4| \leqslant 11$.

33. Плотность распределения случайной величины ξ непрерывна на всей прямой и дается соотношением:

$$f_\xi(t) = \begin{cases} \alpha t^2 + \beta t + \gamma, & -1 \leqslant t \leqslant 1, \\ 0, & t \leqslant -1 \cup t \geqslant 1. \end{cases}$$

Найти параметры α, β, γ, функцию распределения случайной величины ξ, математическое ожидание случайной величины $\eta_1 = 3\xi - 1$, дисперсию $\eta_2 = -\xi - 1$ и вероятность того, что ξ не меньше 1/2.

34. Функция распределения случайной величины ξ непрерывна в $\mathbf{R}$ и дана соотношением

$$F_\xi(x) = \begin{cases} 0, & x \leqslant -1, \\ \dfrac{\beta^2}{\gamma}x^2 + 2\beta x + \gamma, & -1 \leqslant x \leqslant 1 \\ 1, & x \geqslant 1. \end{cases}$$

Найти значения постоянных β и γ, плотность распределения случайной величины ξ, математическое ожидание случайной величины $Y_1 = -3\xi + 4$, дисперсию случайной величины $Y_2 = -2\xi - 8$ и вероятность того, что случайная величина $Y_3 = -4\xi$ заключена в пределах от -8 до 2.

35. Плотность распределения случайной величины X дана соотношением

$$f_x(t) = \exp\{\alpha t^2 + \beta t + c\}$$

и известно, что $MX = 2$, $DX = 9$. Найти значение параметров a, b и c, а также вероятность того, что $|2\xi + 4| < 6$.

36. Прибор проходит испытания на надежность, причем время безотказной работы прибора — экспоненциальная случайная величина, с параметром $\mu = 0{,}5$. Сколько приборов следует испытать, чтобы с вероятностью, не меньшей 0,9, отобрать среди них не менее трех, прошедших испытания без отказов? Известно, что испытания одного прибора проходят 1,5 часа.

37. Функция распределения случайной величины ξ непрерывна в $\mathbf{R}$ и дана соотношением

$$F_\xi(t) = \begin{cases} 0, & x \leqslant 3, \\ \dfrac{\beta^2}{4\gamma}x^2 + \beta x + \gamma, & 3 \leqslant x \leqslant 4, \\ 1, & x \geqslant 4. \end{cases}$$

Найти параметры β и γ, плотность распределения случайной величины ξ, математическое ожидание случайной величины $Y_1 = 8\xi - 3$, дисперсию случайной величины $Y_2 = -4\xi + 2$ и вероятность того, что случайная величина $Y_3 = -3\xi$ заключена в пределах от -11 до -10.

38. Случайная величина ξ имеет пуассоновское распределение с параметром $\lambda = 2{,}5$. Найти закон распределения, математическое ожидание и дисперсию случайной величины η, заданной соотношением

$$\eta = (-1)^\xi.$$

39. Функция распределения случайной величины ξ непрерывна на $\mathbf{R}$ и дана соотношением

$$F_\xi(x) = \begin{cases} 0, & x \leqslant -1, \\ \dfrac{\beta^2}{4\gamma}x^2 + \beta x + \gamma, & -1 \leqslant x \leqslant 2, \\ 1, & x \geqslant 2. \end{cases}$$

Найти значения постоянных β и γ, плотность распределения случайной величины ξ, математическое ожидание случайной величины $Y_1 = 2\xi - 3$, дисперсию случайной величины $Y_2 = -3\xi + 5$ и вероятность того, что случайная величина $Y_3 = -2\xi$ заключена в пределах от -5 до 1.

40. Плотность распределения случайной величины ξ дана соотношением

$$f_\xi(t) = \exp\{\alpha t^2 + \beta t + \gamma\}$$

и известно, что $M\xi = -4$, $D\xi = 9$. Найти параметры α, β и γ, а также вероятность того, что $|3\xi + 4| \leqslant 11$.

41. Функция распределения случайной величины ξ дана соотношением

$$F_\xi(x) = \int_{-\infty}^{x} \exp\{at^2 + bt + c\}\, dt$$

и известно, что $M\xi = -2$, $P\{|\xi + 2| < 1\} = 0{,}668$. Найти параметры a, b, c.

42. Функция распределения случайной величины ξ дана соотношением

$$F_\xi(x) = \int_{-\infty}^{x} \exp\{\alpha t^2 + \beta t + c\}\, dt$$

и известно, что $M\xi = -4$, $P\{|\xi + 4| < 3\} = 0{,}758$. Найти параметры a, b и c.

43. Время безотказной работы электронной лампы — экспоненциальная случайная величина со средним 5. Для увеличения надежности агрегата ставят параллельно несколько ламп. Сколько ламп следует запараллелить, чтобы с вероятностью, не меньшей 0,99, по крайней мере, одна из них не вышла из строя за 10 часов?

44. Прибор проходит испытания на надежность, причем время безотказной работы прибора — экспоненциальная случайная величина, с параметром $\mu = 0{,}5$. Сколько приборов следует испытать, чтобы с вероятностью, не меньшей 0,9, отобрать среди них не менее трех, прошедших испытания без отказов? Известно, что испытания одного прибора проходят 1,5 часа.

45. Случайное отклонение размера детали от заданного является нормальной случайной величиной с параметрами $m = 0$, $\sigma = 5$мк. Сколько необходимо изготовить деталей, чтобы с вероятностью, не меньшей 0,99, среди них была хотя бы одна годная, если для годной детали допустимо отклонение от заданного размера не более, чем на 2 мк?

46. Изделие считается изделием 1-го сорта, если отклонение его размеров от номинала не превышает 3,45 мм. Эти отклонения случайны и подчиняются нормальному закону, причем систематические отклонения отсутствуют, а среднеквадратичное отклонение равно 3 мм. Найти среднее число первосортных изделий, если было изготовлено 4 изделия.

47. Случайное отклонение размера детали от заданного — нормальная случайная величина с параметрами m, σ. Изделие подлежит переделке, если отклонение превышает 10 мм. Найти функцию распределения отклонений для изделий, подлежащих переделке.

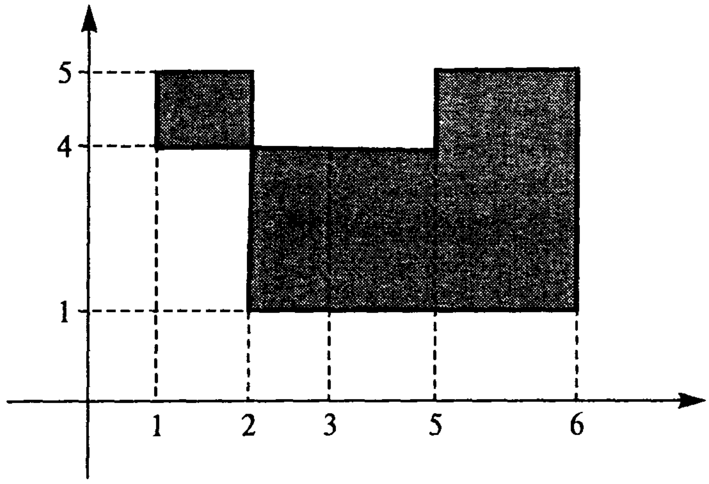
48. Случайная величина ξ имеет пуассоновское распределение с параметром $\lambda = 1$. Найти закон распределения, математическое ожидание и дисперсию случайной величины η, с равными вероятностями принимающей значения $(-1)^\xi$ и $(-1)^{\xi+1}$.

49. Время безотказной работы двигателя — экспоненциальная случайная величина со средним 1000 часов. Найти среднее число двигателей, вышедших из строя за 2000 часов, если в запасе их было 10.

50. Время t безотказной работы ЭВМ распределено по экспоненциальному закону с параметром 1. Решение определенной задачи требует безотказной работы машины в течение 2 минут. Если за это время происходит сбой, то задачу приходится решать заново. Сбой можно обнаружить только после окончания решения. Найти закон распределения времени, необходимого для решения задачи.

51. Браковка шариков для подшипников производится следующим образом: если шарик не проходит через отверстие диаметром 10 мм, но проходит через отверстие диаметром 12 мм, то его размер считать приемлемым. Если не выполнено первое условие — шарик бракуется, если второе — идет в переделку. Известно, что диаметр шарика — нормальная случайная величина с параметрами (11 мм; 0,5 мм). Найти распределение числа шариков, идущих в переделку, если было изготовлено 10 шариков.

52. Двумерная случайная величина (ξ, η) задана своей функцией распределения $F_{\xi\eta}(x, y)$. Найти вероятность попадания этой случайной величины в указанную на рисунке область·



53. Пара случайных величин X и Y имеет совместное нормальное распределение с вектором математических ожиданий $\{2, 1\}$ и ковариационной матрицей **K**

$$\mathbf{K}[X, Y] = \sigma^2 \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}.$$

Известно, что $P\{X + Y < 3\} = 0{,}65$. Найти DX, DY.

54. Найти ковариацию ординаты и абсциссы точки M, равномерно распределенной в треугольнике с вершинами $A(1, 1)$, $B(0, 1)$ и $C(1, 0)$. Зависимы ли эти случайные величины?

55. Найти ковариацию ординаты и абсциссы точки M, равномерно распределенной в треугольнике с вершинами $A(-2, 0)$, $B(0, 1)$ и $C(1, 0)$. Зависимы ли эти случайные величины?

56. Случайные величины ξ, η и ζ независимы и распределены нормально с параметрами соответственно $M\xi = M\eta = M\zeta = 5$.

$$10D\xi = 2D\eta = 5D\zeta = 10\sigma^2.$$

Найти дисперсию случайной величины $2\xi - 3\zeta$, если известно, что $P\{2\xi + \eta < 20\} = 0{,}9973$.

57. Случайные величины X, Y и Z связаны соотношением $4X - 2Y + +3Z = 0$. Известно, что $MX = 0$, $MY = 0$, $MZ = 0$, $DX = DY = DZ = = 1$. Найти ковариационную матрицу этих случайных величин.

58. Случайные величины ξ_1, ξ_2 и ξ_3 имеют совместное нормальное распределение с ковариационной матрицей

$$\mathbf{K}[\xi] = \begin{bmatrix} 3\sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \frac{3}{2}\sigma^2 \end{bmatrix}$$

и вектором математических ожиданий $(2, 2, 1)$. Найти дисперсию случайной величины ξ_1, если известно, что сумма $\xi_1 + \xi_2 + \xi_3$ в 99 случаях из 100 принимает значения не меньше 3,4.

59. Случайные величины ξ, η и ζ независимы и распределены нормально с параметрами соответственно $M\xi = M\eta = M\zeta = 5$.

$$10D\xi = 2D\eta = 5D\zeta = 10\sigma^2.$$

Найти дисперсию случайной величины $\xi - \zeta$, если известно, что $P\{2\xi + \eta < 20\} = 0{,}9973$.

60. Случайные величины ξ_1, ξ_2 и ξ_3 независимы и нормальны с одинаковым средним, равным 3. Известно, что их дисперсии относятся соответственно как $1 : 3 : 2$. Найти ковариационную матрицу этих случайных величин, если известно, что

$$P\{2\xi_1 - \xi_2 + 3\xi_3 > 15\} = 0{,}001.$$

61. Двумерная случайная величина (X, Y) имеет вектор математического ожидания $(7, -1)$ и ковариационную матрицу

$$\mathbf{K}[X, Y] = \begin{pmatrix} 9 & 6 \\ 6 & 4 \end{pmatrix}.$$

Достаточно ли этих данных, чтобы спрогнозировать значения компоненты Y при известных значениях компоненты X?

62. Случайные величины ξ, η и ζ независимы и распределены нормально с параметрами соответственно $M\xi = M\eta = M\zeta = 6$,

$$10D\xi = 2D\eta = 5D\zeta = 10\sigma^2.$$

Найти дисперсию случайной величины $\xi - 4\zeta$, если известно, что

$$P\{2\xi + \eta < 20\} = 0{,}9973.$$

63. При проведении телепатического опыта индуктор независимо от предшествующих опытов выбирает с $p = 0{,}5$ один из двух предметов и думает о нем, а реципиент угадывает, о каком предмете думает индуктор. Опыт был повторен 100 раз и при этом было получено 60 правильных ответов. Какова вероятность правильного ответа в одном опыте, в предположении, что телепатической связи между индуктором и реципиентом нет? Можно ли приписать наблюденный результат случайному совпадению или он говорит о наличии телепатической связи между индуктором и реципиентом?

64. Для лица, дожившего до 20-летнего возраста, вероятность несчастного случая на 21-ом году равна 0,006. Застрахована группа в 10000 человек 20-летнего возраста, причем страховой взнос составляет 1,2 рубля. В случае несчастья страховое агентство выплачивает застрахованному 100 руб. Какова вероятность того, что к концу года страховое агентство разорится? Его доход превысит 4000 рублей?

65. Театр, вмещающий 1000 человек, имеет 2 разных входа. Около каждого из входов свой гардероб. Зрители приходят парами и каждая пара, независимо от прочих, выбирает случайным образом один из входов. Сколько мест следует иметь в каждом из гардеробов, чтобы не менее чем в 99 случаях из 100 зрители могли раздеться в гардеробе того входа, через который они вошли?

66. Для проверки влияния нового лекарства на кровяное давление у 100 пациентов было измерено давление до и после приема лекарства. При этом оказалось, что в 32 случаях давление после приема лекарства повысилось, а в 68 случаях — понизилось. Какова вероятность того, что случайные колебания давления вызовут не меньшее отклонение от 50? Можно ли считать установленным, что лекарство влияет на давление?

67. При составлении статистического отчета надо было сложить 10^4 чисел, каждое из которых округлено с точностью до 10^{-4}. Предполагая, что ошибки округления чисел взаимнонезависимы и равномерно распределены на $(-1/2 \cdot 10^{-4}, 1/2 \cdot 10^{-4})$, найти пределы, в которых с вероятностью, большей 0,997, будет лежать суммарная ошибка.

68. ЭВМ имеет 24-разрядную двоичную сетку, т. е. оперирует с числами с точностью до 10^{-8}. При вводе числа округляются, причем ошибки округления — независимы, равномерно распределенные на промежутке $(-1/2 \cdot 10^{-8}; 1/2 \cdot 10^{-8})$ случайные величины. Находится среднее арифметическое 10000 чисел. В каких пределах лежит суммарная ошибка не менее чем в 99 случаях из 100?

69. При вычислении площади плоской фигуры D методом Монте-Карло поступают следующим образом: бросают n точек внутрь квадрата K, полностью содержащего фигуру, площадь которой ищут. Пусть площадь квадрата известна и равна $S(K)$. Тогда вероятность того, что точка, равномерно распределенная на квадрате K, попадает в область D, равна

$$P\{M \in D\} = \frac{S(D)}{S(K)}.$$

По теореме Бернулли при достаточно большом числе n вероятность $P\{M \in D\}$ может быть приблизительно заменена частотой $V_n = m/n$, где m — число точек, попавших в область D. Тогда искомую площадь определяют по формуле $S(D) = V_n S(K)$.

а) Пусть $S(K) = 1$ м^2. Сколько точек следует использовать для нахождения площади, чтобы не менее чем в 99 случаях из 100 точность была не хуже 1 дм^2?

б) Пусть $S(K) = 100$ см^2. Было брошено 1000 точек. Какую точность можно гарантировать не менее чем в 99 случаях из 100?

70. Игральную кость бросали 1000 раз. Найти пределы, в которых с вероятностью 0,99 будет лежать сумма выпавших очков.

71. Сколько раз следует бросить игральную кость, чтобы не менее чем в 9 случаях из 10 сумма выпавших очков отличалась от средней не более чем на 10?

72. Стрелок попадает при выстреле по мишени в десятку — с вероятностью 0,5; в девятку — 0,3; в восьмерку — 0,1; в семерку — 0,05; в шестерку — 0,05. Стрелок производит 100 выстрелов. Какова вероятность того, что он набрал более 920 очков?

Примеры решения задач

1. *Вероятность выбора точки из дуги полуокружности радиуса 1, с центром в начале координат, расположенной в правой полуплоскости, принимается пропорциональной длине проекции этой дуги на ось OX. Согласуется ли такое правило подсчета вероятностей с требованиями, предъявляемыми к вероятности?*

Решение. Нет, не согласуется. Нарушено правило сложения — для любых двух несовместных событий A и B должно выполняться $P\{A+B\} = P\{A\} + P\{B\}$. Пусть A — событие, состоящее в выборе точки из дуги, представляющей собой четверть окружности, лежащей в первой четверти, B — событие, состоящее в выборе точки из симметричной ей дуги, лежащей в четвертой четверти. Заметим, что $A+B$ — достоверное событие и, следовательно, $P\{A+B\} = 1$. В соответствии с предлагаемым определением, $P\{A\} = P\{B\} = P\{A+B\}$, что возможно лишь если эти вероятности равны нулю, но последнее противоречит отмеченной выше достоверности события $A+B$.

2. *Из множества точек, расположенных на отрезке $[-2, 1]$ случайным образом выбирают точку и полагают вероятность выбора из подотрезка $[\alpha, \beta] \in [-2, 1]$ пропорциональной $\beta^3 - \alpha^3$. Согласуется ли такой способ подсчета вероятностей с требованиями, предъявляемыми к вероятности?*

Решение. Да, согласуется. Действительно, из условия задачи

$$P\{M \in [\alpha, \beta]\} = k \cdot (\beta^3 - \alpha^3).$$

Условие нормировки вероятности позволяет определить значение коэффициента пропорциональности k:

$$P\{M \in [-2, 1]\} = k \cdot (1^3 - (-2)^3) = k \cdot 9 = 1 \quad \Rightarrow \quad k = \frac{1}{9}.$$

Пусть $[\alpha, \beta]$ и $[\beta, \gamma]$ — смежные подпромежутки основного промежутка. Тогда

$$P\{M \in [\alpha, \beta] \cup [\beta, \gamma]\} = P\{M \in [\alpha, \gamma]\} = \frac{1}{9}(\gamma^3 - \alpha^3).$$

В то же время

$$P\{M \in [\alpha, \beta]\} = \frac{1}{9}(\beta^3 - \alpha^3), \quad P\{M \in [\beta, \gamma]\} = \frac{1}{9}(\gamma^3 - \beta^3),$$

и в случае смежных промежутков правило сложения выполняется. Если же $[\alpha, \beta]$ и $[\gamma, \delta]$ — произвольные непересекающиеся промежутки (т. е. — несовместные события), то выполнение правила сложения следует из соотношения

$$P\{M \in [\alpha, \beta] \cup [\beta, \gamma] \cup [\gamma, \delta]\} = P\{M \in [\alpha, \beta]\} + P\{M \in [\beta, \gamma]\} + P\{M \in [\gamma, \delta]\}$$

и следующего из него равенства

$$P\{M \in [\alpha, \beta] \cup [\gamma, \delta]\} = P\{M \in [\alpha, \delta]\} - P\{M \in [\beta, \gamma]\}.$$

Прочие требования, предъявляемые к вероятности, легко устанавливаются.

3. *Вероятность попадания точки $M(p, q)$ в область F, лежащую в квадрате $K = \{(p, q)\colon |p| \leqslant 1,\ |q| \leqslant 1\}$, пропорциональна площади этой области: $P(M \in F) = S(F)/4$. Найти вероятность того, что у уравнения $px^2 - 2qx + 1 = 0$ нет действительных корней.*

Решение. Квадратное уравнение не имеет действительных корней, если его дискриминант отрицателен:

$$D = q^2 - p < 0.$$

Следовательно

$$P\{\text{нет корней}\} = P\{D < 0\} = P\{p > q^2\},$$

а последняя вероятность, в соответствии с условием задачи, пропорциональна площади множества $F = \{(p, q) : |p| \leqslant 1,\ |q| \leqslant 1,\ p > q^2$. Эта площадь равна (рис. 1)

$$S(F) = 2 \int_0^1 \sqrt{p}\, dp = \frac{4}{3}$$

и, следовательно, искомая вероятность может быть найдена из соотношения

$$P\{\text{нет корней}\} = \frac{S(F)}{4} = \frac{1}{3}.$$

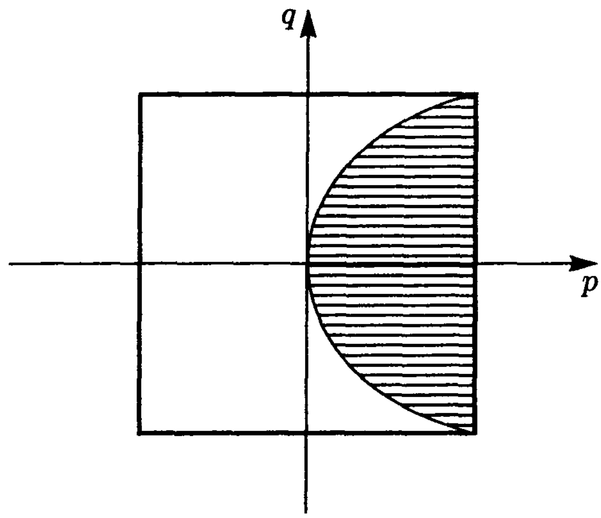


Рис. 1

4. *Из урны, содержащей всего 8 белых и черных шаров, извлекают одновременно два шара. Известно, что вероятность извлечения пары белых шаров в 15 раз больше вероятности извлечения пары черных. Можно ли по этим данным установить состав шаров в урне?*

Решение. Да, можно — в урне лежат 6 белых и 2 черных шара.

Действительно, пусть в урне m белых и $8 - m$ черных шаров. Вероятность извлечения пары белых шаров будет равна

$$P\{\text{бб}\} = \frac{C_m^2}{C_8^2}.$$

Аналогично, вероятность извлечения пары черных

$$P\{\text{чч}\} = \frac{C_{12-m}^2}{C_8^2}.$$

Из условия заключаем, что, поскольку $P\{\text{бб}\} = 15P\{\text{чч}\}$, постольку

$$C_m^2 = 15 \cdot C_{12-m}^2,$$

откуда получаем для m уравнение

$$7m^2 - 112m + 420 = 0,$$

имеющее пару действительных корней $m_1 = 6$ и $m_2 = 10$. Второй корень не удовлетворяет условию, поскольку в урне всего 8 шаров.

5. *Случайная величина ξ имеет пуассоновское распределение и известно, что ее математическое ожидание m и дисперсия D связаны соотношением $6 - m^2 = 5 \cdot D$. Найти вероятность $P(\xi < 3)$.*

Решение. Известно, что математическое ожидание и дисперсия пуассоновского распределения совпадают и равны значению его параметра λ. Условие задачи приводит к уравнению относительно λ:

$$\lambda^2 + 5\lambda - 6 = 0,$$

решениями которого являются числа $\lambda_1 = 1$ и $\lambda_2 = -6$. Последнее значение не может быть параметром пуассоновского распределения в силу положительности параметра. Таким образом, случайная величина ξ имеет ряд распределения

$$P\{\xi = n\} = e^{-1}\frac{1}{n!}, \quad n = 0, 1, 2, \dots .$$

Для искомой вероятности получаем

$$P(\xi < 3) = P(\xi = 0) + P(\xi = 1) + P(\xi = 2) = e^{-1} + e^{-1}\frac{1}{1!} + e^{-1}\frac{1}{2!} \approx 0{,}9197.$$

6. *Плотность распределения случайной величины ξ непрерывна на всей прямой и дается соотношением:*

$$f_\xi(t) = \begin{cases} \alpha t^2 + \beta t + \gamma, & 2 \leqslant t \leqslant 3, \\ 0, & t \leqslant 2 \cup t \geqslant 3. \end{cases}$$

Найти параметры α, β, γ, функцию распределения случайной величины ξ, математическое ожидание случайной величины $\eta_1 = 3\xi - 1$, дисперсию $\eta_2 = -2\xi + 5$ и вероятность того, что ξ не меньше 2,7.

Решение. Непрерывность плотности на всей прямой позволяет сделать заключение о ее непрерывности в точках $x = 2$ и $x = 3$, откуда $f_\xi(2) = f_\xi(3) = 0$. Условие нормировки плотности дает еще одно соотношение

$$\int_2^3 f_\xi(t)\,dt = 1,$$

а вместе они приводят к системе уравнений относительно неизвестных параметров α, β, γ:

$$\begin{cases} 4\alpha + 2\beta + \gamma = 0, \\ 9\alpha + 3\beta + \gamma = 0, \\ \displaystyle\int_2^3 (\alpha t^2 + \beta t + \gamma)\,dt = 1, \end{cases}$$

решение которой

$$\alpha = -6, \quad \beta = 30, \quad \gamma = -36.$$

Функция распределения случайной величины ξ дается соотношением

$$F_\xi(t) = \int_{-\infty}^{t} f_\xi(x)\,dx = \begin{cases} 0, & x < 2, \\ -2t^3 + 15t^2 - 36t + 28, & 2 \leqslant x \leqslant 3, \\ 1, & x \geqslant 3. \end{cases}$$

Заметим, что математическое ожидание случайной величины ξ, в силу симметрии плотности относительно точки $t = 2{,}5$, равно 2,5. Поэтому для математического ожидания случайной величины $\eta_1 = 3\xi - 1$ получаем

$$M\eta_1 = 3M\xi - 1 = 3 \cdot 2{,}5 - 1 = 6{,}5.$$

Для дисперсии случайной величины $\eta_2 = -2\xi + 5$ имеем

$$D\eta_2 = D(-2\xi + 5) = 4D\xi = 4\int_2^3 (t - 2{,}5)^2(-6t^2 + 30t - 36)\,dt = \frac{1}{5}.$$

Наконец, вероятность того, что ξ не меньше 2,7 равна

$$P\{\xi \geqslant 2{,}7\} = 1 - F_\xi(2{,}7) = 1 - \left.(-2t^3 + 15t^2 - 36t + 28)\right|_{t=2{,}7} = 0{,}216.$$

7. *Функция распределения случайной величины ξ непрерывна в* $\mathbf{R}$ *и дана соотношением*

$$F_\xi(x) = \begin{cases} 0, & x \leqslant -2, \\ \alpha x^2 + \beta x + \gamma, & -2 \leqslant x \leqslant 1, \\ 1, & x \geqslant 1. \end{cases}$$

Известно, что случайная величина ξ неотрицательна с вероятностью $1/9$. Найти значения постоянных α, β и γ, плотность распределения случайной величины ξ, математическое ожидание случайной величины $\eta = -2\xi^2$, и вероятность того, что эта случайная величина заключена в пределах от -1 до 0.

Решение. Условие непрерывности функции распределения в точках $x = -2$ и $x = 1$ дает

$$\alpha x^2 + \beta x + \gamma|_{x=-2} = 4\alpha - 2\beta + \gamma = 0$$

и

$$\alpha x^2 + \beta x + \gamma|_{x=1} = \alpha + \beta + \gamma = 1.$$

Присоединяя к ним условие

$$P\{\xi \geqslant 0\} = 1 - F_\xi(0) = 1 - \gamma = \frac{1}{9},$$

находим

$$\alpha = -\frac{1}{9}, \quad \beta = \frac{2}{9}, \quad \gamma = \frac{8}{9}.$$

Плотность, очевидно, равна:

$$\begin{cases} f_\xi(x) = \dfrac{d}{dx}F_\xi(x) = 0, & x < -2 \cup x > 1, \\ -\dfrac{2}{9}x + \dfrac{2}{9}, & -2 < x < 1. \end{cases}$$

Матожидание случайной величины $\eta = -2\xi^2$ получим, используя теорему о математическом ожидании функции от случайной величины:

$$M\eta = \int_{-2}^{1} (-2x^2)\left(-\frac{2}{9}x + \frac{2}{9}\right) dx = -\frac{17}{27}.$$

Ответ на последний вопрос задачи дают следующие рассуждения

$$P\{-1 < \eta < 0\} = P\{-1 < -2\xi^2 < 0\} = P\left\{-\frac{1}{2} < -\xi^2 < 0\right\} =$$

$$= P\left\{0 < \xi^2 < \frac{1}{2}\right\} = P\left\{-\sqrt{\frac{1}{2}} < \xi < \sqrt{\frac{1}{2}}\right\} = F_\xi\left(\sqrt{\frac{1}{2}}\right) - F_\xi\left(-\sqrt{\frac{1}{2}}\right) =$$

$$= \left(-\frac{1}{9}x^2 + \frac{2}{9}x + \frac{8}{9}\right)\Big|_{x=1/\sqrt{2}} - \left(-\frac{1}{9}x^2 + \frac{2}{9}x + \frac{8}{9}\right)\Big|_{x=-1/\sqrt{2}} = \frac{4}{9\sqrt{2}} \approx 0{,}3143.$$

8. *Функция распределения случайной величины ξ дана соотношением*

$$F_\xi(x) = \int_{-\infty}^{x} \exp\{ax^2 + bx + c\}\, dx$$

и известно, что $D\xi = 2$, $P\{\xi < 2\} = 0{,}9207$. Найти параметры a, b и c.

Решение. Выделяя полный квадрат в показателе степени экспоненты, преобразуем выражение для функции распределения следующим образом

$$F_\xi(x) = \int_{-\infty}^{x} \exp\{ax^2 + bx + c\}\, dx = \exp\left\{c - \frac{b^2}{4a^2}\right\} \int_{-\infty}^{x} \exp\left\{a\left(x + \frac{b}{2a}\right)^2\right\} dx.$$

Поскольку

$$\lim_{x\to\infty} F_\xi(x) = \exp\left\{c - \frac{b^2}{4a^2}\right\} \int_{-\infty}^{\infty} \exp\left\{a\left(x + \frac{b}{2a}\right)^2\right\} dx = 1,$$

постольку несобственный интеграл сходится, что возможно лишь если $a < 0$. Полагая $a = -1/\alpha^2$, продолжим преобразования, сделав в интеграле замену переменных

$$a\left(x + \frac{b}{2a}\right)^2 = \frac{1}{\alpha^2}\left(x - \frac{\alpha^2 b}{2}\right)^2 = \frac{u^2}{2}.$$

Получим

$$\int_{-\infty}^{\infty} \exp\left\{a\left(x + \frac{b}{2a}\right)^2\right\} dx = \frac{\alpha}{\sqrt{2}} \int_{-\infty}^{-\infty} \exp\left\{-\frac{u^2}{2}\right\} du,$$

откуда, учитывая, что

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{u^2}{2}\right\} du = 1,$$

заключаем

$$\alpha \cdot \sqrt{\pi} \cdot \exp\left\{c - \frac{b^2}{4a^2}\right\} = 1 \quad \Rightarrow \quad \exp\left\{c - \frac{b^2}{4a^2}\right\} = \frac{1}{\alpha\sqrt{\pi}}.$$

Функция распределения рассматриваемой случайной величины может быть, следовательно, представлена в виде

$$F_\xi(x) = \frac{1}{\alpha\sqrt{\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{1}{\alpha^2}\left(x - \frac{\alpha^2 b}{2}\right)^2\right\} dx,$$

и значит, рассматриваемая случайная величина является *нормальной* с параметрами

$$M\xi = \frac{\alpha^2 b}{2} = -\frac{b}{2a}, \quad D\xi = \frac{\alpha^2}{2}.$$

Из условия задачи получаем:

$$D\xi = 2 \quad \Rightarrow \quad \alpha^2 = 4 \quad \Rightarrow \quad a = -\frac{1}{\alpha^2} = -\frac{1}{4}.$$

$$P\{\xi < 2\} = F_\xi(2) = F\left(\frac{2 - M\xi}{\sigma}\right) = F\left(\frac{2 - M\xi}{\sqrt{2}}\right) = 0{,}9207,$$

где $F(x)$ — функция стандартного нормального распределения с параметрами $(0, 1)$. Отсюда

$$\frac{2 - M\xi}{\sqrt{2}} = F^{-1}(0{,}9207) \approx 1{,}41 \quad \Rightarrow \quad M\xi \approx 0 \quad \Rightarrow \quad b \approx 0.$$

Для нахождения величины c используем полученное выше соотношение

$$\exp\left\{c - \frac{b^2}{4a^2}\right\} = \frac{1}{\alpha\sqrt{\pi}} \quad \Rightarrow \quad c = -\ln 2\sqrt{\pi} \approx -1{,}266.$$

9. *Время безотказной работы некоторого узла сложного агрегата — экспоненциальная случайная величина со средним $M = 2$. Для увеличения надежности агрегата узел дублируется — ставят параллельно несколько одинаковых, но функционирующих независимо узлов. Сколько узлов следует запараллелить, чтобы с вероятностью не меньшей чем 0,9 по крайней мере один из них не вышел из строя за 10 часов работы?*

Решение. T — случайное время безотказной работы узла — имеет экспоненциальное распределение. Это означает, что

$$P\{T < t\} = F_T(t) = \begin{cases} 0, & t < 0, \\ 1 - \exp\{-\mu \cdot t\}, & t \geqslant 0. \end{cases}$$

Известно, что математическое ожидание экспоненциальной случайной величины есть величина, обратная параметру: $M = 1/\mu \Rightarrow \mu = 0{,}5$. Следовательно, вероятность отказа узла в течение 10 часов будет равна

$$P\{T < 10\} = F_T(10) = 1 - \exp\{-0{,}5 \cdot 10\} \approx 0{,}9933.$$

Если запараллелено N идентичных узлов, то событие $A = \{$по крайней мере один из узлов не вышел из строя за 10 часов$\}$ является противоположным событию $\overline{A} = \{$все узлы вышли из строя за 10 часов$\}$. Поэтому $P\{A\} = 1 - P\{\overline{A}\}$. Для последней вероятности (в силу независимости отказов запараллеленных узлов) получаем

$$P\{\overline{A}\} = (P\{T < 10\})^N.$$

Искомое количество N может теперь быть найдено как наименьшее целое решение неравенства

$$P\{A\} = 1 - P\{\overline{A}\} = 1 - (P\{T < 10\})^N = 1 - (0{,}9933)^N \geqslant 0{,}9 \Rightarrow N \geqslant 343.$$

10. *Случайное отклонение размера детали от номинального — нормальная случайная величина с параметрами $m = 0$, $\sigma = 1$. Годными являются те детали, для которых отклонение заключено в пределах от -1 до 1 мм. Найти функцию распределения отклонений для годных изделий.*

Решение. Пусть ξ — случайное отклонение размера детали от номинального: $\xi = N[0, 1]$, η — отклонение размера *годной* детали от номинального:

$$\eta = \xi, \quad -1 \leqslant \xi \leqslant 1.$$

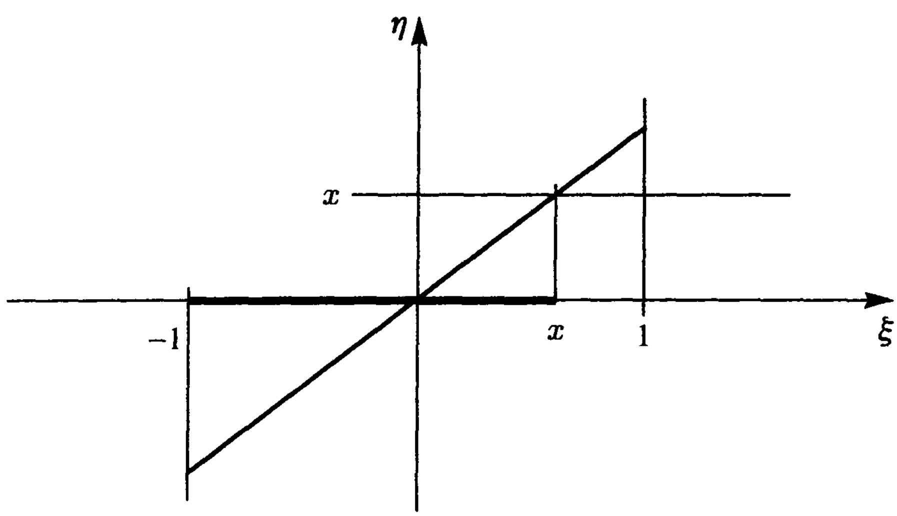


Рис. 2

Для функции распределения случайной величины η получаем по определению (рис. 2)

$$F_\eta(x) = P\{\eta < x\} = \begin{cases} 0, & x < -1, \\ P\{\xi < x | -1 \leqslant \xi \leqslant 1\}, & -1 \leqslant x \leqslant 1, \\ 1, & x > 1. \end{cases}$$

Вероятность же $P\{\xi < x | -1 \leqslant \xi \leqslant 1\}$ равна

$$P\{\xi < x | -1 \leqslant \xi \leqslant 1\} = \frac{P\{-1 \leqslant \xi < x\}}{P\{-1 \leqslant \xi \leqslant 1\}}$$

в соответствии с формулой условной вероятности. Таким образом, искомая функция распределения дается соотношением

$$F_\eta(x) = \begin{cases} 0, & x < -1, \\ \dfrac{F(x) - F(-1)}{F(1) - F(-1)}, & -1 \leqslant x \leqslant 1, \\ 1, & x > 1, \end{cases}$$

где $F(x)$ — функция стандартного нормального распределения с параметрами 0 и 1. Находя значения $F(-1) = 0{,}1587$ и $F(1) = 0{,}8413$, окончательно получаем (рис. 3)

$$F_\eta(x) = \begin{cases} 0, & x < -1, \\ \dfrac{F(x) - 0{,}1587}{0{,}6826}, & -1 \leqslant x \leqslant 1, \\ 1, & x > 1. \end{cases}$$

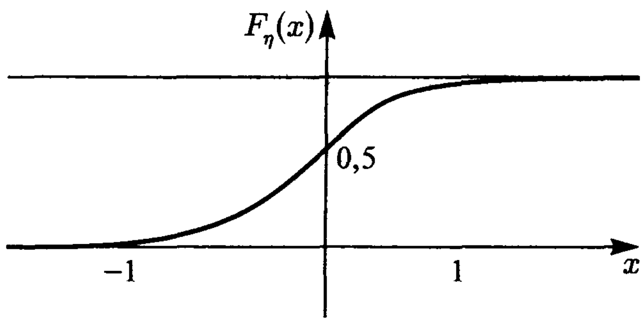


Рис. 3

11. *При сложении тысячи чисел каждое из них было округлено с точностью до* 10^{-3}. *Предполагая, что ошибки округления слагаемых взаимно-независимы и равномерно распределены на промежутке* $[-1/2 \cdot 10^{-3}, 1/2 \cdot 10^{-3}]$, *найти пределы, в которых будет лежать суммарная ошибка с надежностью, не худшей* 0,997.

Решение. Пусть ξ_i, $i = 1, 2, \ldots, 1000$ — взаимно независимые случайные слагаемые, речь о которых идет в условии задачи. Задача состоит в отыскании таких значений ε_1 и ε_2, что

$$P\{\varepsilon_1 \leqslant \xi_1 + \xi_2 + \cdots + \xi_{1000} \leqslant \varepsilon_2\} \geqslant 0{,}997.$$

Учитывая симметрию рассматриваемой суммы относительно нуля, можно положить $-\varepsilon_1 = \varepsilon_2 = \varepsilon > 0$. Таким образом, приходим к следующей задаче:

Найти значение ε такое, что

$$P\{|\xi_1 + \xi_2 + \cdots + \xi_{1000}| \leqslant \varepsilon\} \geqslant 0{,}997.$$

Для вычисления вероятности в левой части неравенства воспользуемся центральной предельной теоремой для одинаково распределенных слагаемых

$$P\left\{ \frac{\sum\limits_{i=1}^{N} \xi_i - N \cdot M\xi}{\sigma\sqrt{N}} < x \right\} \approx F(x),$$

где $F(x)$ — стандартная функция нормального распределения с параметрами $(0, 1)$, N — достаточно большое число[2)].

[2)] Известно, что для равномерно распределенных слагаемых указанное приближение справедливо с высокой степенью точности уже при $N > 12$.

Поскольку каждая из ξ_i равномерна на заданном промежутке $[-1/2 \cdot 10^{-3}, 1/2 \cdot 10^{-3}]$, то

$$M\xi_i = M\xi = 0, \quad D\xi_i = \frac{(b-a)^2}{12} = \frac{10^{-6}}{12}, \quad \sigma_\xi = \sigma = \frac{10^{-3}}{2\sqrt{3}},$$

$$P\left\{\left|\sum_{i=1}^{1000} \xi_i\right| \leqslant \varepsilon\right\} = P\left\{\frac{\left|\sum\limits_{i=1}^{1000} \xi_i\right|}{\sigma\sqrt{1000}} \leqslant \frac{\varepsilon}{\sigma\sqrt{1000}}\right\} \approx 2F\left(\frac{\varepsilon}{\sigma\sqrt{1000}}\right) - 1.$$

Последнее соотношение приводит к неравенству относительно ε

$$2F\left(\frac{\varepsilon}{\sigma\sqrt{1000}}\right) - 1 \geqslant 0{,}997 \quad \Rightarrow \quad F\left(\frac{\varepsilon}{\sigma\sqrt{1000}}\right) \geqslant 0{,}9985,$$

что в силу монотонности функции $F(x)$ дает

$$\frac{\varepsilon}{\sigma\sqrt{1000}} \geqslant F^{-1}(0{,}9985) \approx 2{,}96 \quad \Rightarrow \quad \varepsilon \geqslant 2{,}96 \cdot \frac{10^{-3}}{2\sqrt{3}} \cdot \sqrt{1000} \approx 0{,}027.$$

12. *При вычислении площади плоской фигуры D методом Монте-Карло поступают следующим образом:*

Бросают случайно N точек внутрь квадрата K, полностью содержащего фигуру, площадь которой ищут.

Вероятность того, что точка, равномерно распределенная на квадрате K, попадает в область D, равна

$$P\{M \in D\} = \frac{S(D)}{S(K)}$$

и, следовательно, искомая площадь может быть найдена из соотношения

$$S(D) = P\{M \in D\} \cdot S(K).$$

По теореме Бернулли, при достаточно большом числе N вероятность $P\{M \in D\}$ может быть приблизительно заменена частотой

$$P\{M \in D\} \approx \nu_N = \frac{m}{N},$$

где m — число точек, попавших в область D. Для площади отсюда получаем приближенную формулу

$$S(D) \approx S_N(D) = \nu_N \cdot S(K).$$

Сколько точек следует использовать для нахождения площади, чтобы не менее чем в 9 случаях из 10 точность $\Delta_N = |S(D) - S_N(D)|$ была не хуже 1 дм^2, если известно, что $S(K) = 1$ м^2?

Решение. Задача сводится к определению значения N такого, что

$$P\{|S(D) - S_N(D)| < 1\} \geqslant 0{,}9.$$

Заметим, что (ниже положено для сокращения записи $P\{M \in D\} = p$)

$$P\{|S(D) - S_N(D)| < 1\} = P\{|p \cdot S(K) - \nu_N \cdot S(K)| < 1\} =$$

$$= P\{|p - \nu_N| \cdot S(K) < 1\} = P\{|p - \nu_N| < 0{,}01\}.$$

Теорема Бернулли для последней вероятности при больших значениях N дает оценку

$$P\{|p - \nu_N| < 0{,}01\} \approx 2F\left(\frac{0{,}01 \cdot \sqrt{N}}{\sqrt{p(1-p)}}\right) - 1.$$

Таким образом, определению подлежит величина N (достаточно большая, чтобы последнее приближение было справедливо), удовлетворяющая неравенству

$$2F\left(\frac{0{,}01\cdot\sqrt{N}}{\sqrt{p(1-p)}}\right)-1\geqslant 0{,}9.$$

В таком виде это неравенство для своего решения относительно N требует знания величины $p=P\{M\in D\}$, которая неизвестна, ибо ради ее определения и был задуман рассматриваемый эксперимент!

Мы предлагаем ниже два способа оценки значения N сверху[3], не являющиеся, конечно, точными, однако достаточно удобные с вычислительной точки зрения.

I. Первый способ связан с усилением основного неравенства в соответствии с нижеследующими рассуждениями:

Поскольку $p+(1-p)=1$, то $\sqrt{p(1-p)}\leqslant 0{,}5$ и, следовательно,

$$\frac{0{,}01\cdot\sqrt{N}}{\sqrt{p(1-p)}}\geqslant 2\cdot 0{,}01\cdot\sqrt{N}.$$

Функция $F(\cdot)$ монотонно возрастает, поэтому

$$2F\left(\frac{0{,}01\cdot\sqrt{N}}{\sqrt{p(1-p)}}\right)\geqslant 2F(2\cdot 0{,}01\cdot\sqrt{N}).$$

Отсюда, всякое значение N, удовлетворяющее неравенству

$$2F(2\cdot 0{,}01\cdot\sqrt{N})-1\geqslant 0{,}9,$$

подавно удовлетворяет и неравенству

$$2F\left(\frac{0{,}01\cdot\sqrt{N}}{\sqrt{p(1-p)}}\right)-1\geqslant 0{,}9.$$

Предыдущее же неравенство легко разрешимо относительно N:

$$F(2\cdot 0{,}01\cdot\sqrt{N})\geqslant 0{,}95 \quad\Rightarrow\quad 2\cdot 0{,}01\cdot\sqrt{N}\geqslant F^{-1}(0{,}95)\approx 1{,}645 \quad\Rightarrow$$
$$\Rightarrow\quad \sqrt{N}\geqslant\frac{1{,}645}{2\cdot 0{,}01}\approx 82{,}25 \quad\Rightarrow\quad N\geqslant 6766.$$

II. Второй способ связан с заменой неизвестной величины p ее оценкой ν_N. Основное неравенство при этом принимает вид

$$2F\left(\frac{0{,}01\cdot\sqrt{N}}{\sqrt{\nu_N(1-\nu_N)}}\right)-1\geqslant 0{,}9$$

и легко разрешается

$$F\left(\frac{0{,}01\cdot\sqrt{N}}{\sqrt{\nu_N(1-\nu_N)}}\right)\geqslant 0{,}95 \quad\Rightarrow\quad \frac{0{,}01\cdot\sqrt{N}}{\sqrt{\nu_N(1-\nu_N)}}\geqslant F^{-1}(0{,}95)\approx 1{,}645 \quad\Rightarrow$$
$$\Rightarrow\quad \sqrt{N}\geqslant\frac{1{,}645\cdot\sqrt{\nu_N(1-\nu_N)}}{0{,}01},$$

[3] То есть, мы предлагаем способы определения числа N_0, такого, что искомое N заведомо не превышает его и, следовательно, количество экспериментов большее чем N_0 гарантирует нам требуемую точность.

и окончательный ответ зависит от априорного (может быть и приблизительного) знания величины искомой площади

Так, если искомая площадь порядка половины площади квадрата, то $\nu_N \approx 0{,}5$, и ответ такой же, как и выше.

Если искомая площадь близка к площади квадрата, то $\nu_N \approx 1$ и величина $\nu_N(1-\nu_N)$ мала, а вместе с ней мало и требуемое число экспериментов. Так, если, например $\nu_N \approx 0{,}9$, то $\nu_N(1-\nu_N) \approx 0{,}09$ и для N получаем

$$\sqrt{N} \geqslant \frac{1{,}645 \cdot \sqrt{0{,}09}}{0{,}01} \approx 49{,}35 \quad \Rightarrow \quad N \geqslant 2436.$$

Вообще, приведенные выше выкладки показывают, что подобная процедура нахождения площадей плоских фигур тем эффективнее, чем бóльшую часть объемлющего квадрата занимает множество D или его дополнение $\overline{D} = K - D$.

13. *При проведении телепатического опыта индуктор независимо от предшествующих опытов выбирает с $p = 0{,}5$ один из двух предметов и думает о нем, а реципиент угадывает, о каком предмете думает индуктор. Опыт был повторен 100 раз и при этом было получено 58 правильных ответов. Какова вероятность правильного ответа в одном опыте, в предположении, что телепатической связи между индуктором и реципиентом нет? Можно ли приписать наблюденный результат случайному совпадению или он говорит о наличии телепатической связи между индуктором и реципиентом?*

Решение. Пусть событие H_A состоит в том, что индуктор выбрал предмет A, событие H_B — он выбрал предмет B. Пусть, далее, события A и B означают, что реципиент выбрал соответствующий предмет

Если телепатической связи между индуктором и реципиентом нет, то в каждом из опытов

$$P\{A|H_A\} = P\{A|H_B\} = P\{B|H_A\} = P\{B|H_B\} = \frac{1}{2}.$$

При этом вероятность $P\{У\}$ угадывания реципиентом предмета, выбранного индуктором равна 0,5. Действительно, по формуле полной вероятности легко получить:

$$P\{У\} = P\{A|H_A\}P\{H_A\} + P\{B|H_B\}P\{H_B\} = \frac{1}{2}.$$

Если же телепатическая связь между индуктором и реципиентом есть, то вероятности $P\{A|H_A\}$ и $P\{B|H_B\}$ должны быть больше 0,5 (соответственно, $P\{B|H_A\}$ и $P\{A|H_B\}$ меньше 0,5), при этом и вероятность угадывания будет больше 0,5.

Следовательно, если телепатической связи между индуктором и реципиентом нет, то количество S_{100} правильно угаданных ответов должно быть близко к 50. Если же количество правильно угаданных ответов будет сильно отличаться от 50 (в 100 опытах), то скорее всего всего это должно означать, что телепатическая связь между индуктором и реципиентом есть

Пусть $0 < \alpha < 1$ некоторая вероятность, настолько малая, что событиями, вероятность которых меньше α, можно пренебречь как практически невозможными[4]. Предполагая, что телепатической связи между индуктором и реципиентом нет, определим, сколько же правильно угаданных ответов мы должны в этом случае наблюдать. И хотя логически понятно, что это может быть любое число от 0 до 100, скорее всего все-таки это будут числа, близкие к 50. Пусть $0 < k < 50$ число, такое, что

$$P\{50 - k \leqslant S_{100} \leqslant 50 + k\} \geqslant 1 - \alpha.$$

Тогда почти наверняка (с вероятностью, не меньшей чем $1 - \alpha$, т. е. в $(1-\alpha)\cdot 100$ % случаев) мы будем наблюдать в эксперименте не менее чем $50 - k$ и не более чем $50 + k$ правильных угадываний.

[4] Обычно в качестве α берут одно из значений 0,001, 0,005, 0,01, 0,05, 0,1.

Если наблюденное количество угадываний будет выходить за эти пределы, то объяснений этому может быть два:

— телепатии нет, но случайность процесса угадывания привела к такому значительному отклонению от среднего ожидаемого числа угадываний;

— такое значительное отклонение от среднего ожидаемого числа угадываний есть следствие наличия телепатической связи между индуктором и реципиентом.

И хотя логически оба рассуждения безупречны, мы, тем не менее, выберем второе, потому что выбор первого означает, что мы признаем возможность осуществления события с вероятностью α, а это противоречит нашим исходным посылкам! Вероятность ошибки здесь в точности равна α.

Если же наблюденное количество угадываний будет входить в найденные пределы, то и этому может быть дано два различных объяснения:

— телепатии нет, и именно поэтому мы получили незначительное отклонение наблюденного числа угадываний от ожидаемого;

— телепатия есть, но случайно так получилось.

И здесь оба рассуждения логически безупречны, но мы выберем первое, ибо если наше предположение об отсутствии телепатии справедливо, то наблюденный результат должен наблюдаться почти всегда (с вероятностью $1-\alpha$, т. е. в $(1-\alpha)\cdot 100$ % случаев). Наша вероятность ошибиться будет зависеть в этом случае от того, насколько силен предполагаемый телепатический эффект. Чем он сильнее, тем реже мы будем ошибаться.

Таким образом, для того чтобы получить ответ на вопрос задачи, следует определить пограничное значение k, описывающее возможные в подавляющем большинстве случаев в экспериментальной ситуации отклонения наблюденного числа угадываний от предполагаемого в случае отсутствия телепатической связи между индуктором и реципиентом. Для этого необходимо для заданного значения α решить относительно k неравенство

$$P\{50-k \leqslant S_{100} \leqslant 50+k\} \geqslant 1-\alpha.$$

Заметим, что поскольку опыты проводятся независимо друг от друга и вероятность успеха (угадывания) в каждом опыте одна и та же, то количество успехов S_{100} является биномиальной случайной величиной с параметрами $n = 100$, $p = 0{,}5$ и вероятность, стоящая в левой части неравенства, равна

$$P\{50-k \leqslant S_{100} \leqslant 50+k\} = \sum_{i=50-k}^{50+k} C_{100}^{i}\cdot(0{,}5)^{i}.$$

Применяя теорему Муавра—Лапласа, получаем

$$P\{50-k \leqslant S_{100} \leqslant 50+k\} \approx F\left(\frac{k}{\sqrt{100\cdot 0{,}5\cdot 0{,}5}}\right) - F\left(\frac{-k}{\sqrt{100\cdot 0{,}5\cdot 0{,}5}}\right) =$$

$$= 2F\left(\frac{k}{\sqrt{100\cdot 0{,}5\cdot 0{,}5}}\right) - 1.$$

Последнее соотношение позволяет записать основное неравенство в виде

$$2F\left(\frac{k}{\sqrt{100\cdot 0{,}5\cdot 0{,}5}}\right) - 1 \geqslant 1-\alpha$$

и решить его относительно k

$$F\left(\frac{k}{5}\right) \geqslant \frac{2-\alpha}{2} \quad\Rightarrow\quad \frac{k}{5} \geqslant F^{-1}\left(\frac{2-\alpha}{2}\right) \quad\Rightarrow\quad k \geqslant 5\cdot F^{-1}\left(1-\frac{\alpha}{2}\right).$$

Граничные значения k_α, соответствующие различным значениям α, приведены ниже

α	$F^{-1}\left(1-\frac{\alpha}{2}\right)$	k
0,1	1,645	9
0,01	2,576	13
0,001	3,29	17

Таким образом, если телепатической связи между индуктором и реципиентом нет, то количество случайных угадываний будет в пределах от 41 до 59 с надежностью не худшей 0,9, от 37 до 63 с надежностью не худшей 0,99 и от 33 до 67 с надежностью не худшей 0,999.

Наблюденные в эксперименте 58 угадываний не дают оснований для вывода о наличии телепатической связи — результат может быть объяснен случайным совпадением.

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Приступая к изучению элементов статистики, отметим несколько особенностей в постановке и решении ее задач в сравнении с задачами теоретико-вероятностными.

Теория вероятностей, исходя из известных характеристик совокупности случайных величин, отвечает на вопрос о возможности осуществления того или иного события, обусловленного рассматриваемыми случайными величинами:

знаем закон распределения совокупности случайных величин $\bar{\xi} = \{\xi_1 \dots \xi_n\}$ — хотим уметь находить вероятности событий, которые этими случайными величинами определяются.

В статистике мы решаем задачи, в некотором смысле обратные, а именно: наблюдая некоторые события, о которых известно, что возможность их осуществления или неосуществления обусловливается комплексом случайных величин, хотим определить эти (неизвестные) случайные величины, их вероятностные характеристики:

знаем результаты наблюдений (конкретные значения, принятые случайной величиной) — хотим сделать какие-нибудь заключения о законе распределения (в частности о параметрах и числовых характеристиках) наблюдаемой случайной величины.

Ясно, что в силу принципиальной *непредсказуемости* результатов наблюдения за *случайной* величиной, выводы, сделанные на основе результатов эксперимента, будут информативными только в том случае, когда эти наблюдения «хорошие» — т. е. те значения ξ, которые имеют бóльшую вероятность, будут наблюдаться в эксперименте чаще, а имеющие меньшую вероятность — реже. Законы больших чисел утверждают, что в подавляющем большинстве экспериментов так и будет. Однако, это не гарантирует нам, что *данный конкретный* эксперимент окажется именно таким.

Поэтому всякое статистическое заключение недостоверно: если основа заключения «хороший» эксперимент, то заключение достаточно близко к истине, если «плохой», то ошибочно. При этом «хорош» эксперимент или «плох» определяется не нами, не нашей добросовестностью наблюдателя и тщательностью экспериментатора, а исключительно случаем — природой.

Сказанное хорошо иллюстрируется следующим примером: пусть наблюдаются результаты n-кратного бросания монеты. Если бросать монету достаточно долго, то частота появления, например, герба, как гласит закон больших чисел в форме Бернулли, будет близка к вероятности. Поэтому по частоте (наблюдаемой и вычисляемой величине) можно сделать заключение о вероятности (неизвестной величине). Насколько это заключение соответствует истине?

Пусть монета симметрична, т. е. $P(\Gamma) = P(\text{Р}) = 0{,}5$ (что тем не менее не препятствует асимметрии в количестве появлений герба и решки в конкретном эксперименте)!

Может статься, что в серии из 100 бросаний герб появится 45 раз, а решка 55, а может статься и так, что герб появится 20 раз, а решка — 80. Ясно, что первая серия может быть признана «хорошей» с точки зрения рассматриваемой задачи, а вторая — «плохой». В любом случае мы сделаем заключение о неизвестной вероятности выпадения герба по наблюденной в эксперименте частоте и в первом случае положим $P(\Gamma) = 0{,}45$, а во втором — $0{,}2$. Основанием для оптимизма является то важное обстоятельство, что «плохие» серии будут встречаться тем реже, чем длиннее серия! Значит, при достаточно длинной серии бросаний эксперимент скорее будет «хорошим» чем «плохим», и определенная по результатам такого эксперимента $P(\Gamma)$

будет «похожа» на истинную. Достоверность статистического вывода будет определяться тем, насколько «редки» плохие эксперименты.

Практика использования статистических процедур показывает, что чаще всего решения, принятые на основании подобных выводов, оказываются верными. И именно это обстоятельство (согласованность статистических выводов с экспериментом) делает математическую статистику не бесполезной в практическом отношении наукой.

В дальнейшем мы неоднократно будем употреблять термины «маленькая вероятность», «маловероятное событие» и т. п. Какая же вероятность может считаться маленькой, а какая нет? Не вдаваясь подробно в обсуждение этого вопроса, заметим только, что абсолютная величина вероятности вне связи с конкретной обстановкой не дает нам никаких сведений о ее малости или немалости. Скажем, если нам известно, что вероятность осуществления некоторого события равна 0,01, то эта вероятность будет маленькой, если комплекс условий, обусловливающий рассматриваемое событие, складывается один раз за сто лет. Если же комплекс условий, при котором наблюдается рассматриваемое событие, складывается каждые пять минут, то эта же вероятность должна рассматриваться как значительная. Другими словами, под маленькой вероятностью мы будем понимать вероятность такого события, которое практически не наблюдается, вне зависимости от ее численного значения.

Глава XLIII

ОЦЕНКИ

§ 1. Выборки и выборочные характеристики

Рассмотрим эксперимент Ω, описание которого строится при помощи случайной величины ξ, или, что то же самое, рассмотрим случайную величину ξ, которую мы можем наблюдать в эксперименте Ω. Это означает, что однократный эксперимент Ω дает нам возможность определить одно из возможных значений случайной величины ξ.

Пусть в результате n экспериментов Ω получен набор значений случайной величины ξ

$$X_1, X_2, \ldots, X_n. \tag{1}$$

Если случайная величина ξ в процессе экспериментирования не менялась, если не менялись условия проведения эксперимента Ω и все измерения значений случайной величины ξ проводились независимо друг от друга, то говорят, что ***набор (1) образует выборку объема n из распределения случайной величины ξ***.

Заметим, что если сказано: некоторая совокупность из n чисел образует выборку, то при этом предполагается следующее:

а) эксперимент Ω может быть проведен при неизменных условиях сколько угодно раз;

б) имеет место устойчивость частот, т. е. имеет смысл говорить о вероятности попадания вектора $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$ в некоторое наперед заданное множество A из множества всех возможных совокупностей (1) (обычно $A \subset \mathbb{R}^n$).

Предположение а) означает, что, говоря о выборке объема n, мы говорим не о n конкретных числах, а о целой матрице чисел

$$\begin{matrix} X_1^1 & X_2^1 & \ldots & X_n^1, \\ X_1^2 & X_2^2 & \ldots & X_n^2, \\ \ldots & \ldots & \ldots & \ldots \\ X_1^k & X_2^k & \ldots & X_n^k, \end{matrix}$$

где элемент X_i^j — это значение случайной величины ξ, полученное в i-м эксперименте Ω, который проводился в j-й серии, $1 \leqslant j \leqslant k$. Таким образом, можно дать следующее (уточняющее) определение выборки: ***выборкой объема n из закона распределения случайной величины ξ называется совокупность n штук независимых одинаково распределенных случайных величин, совпадающих с ξ***.

Пример 1. Измеряется n однотипных деталей, изготовленных на одном станке. В результате получена совокупность чисел $X_1, X_2, \ldots, X_n$, которые образуют выборку из распределения случайной величины ξ — размер изготовляемой на данном станке детали.

Пример 2. Измеряется n однотипных деталей, изготовленных на различных станках. Получается совокупность чисел $X_1, X_2, \ldots, X_n$. Эти числа, вообще говоря, выборку не образуют, так как может, например, оказаться, что точность изготовления детали на различных станках различна и, следовательно, числа X_i являются реализациями различных случайных величин.

Отметим некоторую неоднозначность термина «выборка». Иногда под выборкой объема n понимают и конкретный набор n чисел, полученных в результате серии из n экспериментов. Но обычно бывает ясно, в каком смысле говорится о выборке. Скажем, если найдено среднее арифметическое по выборке и оно равно 5,

$$\bar{x} = \frac{1}{n} \sum x_i = 5,$$

то здесь под выборкой понимается конкретный набор чисел. Если же обсуждаются свойства величины

$$\bar{x} = \frac{1}{n} \sum x_i,$$

то под выборкой понимается любой возможный набор значений случайной величины ξ, т. е. совокупность n штук независимых случайных величин.

Первой статистической задачей, которую мы рассмотрим, будет задача нахождения функции распределения и числовых характеристик случайной величины ξ по выборке, полученной в результате эксперимента.

Пусть дана выборка (1) из закона распределения $F_\xi(x)$ случайной величины ξ. Требуется определить функцию распределения $F_\xi(x)$ случайной величины ξ и ее моменты.

Рассмотрим дискретную случайную величину ξ_n, принимающую значения X_1, X_2, ..., X_n, каждое с вероятностью $1/n$.

Эмпирической функцией распределения случайной величины ξ называется функция распределения случайной величины ξ_n

$$F_n(x) = \frac{\text{количество } X_i \text{ таких, что } X_i < x}{n}. \tag{2}$$

Эмпирическая (или выборочная) функция распределения является случайной величиной, так как она определяется по выборке (1) и зависит от того, какой конкретно набор чисел получен в данной серии из n экспериментов Ω. Оказывается, что если последовательность $X_1, X_2, \ldots, X_n$ достаточно длинная, то эмпирическая функция распределения будет очень похожей на теоретическую функцию $F_\xi(x)$ для большинства из возможных наборов чисел $X_1, X_2, \ldots, X_n$. Точнее, имеет место

Теорема (Гливенко—Кантелли). *Пусть $F_\xi(x)$ — теоретическая функция распределения случайной величины ξ, а $F_n(x)$ — эмпирическая. Тогда для произвольного значения $\varepsilon > 0$*

$$\lim_{n\to\infty} \left\{ |F_n(x) - F_\xi(x)| > \varepsilon \right\} = 0.$$

◀ Фиксируем некоторое число x и рассмотрим случайные величины

$$\xi_i(x) = \begin{cases} 0, & X_i \geqslant x, \\ 1, & X_i < x. \end{cases}$$

Заметим, что ряд распределения случайных величин ξ_i имеет вид

$$P\{\xi_i = 0\} = P\{X_i \geqslant x\} = 1 - F_\xi(x),$$
$$P\{\xi_i = 1\} = P\{X_i < x\} = F_\xi(x).$$

Все ξ_t — одинаково распределенные случайные величины с указанным рядом распределения и конечным математическим ожиданием

$$M\xi_t(x) = P\{\xi_t = 0\} \cdot 0 + P\{\xi_t = 1\} \cdot 1 = F_\xi(x).$$

Применяя к последовательности случайных величин $\xi(x)$ закон больших чисел в форме Хинчина, получаем

$$\lim_{n\to\infty} P\left\{\left|\frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n} - M\xi\right| > \varepsilon\right\} = 0,$$

откуда следует утверждение теоремы. ▶

Теорема позволяет сделать вывод: эмпирическая функция распределения, построенная по выборке объема n, тем более похожа на $F_\xi(x)$, чем больше n — объем выборки.

Этот вывод не является достоверным, а, как утверждает теорема, носит вероятностный характер — *вероятность* отклонений $F_\xi(x)$ от $F_n(x)$ при $n \to \infty$ очень мала — как бы ни была велика выборка, всегда существует возможность получить $F_n(x)$, значительно отличающуюся от $F_\xi(x)$. Однако при достаточно больших n этой возможностью можно пренебречь. Обратимся к поясняющему примеру.

Пример 1. Рассмотрим массовое производство некоторых однотипных изделий, изготавливаемых в неизменных условиях. Пусть случайная величина ξ принимает значение 1, если изготовленное изделий доброкачественно, и 0 в противном случае. Отобрано n изделий и среди них оказалось n_1 дефектных. Оценить функцию распределения случайной величины ξ.

◀ Случайная величина ξ дискретна и задача оценки ее функции распределения есть попросту задача оценки вероятности $P\{\xi = 0\}$, т. е. вероятности получить дефектное изделие. Теорема Гливенко—Кантелли в этом случае говорит, что оценкой вероятности $P\{\xi = 0\}$ может служить частота появления дефектных изделий в рассматриваемой выборке

$$F_n(x) = \begin{cases} 0, & x < 0, \\ \dfrac{n_1}{n}, & 0 \leqslant x < 1, \\ 1, & x \geqslant 1. \end{cases}$$

Совершенно ясно, что если нам попалась «хорошая» выборка (которая содержит такой же процент брака, как и вся контролируемая партия), то заменяя неизвестную вероятность $P\{\xi = 0\}$ частотой n_1/n, мы ошибемся мало. Если получена «плохая» выборка, то при подобной замене можно допустить ошибку. Однако «плохие» выборки будут попадаться тем реже, чем больше объем рассматриваемой выборки и, следовательно, ошибаться мы также будем редко. Пусть $n = 10$, $n_1 = 2$ и истинная вероятность $P\{\xi = 0\} = 0{,}4$ Мы же, изучая нашу выборку, положим $P\{\xi = 0\} = n_1/n = 0{,}2$. Вероятность встретить выборку, давшую основание для подобной оценки, будет

$$P\{S_{10} = 2\} = C_{10}^2(0{,}4)^2(0{,}6)^8 \approx 0{,}121,$$

т. е. примерно 12 выборок из 100 в данной ситуации будут плохими Если же заключение о том, что $P = 0{,}2$, мы сделали по выборке $n = 25$, $n_1 = 5$, то вероятность встретить плохую выборку

$$P\{S_{25} = 5\} \approx 0{,}002.$$

Для выборки $n = 100$, $n_1 = 20$ эта вероятность

$$P\{S_{100} = 20\} \approx 0{,}00002,$$

т. е. лишь в двух случаях из 100 000 мы получаем в эксперименте выборку, дающую основание для ложной оценки искомой вероятности. ▶

Поскольку вся информация о случайной величине ξ может быть получена при изучении ее функции распределения, а функции распределения случайных величин ξ и ξ_n оказались похожи, то следует ожидать, что аналогичная картина будет иметь место и для прочих характеристик случайной величины ξ. Таким образом, наблюдаемая в эксперименте случайная величина ξ_n служит своего рода «тенью» изучаемой случайной величины ξ. В дальнейшем мы будем называть ее *эмпирическим, или выборочным, аналогом* случайной величины ξ, а все характеристики величины ξ_n будем именовать *эмпирическими, или выборочными,* характеристиками случайной величины ξ. Так функция распределения случайной величины ξ_n называется эмпирической (выборочной) функцией распределения случайной величины ξ, математическое ожидание ξ_n — эмпирическим математическим ожиданием ξ, дисперсия ξ_n — эмпирической дисперсией ξ и вообще: ***эмпирическими (выборочными) моментами случайной величины ξ будем называть соответственно моменты случайной величины ξ_n.***

Эмпирические моменты будем обозначать той же буквой, что и соответствующие теоретические с добавлением вверху звездочки. Тогда эмпирические начальные моменты случайной величины ξ определяются формулой

$$m_k^* = M(\xi_n)^k = \frac{1}{n}\sum_{i=1}^{n} X_i^k, \tag{3}$$

а эмпирические центральные моменты —

$$\alpha_k^* = M(\xi_n - \bar{x})^k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{x})^k. \tag{4}$$

Для первого эмпирического начального момента (среднего значения) обычно используется обозначение

$$\bar{x} = m_1^* = \frac{1}{n}\sum X_i, \tag{5}$$

а для второго эмпирического центрального момента и среднеквадратичного отклонения обычно используют обозначения

$$S^2 = \alpha_2^* = \frac{1}{n}\sum(X_i - \bar{x})^2, \tag{6}$$

$$S = \sqrt{S^2} = \sqrt{\frac{1}{n}\sum(X_i - \bar{x})^2}. \tag{7}$$

Ими мы и будем пользоваться в дальнейшем.

Имеют место следующие утверждения.

Теорема 1. *Если n достаточно велико, то начальные эмпирические моменты мало отличаются от соответствующих теоретических, точнее*

$$\lim_{n\to\infty} P\left\{|m_k - m_k^*| > \varepsilon\right\} = 0 \quad \forall\, \varepsilon > 0.$$

◄ Пусть k — фиксировано. Рассмотрим последовательность случайных величин

$$\xi_1 = X_1^k, \quad \xi_2 = X_2^k, \quad \ldots, \quad \xi_n = X_n^k.$$

Поскольку все ξ_i — одинаковые независимые случайные величины, совпадающие со случайной величиной ξ^k, то

$$M\xi_i = M\xi^k = m_k, \quad i = 1, 2, \ldots, n.$$

Отсюда и из соотношения (3) следует равенство

$$P\left\{|m_k^* - m_k| > \varepsilon\right\} = P\left\{\left|\frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n} - M\xi\right| > \varepsilon\right\}.$$

Применяя к последовательности случайных величин ξ_i закон больших чисел в форме Хинчина, получаем, что

$$P\left\{|m_k^* - m_k| > \varepsilon\right\} = \lim P\left\{\left|\frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n} - M\xi\right| > \varepsilon\right\} = 0,$$

что и требовалось. ▶

В теореме, конечно, предполагается, что соответствующие теоретические моменты существуют.

Столь же легко может быть доказана и теорема о близости эмпирических центральных моментов к соответствующим теоретическим.

Теорема 2. *Если n достаточно велико, то центральные эмпирические моменты мало отличаются от соответствующих теоретических, точнее*

$$\lim_{n\to\infty} P\left\{|\alpha_k - \alpha_k^*| > \varepsilon\right\} = 0 \quad \forall\, \varepsilon > 0.$$

Пусть теперь в эксперименте наблюдается несколько случайных величин (случайный вектор) $\bar{\xi} = \left\{\xi_1, \xi_2, \ldots, \xi_l\right\}$.

***Выборкой объема** n **из закона распределения случайного вектора** $\bar{\xi}$* будем называть n реализаций (измерений) случайной величины $\bar{\xi}$

$$\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_n,$$

полученных в n независимых экспериментах.

Как и для случая одномерной случайной величины, выборка — это n штук независимых одинаково распределенных векторов. Отметим, что реализацией случайного вектора будет упорядоченный набор l чисел

$$\mathbb{X}_j = \begin{Bmatrix} X_1^j \\ X_2^j \\ \vdots \\ X_l^j \end{Bmatrix}, \quad j = 1,2, \ldots, n.$$

Известно, что важной характеристикой векторной случайной величины является ее ковариационная матрица

$$\mathbf{K}[\bar{\xi}] = M\left[\bar{\xi} - M\bar{\xi}\right]^T\left[\bar{\xi} - M\bar{\xi}\right] = \left(M(\xi_i - m_i)(\xi_j - m_j)\right)_{i,\,j=1}^{l},$$

элементы которой — ковариации компонент ξ_i и ξ_j. Ее эмпирический аналог (т. е. ковариационная матрица выборочного вектора $\bar{\xi}_n$, принимающего значение $\mathbb{X}_j$ с вероятностью $1/n$) называется эмпирической ковариационной матрицей вектора $\bar{\xi}$.

Как следует из вышеизложенного, ее компоненты могут быть найдены по формулам

$$\begin{aligned} k_{ij} &= \mathbf{K}\left[\xi_n^i, \xi_n^j\right] = M\left(\xi_n^i - M\xi_n^i\right)\left(\xi_n^j - M\xi_n^j\right) = \\ &= \frac{1}{n}\sum_{s=1}^{n}\left(X_s^i - \frac{1}{n}\sum_{r=1}^{n} X_r^i\right)\left(X_s^j - \frac{1}{n}\sum_{r=1}^{n} X_r^j\right) = \\ &= \frac{1}{n}\sum_{s=1}^{n} X_s^i X_s^j - \frac{1}{n}\sum_{s=1}^{n} X_s^i \cdot \frac{1}{n}\sum_{s=1}^{n} X_s^j. \end{aligned} \tag{8}$$

В частности, из соотношения (8) с учетом выражения для эмпирического среднеквадратичного отклонения (7) получаем соотношение для расчета эмпирического коэффициента корреляции пары случайных величин ξ и η

$$r_n = r_{\xi_n \eta_n} = \frac{\mathrm{K}[\xi_n, \eta_n]}{\sigma_{\xi_n}\sigma_{\eta_n}} = \frac{\sum\limits_{s=1}^{n}(X_s - \bar{x})(Y_s - \bar{y})}{\sqrt{\sum\limits_{s=1}^{n}(X_s - \bar{x})^2 \cdot \sum\limits_{s=1}^{n}(Y_s - \bar{y})^2}}, \tag{9}$$

здесь (X_s, Y_s), $s = 1,2, \ldots, n$, — выборка из двумерного распределения случайных величин (ξ, η), $\bar{x}$, $\bar{y}$ — эмпирические средние случайных величин ξ и η соответственно. Соотношение (9) может быть переписано в эквивалентной форме

$$r_n = \frac{\sum\limits_{s=1}^{n} X_s Y_s - \bar{x}\bar{y}}{\sqrt{\left(\sum\limits_{s=1}^{n}(X_s^2 - \bar{x}^2)\right)\left(\sum\limits_{s=1}^{n}(Y_s^2 - \bar{y}^2)\right)}}. \tag{10}$$

Как и выше, можно доказать теоремы о близости в подавляющем большинстве случаев эмпирических характеристик многомерной случайной величины к соответствующим характеристикам вектора $\bar{\xi}$. Эти теоремы позволяют высказать более или менее правдоподобное суждение о числовых характеристиках случайной величины ξ по выборке (1). Конечно, заменяя истинные числовых характеристики эмпирическими, можно ошибиться. Однако, как и в случае с функцией распределения, мы хотим надеяться, что «плохие» выборки будут встречаться редко и что в подавляющем большинстве случаев эмпирические моменты будут мало отличаться от теоретических. Хотелось бы научиться оценивать достоверность наших суждений о рассмотренных выше характеристиках случайной величины ξ поточнее. Этим мы займемся позднее, а сейчас попробуем несколько расширить наши представления о случайной величине ξ, составленные по выборке.

§ 2. Параметры распределений. Точечное оценивание

Пусть в эксперименте Ω изучается случайная величина $\bar{\xi} = \{\xi_1, \ldots, \xi_l\}$ с законом распределения $\Phi_{\bar{\xi}}(\mathbf{x}; \bar{\alpha}) = \Phi_{\xi_1\xi_2\ldots\xi_l}(x_1, x_2, \ldots, x_k; \alpha_1, \alpha_2, \ldots, \alpha_r)$, зависящим от некоторых параметров $\alpha_1, \alpha_2, \ldots, \alpha_r$.

Например, если случайная величина ξ — нормальная, то ее закон распределения зависит от двух параметров — m и σ, если ξ — равномерная на промежутке $[a, b]$, то параметрами закона распределения являются концы a и b, и т. д.

Пусть β — один из подобных параметров. Попробуем по выборке, полученной в результате эксперимента, высказать некоторое суждение о возможных значениях параметра β.

Для этого сначала следует указать способ вычисления величины β по выборке

$$\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_n, \tag{11}$$

т. е. функцию b от n векторных переменных такую, что

$$\beta \approx b(\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_n), \tag{12}$$

а потом пояснить, как в соотношении (12) понимать знак приближенного равенства.

Пусть, к примеру, β — это математическое ожидание случайной величины ξ. Тогда, как показано в предыдущем параграфе, в качестве функции b можно взять среднее арифметическое наблюденных значений и при этом понимать равенство в соотношении (12) как «равенство в большинстве случаев», т. е. вероятность больших отличий левой части от правой мала.

Формализуя вышеизложенное, скажем, что *оценкой неизвестного параметра β будем называть функцию $b(\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_n)$ от наблюденных значений такую, что $b \approx \beta$.*

Ясно, что нас будут интересовать не любые оценки параметра β, а только те, которые в некотором смысле на него похожи. Критериев «похожести», т. е. интерпретаций приближенного равенства в соотношении (12), существует много. Мы рассмотрим здесь наиболее употребительные.

Оценка $b(\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_n)$ называется *состоятельной оценкой* неизвестного параметра β, если вероятность отклонений β от $b(\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_n)$ становится малой с ростом n

$$\lim_{n\to\infty} P\left\{\left|\beta - b(\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_n)\right| > \varepsilon\right\} = 0. \tag{13}$$

В соответствии с этим определением (как следует из теорем 1 и 2) эмпирические моменты являются состоятельными оценками соответствующих теоретических моментов. Состоятельность — это *«похожесть в большинстве случаев»*.

Оценка $b(\mathbb{X}_1, \ldots, \mathbb{X}_n)$ называется *несмещенной* оценкой параметра β, если

$$Mb(\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_n) = \beta. \tag{14}$$

Несмещенность оценки есть ее *похожесть на оцениваемый параметр «в среднем»*, т. е., если мы обладаем несколькими выборками

$$\begin{array}{cccc} \mathbb{X}_1^1 & \mathbb{X}_2^1 & \ldots & \mathbb{X}_n^1; \\ \mathbb{X}_1^2 & \mathbb{X}_2^2 & \ldots & \mathbb{X}_n^2; \\ \ldots & \ldots & \ldots & \ldots \\ \mathbb{X}_1^k & \mathbb{X}_2^k & \ldots & \mathbb{X}_n^k \end{array}$$

и по каждой из них найдем оценку $b\left(\mathbb{X}_1^i, \mathbb{X}_2^i, \ldots, \mathbb{X}_n^i\right)$, то эти числа будут одинаково часто как превышать истинное значение оцениваемого параметра β, так и не превосходить его, т. е. отклонение оценки от оцениваемого параметра в случае несмещенности оценки носит несистематический характер.

Пример 1. Эмпирическая оценка математического ожидания является несмещенной оценкой

◀ Мы хотим доказать, что

$$M\bar{x} = M\xi.$$

Рассмотрим

$$M\bar{x} = M\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} MX_i.$$

Так как X_i — независимые в совокупности случайные величины, каждая из которых совпадает с ξ, то $MX_i = M\xi$ и мы получаем

$$M\bar{x} = \frac{1}{n}\sum_{i=1}^{n} M\xi = M\xi. \ ▶$$

Пример 2. Эмпирическая оценка дисперсии

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{x})^2$$

не обладает свойством несмещенности.

◀ Действительно,

$$S^2 = \frac{1}{n}\sum_{i=0}^{n}(X_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2.$$

Поэтому

$$MS^2 = \frac{1}{n} M\left(\sum_{i=1}^{n} X_i^2\right) - \frac{1}{n^2} M\left(\sum_{i=1}^{n} X_i\right)^2 = M\xi^2 - \frac{1}{n} M\xi^2 - \frac{n-1}{n} (M\xi)^2 =$$
$$= \frac{n-1}{n} \left(M\xi^2 - M^2\xi\right) = \frac{n-1}{n} D\xi \neq D\xi. \blacktriangleright$$

Несмещенную оценку дисперсии можно легко построить. Как показано выше,

$$MS^2 = \frac{n-1}{n} D\xi.$$

Поэтому

$$M\left(\frac{n}{n-1} S^2\right) = D\xi$$

и, следовательно, исправленная величина s^2, определяемая соотношением

$$s^2 = \frac{n}{n-1} S^2 = \frac{1}{n-1} \sum_{i=1}^{n}(X_i - \bar{x})^2,$$

есть несмещенная оценка дисперсии.

С практической точки зрения свойство состоятельности очень важно — его наличие позволяет надеяться, что с увеличением объема выборки точность оценивания будет расти (конечно, только в подавляющем большинстве случаев). Несмещенность же играет менее важную роль. Если оценка является несмещенной, то это свидетельствует об отсутствии *систематической* ошибки в оценивании неизвестного параметра. Указанное обстоятельство становится важным в случае малых выборок, когда оценки могут быть далеки от оцениваемого параметра и наличие систематической погрешности оценивания только ухудшает точность оценивания. В случае больших выборок смещение оценки (при наличии состоятельности!) на точность оценивания существенного влияния не оказывает.

Важно также понимать, что вышеизложенное имеет смысл только если выполнены условия применимости законов больших чисел, на выводах из которых базируются наши заключения. Важнейшим из подобных условий является существование математического ожидания исследуемой случайной величины ξ.

Рассмотрим некоторые методы нахождения оценок неизвестных параметров распределения, сделав дополнительное предположение, а именно: пусть вид функции распределения случайной величины ξ известен

$$F_\xi(x) = F_\xi(x;\, \beta_1, \beta_2, \ldots, \beta_k).$$

Итак, в результате эксперимента получена выборка объема n. Требуется по выборке найти оценки неизвестных параметров $\beta_1, \beta_2, \ldots, \beta_n$.

2.1. Метод моментов

Идея метода моментов состоит в приравнивании эмпирических моментов, найденных по выборке, соответствующим теоретическим, которые зависят от неизвестных параметров $\beta_1, \beta_2, \ldots, \beta_k$:

$$
\begin{cases}
m_1^* = \dfrac{1}{n}\sum X_i = M_1(\beta_1, \beta_2, \ldots, \beta_k), \\
m_2^* = \dfrac{1}{n}\sum X_i^2 = M_2(\beta_1, \beta_2, \ldots, \beta_k), \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
m_k^* = \dfrac{1}{n}\sum X_i^k = M_k(\beta_1, \beta_2, \ldots, \beta_k).
\end{cases}
\tag{15}
$$

Система (15) позволяет выразить неизвестные параметры $\beta_1, \beta_2, \ldots, \beta_k$ через выборочные значения $X_1, X_2, \ldots, X_n$:

$$
\beta_1 = b_1(X_1, \ldots, X_n), \quad \beta_2 = b_2(X_1, \ldots, X_n), \quad \ldots, \quad \beta_k = b_k(X_1, \ldots, X_n).
$$

Функции $b_i(X_1, \ldots, X_n)$ и считаются оценками параметров β_i. Близость оценок, найденных по методу моментов, к истинным значениям оцениваемых параметров описывается следующей теоремой.

Теорема. *Пусть решение системы* (15) *существует, причем функции*

$$
b_j(X_1, \ldots, X_n) = \tilde{b}_j(m_1^*, \ldots, m_k^*)
$$

непрерывны в точке $m = (m_1, m_2, \ldots, m_k)$. *Тогда оценки, полученные по методу моментов, состоятельны.*

Оценки, полученные по методу моментов, необязательно являются несмещенными. Однако, если наложить на функции $\tilde{b}_j(m_1, \ldots, m_k)$ некоторые дополнительные ограничения, то можно получить утверждение, касающееся асимптотической несмещенности оценок, найденных методом моментов

$$
Mb_j(X_1, \ldots, X_n) = \beta_j + O\left(\frac{1}{n}\right), \quad j = 1, \ldots, k, \tag{16}
$$

т. е. смещение оценки с ростом объема выборки убывает.

На практике метод моментов приводит к относительно простым вычислениям и, как следует из теоремы, позволяет находить состоятельные оценки параметров. Смещение этих оценок для больших выборок несущественно (16). Кроме того, во всех практически важных случаях это смещение легко устраняется с помощью простых поправок.

Пример 1. Известно, что случайная величина ξ равномерно распределена на отрезке $[\alpha, \beta]$. Получена выборка объема n из распределения случайной величины ξ. Оценить величины α и β.

◀ Произведем оценку неизвестных параметров α и β, пользуясь методом моментов. Имеем

$$
M\xi = \int_\alpha^\beta \frac{x\,dx}{\beta - \alpha} = \frac{\alpha + \beta}{2};
$$

$$
M\xi^2 = \int_\alpha^\beta \frac{x^2\,dx}{\beta - \alpha} = \frac{\alpha^2 + \alpha\beta + \beta^2}{3}.
$$

Система (15) в данном случае принимает вид

$$
\begin{cases}
\dfrac{1}{n}\sum X_i = \dfrac{\alpha + \beta}{2}, \\
\dfrac{1}{n}\sum X_i^2 = \dfrac{\alpha^2 + \alpha\beta + \beta^2}{3}.
\end{cases}
$$

Решая ее относительно α и β, получаем

$$\beta^* = \bar{x} + \sqrt{3}S;$$
$$\alpha^* = \bar{x} - \sqrt{3}S.$$

Учитывая, что $\alpha < \beta$, заключаем, что наша система всегда имеет решение и притом единственное. Полученные оценки состоятельны, однако свойством несмещенности не обладают. ▶

Пример 2. Оценить по выборке параметр μ экспоненциально распределенной случайной величины ξ.

◀ Функция распределения случайной величины ξ имеет вид

$$F_\xi(x) = \begin{cases} 1 - e^{\mu x}, & x \geqslant 0, \\ 0, & x < 0. \end{cases}$$

Следовательно,

$$M\xi = \int_0^\infty \mu e^{-\mu x} x\, dx = \frac{1}{\mu}.$$

Система (15) сводится к одному уравнению

$$\frac{1}{n} \sum X_i = \frac{1}{\mu},$$

откуда

$$\mu^* = \frac{1}{\bar{x}}. \tag{17}$$

Полученная оценка состоятельна. Что касается несмещенности, то поскольку X_i экспоненциально распределены, то $n\bar{x} = \sum x_i$ имеет гамма-распределение с плотностью

$$f_{n\bar{x}}(t) = \begin{cases} \dfrac{\mu^n t^{n-1}}{(n-1)!}\, e^{-\mu t}, & t \geqslant 0, \\ 0, & t < 0, \end{cases}$$

и поэтому

$$M\left(\frac{1}{n\bar{x}}\right) = \int_0^\infty \frac{1}{t} f_{n\bar{x}}(t)\, dt = \frac{\mu}{n-1}.$$

Следовательно, оценка (17) свойством несмещенности не обладает, так как

$$M(\mu^*) = nM\left(\frac{1}{n\bar{x}}\right) = \frac{n}{n-1}\,\mu \neq \mu. \tag{18}$$

Однако, используя соотношение (18), легко можно получить несмещенную оценку параметра μ

$$\mu_1^* = \frac{n-1}{n}\,\mu^* = \frac{n-1}{\sum X_i}. \ ▶$$

2.2. Метод максимального правдоподобия

Пусть ξ — непрерывная случайная величина с плотностью

$$f_\xi(x) = f_\xi(x;\, \beta_1, \beta_2, \ldots, \beta_k).$$

Вид плотности известен, но неизвестны значения параметров β_j.

Функцией правдоподобия называется функция

$$L(X_1, X_2, \ldots, X_n;\ \beta_1, \beta_2, \ldots, \beta_k) = \prod_{i=1}^{n} f_\xi(X_i;\ \beta_1, \beta_2, \ldots, \beta_k).$$

(здесь $X_1, \ldots, X_n$ — выборка объема n из распределения случайной величины ξ). Легко видеть, что функции правдоподобия можно придать вероятностный смысл, а именно: рассмотрим случайный вектор $\Xi = \{\xi_1, \ldots, \xi_n\}$, компоненты которого независимые в совокупности одинаково распределенные случайные величины с законом $f_\xi(x)$. Тогда элемент вероятности вектора Ξ имеет вид

$$P\{\Xi \in \Delta V\} = L(X_1, \ldots, X_n;\ \beta_1, \ldots, \beta_k)\Delta X_1 \ldots \Delta X_n,$$
$$\Delta V = \{\Xi\colon\ X_i \leqslant \xi_i < X_i + \Delta X_i\} \quad i = 1, 2, \ldots, n,$$

т. е. функция правдоподобия связана с вероятностью получения фиксированной выборки в последовательности экспериментов Ω.

Основная идея метода правдоподобия состоит в том, что в качестве оценок параметров β_i предлагается взять такие значения β_j^*, которые доставляют максимум функции правдоподобия при данной фиксированной выборке, т. е. предлагается считать выборку, полученную в эксперименте, наиболее вероятной. Нахождение оценок параметров β_j сводится к решению системы k уравнений (k — число неизвестных параметров):

$$\frac{\partial L(X_1, \ldots, X_n;\ \beta_1, \ldots, \beta_k)}{\partial \beta_i} = 0; \quad i = 1,2, \ldots, k. \tag{19}$$

Поскольку функция $\log L$ имеет максимум в той же точке, что и функция правдоподобия, то часто систему уравнений правдоподобия (19) записывают в виде

$$\frac{\partial \log L}{\partial \beta_i} = \sum_{j=1}^{n} \frac{\partial f_\xi(X_j;\ \beta_1, \ldots, \beta_k)}{\partial \beta_i} \cdot \frac{1}{f_\xi(X_j;\ \beta_1, \ldots, \beta_k)} = 0. \tag{20}$$

В качестве оценок неизвестных параметров β_i следует брать решения системы (19) или (20), действительно зависящие от выборки и не являющиеся постоянными.

В случае, когда ξ дискретна с рядом распределения $P\{\xi = X_i\} = P\{X_i;\ \beta_1, \ldots, \beta_k\}$, $i = 1, \ldots, n$, функцией правдоподобия называют функцию

$$L(X_1, \ldots, X_n;\ \beta_1, \ldots, \beta_k) = \prod_{i=1}^{n} P\{\xi = X_i;\ \beta_1, \ldots, \beta_k\} \tag{21}$$

и оценки ищут как решения системы

$$\frac{\partial L(X_1, \ldots, X_n;\ \beta_1, \ldots, \beta_k)}{\partial \beta_i} = 0, \quad i = 1,2, \ldots, k, \tag{22}$$

или эквивалентной ей

$$\frac{\partial \log L}{\partial \beta_i} = \sum_{j=1}^{n} \frac{\partial P\{\xi = X_j;\ \beta_1, \ldots, \beta_k\}}{\partial \beta_i} \cdot \frac{1}{P\{\xi = X_j;\ \beta_1, \ldots, \beta_k\}} = 0. \tag{23}$$

Можно показать, что оценки максимального правдоподобия обладают свойством состоятельности. Следует отметить, что метод максимального правдоподобия приводит к более сложным вычислениям, нежели метод моментов, но теоретически он более эффективен, так как оценки максимального правдоподобия меньше уклоняются от истинных значений оцениваемых параметров, чем оценки, полученные по методу моментов.

Для наиболее часто встречающихся в приложениях распределений оценки параметров, полученные по методу моментов и по методу максимального правдоподобия, в большинстве случаев совпадают.

Пример 1. Отклонение ξ размера детали от номинала является нормально распределенной случайной величиной. Требуется по выборке определить систематическую ошибку и дисперсию отклонения.

◀ По условию ξ — нормально распределенная случайная величина с математическим ожиданием (систематическая ошибка) и дисперсией, подлежащими оценке по выборке объема n: $X_1, \ldots, X_n$. В этом случае

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} = f_\xi(x;\ m, \sigma).$$

Функция правдоподобия

$$L(X_1, \ldots, X_n; m, \sigma) = \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left\{-\frac{\sum(X_i - m)^2}{2\sigma^2}\right\}.$$

Система (19) имеет вид

$$\begin{cases} \dfrac{\partial L}{\partial m} = \dfrac{1}{\sigma^n (2\pi)^{n/2}} \exp\left\{-\dfrac{\sum (X_i - m)^2}{2\sigma^2}\right\}\left(-\dfrac{1}{\sigma^2}\sum (X_i - m)\right) = 0, \\ \dfrac{\partial L}{\partial \sigma} = \dfrac{1}{\sigma^{n+1}(2\pi)^{n/2}} \exp\left\{-\dfrac{\sum (X_i - m)^2}{2\sigma^2}\right\}\left(-n + \dfrac{1}{\sigma^2}\sum (X_i - m)^2\right) = 0. \end{cases}$$

Отсюда, исключая решения, не зависящие от X_i, получаем

$$m^* = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{x}, \quad (\sigma^*)^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2 = S^2,$$

т. е. оценки максимального правдоподобия в этом случае совпадают с уже известными нам эмпирическими средним и дисперсией ▶

Пример 2. Оценить по выборке параметр μ экспоненциально распределенной случайной величины.

◀ Функция правдоподобия имеет вид

$$L(X_1, \ldots, X_n; \mu) = \mu^n e^{-\mu \sum X_i}.$$

Уравнение правдоподобия

$$\frac{\partial L}{\partial \mu} = n\mu^{n-1} e^{-\mu \sum X_i} - \mu^n \left(\sum X_i\right) e^{-\mu \sum X_i} = 0$$

приводит нас к решению

$$\mu^* = \frac{n}{\sum X_i} = \frac{1}{\bar{x}},$$

совпадающему с оценкой этого же параметра, полученной по методу моментов, см. (17). ▶

Пример 3. Пользуясь методом максимального правдоподобия, оценить вероятность появления герба, если при десяти бросаниях монеты герб появился 8 раз.

◀ Пусть подлежащая оценке вероятность равна p. Рассмотрим случайную величину ξ с рядом распределения

$$P\{\xi = 0; p\} = p; \quad P\{\xi = 1; p\} = 1 - p.$$

Функция правдоподобия (21) имеет вид

$$L(X_1, \ldots, X_n; p) = p^8 (1-p)^2,$$

так как

$$P\{\xi = X_i; p\} = \begin{cases} p, & \text{если } X_i = 0, \\ 1 - p, & \text{если } X_i = 1. \end{cases}$$

Уравнение правдоподобия

$$8p^7(1-p)^2 - 2p^8(1-p) = 0$$

дает в качестве оценки неизвестной вероятности p частоту появления герба в эксперименте

$$p^* = 0{,}8. \ ▶$$

Заканчивая обсуждение методов нахождения оценок, подчеркнем, что, даже имея очень большой объем экспериментальных данных, мы все равно не можем указать точного значения оцениваемого параметра, более того, как уже неоднократно отмечалось, получаемые нами оценки близки к истинным значениям оцениваемых параметров только «в среднем» или «в большинстве случаев». Поэтому важной статистической задачей, которую мы рассмотрим далее, является задача определения точности и достоверности проводимого нами оценивания.

§ 3. Интервальное оценивание

Результаты предыдущего параграфа позволяют по выборке определить оценку неизвестного параметра β-распределения $F_\xi(x)$. Эти оценки носят точечный характер — они указывают число, в некотором смысле похожее на оцениваемый параметр, другими словами, они позволяют определить точку β^*, находящуюся в большей или меньшей близости к истинному значению оцениваемого параметра (рис. 1).

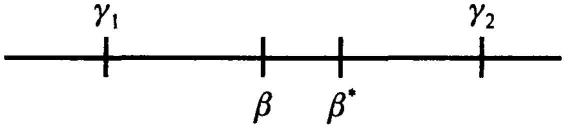


Рис. 1

Пусть нам удалось построить две функции $\gamma_1(X_1, \ldots, X_n)$ и $\gamma_2(X_1, \ldots, X_n)$, удовлетворяющие условию

$$\gamma_1(X_1, \ldots, X_n) \leqslant \gamma_2(X_1, \ldots, X_n)$$

для любых значений $X_1, X_2, \ldots, X_n$. Рассмотрим на числовой оси промежуток (γ_1, γ_2). Его концы γ_1 и γ_2 зависят от выборочных значений и, следовательно, являются случайными величинами. Вследствие этого промежуток (γ_1, γ_2) также является случайным в том смысле, что его длина и положение на числовой прямой зависят от выборки $X_1, X_2, \ldots, X_n$. Истинное значение параметра β — неслучайное число и его положение на числовой оси фиксировано. Поэтому для некоторых выборок случайный интервал (γ_1, γ_2) будет накрывать число β, а для некоторых не будет. Если нам удается подобрать функции $\gamma_1(X_1, \ldots, X_n)$ и $\gamma_2(X_1, \ldots, X_n)$ так, что случайный интервал (γ_1, γ_2) «часто» накрывает истинное значение неизвестного параметра β, то мы можем в качестве оценки этого параметра взять любую точку интервала (γ_1, γ_2). При этом можно утверждать, что «довольно часто» наша оценка отличается от оцениваемого параметра не более чем на длину интервала (γ_1, γ_2).

Введем новое понятие, формализующее выше приведенные рассуждения.

Доверительным интервалом для параметра β называется случайный интервал (γ_1, γ_2) такой, что

$$P\{(\gamma_1, \gamma_2) \ni \beta\} \geqslant \varkappa. \tag{24}$$

Число $\varkappa$ при этом называется ***уровнем доверия*** или ***доверительной вероятностью***.

Выбор числа $\varkappa$ — уровня доверия — зависит от того, что мы понимаем под словами «довольно часто», и от того, какой точности в определении параметра β мы хотим достичь. Поскольку добиться абсолютной достоверности (чтобы ошибка не превышала длины интервала всегда) мы не можем[1)], то поступимся достоверностью, чтобы получить нетривиальную информацию о точности. Выбирая $\varkappa$ очень маленьким, мы, конечно, можем добиться того, чтобы длина интервала (γ_1, γ_2) была сколь угодно малой, однако в этом случае (из-за малости $\varkappa$ и неравенства (24)) мы крайне редко будем получать доверительный интервал, накрывающий истинное значение параметра β, т. е. найденные нами по конкретным выборкам интервалы будут ненадежны. Выбор же $\varkappa$ очень близкого к единице, неоправданно расширяет границы доверительного интервала и тем самым понижает точность определения параметра β. Поскольку чаще всего нас интересует вопрос, как сильно мы можем ошибиться, заменяя истинное значение параметра β его оценкой β^*, то обычно доверительную вероятность $\varkappa$ выбирают настолько близкой к единице, чтобы с событиями, вероятность которых меньше, чем $1 - \varkappa$, можно было практически не считаться. Соответствующий этой вероятности доверительный интервал дает надежную (с вероятностью $\varkappa$) оценку отличия приближенного значения β^* от неизвестного точного β. На практике в качестве $\varkappa$ в зависимости от конкретной ситуации выбирают одно из чисел — 0,9; 0,95; 0,99; 0,999.

Возвращаясь к задаче определения точности и достоверности оценки β^* неизвестного параметра β, отметим, что определить точность оценки — значит указать, как велика может быть разница

$$\beta - \beta^*. \tag{25}$$

[1)] Из неравенства (24) следует, что при этом (γ_1, γ_2) расползается на всю числовую прямую и мы получаем тривиальный результат: ошибка $\gamma_2 - \gamma_1$ в распределении параметра может быть любой.

Но в силу того, что β^* — случайная величина, разность (25) — также случайная величина и может принимать любые значения, причем одни чаще, другие реже. Поэтому тесно связанной с определением точности является задача определения достоверности оценки, т. е. указания той доли случаев, когда величина $\beta - \beta^*$ не превосходит некоторой величины ε. Суммируя, получаем, что определить точность и достоверность оценки β^* — значит указать числа ε_1, ε_2 и $\varkappa$ такие, что

$$P\{\varepsilon_1 \leqslant \beta - \beta^* \leqslant \varepsilon_2\} \geqslant \varkappa, \tag{26}$$

т. е. задача определения точности и достоверности оценки β^* — это задача построения доверительного интервала для параметра β.

Заметим, что при $\varepsilon_1 = -\varepsilon_2$ доверительный интервал оказывается симметричным относительно точечной оценки β^*.

3.1. Точность и надежность оценивания математического ожидания нормальной случайной величины

Пусть ξ — нормальная случайная величина с параметрами (m, σ), $\bar{x}$ — эмпирическая оценка параметра $m = M\xi$. Существенным для дальнейшего является вопрос о том, известна или нет дисперсия.

1. Пусть σ известна.

В силу нормальности ξ отклонение оценки $\bar{x}$ от $M\xi$ также является нормальной случайной величиной с параметрами

$$M(\bar{x} - M\xi) = 0; \quad D(\bar{x} - M\xi) = \frac{D\xi}{n} = \frac{\sigma^2}{n}; \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

При $-\varepsilon_1 = \varepsilon_2 = \varepsilon > 0$ соотношение (26) примет вид

$$\varkappa \leqslant P\{|\bar{x} - M\xi| < \varepsilon\} = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} \int_{-\varepsilon}^{\varepsilon} e^{-t^2 n/2\sigma^2}\, dt = 2F\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - 1. \tag{27}$$

Здесь $F(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} e^{-x^2/2}\, dx$ — функция Лапласа. Таким образом, задача свелась к решению уравнения

$$2F\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - 1 = \varkappa \tag{28}$$

относительно ε при заданном уровне доверия $\varkappa$. Обозначим решение уравнения

$$2F(x) - 1 = \varkappa$$

через $x_\varkappa$. Тогда

$$\varepsilon_\varkappa = x_\varkappa \frac{\sigma}{\sqrt{n}} \tag{29}$$

— решение уравнения (28). Искомый доверительный интервал —

$$\bar{x} - x_\varkappa \frac{\sigma}{\sqrt{n}} \leqslant M\xi \leqslant \bar{x} + x_\varkappa \frac{\sigma}{\sqrt{n}}, \tag{30}$$

так что в $\varkappa \cdot 100$ % случаев неизвестное значение $M\xi$ накрывается интервалом (30), т. е. точность в определении $M\xi$ не превышает по модулю величины $x_\varkappa \sigma/\sqrt{n}$ в $\varkappa \cdot 100$ % случаев. Правда, в $(1 - \varkappa) \cdot 100$ % случаев найденное нами среднее арифметическое $\bar{x}$

может отличаться от $M\xi$ на сколь угодно большую величину, однако за счет того, что события с вероятностью $1-\varkappa$ практически невозможны, этим можно пренебречь.

Отметим, что соотношение (30) — точное, т. е. справедливо для любых объемов экспериментальных данных, в том числе и для малых выборок.

2. Пусть теперь σ неизвестна.

В этом случае рассуждения предыдущего пункта мы применить не можем, так как в соотношении (27) значение параметра σ нам неизвестно, и мы получим одно уравнение с двумя неизвестными ε и σ.

Рассмотрим величину

$$t=\frac{\bar{x}-M\xi}{s}\sqrt{n}. \tag{31}$$

Здесь s — исправленная оценка среднеквадратичного отклонения

$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{x})^2. \tag{32}$$

Отметим следующее, важное для дальнейшего, обстоятельство: *случайные величины $\bar{x}$ и s^2 — статистически независимы.*

◀ Действительно, случайный вектор

$$\overline{X}=\begin{pmatrix} x_1-\bar{x}\\ x_2-\bar{x}\\ \cdots\cdots\\ x_n-\bar{x}\end{pmatrix}=\bar{x}-M\xi$$

имеет нормальное распределение, при этом

$$\begin{aligned}\forall j:\ M\left[(x_j-\bar{x})(\bar{x}-M\xi)\right]&=M(X_j-m)\,\frac{1}{n}\sum_{i=1}^{n}(X_i-m)-M(\overline{X}-m)^2=\\ &=M(X_j-m)\,\frac{1}{n}\sum_{i\neq j}(x_i-m)+\frac{1}{n}\,M(X_j-m)^2-\frac{1}{n}\,\sigma^2=\\ &=\frac{1}{n}\,\sigma^2-\frac{1}{n}\,\sigma^2=0.\end{aligned}$$

Откуда и следует искомое, так как независимость случайного вектора $\overline{X}$ и разности $\bar{x}-m$ влечет независимость случайных величин, являющихся их непрерывными функциями. ▶

Для величины t, задаваемой соотношением (31), докажем теперь следующую теорему.

Теорема. *Случайная величина*

$$t=\frac{x-M\xi}{s}\sqrt{n}$$

подчиняется распределению Стьюдента с $n-1$ степенью свободы.

◀ Заметим, что

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - M\xi)^2 - \frac{1}{n-1}(\bar{x} - M\xi)^2,$$

$$(\bar{x} - M\xi)\sqrt{n} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i - M\xi).$$

Поэтому

$$t = \frac{\bar{x} - M\xi}{S}\sqrt{n} = \frac{\bar{x} - M\xi}{S}\sqrt{n}\,\frac{\sqrt{n-1}}{\sqrt{n-1}}\,\frac{\sigma}{\sigma} = \frac{\dfrac{\bar{x} - M\xi}{\sigma}\sqrt{n}}{\dfrac{S\sqrt{n-1}}{\sigma}}\sqrt{n-1} =$$

$$= \frac{\dfrac{1}{\sqrt{n}}\sum\left(\dfrac{X_i - M\xi}{\sigma}\right)\sqrt{n-1}}{\sqrt{\sum\left(\dfrac{X_i - M\xi}{\sigma}\right)^2 - \dfrac{1}{n}\left(\sum\left(\dfrac{X_i - M\xi}{\sigma}\right)\right)^2}} = \frac{\left(\dfrac{1}{\sqrt{n}}\sum Y_i\right)\sqrt{n-1}}{\sqrt{\sum Y_i^2 - \left(\dfrac{1}{\sqrt{n}}\sum Y_i\right)^2}}. \tag{33}$$

В соотношении (33) через Y_i обозначены независимые нормально распределенные случайные величины с параметрами $MY_i = 0$ и $DY_i = 1$

$$Y_i = \frac{X_i - M\xi}{\sigma}.$$

Рассмотрим в n-мерном координатном пространстве $\mathbb{R}^n$ гиперплоскость π, задаваемую уравнением

$$Y_1 + Y_2 + \ldots + Y_n = 0. \tag{34}$$

Сделаем поворот осей в $\mathbb{R}^n$ таким образом, чтобы одна из новых координатных осей (для определенности последняя) была бы ортогональна плоскости π. При этом координаты Y_i перейдут в Z_i и

$$\sum_{i=1}^{n} Y_i^2 = \sum_{i=1}^{n} Z_i^2. \tag{35}$$

Поскольку Y_i — независимые нормально распределенные случайные величины с параметрами 0 и 1, то и Z_i также будут независимыми нормально распределенными случайными величинами с параметрами 0 и 1. Из условия ортогональности одной из новых осей плоскости π вытекает, что соответствующая ей новая координата Z_n будет иметь вид

$$Z_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i \tag{36}$$

и

$$t = \frac{Z_n}{\sqrt{\sum\limits_{i=1}^{n-1} Z_i^2}}\sqrt{n-1}, \tag{37}$$

так как

$$\sum_{i=1}^{n} Y_i^2 - \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i\right)^2 = \sum_{i=1}^{n-1} Z_i^2.$$

Учитывая независимость Z_j, заключаем, что t имеет распределение Стьюдента с $n - 1$ степенью свободы. ▶

Возвратимся к определению доверительного интервала для $M\xi$,

$$P\{|\bar{x} - M\xi| < \varepsilon\} = P\left\{\frac{|\bar{x} - M\xi|}{S}\sqrt{n} < \frac{\varepsilon\sqrt{n}}{S}\right\} = S_{n-1}\left(-\frac{\varepsilon\sqrt{n}}{S}\right) - S_{n-1}\left(-\frac{\varepsilon\sqrt{n}}{S}\right).$$

Здесь $S_{n-1}(t)$ — функция распределения Стьюдента. Последнее уравнение, используя свойство функции S_{n-1}

$$S_{n-1}(-t) = 1 - S_{n-1}(t),$$

перепишем в виде

$$P\{|\bar{x} - M\xi| < \varepsilon\} = 2S_{n-1}\left(\frac{\varepsilon\sqrt{n}}{S}\right) - 1.$$

Задавая уровень доверия $\varkappa$ и обозначая решение уравнения

$$2S_{n-1}(t) - 1 = \varkappa$$

через $t_\varkappa$, получаем доверительные границы

$$\bar{x} - t_\varkappa\frac{S}{\sqrt{n}} \leqslant M\xi \leqslant \bar{x} + t_\varkappa\frac{S}{\sqrt{n}}. \tag{38}$$

Заметим, что доверительные границы для математического ожидания в случае известной дисперсии σ^2 имеют такой же вид. В соотношении (38) вместо среднеквадратического отклонения σ стоит оценка (32) среднеквадратического отклонения s, вместо $x_\varkappa$ — решения уравнения $2F(x) - 1 = \varkappa$ — стоит $t_\varkappa$ — решение уравнения $2S_{n-1}(t) - 1 = \varkappa$. Доверительные границы (38), вообще говоря, шире доверительных границ (30), что объясняется большей долей неопределенности при нахождении $\bar{x}$ по выборке в случае, когда дисперсия неизвестна, по сравнению со случаем, когда дисперсия известна.

Как и в случае известной σ, интервал (38) — точный и может быть использован для оценивания математического ожидания по выборкам любого объема, в том числе и по малым выборкам.

3.2. Точность и надежность оценивания дисперсии нормальной случайной величины

Пусть ξ — нормальная случайная величина с параметрами (m, σ), которые оцениваются по выборке объема n

$$m \approx \bar{x} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \sigma^2 \approx s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{x})^2. \tag{39}$$

Величина $Z_j = \frac{X_j - \bar{x}}{\sigma\sqrt{n-1}}\sqrt{n}$ нормальна с параметрами $(0,1)$. Действительно, Z_j является линейной комбинацией нормальных величин X_s, $s = 1, \dots, n$,

$$Z_j = \frac{\sqrt{n}}{\sqrt{n-1}\sigma}\left(X_j - \frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{\sqrt{n}}{\sqrt{n-1}\sigma}\left(\frac{n-1}{n}X_j - \frac{1}{n}\sum_{i\neq j} X_i\right)$$

с параметрами $MX_s = m$, $DX_s = \sigma^2$, а потому Z_j — нормальна. Далее, очевидно, $MZ_j = 0$ и

$$DZ_j = \frac{n}{(n-1)\sigma^2}\left[\left(\frac{n-1}{n}\right)^2 DX_j + \frac{1}{n^2}\sum_{i\neq j} DX_i\right] = \frac{n-1}{n} + \frac{1}{n} = 1.$$

Исправленная оценка дисперсии s^2 представляется в виде

$$s^2 = \frac{\sigma^2}{n} \sum_{j=1}^{n} Z_j^2.$$

Можно установить, что величина $\frac{n-1}{n} \sum\limits_{j=1}^{n} Z_j^2$ имеет распределение χ^2 с $n-1$ степенью свободы, т. е. распределена как сумма квадратов $n-1$ независимых нормальных $(0,1)$ случайных величин, что дает возможность вычислять вероятности

$$P\left\{\frac{s^2}{\sigma^2}(n-1) < t\right\} = G_{n-1}(t). \tag{40}$$

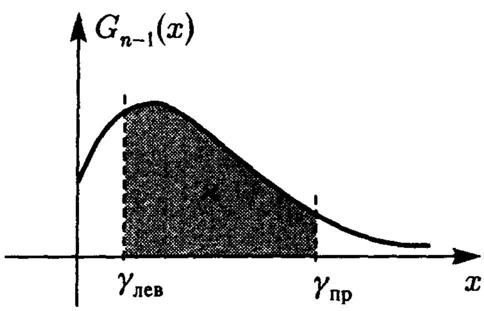


Рис. 2

Зададим некоторую вероятность $\varkappa$, близкую к единице, и найдем числа $\gamma_{\text{пр}} \leqslant \gamma_{\text{лев}}$ такие, что

$$P\left\{\frac{s^2}{\sigma^2}(n-1) < \gamma_{\text{лев}}\right\} \leqslant \frac{1-\varkappa}{2},$$

$$P\left\{\frac{s^2}{\sigma^2}(n-1) > \gamma_{\text{пр}}\right\} \leqslant \frac{1-\varkappa}{2}$$

(рис. 2). При этом для неизвестной дисперсии σ^2 получим

$$\frac{(n-1)s^2}{\gamma_{\text{пр}}} < \sigma^2 < \frac{(n-1)s^2}{\gamma_{\text{лев}}}. \tag{41}$$

Числа $\gamma_{\text{пр}}$ и $\gamma_{\text{лев}}$ легко определяются как решения уравнений

$$G_{n-1}(\gamma_{\text{лев}}) = \frac{1-\varkappa}{2}; \quad G_{n-1}(\gamma_{\text{пр}}) = \frac{1+\varkappa}{2}.$$

Теперь в качестве точечной оценки неизвестной дисперсии можно взять любое число ${\sigma^*}^2$ из промежутка (41) и с надежностью $\varkappa$ точность такого оценивания будет не хуже, чем

$$\min\left\{\left|\frac{(n-1)s^2}{\gamma_{\text{пр}}} - {\sigma^*}^2\right|, \left|\frac{(n-1)s^2}{\gamma_{\text{лев}}} - {\sigma^*}^2\right|\right\}.$$

Для симметричной оценки,

$${\sigma^*}^2 = \frac{1}{2}(n-1)s^2 \frac{\gamma_{\text{пр}} + \gamma_{\text{лев}}}{\gamma_{\text{пр}}\gamma_{\text{лев}}},$$

точность $\varepsilon = \dfrac{1}{2}(n-1)s^2 \dfrac{\gamma_{\text{пр}} - \gamma_{\text{лев}}}{\gamma_{\text{пр}}\gamma_{\text{лев}}}$.

Отметим, что если в качестве точечной оценки взять исправленную оценку дисперсии s^2, то интервал (41) относительно этой оценки симметричным не будет.

3.3. Точность и надежность оценивания для негауссовских распределений

Результаты предыдущего параграфа позволяют найти точные доверительные интервалы для математического ожидания и дисперсии нормальной случайной величины, при условии ее нормальности. Если же случайная величина ξ имеет произвольную функцию распределения, то это удается уже не всегда. Однако, если объем выборки достаточно велик, то можно указать приближенные доверительные интервалы границы для моментов случайной величины ξ, используя следующее утверждение.

Теорема. *Пусть*

$$m_s^* = \frac{1}{n}\sum_{i=1}^{n} X_i^s$$

— эмпирическая оценка момента s-го порядка случайной величины ξ, существование которого предполагается. Тогда случайная величина

$$\eta_n = \frac{m_s^* - m_s}{\sqrt{D\xi^s}}\sqrt{n} \tag{42}$$

распределена асимптотически нормально с параметрами 0 *и* 1,

$$\lim_{n\to\infty} P\{\eta_n < x\} = F(x). \tag{43}$$

Отсюда следует, что

$$P\{|m_s^* - m_s| < \varepsilon\} \approx 2F\left(\frac{\varepsilon\sqrt{n}}{\sqrt{D\xi^s}}\right) - 1.$$

Задавая уровень доверия $\varkappa$ и решая уравнение

$$2F(x) - 1 = \varkappa,$$

получаем значение $x_\varkappa$ такое, что

$$m_s^* - x_\varkappa\frac{\sqrt{D\xi^s}}{\sqrt{n}} \leqslant m_s \leqslant m_s^* + x_\varkappa\frac{\sqrt{D\xi^s}}{\sqrt{n}} \tag{44}$$

с вероятностью $\varkappa$. И, если известна $D\xi^s$, то соотношение (44) дает искомый доверительный интервал.

Замечание. Нормальное распределение дает плохое приближение к истинному распределению суммы ξ_i в области вероятностей очень малых или очень близких к единице. Поэтому доверительные интервалы (44) очень грубы уже при $p > 0{,}999$. В некоторых случаях удается оценить относительную ошибку, которая получается при замене истинного распределения суммы большого числа случайных величин нормальным распределением, и построить более точные интервалы.

Если же $D\xi^s$ неизвестна, то обычно ее заменяют оценкой

$$D^*\xi^s = m_{2s}^* - (m_s^*)^2,$$

еще более снижая точность доверительного интервала (44).

Пример 1. Построить доверительный интервал для математического ожидания нормальной случайной величины, если по выборке объема $n = 21$ построены оценки $\bar{x} = 5$ и $s^2 = 1{,}21$. Уровень доверия $\varkappa = 0{,}999$.

◀ Точный доверительный интервал имеет вид

$$\left(\bar{x} - t_\varkappa\frac{s}{\sqrt{n}},\ \bar{x} + t_\varkappa\frac{s}{\sqrt{n}}\right),$$

где $t_\varkappa$ — решение уравнения

$$2S_{n-1}(t) - 1 = 0{,}999.$$

В нашем случае $n - 1 = 20$, $\varkappa = 0{,}999$. По таблице распределения Стьюдента (см. ниже) определяем $t_\varkappa = 3{,}85$. Искомый доверительный интервал имеет границы

$$4{,}053 \leqslant M\xi \leqslant 5{,}947.$$

Распределение Стьюдента

$n-1$	$\varkappa$								
	0,90	0,95	0,975	0,980	0,990	0,995	0,997	0,998	0,999
1	6,314	12,706	25,452	31,821	63,657	127,3	212,2	318,3	636,6
2	2,920	4,303	6,205	6,965	9,925	14,089	18,216	22,327	31,598
3	2,353	3,182	4,177	4,541	5,841	7,453	8,891	10,214	12,941
4	2,132	2,776	3,495	3,747	4,604	5,597	6,435	7,173	8,610
5	2,015	2,571	3,163	3,365	4,032	4,773	5,376	5,893	6,859
6	1,943	2,447	2,969	3,143	3,707	4,317	4,800	5,208	5,959
7	1,895	2,365	2,841	2,998	3,499	4,029	4,442	4,785	5,405
8	1,860	2,306	2,752	2,896	3,355	3,833	4,199	4,501	5,041
9	1,833	2,262	2,685	2,821	3,250	3,690	4,024	4,297	4,781
10	1,812	2,228	2,634	2,764	3,169	3,581	3,892	4,144	4,587
12	1,782	2,179	2,560	2,681	3,055	3,428	3,706	3,930	4,318
14	1,761	2,145	2,510	2,624	2,977	3,326	3,583	3,787	4,140
16	1,746	2,120	2,473	2,583	2,921	3,252	3,494	3,686	4,015
18	1,734	2,101	2,445	2,552	2,878	3,193	3,428	3,610	3,922
20	1,725	2,086	2,423	2,528	2,845	3,153	3,376	3,552	3,849
22	1,717	2,074	2,405	2,508	2,819	3,119	3,335	3,505	3,792
24	1,711	2,064	2,391	2,492	2,797	3,092	3,302	3,467	3,745
26	1,706	2,056	2,379	2,479	2,779	3,067	3,274	3,435	3,707
28	1,701	2,048	2,369	2,467	2,763	3,047	3,250	3,408	3,674
30	1,697	2,042	2,360	2,457	2,750	3,030	3,230	3,386	3,646
∞	1,645	1,960	2,241	2,326	2,576	2,807	2,968	3,090	3,291

Приближенный доверительный интервал с тем же уровнем доверия выглядит так:

$$\left(\bar{x} - x_\varkappa \frac{s}{\sqrt{n}},\ \bar{x} + x_\varkappa \frac{s}{\sqrt{n}}\right),$$

где $x_\varkappa$ — решение уравнения

$$2F(x) - 1 = 0{,}999.$$

По таблице нормального распределения (см. с. 69) определяем $x_\varkappa = 3{,}3$ и искомый доверительный интервал

$$4{,}19 \leqslant M\xi \leqslant 5{,}81.\ \blacktriangleright$$

Заметим, что точный доверительный интервал оказался более осторожным, чем приближенный. Этого и следовало ожидать.

С увеличением же объема выборки приближенный доверительный интервал становится более близок к точному. Действительно, пусть данные нашей задачи получены по выборке объема $n = 50$. Тогда $t_\varkappa = 3{,}5$ и $4{,}45 \leqslant M\xi \leqslant 5 < 55$. В то же время, $x_\varkappa = 3{,}3$ и $4{,}48 \leqslant M\xi \leqslant 5{,}52$ и разница между точным и приближенным доверительными интервалами уменьшилась.

3.4. Эффективность оценивания. Неравенство Рао—Крамера

При оценивании естественно считать дисперсию оценок мерилом того, насколько хороша или плоха принятая процедура. Если ξ — случайная величина с законом распределения $\Phi_\xi(x;\beta)$ и для оценивания параметра β мы имеем две различные процедуры $\beta \approx \beta_1^*(X_1\ \ldots\ X_n)$ и $\beta \approx \beta_2^*(X_1, X_2\ \ldots\ X_n)$ с дисперсиями $\sigma^2_{\beta_1} = D\beta_1^*$ и $\sigma^2_{\beta_2} = D\beta_2^*$ соответственно, то при $\sigma^2_{\beta_1} < \sigma^2_{\beta_2}$ оценка β_1^* считается лучше оценки β_2^*. Для больших выборок этот показатель не очень существен, ибо, как следует из полученных выше соотношений (30), (38) и (44), при $n \to \infty$ точность оценивания убывает как $1/\sqrt{n}$ и стремится к нулю независимо от σ_{β_i}.

Однако для малых выборок вопрос о выборе наилучшей в указанном смысле оценки приобретает важное значение. Мы получим ответ на него при дополнительном предположении о несмещенности рассматриваемых оценок.

Пусть ξ — случайная величина, плотность распределения которой $f_\xi(x;\,\beta)$ зависит от одного параметра β. Пусть, далее, β оценивается по выборке объема n несмещенным образом

$$\beta \approx \beta^*(X_1, X_2, \ldots, X_n), \qquad M\beta^* = \beta \tag{45}$$

И (см. п. 1.2.2.) функция правдоподобия выборки $L = L(X_1\ \ldots\ X_n;\,\beta)$ определяется соотношением

$$L(X_1, \ldots, X_n;\,\beta) = \prod_{i=1}^{n} f_\xi(X_i;\,\beta).$$

Тогда имеет место утверждение

Теорема (Неравенство Рао—Крамера)

$$\sigma^2_{\beta^*} \geqslant \frac{1}{M\left(\dfrac{1}{L}\,\dfrac{\partial L}{\partial \beta}\right)^2}. \tag{46}$$

◀ Для сокращения записи положим

$$\overline{X} = (X_1, \ldots, X_n); \qquad \underbrace{\int\!\ldots\!\int}_{n} \varphi(X_1, \ldots, X_n)\, dX_1 \ldots dX_n = \int\limits_{\mathbb{R}^n} \varphi(\overline{X})\, d\overline{X}.$$

И с учетом этих обозначений получим

$$\int\limits_{\mathbb{R}^n} L(\overline{X};\,\beta)\, d\overline{X} = 1 \quad \text{и} \quad \int\limits_{\mathbb{R}^n} \beta^*(\overline{X}) L(\overline{X};\,\beta)\, d\overline{X} = \beta.$$

Первое из соотношений следует из того, что функция правдоподобия есть плотность распределения вектора $\overline{X}$ выборочных значений $(X_1, \ldots, X_n)$, второе — из несмещенности оценки β^*. Дифференцируя эти соотношения по β и вычитая результаты дифференцирования, получаем

$$\int\limits_{\mathbb{R}^n} \left(\beta^*(\overline{X}) - \beta\right)\left(\frac{1}{L}\,\frac{\partial L}{\partial \beta}\right) L\, d\overline{X} = 1 \tag{47}$$

или

$$M(\beta^* - \beta)\left(\frac{1}{L}\,\frac{\partial L}{\partial \beta}\right) = 1.$$

Из неравенства Коши—Буняковского для математических ожиданий заключаем, что

$$M(\beta^* - \beta)^2 M\left(\frac{1}{L}\,\frac{\partial L}{\partial \beta}\right)^2 \geqslant 1,$$

чем доказательство и завершается. ▶

Неравенство (46) при помощи несложных выкладок может быть переписано в несколько более удобном для практического использования виде. А именно, поскольку $L(\overline{X};\beta)=\prod\limits_{j=1}^{n} f(X_j;\beta)$, то

$$\frac{1}{L}\,\frac{\partial L}{\partial\beta}=\sum_{j=1}^{n}\frac{1}{f_\xi(X_j;\beta)}\,\frac{\partial f_\xi}{\partial\beta}(X_j;\beta)=\sum_{j=1}^{n}\frac{\partial\ln f_\xi(X_j;\beta)}{\partial\beta}.$$

Функции $\dfrac{\partial\ln f_\xi(X_j;\beta)}{\partial\beta}$ независимы в силу независимости выборочных значений X_j. Поэтому $M\left(\dfrac{\partial}{\partial\beta}\ \ln f_\xi(X_i;\beta)\ \dfrac{\partial}{\partial\beta}\ \ln f_\xi(X_j;\beta)\right)=0$. С учетом этого замечания

$$M\left[\frac{1}{L}\,\frac{\partial L}{\partial\beta}\right]^2=M\left[\sum_{j=1}^{n}\left(\frac{\partial\ln f_\xi(Y_j;\beta)}{\partial\beta}\right)^2+2\sum_{i\neq j}\left(\frac{\partial\ln f_xi(X_j;\beta)}{\partial\beta}\;\frac{\partial\ln f_\xi(X_j;\beta)}{\partial\beta}\right)\right]=$$

$$=\sum_{j=1}^{n}M\left(\frac{\partial\ln f_xi(X_j;\beta)}{\partial\beta}\right)^2=n\cdot M\left(\frac{\partial\ln f_xi(\xi;\beta)}{\partial\beta}\right)^2,$$

и неравенство (46) принимает вид

$$\sigma^2_{\beta^*}\geqslant\frac{1}{nM\left(\dfrac{\partial\ln f_xi(\xi;\beta)}{\partial\beta}\right)^2}. \tag{48}$$

Пример 1. В важном для приложений случае оценивания математического ожидания нормальной случайной величины ξ получаем

$$f_\xi(\xi;m)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(\xi-m)^2}{2\sigma^2}\right\}$$

и (48) записывается в виде

$$\sigma^2_{m^*}\geqslant\frac{\sigma^2}{n},$$

где $\sigma^2=D\xi$. Заметим, что для эмпирической оценки математического ожидания $m\approx\bar{x}=n^{-1}\sum X_i$ выполняется

$$\sigma^2_{\bar{x}}=\frac{1}{n^2}D\left(\sum X_i\right)=\frac{1}{n^2}\,nDX_i=\frac{\sigma^2}{n}$$

и, следовательно, с рассматриваемой точки зрения эта оценка наилучшая.

Такие оценки в статистике называются *эффективными*.

Аналогичная теорема может быть доказана и для случая совместного оценивания нескольких неизвестных параметров распределения. Здесь мы ограничимся только формулировкой указанной теоремы.

Пусть, как и выше, ξ — непрерывная случайная величина с плотностью $f_\xi(X;\beta_1,\ldots,\beta_k)$, зависящей от вектора параметров $\bar{\beta}=(\beta_1,\ldots,\beta_k)$, несмещенная оценка которого дается соотношениями

$$\beta_j\approx\beta_j^*=\beta_j^*(\bar{x}),\qquad M\beta_j^*=\beta,\qquad j=1,\ldots,k,$$

и $L=L(\overline{X};\bar{\beta})$ — функция правдоподобия выборки $\overline{X}$.

Пусть далее

$$\mathbf{K}[\bar{\beta}^*]=M(\beta^*-\beta)(\beta^*-\beta)^T$$

и

$$\mathbf{I} = \mathbf{K}\left[\frac{\partial \ln L}{\partial \beta}\right] = M\left[\frac{\partial \ln L}{\partial \beta}\left(\frac{\partial \ln L}{\partial \beta}\right)^T\right]$$

— корреляционные матрицы вектора оценок β^* и градиента логарифма функции правдоподобия, соответственно:

$$\mathbf{K}[\beta^*] = \left(K_{ij}^{\beta}\right)_{i,j=1}^{k} = M(\beta_i - \beta_i^*)(\beta_j - \beta_j^*),$$

$$\mathbf{I} = \mathbf{K}\left[\frac{\partial \ln L}{\partial \beta}\right] = \left(K_{ij}^{L}\right)_{i,j=1}^{k} = \left(M\left(\frac{\partial \ln L(\overline{X}, \beta)}{\partial \beta_i} \cdot \frac{\partial \ln L(\overline{X}, \beta)}{\partial \beta_j}\right)\right).$$

Тогда, в предположении, что матрица $\mathbf{I}$ обратима, имеет место неравенство

$$\mathbf{K}[\beta^*] \geqslant \mathbf{I}^{-1}, \tag{49}$$

которое следует понимать как неотрицательную определенность матрицы $\mathbf{S} = \mathbf{K} - \mathbf{I}^{-1}$, т. е. $\forall\, \mathbf{x} = (x_i)_{i=1}^{k}$, $\mathbf{x} \neq 0$, выполняется неравенство

$$\sum_{i,j} S_{ij} x_i x_j \geqslant 0. \tag{50}$$

Глава XLIV

СТАТИСТИЧЕСКАЯ ПРОВЕРКА ГИПОТЕЗ

В этой главе мы познакомимся с элементами статистической проверки гипотез, т. е. с процедурой построения некоторых правил, позволяющих по результатам эксперимента высказывать суждение о природе явлений, обусловливающих изучаемый эксперимент.

§ 1. Общие сведения

Пусть высказано некоторое предположение (гипотеза) H о природе явления, которое мы наблюдаем в эксперименте. Чтобы проверить справедливость H, следует либо изучить всю совокупность следствий, которые должны иметь место, если гипотеза H верна, либо указать некоторое событие S, невозможное при верной гипотезе H. В первом случае (если все эти следствия наблюдаются) гипотезу H можно считать верной, во втором (если событие наблюдается в эксперименте) гипотеза H неверна. Это самая простая ситуация и рассуждения здесь проводятся по следующей схеме: гипотеза H эквивалентна полному набору следствий, поэтому

$$\textit{все следствия имеют место} \Longrightarrow H \textit{ верна} \tag{1}$$

или: если H верна, то событие S невозможно; поэтому

$$S \textit{ имеет место} \Longrightarrow H \textit{ неверна.} \tag{2}$$

Проверка гипотез подобного рода не представляет для исследователя никаких затруднений, но на практике такая ситуация встречается редко.

Первая сложность, с которой приходится сталкиваться, состоит в том, что в большинстве действительно интересных для исследователя случаев проверить все следствия из гипотезы H не представляется возможным и приходится ограничиваться проверкой только части следствий. Но заключение о справедливости гипотезы, сделанное по неполному набору следствий из нее по схеме (1), уже не является достоверным. В то же время заключение о несправедливости гипотезы H, сделанное по схеме (2), все еще достоверно. Поэтому, находясь в указанной выше ситуации, можно только отвергнуть гипотезу по схеме (2), наблюдая событие S, невозможное в случае ее справедливости, но нельзя гипотезу подтвердить. Можно лишь высказать суждение о правдоподобии гипотезы. Причем степень нашей уверенности в справедливости высказанного суждения будет тем выше, чем больший набор следствий из гипотезы H удалось проверить.

Классическим примером подобных гипотез являются естественно-научные гипотезы, которые всегда подвергаются указанной выше проверке и либо становятся теориями (если нет противоречащих рассматриваемой гипотезе явлений), либо отвергаются (если таковые есть).

Хотелось бы подчеркнуть вот какое обстоятельство: до тех пор, пока не обнаружено явление, противоречащее проверяемой гипотезе, ее отвергнуть нельзя. Поэтому если мы располагаем двумя гипотезами, одинаково подтверждающимися в эксперименте, то у нас нет никаких оснований для предпочтения одной из гипотез другой, и в то же время мы не в состоянии (поскольку располагаем неполным набором следствий) утверждать, что обе гипотезы справедливы!

Дальнейшее усложнение связано с тем, что в основе изучаемых нами явлений могут лежать случайные воздействия, и мало того, что мы располагаем неполным набором следствий и не можем достоверно подтвердить гипотезу, мы теперь не можем ее и отвергнуть, ибо довольно трудно указать событие S, невозможное в случае справедливости гипотезы H. Можно лишь указать событие S такое, которое происходит редко, если гипотеза H верна. Схема (2) в этом случае уже неприменима, ибо из того, что гипотеза H верна, мы можем сделать заключение лишь о редкости события S, но не о его возможности. Поэтому наблюдение события S в эксперименте гипотезу H не опровергает.

Рассмотрим пример. Пусть производится контроль качества партии продукции, причем характер продукции таков, что сплошной контроль невозможен или нерационален Для решения вопроса о качестве всей партии, содержащей N изделий, отберем $n < N$ изделий и тщательно исследуем их качество. Пусть в выборке оказалось n_1 дефектных изделий. Какое заключение можно сделать по этой выборке о качестве всей исследуемой партии? Видимо, единственное, что можно сказать наверняка, так это то, что исследуемая партия содержит не менее, чем n_1, и не более, чем $N - n + n_1$, дефектных изделий. Результаты произведенного исследования выборки, однако, позволяют надеяться, что доля дефектных изделий в партии близка к n_1/n. Утверждать же это наверняка нельзя, ибо совершенно ясно, что и любое другое допустимое (не меньшее n_1 и не большее $N - n + n_1$) количество дефектных изделий в партии может привести к полученной нами выборке. Пусть гипотеза H_0 состоит в том, что исследуемая партия содержит долю $q_0 = n_1/n$ дефектных изделий. Для проверки этой гипотезы рассмотрим еще одну выборку из совокупности в N изделий. Пусть доля дефектных изделий в этой выборке оказалась равной $q_1 \neq q_0$. Если разница между q_0 и q_1 не очень велика, то отсюда еще не следует, что проверяемая гипотеза верна, хотя можно ожидать, что в большинстве случаев так оно и будет. Точно также, значительное различие q_0 и q_1 не обусловливает неверности гипотезы H_0, но приводит нас к мысли, что гипотеза H_0 все же малоправдоподобна. Это связано с тем, что при верной гипотезе H_0 мы должны чаще получать выборки, доля дефектных изделий в которых близка к q_0, нежели выборки, доля дефектных изделий в которых значительно отличается от q_0.

Возвращаясь к обсуждению общей ситуации, несколько видоизменим правила (1) и (2) принятия решений, предварительно формализовав рассматриваемые понятия.

Пусть в эксперименте наблюдается случайная величина ξ (или несколько случайных величин $\xi_1, \ldots, \xi_n$).

Любой непротиворечивый набор суждений о законе распределения случайной величины ξ (или совокупности $\xi_1, \ldots, \xi_n$) будем называть гипотезой. Гипотезу будем называть простой, если она однозначно указывает закон распределения случайной величины ξ (или совокупности $\xi_1, \ldots, \xi_n$). В противном случае гипотеза называется сложной.

Пример 1. Пусть случайная величина ξ — количество дефектных изделий в партии. Гипотеза H_0 состоит в том, что доля дефектных изделий в партии равна q_0. Это простая гипотеза. Примером сложной гипотезы в данной ситуации может служить гипотеза о том, что доля брака в партии не превышает q_0.

Пример 2. По выборке $X_1, \ldots, X_n$ получена оценка неизвестного математического ожидания случайной величины ξ. Гипотеза о равенстве $M\xi$ некоторому числу α является простой.

Пример 3. Пусть в эксперименте рассматривается пара независимых случайных величин. Гипотеза о равенстве их математических ожиданий является сложной.

Пример 4. Пусть закон распределения случайной величины ξ известен, но неизвестны значения параметров, его определяющих, $F_\xi(x) = F_\xi(x;\ \beta_1, \ldots, \beta_k)$. Тогда гипотеза H_0 о том, что параметры принимают известные значения

$$\beta_1 = b_1; \quad \beta_2 = b_2; \quad \ldots; \quad \beta_k = b_k,$$

является простой. Гипотеза же, указывающая только возможную область значений параметров

$$b_1' \leqslant \beta_1 \leqslant b_2; \quad \ldots; \quad b_k' \leqslant \beta_k \leqslant b_k,$$

будет сложной.

Критерием проверки гипотезы будем называть любое правило, позволяющее по выборке делать заключение о справедливости или несправедливости проверяемой гипотезы.

Как уже было отмечено выше, мы не можем построить логически безупречного критерия в случае гипотезы, связанной с законом распределения случайной величины. Поступать в этом случае будем следующим образом: пусть $\mathcal{M}$ — множество событий наблюдаемого эксперимента. Выделим в $\mathcal{M}$ множество S событий, происходящих редко в случае справедливости проверяемой гипотезы H. Пусть ω — результат эксперимента. Тогда

если $\omega \in S$, *считаем гипотезу малоправдоподобной*,

если $\omega \notin S$, *считаем гипотезу правдоподобной.*

Множество S называется *критическим множеством* критерия. Здесь возможны четыре случая.

I. *Гипотеза H верна и признана согласно критерию правдоподобной.*

II. *Гипотеза H неверна и признана согласно критерию неправдоподобной.*

III. *Гипотеза H верна, но согласно критерию признана неправдоподобной.*

IV. *Гипотеза H неверна, но согласно критерию признана правдоподобной.*

Случаи III и IV описывают ошибки, возможные при проверке гипотезы статистическими критериями. Они носят название соответственно *ошибок* 1 *и* 2*-го рода.*

Хотелось бы, чтобы применяемые нами критерии как можно чаще приводили к случаям I или II и как можно реже к ошибкам (случаи III и IV). Поэтому критическое множество S обычно выбирают так, чтобы при правильной гипотезе H вероятность получения в эксперименте исхода $\omega \in S$ была как можно меньше. Эта вероятность (вероятность ошибки 1-го рода) носит название *уровня значимости* критерия. Как следует из вышеизложенного, мы не можем указать множество S, соответствующее нулевому уровню значимости. Поэтому будем довольствоваться критическими множествами, соответствующими хоть и не нулевому, но довольно близкому к нулю уровню значимости. Обычно в качестве уровня значимости берут значения 0,05; 0,01; 0,001, хотя в зависимости от конкретной ситуации могут употребляться и другие близкие к нулю вероятности.

Для того чтобы свести к минимуму ошибки 2-го рода, следует, наряду с исследуемой гипотезой H, рассмотреть конкурирующие с ней гипотезы. Действительно, пусть верна какая-либо из альтернативных простых гипотез H_1. Тогда неверная гипотеза H

будет признана верной в том случае, когда множество событий, имеющих место в случае справедливости гипотезы H_1, пересекается с множеством событий, частых в случае справедливости проверяемой гипотезы H.

Вероятность принять гипотезу H в случае, когда верна гипотеза H_1, называется оперативной характеристикой критерия относительно гипотезы H_1.

Вероятность отвергнуть гипотезу H в случае, когда верна гипотеза H_1, называется мощностью критерия относительно гипотезы H_1.

Таким образом, выбор критической области S диктуется минимизацией вероятностей ошибок первого и второго рода. Если удается построить критическую область так, что мощность критерия принимает наибольшее значение для данной простой альтернативной гипотезы H_1, то соответствующий *критерий называется наиболее мощным при данном уровне значимости.*

Равномерно наиболее мощным критерием называется критерий, наиболее мощный относительно всех допустимых альтернативных гипотез при данном уровне значимости.

Рассмотрим пример, иллюстрирующий введенные выше понятия.

Пусть ξ — случайная величина, описывающая число появлений герба в n последовательных независимых испытаниях, вероятность появления герба в каждом из которых неизменна. Гипотеза, которую мы хотим проверить, состоит в том, что вероятность появления герба в отдельном испытании равна 0,5. Альтернативные гипотезы H_p — вероятность выпадения герба в отдельном испытании равна $p \neq 0{,}5$. Легко видеть, что как проверяемая, так и альтернативные гипотезы являются простыми. Для проверки гипотезы $H_{0,5}$ проведено n экспериментов и отмечено, что герб появился n_1 раз. Множество $\mathcal{M}$ исходов эксперимента состоит из всех возможных наборов $\{n_1,\ n_1 \leqslant n\}$, описывающих число появления герба n_1, $n_1 = 0, 1, 2, 3, \ldots, n$. Критическое множество S, определяющее критерий, будет подмножеством множества исходов $\mathcal{M}$. Зададим уровень значимости α и определим S так, что если гипотеза $H_{0,5}$ верна, то

$$P_{0,5}\{\xi \in S\} \leqslant \alpha, \tag{3}$$

или

$$\sum C_n^k (0{,}5)^n \leqslant \alpha.$$

Суммирование ведется здесь по всем k таким, что значение $\xi = k$ принадлежит критическому множеству S. Легко видеть, что при заданном уровне значимости можно указать довольно много различных множеств S, удовлетворяющих соотношению (3). Каждое из этих множеств будет определять критерий для проверки нашей гипотезы. Возьмем, к примеру, в качестве S множество $S_1 = \{n_1\colon\ n_1 < k_0\}$, где k_0 однозначно определяется из соотношения

$$\sum_{k=0}^{k_0} C_n^k (0{,}5)^n \leqslant \alpha \tag{4}$$

как наибольшее из возможных k_0. Критерий K_1, построенный на основании S_1, будет признавать гипотезу $H_{0,5}$ неверной, если $\xi \leqslant k_0$, и верной в противном случае. Ясно, что это должен быть не очень хороший критерий. Критерий K_2 построим на основании множества $S_2 = \{n_1\colon\ n_1 < k_1 \ \text{ или } \ n_1 \geqslant n - k_1\}$. Этот критерий будет признавать гипотезу $H_{0,5}$ верной, если $k_1 \leqslant \xi \leqslant n - k_1$, и неверной в противном случае. Он уже кажется лучшим, чем K_1.

Действительно, рассмотрим мощности критериев K_1 и K_2 относительно какой-либо из альтернативных гипотез H_p, $p \neq 0{,}5$. Пусть верна гипотеза H_p. Мощность критерия K_1 относительно гипотезы H_p

$$P_p\{\xi \in S_1\} = \sum_{k=0}^{k_0} C_n^k p^k (1-p)^{n-k}. \tag{5}$$

Для критерия K_2

$$P_p\{\xi \in S_2\} = \sum_{k \in S_2} C_n^k p^k (1-p)^{n-k}. \tag{6}$$

В этом равенстве k_1 определяется из соотношения

$$\sum_{k=n-k_1}^{k_1} C_n^k (0{,}5)^n \leqslant \alpha$$

как наибольшее из возможных k_1.

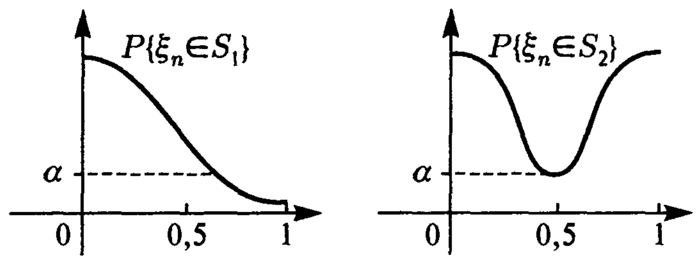


Рис. 1

Зависимость мощности (5) и (6) критериев K_1 и K_2 соответственно от альтернативной гипотезы H_p схематично представлена на рис. 1. Отсюда легко усмотреть, что критерий K_1 будет неплох, если альтернативная гипотеза H_p такова, что $p < 0{,}5$. Если же $p > 0{,}5$, то согласно критерию K_1 мы будем почти всегда проверяемую гипотезу $H_{0,5}$ считать верной. Впрочем, это было очевидно с самого начала: выбранная нами критическая область S_1 совершенно нечувствительна к отклонениям числа появившихся в эксперименте гербов в сторону чисел, бóльших $0{,}5n$. Критерий же K_2 строился на основании отклонений как в ту, так и в другую сторону от наиболее вероятного при верной гипотезе $H_{0,5}$ числа $0{,}5n$ и потому оказался чувствительным ко всем альтернативным гипотезам. Однако и он не лишен недостатков. Его чувствительность падает с приближением p к $0{,}5$ (см. рис. 1). Но (ясно из постановки задачи) это вполне естественно, и ничего лучшего в данной ситуации предложить нельзя.

Легко проверить, что критерий K_2 будет более мощным, чем критерий K_1, для любой альтернативной гипотезы H_p такой, что $p > 0{,}5$.

В дальнейшем мы не будем останавливаться на исследовании мощности того или иного критерия, ибо сама постановка задачи обычно определяет, какая из возможных при данном уровне значимости критических областей S будет наилучшей.

В заключение отметим важное обстоятельство: проверяемая нами при помощи статистических критериев гипотеза ***не подлежит вероятностной оценке***. Поскольку она описывает некоторые объективные стороны исследуемого процесса, то может быть либо верной, либо неверной, и высказывание типа: ***«Гипотеза верна с вероятностью такой-то»*** бессмысленно. В связи с этим полезно иметь в виду, что уровень значимости критерия, мощность критерия, оперативная характеристика критерия не являются условными вероятностями описанных выше событий ***«при условии, что верна гипотеза H_0». Эти характеристики критерия описывают вероятность встретить в эксперименте ту или иную выборку в предположении, что истинная природа явлений, наблюдаемых нами, описывается гипотезой H_0 или какой-нибудь альтернативной гипотезой H.*** Мы не можем говорить об условной вероятности $P\{\xi \in S \mid H_0\}$, так как не в состоянии осмысленно приписать какую-либо вероятность гипотезе H_0.

§ 2. Параметрические гипотезы. Лемма Неймана—Пирсона

Пусть случайная величина ξ имеет распределение $F_\xi(x;\ \beta_1, \beta_2, \ldots, \beta_k)$, известное с точностью до вектора параметров $\beta = \{\beta_1, \ldots, \beta_k\}$. Назовем гипотезу H_0 *параметрической*, если она состоит в предположении, что вектор β принимает значения из некоторого множества W,

$$H_0\colon\ \beta \in W \subseteq \mathbb{R}^k.$$

При построении критериев проверки параметрических гипотез важную роль играет *принцип отношения правдоподобия*, позволяющий в подавляющем большинстве важных для приложений ситуаций строить критические области критериев.

Для упрощения дальнейшего изложения будем считать ξ непрерывной с плотностью $f_\xi(x;\ \beta)$.

Напомним, что процедура проверки подобной гипотезы против альтернативы $H\colon\ \beta \notin W$ требует указания критического множества S такого, что если $\bar{x} \in S$ — гипотеза принимается, в противном же случае — отвергается.

Положим

$$\mathcal{L}(W) = \sup_{\beta \in W} L(\overline{X};\ \beta), \qquad \mathcal{L} = \sup_{\beta} L(\overline{X};\ \beta),$$

где $L = \prod\limits_{i=1}^{n} f_\xi(X_i;\ \beta)$ — функция правдоподобия выборки $\overline{X}$, и рассмотрим отношение

$$\lambda_{H_0} = \frac{\mathcal{L}(W)}{\mathcal{L}}, \tag{7}$$

которое называется *отношением правдоподобия*. Ясно, что λ_{H_0} находится в пределах от 0 до 1. Далее заметим, что при фиксированной выборке $\overline{X}$ предпочтительными являются те значения параметров β, для которых $L(\overline{X};\ \beta)$ больше; поэтому чем ближе величина λ_{H_0} к единице, тем «более правдоподобно», что гипотеза H_0 верна, если же значения λ_{H_0} — маленькие, то скорее всего гипотеза H_0 неверна, так как более «весомой» представляется одна из альтернативных гипотез, значительно увеличивающая знаменатель отношения правдоподобия в сравнении с числителем.

Приведенные выше интуитивные соображения удается аккуратно формализовать в виде следующего утверждения.

Теорема (принцип отношения правдоподобия Неймана—Пирсона). *Для любого* $0 < \alpha < 1$ *критическое множество S критерия проверки параметрической гипотезы* $H_0\colon\ \beta \in W$ *с уровнем значимости α дается соотношением*

$$S = \left\{\overline{X} \in \mathbb{R}^n : \lambda_{H_0} < C_\alpha\right\}, \tag{8}$$

где C_α — постоянная, определяемая условием

$$P_{H_0}\left\{\overline{X} \in S\right\} \leqslant \alpha. \tag{9}$$

Можно доказать, что так построенный критерий обладает определенными оптимальными свойствами, в частности, если гипотеза H_0 — простая и строится критерий проверки против гипотезы H — также простой, то критерий отношения правдоподобия оказывается равномерно наиболее мощным критерием.

В качестве примера использования сформулированного выше принципа рассмотрим процедуры построения критического множества S для проверки различных, часто встречающихся гипотез.

2.1. Проверка гипотезы о равенстве математического ожидания нормальной случайной величины числу a

Постановка задачи. В эксперименте наблюдается случайная величина ξ, распределенная по нормальному закону с неизвестными параметрами $M\xi$ и $D\xi$. Получена выборка из распределения случайной величины ξ. Требуется выяснить, справедлива ли гипотеза о равенстве $M\xi = a$.

Вектор параметров β в этой задаче двумерен

$$\beta = (m, \sigma),$$

нулевая гипотеза H_0 состоит в том, что $\beta \in W$, где $W = \{(m, \sigma)\colon\ m = a,\ \sigma > 0\}$ — полупрямая на полуплоскости $\mathbb{R}_+^2 = \{(m, \sigma)\colon\ \sigma > 0\}$. Функция правдоподобия выборки $\overline{X} = (X_1 \ldots X_n)$ будет иметь вид

$$L(\overline{X}, \beta) = L(\overline{X}, m, \sigma) = \frac{1}{(2\pi)^{n/2}} \cdot \frac{1}{\sigma^n} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (X_i - m)^2\right]. \tag{10}$$

Для $\mathcal{L}(W)$ и $\mathcal{L}$ получим соответственно

$$\begin{aligned} \mathcal{L}(W) &= \sup_{(m,\,\sigma)\in W} L(\overline{X}, m, \sigma), \\ \mathcal{L} &= \sup_{(m,\,\sigma)\in \mathbb{R}_+^2} L(\overline{X}, m, \sigma). \end{aligned} \tag{11}$$

Несложные выкладки по нахождению экстремумов (11) приводят к формулам

$$\operatorname{argextr} \mathcal{L}(W) = \frac{1}{n} \sum_{i=1}^{n} (X_i - a)^2,$$

$$\operatorname{argextr} \mathcal{L} = \left\{ m_{\text{экстр}} = \frac{1}{n} \sum X_i = \bar{x},\ \sigma^2_{\text{экстр}} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{x})^2 \right\}.$$

Отношение правдоподобия (7) принимает вид

$$\lambda_{H_0} = \left(\frac{\sum (X_i - \bar{x})^2}{\sum (X_i - a)^2} \right)^{n/2}.$$

Заметим, что так как

$$(X_i - \bar{x})^2 = (X_i - a + a - \bar{x})^2 = (X_i - a)^2 + (a - \bar{x})^2 + 2(X_i - \bar{a})(a - \bar{x}),$$

то

$$\sum_{i=1}^{n} (X_i - \bar{x})^2 = \sum (X_i - a)^2 + n(\bar{x} - a)^2 - 2n(\bar{x} - a)^2 = \sum_{i=1}^{n} (X_i - a)^2 - n(\bar{x} - a)^2.$$

Отсюда, разделив последнее соотношение на его левую часть, получим

$$1 = \left(\frac{1}{\lambda_{H_0}}\right)^{2/n} - \frac{t_{n-1}^2}{n-1},$$

где $t_{n-1} = \frac{\bar{x}-a}{s}\sqrt{n}$ — случайная величина, имеющая распределение Стьюдента с $n-1$ степенью свободы.

Поэтому критическая область (8) в рассматриваемом случае имеет вид

$$S = \left\{\bar{X} \in \mathbb{R}^n\colon\ \lambda_{H_0} < C_\alpha\right\} = \left\{t_{n-1} \in \mathbb{R}\colon\ |t_{n-1}| > t_\alpha\right\}, \tag{12}$$

где значение t_α дается соотношением

$$P\{|t_{n-1}| > t_\alpha\} \leqslant \alpha.$$

Мы пришли к хорошо известному критерию Стьюдента проверки рассматриваемой гипотезы, который, впрочем, легко мог бы быть получен прямыми рассуждениями, не связанными с использованием отношения правдоподобия.

Действительно, по выборке, полученной в результате эксперимента, мы можем построить точечную оценку $\bar{x}$ неизвестного параметра $M\xi$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{13}$$

Результаты предыдущей главы позволяют утверждать, что оценка (13) похожа на истинное значение $M\xi$, но не обязательно с ним совпадает. Поэтому из того, что $\bar{x} \neq a$, мы еще не можем сделать заключения, что $M\xi \neq a$. Если вспомнить аналогию с пещерным человеком Платона, то $\bar{x}$ — это наблюдаемая нами «тень» $M\xi$ и мы должны, сравнивая «тень» и известное нам число a, высказать суждение, верна гипотеза $M\xi = a$ или неверна.

Если принять, что гипотеза $M\xi = a$ верна, то величина

$$t = \frac{\bar{x}-a}{s}\sqrt{n} \tag{14}$$

оказывается распределенной по закону Стьюдента с $n-1$ степенью свободы.

Зададим некоторый уровень значимости a и определим критическое множество S как множество таких отклонений $\bar{x}$ от a, вероятность встретить которые в эксперименте (в случае справедливости гипотезы $M\xi = a$) не превышает α. Здесь заложена следующая идея: если гипотеза верна, то отклонения чаще будут малыми, а реже большими. *Поэтому малыми считаем те отклонения, которые встречаются часто!*

$$S = \left\{\bar{x} - a\colon\ P\{|\bar{x} - a| > \Delta_\alpha\} \leqslant \alpha\right\}.$$

Поскольку

$$P\{|\bar{x} - a| < \Delta_\alpha\} = P\left\{|t_{n-1}| < \frac{\Delta_\alpha\sqrt{n}}{s}\right\} = 2S_{n-1}\left(\frac{\Delta_\alpha\sqrt{n}}{s}\right) - 1,$$

где $S_{n-1}(\tau)$ — функция распределения Стьюдента с $n-1$ степенью свободы, то пограничная величина Δ_α может быть определена из соотношения

$$2S_{n-1}\left(\frac{\Delta_\alpha\sqrt{n}}{s}\right) - 1 \geqslant 1 - \alpha.$$

Для проверки гипотезы по конкретному набору $X_1, X_2, \ldots, X_n$, полученному в эксперименте, вычислим оценку $\bar{x}$ и найдем отклонение $\bar{x} - a$. Если оно попадает во множество S, то гипотеза о равенстве $M\xi = a$ считается несогласующейся с экспериментом и отвергается на уровне значимости a, в противном случае гипотеза принимается на уровне значимости a.

Пример. Станок настроен на выпуск деталей размером d. Размеры деталей, изготавливаемых на данном станке, не будут в точности равны d, а будут иметь размер

$$\bar{d} = d + \xi,$$

где ξ можно считать нормальной случайной величиной с математическим ожиданием 0 и некоторой дисперсией σ^2. Деталь считается бездефектной, если отклонение ее размера от заданного заключены в пределах

$$-h \leqslant \xi \leqslant h.$$

Таким образом, хорошо настроенный станок будет в среднем давать долю q бездефектных изделий, где

$$q = 2\Phi\left(\frac{h}{\sigma} - 1\right).$$

В процессе изготовления деталей станок может разладиться — центр настройки d может сместиться, при этом размер детали будет выражаться соотношением

$$\bar{d} = d + a + \xi = d + \xi_1.$$

Здесь a — смещение центра настройки станка. Отклонение ξ_1 размера детали от заданного будет в этом случае случайной величиной с нормальным законом распределения и с математическим ожиданием $M\xi_1 = M(\xi + a) = a$. Доля бездефектных изделий, изготовленных на станке, при этом уменьшится

$$q_1 = \Phi\left(\frac{h-a}{\sigma}\right) - \Phi\left(\frac{-h-a}{\sigma}\right),$$

т. е. увеличится доля брака (рис. 2). Важная задача — своевременно установить момент смещения центра настройки.

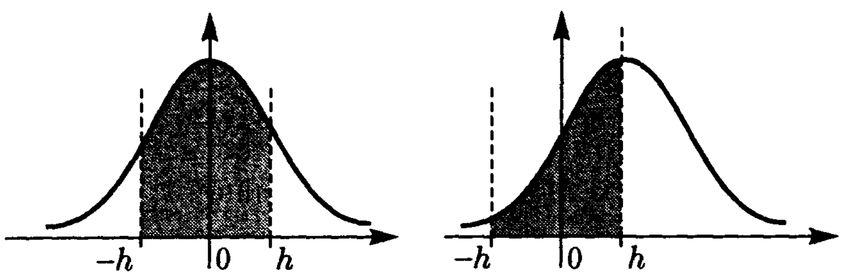


Рис. 2

Берется некоторое количество деталей (обычно 3–4) и находится средний размер, а затем отклонение этого среднего размера от предполагаемого d. Сравнив полученное отклонение с границами для отклонения, которые должны иметь место в случае, если смещение центра настройки не произошло, можно выяснить, справедлива ли гипотеза о смещении центра настройки станка.

2.2. Проверка гипотезы о равенстве дисперсии нормальной случайной величины ξ числу b

Постановка задачи. В эксперименте наблюдается случайная величина ξ, распределенная по нормальному закону с параметрами $M\xi = m$ и $D\xi = \sigma^2$, которые неизвестны. Получена выборка из распределения случайной величины ξ: $X_1, X_2, \ldots, X_n$. Требуется выяснить, равно ли число $D\xi$ некоторому наперед заданному числу b?

Аналогично тому, как это было сделано в предыдущем пункте, можно показать, что принцип отношения правдоподобия приводит к критическому множеству S, определяемому статистикой

$$\chi^2_{n-1} = \frac{s^2}{b}(n-1), \tag{15}$$

где s^2 — несмещенная оценка дисперсии. А именно, если гипотеза $\sigma^2 = b$ справед іва, то величина (15) имеет распределения χ^2 с $n - 1$ степенью свободы и для у значимости α критическое множество S может быть определено как множество таких значений s^2, вероятность встретить которые в эксперименте не больше α,

$$S = \left\{s^2 \colon P\left\{s^2 \in S\right\} \leqslant \alpha\right\}.$$

В силу несимметричности распределения χ^2 мы используем для построения области S несимметричные доверительные границы.

Получаем

$$P\left\{s^2 < \Delta_1\right\} = P\left\{\delta < \frac{\Delta_1}{b}(n-1)\right\} \quad \text{и} \quad P\left\{s^2 > \Delta_2\right\} = P\left\{\delta > \frac{\Delta_2}{b}(n-1)\right\}.$$

Здесь $P\{\delta < x\}$ — функция χ^2-распределения с $n-1$ степенью свободы.

Решая уравнения

$$G_{n-1}(t) = \frac{\alpha}{2} \quad \text{и} \quad 1 - G_{n-1}(t) = \frac{\alpha}{2},$$

находим числа t'_α и t''_α такие, что

$$P\left\{s^2 < t'_\alpha \frac{b}{n-1}\right\} = \frac{\alpha}{2} \quad \text{и} \quad P\left\{s^2 > t''_\alpha \frac{b}{n-1}\right\} = \frac{\alpha}{2}, \tag{16}$$

откуда критическое множество S имеет вид

$$\left(t'_\alpha \frac{b}{n-1},\ t''_\alpha \frac{b}{n-1}\right).$$

Для проверки гипотезы по конкретному набору $X_1, X_2, \ldots, X_n$ значений случайной величины ξ, полученному в эксперименте, вычислим оценку s^2 дисперсии. Если полученное число попадает в критическую область S, гипотеза отвергается на уровне значимости α, в противном случае гипотеза принимается.

Пример. Станок настроен на выпуск деталей некоторого наперед заданного размера d, причем точность работы станка описывается дисперсией $D\xi$ случайной величины ξ — отклонения размера $\bar{d}$ детали от заданного среднего d:

$$\bar{d} - d = \xi,$$

где ξ — нормально распределенная случайная величина с $M\xi = 0$ и $D\xi = \sigma^2$. Деталь считается бездефектной, если отклонение ξ удовлетворяет условию

$$-h \leqslant \xi \leqslant h.$$

Если смещение центра настройки не наблюдается, то в среднем мы будем получать долю q бездефектных изделий

$$q = 2\Phi\left(\frac{h}{\sigma}\right) - 1.$$

В процессе изготовления деталей точность может снизиться, т. е. может увеличиться дисперсия наблюдаемых отклонений ξ от заданного размера d. Если смещение центра при этом не произошло, то отклонение будет описываться случайной величиной ξ_1, дисперсия которой $D\xi_1 > D\xi$. Доля дефектных изделий при этом увеличится (рис. 3).

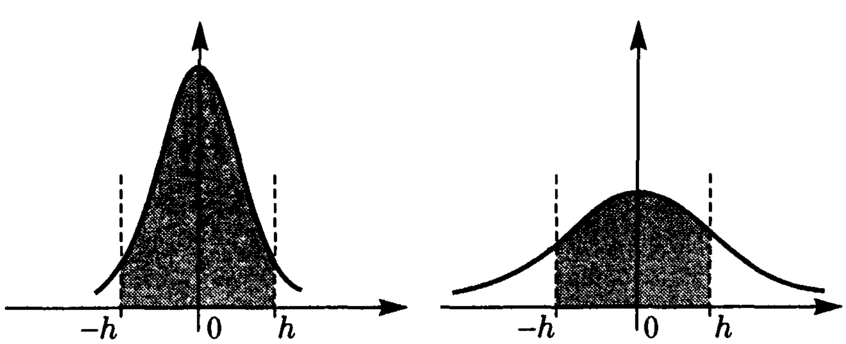


Рис. 3

Для того чтобы вовремя обнаружить разладки станка, возьмем некоторое количество деталей (3–4) и найдем оценку s^2. Сравнив полученную оценку с границами, которые должны иметь место, если разладки нет, мы сможем выяснить справедливость наших подозрений относительно снижения точности изготовления деталей на данном станке.

В заключение этого пункта отметим, что при проверке гипотезы о равенстве дисперсий $\sigma_1^2 = \sigma_2^2$ пары независимых нормальных случайных величин по независимым выборкам объемов n и m соответственно, принцип отношения правдоподобия в качестве статистики для построения критической области дает величину Z — отношение Фишера—Снедекера

$$Z = \frac{(n-1)s_1^2}{(m-1)s_2^2} = Z[n-1; m-1], \tag{17}$$

имеющую, в случае справедливости гипотезы о равенстве дисперсий, распределение Фишера с $(n-1, m-1)$ степенями свободы.

Зададим уровень значимости α и определим критическое множество S как множество таких значений s_1^2/s_2^2, вероятность встретить которые в эксперименте не больше α,

$$P\left\{\frac{s_1^2}{s_2^2} \in S\right\} \leqslant \alpha.$$

Напомним, что распределение Фишера асимметрично и при $n > 2$ унимодально.

Если наша гипотеза справедлива, то в большинстве случаев отношение s_1^2/s_2^2 должно быть близко к единице, т. е. отношение (17) должно быть близко к $(n-1)/(m-1)$. Учитывая асимметрию, выберем числа $\varkappa_1$ и $\varkappa_2$ так, чтобы

$$P\left\{\frac{s_1^2}{s_2^2} < \varkappa_1\right\} = \frac{\alpha}{2} \quad \text{и} \quad P\left\{\frac{s_1^2}{s_2^2} > \varkappa_2\right\} = \frac{\alpha}{2}.$$

Если $F_{n-1,m-1}(x)$ — функция распределения случайной величины Z, то числа $\varkappa_1$ и $\varkappa_2$ являются решениями уравнений

$$F_{n-1,m-1}\left(\frac{n-1}{m-1}\varkappa_1\right) = \frac{\alpha}{2}, \quad \text{и} \quad 1 - F_{n-1,m-1}\left(\frac{n-1}{m-1}\varkappa_2\right) = \frac{\alpha}{2}.$$

Для проверки гипотезы по выборкам, полученным в результате эксперимента, находим отношение s_1^2/s_2^2. Если оно попадает в критическую область, гипотеза о равенстве дисперсий считается несогласующейся с опытными данными на уровне значимости α, в противном случае гипотеза принимается.

2.3. Проверка гипотезы о равенстве средних нормальных случайных величин

Постановка задачи. Рассмотрим пару независимых нормально распределенных случайных величин ξ_1 и ξ_2 с параметрами $M\xi_1$, $D\xi_1$ и $M\xi_2$, $D\xi_2$ соответственно. В результате эксперимента получены две независимые выборки из распределения случайных величин ξ_1 и ξ_2

$$X_1, X_2, \ldots, X_n \quad \text{и} \quad Y_1, Y_2, \ldots, Y_n.$$

Требуется выяснить, совпадают ли математические ожидания $M\xi_1$ и $M\xi_2$.

1. По выборкам строим оценки s_1^2 и s_2^2 дисперсий $D\xi_1$ и $D\xi_2$ и, как это указано в предыдущем пункте, проверяем гипотезу о равенстве $D\xi_1 = D\xi_2$. Пусть гипотеза о равенстве дисперсий согласуется с экспериментальными данными: $\sigma_1^2 = \sigma_2^2$.

Лемма. *Случайная величина*

$$S^2_{n,m} = \frac{\sum\limits_{i=1}^{n}(X_i - \bar{x})^2 + \sum\limits_{i=1}^{m}(Y_i - \bar{y})^2}{n+m-2} \tag{18}$$

является несмещенной оценкой общей неизвестной дисперсии случайных величин ξ_1 и ξ_2.

◀ Заметим, что

$$S^2_{n,m} = \frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2}.$$

Поэтому

$$MS^2 = \frac{n-1}{n+m-2}\, Ms_1^2 + \frac{m-1}{n+m-2}\, Ms_2^2.$$

Но s_1^2 и s_2^2 — несмещенные оценки $D\xi_i = \sigma^2$, а потому $Ms_1^2 = \sigma^2$ и $Ms_2^2 = \sigma^2$, откуда

$$MS^2 = \sigma^2,$$

что и требовалось доказать. ▶

Теорема. *Случайная величина*

$$t_{n+m-2} = \frac{\bar{x} - \bar{y}}{S\sqrt{n+m}}\sqrt{nm}$$

имеет распределение Стьюдента с $n + m - 2$ степенями свободы, если только верна гипотеза о том, что $M\xi_1 = M\xi_2$.

Здесь $\bar{x} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$, $\bar{y} = \frac{1}{n}\sum\limits_{i=1}^{m} Y_i$, а S^2 — несмещенная оценка (18) общей дисперсии $D\xi_1 = D\xi_2 = \sigma^2$.

◀ Поскольку выборки $X_1, X_2, \ldots, X_n$ и $Y_1, Y_2, \ldots, Y_n$ из нормальных законов, то разность $\bar{x} - \bar{y}$ распределена по нормальному закону с $M(\bar{x} - \bar{y}) = 0$ (в случае верной гипотезы $M\xi_1 = M\xi_2$). Поэтому величина $(\bar{x} - \bar{y})/\sigma$ распределена нормально с параметрами 0 и $\sqrt{\frac{1}{n} + \frac{1}{m}}$. Случайные величины $(n-1)s_1^2/\sigma^2$ и $(n-1)s_2^2/\sigma^2$ не зависят от $\bar{x}$ и $\bar{y}$ и распределены каждая по закону χ^2 с $n-1$ и $m-1$ степенями свободы соответственно. Поэтому величина

$$\frac{S^2}{\sigma^2}(n+m-2) = \frac{s_1^2}{\sigma^2}(n-1) + \frac{s_2^2}{\sigma^2}(m-1)$$

имеет распределение χ^2 с $n+m-2$ степенями свободы и не зависит от $\bar{x} - \bar{y}$. Отсюда отношение

$$t = \frac{\bar{x} - \bar{y}}{S\sqrt{n+m}}\sqrt{nm} = \frac{\dfrac{\bar{x} - \bar{y}}{\sigma\sqrt{m+n}}\sqrt{mn}}{S/\sigma} \tag{19}$$

имеет распределение Стьюдента с $n + m - 2$ степенями свободы. ▶

Как и выше, можно показать, что принцип отношения правдоподобия приводит к критической области, определяемой статистикой (19) (для фиксированного уровня значимости α область S определяется из условия

$$P\{|\bar{x} - \bar{y}| > \Delta_\alpha\} \leqslant \alpha).$$

2. Если же гипотеза о равенстве дисперсий $D\xi_1 = D\xi_2$ не подтверждается, то случайную величину, описывающую отклонения $\bar{x}$ и $\bar{y}$, закон распределения которой не зависит от параметров распределения случайных величин ξ_1 и ξ_2, построить уже не удается.

Проверку гипотезы о равенстве средних двух независимых нормальных совокупностей проводят в этом случае следующим образом: рассматривается случайная величина

$$\tau = \frac{\bar{x} - \bar{y}}{\sqrt{\dfrac{D\xi_1}{n} + \dfrac{D\xi_2}{m}}},$$

которая имеет нормальное распределение с параметрами 0 и 1. Если n и m достаточно велики, то замена точных значений $D\xi_1$ и $D\xi_2$ их оценками s_1^2 и s_2^2 не очень нарушает распределения случайной величины и можно считать, что случайная величина

$$\tau' = \frac{\bar{x} - \bar{y}}{\sqrt{ms_1^2 + ns_2^2}} \sqrt{nm}$$

имеет приблизительно нормальное распределение с параметрами 0 и 1 (при верной гипотезе $M\xi_1 = M\xi_2$). Этим удобно воспользоваться при построении критической области S, описывающей редкие отклонения $\bar{x}$ и $\bar{y}$.

§ 3. Критерии согласия

Другую важную группу гипотез образуют непараметрические гипотезы, из которых мы остановимся здесь на гипотезах о законах распределения.

Очень часто из тех или иных соображений может быть высказана гипотеза о характере закона распределения наблюдаемой случайной величины. К примеру, если случайная величина ξ обусловлена суммарным воздействием большого числа приблизительно одинаковых факторов, то, руководствуясь центральной предельной теоремой, разумно предполагать, что ξ имеет нормальное распределение.

Как мы уже знаем, представление об истинной функции распределения случайной величины ξ можно составить по эмпирической функции распределения. Поэтому если высказана гипотеза о том, что истинная функция распределения случайной величины ξ есть $F_\xi(x)$, то естественно изучать поведение отклонения предполагаемой функции $F_\xi(x)$ от наблюдаемой эмпирической $F_n(x)$. Если отклонение $F_\xi(x)$ от $F_n(x)$ окажется значительным, то $F_\xi(x)$ не может быть функцией распределения случайной величины ξ. Причем значительными будем считать такие отклонения, вероятность наблюдения которых в эксперименте при верной гипотезе очень мала.

Построим случайную величину $d = d\{F_\xi(x), F_n(x)\}$, описывающую различие между гипотетической функцией $F_\xi(x)$ и наблюдаемой $F_n(x)$. Задавая уровень значимости α, определяем число d_α такое, что

$$P\{d > d_\alpha\} \leqslant \alpha.$$

Тогда гипотеза о виде закона распределения считается согласующейся с экспериментальными данными, если $d < d_\alpha$. В противном случае гипотеза считается плохо согласующейся с экспериментом и отвергается на уровне значимости α.

Выбирая ту или иную меру $d\{F_\xi(x), F_n(x)\}$ отличия $F_\xi(x)$ от $F_n(x)$, будем получать для проверки изучаемой гипотезы различные критерии.

7 Зак. 244

3.1. Критерий Колмогорова—Смирнова

Пусть в качестве d взята величина

$$D_n = \sup_x \left|F_n(x) - F_\xi(x)\right|. \tag{20}$$

Теорема Гливенко—Кантелли утверждает, что $D_n \to 0$, если объем выборки неограниченно возрастает. Рассмотрим

$$\lambda_n = \sqrt{n} D_n.$$

Имеет место следующая

Теорема. *Пусть гипотетическая функция $F_\xi(x)$ непрерывна. Тогда функция распределения случайной величины λ_n не зависит от вида $F_\xi(x)$.*

◄ Пусть $X_1, X_2, \ldots, X_n$ — выборка объема n из закона распределения случайной величины ξ. Рассмотрим набор случайных величин $Y_i = F_\xi(X_i)$, $i = 1, 2, \ldots, n$.

Лемма. *Если ξ — случайная величина с законом распределения $F_\xi(x)$, причем $F_\xi(x)$ непрерывна, то случайная величина $\eta = F_\xi(\xi)$ равномерно распределена на $[0,1]$, т. е.*

$$F_\eta(y) = P\{\eta < y\} = \begin{cases} 0, & y < 0, \\ y, & 0 \leqslant y < 1, \\ 1, & y \geqslant 1. \end{cases}$$

$$◄\ P\{\eta < y\} = P\left\{F_\xi(\xi) < y\right\}.$$

Так как функция распределения $F_\xi(x)$ монотонно возрастает и $0 \leqslant F_\xi(x) \leqslant 1$, то

$$P\{\eta < y\} = \begin{cases} 0, & y < 0, \\ 1, & y \geqslant 1. \end{cases}$$

На отрезке $0 \leqslant y < 1$

$$P\left\{F_\xi(\xi) < y\right\} = P\left\{\xi < F_\xi^{-1}(y)\right\} = F_\xi\left(F_\xi^{-1}(y)\right) = y. ►$$

На основании леммы, набор случайных величин Y_i, $i = 1,2, \ldots, n$, образует выборку объема n из равномерного распределения. Пусть $F_\eta(y)$ — функция равномерного распределения, $F_\eta^n(y)$ — эмпирическая функция равномерного распределения, построенная на выборке $Y_i = F_\xi(X_i)$,

$$F_\eta^n(y) = \frac{\text{количество } Y_i, \text{ которые меньше } y}{n},$$

и x таково, что $F_\xi(x) = y$. Рассмотрим

$$F_\eta^n(y) - F_\eta(y) = \frac{\text{количество } Y_i, \text{ которые меньше } y}{n} - y =$$

$$= \frac{\text{количество } F_\xi(X_i), \text{ которые меньше } y}{n} - y.$$

Вследствие монотонности функции $F_\xi(x)$

$$\left\{\text{количество } F_\xi(X_i), \text{ которые меньше } y\right\} = \left\{\text{количество } X_i, \text{ которые меньше } x\right\}.$$

Учитывая это, получаем

$$F^n_\eta(y) - F_\eta(y) = \frac{\text{количество } X_i, \text{ которые меньше } x}{n} - F_\xi(x) = F_n(x) - F_\xi(x).$$

Отсюда

$$\sup_y \left| F^n_\eta(y) - F_\eta(y) \right| = \sup_x \left| F_n(x) - F_\xi(x) \right|.$$

Но левая часть последнего соотношения не зависит от вида функции $F_\xi(x)$, следовательно, не зависит и правая. ▶

Таким образом, введенная нами мера λ_n, описывающая различия эмпирической и гипотетической функций распределения, не зависит от вида гипотетической функции распределения, а определяется лишь объемом выборки n.

Если объем выборки неограниченно возрастает, то функция распределения случайной величины λ_n мало отличается от некоторой фиксированной функции. А именно, имеет место *теорема Колмогорова*

$$K(x) = \lim_{n\to\infty} P\{\lambda_n < x\} = 1 - 2\sum_{k=1}^{\infty}(-1)^{k+1}e^{-2k^2x^2}, \qquad x \geqslant 0.$$

Независимость предельного распределения λ_n от гипотетической функции $F_\xi(x)$ позволяет построить критерий для проверки гипотезы о согласованности эмпирических данных с гипотетическим распределением.

Пусть гипотеза верна, тогда (если n достаточно велико, $n \geqslant 50$)

$$P\{\lambda_n \geqslant x\} \approx 1 - K(x).$$

Задавая уровень значимости α, определяем λ_α из уравнения

$$P\{\lambda_n \geqslant \lambda_\alpha\} = \alpha.$$

В соответствии с общей установкой гипотезу считаем согласующейся с эмпирическими данными, если полученное по конкретным данным значение

$$\lambda_n = \sqrt{n}\,\sup\left| F_n(x) - F_\xi(x) \right|$$

не превышает λ_α, в противном случае гипотезу отвергаем.

3.2. Критерий χ^2 Пирсона

Одним из наиболее часто употребляемых на практике критериев согласия является критерий χ^2 Пирсона. В качестве меры несогласованности гипотетического и эмпирического распределений рассмотрим следующую величину.

Пусть $X_1, X_2, \ldots, X_n$ — выборка из закона распределения случайной величины ξ.

Разобьем числовую прямую на s разрядов и найдем частоту ν_i попадания случайной величины ξ в i-й разряд разбиения Δ_i:

$$\nu_i = \frac{\text{количество } X_j, \text{ принадлежащих } \Delta_i}{n} = \frac{n_i}{n}, \qquad i = 1, 2, \ldots, s.$$

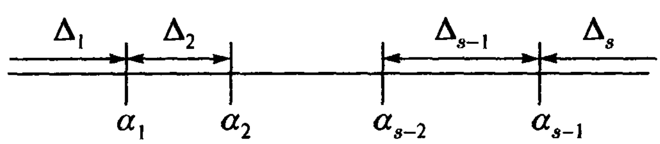


Рис. 4

Пусть $F_\xi(x)$ — гипотетическое распределение случайной величины ξ. Тогда вероятность того, что случайная величина ξ принимает значения в i-м разряде разбиения Δ_i, равна

$$p_i = F_\xi(a_i) - F_\xi(a_{i-1}).$$

Здесь a_i, a_{i-1} — концы i-го разряда разбиения ($a_0 = -\infty$, $a_s = \infty$) (рис. 4).

Рассмотрим величину

$$\chi^2_{s-1} = n\sum_{i=1}^{s}\frac{(\nu_i - p_i)^2}{p_i} = \sum_{i=1}^{s}\frac{(n_i - np_i)^2}{np_i}. \tag{21}$$

Если наша гипотеза верна, то отклонения $\nu_i - p_i$ в большинстве случаев должны быть малы, поэтому в качестве меры различия эмпирического и теоретического законов распределения целесообразно взять величину (21).

Имеет место теорема о независимости предельного распределения для χ^2_{s-1} от вида гипотетической функции распределения.

Теорема Пирсона.

$$\lim_{n\to\infty} P\{\chi^2_{s-1} < x\} = G_{s-1}(x), \qquad x \geqslant 0.$$

Здесь $G_{s-1}(x)$ — χ^2-распределение с $s-1$ степенью свободы.

При достаточно большом числе наблюдений эта теорема может быть использована для построения критерия согласия. Пусть α — уровень значимости. Решив уравнение

$$P\{\chi^2_{s-1} \geqslant x\} = \alpha,$$

найдем пограничное значение χ^2_α, сравнивая с которым экспериментальное значение (21), будем делать заключение о согласованности или несогласованности нашей гипотезы с экспериментом.

Остановимся на чувствительности критерия χ^2 несколько подробнее. Пусть наша гипотеза ошибочна и истинные значения q_i вероятностей попадания в i-й разряд разбиения отличны от найденных нами вероятностей p_i. Тогда случайная величина (21) уже не будет следовать распределению χ^2 с $s-1$ степенью свободы и для математического ожидания величины (21) получим

$$M\chi^2_{s-1} = n\sum_{i=1}^{s}\frac{M(\nu_i - p_i)^2}{p_i} = n\sum_{i=1}^{s}\frac{M(\nu_i - p_i + q_i - q_i)^2}{p_i} =$$

$$= n\sum_{i=1}^{s}\frac{M(\nu_i - q_i)^2}{p_i} + 2n\sum_{i=1}^{s}\frac{M(\nu_i - q_i)}{p_i} + n\sum_{i=1}^{s}\frac{(q_i - p_i)^2}{p_i}.$$

Но $M(\nu_1 - q_i)^2 = q_i(1-q_i)$, $M(\nu_1 - q_i) = 0$, и поскольку хотя бы одна из вероятностей $q_i \neq p_i$, то $\sum(q_i - p_i)^2/p_i \neq 0$. Поэтому

$$M\chi^n_{s-1} \geqslant n\sum_{i=1}^{s}\frac{(q_i - p_i)^2}{p_i}.$$

Тем самым, с ростом объема выборки указанная величина неограниченно возрастает, если только наша гипотеза неверна. Таким образом, на практике, если число наблюдений достаточно велико, неверная гипотеза будет отвергнута.

Практические рекомендации к применению критерия Пирсона следующие: желательно, чтобы разбиения на разряды проводились таким образом, чтобы $np_i \geqslant 10$. Число разрядов разбиения при этом должно быть не менее 7–8. Если же эмпирических данных очень много (скажем, число разрядов превышает $s = 30$), то целесообразно воспользоваться для построения критерия не распределением χ^2_{s-1}, а предельным для него при $s \to \infty$ нормальным.

Сравнивая критерий Колмогорова и критерий Пирсона, заметим, что первый более точен и приводит на практике к менее громоздким вычислениям, чем второй.

Следует, однако, отметить, что в практической ситуации гипотетический закон распределения $F_\xi(x)$ может быть точно указан крайне редко. Более реальной является такая ситуация, когда можно лишь высказать предположение о целой группе гипотетических законов $F_\xi(x;\ \beta_1, \ldots, \beta_k)$, каждый из которых определяется фиксированным набором параметров $\beta_1, \ldots, \beta_k$. В этом случае гипотеза выглядит следующим образом:

распределение случайной величины ξ описывается законом $F_\xi(x;\ \beta_1, \ldots, \beta_k)$ при некотором наборе параметров β_j.

При замене неизвестных параметров их оценками, найденными по выборке, следует иметь в виду, что для одного и того же параметра можно указать очень много различных оценок. Поэтому, заменяя истинные значения неизвестных параметров их оценками, мы портим предельные распределения рассмотренных нами мер отличия — основные теоремы предыдущего и настоящего параграфов становятся неверными.

В этом случае описанными критериями пользоваться, вообще говоря, уже нельзя. Так, например, применение в указанной ситуации критерия Колмогорова приводит к тому, что чем больше параметров мы оценили по выборке, тем лучшее согласие он покажет даже при неверной гипотезе, тогда как критерий Пирсона допускает некоторое видоизменение таким образом, что он остается пригодным и в описанной выше ситуации.

Теорема. *Пусть $\beta_1^*, \beta_2^*, \ldots, \beta_k^*$ — оценки максимального правдоподобия или оценки, полученные по методу моментов. Тогда случайная величина* (21) *имеет распределение χ^2 с $s - k - 1$ степенями свободы, т. е. число степеней свободы распределения случайной величины* (21) *уменьшается на число оцениваемых по выборке параметров.*

Пример. Пусть в эксперименте получена выборка объема n из распределения случайной величины ξ

$$X_1, X_2, \ldots, X_n$$

и высказана гипотеза о нормальности распределения случайной величины ξ.

Применим критерий χ^2. Производя разбиение числовой прямой на разряды, вычисляем значения эмпирических частот ν_i попадания случайной величины ξ в i-й разряд разбиения. Для подсчета теоретических вероятностей попадания в i-й разряд разбиения необходимо знать параметры m и σ нормального распределения. Заменим их оценками

$$m \approx \bar{x} = \frac{1}{n}\sum_{i=1}^{n} X_i; \qquad \sigma^2 \approx s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{x})^2$$

и вычислим вероятности

$$p_i = \frac{1}{s\sqrt{2\pi}} \int\limits_{\Delta_i} \exp\left\{-\frac{(x-\bar{x})^2}{2s^2}\right\} dx.$$

Находим

$$\chi^2 = n\sum_{i=1}^{s} \frac{(\nu_i - p_i)^2}{p_i}.$$

Для того чтобы определить пограничное значение χ^2_α, заметим, что число степеней свободы случайной величины χ^2 равно $s - 3$, так как мы оценили по выборке два неизвестных параметра распределения: m и σ^2. Поэтому χ^2_α ищем из уравнения

$$1 - G_{s-3}(x) = \alpha.$$

Здесь $G_{s-3}(x)$ — χ^2-распределение с $s - 3$ степенями свободы.

В заключение отметим, что если по выборке оценено значительное количество неизвестных параметров, то тем самым гипотетическая функция распределения искусственно приближена к эмпирической и критерий в этих случаях дает неоправданно высокую степень согласованности. Поэтому, если число степеней свободы оказывается малым (3–4 и меньше), то для повышения уровня достоверности допускаемых нами выводов наряду с критерием χ^2 следует использовать и другие критерии и оценки.

Глава XLV

СТАТИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ ЗАВИСИМОСТЕЙ

В этой главе мы рассмотрим несколько часто встречающихся в приложениях задач, связанных с восстановлением зависимостей между наблюдаемыми в эксперименте переменными.

Общая структура задач, о которых пойдет речь ниже, такова: наблюдаются две группы переменных — объясняющая (или исходная), описываемая вектором $\mathbf{x} = (x_i)_{i=1}^n$ и итоговая (или выходная), описываемая скаляром y. Следует дать ответы на вопросы:

— есть ли связь между переменными $\mathbf{x}$ и y?
— если связь есть, то какова возможная форма этой связи?
— каковы качественные характеристики этой связи?

Ответы на перечисленные вопросы (да и сами вопросы) могут выглядеть по разному в зависимости как от природы переменных $\mathbf{x}$ и y, наблюдаемых в эксперименте, так и от условий проведения измерений.

Различают следующие три основные модели.

1. Переменные $\mathbf{x}$ — *детерминированные*, а переменная $y = \eta$ — *случайная величина*.

Мы будем говорить что y зависит от $\mathbf{x}$, если изменение переменных $\mathbf{x}$ влечет за собой изменение *закона распределения* случайной величины $y = \eta$

$$\Phi_\eta(A) = P\{\eta \in A\} = P_\eta(\mathbf{x}). \tag{1}$$

Если при этом математическое ожидание случайной величины η может быть описано некоторой функцией переменных $\mathbf{x}$

$$M(\eta) = \varphi(\mathbf{x}) = \varphi(x_1, x_2, \dots, x_n), \tag{2}$$

то мы, следуя сложившейся традиции, будем называть такую зависимость *регрессионной*.

Широкий класс регрессионных моделей описывается следующим примером.

Пример. Пусть $\mathbf{x} = (x_i)_{i=1}^n$ — неслучайный вектор, $y = f(\mathbf{x})$. В эксперименте измеряются значения y так, что

$$\eta = y + \xi = f(\mathbf{x}) + \xi. \tag{3}$$

Здесь η — измеренное значение переменной y, ξ — ошибка измерения с $M\xi = 0$ (последнее означает отсутствие систематических ошибок измерения).

Соотношение (3) — модель измерений — доставляет нам пример регрессионной зависимости, так как из сделанных выше предположений следует, что

$$M\eta = f(\mathbf{x}). \tag{4}$$

Как правило, в рассматриваемой ситуации вид линии регрессии $f(\mathbf{x}) = f(\mathbf{x}; \bar{\alpha})$ бывает известен с точностью до вектора параметров $\bar{\alpha} = (\alpha_i)_{i=1}^k$ и задача установления *наличия зависимости* между $\mathbf{x}$ и y не стоит. Требуется только определить возможные значения параметров $\bar{\alpha}$ и установить «качество» описания величины y посредством функции $f(\mathbf{x}, \bar{\alpha})$

2. Переменные $\mathbf{x} = \bar{\xi}$ и $y = \eta$ — *случайные величины*, совместный закон распределения которых дается функцией $\Phi_{\bar{\xi},\eta}(A, B) = P\{\bar{\xi} \in A; \eta \in B\}$. Известно, что случайные величины $\bar{\xi}$ и η независимы тогда и только тогда, когда выполняется соотношение

$$\Phi_{\bar{\xi},\eta}(A, B) = \Phi_{\bar{\xi}}(A) \cdot \Phi_\eta(B). \tag{5}$$

Если же равенство (5) не имеет места, то случайные величины зависимы. При этом, как и выше, выделим специальный случай, когда изменение математического ожидания случайной величины η описывается некоторой функцией от $\bar{\xi}$,

$$M[\eta/\bar{\xi}] = f(\bar{\xi}) \not\equiv \text{const}, \tag{6}$$

и будем говорить, что соотношение (6) описывает регрессию η на $\bar{\xi}$. (Естественно, предполагается, что условные средние (6) существуют.)

Идентификация *наличия* или *отсутствия* зависимости между η и $\bar{\xi}$ представляет в рассматриваемом случае содержательную задачу, равно как и проблема установления *формы* регрессионной связи (6).

Важным для приложений частным случаем рассмотренной ситуации является случай совместной нормальности $\bar{\xi}$ и η. В этом случае проблема определения *формы* связи (6) не стои́т — регрессия η на $\bar{\xi}$ является линейной и дается соотношением

$$M[\eta/\bar{\xi}] = \mathbf{K}^T[\eta, \bar{\xi}] \cdot \mathbf{K}^{-1}[\bar{\xi}, \bar{\xi}](\bar{\xi} - M\bar{\xi}) + M\eta. \tag{7}$$

Здесь $\mathbf{K}[\bar{\xi}, \bar{\xi}] = M(\bar{\xi} - M\bar{\xi})(\bar{\xi} - M\bar{\xi})^T$ — ковариационная матрица вектора $\bar{\xi}$, а $\mathbf{K}[\eta, \bar{\xi}] = M(\eta - M\eta)(\bar{\xi} - M\bar{\xi})^T$ — ковариационная матрица η и $\bar{\xi}$.

Как следует из соотношения (7), регрессия $M[\eta/\bar{\xi}]$ полностью определяется корреляционными характеристиками совместного распределения величин η и $\bar{\xi}$.

Конечно, регрессия может оказаться линейной не только для совместно нормальных случайных величин. В любом случае мы будем называть соответствующую связь между переменными *корреляционной*.

3. Переменные $\mathbf{x}$ и y — *детерминированные, неслучайные* переменные и в эксперименте измеряются абсолютно точно. При этом уже факт наблюдения пар конкретных значений $(\mathbb{X}, Y)$ в одном и том же опыте определяет положительный ответ на первый вопрос, и речь может идти только об установлении формы зависимости y и $\mathbf{x}$.

Здесь возможны различные постановки и способы решения — интерполяция, сглаживание и т. п. Это задачи классического анализа и мы на них останавливаться не будем.

Ниже будут рассмотрены статистические процедуры, исследующие ситуации **1** и **2**. Учитывая важность для приложений случая совместной нормальности изучаемых случайных величин и их линейных описаний, основное внимание будет уделено исследованию корреляционных связей.

§ 1. Случайные переменные. Корреляционные связи

Пусть в эксперименте Ω наблюдаются случайные переменные $\bar{\xi} = (\xi_1\ \xi_2\ \dots\ \xi_n)$ и η. Проведено N опытов и получена выборка из совместного закона распределения величин $(\bar{\xi}, \eta)$,

$$(\mathbb{X}_1, Y_1), \quad \dots, \quad (\mathbb{X}_N, Y_N), \tag{1}$$

здесь $\mathbb{X}_i = (X_i^s)_{i=1,\dots,N,\ s=1,\dots,n}$.

Требуется по выборке (1) сделать заключение о наличии или отсутствии зависимости между переменными $\bar{\xi}$ и η и оценить эту зависимость.

1.1. Значимость коэффициента корреляции

Исследование вопроса о наличии связей между η и $\bar{\xi}$ мы начнем с простейшего случая $n = 1$. Тогда $\bar{\xi} = \xi$ и η — скалярные случайные величины, а выборка (1) — это выборка $(X_1, Y_1), \ldots, (X_N, Y_N)$ из совместного закона распределения пары (ξ, η).

Отличие от нуля коэффициента корреляции $\rho_{\xi\eta}$ пары (ξ, η) является достаточно надежным показателем *наличия* зависимости между ξ и η. Эта зависимость носит корреляционный характер, т. е. может быть удовлетворительно описана линейной функцией в том смысле, что имеется четко выраженная тенденция к линейному изменению η относительно ξ. Точнее, в этом случае величина η представима в виде

$$\eta = M\eta + \rho_{\xi\eta} \frac{\sigma_\eta}{\sigma_\xi} (\xi - M\xi) + \zeta, \tag{2}$$

и вклад остатка ζ в рассеяние η тем меньше, чем ближе $|\rho_{\xi\eta}|$ к 1. В случае $|\rho_{\xi\eta}| = 1$ между ξ и η имеется жесткая линейная связь, позволяющая с вероятностью 1 восстанавливать неизвестные значения η по измеренным значениям ξ.

Если $\rho_{\xi\eta} = 0$, то это, вообще говоря, еще не свидетельствует об отсутствии зависимости между ξ и η, но является надежным свидетельством в пользу отсутствия *линейной* зависимости.

Значит, для идентификации наличия или отсутствия *линейной связи* между скалярными случайными величинами достаточно уметь по выборке (1) делать заключение о равенстве или неравенстве нулю коэффициента корреляции $\rho_{\xi\eta}$ пары (ξ, η).

Пусть r_N — выборочный (эмпирический) коэффициент корреляции пары (ξ, η)

$$r_N = \frac{\sum_1^N (X_i - \bar{x})(Y_i - \bar{y})}{\sqrt{\sum_1^N (X_i - \bar{x})^2 \cdot \sum_1^N (Y_i - \bar{y})^2}}. \tag{3}$$

Он является оценкой истинного, но неизвестного коэффициента корреляции $\rho_{\xi\eta}$. Если $\rho_{\xi\eta} = 0$, то чисто случайные колебания r_N могут привести к значению $r_N \neq 0$. Как же выяснить, равен или не равен нулю коэффициент корреляции ρ по наблюдению за значением r_N, вычисленному по формуле (3)?

Теорема. *Пусть (ξ, η) — двумерная нормальная случайная величина, $(X_i, Y_i)_{i=1}^N$ — выборка значений (ξ, η), полученная в эксперименте, и r_N — выборочный коэффициент корреляции (3). Тогда, если $\rho = 0$, то величина t^r_{N-2}, даваемая соотношением*

$$t^r_{N-2} = \frac{r_N}{\sqrt{1 - r_N^2}} \sqrt{N - 2}, \tag{4}$$

имеет распределение Стьюдента с $N - 2$ степенями свободы.

Эта теорема дает возможность построить процедуру статистической проверки гипотезы о значимости коэффициента корреляции (т. е. гипотезы об отличии ρ от нуля) следующим образом. Так как при фиксированном значении $N > 2$ величина (4) монотонно возрастает (рис. 1), то «маленьким» значениям r_N отвечают «маленькие» значения t^r. Если истинный коэффициент корреляции ρ равен нулю, то его эмпирический аналог r_N в подавляющем большинстве случаев будет «маленьким», а, следовательно, в подавляющем большинстве случаев должна быть маленькой и величина t^r. Зададим некоторый (близкий к нулю) уровень значимости $\varkappa$ и найдем значение $t_\varkappa$ такое, что

$$P\left\{\left|t^r_{N-2}\right| < t_\varkappa\right\} \geqslant 1 - \varkappa.$$

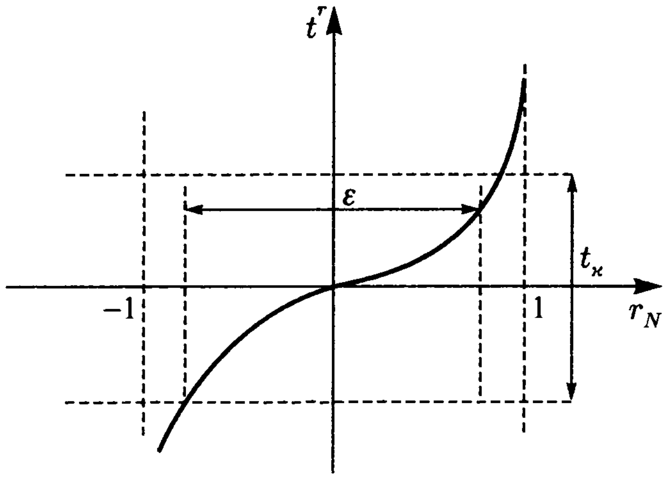


Рис. 1

При этом для r_N будет иметь место соотношение

$$P\left\{|r_N| < \frac{t_\varkappa}{\sqrt{N-2+t_\varkappa^2}}\right\} \geqslant 1 - \varkappa.$$

Если окажется, что $|r_N| > t_\varkappa/\sqrt{N-2+t_\varkappa^2}$, то гипотеза о равенстве нулю коэффициента корреляции должна быть признана не согласующейся с экспериментальными данными. Если же $|r_N| < t_\varkappa/\sqrt{N-2+t_\varkappa^2}$, то можно считать, что $\rho = 0$ и, следовательно, величины ξ и η независимы (напомним, что мы рассматриваем случай совместной нормальности ξ и η, когда равенство нулю коэффициента корреляции эквивалентно независимости).

Если совместное распределение (ξ, η) не является нормальным, то основная теорема, позволившая построить критерий значимости для коэффициента корреляции, уже оказывается неверной. В этом случае рекомендуется использовать статистику Фишера $\mathcal{Z}$, определяемую соотношением

$$\mathcal{Z} = \frac{1}{2}\ \ln\frac{1+r_N}{1-r_N}. \tag{5}$$

Установлено, что уже при достаточно небольших значениях N величина $\mathcal{Z}$ приближенно нормальна с параметрами

$$M\mathcal{Z} = \frac{1}{2}\ \ln\frac{1+\rho}{1-\rho} + \frac{\rho}{2(N-1)}, \qquad D\mathcal{Z} \approx \frac{1}{N-3},$$

где ρ — истинный коэффициент корреляции величин ξ и η. Учитывая монотонность функции $\mathcal{Z} = \mathcal{Z}(r)$ на промежутке $(-1, 1)$, мы можем, как и выше, построить процедуру статистической проверки гипотезы о значимости коэффициента корреляции.

Пусть $\rho = 0$. Тогда $\mathcal{Z} = N\left[0;\ 1/\left(\sqrt{N-3}\right)\right]$. Для $\varkappa \ll 1$ определим $\mathcal{Z}_\varkappa > 0$ так, что

$$P\{|z| < \mathcal{Z}_\varkappa\} \geqslant 1 - \varkappa.$$

При этом[1)]

$$P\{|r| < \operatorname{th} \mathcal{Z}_\varkappa\} \geqslant 1 - \varkappa.$$

[1)] Если $(1/2)\ \ln(1+r)/(1-r) = z$, то $r = (e^z - e^{-z})/(e^z + e^{-z}) = \operatorname{th} z$ — тангенс гиперболический z. В силу нечетности и монотонности функции $\operatorname{th} z$ неравенство $|z| < \mathcal{Z}_\varkappa$ эквивалентно неравенству $|r_N| < \operatorname{th} \mathcal{Z}_\varkappa$.

Гипотеза $\rho = 0$ принимается на уровне значимости $\varkappa$, если выполняется неравенство $|r_N| < \operatorname{th} \mathcal{Z}_\varkappa$; в противном случае признается значимым отличие коэффициента корреляции от нуля.

Таким образом, если предложенные выше критерии признали коэффициент корреляции значимо отличным от нуля, то можно сделать вывод о наличии зависимости между случайными величинами η и ξ, и чем ближе значения коэффициента корреляции к единице, тем лучше идентифицируемая зависимость описывается линейным соотношением (2).

Если отличие от нуля коэффициента корреляции признано незначимым, то в случае совместной нормальности это, как уже отмечалось выше, свидетельствует о независимости наблюдаемых случайных величин. В случае же, когда совместное распределение отличается от нормального, между наблюдаемыми случайными величинами возможно наличие зависимости, причем не исключается даже жесткая функциональная связь.

1.2. Множественная регрессия. Метод наименьших квадратов

Пусть теперь $\bar{\xi} = (\xi_i)_{i=1}^n$ — векторная случайная величина.

Если $f(\bar{\xi})$ — некоторая неслучайная функция n переменных, описывающая связь между η и $\bar{\xi}$ так, что η представима в виде

$$\eta = f(\bar{\xi}) + \zeta, \quad M\zeta = 0, \tag{6}$$

то естественно оценивать качество описания случайной величины η функцией $f(\bar{\xi})$ с помощью «остатка» ζ — чем меньше (в точно определенном смысле) ζ, тем лучше $f(\bar{\xi})$ описывает величину η. Если в качестве критерия малости ζ взять ее дисперсию, то наилучшей из всех функций $f(\bar{\xi})$ будет удовлетворяющая условию

$$\min D\zeta = \min M\left(\eta - f(\bar{\xi})\right)^2.$$

Оказывается, что при достаточно естественных предположениях о законе распределения совокупности $(\eta, \bar{\xi})$, наилучшая (в смысле минимума дисперсии остатка) функция $f(\bar{\xi})$ существует и, как и в одномерном случае, совпадает с условным средним случайной величины η относительно $\bar{\xi}$.

Теорема. *Если существует условное среднее $\varphi(\bar{\xi}) = M[\eta/\bar{\xi}]$, то*

$$\operatorname{argmin} M[\eta - f(\bar{\xi})]^2 = M[\eta/\bar{\xi}] = \varphi(\bar{\xi}).$$

◀ Рассмотрим

$$\begin{aligned} M[\eta - f(\bar{\xi})]^2 &= M[\eta - \varphi(\bar{\xi}) + \varphi(\bar{\xi}) - f(\bar{\xi})]^2 = \\ &= M\left[(\eta-\varphi)^2 + 2(\eta-\varphi)(\varphi - f) + (\varphi - f)^2\right] = \\ &= M(\eta-\varphi)^2 + M(\varphi - f)^2 + 2M(\eta-\varphi)(\varphi - f). \end{aligned} \tag{7}$$

Последнее слагаемое равно нулю. Чтобы не загромождать изложение выкладками, покажем это для частного случая, когда $n = 1$ и (ξ, η) непрерывна.

Пусть $g_{\xi\eta}(x, y)$ — совместная плотность величин ξ и η и $g_{\eta/\xi}(y/x)$ — условная плотность η относительно ξ. Тогда

$$\varphi(\xi) = M[\eta/\xi] = \int\limits_{\mathbb{R}} g_{\eta/\xi}(y/\xi)y\, dy. \tag{8}$$

С учетом $g_{\xi\eta}(x, y) = g_{\eta/\xi}(y/x)g_{\xi}(x)$ получаем, что

$$\begin{aligned} M(\eta - \varphi(\xi))(\varphi(\xi) - f(\xi)) &= \iint\limits_{\mathbb{R}^2} g_{\xi\eta}(x, y)(y - \varphi(x))(\varphi(x) - f(x))\, dx\, dy = \\ &= \iint\limits_{\mathbb{R}^2} g_{\eta/\xi}(y/x)g_{\xi}(x)(y - \varphi(x))(\varphi(x) - f(x))\, dx\, dy = \\ &= \int\limits_{\mathbb{R}} g_{\xi}(x)(\varphi - f)\, dx \int\limits_{\mathbb{R}} g_{\eta/\xi}(y/x)(y - \varphi)\, dy. \end{aligned}$$

Но в силу соотношения (8) внутренний интеграл равен нулю, откуда и следует искомое.

Теперь соотношение (7) принимает вид

$$M[\eta - f(\xi)]^2 = M\left[\eta - \varphi(\bar{\xi})\right]^2 + M\left[\varphi(\bar{\xi}) - f(\bar{\xi})\right]^2, \tag{9}$$

откуда (в силу неотрицательности $M\left[\varphi(\bar{\xi}) - f(\bar{\xi})\right]^2$) немедленно заключаем, что

$$M\left[\eta - f(\bar{\xi})\right]^2 \geqslant M\left[\eta - \varphi(\bar{\xi})\right]^2. \blacktriangleright$$

Таким образом, с точки зрения критерия минимума дисперсии остатка ζ, наилучшее описание зависимости η от $\bar{\xi}$ дается линией регрессии η на $\bar{\xi}$

$$\eta = \varphi(\bar{\xi}) + \zeta = M(\eta/\xi) + \zeta. \tag{10}$$

Отметим, что соотношению (10) отвечает дисперсионное соотношение, позволяющее оценивать степень тесноты связи между η и $\bar{\xi}$. Имеет место следующее утверждение.

Теорема (о разложении дисперсии). *В условиях предыдущей теоремы*

$$D\eta = D\varphi + D\zeta, \tag{11}$$

где $D\eta = M[\eta - M\eta]^2$; $D\varphi = M\left[\varphi(\bar{\xi}) - M\eta\right]^2$; $D\zeta = M\left[\eta - \varphi(\bar{\xi})\right]^2$.

◀ Для доказательства соотношения (11) подставим в тождество (9) $f(\xi) \equiv M\eta$ и заметим, что

$$\begin{aligned} M\varphi(\xi) &= \int\limits_{\mathbb{R}} \varphi(x)g_{\xi}(x)\, dx = \int\limits_{\mathbb{R}} g_{x}(x)\, dx \int\limits_{\mathbb{R}} g_{\eta/\xi}(y/x)y\, dy = \\ &= \iint\limits_{\mathbb{R}^2} g_{\xi\eta}(x, y)y\, dx\, dy = M\eta. \blacktriangleright \end{aligned}$$

Разделим обе части соотношения (11) на $D\eta$

$$1 = \frac{D\varphi}{D\eta} + \frac{D\zeta}{D\eta}$$

и введем величину $\Theta^2_{\eta/\bar{\xi}}$ — *корреляционное отношение величины* η *относительно* $\bar{\xi}$ — равенством

$$\Theta^2_{\eta/\bar{\xi}} = 1 - \frac{D\zeta}{D\eta} = \frac{D\varphi}{D\eta}. \tag{12}$$

Тогда соотношение (11) запишется в виде

$$D\eta = \Theta^2_{\eta/\bar{\xi}} \cdot D\eta + \left(1 - \Theta^2_{\eta/\bar{\xi}}\right) D\eta, \tag{13}$$

откуда видно, что если показатель (12) близок к единице — связь между η и $\bar{\xi}$ достаточно тесная в том смысле, что среднее значение η закономерно меняется с изменением $\bar{\xi}$, если же $\Theta^2_{\eta/\bar{\xi}}$ близко к нулю, то такой тенденции нет — основной вклад в изменение η вносит случайный остаток ζ.

Важным случаем регрессионных связей (10) являются линейные зависимости

$$l(\bar{\xi}) = \alpha_1(\xi_1 - M\xi_1) + \alpha_2(\xi_2 - M\xi_2) + \ldots + \alpha_n(\xi_n - M\xi_n) + M\eta. \tag{14}$$

Как уже отмечалось выше (см. соотношение (7)), в случае совместной нормальности $(\bar{\xi}, \eta)$ регрессия всегда линейна. Она может оказаться таковой и в других случаях.

Поскольку любая регрессия $\varphi(\bar{\xi})$, независимо от того, линейна она или нет, должна удовлетворять условию $\varphi(\bar{\xi}) = \operatorname{argmin} M\left[\eta - f(\bar{\xi})\right]^2$, то коэффициенты $\bar{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ линейной регрессии (14) могут быть найдены следующим образом. Положив $\bar{\alpha} = \underset{S_i}{\operatorname{argmin}}\ M\left[\eta - M\eta - \sum\limits_{i=1}^{n} S_i(\xi_i - M\xi_i)\right]^2$, заметим, что

$$M\left[\eta - M\eta - \sum_1^n S_i(\xi_i - M\xi_i)\right]^2 = M[\eta - M\eta]^2 - 2\sum_1^n S_i M(\eta - M\eta)(\xi_i - M\xi_i) + \\ + \sum S_i S_j M(\xi_i - M\xi_i)(\xi_j - M\xi_j) = M(S_1, \ldots, S_n).$$

Необходимое условие экстремума описывается набором условий

$$\frac{\partial M(S_1, \ldots, S_n)}{\partial S_i} = -2\mathbf{K}[\eta, \xi_i] + 2\sum_{j=1}^{n} \mathbf{K}[\xi_i \xi_j] S_j = 0, \quad i = 1, 2, \ldots, n. \tag{15}$$

Если воспользоваться матричными обозначениями, то система линейных уравнений (15) может быть представлена в виде

$$\mathbf{K}[\bar{\xi}, \bar{\xi}]\bar{S} = \mathbf{K}[\eta, \bar{\xi}], \tag{16}$$

где $\mathbf{K}[\bar{\xi}, \bar{\xi}]$ — ковариационная матрица вектора $\bar{\xi}$, а $\mathbf{K}[\eta, \bar{\xi}]$ — матрица ковариаций η и $\bar{\xi}$.

Лемма. *Пусть компоненты вектора* $\bar{\xi}$ *линейно независимы. Тогда ковариационная матрица* $\mathbf{K}[\bar{\xi}, \bar{\xi}]$ *невырождена и вектор коэффициентов линейной регрессии дается соотношением*

$$\bar{\alpha} = \mathbf{K}^{-1}[\bar{\xi}, \bar{\xi}]\mathbf{K}[\eta, \bar{\xi}]. \tag{17}$$

◀ Пусть компоненты вектора $\bar{\xi}$ линейно зависимы, тогда существует ненулевой вектор $\bar{\lambda}^\circ = (\lambda_i^\circ)_{i=1}^n$ такой, что

$$\lambda_1^\circ \xi_1 + \ldots + \lambda_n^\circ \xi_n \equiv 0.$$

При этом

$$D\left[\sum_{i=1}^n \lambda_i^\circ \xi_i\right] = \sum_{i,j} \lambda_i^\circ \lambda_j^\circ \mathbf{K}[\xi_i, \xi_j] = \left(\bar{\lambda}^\circ\right)^T \mathbf{K}\left[\bar{\xi}, \bar{\xi}\right] \bar{\lambda}^\circ = \mathbf{0}. \tag{18}$$

Заметим, что функция $D(\lambda_1, \ldots, \lambda_n) = D\left[\sum \lambda_i \xi_i\right]$ неотрицательна и (в силу (18)) достигает в ненулевой точке $\bar{\lambda}^\circ$ своего наименьшего значения. Поэтому однородная система (необходимое условие экстремума!)

$$\frac{\partial D}{\partial \bar{\lambda}} = \mathbf{K}\left[\bar{\xi}, \bar{\xi}\right] \bar{\lambda} = \mathbf{0} \tag{19}$$

имеет ненулевое решение, а отсюда следует вырожденность ковариационной матрицы $\mathbf{K}\left[\bar{\xi}, \bar{\xi}\right]$.

Пусть теперь матрица $\mathbf{K}\left[\bar{\xi}, \bar{\xi}\right]$ вырождена. Тогда система (19) имеет ненулевое решение $\bar{\lambda}^\circ$. Умножая соотношение (19) на $\left(\bar{\lambda}^\circ\right)^T$, заключаем, что

$$(\bar{\lambda}^\circ)^T \mathbf{K}\left[\bar{\xi}, \bar{\xi}\right] \bar{\lambda}^\circ = \mathbf{0},$$

откуда (см. соотношение (18))

$$D\left(\sum \lambda_i^\circ \xi_i\right) = \mathbf{0}$$

и значит с вероятностью единица компоненты ξ_i линейно зависимы. ▶

Еще раз отметим, что независимо от того, линейна регрессия или нет, мы можем попытаться описать связь величины η с величинами $\{\xi_j\}_{j=1}^n$ линейным соотношением (14), коэффициенты которого даются равенством (17). Это будет наилучшее с рассматриваемой точки зрения линейное приближение к линии регрессии η на $\bar{\xi}$, совпадающее с ней в случае, когда регрессия линейна.

Пусть $\varphi\left(\bar{\xi}\right)$ — регрессия η на $\bar{\xi}$, $l\left(\bar{\xi}\right)$ — линейная функция (14) с коэффициентами (17). Рассмотрим корреляцию η и $\varphi\left(\bar{\xi}\right)$. Оказывается, функция регрессии $\varphi\left(\bar{\xi}\right)$ дает наилучшее описание случайной величины η не только с точки зрения минимума дисперсии остатка, но и имеет наибольший коэффициент корреляции с η.

Теорема. *Пусть $f\left(\bar{\xi}\right)$ — произвольная функция, $\varphi\left(\bar{\xi}\right)$ — регрессия η на $\bar{\xi}$. Тогда*

$$\left|\rho_{\eta\overline{\varphi}}\right| \geqslant \left|\rho_{\eta f}\right|$$

при этом $\rho^2_{\eta\overline{\varphi}} = \Theta^2_{\eta/\bar{\xi}}$ (см. соотношение (12)).

Рассмотрим корреляцию случайной величины η и линейной функции $l(\bar{\xi})$ (соотношение (14)) и заметим, что если регрессия $\varphi(\bar{\xi})$ — линейна, т. е. $\varphi(\bar{\xi}) \equiv l(\bar{\xi})$, то

$$\Theta^2_{\eta/\xi} = \rho^2_{\eta\varphi} = \rho^2_{\eta l} = 1 - \frac{D\zeta}{D\eta} = \frac{\mathbf{K}^T[\eta, \bar{\xi}] \cdot \mathbf{K}[\bar{\xi}, \bar{\xi}] \cdot \mathbf{K}[\eta, \bar{\xi}]}{D\eta}. \tag{20}$$

Если же регрессия $\varphi(\bar{\xi})$ линейной не является, то, как уже было отмечено выше, $\rho^2_{\eta\varphi} \geqslant \rho^2_{\eta l}$. Величина $\rho^2_{\eta l}$, определяемая правой частью соотношения (20), в этом случае описывает качество представления случайной величины η линейной функцией переменных $\bar{\xi}$ и называется ***множественным коэффициентом корреляции***.

Таким образом, множественный коэффициент корреляции $\rho^2_{\eta l}$ в случае линейной регрессии $(\varphi(\bar{\xi}) = l(\bar{\xi}))$ совпадает с корреляционным отношением $\Theta^2_{\eta/\bar{\xi}}$. Если же регрессия нелинейна, то он отличен от $\Theta^2_{\eta/\bar{\xi}}$, заключен между 0 и 1 и характеризует степень представимости η линейной комбинацией величин $\bar{\xi}$ — при $\rho^2_{\eta l} = 0$ корреляционная связь между η и линейными комбинациями $\bar{\xi}$ отсутствует, при $\rho^2_{\eta l} = 1$ — имеется жесткая функциональная связь между η и компонентами $\bar{\xi}$: с вероятностью 1 η является линейной комбинацией компонент $\bar{\xi}$.

1.3. Значимость множественного коэффициента корреляции

В практической ситуации, имея дело с выборочными значениями из закона распределения совокупности $(\eta, \bar{\xi})$, мы не имеем возможности определить точное значение множественного коэффициента корреляции $\rho^2_{\eta l}$, а можем лишь найти его выборочный аналог. При этом, как и в случае с парным коэффициентом корреляции, возникает задача установления значимости выборочного коэффициента множественной корреляции, для решения которой необходимо знание закона распределения последнего.

Пусть в эксперименте получена выборка из закона распределения совокупности $(\eta, \bar{\xi}) = (\eta, \xi_1, \xi_2, \ldots, \xi_n)$

$$(Y_1, \mathbb{X}_1), \quad (Y_2, \mathbb{X}_2), \quad \ldots, \quad (Y_N, \mathbb{X}_N). \tag{21}$$

Эмпирический аналог $r_{\eta l}$ множественного коэффициента корреляции $\rho_{\eta l}$ может быть построен на основе соотношения (20) заменой фигурирующих там дисперсий их выборочными аналогами, найденными по выборке (21)

$$r_{\eta l} = 1 - \frac{\dfrac{1}{N-n} \sum\limits_{j=1}^{N} \left(Y_j - \overline{Y} - \sum\limits_{i=1}^{n} a_i (X^i_j - \overline{X}_j)\right)^2}{\dfrac{1}{N} \sum\limits_{j=1}^{N} (Y_j - \overline{Y})^2},$$

где $\overline{Y} = \frac{1}{N} \sum\limits_{j=1}^{N} Y_j$ и a_i — оценки (17) коэффициентов регрессии (14), полученные методом наименьших квадратов.

В предположении совместной нормальности (η, ξ) установлено, что если множественный коэффициент корреляции равен нулю, то квадрат его эмпирического аналога имеет известное распределение. Этим можно воспользоваться для построения процедуры проверки значимости $r_{\eta l}$.

На практике для установления значимости отличия эмпирического коэффициента множественной корреляции от нуля пользуются тем, что величина

$$F_r = \frac{r^2(N-n-1)}{(1-r^2)n}$$

имеет распределение Фишера с $(n,\ N-n-1)$ степенями свободы при условии справедливости сделанных выше предположений о равенстве нулю $\rho_{\eta l}$ и совместной нормальности $(\eta, \bar{\xi})$.

По заданному уровню значимости $\varkappa$ определяют величину $F_\varkappa$ такую, что $P\{F_r < F_\varkappa\} > 1-\varkappa$. Если расчетное значение F_r оказывается меньше табличного $F_\varkappa$ — гипотеза о незначимом отличии $r_{\eta l}$ от нуля считается согласующейся с результатами эксперимента и $\rho_{\eta l}$ полагается равным нулю. В противном случае считается, что коэффициент $\rho_{\eta l}$ отличен от нуля и множественная регрессия (14) дает представление о характере изменения величины η с изменением $\bar{\xi}$.

§ 2. Случайные переменные. Нелинейные зависимости

Пусть $(\eta, \bar{\xi})$ — наблюдаемые в эксперименте случайные переменные. Предположим, что на основе рассмотрений предыдущих пунктов сделан вывод о *незначимом* отличии от нуля выборочного коэффициента корреляции (парного или множественного). Как уже неоднократно было отмечено, это дает основание для вывода об отсутствии *линейной* зависимости между η и $\bar{\xi}$, но не исключает возможного наличия зависимостей нелинейных (в том числе и жестких функциональных) между ними.

Для идентификации этих зависимостей изучим корреляционное отношение $\Theta^2_{\eta/\bar{\xi}}$, введенное соотношением (12). Оно обладает следующими свойствами:

1. Величина корреляционного отношения $\Theta^2_{\eta/\bar{\xi}}$ заключена между нулем и единицей. Если отношение $\Theta^2_{\eta/\bar{\xi}}$ близко к нулю — закономерное изменение переменной η в зависимости от $\bar{\xi}$ отсутствует. Если же $\Theta^2_{\eta/\bar{\xi}}$ близко к 1 — средние значения η с высокой степенью надежности могут быть найдены по известным значениям $\bar{\xi}$.

2. Корреляционное отношение $\Theta^2_{\eta/\bar{\xi}}$ не меньше квадрата коэффициента корреляции (теорема пункта 3.1.2)

$$\Theta^2_{\eta/\bar{\xi}} \geqslant \rho^2_{\eta l}.$$

Знак равенства достигается тогда и только тогда, когда между переменными η и $\bar{\xi}$ имеется корреляционная (т.е. хорошо описываемая линейными соотношениями) связь.

Как и при анализе линейных зависимостей, достаточно уметь строить эмпирический аналог $H^2_{\eta/\xi}$ величины $\Theta^2_{\eta/\bar{\xi}}$ по выборке (21) и оценивать значимость отличия этого выборочного аналога от нуля. Обе задачи оказываются технически более сложными, чем аналогичные для парного и множественного коэффициентов корреляции. Уже процедура построения эмпирического аналога $\Theta^2_{\eta/\bar{\xi}}$ предъявляет специальные требования к выборке (21) — в соответствии с (12) мы должны быть в состоянии вычислить эмпирический аналог (оценку) величины $D\varphi$, описывающей рассеяние η относительно линии регрессии $\varphi(\bar{\xi})$. А для этого необходимо, чтобы экспериментальные данные позволяли строить линию условных средних $\overline{Y}(\bar{\xi})$. Возможность построения линии условных средних может быть обеспечена, например, наличием в выборке (21) повторных измерений величины η: каждому из наблюденных значений $\bar{\xi} = \mathbb{X}_i$ отвечает несколько измерений η

$$(\mathbb{X}_1, Y_1^1), \ldots, (\mathbb{X}_1, Y_1^{r_1}), (\mathbb{X}_2, Y_2^1), \ldots, (\mathbb{X}_2, Y_2^{r_2}), \ldots, (\mathbb{X}_N, Y_N^1), \ldots, (\mathbb{X}_N, Y_N^{r_N}).$$

Либо выборочные данные должны допускать объединение наблюденных значений переменной $\bar{\xi} = \mathbb{X}_i$ в группы так, чтобы каждому групповому выборочному среднему отвечало несколько значений переменной η.

В любом из указанных случаев эмпирическое корреляционное отношение может быть подсчитано по формуле

$$H_{\eta/\xi}^2 = \frac{\sum\limits_{i=1}^{N} r_i\left(\overline{Y}_i - \overline{Y}\right)^2}{\sum\limits_{i,j=1}^{N,r_i} \left(Y_i^j - \overline{Y}\right)^2},$$

где $\overline{Y}_i$ — среднее наблюденных в i-й точке значений η (либо групповое среднее игреков),

$$\overline{Y}_i = \frac{1}{r_i}\sum_{j=1}^{r_i} Y_i^j,$$

$\overline{Y}$ — среднее всех наблюденных значений Y_i^j

$$\overline{Y} = \frac{1}{\sum r_i}\sum_{i,j=1}^{N,r_i} Y_i^j.$$

Для проверки гипотезы о значимом отличии от нуля величины $H_{\eta/\xi}^2$ воспользуемся тем фактом, что статистика F_H, задаваемая соотношением

$$F_H = \frac{H_{\eta/\xi}^2\left(\sum r_i - N\right)}{\left(1 - H_{\eta/\xi}^2\right)(N-1)},$$

имеет *приближенно* распределение Фишера с $\left(N-1, \sum r_i - N\right)$ степенями свободы при условии, что все сечения (групповые данные) нормальны с одинаковой дисперсией, и в предположении, что $\Theta_{\eta/\xi}^2 = 0$. В силу монотонности F_H как функции переменной $H_{\eta/\xi}^2$ это обстоятельство дает возможность обычным образом строить процедуру проверки интересующей нас гипотезы.

По заданному уровню значимости $\varkappa$ определяем $F_\varkappa$ так, что

$$P\left\{F_H \geqslant F_\varkappa\right\} \leqslant \varkappa.$$

При этом

$$P\left\{H_{\eta/\xi}^2 \leqslant \frac{F_\varkappa(N-1)}{\sum r_i - N + F_\varkappa(N-1)}\right\} \geqslant 1 - \varkappa,$$

и гипотеза о значимом отличии от нуля эмпирического корреляционного отношения признается согласующейся с опытными данными, если

$$H_{\eta/\xi}^2 \geqslant \frac{F_\varkappa(N-1)}{\sum r_i - N + F_\varkappa(N-1)}.$$

В этом случае можно считать, что переменные η и $\bar{\xi}$ связаны некоторой зависимостью, вообще говоря, нелинейной. Представление о ней может дать линия условных средних, построенная по экспериментальным точкам

$$(\mathbb{X}_1, \overline{Y}_1), \quad (\mathbb{X}_2, \overline{Y}_2), \quad \ldots, \quad (\mathbb{X}_N, \overline{Y}_N).$$

В противном случае корреляционное отношение признается равным нулю и, следовательно, можно считать, что закономерного изменения среднего значения переменной η в связи с изменением значений переменных $\bar{\xi}$ нет.

§ 3. Неслучайные переменные. Линейные по параметрам регрессионные модели

3.1. Основные допущения

Пусть теперь $x_1, x_2, \ldots, x_n$ — неслучайные переменные, связанные с неслучайной переменной y соотношением

$$y = f(x_1, x_2, \ldots, x_n;\ \alpha_1, \alpha_2, \ldots, \alpha_k), \tag{1}$$

где функция $f(\mathbf{x}, \bar{\alpha})$ известна с точностью до параметров. В эксперименте получены значения $(Y, \mathbb{X})$ переменных $(y, \mathbf{x})$. Предполагается, что модель измерений аддитивна относительно ошибки измерений ε и каждое измерение Y складывается из значения $f(\mathbb{X})$ и ошибки измерения ε

$$Y = f(\mathbb{X}; \bar{\alpha}) + \varepsilon. \tag{2}$$

При этом допускается проведение нескольких измерений в одной и той же точке $\mathbb{X}$, так что выборка измеренных значений имеет вид

$$Y_i^j = f(\mathbb{X}_i; \bar{\alpha}) + \varepsilon_i^j. \tag{3}$$

Здесь нижний индекс показывает точку, в которой проведено измерение ($i=1,2,\ldots,N$), верхний — номер измерения в i-й точке $\left(j = 1, 2, \ldots, r_i,\ \sum\limits_{i=1}^{N} r_i = m\right)$.

Требуется найти оценки неизвестных параметров α_s, $s = 1, 2, \ldots, k$, и высказать некоторое обоснованное суждение о качестве найденного описания зависимости величины y от переменных $x_1, \ldots, x_n$.

Последняя задача ниже будет уточнена, однако сразу заметим, что проблему подбора наилучшего *класса параметрических функций* $f(\mathbf{x}; \bar{\alpha})$, выбираемых для построения зависимости (1), мы здесь не обсуждаем. Предполагается, что этот класс определен из внестатистических соображений.

В дальнейшем ограничимся, чтобы избежать технических осложнений в выкладках, линейными по параметрам функциями $f(\mathbf{x}; \bar{\alpha}) = \alpha_1\varphi_1(\mathbf{x}) + \ldots + \alpha_k\varphi_k(\mathbf{x})$. Функции $\varphi_s(\mathbf{x})$, $s = 1, 2, \ldots, k$, будем предполагать линейно независимыми. Тогда модель измерений (3) запишется в виде

$$Y_i^j = \sum_{s=1}^{k} \alpha_s\varphi_s(\mathbb{X}_i) + \varepsilon_i^j. \tag{4}$$

Считая ошибки ε_i^j попарно некоррелированными случайными величинами (что соответствует предположению о независимости измерений) с нулевым математическим ожиданием (отсутствие систематических ошибок) и одинаковой дисперсией σ^2

(равноточность измерений), будем искать неизвестные параметры $\{\alpha_s\}_{s=1}^{k}$, исходя из принципа минимизации суммы квадратов ошибок[2)]

$$\bar{\alpha} = \operatorname{argmin} \sum_{i=1}^{N} \sum_{j=1}^{r_i} \left(Y_i^j - \sum_{s=1}^{k} \alpha_s \varphi_s \right)^2. \tag{5}$$

Заметим, что в силу предположения об отсутствии систематических ошибок измерения ($M\varepsilon_i^j = 0$) из соотношения (2) следует

$$MY = f(\mathbb{X};\, \bar{\alpha}). \tag{6}$$

В соответствии с принятой терминологией соотношение (6) позволяет назвать функцию $f(\mathbb{X};\, \bar{\alpha})$ функцией регрессии, параметры $\bar{\alpha}$ — параметрами регрессии; в частности, для линейной по параметрам модели (4), параметры $\bar{\alpha}$ будем называть коэффициентами регрессии.

3.2. Оценивание коэффициентов регрессии методом наименьших квадратов

Для удобства дальнейшего изложения и обозримости выкладок перейдем от скалярной формы записи задачи (5) к векторно-матричной, для чего введем следующие обозначения: $\mathbb{Y}$ — вектор измерений значений переменной y, $\bar{\varepsilon}$ — вектор ошибок измерений, $\bar{\varphi}(\mathbf{x})$ — вектор параметров. Все векторы — матрицы-столбцы,

$$\mathbb{Y} = \begin{pmatrix} Y_1^1 \\ \vdots \\ Y_1^{r_1} \\ Y_2^1 \\ \vdots \\ Y_2^{r_2} \\ \dots \\ Y_N^1 \\ \vdots \\ Y_N^{r_N} \end{pmatrix}; \quad \bar{\varepsilon} = \begin{pmatrix} \varepsilon_1^1 \\ \vdots \\ \varepsilon_1^{r_1} \\ \varepsilon_2^1 \\ \vdots \\ \varepsilon_2^{r_2} \\ \dots \\ \varepsilon_N^1 \\ \vdots \\ \varepsilon_N^{r_N} \end{pmatrix}; \quad \bar{\varphi}(\mathbf{x}) = \begin{pmatrix} \varphi_1(\mathbf{x}) \\ \varphi_2(\mathbf{x}) \\ \vdots \\ \varphi_k(\mathbf{x}) \end{pmatrix}; \quad \bar{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{pmatrix}.$$

Тогда исследуемая зависимость (1) представима в виде

$$y = \bar{\varphi}^T(\mathbf{x}) \cdot \bar{\alpha}, \tag{7}$$

где символом $\mathbf{A}^T$ обозначена матрица, транспонированная по отношению к $\mathbf{A}$ (в данном случае φ^T — матрица-строка).

Модель измерений (4) запишется так

$$\mathbb{Y} = \mathbf{F} \cdot \bar{\alpha} + \bar{\varepsilon}, \tag{8}$$

[2)] В предположении совместной нормальности ошибок ε_i^j метод наименьших квадратов (5) является естественным следствием метода максимального правдоподобия.

где $\mathbf{F}$ — матрица формата $m \times k$ $\left(m = \sum\limits_{i=1}^{N} r_i\right)$, имеющая следующую структуру

$$\mathbf{F} = \begin{pmatrix} \varphi_1(\mathbb{X}_1) & \dots & \varphi_k(\mathbb{X}_1) \\ \dots & \dots & \dots \\ \varphi_1(\mathbb{X}_1) & \dots & \varphi_k(\mathbb{X}_1) \\ \varphi_1(\mathbb{X}_2) & \dots & \varphi_k(\mathbb{X}_2) \\ \dots & \dots & \dots \\ \varphi_1(\mathbb{X}_2) & \dots & \varphi_k(\mathbb{X}_2) \\ \dots & \dots & \dots \\ \varphi_1(\mathbb{X}_N) & \dots & \varphi_k(\mathbb{X}_N) \\ \dots & \dots & \dots \\ \varphi_1(\mathbb{X}_N) & \dots & \varphi_k(\mathbb{X}_N) \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_1 \text{ — строк} \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_2 \text{ — строк} \\ \dots\dots\dots \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_N \text{ — строк} \end{matrix}$$

Метод наименьших квадратов (5) в матричной форме примет вид

$$\bar{\varepsilon}^T \cdot \bar{\varepsilon} = (\mathbb{Y} - \mathbf{F} \cdot \bar{\alpha})^T (\mathbb{Y} - \mathbf{F} \cdot \bar{\alpha}) \to \min, \tag{9}$$

а необходимое условие экстремума

$$\operatorname{grad} \bar{\varepsilon}^T \cdot \bar{\varepsilon} = 0$$

после несложных выкладок приводит к соотношению

$$\mathbf{F}^T \mathbf{F} \cdot \bar{\alpha} = \mathbf{F}^T \cdot \mathbb{Y}. \tag{10}$$

Отметим, что функция $\bar{\varepsilon}^T \bar{\varepsilon}$ как функция параметров $\{\alpha_s\}$ является неотрицательно определенной квадратичной функцией, а потому достигает своего наименьшего значения. Поэтому $\operatorname{argmin} \bar{\varepsilon}^T \bar{\varepsilon}$ содержится среди решений системы (10). Система линейных относительно $\bar{\alpha}$ уравнений (10) носит название *системы нормальных уравнений* метода наименьших квадратов. Если матрица $\mathbf{F}^T \mathbf{F}$ оказывается невырожденной, то (10) имеет единственное решение, задаваемое соотношением

$$\bar{\alpha} = \left(\mathbf{F}^T \mathbf{F}\right)^{-1} \mathbf{F}^T \mathbb{Y}, \tag{11}$$

которое определяет оценки метода наименьших квадратов коэффициентов регрессии.

Отметим, что линейная независимость функций $\varphi_j(\mathbf{x})$ не гарантирует невырожденности матрицы $\mathbf{F}^T \mathbf{F}$, в то время как их линейная зависимость приводит к вырожденности указанной матрицы. Связано это с тем, что вырожденность или невырожденность $\mathbf{F}^T \mathbf{F}$ определяется наличием или отсутствием линейной зависимости системы столбцов матрицы $\mathbf{F}$, которые представляют собой упорядоченные значения базисных функций $\varphi_j(\mathbf{x})$ в точках, где проводятся измерения, с учетом их кратности

$$\bar{f}_1 = \begin{pmatrix} \varphi_1(\mathbf{x}_1) \\ \dots \\ \varphi_1(\mathbf{x}_1) \\ \varphi_1(\mathbf{x}_2) \\ \dots \\ \varphi_1(\mathbf{x}_2) \\ \dots \\ \varphi_1(\mathbf{x}_N) \\ \dots \\ \varphi_1(\mathbf{x}_N) \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_1 \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_2 \\ \dots \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_N \end{matrix} \quad \bar{f}_2 = \begin{pmatrix} \varphi_2(\mathbf{x}_1) \\ \dots \\ \varphi_2(\mathbf{x}_1) \\ \varphi_2(\mathbf{x}_2) \\ \dots \\ \varphi_2(\mathbf{x}_2) \\ \dots \\ \varphi_2(\mathbf{x}_N) \\ \dots \\ \varphi_2(\mathbf{x}_N) \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_1 \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_2 \\ \dots \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_N \end{matrix} \quad \dots \quad \bar{f}_k = \begin{pmatrix} \varphi_1(\mathbf{x}_1) \\ \dots \\ \varphi_k(\mathbf{x}_1) \\ \varphi_k(\mathbf{x}_2) \\ \dots \\ \varphi_k(\mathbf{x}_2) \\ \dots \\ \varphi_k(\mathbf{x}_N) \\ \dots \\ \varphi_k(\mathbf{x}_N) \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_1 \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_2 \\ \dots \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_N. \end{matrix}$$

Матрица $\mathbf{F}^T\mathbf{F}$ является матрицей Грама векторов $\bar{f}_j$ — ее элементами являются скалярные произведения $(\bar{f}_i, \bar{f}_j)$. Она невырождена тогда и только тогда, когда векторы $\bar{f}_j$ линейно независимы.

Ясно, что для линейно зависимых функций $\varphi_j(\mathbf{x})$ векторы $\bar{f}_j$ линейно зависимы. Линейная независимость $\varphi_j(\mathbf{x})$, как показывает следующий простой пример

$$\varphi_1(x) = x; \quad \varphi_2(x) = x^2; \quad \varphi_3(x) = x^3,$$
$$x_1 = 0; \quad x_2 = 1; \quad r_1 = r_2 = r_3 = 2,$$
$$\bar{f}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}; \quad \bar{f}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}; \quad \bar{f}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},$$

линейной независимости векторов $\bar{f}_j$ не гарантирует.

Однако, даже если матрица $\mathbf{F}^T\mathbf{F}$ — вырождена, система нормальных уравнений (10) все равно разрешима, и ее решение (т. е. оценки коэффициентов регрессии) могут быть найдены с помощью обобщенного обращения матрицы $\mathbf{F}^T\mathbf{F}$.

3.3. Свойства оценок коэффициентов регрессии

Пусть ошибки измерений ε_i^j совместно нормальны с нулевым средним и ковариационной матрицей $\mathbf{K}[\bar{\varepsilon}, \bar{\varepsilon}] = \sigma^2\mathbf{I}$, где σ^2 — величина, характеризующая точность каждого отдельного измерения, $\mathbf{I}$ — единичная матрица порядка $m = \sum_{i=1}^{N} r_i$. Тогда статистические свойства оценок (11) коэффициентов регрессии описываются следующим утверждением.

Теорема. *Оценки* (11) *коэффициентов регрессии совместно нормальны с вектором средних* $M\bar{a} = \bar{\alpha}$ *и ковариационной матрицей* $\mathbf{K}[\bar{a}, \bar{a}] = \sigma^2(\mathbf{F}^T\mathbf{F})^{-1}$.

◀ В силу модели (8) из формулы (11) получаем, что

$$\bar{a} = (\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T\mathbb{Y} = (\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T(\mathbf{F}\,\bar{\alpha} + \bar{\varepsilon}) = \bar{\alpha} + (\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T\cdot\bar{\varepsilon},$$

откуда следует совместная нормальность компонент вектора оценок $\bar{a}$, так как они являются линейными комбинациями нормальных величин. Далее

$$M\bar{a} = \bar{\alpha} + (\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T\cdot M\bar{\varepsilon} = \bar{\alpha},$$
$$K[\bar{a},\bar{a}] = M[\bar{a} - M\bar{a}][\bar{a} - M\bar{a}]^T = M\left[(\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T\cdot\bar{\varepsilon}\right]\left[(\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T\cdot\bar{\varepsilon}\right]^T =$$
$$= (\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T M\left[\bar{\varepsilon}\cdot\bar{\varepsilon}^T\right]\mathbf{F}(\mathbf{F}^T\mathbf{F})^{-1} = \sigma^2(\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T\mathbf{F}(\mathbf{F}^T\mathbf{F})^{-1} = \sigma^2(\mathbf{F}^T\mathbf{F})^{-1}. \blacktriangleright$$

Заметим, что утверждения теоремы о несмещенности оценок (11) и о корреляционной матрице не зависит от предположения о нормальности ошибок. Точно так же не зависит от этого предположения некоторое свойство оптимальности оценок (4), описываемое теоремой Гаусса—Маркова.

Теорема (Гаусса—Маркова). *Среди всех линейных несмещенных оценок коэффициентов регрессии оценки* (11) *— наилучшие в том смысле, что обладают наименьшим рассеянием.*

3.4. Оценивание параметра σ^2

Как правило, параметр σ^2, характеризующий точность каждого отдельного измерения при построении регрессионных зависимостей не известен и подлежит оцениванию по тем же исходным данным, по которым была построена модель. Это можно сделать различными способами. Некоторые из них мы приведем ниже.

Если $\bar{a}$ — вектор оценок (11), то модель принимает вид

$$y = \varphi^T(\mathbf{x})\,\bar{a}, \tag{12}$$

при этом сумма квадратов ошибок, допущенных при замене экспериментальных данных Y_i^j, данными, рассчитанными по модельному соотношению (7), дается соотношением $(Y - \mathbf{F}\bar{a})^T(Y - \mathbf{F}\bar{a})$ — минимальным значением суммы квадратов отклонений модели от расчетных данных. Имеет место следующая теорема.

Теорема (об оценивании σ^2). *Величина*

$$s^2 = \frac{(\mathbb{Y} - \mathbf{F}\,\bar{a})^T(\mathbb{Y} - \mathbf{F}\,\bar{a})}{m - k} \tag{13}$$

является несмещенной оценкой неизвестной дисперсии σ^2.

◀ В силу

$$\mathbb{Y} = \mathbf{F}\,\bar{\alpha} + \bar{\varepsilon}$$

и

$$\bar{a} - \bar{\alpha} = \mathbf{F}(\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T\,\bar{\varepsilon}$$

числитель соотношения (13) может быть представлен в виде

$$(\mathbb{Y} - \mathbf{F}\,\bar{a})^T(\mathbb{Y} - \mathbf{F}\,\bar{a}) = \left(\bar{\varepsilon} - \mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T\,\bar{\varepsilon}\right)^T\left(\bar{\varepsilon} - \mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T\,\bar{\varepsilon}\right) =$$
$$= \bar{\varepsilon}^T\left(\mathbf{I} - \mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T\right)^T\left(\mathbf{I} - \mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T\right)\bar{\varepsilon},$$

где $\mathbf{I}$ — единичная матрица формата $m \times m$, $m = \sum\limits_{i=1}^{N} r_i$.

Далее

$$\left(\mathbf{I} - \mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T\right)^T = \mathbf{I} - \mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T$$

и

$$\left(\mathbf{I} - \mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T\right)^2 = \mathbf{I} - 2\mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T + \mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T\mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T =$$
$$= \mathbf{I} - \mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T,$$

поэтому

$$(\mathbb{Y} - \mathbf{F}\,\bar{a})^T(\mathbb{Y} - \mathbf{F}\,\bar{a}) = \bar{\varepsilon}^T\left(\mathbf{I} - \mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T\right)\bar{\varepsilon}.$$

Вычисляя математическое ожидание обеих частей последнего соотношения и учитывая некоррелированность компонент вектора ошибок $\bar{\varepsilon}$, получим, что

$$M(\mathbb{Y} - \mathbf{F}\,\bar{a})^T(\mathbb{Y} - \mathbf{F}\,\bar{a}) = \sigma^2\,\mathrm{Tr}\left(\mathbf{I} - \mathbf{F}\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T\right),$$

где $\mathrm{Tr}\,\mathbf{A}$ — след матрицы $\mathbf{A}$, обладающий следующими свойствами

$$\mathrm{Tr}(\mathbf{A} + \mathbf{B}) = \mathrm{Tr}\,\mathbf{A} + \mathrm{Tr}\,\mathbf{B}, \qquad \mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$$

последнее — (для согласованных относительно умножения матриц).

Отсюда немедленно следует, что

$$\operatorname{Tr}\left(\mathbf{I}-\mathbf{F}\left(\mathbf{F}^{T}\mathbf{F}\right)^{-1}\mathbf{F}^{T}\right)=\sum_{i=1}^{N} r_i-\operatorname{Tr}\left(\mathbf{F}(\mathbf{F}^{T}\mathbf{F})^{-1}\mathbf{F}^{T}\right)=$$

$$=\sum_{i=1}^{N} r_i-\operatorname{Tr}\left(\left(\mathbf{F}^{T}\mathbf{F}\right)^{-1}\mathbf{F}^{T}\mathbf{F}\right)=\sum_{i=1}^{N} r_i-k. \blacktriangleright$$

Если дополнительно вектор ошибок измерений считать нормальным с нулевым вектором средних и ковариационной матрицей $\sigma^2\,\mathbf{I}$, то, используя метод максимального правдоподобия, можно получить отличную от (13) оценку параметра σ^2

$$S^2=\frac{(\mathbb{Y}-\mathbf{F}\,\bar{a})^T(\mathbb{Y}-\mathbf{F}\,\bar{a})}{m}. \tag{14}$$

Оценка (14) свойством несмещенности обладать уже не будет, однако можно показать, что она несмещена асимптотически.

3.5. Адекватность модели

Пусть $\bar{a}$ — оценки коэффициентов регрессии (11), а зависимость переменной y от переменных $\bar{\mathbf{x}}$ принимается в виде

$$y=\bar{\varphi}^T(\bar{\mathbf{x}})\,\bar{a}. \tag{15}$$

Важный вопрос (о соответствии построенной нами модели (15) исследуемому процессу) — насколько удачно соотношение (15) позволяет прогнозировать значения переменной y по известным значениям переменных $\bar{\mathbf{x}}$?

Мы не задаемся целью установить, является ли найденная нами зависимость «истинной» или «правильной» — исследование этого вопроса лежит вне плоскости наших рассмотрений, а мы просто *хотим получить ответ на вопрос о том, насколько хорошо найденная нами зависимость заменяет реальный эксперимент.*

Уточним постановку задачи. Будем называть *остаточной ошибкой* модели величину $(\mathbb{Y}-\mathbf{F}\bar{a})^T(\mathbb{Y}-\mathbf{F}\bar{a})$, т. е. сумму квадратов отклонений наблюденных в эксперименте значений $\mathbb{Y}$ переменной y от прогнозируемых в этих же точках значений, полученных с помощью соотношения (15). Если через $\overline{\mathbb{Y}}$ обозначить вектор, составленный из средних значений измерений переменной y в точках $\bar{\mathbf{x}}_i$, компоненты которого

$$\overline{Y}_i=\frac{\sum\limits_{j=1}^{r_i} Y_i^j}{r_i}$$

повторяются столько раз, какова кратность (т. е. количество r_i) измерений в i-й точке, то остаточная ошибка может быть представлена в виде

$$(\mathbb{Y}-\mathbf{F}\bar{a})^T(\mathbb{Y}-\mathbf{F}\bar{a})=\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)^T\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)+\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)^T\left(\mathbb{Y}-\overline{\mathbb{Y}}\right). \tag{16}$$

◀ Положим $\mathbb{Y}-\mathbf{F}\bar{a}=\mathbb{Y}-\overline{\mathbb{Y}}+\overline{\mathbb{Y}}-\mathbf{F}\bar{a}$ и рассмотрим

$$\left(\mathbb{Y}-\mathbf{F}\bar{a}\right)^T\left(\mathbb{Y}-\mathbf{F}\bar{a}\right)=\left[\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)+(\mathbb{Y}-\overline{\mathbb{Y}}\right]^T\left[\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)+\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)\right]=$$

$$=\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)^T\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)+\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)^T\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)+\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)^T\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)+\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)^T\left(\mathbb{Y}-\overline{\mathbb{Y}}\right).$$

Легко убедиться в том, что два последних слагаемых равны нулю. Например,

$$\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)^T\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)=\sum_{i,s=1}^{N,r_i}\left(Y_i^s-\overline{Y_i}\right)\left(\overline{Y}_i-\sum_{l=1}^{k}a^l\varphi_l(\mathbb{X}_i)\right)=$$

$$=\sum_{i=1}^{N}\left(\overline{Y_i}-\sum_{l=1}^{k}a^l\varphi_l(\mathbb{X}_i)\right)\sum_{s=1}^{r_i}\left(Y_i^s-\overline{Y}_i\right)=0. \blacktriangleright$$

В полученном разложении остаточной ошибки модели первое слагаемое описывает отличие найденной нами зависимости от эмпирической регрессии, второе — рассеяние экспериментальных данных относительно эмпирической регрессии, т. е. ошибку эксперимента. Ясно, что модель тем лучше, чем меньше первое слагаемое и чем ближе остаточная ошибка модели ко второму слагаемому. Остаточная ошибка, описывает рассеяние экспериментальных данных относительно модели и не может быть меньше второго слагаемого — ошибки эксперимента. Чем меньше разница между остаточной ошибкой и рассеянием экспериментальных данных относительно эмпирической регрессии, тем лучше модель представляет наблюденные в эксперименте значения y.

Будем говорить что модель *адекватно* описывает результаты эксперимента, или просто, что модель *адекватна*, если в подавляющем большинстве случаев остаточная ошибка модели близка к ошибке эксперимента.

Статистические свойства слагаемых из соотношения (16) описываются следующей теоремой.

Теорема. *Пусть ошибки измерений совместно нормальны с нулевым вектором средних и ковариационной матрицей* $\mathbf{K}[\bar{\varepsilon},\bar{\varepsilon}]=\sigma^2\mathbf{I}$.

Тогда

— *статистики* $\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)^T\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)$ *и* $\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)^T\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)$ *независимы;*

— *статистика* $\dfrac{\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)^T\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)}{\sigma^2}$ *имеет распределение* χ^2 *с* $f=\sum\limits_{i=1}^{N}r_i-N$ *степенями свободы, при этом величина*

$$s^2_{\text{эксп}}=\frac{\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)^T\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)}{\sum\limits_{i=1}^{N}r_i-N} \tag{17}$$

является несмещенной оценкой параметра σ^2.

Если гипотеза об адекватности модели (15) *справедлива, то дополнительно статистика*

$$\frac{\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)^T\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)}{\sigma^2}$$

имеет χ^2*-распределение с* $N-k$ *степенями свободы.*

Сформулированная теорема позволяет установить, что величина $F_{ад}$, описываемая отношением

$$F_{ад}=\frac{\left(\overline{\mathbb{Y}}-\mathbf{F}\bar{a}\right)^T\left(\sum\limits_{i=1}^{N}r_i-N\right)}{\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)^T\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)(N-k)}, \tag{45}$$

в случае справедливости гипотезы об адекватности имеет распределение Фишера с $\left(N - k,\ \sum_{i=1}^{N} r_i - N\right)$ *степенями свободы.*

Задавая уровень значимости $\varkappa$, определяем величину $F_\varkappa$ так, что $P\{F_{a\partial} \geqslant F_\varkappa\} \leqslant \varkappa$. Если рассчитанная по результатам эксперимента величина (18) окажется меньше, чем $F_\varkappa$, то гипотезу об адекватности следует признать согласующейся с опытными данными, так как в этом случае с надежностью, не худшей чем $1 - \varkappa$, остаточная ошибка модели ненамного превышает ошибку эксперимента. В противном случае гипотезу об адекватности следует признать плохо согласующейся с опытными данными, а в постулируемую модель внести изменения.

3.6. Точность и надежность оценивания коэффициентов регрессии

Для адекватных моделей представляет интерес вопрос о качестве оценок (11) коэффициентов регрессии: как велик может быть диапазон их надежного (для заданной степени надежности) варьирования? Поскольку оценки коэффициентов регрессии являются, вообще говоря, зависимыми случайными величинами, то желательно уметь получать информацию не только об индивидуальной, но и о *совместной точности* их оценивания.

Мы рассмотрим ниже процедуры определения точности оценивания коэффициентов регрессии как в предположении, что параметр σ^2 известен точно, так и считая, что он оценен по результатам эксперимента.

1. Из результатов п. 3.2.3 следует, что вектор оценок $\bar{a}$ нормален с вектором средних $\bar{\alpha}$ и ковариационной матрицей $\mathbf{K}[\bar{a}, \bar{a}] = \sigma^2\left(\mathbf{F}^T\mathbf{F}\right)^{-1}$. При точно известной величине σ^2 этой информации достаточно для построения доверительных интервалов для каждого из коэффициентов регрессии в отдельности и для совместного определения точности их оценивания.

Индивидуальные доверительные интервалы

Каждая из компонент a_i вектора $\bar{a}$ нормальна со средним α_i и дисперсией $\sigma_i^2 = \sigma^2\left(\mathbf{F}^T\mathbf{F}\right)_{ii}^{-1}$, где $\left(\mathbf{F}^T\mathbf{F}\right)_{ii}^{-1}$ — i-й диагональный элемент матрицы $\left(\mathbf{F}^T\mathbf{F}\right)^{-1}$. Задавая уровень доверия $\varkappa$, легко находим (см. п. 1.3.1) доверительные границы для коэффициентов регрессии α_i

$$a_i - \delta_i \leqslant \alpha_i \leqslant a_i + \delta_i;$$

величины δ_i определяются из условия $P\{|a_i - \alpha_i| < \delta_i\} = \varkappa$ и задаются равенством

$$\delta_i = x_\varkappa \sigma \sqrt{\left(\mathbf{F}^T\mathbf{F}\right)_{ii}^{-1}}, \tag{19}$$

где $x_\varkappa$ — решение уравнения $F(x) = (1 + \varkappa)/2$, $F(x)$ — функция стандартного нормального с параметрами (0,1) распределения.

Совместные доверительные границы

Вообще говоря (если матрица $\left(\mathbf{F}^T\mathbf{F}\right)^{-1}$ не является диагональной), оценки коэффициентов регрессии зависимы.

Назовем *главным эллипсоидом рассеяния* случайного вектора $\bar{a}$ относительно $\bar{\alpha}$ эллипсоид E_ε, задаваемый неравенством

$$E_\varepsilon = \left\{\bar{a} \in \mathbb{R}^k,\ (\bar{a} - \bar{\alpha})^T\left(\mathbf{F}^T\mathbf{F}\right)(\bar{a} - \bar{\alpha}) \leqslant \varepsilon^2\right\}. \tag{20}$$

Величина ε называется *радиусом* эллипсоида. Заметим, что уравнение

$$P\{\bar{a} \in E_\varepsilon\} = \varkappa \tag{21}$$

при заданной $\varkappa$ однозначно разрешимо относительно ε.

◀ В силу совместной нормальности компонент вектора $\bar{a}$, уравнение (21) имеет вид

$$\frac{\sqrt{|(\mathbf{F}^T\mathbf{F})|}}{(2\pi)^{\frac{k}{2}}}\int\limits_{E_\varepsilon}\exp\left\{-\frac{1}{2}(\bar{a}-\bar{\alpha})^T(\mathbf{F}^T\mathbf{F})(\bar{a}-\bar{\alpha})\right\}d\bar{a}=\varkappa.$$

Сделаем в интеграле замену переменных, приводящую квадратичную форму, стоящую в показателе степени у экспоненты, к сумме квадратов[3]. Последнее соотношение примет вид

$$\frac{1}{(2\pi)^{k/2}}\int\limits_{\sum\limits_{i=1}^{k}y_i^2\leqslant\varepsilon^2}\exp\left(-\frac{1}{2}\sum_{i=1}^{k}y_i^2\right)dy_1\,dy_2\ldots dy_k=\varkappa. \tag{22}$$

Переходя к сферическим координатам в $\mathbb{R}^k$ и полагая

$$\begin{aligned}
y_1&=r\sin\varphi_{k-1}\sin\varphi_{k-2}\ldots\sin\varphi_2\sin\varphi_1,\\
y_2&=r\sin\varphi_{k-1}\sin\varphi_{k-2}\ldots\sin\varphi_2\cos\varphi_1,\\
y_3&=r\sin\varphi_{k-1}\sin\varphi_{k-2}\ldots\cos\varphi_2,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
y_{k-1}&=r\sin\varphi_{k-1}\cos\varphi_{k-2},\\
y_k&=r\cos\varphi_{k-1},
\end{aligned}$$

где $0\leqslant r\leqslant\varepsilon$, $0\leqslant\varphi_1\leqslant 2\pi$, $0\leqslant\varphi_j\leqslant\pi$, $j=2,3,\ldots,k$, из соотношения (22) получим

$$\frac{1}{(2\pi)^{k/2}}\int\limits_0^{2\pi}d\varphi_1\int\limits_0^{\pi}\sin\varphi_2\,d\varphi_2\ldots\int\limits_0^{\pi}\sin^{k-2}\varphi_{k-1}\,d\varphi_{k-1}\int\limits_0^{\varepsilon}r^{k-1}\exp\left(-\frac{r^2}{2}\right)dr=\varkappa.$$

Учитывая, что

$$\int\limits_0^{\pi}\sin^{s-1}\varphi_s\,d\varphi_s=\frac{\Gamma\left(\frac{s}{2}\right)}{\Gamma\left(\frac{s+1}{2}\right)}\sqrt{\pi},$$

где $\Gamma(s)$ — гамма-функция Эйлера, для определения радиуса главного эллипсоида рассеяния ε получаем соотношение

$$2\frac{\pi^{k/2}}{\Gamma\left(\frac{k}{2}\right)}\int\limits_0^{\varepsilon}r^{k-1}\exp\left(-\frac{r^2}{2}\right)dr=\varkappa, \tag{23}$$

которое доказывает утверждение. ▶

Если задать доверительную вероятность $\varkappa$ ($\varkappa$ близка к единице) и из уравнения (23) найти соответствующий ей радиус ε, то совместное рассеяние оценок $\bar{a}$ коэффициентов регрессии будет описываться главным эллипсоидом E_ε.

[3] Это всегда возможно в силу симметричности матрицы $\mathbf{F}^T\mathbf{F}$.

2. Если величина параметра σ^2 неизвестна, то, заменив в полученных выше соотношениях для индивидуальных или совместных доверительных областей значение параметра σ^2 какой-нибудь его оценкой, мы получим *приближенные* индивидуальные или совместные доверительные области.

Построение *точных* доверительных областей в этом случае может быть осуществлено следующим образом.

Индивидуальные доверительные границы

Рассмотрим величины t^i_{m-k}, $i = 1, 2, \ldots, k$, определяемые отношением

$$t^i_{m-k} = \frac{a_i - \alpha_i}{s\sqrt{\left(\mathbf{F}^T\mathbf{F}\right)^{-1}_{ii}}},$$

где s — оценка параметра σ^2, даваемая соотношением (13). Так же, как и в п. 1.3.1, можно установить, что *каждая из этих величин имеет распределение Стьюдента с* $\left(\sum\limits_{i=1}^{N} r_i - k\right)$ *степенями свободы.* Отсюда для заданной доверительной вероятности $\varkappa$ обычным образом получаем, что точность оценивания i-го коэффициента регрессии α_i описывается двойным неравенством

$$a_i - s\Delta_\varkappa\sqrt{\left(\mathbf{F}^T\mathbf{F}\right)^{-1}_{ii}} \leqslant \alpha_i \leqslant a_i + s\Delta_\varkappa\sqrt{\left(\mathbf{F}^T\mathbf{F}\right)^{-1}_{ii}}. \tag{24}$$

Величина $\Delta_\varkappa$ находится по заданному значению $\varkappa$ из условия

$$P\left\{\left|t^i_{m-k}\right| < \Delta_\varkappa\right\} = \varkappa.$$

Совместные доверительные границы

Индивидуальные доверительные интервалы (24) не дают исчерпывающей информации *о совместной оценке точности* определения коэффициентов регрессии из-за возможной зависимости последних.

Для получения доверительной области в этом случае заметим, что в силу совместной нормальности компонент a_i вектора оценок $\bar{a}$ величина

$$\frac{(\bar{a} - \bar{\alpha})^T\mathbf{F}^T\mathbf{F}(\bar{a} - \bar{\alpha})}{\sigma^2} \tag{25}$$

имеет распределение χ^2 с k степенями свободы. По теореме из предыдущего пункта величина

$$\frac{(\mathbb{Y} - \mathbf{F}\bar{a})^T(\mathbb{Y} - \mathbf{F}\bar{a})}{\sigma^2}$$

имеет распределение χ^2 с $\left(\sum\limits_{i=1}^{N} r_i - k\right)$ степенями свободы и не зависит от величины, даваемой формулой (25).

Следовательно их отношение

$$F_{k,M-k} = \frac{(\bar{a} - \bar{\alpha})^T\mathbf{F}^T\mathbf{F}(\bar{a} - \bar{\alpha})\left(\sum\limits_{i=1}^{N} r_i - k\right)}{(\mathbb{Y} - \mathbf{F}\bar{a})^T(\mathbb{Y} - \mathbf{F}\bar{a})k} \tag{26}$$

имеет распределение Фишера с $\left(k, \sum\limits_{i=1}^{N} r_i - k\right)$ степенями свободы. Зададим доверительную вероятность $\varkappa$ и определим величину $F_\varkappa$ так, что

$$P\left\{F_{k,M-k} < F_\varkappa\right\} \geqslant \varkappa. \tag{27}$$

Перепишем неравенство (27) с учетом (26)

$$P\{F_{k,M-k} < F_\varkappa\} = P\{(\bar a - \bar\alpha)^T \mathbf{F}^T \mathbf{F}(\bar a - \bar\alpha) < s^2 k F_\varkappa\} = P\{\bar a \in E_\varepsilon\} \geqslant \varkappa$$

(здесь s^2 — оценка (13) параметра σ^2). Искомой совместной доверительной областью для коэффициентов регрессии $\bar\alpha$ будет главный эллипсоид рассеяния E_ε радиуса $\varepsilon = s\sqrt{kF_\varkappa}$.

3.7. Прогнозирование результатов эксперимента. Точность и надежность прогноза

В предыдущих разделах мы показали, как по результатам эксперимента можно найти зависимость (1) между переменными y и $\mathbf{x} = (x_i)_{i=1}^n$, и описали эту зависимость соотношением (12)

$$y = \varphi^T(\mathbf{x})\,\bar a = \sum_{i=1}^{k} a_i\varphi_i(x_1, x_2, \ldots, x_n),$$

где $\varphi_i(\mathbf{x})$ — известные функции, а оценки a_i даются формулой (11)

$$\bar a = \left(\mathbf{F}^T\mathbf{F}\right)^{-1}\mathbf{F}^T\mathbb{Y}.$$

Эти формулы могут быть использованы теперь для прогнозирования результатов эксперимента. При этом, конечно, предполагается, что новые измерения подчиняются тем же закономерностям и структура взаимодействия переменных в новых экспериментах такая же, какой она была в экспериментах, послуживших источником информации для построения модели.

Пусть $\mathbf{x}_0$ — точка, в которой мы хотим спрогнозировать значение переменной $y(\mathbf{x}_0)$. Отметим, что $\widehat{y}$ — предсказываемое моделью (12) значение переменной y в точке $\mathbf{x}_0$,

$$\widehat{y}(\mathbf{x}_0) = \varphi^T(\mathbf{x}_0)\,\bar a,$$

является несмещенной оценкой математического ожидания прогнозируемого измерения $Y(\mathbf{x}_0)$. Дисперсия этой оценки легко находится

$$D[\widehat{y}(\mathbf{x}_0)] = M\left[\varphi^T(\mathbf{x}_0)\,\bar a\,\bar a^T\varphi(\mathbf{x}_0)\right] = \varphi^T(\mathbf{x}_0)M\left[\bar a\,\bar a^T\right]\varphi(\mathbf{x}_0),$$

откуда, учитывая равенство $\mathbf{K}[\bar a, \bar a]\sigma^2\left(\mathbf{F}^T\mathbf{F}\right)^{-1}$, заключаем, что

$$D\left[\widehat{y}(\mathbf{x}_0)\right] = \sigma^2\varphi^T\left(\mathbf{x}_0\right)\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\varphi(\mathbf{x}_0).$$

В соответствии с принятыми допущениями прогнозируемое измерение $Y(\mathbf{x}_0)$ складывается из значения $y(\mathbf{x}_0)$ и ошибки измерения ε

$$Y(\mathbf{x}_0) = y(\mathbf{x}_0) + \varepsilon.$$

Разница между прогнозируемым значением и предсказанным с помощью модели (12) является нормальной случайной величиной с нулевым средним

$$M[\Delta Y] = M[Y(\mathbf{x}_0) - \widehat{y}(\mathbf{x}_0)] = y(\mathbf{x}_0) - M[\widehat{y}(\mathbf{x}_0)] = 0$$

и дисперсией

$$D[\Delta Y] = D[Y(\mathbf{x}_0)] + D[\widehat{y}(\mathbf{x}_0)] = \sigma^2\left(1 + \varphi^T(\mathbf{x}_0)\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\varphi(\mathbf{x}_0)\right).$$

Отсюда, как и выше, заключаем, что отношение

$$t_{\Delta Y} = \frac{Y(\mathbf{x}_0) - \widehat{y}(\mathbf{x}_0)}{s\sqrt{1 + \varphi^T(\mathbf{x}_0)\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\varphi(\mathbf{x}_0)}}$$

имеет распределение Стьюдента с $\sum\limits_{i=1}^{N} r_i - k$ степенями свободы. Это позволяет нам оценивать точность прогноза стандартным образом — задаем надежность $\varkappa$, близкую к единице, и находим значение $t_\varkappa$ такое, что $P\{|t_{\Delta Y}| < t_\varkappa\} = \varkappa$. С вероятностью, не меньшей $\varkappa$, при этом выполняется соотношение

$$\varphi^T(\mathbf{x}_0)\,\bar{a} - t_\varkappa \widehat{D} \leqslant Y(\mathbf{x}_0) \leqslant \varphi^T(\mathbf{x}_0)\,\bar{a} + t_\varkappa \widehat{D},$$

указывающее границы, в которых с надежностью, не худшей $\varkappa$, находится прогнозируемое значение $y(\mathbf{x}_0)$ переменной y.

Здесь $\widehat{D} = s\sqrt{1 + \varphi^T(\mathbf{x}_0)\left(\mathbf{F}^T\mathbf{F}\right)^{-1}\varphi(\mathbf{x}_0)}$ — оценка дисперсии $D[\Delta Y]$.

Заключение

Завершая обсуждение простейших задач статистического анализа экспериментальных данных отметим, что рассмотренные нами проблемы допускают широкое обобщение, как с точки зрения постановок, так и с точки зрения используемых методов исследования. Одним из важнейших обстоятельств, определяющих успех в постановке и решении статистических задач, является четкое осознание исследователем того, каким фактическим исходным материалом он обладает, какие цели он перед собой ставит и чего в конечном итоге хочет добиться. Постановка задачи определяет, как правило, не только аппарат, необходимый для ее решения, но, зачастую, и способы получения экспериментального материала.

Не менее существенной является и интерпретация результатов применения тех или иных статистических процедур.

Только компетентность исследователя и корректность статистика являются гарантией содержательной и безошибочной интерпретации. Сами по себе статистические процедуры не решают реальных прикладных проблем, однако правильно понятые и объясненные результаты их применения являются надежным ориентиром для прикладника.

Примеры решения задач

1. *Случайные величины ξ_1, ξ_2 и ξ_3 независимы и нормальны с одинаковым средним, равным 2. Известно, что их дисперсии относятся соответственно как 3 : 4 : 2. Найти ковариационную матрицу этих случайных величин, если известно, что*

$$P\{2\xi_1 - \xi_2 + 3\xi_3 > 5\} = 0{,}95.$$

Решение. Поскольку случайные величины ξ_1, ξ_2 и ξ_3 — независимы, то они некоррелированы. Следовательно, искомая ковариационная матрица диагональна и ее диагональные элементы — дисперсии случайных величин ξ_1, ξ_2 и ξ_3 соответственно:

$$\mathbf{K}[\xi_1, \xi_2, \xi_3] = \begin{pmatrix} D_{\xi_1} & 0 & 0 \\ 0 & D_{\xi_2} & 0 \\ 0 & 0 & D_{\xi_3} \end{pmatrix}.$$

Далее, из независимости и нормальности заключаем, что любая линейная комбинация этих случайных величин — нормальная случайная величина. В частности

$$2\xi_1 - \xi_2 + 3\xi_3 = N[m, D],$$

где

$$m = M(2\xi_1 - \xi_2 + 3\xi_3) = 2M\xi_1 - M\xi_2 + 3M\xi_3 = 8$$

и

$$D = D(2\xi_1 - \xi_2 + 3\xi_3) = 4D\xi_1 + D\xi_2 + 9D\xi_3.$$

Пусть дисперсия случайной величины ξ_1 равна σ^2. Тогда из условия задачи получим, что $D\xi_2 = 4\sigma^2/3$, $D\xi_3 = 2\sigma^2/3$, откуда $D = 34\sigma^2/3$.

Для нахождения величины σ^2 используем последнее условие задачи

$$P\{2\xi_1 - \xi_2 + 3\xi_3 > 5\} = 1 - F\left(\frac{5-8}{\sigma\sqrt{34/3}}\right) = 0{,}95.$$

Отсюда

$$F\left(\frac{5-8}{\sigma\sqrt{34/3}}\right) = 0{,}05 \quad \Rightarrow \quad \frac{5-8}{\sigma\sqrt{34/3}} = -1{,}645 \quad \Rightarrow \quad \sigma \approx 0{,}542.$$

Искомая ковариационная матрица

$$\mathbf{K}[\xi_1, \xi_2, \xi_3] = \begin{pmatrix} 0{,}293 & 0 & 0 \\ 0 & 0{,}391 & 0 \\ 0 & 0 & 0{,}195 \end{pmatrix}.$$

2. *Пара случайных величин ξ и η имеет совместное нормальное распределение с вектором математических ожиданий $\{-2, -1\}$ и ковариационной матрицей K*

$$\mathbf{K}[\xi, \eta] = \sigma^2 \begin{bmatrix} 2 & 3 \\ 3 & 7 \end{bmatrix}.$$

Известно, что $P\{\xi - 2\eta < 3\} = 0{,}65$. Найти $D\xi$, $D\eta$.

Решение. Совместная нормальность пары случайных величин ξ и η обеспечивает нормальность каждой из них и любой их линейной комбинации, в частности величина $\zeta = \xi - 2\eta$ нормальна с параметрами

$$M\zeta = M\xi - 2M\eta = -2 - 2(-1) = 0, \quad D\zeta = D\xi + 4D\eta - 4\operatorname{cov}(\xi, \eta).$$

Подставляя в последнее соотношение элементы ковариационной матрицы:

$$D\xi = 2\sigma^2, \quad D\eta = 7\sigma^2, \quad \operatorname{cov}(\xi, \eta) = 3\sigma^2,$$

получим

$$D\zeta = 2\sigma^2 + 4\cdot 7\sigma^2 - 4\cdot 3\sigma^2 = 18\sigma^2.$$

По условию $P\{\zeta < 3\} = 0{,}65$, откуда, используя нормальность ζ,

$$F\left(\frac{3}{\sigma\sqrt{18}}\right) = 0{,}65 \quad \Rightarrow \quad \frac{3}{\sigma\sqrt{18}} = 0{,}385 \quad \Rightarrow \quad \sigma \approx 1{,}837.$$

Искомые дисперсии равны, соответственно,

$$D\xi = 2\sigma^2 \approx 6{,}747, \quad D\eta = 7\sigma^2 \approx 23{,}622.$$

3. *Найти ковариацию ординаты и абсциссы точки $M(\xi, \eta)$, равномерно распределенной в квадрате K с вершинами $A(-1, 0)$, $B(0, 1)$, $C(1, 0)$ и $D(0, -1)$. Зависимы ли эти случайные величины?*

Решение. Пара (ξ, η) равномерно распределена в квадрате, значит, ее плотность задается соотношением

$$f_{\xi\eta}(x, y) = \begin{cases} \dfrac{1}{2}, & (x, y) \in K, \\ 0, & \text{в противном случае.} \end{cases}$$

Для ковариации получаем

$$\operatorname{cov}(\xi, \eta) = \iint\limits_K f_{\xi\eta}(x, y)(x - M\xi)(y - M\eta)\, dx\, dy.$$

Поскольку

$$M\xi = M\eta = \iint\limits_K f_{\xi\eta}(x, y)(x - M\xi)\, dx = \iint\limits_K f_{\xi\eta}(x, y)(y - M\eta)\, dy = 0,$$

постольку

$$\operatorname{cov}(\xi, \eta) = \iint\limits_K f_{\xi\eta}(x, y) xy\, dx\, dy = \iint\limits_K \frac{1}{2} xy\, dx\, dy = 0.$$

Тем не менее случайные величины (ξ, η) зависимы, так как изменение значения одной из них вызывает изменение диапазона значений другой[4].

4. *Двумерная случайная величина (ξ, η) имеет вектор математических ожиданий $\{0, -1\}$ и ковариационную матрицу*

$$\mathbf{K}[\xi, \eta] = \begin{pmatrix} 2 & -4 \\ -4 & 8 \end{pmatrix}.$$

Достаточно ли этих данных, чтобы спрогнозировать значения компоненты η при известных значениях компоненты ξ?

Решение. Заметим, что

$$\operatorname{cov}(\xi, \eta) = \rho\cdot\sigma_\xi\sigma_\eta \quad \Rightarrow \quad -4 = \rho\cdot\sqrt{2}\cdot\sqrt{8}.$$

[4] Впрочем, можно чисто формально установить интегрированием, что закон распределения одной из компонент зависит от значения, принятого другой компонентой.

Отсюда $\rho = -1$, и это значит, что между случайными величинами ξ и η имеется линейная функциональная зависимость, описываемая соотношением

$$\eta - M\eta = \rho\frac{\sigma_\eta}{\sigma_\xi}(\xi - M\xi) \quad \Rightarrow \quad \eta + 1 = -\frac{\sqrt{8}}{\sqrt{2}}\xi = -\frac{\xi}{2}.$$

Следовательно, при известных значениях компоненты ξ значения компоненты η могут быть спрогнозированы с вероятностью 1.

5. *Случайные величины ξ, η и ζ связаны соотношением*

$$2\xi - 3\eta + \zeta = 0.$$

Известно, что $M\xi = 0$, $M\eta = 0$, $M\zeta = 0$, $D\xi = 2$, $D\eta = 1$, $D\zeta = 5$. Найти ковариационную матрицу этих случайных величин.

Решение. Умножая данное в условии линейное соотношение последовательно на ξ, η и ζ и находя математические ожидания от обеих частей получающихся равенств, получим

$$\begin{cases} 2\xi^2 - 3\eta\xi + \zeta\xi = 0 \\ 2\xi\eta - 3\eta^2 + \zeta\eta = 0 \\ 2\xi\zeta - 3\eta\zeta + \zeta^2 = 0 \end{cases} \quad \Rightarrow \quad \begin{cases} 2M\xi^2 - 3M\eta\xi + M\zeta\xi = 0, \\ 2M\xi\eta - 3M\eta^2 + M\zeta\eta = 0, \\ 2M\xi\zeta - 3M\eta\zeta + M\zeta^2 = 0. \end{cases}$$

В силу равенства нулю математических ожиданий случайных величин, математические ожидания квадратов будут равны дисперсиям, а математические ожидания произведений — ковариациям. Для элементов ковариационной матрицы приходим к системе

$$\begin{cases} 2D_\xi - 3\operatorname{cov}(\eta\xi) + \operatorname{cov}(\zeta\xi) = 0, \\ 2\operatorname{cov}(\xi\eta) - 3D_\eta + \operatorname{cov}(\zeta\eta) = 0, \\ 2\operatorname{cov}(\xi\zeta) - 3\operatorname{cov}(\eta\zeta) + D_\zeta = 0, \end{cases}$$

или, подставляя известные дисперсии рассматриваемых случайных величин

$$\begin{cases} 4 = \ 3\operatorname{cov}(\eta\xi) - \operatorname{cov}(\zeta\xi), \\ 3 = \ 2\operatorname{cov}(\xi\eta) + \operatorname{cov}(\zeta\eta), \\ 5 = -2\operatorname{cov}(\xi\zeta) + 3\operatorname{cov}(\eta\zeta), \end{cases}$$

Решая эту систему, получим искомую ковариационную матрицу

$$\mathbf{K}[\xi, \eta, \zeta] = \begin{pmatrix} 2 & \frac{8}{9} & -\frac{4}{3} \\ \frac{8}{9} & 1 & \frac{11}{9} \\ -\frac{4}{3} & \frac{11}{9} & 5 \end{pmatrix}.$$

ВВЕДЕНИЕ В ТЕОРИЮ ИГР

В практической деятельности весьма часто приходится рассматривать явления и ситуации, в которых участвуют две или более стороны, имеющие различные интересы и обладающие возможностями применять для достижения своих целей разнообразные действия. Подобные явления и ситуации принято называть *конфликтными*, или просто *конфликтами*.

Типичный конфликт характеризуется тремя основными составляющими:

1) заинтересованными сторонами,
2) возможными действиями этих сторон,
3) интересами сторон.

Конфликтная ситуация, взятая из реальной жизни, как правило, довольно сложна. К тому же ее изучение затруднено наличием многих разных обстоятельств, часть из которых не оказывает сколь-либо существенного влияния ни на развитие конфликта, ни на его исход. Поэтому для того, чтобы анализ конфликтной ситуации оказался возможным, необходимо отвлечение от этих второстепенных факторов, при удачном стечении обстоятельств позволяющее построить упрощенную формализованную модель конфликта, которую и принято называть *игрой*. От реальной конфликтной ситуации игра отличается еще и тем, что ведется по вполне определенным правилам.

Необходимость изучения и анализа конфликтов, представляемых в виде упрощенных математических моделей (игр), вызвала к жизни специальный математический аппарат — *теорию игр*.

Опишем некоторые основные понятия, используемые в этой теории.

Заинтересованные стороны называются *игроками*. Любое возможное для игрока действие (в рамках заданных *правил игры*) называется его *стратегией*. В условиях конфликта каждый игрок выбирает свою стратегию, в результате чего складывается набор стратегий, называемый *ситуацией*. Заинтересованность игроков в ситуации проявляется в том, что каждому игроку в каждой ситуации приписывается число, выражающее степень удовлетворения его интересов в этой ситуации и называемое его *выигрышем* в ней.

В этих условиях протекание конфликта состоит в выборе каждым игроком своей стратегии и в получении им в сложившейся ситуации выигрыша из некоторого источника. На этом пути создается *теория игр с выигрышами*.

Однако оценка игроком ситуации путем указания его выигрыша, вообще говоря, не всегда возможна практически и даже не всегда имеет смысл. В подобных случаях иногда удается вместо прямых численных оценок ситуаций указывать на их сравнительную предпочтительность для отдельных игроков. На этом пути создается *теория игр с предпочтениями*, включающая в себя как частный случай и теорию игр с выигрышами. В дальнейшем мы ограничимся рассмотрением только игр с выигрышами.

Изучение игр можно проводить с различных точек зрения. Мы будем стремиться к

— выработке принципов оптимальности, то есть того, какое поведение игроков следует считать оптимальным (разумным, целесообразным),

— выяснению реализуемости этих принципов, то есть установлению существования оптимальных в выработанном смысле ситуаций,

и

— отысканию этих реализаций.

Одной из плодотворных форм реализации представлений об оптимальности можно считать понятие *равновесия*, при котором складывается такая (равновесная) ситуация, в нарушении которой не заинтересован ни один из игроков.

Именно ситуации равновесия могут быть предметом устойчивых договоров между игроками (ни у одного из игроков не будет мотивов к нарушению договора). Кроме того, ситуации равновесия являются выгодными для каждого игрока: в равновесной ситуации каждый игрок получает наибольший выигрыш (разумеется, в той мере, в какой это от него зависит).

Если в игре ситуации равновесия (в пределах отпущенных возможностей) нет, то, оставаясь в условиях стратегий, имеющихся у игроков, мы сталкиваемся с неразрешимой задачей. При возникновении подобных случаев естественно ставить вопрос о таком расширении первоначального понятия стратегии, чтобы среди ситуаций, составленных из новых, обобщенных стратегий, заведомо нашлись бы равновесные. Если такие обобщенные стратегии существуют, то обычно они представляются некоторыми комбинациями исходных стратегий (при этом, естественно, предполагается, что игра повторяется многократно). Для того, чтобы отличать прежние стратегии от новых, первые называют *чистыми*, а вторые — *смешанными* стратегиями.

Весьма плодотворным является представление смешанной стратегии как случайного выбора игроками их чистых стратегий, при котором случайные выборы различных игроков независимы в совокупности, а выигрыш каждого из них определяется как математическое ожидание случайного выигрыша. Игра, преобразованная таким образом, обычно называется *смешанным расширением* исходной игры.

Проиллюстрируем сказанное на примере одного из самых простых, но одновременно и наиболее изученных классов игр, на так называемых *матричных играх*. Исследование матричных игр интересно еще и потому, что к ним могут быть приближенно сведены многие игры более общего вида.

Затем мы кратко остановимся на вопросе классификации игр и рассмотрим еще два вида игр — *позиционные игры* и *биматричные игры*.

Глава XLVI

МАТРИЧНЫЕ ИГРЫ

Рассмотрим игру, в которой участвуют два игрока, причем каждый из игроков имеет конечное число стратегий. Обозначим для удобства одного из игроков через A, в другого — через B.

Предположим, что игрок A имеет m стратегий — $A_1, A_2, \ldots, A_m$, а игрок B имеет n стратегий $B_1, B_2, \ldots, B_n$.

Пусть игрок A выбрал стратегию A_i, а игрок B стратегию B_k. Будем считать, что выбор игроками стратегий A_i и B_k однозначно определяет исход игры — выигрыш a_{ik} игрока A и выигрыш b_{ik} игрока B, причем эти выигрыши связаны равенством

$$\boxed{b_{ik} = -a_{ik}}$$

(отрицательный выигрыш на бытовом языке обычно называют проигрышем).

Последнее условие показывает, что в рассматриваемых обстоятельствах выигрыш одного из игроков равен выигрышу другого, взятому с противоположным знаком. Поэтому при анализе такой игры можно рассматривать выигрыши только одного из игроков. Пусть это будут, например, выигрыши игрока A.

Если нам известны значения a_{ik} выигрыша при каждой паре стратегий (в каждой ситуации) $\{A_i, B_k\}$, $i = 1, 2, \ldots, m$, $k = 1, 2, \ldots, n$, то их удобно записывать или в виде прямоугольной таблицы, строки которой соответствуют стратегиям игрока A, а столбцы — стратегиям игрока B:

	B_1	B_2	$\ldots$	B_n
A_1	a_{11}	a_{12}	$\ldots$	a_{1n}
A_2	a_{21}	a_{22}	$\ldots$	a_{2n}
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
A_m	a_{m1}	a_{m2}	$\ldots$	a_{mn}

или в виде матрицы

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}.$$

Полученная матрица имеет размер $m \times n$ и называется *матрицей игры*, или *платежной матрицей* (отсюда и название игры — *матричная*).

Рассматриваемую игру часто называют *игрой* $m \times n$ или $m \times n$ *игрой*.

Замечание. Матричные игры относятся к разряду так называемых антагонистических игр, то есть игр, в которых интересы игроков прямо противоположны.

Пример 1. Каждый из двух игроков A и B одновременно и независимо друг от друга зписывает нв листе бумаги любое целое число. Если выписанные числа имеют одинаковую четность, то игрок A получает от игрока B 1 рубль, а если разную, то наоборот — игрок A платит 1 рубль игроку B.

У игрока A две стратегии: A_1 — записать четное число и A_2 — записать нечетное число. У игрока B такие же две стратегии. B_1 — записать четное число и B_2 — записать нечетное число. Выбор игроками соответственно стратегий A_i и B_k однозначно определяет исход игры — выигрыш игрока A.

Матрица этой 2×2 игры имеет следующий вид

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

(здесь строки соответствуют стратегиям игрока A, а столбцы — стратегиям игрока B).

§ 1. Равновесная ситуация

Рассмотрим следующий пример.

Пример 2. Два игрока A и B, не глядя друг на друга, кладут на стол по картонному кружку красного (r), зеленого (g) или синего (b) цветов, сравнивают цвета кружков и расплачиваются друг с другом так, как показано в матрице игры

$$\begin{pmatrix} -2 & 2 & -1 \\ 2 & 1 & 1 \\ 3 & -3 & 1 \end{pmatrix}$$

(напомним, что у этой 3×3-матрицы строки соответствуют стратегиям игрока A, а столбцы — стратегиям игрока B).

Считая, что эта 3×3 игра повторяется многократно, попробуем определить оптимальные стратегии каждого из игроков.

Начнем с последовательного анализа стратегий игрока A, не забывая о том, что, выбирая стратегию игрока A, должно принимать в расчет, что его противник B может ответить на нее той из своих стратегий, при которой выигрыш игрока A будет минимальным.

Так, на стратегию A_r он ответит стратегией B_r (минимальный выигрыш равен -2, что на самом деле означает проигрыш игрока A, равный 2), на стратегию A_g — стратегией B_g или B_b (минимальный выигрыш игрока A равен 1), а на стратегию A_b — стратегией B_g (минимальный выигрыш игрока A равен -3).

Запишем эти минимальные выигрыши в правом столбце таблицы:

	B_r	B_g	B_b	
A_r	-2	2	-1	-2
A_g	2	1	1	1
A_b	3	-3	1	-3

maxmin. Неудивительно, что игрок A останавливает свой выбор на стратегии A_g, при которой его минимальный выигрыш максимален (из трех чисел -2, 1 и -3 максимальным является 1):

	B_r	B_g	B_b	
A_r	-2	2	-1	-2
$\boldsymbol{A_g}$	2	1	1	$\boxed{1}$
A_b	3	-3	1	-3

$$\boxed{\text{maxmin} = 1.}$$

Если игрок A будет придерживаться этой стратегии, то ему гарантирован выигрыш, не меньший 1, при любом поведении противника в игре.

Аналогичные рассуждения можно провести и за игрока B. Так как игрок B заинтересован в том, чтобы обратить выигрыш игрока A в минимум, то ему нужно проанализировать каждую свою стратегию с точки зрения максимального выигрыша игрока A.

Выбирая свою стратегию, игрок B должен учитывать, что при этом стратегией его противника A может оказаться та, при которой выигрыш игрока A будет максимальным. Так, на стратегию B_r он ответит стратегией A_b (максимальный выигрыш игрока A равен 3), на стратегию B_g — стратегией A_r (максимальный выигрыш игрока A равен 2), а на стратегию B_b — стратегией A_g или A_b (максимальный выигрыш игрока A равен 1). Эти максимальные выигрыши записаны в нижней строке таблицы

	B_r	B_g	B_b	
A_r	−2	2	−1	−2
A_g	2	1	1	1
A_b	3	−3	1	−3
	3	2	1	

minmax. Неудивительно, если игрок B остановит свой выбор на стратегии B_b, при которой максимальный выигрыш игрока A минимален (из трех чисел 3, 2 и 1 минимальным является 1):

	B_r	B_g	$\boldsymbol{B_b}$	
A_r	−2	2	−1	−2
A_g	2	1	1	1
A_b	3	−3	1	−3
	3	2	$\boxed{1}$	

$$\boxed{\mathrm{minmax} = 1.}$$

Если игрок B будет придерживаться этой стратегии, то при любом поведении противника он проиграет не больше 1.

В рассматриваемой игре числа maxmin и minmax совпали:

$$\mathrm{maxmin} = \mathrm{minmax} = 1$$

(соответствующие элементы в таблице

	B_r	B_g	$\boldsymbol{B_b}$
A_r	−2	2	−1
$\boldsymbol{A_g}$	2	1	$\boxed{1}$
A_b	3	−3	1

выделены жирным шрифтом).

Выделенные стратегии A_g и B_b являются оптимальными стратегиями игроков A и B,

$$\boxed{A_g = A_{\mathrm{opt}}, \qquad B_b = B_{\mathrm{opt}},}$$

в следующем смысле:

при многократном повторении игры отказ от выбранной стратегии любым из игроков уменьшает его шансы на выигрыш (увеличивает шансы на проигрыш).

В самом деле, если игрок A будет придерживаться не стратегии A_{opt}, а выберет иную стратегию, например, A_r, то вряд ли стоит рассчитывать на то, что игрок B этого не заметит. Конечно, заметит и не преминет воспользоваться своим наблюдением. Ясно, что в этом случае он отдаст предпочтение стратегии B_r. А на выбор A_b игрок B ответит, например, так — B_g. В результате отказа от стратегии A_g выигрыш игрока A уменьшится.

Если же от стратегии B_{opt} отказывается игрок B, выбирая, например, стратегию B_r, то игрок A может ответить на это стратегией A_b и, тем самым, увеличить свой выигрыш. В случае стратегии B_g ответ игрока A — A_r.

Тем самым, ситуация $\{A_g, B_b\}$ оказывается равновесной.

Еще раз подчеркнем, что элементами матрицы игры являются числа, описывающие выигрыш игрока A. Более точно, выигрыш соответствует положительному элементу платежной матрицы, а отрицательный указывает на проигрыш игрока A.

Матрица выплат игроку B получается из матрицы игры заменой каждого ее элемента на противоположный по знаку.

Рассмотрим теперь произвольную матричную игру

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

(строки заданной $m \times n$-матрицы соответствуют стратегиям игрока A, а столбцы — стратегиям игрока B) и опишем общий алгоритм, посредством которого можно определить, есть ли в этой игре ситуация равновесия или ее нет.

В теории игр предполагается, что оба игрока действуют разумно, то есть стремятся к получению максимального выигрыша, считая, что соперник действует наилучшим (для себя) образом.

Действия игрока A

1-й шаг. В каждой строке матрицы A ищется минимальный элемент

$$\boxed{\alpha_i = \min_k a_{ik}, \quad i = 1, 2, \dots, m.}$$

Полученные числа

$$\alpha_1, \alpha_2, \dots, \alpha_m$$

приписываются к заданной таблице в виде правого добавочного столбца

$$\begin{array}{cccc|c}
a_{11} & a_{12} & \dots & a_{1n} & \alpha_1 \\
a_{21} & a_{22} & \dots & a_{2n} & \alpha_2 \\
\dots & \dots & \dots & \dots & \dots \\
a_{m1} & a_{m2} & \dots & a_{mn} & \alpha_m
\end{array}$$

Пояснение. Выбирая стратегию A_i, игрок A вправе рассчитывать на то, что в результате разумных действий противника (игрока B) он выиграет не меньше чем α_i.

2-й шаг. Среди чисел

$$\alpha_1, \alpha_2, \dots, \alpha_m$$

выбирается максимальное число

$$\boxed{\alpha = \max_i \alpha_i}$$

или, подробнее,

$$\boxed{\alpha = \max_i \min_k a_{ik}.}$$

Специально отметим, что выбранное число α является одним из элементов заданной матрицы A.

Пояснение. Действуя наиболее осторожно и рассчитывая на наиболее разумное поведение противника, игрок A должен остановиться на той стратегии A_i, для которой число α_i является максимальным.

Если игрок A будет придерживаться стратегии, выбранной описанным выше способом, то при любом поведении игрока B игроку A гарантирован выигрыш, не меньший α.

Число α называется *нижней ценой игры.*

Принцип построения стратегии игрока A, основанный на максимизации минимальных выигрышей, называется *принципом максимина*, а выбираемая в соответствии с этим принципом стратегия A_{i0} — *максиминной стратегией* игрока A.

Действия игрока B

1-й шаг. В каждом столбце матрицы А ищется максимальный элемент

$$\boxed{\beta_k = \max_i a_{ik}, \quad k = 1, 2, \ldots, n.}$$

Полученные числа

$$\beta_1, \beta_2, \ldots, \beta_n$$

приписываются к заданной таблице в виде нижней добавочной строки

$$\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & \alpha_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & \alpha_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & \alpha_m \\ \hline \beta_1 & \beta_2 & \ldots & \beta_n & \end{array}$$

Пояснение. Выбирая стратегию B_k, игрок B должен рассчитывать на то, что в результате разумных действий противника (игрока A) он проиграет не больше чем β_k.

2-й шаг. Среди чисел

$$\beta_1, \beta_2, \ldots, \beta_n$$

выбирается минимальное число

$$\boxed{\beta = \min_k \beta_k}$$

или, подробнее,

$$\boxed{\beta = \min_k \max_i a_{ik}.}$$

Выбранное число β также является одним из элементов заданной матрицы А.

Пояснение. Действуя наиболее осторожно и рассчитывая на наиболее разумное поведение противника, игрок B должен остановиться на той стратегии B_k, для которой число β_k является минимальным.

Если игрок B будет придерживаться стратегии, выбранной описанным выше способом, то при любом поведении игрока A игроку B гарантирован проигрыш, не больший β.

Число β называется *верхней ценой игры*.

Принцип построения стратегии игрока B, основанный на минимизации максимальных потерь, называется *принципом минимакса*, а выбираемая в соответствии с этим принципом стратегия B_{k^0} — *минимаксной стратегией* игрока B.

Нижняя цена игры α и верхняя цена игры β всегда связаны неравенством

$$\boxed{\alpha \leqslant \beta.}$$

Замечание. Реализация описанного алгоритма требует $2mn - 1$ сравнений элементов матрицы А:

$$(n-1)m + m - 1 = mn - 1$$

сравнений для определения α,

$$(n-1)m + m - 1 = mn - 1$$

сравнений для определения β и одно сравнение полученных чисел α и β.

Если

$$\boxed{\alpha = \beta,}$$

или, подробнее,

$$\boxed{\max_i \min_k a_{ik} = a_{i^0k^0} = \min_k \max_i a_{ik},}$$

то ситуация $\{A_{i^0}, B_{k^0}\}$ оказывается равновесной, и ни один из игроков не заинтересован в том, чтобы ее нарушить (в этом нетрудно убедиться путем рассуждений, подобных проведенным при анализе игры в примере 2)

В том случае, когда нижняя цена игры равна верхней цене игры, их общее значение называется просто *ценой игры* и обозначается через ν

Цена игры совпадает с элементом $a_{i^0k^0}$ матрицы игры A, расположенным на пересечении i^0-й строки (стратегия A_{i^0} игрока A) и k^0-го столбца (стратегия B_{k^0} игрока B) — минимальным в своей строке и максимальным в своем столбце

Этот элемент называют *седловой точкой матрицы* A, или *точкой равновесия*, а про игру говорят, что она *имеет седловую точку*

Стратегии A_{i^0} и B_{k^0}, соответствующие седловой точке, называются *оптимальными*, а совокупность оптимальных ситуаций и цена игры — *решением матричной игры с седловой точкой*

Замечание Седловых точек в матричной игре может быть несколько, но все они имеют одно и то же значение

Матричные игры с седловой точкой важны и интересны, однако более типичным является случай, когда применение описанного алгоритма приводит к неравенству

$$\boxed{\alpha < \beta}$$

Как показывает следующий пример, в этом случае предложенный выбор стратегий уже, вообще говоря, к равновесной ситуации не приводит, и при многократном ее повторении у игроков вполне могут возникнуть мотивы к нарушению рекомендаций, основанных на описанном алгоритме действий игроков A и B

Пример 3 Рассмотрим 3×3 игру, заданную матрицей

$$\mathsf{A} = \begin{pmatrix} 4 & -1 & -3 \\ -2 & 1 & 3 \\ 0 & 2 & -3 \end{pmatrix}$$

Применив предложенный алгоритм

$$\begin{array}{ccc|c} 4 & 1 & -3 & -3 \\ \boxed{-2} & 1 & 3 & -2 \\ 0 & \boxed{2} & -3 & -3 \\ \hline 4 & 2 & 3 & \end{array}$$

находим нижнюю цену игры $\alpha = -2$ и соответствующую ей стратегию A_2, и верхнюю цену игры $\beta = 2$ и соответствующую ей стратегию B_2

Нетрудно убедиться в том что пока игроки придерживаются этих стратегий, средний выигрыш при многократном повторении игры равен 1 Он больше нижней цены игры, но меньше верхней цены

Однако если игроку B станет известно, что игрок A придерживается стратегии A_2, он немедленно ответит стратегией B_1 и сведет его выигрыш к проигрышу -2 В свою очередь, на стратегию B_1 у игрока A имеется ответная стратегия A_1, дающая ему выигрыш 4

Тем самым, ситуация $\{A_2, B_2\}$ равновесной не является

§ 2. Смешанные стратегии

В случае, когда нижняя цена игры α и верхняя цена игры β не совпадают,

$$\alpha < \beta,$$

игрок A может обеспечить себе выигрыш, не меньший α, а игрок B имеет возможность не дать ему больше, чем β Возникает вопрос — а как разделить между игроками разность

$$\beta - \alpha ?$$

Предыдущие построения на этот вопрос ответа не дают — тесны рамки возможных действий игроков. Поэтому довольно ясно, что механизм, обеспечивающий получение каждым из игроков как можно большей доли этой разности, следует искать в определенном расширении стратегических возможностей, имеющихся у игроков изначально.

Оказывается, что компромиссного распределения разности $\beta - \alpha$ между игроками и уверенного получения каждым игроком своей доли при многократном повторении игры можно достичь путем случайного применения ими своих первоначальных, чистых стратегий. При таких действиях

- во-первых, обеспечивается наибольшая скрытность выбора стратегии (результат выбора не может стать известным противнику, поскольку он неизвестен самому игроку),
- во-вторых, при разумном построении механизма случайного выбора стратегий, последние оказываются оптимальными.

Случайная величина, значениями которой являются стратегии игрока, называется его *смешанной стратегией*.

Тем самым, задание смешанной стратегии игрока состоит в указании тех вероятностей, с которыми выбираются его первоначальные стратегии.

2.1. Основные определения

Рассмотрим произвольную $m \times n$ игру, заданную $m \times n$-матрицей

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Так как игрок A имеет m чистых стратегий, то его смешанная стратегия может быть описана набором m неотрицательных чисел

$$\boxed{p_1 \geqslant 0,\ \ p_2 \geqslant 0,\ \ \dots,\ \ p_m \geqslant 0,}$$

сумма которых равна 1,

$$\boxed{\sum_{i=1}^{m} p_i = 1.}$$

Смешанная стратегия второго игрока B, имеющего n чистых стратегий, описывается набором n неотрицательных чисел

$$\boxed{q_1 \geqslant 0,\ \ q_2 \geqslant 0,\ \ \dots,\ \ q_n \geqslant 0,}$$

сумма которых равна 1,

$$\boxed{\sum_{k=1}^{n} q_k = 1.}$$

Замечание. Каждая чистая стратегия является частным случаем смешанной стратегии: в частности, чистая стратегия A_i является смешанной стратегией, описываемой набором чисел $p_1, p_2, \dots, p_m$, в котором

$$p_i = 1, \quad p_j = 0 \quad (j \neq i).$$

Подчеркнем, что для соблюдения секретности каждый из игроков применяет свои стратегии независимо от другого игрока.

Таким образом, задав два набора

$$\boxed{P = \{p_1, p_2, \ldots, p_m\}, \qquad Q = \{q_1, q_2, \ldots, q_n\},}$$

мы оказываемся *в ситуации в смешанных стратегиях*. В этих условиях каждая обычная ситуация (в чистых стратегиях) $\{A_i, B_k\}$ по определению является случайным событием и, ввиду независимости наборов P и Q, реализуется с вероятностью $p_i q_k$. В этой ситуации $\{A_i, B_k\}$ игрок A получает выигрыш a_{ik}. Тем самым, математическое ожидание выигрыша игрока A в условиях ситуации в смешанных стратегиях (P, Q) равно

$$\boxed{H_A(P, Q) = \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik} p_i q_k.}$$

Это число и принимается за *средний выигрыш* игрока A при смешанных стратегиях

$$P = \{p_1, p_2, \ldots, p_m\}, \qquad Q = \{q_1, q_2, \ldots, q_n\}.$$

Определение. Стратегии

$$P^0 = \left\{p_1^0, p_2^0, \ldots, p_m^0\right\} \quad \text{и} \quad Q^0 = \left\{q_1^0, q_2^0, \ldots, q_n^0\right\},$$

называются ***оптимальными смешанными стратегиями*** игроков A и B соответственно, если выполнено следующее соотношение

$$H_A\left(P, Q^0\right) \leqslant H_A\left(P^0, Q^0\right) \leqslant H_A\left(P^0, Q\right).$$

Пояснение. Выписанные неравенства означают следующее:

левое неравенство — отклонение игрока A от оптимальной стратегии P^0 при условии, что игрок B придерживается стратегии Q^0, приводит к тому, что выигрыш отклонившегося игрока A может только уменьшиться,

правое неравенство — отклонение игрока B от оптимальной стратегии Q^0 при условии, что игрок A придерживается стратегии P^0, приводит к тому, что выигрыш игрока A может только возрасти, и значит, выигрыш игрока B — только уменьшиться.

Приведенное условие оптимальности равносильно тому, что[1)]

$$\boxed{\max_P \min_Q H_A(P, Q) = H_A\left(P^0, Q^0\right) = \min_Q \max_P H_A(P, Q).}$$

Величина

$$\boxed{\nu = H_A\left(P^0, Q^0\right),}$$

определяемая последней формулой, называется *ценой игры*.

Набор $\left(P^0, Q^0, \nu\right)$, состоящий из оптимальных смешанных стратегий игроков A и B и цены игры, называется *решением матричной игры*.

Пример 4. Рассмотрим 2×2 матричную игру из примера 1. Матрица этой игры имеет следующий вид

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Как нетрудно убедиться, седловой точки у нее нет.

[1)] Экстремальные величины $\max\limits_P \min\limits_Q H_A(P, Q)$ и $\min\limits_Q \max\limits_P H_A(P, Q)$ всегда существуют вследствие того, что функция $H_A(P, Q)$ является непрерывной на замкнутом множестве

$$p_i \geqslant 0, \quad \sum p_i = 1, \qquad q_k \geqslant 0, \quad \sum q_k = 1.$$

Смешанные стратегии игроков A и B могут быть описаны парами чисел

$$P = \{p, 1-p\} \quad \text{и} \quad Q = \{q, 1-q\}$$

соответственно.

Средний выигрыш игрока A вычисляется так:

$$H_A(P,Q) = 1\cdot p\cdot q + (-1)\cdot p\cdot(1-q) + (-1)p\cdot(1-p)\cdot q + 2\cdot(1-p)(1-q),$$

откуда легко следует, что

$$H_A(P,Q) = 4pq - 2p - 2q + 1 = (2p-1)(2q-1).$$

Последнее удобно записать так

$$H_A(P,Q) = 4\left(p - \tfrac{1}{2}\right)\left(q - \tfrac{1}{2}\right).$$

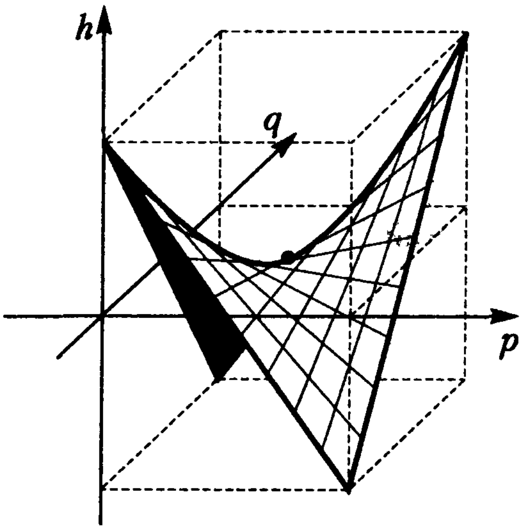


Рис. 1

Полученная формула показывает, что если игрок A в половине случаев записывает на листе бумаге четное (нечетное) число (выбирает $p = 1/2$), то независимо от того, что делает игрок B, ожидаемый (средний) выигрыш игрока A в каждой партии будет нулевым.

Если же игрок A выберет $p > 1/2$ (так что разность $p - 1/2$ будет положительной), то узнвв об этом, игрок B может выбрать $q < 1/2$ (так что разность $q - 1/2$ будет отрицательной) и, тем самым, сделвть средний выигрыш игрока A отрицательным, то есть заставит его проиграть. Если же игрок A выберет $p < 1/2$ (так что разность $p - 1/2$ будет отрицательной) и игрок B узнает об этом, то он может выбрать $q > 1/2$ (так что разность $q - 1/2$ будет положительной) и вновь сделать выигрыш игрока A отрицательным, то есть опять заставит его проиграть.

Исследуем теперь эту формулу с точки зрения игрока B.

Если игрок A выбирает $p = 1/2$, то ожидаемый (средний) выигрыш игрока B независимо от его действий будет нулевым в каждой партии. Но если игрок B выберет $q > 1/2$ (так что разность $q - 1/2$ будет положительной), то, узнав об этом, игрок A может выбрать $p < 1/2$ (так что разность $p - 1/2$ будет отрицательной), и тогда игрок B будет проигрывать в каждой партии. Если же игрок B выберет $q < 1/2$ (так что разность $q - 1/2$ будет отрицательной) и игрок A узнает об этом, то он может выбрать $p > 1/2$ (так что разность $p - 1/2$ будет положительной) и вновь заставит игрока B проиграть.

Тем самым, наборы

$$P^0 = \left\{\tfrac{1}{2}, \tfrac{1}{2}\right\}, \quad Q^0 = \left\{\tfrac{1}{2}, \tfrac{1}{2}\right\}$$

являются оптимальными, а исход игры ничейным:

$$\boxed{\nu = 0.}$$

Замечание. На рисунке 1 показано, как устроена поверхность, описываемая функцией

$$h(p,q) = 4pq - 2p - 2q + 1 = (2p-1)(2q-1).$$

Точка

$$\left(\tfrac{1}{2}, \tfrac{1}{2}, 0\right)$$

является седловой точкой (точкой перевала) этой поверхности. Именно эта точка и дает решение рассматриваемой матричной игры.

Естественно возникают два *ключевых вопроса*:

1-й — какие матричные игры имеют решение в смешанных стратегиях?

2-й — как находить решение матричной игры, если оно существует?

Ответы на эти вопросы дают следующие две теоремы.

2.2. Основная теорема теории матричных игр

Теорема 1 (Дж. фон Нейман). *Для матричной игры с любой матрицей* A *величины*

$$\max_P \min_Q H_A(P, Q), \qquad \min_Q \max_P H_A(P, Q)$$

равны между собой,

$$\max_P \min_Q H_A(P, Q) = \min_Q \max_P H_A(P, Q).$$

Более того, существует хотя бы одна ситуация в смешанных стратегиях $\left(P^0, Q^0\right)$, для которой выполняется соотношение

$$H_A\left(P^0, Q^0\right) = \max_P \min_Q H_A(P, Q) = \min_Q \max_P H_A(P, Q).$$

Иными словами, любая матричная игра имеет решение в смешанных стратегиях. Поиск этого решения опирается на следующие установленные факты.

2.3. Основные свойства оптимальных смешанных стратегий

Теорема 2. *Пусть*

$$P^0 = \left\{p_1^0, p_2^0, \ldots, p_m^0\right\} \quad \text{и} \quad Q^0 = \left\{q_1^0, q_2^0, \ldots, q_n^0\right\}$$

— оптимальные смешанные стратегии и ν — цена игры.

Оптимальная смешанная стратегия P^0 игрока A смешивается только из тех чистых стратегий A_i, $i = 1, 2, \ldots, m$, то есть отличными от нуля могут быть вероятности p_i только с теми номерами $i = 1, 2, \ldots, m$, для которых выполнены равенства

$$\sum_{k=1}^{n} a_{ik} q_k^0 = \nu.$$

Это означает, что смешиваются не все чистые стратегии. Аналогично,

в оптимальной смешанной стратегии Q^0 игрока B отличных от нуля могут быть только те вероятности q_k, для номеров $k = 1, 2, \ldots, n$, которых выполнены равенства

$$\sum_{i=1}^{m} a_{ik} p_i^0 = \nu.$$

Кроме того, имеют место соотношения

$$\boxed{\begin{aligned}\nu &= \min_{1\leqslant k\leqslant n} \sum_{i=1}^{m} a_{ik} p_i^0 = \max_P \min_{1\leqslant k\leqslant n} \sum_{i=1}^{m} a_{ik} p_i = \\ &= \min_Q \max_{1\leqslant i\leqslant m} \sum_{k=1}^{n} a_{ik} q_k = \max_{1\leqslant i\leqslant m} \sum_{k=1}^{n} a_{ik} q_k^0 = \nu.\end{aligned}}$$

В этом последнем скоплении равенств, по существу, и лежат истоки, питающие методы построения решений матричных игр. Опишем простейшие из них.

§ 3. Методы решения матричных игр

Наши рассмотрения мы начнем с матричных игр, число стратегий хотя бы одного из игроков в которых равно двум.

Для построения решений $2 \times n$ и $m \times 2$ игр существует эффективный метод, основанный на простых геометрических соображениях. Этот метод называют *графическим*.

3.1. $2 \times n$ игры

Пусть

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \end{pmatrix}$$

— платежная матрица $2 \times n$ игры.

Согласно теореме о двойном описании игры (теорема 2) нахождение цены игры и оптимального значения p^0 для игрока A равносильно разрешению уравнения

$$\boxed{\nu = \min_{1 \leqslant k \leqslant n} \left(a_{1k}p^0 + a_{2k}\left(1 - p^0\right)\right) = \max_{0 \leqslant p \leqslant 1} \min_{1 \leqslant k \leqslant n} \left(a_{1k}p + a_{2k}(1 - p)\right).}$$

Опишем общую схему, приводящую к искомому результату.

Максимум функции

$$\min_{1 \leqslant k \leqslant n} \left(a_{1k}p + a_{2k}(1 - p)\right) \tag{$\star$}$$

проще всего найти, построив ее график. Для этого поступают следующим образом.

Предположим, что игрок A выбрал смешанную стратегию $P = \{p,\ 1 - p\}$, а игрок B — k-ю чистую стратегию, $k = 1, 2, \dots, n$. Тогда средний выигрыш игрока A в ситуации $\{P,\ k\}$ оказывается равным

$$(k): \quad w = a_{1k}p + a_{2k}(1 - p).$$

На плоскости $(p,\ w)$ уравнение (k) описывает прямую. Тем самым, каждой чистой стратегии игрока B на этой плоскости соответствует своя прямая.

Поэтому сначала на плоскости $(w,\ p)$ последовательно и аккуратно рисуются все прямые

$$(k): \quad w = a_{1k}p + a_{2k}(1 - p), \qquad k = 1, 2, \dots, n$$

(рис. 2). Затем для каждого значения p, $0 \leqslant p \leqslant 1$, путем визуального сравнения соответствующих ему значений w на каждой из построенных прямых определяется и отмечается минимальное из них (рис. 3). В результате описанной процедуры получается ломаная, которая и является графиком функции $(\star)$ (выделена жирным на рис. 4). Эта ломаная как бы огибает снизу все семейство построенных прямых,

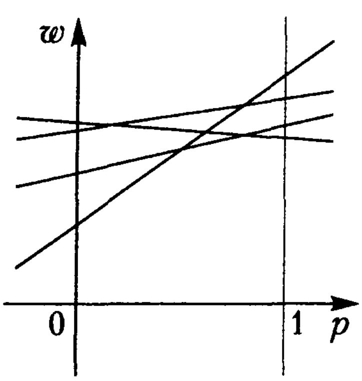


Рис. 2

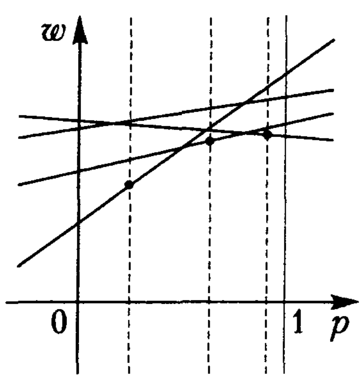


Рис. 3

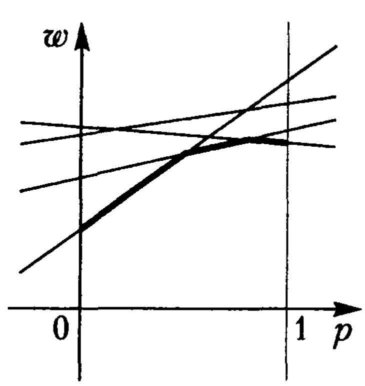


Рис. 4

и по этой причине ее принято называть ***нижней огибающей*** этого семейства. Верхняя точка построенной нижней огибающей определяет и цену игры — ν и оптимальную стратегию игрока A — $P^0 = \{p^0, 1 - p^0\}$ (рис. 5).

Замечание. Описанная процедура может рассматриваться как некоторый аналог максиминного подхода при отсутствии седловой точки.

Опробуем описанную схему решения $2 \times n$ игры на конкретном примере.

Пример 5. Рассмотрим игру, заданную 2×6 матрицей

$$\begin{pmatrix} 6 & 4 & 3 & 1 & -1 & 0 \\ -2 & -1 & 1 & 0 & 5 & 4 \end{pmatrix}.$$

Решение.

1-й шаг. *Анализ игры на наличие седловой точки.*

Нижняя цена игры равна -1, верхняя цена игры равна 1. Седловой точки нет. Решение игры нужно искать в смешанных стратегиях.

2-й шаг. *Вычисление средних выигрышей игрока A* (проводится при условии, что игрок B выбирает только чистые стратегии).

Из таблицы

$$\begin{array}{r|rrrrrr} p & 6 & 4 & 3 & 1 & -1 & 0 \\ 1-p & -2 & -1 & 1 & 0 & 5 & 4 \end{array}$$

легко получаем:

$$\begin{aligned}
&(1): && w = 6p - 2(1-p), \\
&(2): && w = 4p - (1-p), \\
&(3): && w = 3p + (1-p), \\
&(4): && w = p, \\
&(5): && w = -p + 5(1-p), \\
&(6): && w = 4(1-p).
\end{aligned}$$

3-й шаг. *Построение нижней огибающей.*

Аккуратно строим на координатной плоскости (p, w) все шесть прямых, уравнения которых получены на 2-м шаге (рис. 6), и находим их нижнюю огибающую.

4-й шаг. *Отыскание цены игры и оптимальной смешанной стратегии игрока A.*

При аккуратном построении нижней огибающей, нетрудно определить, точкой пересечения каких двух из построенных шести прямых является ее наивысшая точка. В данном случае это прямые (4) и (5), заданные уравнениями $w = p$ и $w = -p + 5(1-p)$ соответственно. Решая систему уравнений

$$\begin{aligned} w &= p, \\ w &= -p + 5(1-p), \end{aligned}$$

получаем

$$p^0 = \tfrac{5}{7}, \quad w^0 = \tfrac{5}{7}$$

(рис. 7).

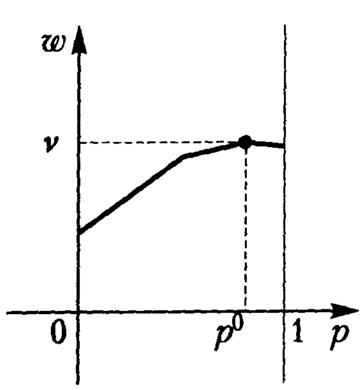


Рис. 5

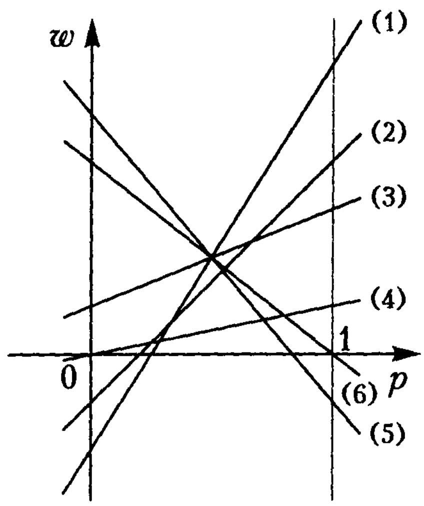


Рис. 6

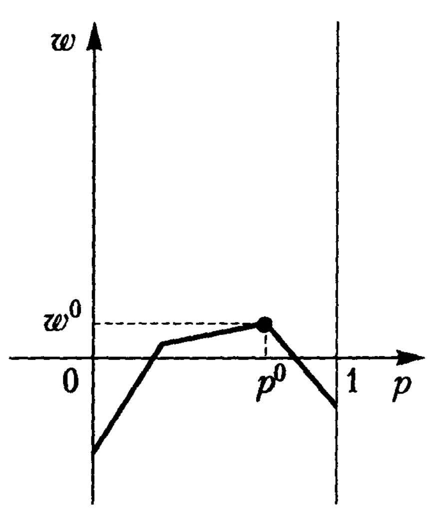


Рис. 7

Тем самым, цена игры ν и оптимальная стратегия P^0 игрока A соответственно равны

$$\nu = \frac{5}{7}, \quad P^0 = \left\{\frac{5}{7}, \frac{2}{7}\right\}.$$

Собственно, этим и заканчивается решение игры для игрока A, поскольку его в первую очередь интересует отыскание собственной оптимальной стратегии и ожидаемого наилучшего гарантированного результата.

Замечание. Решающий матричную игру обычно отождествляет себя с одним из игроков (как правило, это игрок A), считая другого своим противником. Это связано еще и с тем, что в некоторых случаях основное внимание уделяется поиску оптимальных стратегий только игрока A, а стратегии противника могут вообще не интересовать исследователя.

Однако в целом ряде случаев оказывается важным знать оптимальные смешанные стратегии обоих игроков.

Как ищется оптимальная смешанная стратегия игрока A, мы уже описали. Покажем теперь, как отыскать оптимальную смешанную стратегию игрока B.

Здесь, в зависимости от формы нижней огибающей, может представиться несколько случаев.

А. Нижняя огибающая имеет ровно одну наивысшую точку (p^0, w^0):

1) Если $p^0 = 0$ (оптимальная стратегия игрока A — чистая стратегия A_2), то игроку B выгодно применять чистую стратегию, соответствующую номеру прямой, проходящей через точку $(0, w^0)$ и имеющей наибольший отрицательный наклон (рис. 8).

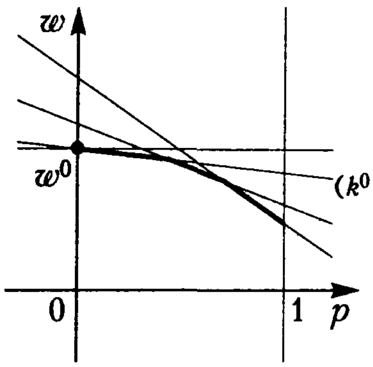


Рис. 8

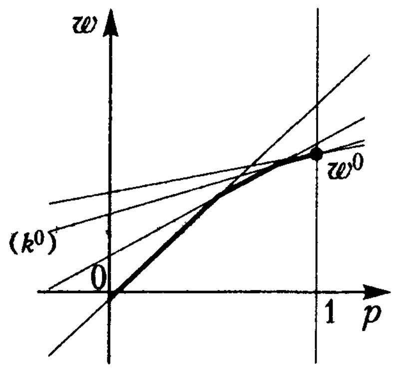


Рис. 9

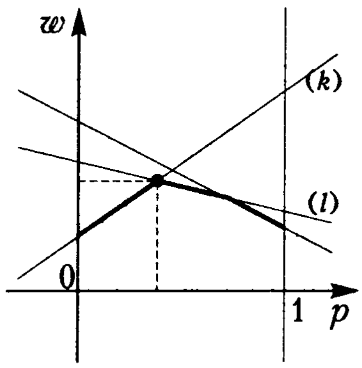


Рис. 10

2) Если $p^0 = 1$ (оптимальная стратегия игрока A — чистая стратегия A_1), то оптимальной для игрока B является чистая стратегия, соответствующая номеру прямой, проходящей через точку $(1, w^0)$ и имеющей наименьший положительный наклон (рис. 9).

3) Если $0 < p^0 < 1$, то в наивысшей точке нижней огибающей пересекаются, по меньшей мере, две прямые, одна из которых (k-я) имеет положительный наклон, а другая (l-я) — отрицательный (рис. 10), и оптимальная смешанная стратегия игрока B получается, если положить

$$q_k = q, \ q_l = 1 - q, \ q_j = 0, \ j \neq k, l,$$

где q — решение уравнения

$$a_{1k}q + a_{1l}(1 - q) = a_{2k}q + a_{2l}(1 - q).$$

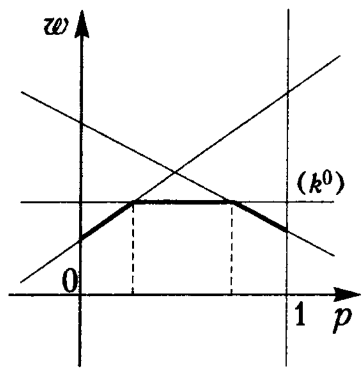


Рис. 11

Б. Нижняя огибающая содержит горизонтальный участок, соответствующий чистой стратегии k^0 игрока B, которая и является оптимальной для него (рис. 11).

Пример 5 (продолжение). Покажем теперь, как найти полное решение игры, то есть еще и оптимальную смешанную стратегию

$$Q^0 = \left\{q_1^0, q_2^0, q_3^0, q_4^0, q_5^0, q_6^0\right\}$$

игрока B.

Для этого поступают так:

1) полагают

$$q_1^0 = 0, \quad q_2^0 = 0, \quad q_3^0 = 0, \quad q_4^0 = q, \quad q_5^0 = 1 - q, \quad q_6^0 = 0$$

(выделяя тем самым из шести чистых стратегий игрока B стратегии B_4 и B_5, которым соответствуют прямые (4) и (5), определяющие наивысшую точку нижней огибающей),

2) приравнивают любой из двух средних выигрышей игрока B (игрок A выбирает только чистые стратегии), отвечающих предложенной смешанной стратегии

0	0	0	q	$1-q$	0
6	4	3	1	-1	0
-2	-1	1	0	5	4

к цене игры

$$q - (1 - q) = \tfrac{5}{7}, \qquad 5(1 - q) = \tfrac{5}{7}$$

и

3) получают (в обоих случаях), что

$$q^0 = \tfrac{6}{7}.$$

Полное решение игры имеет следующий вид

$$P^0 = \left\{\tfrac{5}{7}, \tfrac{2}{7}\right\}, \quad Q^0 = \left\{0, 0, 0, \tfrac{6}{7}, \tfrac{1}{7}, 0\right\}, \quad \nu = \tfrac{5}{7}.$$

Замечание. Ситуация с наличием лишь двух конкурирующих стратегий игрока A нельзя считать надуманной. Она возникает сравнительно часто. Например, в случае, если нужно сравнить два образца некоторого изделия (скажем, старого и модернизированного) с целью выяснения возможности замены, это весьма удобно сделать при помощи платежной матрицы $2 \times n$.

3.2. $m \times 2$ игры

Пусть теперь в матричной игре две чистые стратегии имеет игрок B, а число чистых стратегий у игрока A произвольно (равно m). Это означает, что платежная матрица такой игры имеет вид

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \vdots & \vdots \\ a_{m1} & a_{m2} \end{pmatrix}.$$

Анализ такой игры во многом напоминает рассуждения, описанные для игры $2 \times n$.

Пусть $Q = \{q, 1-q\}$ — произвольная смешанная стратегия игрока B. Если игрок A выбирает i-ю чистую стратегию, $i = 1, 2, \ldots, m$, то средний выигрыш игрока B в ситуации $\{i, Q\}$ будет равным

$$(i): \quad w = a_{i1}q + a_{i2}(1 - q), \qquad i = 1, 2, \ldots, m. \qquad (*)$$

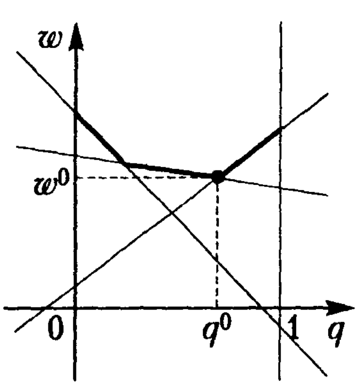


Рис. 12

Зависимость этого выигрыша от переменной q описывается прямой.

Графиком функции

$$\max_{1 \leqslant i \leqslant m} \left(a_{i1}q + a_{i2}(1 - q)\right)$$

является верхняя огибающая семейства прямых (*), соответствующих чистым стратегиям игрока A (рис. 12).

Абсциссой нижней точки полученной ломаной будет значение q^0, определяющее оптимальную смешанную стратегию игрока B, а ординатой w^0 — цена игры.

Замечание. Отыскание оптимальной смешанной стратегии игрока A проводится по той же схеме, которая позволяет находить оптимальную смешанную стратегию игрока B в игре $2 \times n$.

Рассмотрим конкретный пример.

Пример 6. 3×2 игра задана матрицей

$$\begin{pmatrix} 3 & -1 \\ -1 & 3 \\ 1 & 0 \end{pmatrix}.$$

Решение.

1-й шаг. *Анализ игры на наличие седловой точки.*

Нижняя цена игры равна 0, верхняя цена игры равна 3. Седловой точки нет. Решение игры нужно искать в смешанных стратегиях.

2-й шаг. *Вычисление средних выигрышей игрока B* (проводится при условии, что игрок A выбирает только чистые стратегии).

Из таблицы

q	$1-q$
3	−1
−1	3
1	0

получаем:

$$\begin{aligned} (1):&\quad w = 3q - (1-q), \\ (2):&\quad w = -q + 3(1-q), \\ (3):&\quad w = q. \end{aligned}$$

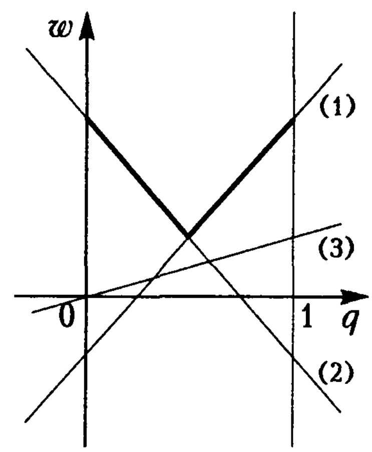


Рис. 13

3-й шаг. *Построение верхней огибающей.*

Построим на координатной плоскости (q, w) все три прямых, а затем и их верхнюю огибающую (рис. 13).

4-й шаг. *Отыскание цены игры и оптимальной смешанной стратегии игрока B.*

Нижняя точка верхней огибающей является точкой пересечения прямых (1) и (2). Решая систему уравнений

$$\begin{aligned} w &= 3q - (1-q), \\ w &= -q + 3(1-q), \end{aligned}$$

получаем

$$q^0 = \frac{1}{2}, \quad w^0 = 1.$$

5-й шаг. *Отыскание оптимальной смешанной стратегии игрока A.*

Полагая

$$p_1^0 = p, \quad p_2^0 = 1 - p, \quad p_3^0 = 0,$$

приравниваем средние выигрыши игрока A, соответствующие чистым выигрышам игрока B,

$$3p - (1-p) = -p + 3(1-p),$$

и находим $p^0 = 1/2$.

Таким образом, цена игры и оптимальные смешанные стратегии игроков A и B соответственно равны

$$\nu = 1, \quad P^0 = \left\{\frac{1}{2}, \frac{1}{2}, 0\right\}, \quad Q^0 = \left\{\frac{1}{2}, \frac{1}{2}\right\}.$$

3.3. $m \times n$ игры

В принципе решение любой матричной игры сводится к решению стандартной задачи линейного программирования и, тем самым, может быть найдено методами линейного программирования. При этом требуемый объем вычислений напрямую зависит от числа чистых стратегий игроков (растет с его увеличением и, значит, с увеличением размеров матрицы игры). Поэтому любые приемы предварительного анализа игры, позволяющие уменьшать размеры ее платежной матрицы или еще как-то упрощать эту матрицу, не нанося ущерба решению, играют на практике весьма важную роль.

Правило доминирования

В целом ряде случаев анализ платежной матрицы обнаруживает, что некоторые чистые стратегии не могут внести никакого вклада в искомые оптимальные смешанные стратегии. Отбрасывание подобных стратегий позволяет заменить первоначальную матрицу на матрицу выигрышей меньших размеров.

Опишем одну из таких возможностей более подробно.

Пусть

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

— произвольная $m \times n$-матрица. Будем говорить, что i-я строка матрицы A

$$a_{i1} \quad a_{i2} \quad \dots \quad a_{in}$$

не больше j-й строки этой матрицы

$$a_{j1} \quad a_{j2} \quad \dots \quad a_{jn},$$

если одновременно выполнены следующие n неравенств

$$\boxed{a_{i1} \leqslant a_{j1}, \quad a_{i2} \leqslant a_{j2}, \quad \dots, \quad a_{in} \leqslant a_{jn}.}$$

При этом говорят также, что j-я строка *доминирует* i-ю строку, или что стратегия A_j игрока *A доминирует* стратегию A_i.

Замечание. Игрок A поступит разумно, если будет избегать стратегий, которым в матрице игры отвечают доминируемые строки.

Если в матрице A одна из строк (j-я) доминирует другую строку (i-ю), то число строк в матрице A можно уменьшить путем отбрасывания доминируемой строки (i-й).

Далее, будем говорить, что k-й столбец матрицы A

$$\begin{matrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{matrix}$$

не меньше l-го столбца этой матрицы

$$\begin{matrix} a_{1l} \\ a_{2l} \\ \vdots \\ a_{ml}, \end{matrix}$$

если одновременно выполнены следующие m неравенств

$$\boxed{a_{1k} \geqslant a_{1l}, \quad a_{2k} \geqslant a_{2l}, \quad \dots, \quad a_{mk} \geqslant a_{ml}.}$$

При этом говорят также, что l-й столбец *доминирует* k-й столбец, или что стратегия B_l игрока *B доминирует* стратегию B_k.

Замечание. Игрок B поступит разумно, если будет избегать стратегий, которым в матрице игры отвечают доминируемые столбцы.

Если в матрице A один из столбцов (l-й) доминирует другой столбец (k-й), то число столбцов в матрице A можно уменьшить путем отбрасывания доминируемого столбца (k-го).

Важное замечание. Оптимальные смешанные стратегии в игре с матрицей, полученной усечением исходной за счет доминируемых строк и столбцов, дадут оптимальное решение в исходной игре: доминируемые чистые стратегии игроков в смешении не участвуют — соответствующие им вероятности следует взять равными нулю.

Пример 7. Рассмотрим игру с матрицей

$$\begin{pmatrix} -1 & 0 & 2 & 1 \\ -2 & 0 & 1 & 0 \\ 2 & 1 & -1 & -2 \\ -1 & 0 & 2 & 1 \end{pmatrix}.$$

Сравнивая строки матрицы, видим, что 1-я строка совпадает с 4-й строкой, или, что то же, стратегия A_4 дублирует стратегию A_1. Тем самым, одну из этих строк можно вычеркнуть, не нанося ущерба решению

$$\begin{pmatrix} -1 & 0 & 2 & 1 \\ -2 & 0 & 1 & 0 \\ 2 & 1 & -1 & -2 \end{pmatrix}.$$

Поэлементно сравнивая 1-ю и 2-ю строки, замечаем, что 1-я строка доминирует 2-ю строку, или, что то же, стрвтегия A_1 доминирует стратегию A_2. Это вновь позволяет уменьшить число строк матрицы

$$\begin{pmatrix} -1 & 0 & 2 & 1 \\ 2 & 1 & -1 & -2 \end{pmatrix}.$$

Замечая, что 4-й столбец полученной матрицы доминирует ее 3-й столбец, приходим к игре с 2×3-матрицей

$$\begin{pmatrix} -1 & 0 & 1 \\ 2 & 1 & -2 \end{pmatrix}.$$

Решвя эту 2×3 игру графическим методом, находим ее решение — цену игры и оптимальные смешанные стратегии игроков A и B:

$$\nu = 0, \quad P^0 = \left\{\frac{2}{3}, \frac{1}{3}\right\}, \quad Q^0 = \left\{\frac{1}{2}, 0, \frac{1}{2}\right\}.$$

Возвращаясь к исходной 4×4 игре, получаем окончательный ответ:

$$\nu = 0, \quad \left\{\frac{2}{3}, 0, \frac{1}{3}, 0\right\}, \quad \left\{\frac{1}{2}, 0, 0, \frac{1}{2}\right\}.$$

Замечание. При отбрасывании доминируемых строк и столбцов некоторые из оптимальных стратегий могут быть потеряны. Однако цена игры не изменится, и по усеченной матрице может быть найдена хотя бы одна пара оптимальных смешанных стратегий.

Аффинное правило

При поиске решения матричных игр часто оказывается полезным следующее свойство.

Допустимые преобразования матрицы игры и ее цена. *Оптимальные стратегии у матричных игр, элементы матриц* A *и* C *которых связаны равенствами*

$$c_{ik} = \lambda a_{ik} + \mu, \qquad i = 1, 2, \dots, m; \qquad k = 1, 2, \dots, n,$$

где $\lambda > 0$, а μ — произвольно, имеют одинаковые равновесные ситуации (либо в чистых либо в смешанных стратегиях), а их цены удовлетворяют следующему условию

$$\nu_{\mathrm{C}} = \lambda \nu_{\mathrm{A}} + \mu.$$

Пример 8. Элементы матриц

$$\mathrm{A} = \begin{pmatrix} -1 & 0 & 1 \\ 2 & 1 & -2 \end{pmatrix} \quad \text{и} \quad \mathrm{C} = \begin{pmatrix} 2 & 5 & 8 \\ 11 & 8 & -1 \end{pmatrix}$$

связаны равенством

$$c_{ik} = 3 \cdot a_{ik} + 5, \qquad i = 1, 2; \qquad k = 1, 2, 3.$$

Поэтому цена игры с матрицей C легко вычисляется

$$\nu_{\mathrm{C}} = 3 \cdot \nu_{\mathrm{A}} + 5 = 3 \cdot 0 + 5 = 5$$

(см. пример 7).

Основные этапы поиска решения матричной игры

1-й этап — проверка наличия (или отсутствия) равновесия в чистых стратегиях (при наличии равновесной ситуации указываются соответствующие оптимальные стратегии игроков и цена игры).

2-й этап — поиск доминирующих стратегий (в случае успеха этого поиска — отбрасывание доминируемых строк и столбцов в исходной матрице игры).

3-й этап — замена игры на ее смешанное расширение и отыскание оптимальных смешанных стратегий и цены игры.

3.4. Итерационный метод решения матричных игр

Опишем метод отыскания решения матричной игры — цены игры и оптимальных смешанных стратегий, в известной степени верно отражающий некоторую реальную ситуацию накопления опыта постепенной выработки игроками *хороших* стратегий в результате многих повторений конфликтных ситуаций. Основная идея этого метода заключается в том, чтобы мысленно как бы смоделировать реальное практическое «обучение» игроков в ходе самой игры, когда каждый из игроков на собственном опыте прощупывает способ поведения противника и старается отвечать на него наиболее выгодным для себя образом. Иными словами, всякий раз при возобновлении игры игрок выбирает наиболее выгодную для себя стратегию, опираясь на предыдущий выбор противника.

Проиллюстрируем этот метод на примере игры, заданной матрицей

$$\begin{pmatrix} 2 & 0 & 3 \\ 1 & 3 & -3 \end{pmatrix}$$

(здесь $\max\min = 0$, $\min\max = 2 \Rightarrow$ седловой точки нет).

Опишем правила выбора ходов игроками, предположив, для определенности, что начинает игрок A:

$$\text{ход игрока } A \text{ — стратегия } A_1 \text{ — } (2 \quad \mathbf{0} \quad 3);$$

игрок B выбирает свою стратегию так, чтобы выигрыш игрока A был минимален (отмечен полужирным шрифтом):

$$\text{ход игрока } B \text{ — стратегия } B_2 \text{ — } \begin{pmatrix} 0 \\ \mathbf{3} \end{pmatrix};$$

игрок A выбирает свою стратегию так, чтобы его выигрыш при стратегии B_2 игрока B был максимален (отмечен полужирным шрифтом):

$$\text{ход игрока } A \text{ — стратегия } A_2 \text{ — } (1 \quad 3 \quad -3),$$

игрок B выбирает свою стратегию так, чтобы «накопленный» выигрыш игрока A при стратегиях A_1 и A_2,

$$(2 \quad 0 \quad 3) + (1 \quad 3 \quad -3) = (3 \quad 3 \quad \mathbf{0}),$$

был минимален:

$$\text{ход игрока } B \text{ — стратегия } B_3 \text{ — } \begin{pmatrix} 3 \\ -3 \end{pmatrix};$$

игрок A выбирает свою стратегию так, чтобы его «накопленный» выигрыш при стратегиях B_2 и B_3 игрока B,

$$\begin{pmatrix} 0 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 \\ -3 \end{pmatrix} = \begin{pmatrix} \mathbf{3} \\ 0 \end{pmatrix},$$

был максимален:

$$\text{ход игрока } A \text{ — стратегия } A_1 \text{ — } (2 \quad 0 \quad 3),$$

игрок B выбирает свою стратегию так, чтобы «накопленный» выигрыш игрока A при стратегиях A_1, A_2 и A_1,

$$(3 \quad 3 \quad 0) + (2 \quad 0 \quad 3) = (5 \quad \mathbf{3} \quad 3),$$

был минимален:

$$\text{ход игрока } B \text{ — стратегия } B_2 \text{ — } \begin{pmatrix} 0 \\ 3 \end{pmatrix}$$

и т. д.

Разобьем последовательные ходы игроков A и B на пары

(*ход игрока* A, *ход игрока* B)

и запишем результаты в таблице

n	i	B_1	B_2	B_3	$\nu_*(n)$	k	A_1	A_2	$\nu^*(n)$	$\nu(n)$
1	1	2	**0**	3	0,00	2	0	**3**	3,00	1,50
2	2	3	3	**0**	0,00	3	**3**	0	1,50	0,75
3	1	5	**3**	3	1,00	2	**3**	3	1,00	1,00
4	1	7	**3**	6	0,75	2	3	**6**	1,50	1,12
5	2	8	6	**3**	0,60	3	**6**	3	1,20	0,90
6	1	10	**6**	6	1,00	2	**6**	6	1,00	1,00
7	1	12	**6**	9	0,86	2	6	**9**	1,44	1,15
8	2	13	9	**6**	0,75	3	**9**	6	1,13	0,93
9	1	15	**9**	9	1,00	2	**9**	9	1,00	1,00
10	1	17	**9**	12	0,90	2	9	**12**	1,20	1,05
11	2	18	12	**9**	0,82	3	**12**	9	1,09	0,96
12	1	20	**12**	12	1,00	2	**12**	12	1,00	1,00
...	...	...	...	...	...	...	...	...	...	...

требующей некоторых пояснений.

Описание таблицы

1-й столбец — номер n шага (пары последовательных ходов игроков A и B),

2-й столбец — номер i стратегии, выбранной игроком A,

3-й столбец — «накопленный» суммарный выигрыш игрока A за первые n шагов при стратегии B_1 игрока B,

4-й столбец — «накопленный» суммарный выигрыш игрока A за первые n шагов при стратегии B_2 игрока B,

5-й столбец — «накопленный» суммарный выигрыш игрока A за первые n шагов при стратегии B_3 игрока B,

} минимальный из этих выигрышей выделяется полужирным шрифтом,

6-й столбец — минимальный средний выигрыш игрока A, равный минимальному накопленному им выигрышу за первые n шагов, деленному на число этих шагов,

7-й столбец — номер k стратегии, выбранной игроком B,

8 й столбец — «накопленный» суммарный выигрыш игрока A за первые n шагов при стратегии A_1,

9-й столбец — «накопленный» суммарный выигрыш игрока A за первые n шагов при стратегии A_2,

} максимальный из этих выигрышей выделяется полужирным шрифтом,

10-й столбец — максимальный средний выигрыш игрока A, равный максимальному накопленному им выигрышу за первые n шагов, деленному на число этих шагов,

11-й столбец — среднее арифметическое минимального среднего выигрыша и максимального среднего выигрыша игрока A.

Решение игры определяется приближенно по окончании любого из шагов.

Например, за приближенную цену игры можно взять среднее арифметическое $\nu(n)$, полученное на n-м шаге. Смешанные стратегии противников определяются частотами появления чистых стратегий.

После 9-го шага имеем

$$\nu(9) = 1{,}00.$$

При этом игрок A 6 раз использовал стратегию A_1 и 3 раза стратегию A_2. В свою очередь игрок B 6 раз применял стратегию B_2, 3 раза стратегию B_3, а стратегией B_1 не пользовался вообще. Отсюда получаем, что

$$P_9 = \left\{\frac{6}{9},\ \frac{3}{9}\right\} \approx \{0{,}67;\ 0{,}33\}, \quad Q_9 = \left\{0,\ \frac{6}{9},\ \frac{3}{9}\right\} \approx \{0;\ 0{,}67;\ 0{,}33\}.$$

Соответственно, после 10-го шага получаем —

$$\nu(10) = 1{,}05, \quad P_{10} = \left\{\frac{7}{10},\ \frac{3}{10}\right\} = \{0{,}7;\ 0{,}3\}, \quad Q_{10} = \left\{0,\ \frac{7}{10},\ \frac{3}{10}\right\} = \{0;\ 0{,}7;\ 0{,}3\}.$$

Данная игра легко решается графически. Вот точный ответ:

$$\nu = 1, \quad P = \left\{\frac{2}{3}, \frac{1}{3}\right\}, \quad Q = \left\{0, \frac{2}{3}, \frac{1}{3}\right\}.$$

Сравнивая результаты, полученные на 9-м, 10-м, а также 11-м и 12-м шагах:

$$\nu(11) = 0{,}96, \qquad \nu(12) = 1{,}00,$$

$$P_{11} = \left\{\frac{7}{11}, \frac{4}{11}\right\} \approx \{0{,}64; 0{,}36\}, \qquad P_{12} = \left\{\frac{8}{12}, \frac{4}{12}\right\} \approx \{0{,}67; 0{,}33\},$$

$$Q_{11} = \left\{0, \frac{7}{11}, \frac{4}{11}\right\} \approx \{0; 0{,}64; 0{,}36\}, \qquad Q_{12} = \left\{0, \frac{8}{12}, \frac{4}{12}\right\} \approx \{0; 0{,}67; 0{,}33\},$$

замечаем, что по мере увеличения числа шагов значения все меньше отличаются от точных.

Сделаем несколько замечаний.

Замечание 1. При увеличении числа шагов все три величины $\nu_*(n)$, $\nu^*(n)$ и $\nu(n)$ будут приближаться к цене игры ν, но среднее арифметическое $\nu(n)$ будет приближаться к ν сравнительно быстрее.

Замечание 2. Хотя сходимость итераций весьма медленна, тем не менее, даже такой небольшой расчет всегда дает возможность находить ориентировочное значение цены игры и доли чистых стратегий.

Замечание 3. Сравнительно медленная скорость сходимости можно объяснить целым рядом причин. Укажем одну из них, психологически наиболее интересную. Если, к примеру, игрок A уже получил оптимальную смешанную стратегию, то он не склонен останавливаться на ней. Отнюдь нет — он продолжит попытки выиграть у противника B побольше, особенно если последний еще не достиг оптимальной смешанной стратегии. Тем самым, игрок A может невольно ухудшить свое положение.

Замечание 4. Отметим два основных преимущества описанного метода:

1) итерационный метод прост и одновременно универсален (при его помощи можно легко найти приближенное решение любой матричной игры),

2) объем и сложность вычислений сравнительно слабо растут по мере увеличения числа стратегий игроков (размеров матрицы игры).

3.5. Сведение матричной игры к задаче линейного программирования

Рассмотрим $m \times n$ игру с платежной матрицей

$$\mathsf{A} = (a_{ik}).$$

Без ограничения общности будем считать, что все элементы матрицы A положительны (этого всегда можно добиться, пользуясь аффинным правилом, преобразующим заданную матрицу игры, но не изменяющим оптимальных смешанных стратегий игроков). Тем самым, искомая цена игры ν — положительное число.

Интересы игрока A

Из теоремы о свойствах оптимальных смешанных стратегий игроков вытекает, что при любой чистой стратегии B_k игрока B, $k = 1, 2, \ldots, n$, оптимальная смешанная стратегия $P = \{p_1, p_2, \ldots, p_m\}$ игрока A обеспечивает его средний выигрыш, не меньший ν. Иными словами, выполняются соотношения

$$\sum_{i=1}^{m} a_{ik} p_i \geqslant \nu, \quad k = 1, 2, \ldots, n, \quad \sum_{i=1}^{m} p_i = 1, \quad p_i \geqslant 0, \quad i = 1, 2, \ldots, m,$$

которые с учетом обозначений

$$x_i = \frac{p_i}{\nu}, \quad i = 1, 2, \ldots, m,$$

можно записать так

$$\boxed{\sum_{i=1}^{m} a_{ik} x_i \geqslant 1, \; k = 1, 2, \ldots, n, \quad \sum_{i=1}^{m} x_i = \frac{1}{\nu}, \; x_i \geqslant 0, \quad i = 1, 2, \ldots, m.}$$

Поскольку игрок A стремится сделать свой гарантированный выигрыш максимально возможным, то задача отыскания решения матричной игры сводится к следующей задаче:

найти неотрицательные величины $x_1, x_2, \dots, x_m$, удовлетворяющие неравенствам

$$\sum_{i=1}^{m} a_{ik}x_i \geqslant 1, \quad k = 1, 2, \dots, n,$$

и такие, что их сумма минимальна

$$\sum_{i=1}^{m} x_i \to \min.$$

Интересы игрока B

Аналогичным образом заключаем, что оптимальная смешанная стратегия $Q = \{q_1, q_2, \dots, q_n\}$ игрока B при любой чистой стратегии A_i игрока A, $i = 1, 2, \dots, m$, обеспечивает его средний проигрыш, не больший ν. Иными словами, выполняются соотношения

$$\sum_{k=1}^{n} a_{ik}q_k \leqslant \nu, \quad i = 1, 2, \dots, m, \quad \sum_{k=1}^{n} q_k = 1, \quad q_k \geqslant 0, \quad k = 1, 2, \dots, n,$$

которые с учетом обозначений

$$y_k = \frac{q_k}{\nu}, \quad k = 1, 2, \dots, n,$$

можно записать так

$$\boxed{\sum_{k=1}^{n} a_{ik}y_k \leqslant 1, \quad i = 1, 2, \dots, m, \quad \sum_{k=1}^{n} y_k = \frac{1}{\nu}, \quad y_k \geqslant 0, \quad k = 1, 2, \dots, n.}$$

Поскольку игрок B стремится сделать свой гарантированный проигрыш минимально возможным, то задача отыскания решения матричной игры сводится к следующей задаче:

найти неотрицательные величины $y_1, y_2, \dots, y_n$, удовлетворяющие неравенствам

$$\sum_{k=1}^{n} a_{ik}y_k \leqslant 1, \quad i = 1, 2, \dots, m,$$

и такие, что их сумма максимальна

$$\sum_{k=1}^{n} y_k \to \max.$$

Тем самым, мы получаем следующий важный результат.

Теорема 3. *Решение матричной игры с положительной платежной матрицей (a_{ik}) равносильно решению двойственных задач линейного программирования*

$$(A) \qquad \sum_{i=1}^{m} x_i \to \min, \quad \sum_{i=1}^{m} a_{ik}x_i \geqslant 1, \ k = 1, 2, \dots, n, \quad x_i \geqslant 0, \ i = 1, 2, \dots, m,$$

$$(B) \qquad \sum_{k=1}^{n} y_k \to \max, \quad \sum_{k=1}^{n} a_{ik}y_k \leqslant 1, \ i = 1, 2, \dots, m, \quad y_k \geqslant 0, \ k = 1, 2, \dots, n.$$

При этом цена игры

$$\boxed{\nu = \frac{1}{\Theta},}$$

где Θ — величина, обратная общему значению оптимальных сумм,

$$\boxed{\Theta = \sum_{i=1}^{m} x_i^0 = \sum_{k=1}^{n} y_k^0,}$$

а оптимальные значения p_i^0 и q_k^0 связаны с оптимальными x_i^0 и y_k^0 посредством равенств

$$\boxed{p_i^0 = \frac{x_i^0}{\Theta}, \quad i = 1, 2, \dots, m, \qquad q_k^0 = \frac{y_k^0}{\Theta}, \quad k = 1, 2, \dots, n.}$$

Алгоритм решения матричной игры

1-й шаг. Ко всем элементам исходной матрицы игры прибавляется одно и то же положительное число γ так, чтобы все элементы новой матрицы были строго положительны.

2-й шаг. Решаются двойственные задачи линейного программирования (A) и (B) (например, симплекс-методом, или как-нибудь иначе). Находятся наборы x_i^0, y_k^0 и число Θ.

3-й шаг. Строятся оптимальные смешанные стратегии игроков A и B соответственно

$$p_i^0 = \frac{x_i^0}{\Theta}, \qquad q_k^0 = \frac{y_k^0}{\Theta}.$$

4-й шаг. Вычисляется цена игры

$$\nu_A = \frac{1}{\Theta} - \gamma.$$

Пример 9. Рассмотрим 2×2 игру с матрицей

$$\begin{pmatrix} 5 & 1 \\ 3 & 4 \end{pmatrix}.$$

Соответствующие ей задачи линейного программирования имеют вид

$$\begin{array}{ll} x_1 + x_2 \to \min & \\ 5x_1 + 3x_2 \geqslant 1, & x_1 \geqslant 0, \\ x_1 + 4x_2 \geqslant 1, & x_2 \geqslant 0, \end{array} \qquad \begin{array}{ll} y_1 + y_2 \to \max & \\ 5y_1 + y_2 \leqslant 1, & y_1 \geqslant 0, \\ 3y_1 + 4y_2 \leqslant 1, & y_2 \geqslant 0. \end{array}$$

Решение.

1-й шаг. Все элементы платежной матрицы положительны.

2-й шаг. Строим решения обеих задач линейного программирования, пользуясь графическим методом. В результате получаем, что

$$x_1^0 = \frac{1}{17}, \quad x_2^0 = \frac{4}{17}, \quad y_1^0 = \frac{3}{17}, \quad y_2^0 = \frac{2}{17}, \quad \Theta = \frac{5}{17}.$$

3-й шаг.

$$p_1^0 = \frac{1}{5}, \quad p_2^0 = \frac{4}{5}, \quad q_1^0 = \frac{3}{5}, \quad q_2^0 = \frac{2}{5}.$$

4-й шаг.

$$\nu_A = \frac{17}{5} = 3\frac{2}{5}.$$

§4. Примеры задач, сводимых к матричным играм

В чистом виде антагонистические конфликты встречаются редко (разве только в боевых действиях и в спортивных состязаниях). Однако довольно часто конфликты, в которых интересы сторон противоположны, при допущении, что множество способов действия сторон конечно, можно моделировать матричными играми.

Рассмотрим несколько конкретных ситуаций.

Пример 10. «Планирование посева». Сельскохозяйственное предприятие имеет возможность выращивать две культуры — A_1 и A_2. Необходимо определить, как сеять эти культуры, если при прочих равных условиях их урожаи зависят от погоды, а план посева должен обеспечить наибольший доход (прибыль от реализации выращенной культуры определяется полученным объемом). В зоне рискованного земледелия (а таковой является большая часть России) планирование посева должно осуществляться с учетом наименее благоприятного состояния погоды.

Таким образом, одной из сторон выступает сельскохозяйственное предприятие, заинтересованное в том, чтобы получить наибольший доход (игрок A), а другой стороной — природа, способная навредить сельскохозяйственному предприятию в максимальной степени (от нее зависят погодные условия) и преследующая тем самым прямо противоположные цели (игрок B).

Принятие природы за противника равносильно планированию посева с учетом наиболее неблагоприятных условий; если же погодные условия окажутся благоприятными, то выбранный план даст возможность увеличить доход.

Налицо антагонистический конфликт, в котором у игрока A две стратегии — A_1 и A_2, а у игрока B три — B_1 (засушливое лето), B_2 (нормальное лето) и B_3 (дождливое лето).

В качестве выигрыша игрока A возьмем прибыль от реализации и будем считать, что расчеты прибыли сельскохозяйственного предприятия (в млрд руб.) в зависимости от состояний погоды сведены в следующую матрицу

$$\begin{pmatrix} 8 & 5 & 3 \\ 2 & 3 & 6 \end{pmatrix}.$$

Нетрудно заметить, что седловой точки у этой матрицы нет. Поэтому оптимальная стратегия игрока A будет смешанной. Применяя графический метод, получаем

$$P = \left\{\frac{3}{5}, \frac{2}{5}\right\}, \qquad \nu = 4\frac{1}{5}.$$

Замечание. Здесь мы столкнулись со сравнительно редкой ситуацией, когда оптимальная смешанная стратегия одного из игроков допускает так называемую «физическую» реализацию. Полученное решение сельскохозяйственное предприятие может использовать так:

на $\frac{3}{5}$ всех площадей выращивать культуру A_1,

на $\frac{2}{5}$ всех площадей выращивать культуру A_2

и получать прибыль в размере, не меньшем $4\frac{1}{5}$ млрд руб.

Пример 11. «Переговоры о заключении контракта между профсоюзом и администрацией». Рассмотрим фирму, администрация которой ведет переговоры с профсоюзом рабочих и служащих о заключении контракта.

Предположим, что платежная матрица, отражающая интересы договаривающихся сторон, имеет следующий вид

$$\begin{pmatrix} 75 & 105 & 65 & 45 \\ 70 & 60 & 55 & 40 \\ 80 & 90 & 35 & 50 \\ 95 & 100 & 50 & 55 \end{pmatrix}.$$

Выплаты указаны в центах в час и представляют собой среднюю зарплату служащего фирмы вместе со всеми добавками. Тем самым, заданная матрица описывает прибыль профсоюза (игрок A) и затраты администрации фирмы (игрок B).

Ясно, что профсоюз стремится максимизировать доходы рабочих и служащих, в то время как администрации хотелось бы минимизировать собственные потери.

Нетрудно заметить, что седловой точки у платежной матрицы нет. Кроме того, для дальнейшего анализа существенными являются лишь стратегии A_1 и A_4 игрока A и стратегии B_3 и B_4 игрока B (в этом нетрудно убедиться, воспользовавшись правилом доминирования стратегий). В результате соответствующего усечения получим матрицу

$$\begin{pmatrix} 65 & 45 \\ 50 & 55 \end{pmatrix}.$$

Элементы матрицы

$$\begin{pmatrix} 4 & 0 \\ 1 & 2 \end{pmatrix},$$

связаны с элементами предыдущей матрицы соотношениями

$$65 = 5 \cdot 4 + 45, \quad 45 = 5 \cdot 4 + 45, \quad 50 = 5 \cdot 1 + 45, \quad 55 = 5 \cdot 2 + 45.$$

Воспользовавшись графическим методом, в итоге получим

$$P = \left\{\frac{1}{5}, 0, 0, \frac{4}{5}\right\}, \quad Q = \left\{0, 0, \frac{2}{5}, \frac{3}{5}\right\}, \quad \nu = 53.$$

Тем самым, профсоюзу следует выбирать стратегию A_1 в 20 % случаев и стратегию A_4 в 80 %. Что касается администрации, то ей следует выбирать стратегию B_3 с вероятностью 0,4 и стратегию B_4 с вероятностью 0,6. При этом ожидаемая цена игры равна 53.

Замечание. Следует отметить, что если процесс переговоров будет повторяться много раз, то среднему должно сходиться к ожидаемому значению 53. Если же переговоры пройдут лишь единожды, то реальный результат получится при выборе каждым игроком некоторой своей чистой стратегии. Поэтому один из игроков, профсоюз или администрация, будет неудовлетворен.

Пример 12. «Локальный конфликт». Рассмотрим войну между двумя небольшими государствами A и B, которая ведется в течение 30 дней.

Для бомбардировки небольшого моста — важного военного объекта страны B — страна A использует оба имеющихся у нее самолета. Разрушенный мост восстанавливается в течение суток, а каждый самолет совершает один полет в день по одному из двух воздушных маршрутов, соединяющих эти страны. У страны B есть два зенитных орудия, при помощи которых можно сбивать самолеты страны A. Если самолет собьют, то некая третья страна в течение суток поставит стране A новый самолет.

Страна A может послать самолеты либо по одному маршруту, либо по разным. Страна B может поместить либо обе зенитки на одном маршруте, либо по одной зенитке на каждый маршрут.

Если один самолет летит по маршруту, на котором расположена одна зенитка, то этот самолет будет сбит. Если два самолета летят по маршруту, на котором расположены две зенитки, то оба самолета будут сбиты. Если два самолета летят по маршруту, на котором расположена одна зенитка, то сбит будет только один самолет. Если самолет доберется до цели, то мост будет уничтожен.

У страны A есть две стратегии:

послать самолеты по разным маршрутам — A_1,
послать самолеты по одному маршруту — A_2.

У страны B — также две стратегии:

поместить зенитки на разных маршрутах — B_1,
поместить зенитки на одном маршруте — B_2.

Если страна A выберет стратегию A_1, а страна B — стратегию B_1, то страна A получит нулевой выигрыш, так как ни один из самолетов не достигнет цели.

Если страна A выберет стратегию A_2, а страна B — стратегию B_1, то хотя бы один самолет достигнет цели и вероятность разрушения моста будет равна 1.

Если страна A выберет стратегию A_1, а страна B — стратегию B_2, то вновь хотя бы один самолет достигнет цели и вероятность разрушения моста будет равна 1.

Если страна A выберет стратегию A_2, а страна B — стратегию B_2, то страна A с вероятностью $1/2$ выберет маршрут, на котором установлены зенитки, и, следовательно, цель будет уничтожена с вероятностью $1/2$.

Оформим результаты проведенного анализа в стандартной игровой форме:

$$\begin{pmatrix} 0 & 1 \\ 1 & 1/2 \end{pmatrix}.$$

При помощи графического метода получаем оптимальные смешанные стратегии игроков и цену игры

$$P = \left\{\frac{1}{3}, \frac{2}{3}\right\}, \quad Q = \left\{\frac{1}{3}, \frac{2}{3}\right\}, \quad \nu = \frac{2}{3}.$$

Это означает, что если страна A будет посылать самолеты по разным маршрутам в течение десяти дней из тридцати, отпущенных на войну (и, значит, по одному маршруту в течение двадцати дней), то в среднем страна A будет иметь 66,7 % удачных случаев (мост будем находиться в нерабочем состоянии). Воспользовавшись для своих зениток предложенным выбором, страна B не позволит бомбить мост чаще, чем в 66,7 % случаев.

§ 5. Несколько слов в заключение

Матричные игры моделируют конфликтные ситуации, в которых каждая из сторон-участниц делает свой ход одновременно со второй стороной. При этом наибольший интерес представляет случай, когда игра не заканчивается сразу же после совершения игроками одной такой пары одновременных ходов, а повторяется многократно. Причем считается, что перед каждым возобновлением игры игроки не получают никаких *новых* сведений ни о конфликте, ни о возможных действиях противной стороны. Иными словами, при многократном повторении матричной игры каждая из сторон всякий раз оказывается перед выбором некоторой стратегии из одного и того же множества стратегий, неизменного у каждого из игроков.

Тем не менее, в таких многократно повторяющихся обстоятельствах большую роль играет анализ игры, как предварительный, так и промежуточный.

В результате разумно проведенного предварительного анализа матричной игры заинтересованная в анализе сторона может определить свою линию поведения (правило выбора стратегий) на всю серию игр. Разумеется, описанный нами выше максиминный подход является далеко не единственным средством. Однако не следует забывать, что принципиальной особенностью этого подхода является то обстоятельство, что игрок, придерживающийся выводимого на его основе правила выбора стратегий, заранее может довольно точно оценить нетривиальные размеры своего гарантированного выигрыша. Кроме того, максиминный подход позволяет сводить задачу поиска решения игры к рассмотрению сравнительно несложных задач линейного программирования и, тем самым, получать эффективные рекомендации по тому, как лучше выбирать стратегии в конкретной игре при многократном ее повторении.

Если игра повторяется много раз, то некоторые дополнительные сведения — какие именно стратегии выбирает противная сторона и какими правилами выбора стратегий она руководствуется — игрок все же получает. На основании этих сведений и результатов предварительного анализа игры он может довольно точно оценить противника и, если тот не придерживается компромиссного максиминного подхода, внести соответствующие изменения в собственную линию поведения и увеличить выигрыш.

§ 6. О классификации игр

Реальные конфликтные ситуации приводят к различным видам игр. К настоящему времени общепризнанной классификации игр пока не сложилось. Тем не менее, легко заметить, что игры различаются по целому ряду признаков: по количеству участвующих в них игроков, по количеству возможных стратегий, по характеру взаимоотношений между игроками, по характеру выигрышей, по виду функций выигрышей, по количеству ходов, по характеру информационной обеспеченности игроков и т. д.

Остановимся на этих различиях чуть подробнее.

В зависимости от количества игроков определяют игры трех типов: игры *одного игрока* (в теории игр, как правило, не рассматриваются), игры *двух игроков* (наиболее изученный класс игр) и игры *n игроков* (успехов в изучении которых сравнительно немного вследствие возникающих принципиальных трудностей).

По количеству стратегий игроков игры делятся на *конечные* (каждый из игроков имеет конечное число возможных стратегий) и *бесконечные* (где хотя бы один из игроков имеет бесконечное количество возможных стратегий).

По характеру взаимоотношений между игроками игры делятся на *бескоалиционные* (в которых игроки не имеют права вступать в соглашения и образовывать коалиции), *коалиционные* (в которых игроки могут вступать в соглашения и образовывать коалиции) и *кооперативные* (в которых соглашения, связывающие игроков, определены наперед и обязательны).

По характеру выигрышей различают игры *с нулевой суммой* (общий капитал игроков не изменяется, а просто перераспределяется между игроками в зависимости от получающихся исходов) и игры *с ненулевой суммой*.

Игру двух игроков с нулевой суммой часто называют *антагонистической*, вследствие того, что цели игроков в них прямо противоположны: выигрыш одного из игроков происходит только за счет проигрыша другого.

По виду функций выигрышей игры делятся на: *матричные* игры (рассмотренные выше), *биматричные* игры, игры *типа дуэлей*, *непрерывные* игры, *выпуклые* игры и др.

По количеству ходов игры делятся на *одношаговые* (завершающиеся после одного хода каждого из игроков) и *многошаговые*, которые, в свою очередь, делятся на *позиционные* игры (каждый из игроков может последовательно во времени делать несколько ходов), *стохастические* игры (где при выборе новых позиций имеется определенная вероятность возврата на предшествующую позицию), *дифференциальные* игры (в которых допускается делать ходы непрерывно и подчинять поведение игроков условиями, описываемыми дифференциальными уравнениями), игры *типа дуэлей* (характеризующиеся моментом выбора хода и вероятностями получения выигрышей в зависимости от времени, прошедшего от начала игры до момента выбора).

По характеру информационной обеспеченности игроков различают игры *с полной информацией* (на каждом ходе игры каждому игроку известно, какие выборы были сделаны ранее всеми игроками) и игры *с неполной информацией* (если в игре не все известно о предыдущих выборах).

Существуют, разумеется, и другие виды игр.

В зависимости от вида игры разрабатывается и метод ее решения. Однако читатель, видимо, уже заметил, что структура предложенного выше описания основных направлений различения игр не является древовидной: одни и те же виды игр оказываются в разных классах.

Впрочем, возможны и иные подходы к разбиению игр.

Далее мы рассмотрим некоторые из указанных видов игр, а именно позиционные и биматричные игры.

Глава XLVII

ПОЗИЦИОННЫЕ ИГРЫ

Во многих практически важных конфликтных ситуациях, располагая той или иной информацией об их прошлом развитии, стороны-участницы совершают свой выбор не раз и навсегда, а последовательно во времени, шаг за шагом. Тем самым, они используют стратегии, отражающие как динамику конфликта, так и степень собственной информированности о фактически складывающейся обстановке в развитии этого конфликта.

§ 1. Структура позиционной игры

Одним из классов игр, описывающих конфликты, динамика которых оказывает влияние на поведение участников, являются так называемые позиционные игры.

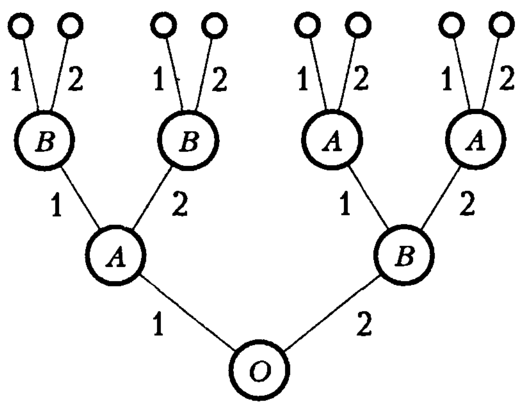


Рис. 1

Позиционная игра — это бескоалиционная игра, моделирующая процессы последовательного принятия решений игроками в условиях меняющейся во времени и, вообще говоря, неполной информации.

Процесс самой игры состоит в последовательном переходе от одного состояния игры к другому состоянию, который осуществляется либо путем выбора игроками одного из возможных действий в соответствии с правилами игры, либо случайным образом (*случайный ход*).

В качестве примеров позиционных игр можно привести крестики-нолики, шашки, шахматы, карточные игры, домино и др. Интересно, что право выбора первого хода в этих играх часто определяется случайным образом.

Состояния игры принято называть *позициями* (отсюда и название — позиционные игры), а возможные выборы в каждой позиции — *альтернативами*.

Характерной особенностью позиционной игры является возможность представления множества позиций в виде древовидного упорядоченного множества, которое называется *деревом игры* (рис. 1).

Для определенности мы будем рассматривать позиционные игры, в каждой позиции которых, кроме окончательных, ровно две альтернативы — первая и вторая.

Замечание. Символ O, A или B в кружке указывает, кто из игроков (O, A или B) делает очередной ход в заданной позиции. При этом символом O обычно обозначается ход в игре, осуществляемый не игроком, а каким-нибудь случайным механизмом (иногда его называют *природой*). Например, в позиционной игре, представленной на рис. 1 своим деревом, первый ход производится случайно.

Пользуясь графическим описанием игры, можно сказать, что процесс игры состоит в переходе от начальной позиции к окончательной через непосредственно следующие одна за другой промежуточные позиции.

Каждая окончательная вершина определяет единственную цепь (последовательность идущих друг за другом звеньев), связывающую начальную вершину с данной (рис. 2). Такая цепь называется *партией*. На рис. 2 одна из партий выделена жирными линиями. Число различных партий равно числу окончательных вершин (позиций).

В каждой окончательной позиции задан числовой выигрыш игрока A.

Замечание. Мы будем рассматривать здесь только антагонистические позиционные игры.

В шахматах функция выигрышей игрока A (белых) определяется так:

$$\begin{aligned} &+1 \quad \text{на выигрываемых партиях,}\\ &\;\;0 \quad \text{на ничейных партиях,}\\ &-1 \quad \text{на проигрываемых партиях.}\end{aligned}$$

Функция выигрышей игрока B (черных) отличается от функции выигрышей белых только знаком.

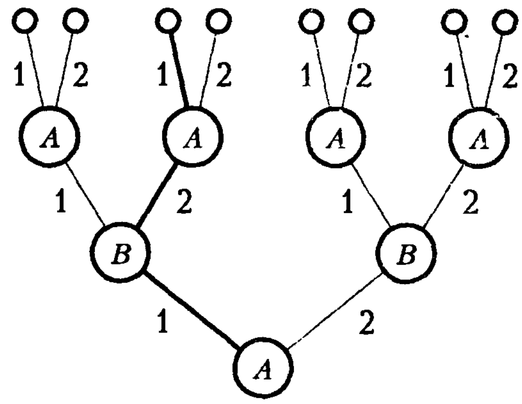


Рис. 2

Различают позиционные игры *с полной информацией* и позиционные игры *с неполной информацией*.

В позиционных играх с полной информацией (пример — шашки, шахматы) каждый игрок при своем ходе знает ту позицию дерева игры, в которой он находится.

В позиционных играх с неполной информацией (пример — домино) игрок, делающий ход, не знает точно, в какой именно позиции дерева игры он фактически находится. Этому игроку известно лишь некоторое множество позиций, включающее в себя его фактическую позицию. Такое множество позиций называется *информационным множеством*.

Таким образом, в игре с неполной информацией игрок при своем ходе знает, в каком информационном множестве он находится, но ему неизвестно, в какой именно позиции этого множества.

Позиции, принадлежащие одному и тому же информационному множеству, объединяются пунктирными линиями.

Рассмотрим примеры двух игр, состоящих из двух ходов, которые последовательно делают участвующие в ней игроки A и B. Начинает игрок A: он выбирает одну из двух возможных альтернатив — число x, равное либо 1 (первая альтернатива), либо 2 (вторая альтернатива). На ход игрока A игрок B отвечает своим ходом, выбирая одну из двух возможных альтернатив — число y, равное либо 1 (первая альтернатива), либо 2 (вторая альтернатива).

И в результате игрок A получает вознаграждение или вынужден платить штраф.

Пример 13.

1-й ход. Игрок A выбирает число x из множества двух чисел $\{1, 2\}$.

2-й ход. Игрок B выбирает число y из множества двух чисел $\{1, 2\}$, зная выбор числа x игроком A.

Функция $W(x, y)$ выплат игроку A за счет игрока B задается так

$$W(1,1) = 1, \qquad W(2,1) = -2,$$
$$W(1,2) = -1, \qquad W(2,2) = 2.$$

На рис. 3 показаны дерево игры и информационные множества.

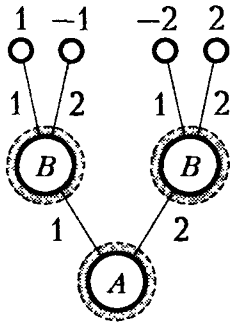


Рис. 3

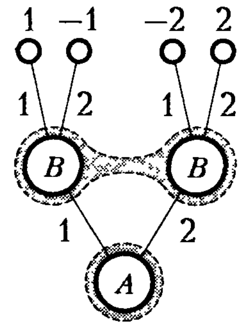


Рис. 4

Пример 14. В случае, если выполнены все условия предыдущего примера, кроме одного — хода игрока B,

2-й ход — игрок B выбирает число y из множества двух чисел $\{1, 2\}$, не зная выбора числа x игроком A, информационные множества выглядят так, как показано на рис. 4.

§ 2. Нормализация позиционной игры

Заранее определенную последовательность ходов игрока, выбранную им в зависимости от информации о ходах другого игрока и ходах игрока O (природы), будем называть *чистой стратегией* этого игрока.

В том случае, если в игре нет случайных ходов (игрок O в игре не участвует), выбор игроком A и игроком B чистых стратегий однозначно определяет исход игры — приводит к окончательной позиции, где игрок A и получает свой выигрыш. Это обстоятельство позволяет сводить позиционную игру к матричной игре.

Процесс сведения позиционной игры к матричной называется *нормализацией позиционной игры*.

Покажем на нескольких примерах, как это делается.

Пример 13 (продолжение). Опишем стратегии игроков.

Стратегию игрока A можно задать одним числом x, показывающим, какую альтернативу, первую или вторую, выбрал игрок.

Тем самым, у игрока A две чистых стратегии:

$$A_1 \text{ — выбрать } x = 1, \quad A_2 \text{ — выбрать } x = 2.$$

Стратегию игрока B, принимая во внимание, что выбор игрока A на 1-м ходе ему известен, удобно описывать упорядоченной парой

$$[y_1, y_2].$$

Здесь y_1 ($y_1 = 1, 2$) — альтернатива, выбираемая игроком B при условии, что игрок A выбрал первую альтернативу, $x = 1$, а y_2 ($y_2 = 1, 2$) — альтернатива, выбираемая игроком B при условии, что игрок A выбрал вторую альтернативу, $x = 2$.

Например, выбор игроком B стратегии $[2, 1]$ означает, что если на 1-м ходе игрок A выбрал $x = 1$, то игрок B на своем ходе должен выбрать $y = 2$. Если же на 1-м ходе игрок A выбрал $x = 2$, то согласно этой стратегии игрок B на своем ходе должен выбрать $y = 1$.

Таким образом, у игрока B четыре чистых стратегии:

$$B_1 \text{ — } [1, 1], \quad y = 1 \text{ при любом выборе } x;$$

$$B_2 \text{ — } [1, 2], \quad y = x \text{ при любом выборе } x;$$

$$B_3 \text{ — } [2, 1], \quad y \neq x \text{ при любом выборе } x;$$

$$B_4 \text{ — } [2, 2], \quad y = 2 \text{ при любом выборе } x.$$

Покажем теперь, как рассчитываются выигрыши игрока A в зависимости от примененных стратегий.

Пусть, например, игрок A выбрал стратегию A_1 — (1), а игрок B — стратегию B_2 — $[1, 2]$. Тогда $x = 1$, а из стратегии $[1, 2]$ вытекает, что $y = 1$. Отсюда

$$W(x, y) = W(1, 1) = 1.$$

Остальные выигрыши рассчитываются совершенно аналогично.

Результаты расчетов записываются обычно или в виде таблицы выигрышей игрока A

		B_1	B_2	B_3	B_4
		[1, 1]	[1, 2]	[2, 1]	[2, 2]
A_1	$x = 1$	$W(1, 1)$	$W(1, 1)$	$W(1, 2)$	$W(1, 2)$
A_2	$x = 2$	$W(2, 1)$	$W(2, 2)$	$W(2, 1)$	$W(2, 2)$

или в виде матрицы игры

$$\begin{pmatrix} 1 & 1 & -1 & -1 \\ -2 & 2 & -2 & 2 \end{pmatrix},$$

где, как обычно, строки соответствуют стратегиям игрока A, а столбцы — стратегиям игрока B.

Полученная матрица имеет седловую точку. Оптимальные стратегии игроков: A_1 — (1) и B_3 — [2, 1]. Тем самым, игрок A на 1-м ходе выбирает $x = 1$, а игрок B на 2-м ходе выбирает $y = 2$. Цена игры $\nu = -1$.

Пример 14 (продолжение). Опишем стратегии игроков.

У игрока A они те же, что и в предыдущем примере:

$$A_1 \text{ — выбрать } x = 1, \quad A_2 \text{ — выбрать } x = 2.$$

Так как игроку B выбор игрока A неизвестен, то есть игрок B не знает, в какой именно из двух позиций он находится (см. рис. 4), то у него те же две стратегии:

$$B_1 \text{ — выбрать } y = 1, \quad B_2 \text{ — выбрать } y = 2.$$

Соответствующие таблица выигрышей игрока A и матрица игры имеют следующий вид

		B_1	B_2
		$y = 1$	$y = 2$
A_1	$x = 1$	$W(1, 1)$	$W(1, 2)$
A_2	$x = 2$	$W(2, 1)$	$W(2, 2)$

$$\begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix}.$$

Полученная матрица седловой точки не имеет. Оптимальные смешанные стратегии игроков: $P = \{2/3, 1/3\}$ и $Q = \{1/2, 1/2\}$. Цена игры $\nu = 0$.

Замечание 1. На этих двух примерах хорошо видно, что результат сведения позиционной игры к матричной напрямую зависит от степени информированности игроков. В частности, отсутствие у игрока B сведений о выборе, сделанном игроком A, приводит к уменьшению количества его возможных стратегий. Сравнивая ответы, полученные в примерах 13 и 14, замечаем, что снижение уровня информированности игрока (в данном случае — игрока B) делает для него исход игры менее благоприятным.

Замечание 2. Приведенные выше примеры не исчерпывают всех возможных вариантов даже в этом, самом простом, случае двухходовых позиционных игр.

Рассмотрим теперь несколько примеров сведения к матричным играм позиционных игр, состоящих из трех ходов, сосредоточив при этом основное внимание на одном из наиболее ответственных шагов нормализации — описании стратегий игроков.

Пример 15.

1-й ход делает игрок A: он выбирает число x из множества двух чисел $\{1, 2\}$.

2-й ход делает игрок B: зная выбранное игроком A число x, он выбирает число y из множества двух чисел $\{1, 2\}$.

3-й ход делает игрок A: не зная о выбранном игроке B числе y на 2-м ходе и забыв выбранное им самим на 1-м ходе число x, он выбирает число z из множества двух чисел $\{1, 2\}$.

После этого игрок A получает вознаграждение $W(x, y, z)$ за счет игрока B, например, такое:

$$W(1,1,1) = -2, \quad W(2,1,1) = 3,$$
$$W(1,1,2) = 4, \quad W(2,1,2) = 0.$$
$$W(1,2,1) = 1, \quad W(2,2,1) = -3,$$
$$W(1,2,2) = -4, \quad W(2,2,2) = 5.$$

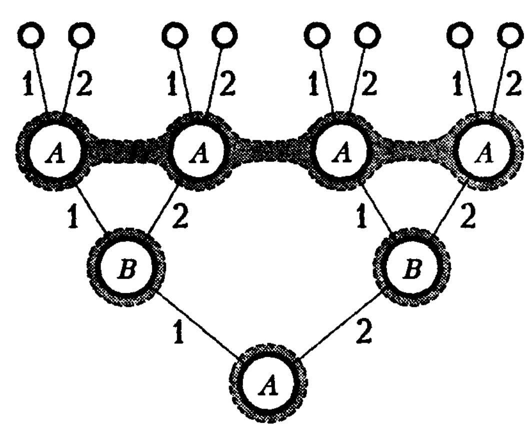


Рис. 5

На рис. 5 показаны дерево игры и информационные множества.

Нормализуем эту игру.

Поскольку игроку B выбор игрока A на 1-м ходе известен, то у игрока B те же четыре стратегии, что и в примере 13:

$$B_1 — [1, 1], \quad B_2 — [1, 2], \quad B_3 — [2, 1], \quad B_4 — [2, 2].$$

Игрок A на 3-м ходе не знает предыдущих выборов — ни значения x, ни значения y. Поэтому каждая его стратегия состоит просто из пары чисел (x, z), где x $(x = 1, 2)$ — альтернатива, выбираемая игроком A на 1-м ходе, а z $(z = 1, 2)$ — альтернатива, выбираемая игроком A на 3-м ходе.

Например, выбор игроком A стратегии $(2, 1)$ означает, что на 1-м ходе он выбирает $x = 2$, а на 3-м ходе — $z = 1$.

Таким образом, у игрока A четыре стратегии:

$$A_1 — (1, 1), \quad A_2 — (1, 2), \quad A_3 — (2, 1), \quad A_4 — (2, 2).$$

Покажем теперь, как рассчитываются выигрыши игрока A в зависимости от стратегий, применяемых в данной игре. Пусть, например, игрок A выбрал стратегию A_2 — $(1, 2)$, а игрок B — стратегию B_3 — $[2, 1]$. Тогда $x = 1$, откуда вытекает, что $y = 2$. Значение $z = 2$ выбрано игроком A независимо от выбора игрока B. Вычисляя значение функции выигрышей для этого набора, получаем

$$W(x, y, z) = W(1, 2, 2) = -4.$$

В результате подобных рассуждений получаются и остальные пятнадцать выигрышей. Это позволяет построить таблицу выигрышей игрока A. Имеем

		B_1	B_2	B_3	B_4
		$[1, 1]$	$[1, 2]$	$[2, 1]$	$[2, 2]$
A_1	$(1, 1)$	$W(1, 1, 1)$	$W(1, 1, 1)$	$W(1, 2, 1)$	$W(1, 2, 1)$
A_2	$(1, 2)$	$W(1, 1, 2)$	$W(1, 1, 2)$	$W(1, 2, 2)$	$W(1, 2, 2)$
A_3	$(2, 1)$	$W(2, 1, 1)$	$W(2, 2, 1)$	$W(2, 1, 1)$	$W(2, 2, 1)$
A_4	$(2, 2)$	$W(2, 1, 2)$	$W(2, 2, 2)$	$W(2, 1, 2)$	$W(2, 2, 2)$

или

$$\begin{pmatrix} -2 & -2 & 1 & 1 \\ 4 & 4 & -4 & -4 \\ 3 & -3 & 3 & -3 \\ 0 & 5 & 0 & 5 \end{pmatrix}.$$

Пример 16.

1-й ход делает игрок A: он выбирает число x из множества двух чисел $\{1, 2\}$.

2-й ход делает игрок B: не зная о выборе игрока A на 1-м ходе, он выбирает число y из множества двух чисел $\{1, 2\}$.

3-й ход делает игрок A: он выбирает число z из множества двух чисел $\{1, 2\}$, не зная ни значения x, ни значения y.

После этого игроки расплачиваются по правилу, указанному в примере 15.

Графическое представление этой игры показано на рис. 6.

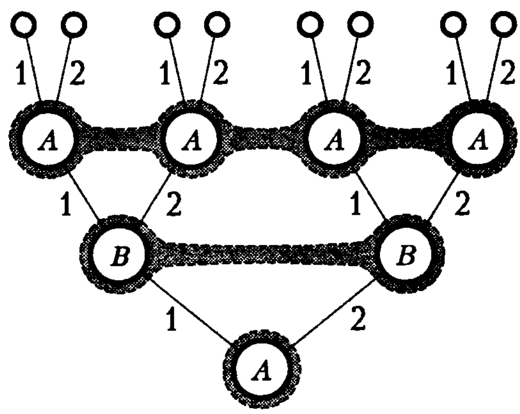


Рис. 6

Ясно, что у игрока A те же четыре стратегии, что и в примере 15:

$$A_1 — (1, 1), \quad A_2 — (1, 2), \quad A_3 — (2, 1), \quad A_4 — (2, 2).$$

У игрока B всего две стратегии:

$$B_1 — \text{выбрать } y = 1, \quad B_2 — \text{выбрать } y = 2.$$

В этом случае (весьма слабой информированности игроков) таблица выигрышей игрока A и соответствующая матрица строятся совсем просто. Имеем

		B_1	B_2
		$y=1$	$y=2$
A_1	(1, 1)	$W(1,1,1)$	$W(1,2,1)$
A_2	(1, 2)	$W(1,1,2)$	$W(1,2,2)$
A_3	(2, 1)	$W(2,1,1)$	$W(2,2,1)$
A_4	(2, 2)	$W(2,1,2)$	$W(2,2,2)$

$$\begin{pmatrix} -2 & 1 \\ 4 & -4 \\ 3 & -3 \\ 0 & 5 \end{pmatrix}.$$

Оптимальные смешанные стратегии игроков и цена игры соответственно равны.

$$P = \left\{0, \frac{9}{13}, 0, \frac{4}{13}\right\}, \quad Q = \left\{\frac{5}{13}, \frac{8}{13}\right\}, \quad \nu = \frac{20}{13}.$$

В следующем примере информационные множества выглядят немного иначе.

Пример 17.

1-й ход делает игрок A: он выбирает число x из множества двух чисел $\{1, 2\}$.

2-й ход делает игрок B: не зная о выборе игрока A на 1-м ходе, он выбирает число y из множества двух чисел $\{1, 2\}$.

3-й ход делает игрок A: он выбирает число z из множества двух чисел $\{1, 2\}$, зная выбор y игрока B на 2-м ходе, но не помня собственного выбора x на 1-м ходе.

После этого игроки расплачиваются по правилу, указанному в примере 15.

Графическое представление этой игры показано на рис. 7.

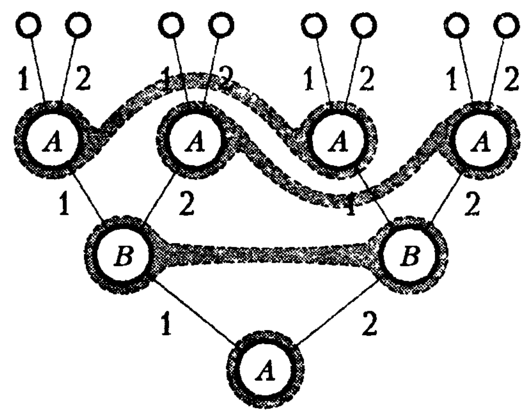


Рис. 7

Поскольку игроку B неизвестен выбор игрока A на 1-м ходе, то, выполняя свой ход, он не знает, в какой именно из двух возможных позиций он находится. Поэтому у игрока B всего две стратегии:

$$B_1 \text{ — выбрать } y = 1, \quad B_2 \text{ — выбрать } y = 2.$$

При описании стратегий игрока A нужно исходить из того, что к 3-му ходу игрок A утрвтил сведения о собственном выборе на 1-м ходе, но ему известен выбор игрока B на 2-м ходе. Поэтому выбор числа z игроку A следует связать с известным ему к 3-му ходу значением y. Удобнее всего это сделать по аналогии с расчетом стратегий игрока B в примерах 13 и 15, т. е. при помощи упорядоченной пары

$$[z_1, z_2].$$

Здесь z_1 $(z_1 = 1, 2)$ — альтернатива, выбираемая игроком A при условии, что игрок B выбрал первую альтернативу, $y = 1$, а z_2 $(z_2 = 1, 2)$ — альтернатива, выбираемая игроком A при условии, что игрок B выбрал вторую альтернативу, $y = 2$.

Чистую стратегию игрока A в данной игре можно записать так

$$(x, [z_1, z_2]).$$

Здесь x $(x = 1, 2)$ — альтернатива, которую игрок A выбирает на 1-м ходе, z_1 $(z_1 = 1, 2)$ — альтернатива, которую игрок A выбирает на 3-м ходе, если на 2-м ходе игрок B выбрал выбрал первую альтернативу $(y = 1)$ и z_2 $(z_2 = 1, 2)$ — альтернатива, которую игрок A выбирает на 3-м ходе, если на 2-м ходе игрок B выбрал вторую альтернативу $(y = 2)$.

Например, выбор игроком A стратегии $(2, [2, 1])$ означает, что на 1-м ходе игрок A выбирает $x = 2$, а на 3-м $z = 2$, если игрок B выбрал $y = 1$, и $z = 1$, если игрок B выбрал $y = 2$.

Тем самым, у игрока A восемь чистых стратегий:

$$A_1 - (1, [1, 1]), \quad A_2 - (1, [1, 2]), \quad A_3 - (1, [2, 1]), \quad A_4 - (1, [2, 2]),$$
$$A_5 - (2, [1, 1]), \quad A_6 - (2, [1, 2]), \quad A_7 - (2, [2, 1]), \quad A_8 - (2, [2, 2]).$$

Покажем теперь, как в зависимости от применяемых стратегий определяются элементы таблицы выигрышей игрока A.

Пусть, например, игрок A выбрал стратегию A_3 — $(1, [2, 1])$, а игрок B — стратегию B_2 — (2) Тогда $x = 1$, $y = 2$, а из $[2, 1]$ вытекает, что $z = 1$. Отсюда

$$W(x, y, z) = W(1, 2, 1) = 1.$$

По этой же схеме вычисляются и остальные элементы таблицы.

В результате получаем

		B_1	B_2
		(1)	(2)
A_1	(1, [1, 1])	$W(1, 1, 1)$	$W(1, 2, 1)$
A_2	(1, [1, 2])	$W(1, 1, 1)$	$W(1, 2, 2)$
A_3	(1, [2, 1])	$W(1, 1, 2)$	$W(1, 2, 1)$
A_4	(1, [2, 2])	$W(1, 1, 2)$	$W(1, 2, 2)$
A_5	(2, [1, 1])	$W(2, 1, 1)$	$W(2, 2, 1)$
A_6	(2, [1, 2])	$W(2, 1, 1)$	$W(2, 2, 2)$
A_7	(2, [2, 1])	$W(2, 1, 2)$	$W(2, 2, 1)$
A_8	(2, [2, 2])	$W(2, 1, 2)$	$W(2, 2, 2)$

$$\begin{pmatrix} -2 & 1 \\ -2 & -4 \\ 4 & 1 \\ 4 & -4 \\ 3 & -3 \\ 3 & 5 \\ 0 & -3 \\ 0 & 5 \end{pmatrix}.$$

Оптимальные смешанные стратегии игроков и цена игры соответственно равны:

$$P = \left\{0, 0, \tfrac{2}{5}, 0, 0, \tfrac{3}{5}, 0, 0\right\}, \quad Q = \left\{\tfrac{4}{5}, \tfrac{1}{5}\right\}, \quad \nu = \tfrac{17}{5}.$$

Рассмотрим позиционную игру со случайным ходом.

Пример 18.

1-й ход производится случайно: игрок O выбирает число x, равное 1, с вероятностью 0,5 и равное 2 с такой же вероятностью.

2-й ход делает игрок A: он выбирает число y из множества двух чисел $\{1, 2\}$, не зная результатов случайного выбора на 1-м ходе.

3-й ход делает игрок B: он выбирает число z из множества двух чисел $\{1, 2\}$, зная о том, какое именно число x случайно выбрано игроком O на 1-м ходе и не зная выбора y игрока A на 2-м ходе.

После этого игроки расплачиваются, используя функцию $W(x, y, z)$, ту же, что и в предыдущих примерах.

Графическое представление этой игры показано на рис. 8.

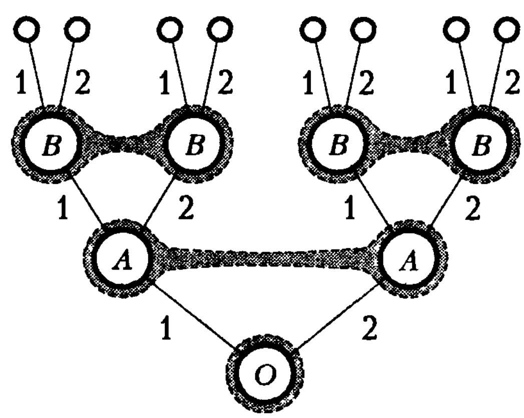


Рис. 8

Опишем стратегии игроков

Поскольку игроку A исход случайного испытания неизвестен, то он имеет всего две стратегии:

$$A_1 — (1), \quad A_2 — (2).$$

При построении своих стратегий игроку B естественно воспользоваться имеющейся у него информацией о результате 1-го хода. Это позволит ему описать свою стратегию упорядоченной парой

$$[z_1, z_2].$$

Здесь z_1 ($z_1 = 1, 2$) — альтернатива, выбираемая игроком B при условии, что $x = 1$, а z_2 ($z_2 = 1, 2$) — альтернатива, выбираемая игроком B при условии, что $x = 2$.

Тем самым, у игрока B четыре стратегии:

$$B_1 — [1, 1], \quad B_2 — [1, 2], \quad B_3 — [2, 1], \quad B_4 — [2, 2].$$

Покажем теперь, как определяются элементы таблицы выигрышей игрока A.

Пусть, например, игрок A выбрал стратегию A_1 — (1), а игрок B — стратегию B_3 — [2, 1].

Различаются два случая

$$1)\ x = 1 \quad \text{и} \quad 2)\ x = 2.$$

Если $x = 1$, то стратегия B_3 указывает игроку B его выбор $z = 2$. А так как $y = 1$, то в результате имеем

$$W(x, y, z) = W(1, 1, 2) = 4.$$

Если $x = 2$, то стратегия B_3 указывает игроку B его выбор $z = 1$. А так как $y = 1$, то в результате имеем

$$W(x, y, z) = W(2, 1, 1) = 3.$$

Поскольку первая и вторая альтернативы на 1-м ходе выбираются с вероятностями 0,5 и 0,5, то и вышеуказанные выигрыши появляются с теми же вероятностями и, следовательно, средний выигрыш игрока A при этих стратегиях определяется так

$$4 \cdot 0{,}5 + 3 \cdot 0{,}5 = 3{,}5.$$

Аналогичным образом рассчитывая остальные средние выигрыши, получаем при $x = 1$

		B_1	B_2	B_3	B_4
		[1, 1]	[1, 2]	[2, 1]	[2, 2]
A_1	(1)	$W(1,1,1)$	$W(1,1,1)$	$W(1,1,2)$	$W(1,1,2)$
A_2	(2)	$W(1,2,1)$	$W(1,2,1)$	$W(1,2,2)$	$W(1,2,2)$

или

$$\begin{pmatrix} -2 & -2 & 4 & 4 \\ 1 & 1 & -4 & -4 \end{pmatrix},$$

при $x = 2$

		B_1	B_2	B_3	B_4
		[1, 1]	[1, 2]	[2, 1]	[2, 2]
A_1	(1)	$W(2,1,1)$	$W(2,1,2)$	$W(2,1,1)$	$W(2,1,2)$
A_2	(2)	$W(2,2,1)$	$W(2,2,2)$	$W(2,2,1)$	$W(2,2,2)$

или

$$\begin{pmatrix} 3 & 0 & 3 & 0 \\ -3 & 5 & -3 & 5 \end{pmatrix}.$$

Искомая матрица игры имеет следующий вид

$$\begin{pmatrix} 0{,}5 & -1 & 3{,}5 & 0 \\ -1 & 3 & -3{,}5 & 0{,}5 \end{pmatrix}.$$

Наконец, рассмотрим пример позиционной игры со случайным разыгрыванием права первого хода.

Пример 19.

1-й ход делает игрок O, выбирая число x, равное 1 с вероятностью $2/3$ и равное 2 с вероятностью $1/3$.

Если $x = 1$, то на **2-м ходе** игрок A выбирает число y из множества двух чисел $\{1, 2\}$, зная результат случайного выбора на 1-м ходе, а на **3-м ходе** игрок B выбирает число z из множества двух чисел $\{1, 2\}$, зная x, но не зная y.

Если $x = 2$, то на **2-м ходе** игрок B выбирает число y из множества двух чисел $\{1, 2\}$, зная результат случайного выбора на 1-м ходе, а на **3-м ходе** игрок A выбирает число z из множества двух чисел $\{1, 2\}$, зная x, но не зная y.

После этого игроки расплачиваются, используя функцию $W(x, y, z)$, ту же, что и в предыдущих примерах.

Графическое представление этой игры показано на рис. 9.

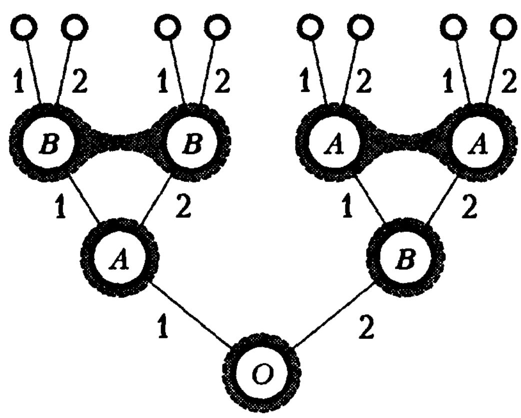


Рис. 9

Чистую стратегию игрока A в данной игре можно описать упорядоченной парой

$$|y, z|,$$

где y ($y = 1, 2$) — выбор игрока A на 2-м ходе, если на 1-м ходе выбрано $x = 1$, а z ($z = 1, 2$) — выбор игрока A на 3-м ходе, если на 1-м ходе выбрано $x = 2$.

Например, стратегия $|1, 2|$ означает, что на 2-м ходе игрок A выбирает $y = 1$, а на 3-м ходе — $z = 2$.

Тем самым, у игрока A четыре стратегии:

$$A_1 - |1, 1|, \quad A_2 - |1, 2|, \quad A_3 - |2, 1|, \quad A_4 - |2, 2|.$$

У игрока B те же четыре стратегии:

$$B_1 - |1, 1|, \quad B_2 - |1, 2|, \quad B_3 - |2, 1|, \quad B_4 - |2, 2|.$$

Покажем теперь, как находятся элементы матрицы выигрышей игрока A.

Пусть, например, игрок A применяет стратегию A_2 — $|1, 2|$, а игрок B — стратегию B_3 — $|2, 1|$. Различаются два случая

$$1)\ x = 1 \qquad \text{и} \qquad 2)\ x = 2.$$

По условию при $x = 1$ игрок A имеет возможность сделать только 2-й ход (выбрать y), а игрок B — только 3-й (выбрать z). При $x = 2$ их возможности меняются местами: игроку B предоставлено право 2-го хода (выбрать y), а игроку A — 3-го (выбрать z).

Если $x = 1$, то стратегия A_2 указывает игроку A при 2-м ходе взять $y = 1$, а стратегия B_3 указывает игроку B при 3-м ходе взять $z = 1$. В результате

$$W(x, y, z) = W(1, 1, 1) = -2.$$

Если $x = 2$, то стратегия B_3 указывает игроку B при 2-м ходе взять $y = 2$, а стратегия A_2 указывает игроку A при 3-м ходе взять $z = 2$. В результате

$$W(x, y, z) = W(2, 2, 2) = 5.$$

Поскольку первая и вторая альтернативы на 1-м ходе выбираются соответственно с вероятностями $2/3$ и $1/3$, то и найденные выигрыши появляются с теми же вероятностями. Следовательно, математическое ожидание выигрыша игрока A при таких стратегиях рассчитывается так

$$-2 \cdot \frac{2}{3} + 5 \cdot \frac{1}{3} = \frac{1}{3}.$$

Итак,

при $x = 1$

		B_1	B_2	B_3	B_4
		$\|1, 1\|$	$\|1, 2\|$	$\|2, 1\|$	$\|2, 2\|$
A_1	$\|1, 1\|$	$W(1, 1, 1)$	$W(1, 1, 2)$	$W(1, 1, 1)$	$W(1, 1, 2)$
A_2	$\|1, 2\|$	$W(1, 1, 1)$	$W(1, 1, 2)$	$W(1, 1, 1)$	$W(1, 1, 2)$
A_3	$\|2, 1\|$	$W(1, 2, 1)$	$W(1, 2, 2)$	$W(1, 2, 1)$	$W(1, 2, 2)$
A_4	$\|2, 2\|$	$W(1, 2, 1)$	$W(1, 2, 2)$	$W(1, 2, 1)$	$W(1, 2, 2)$

или

$$\begin{pmatrix} -2 & 4 & -2 & 4 \\ -2 & 4 & -2 & 4 \\ 1 & -4 & 1 & -4 \\ 1 & -4 & 1 & -4 \end{pmatrix},$$

при $x = 2$

		B_1	B_2	B_3	B_4
		$\|1, 1\|$	$\|1, 2\|$	$\|2, 1\|$	$\|2, 2\|$
A_1	$\|1, 1\|$	$W(2, 1, 1)$	$W(2, 1, 1)$	$W(2, 2, 1)$	$W(2, 2, 1)$
A_2	$\|1, 2\|$	$W(2, 1, 2)$	$W(2, 1, 2)$	$W(2, 2, 2)$	$W(2, 2, 2)$
A_3	$\|2, 1\|$	$W(2, 1, 1)$	$W(2, 1, 1)$	$W(2, 2, 1)$	$W(2, 2, 1)$
A_4	$\|2, 2\|$	$W(2, 1, 2)$	$W(2, 1, 2)$	$W(2, 2, 2)$	$W(2, 2, 2)$

или

$$\begin{pmatrix} 3 & 3 & -3 & -3 \\ 0 & 0 & 5 & 5 \\ 3 & 3 & -3 & -3 \\ 0 & 0 & 5 & 5 \end{pmatrix}.$$

Отсюда получаем искомую матрицу игры

$$\frac{1}{3}\begin{pmatrix} -1 & 11 & -7 & 5 \\ -4 & 8 & 1 & 11 \\ 5 & -5 & -1 & -11 \\ 2 & -8 & 7 & -5 \end{pmatrix}.$$

Замечание. Графическое представление и функция выигрышей полностью определяют позиционную игру. В рассмотренных выше примерах 16-19 мы пользовались одной и той же функцией и одним и тем же деревом. Отличие было только в маркировке вершин дерева и информационных множествах. При построении последних необходимо соблюдать два правила:

1) в одно информационное множество могут входить позиции только одного игрока,

2) цепь, определяющая партию игры, может иметь с информационным множеством не более одной общей позиции.

Как показывает рис. 10, и при таких ограничениях информационные множества могут выглядеть довольно необычно.

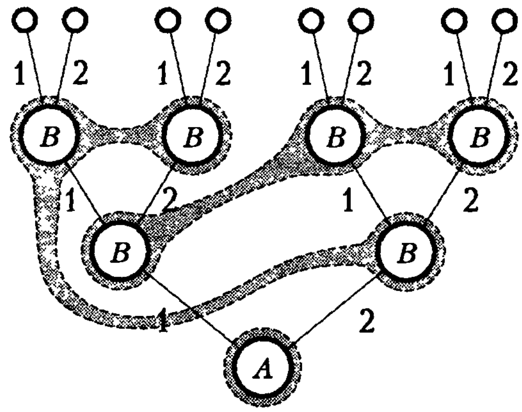


Рис. 10

§ 3. Позиционные игры с полной информацией

Позиционная игра называется *игрой с полной информацией*, если в каждой позиции любой ее партии игрок, делающий ход, знает, какие альтернативы были выбраны на предыдущих ходах. В графическом описании каждая вершина дерева такой игры представляет собой отдельное информационное множество.

Примерами позиционных игр с полной информацией могут служить крестики-нолики, шашки и шахматы.

Основная особенность позиционной игры с полной информацией состоит в том, что соответствующая ей матрица выигрышей всегда имеет седловую точку, то есть в игре с полной информацией существуют оптимальные чистые стратегии и, значит, равновесная ситуация.

Сказанное означает, что в шахматах (крестиках-ноликах, шашках) уже в начальной позиции либо имеется способ выигрыша за белых, либо способ выигрыша за черных, либо как та, так и другая сторона способна форсировать ничью.

Однако известное доказательство существования равновесной ситуации неконструктивно и не дает эффективных приемов фактического нахождения решения игры.

И такие способы (стратегии) в шахматах не найдены до сих пор, и даже неизвестно, какая из перечисленных возможностей имеет место на самом деле.

Иное дело с игрой крестики-нолики: стратегий в ней немного и она разобрана до самого конца — существуют оптимальные чистые стратегии, ведущие игроков к ничьей.

Рассмотрим несколько **примеров.**

1. Как нетрудно заметить, двухходовая игра из примера 11 является игрой с полной информацией. Ее нормализация приводит к матрице с седловой точкой (см. пример 13).

2. **«Выкладывание монет на стол».** Два игрока поочередно кладут монеты одинаковых размеров на обыкновенный стол, всякий раз выбирая произвольное доступное место для монеты (взаимное накрывание монет не допускается). Тот из игроков, кто положит монету, не оставляющую места для новых монет, выигрывает.

Это игра с полной информацией. Существует вполне определенная стратегия, обеспечивающая выигрыш тому из игроков, кто начинает игру. А именно, начинающий игру должен положить первую монету точно в центр стола и на каждый ход противника отвечать симметричным ходом. Исход игры от стратегии второго игрока не зависит.

3. **«Переговоры».** В переговорах участвуют две стороны A и B. В слегка идеализированном варианте это может выглядеть, например, так.

Сначала сторона A высказывает одно из нескольких предложений, способных заинтересовать сторону B. Затем сторона B, ознакомившись с предложением стороны A, высказывает одно из нескольких встречных предложений, способных, по ее мнению, заинтересовать сторону A. В свою очередь, сторона A, ознакомившись с реакцией стороны B на сделанные предложения, высказывает ей новое предложение, внеся одну из нескольких возможных корректировок в свое первоначальное предложение с учетом мнения стороны B и т. д.

Если предмет переговоров сложен, то подобный обмен ходов может затянуться. Однако любые переговоры непременно заканчиваются. И там, на финише, ждет функция выигрышей.

Попробуем смоделировать короткий переговорный процесс трехходовой позиционной игрой.

Предположим, что переговоры заканчиваются через три хода, на каждом из которых соответствующая сторона имеет возможность выбора из двух альтернатив, и опишем соответствующую позиционную игру.

1-й ход делает сторона A. она выбирает одно из двух возможных предложений — число x из множества двух чисел $\{1, 2\}$.

2-й ход делает сторона B: она выбирает число y из множества двух чисел $\{1, 2\}$, зная число x, предложенное стороной A.

3-й ход делает сторона A: она выбирает число z из множества двух чисел $\{1, 2\}$, зная о предложении стороны B на 2-м ходе и помня собственное предложение на 1-м ходе.

После этого сторона A либо получает вознаграждение (например, в виде кредита от стороны B), либо выплачивает стороне B штраф.

Все эти возможности описываются функцией выигрышей $W(x, y, z)$, заданной следующей таблицей

$$\begin{aligned} &W(1,1,1)=a, \qquad &W(2,1,1)=e,\\ &W(1,1,2)=b, \qquad &W(2,1,2)=f.\\ &W(1,2,1)=c, \qquad &W(2,2,1)=g,\\ &W(1,2,2)=d, \qquad &W(2,2,2)=h. \end{aligned}$$

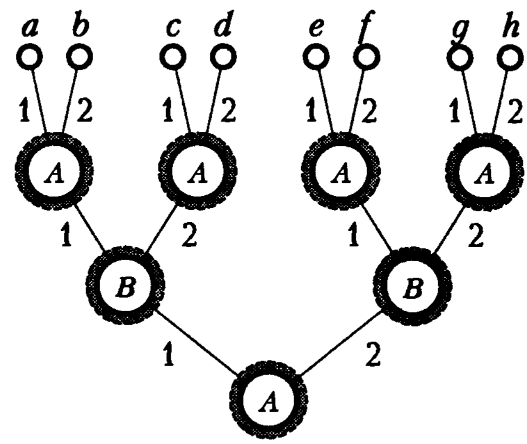


Рис. 11

Графическое представление этой игры показано на рис. 11.

Ясно, что описанная позиционная игра является игрой с полной информацией.

Начнем с описания возможных стратегий игрока B.

Поскольку игроку B выбор игрока A на 1-м ходе известен, то у игрока B те же четыре стратегии, что и в примере 13.

$$B_1 — [1, 1], \quad B_2 — [1, 2], \quad B_3 — [2, 1], \quad B_4 — [2, 2].$$

С описания возможных стратегий игрока A дело обстоит немного посложнее — их восемь.

Чистая стратегия игрока A в данной игре описывается упорядоченной тройкой

$$(x, [z_1, z_2]).$$

Здесь x ($x = 1, 2$) — альтернатива, которую игрок A выбирает на 1-м ходе, z_1 ($z_1 = 1, 2$) — альтернатива, которую игрок A выбирает на 3-м ходе, если на 2-м ходе игрок B выбрал выбрал первую альтернативу ($y = 1$) и z_2 ($z_2 = 1, 2$) — альтернатива, которую игрок A выбирает на 3-м ходе, если на 2-м ходе игрок B выбрал вторую альтернативу ($y = 2$).

Например, выбор игроком A стратегии $(1, [2, 1])$ означает, что на 1-м ходе игрок A выбирает $x = 1$, а на 3-м $z = 2$, если игрок B выбрал $y = 1$, и $z = 1$, если игрок B выбрал $y = 2$.

Тем самым, у игрока A восемь чистых стратегий:

$$A_1 — (1, [1, 1]), \quad A_2 — (1, [1, 2]), \quad A_3 — (1, [2, 1]), \quad A_4 — (1, [2, 2]),$$
$$A_5 — (2, [1, 1]), \quad A_6 — (2, [1, 2]), \quad A_7 — (2, [2, 1]), \quad A_8 — (2, [2, 2]).$$

Покажем теперь, как в зависимости от применяемых стратегий определяются элементы таблицы выигрышей игрока A.

Пусть, например, игрок A выбрал стратегию A_6 — $(2, [1, 2])$, а игрок B — стратегию B_3 — $[2, 1]$. Тогда $x = 2$. Из $[2, 1]$ вытекает, что $y = 1$, а из $(2, [1, 2])$, что $z = 1$. Отсюда

$$W(x, y, z) = W(2, 1, 1) = e.$$

Рассчитывая по этой же схеме все остальные элементы таблицы выигрышей, в итоге получим

		B_1	B_2	B_3	B_4
		[1, 1]	[1, 2]	[2, 1]	[2, 2]
A_1	(1, [1, 1])	$W(1,1,1)$	$W(1,1,1)$	$W(1,2,1)$	$W(1,2,1)$
A_2	(1, [1, 2])	$W(1,1,1)$	$W(1,1,1)$	$W(1,2,2)$	$W(1,2,2)$
A_3	(1, [2, 1])	$W(1,1,2)$	$W(1,1,2)$	$W(1,2,1)$	$W(1,2,1)$
A_4	(1, [2, 2])	$W(1,1,2)$	$W(1,1,2)$	$W(1,2,2)$	$W(1,2,2)$
A_5	(2, [1, 1])	$W(2,1,1)$	$W(2,2,1)$	$W(2,1,1)$	$W(2,2,1)$
A_6	(2, [1, 2])	$W(2,1,1)$	$W(2,2,2)$	$W(2,1,1)$	$W(2,2,2)$
A_7	(2, [2, 1])	$W(2,1,2)$	$W(2,2,1)$	$W(2,1,2)$	$W(2,2,1)$
A_8	(2, [2, 2])	$W(2,1,2)$	$W(2,2,2)$	$W(2,1,2)$	$W(2,2,2)$

$$\begin{pmatrix} a & a & c & c \\ a & a & d & d \\ b & b & c & c \\ b & b & d & d \\ e & g & e & g \\ e & h & e & h \\ f & g & f & g \\ f & h & f & h \end{pmatrix}.$$

Вследствие того, что рассматриваемая позиционная игра является игрой с полной информацией, полученная матрица имеет седловую точку при любой функции выигрышей. В этом легко убедиться, произвольно выбирая значения параметров a, b, c, d, e, f, g и h.

При увеличении числа ходов стратегии в позиционной игре с полной информацией строятся по аналогичной схеме.

Рассмотрим, например, четырехходовую позиционную игру.

1-й ход делает игрок A: он выбирает число x из множества двух чисел $\{1, 2\}$.

2-й ход делает игрок B: он выбирает число y из множества двух чисел $\{1, 2\}$, зная число x, выбранное игроком A на 1-м ходе.

3-й ход делает игрок A: он выбирает число z из множества двух чисел $\{1, 2\}$, зная число y, выбранное игроком B на 2-м ходе, и помня свой выбор числа x на 1-м ходе.

4-й ход делает игрок B: он выбирает число u из множества двух чисел $\{1, 2\}$, зная число z, выбранное игроком A на 3-м ходе, помня свой выбор числа y на 2-м ходе и зная выбор игрока A на 1-м ходе — число x.

После этого игроки A и B расплачиваются в соответствии с заданной функцией выигрышей $W(x, y, z, u)$.

В этой игре стратегии у игрока A те же, что и в задаче, рассмотренной выше: каждая из них задается тройкой вида

$$(x, [z_1, z_2]),$$

и общее их число равно восьми.

Что касается стратегий игрока B, то в этой игре их шестнадцать и каждая из них задается четверкой вида

$$([y_1, y_2], [u_1, u_2]).$$

Матрица выигрышей игрока A в данной игре имеет размер 8×16. Покажем, как определяются ее элементы в зависимости от применяемых стратегий игроков.

Пусть, например, игрок A выбрал стратегию A' — $(2, [2, 1])$, а игрок B стратегию B' — $([2, 1], [1, 2])$. Тогда $x = 1$, $y = 2$, а $z = 1$. Из того, что $[u_1, u_2] = [1, 2]$, получаем, что $u = 1$. Отсюда следует, что искомый элемент матрицы выплат равен

$$W(1, 2, 1, 1).$$

Остальные элементы матрицы вычисляются аналогично.

Так как эта позиционная игра также является игрой с полной информацией, то получаемая матрица будет иметь седловую точку.

§ 4. Несколько слов в заключение

В рассмотренных примерах основное внимание было уделено описанию процесса нормализации позиционной игры — построению дерева игры и информационных множеств, выработке стратегий игроков и вычислению элементов платежной матрицы. Следующий естественный шаг — отыскание цены игры и оптимальных стратегий игроков — проводится методами, о которых рассказывалось в главе, посвященной матричным играм.

Мы достаточно подробно остановились на позиционных играх двух лиц, где были явно выражены интересы одного из игроков (игрока A). Следует, однако, иметь в виду, что в одних случаях интересы игрока B могут быть полностью противоположными интересам игрока A, в то время как в других вполне может оказаться, что то, что хорошо для одного игрока, не обязательно плохо для другого. Приведем два простых примера.

Пример А.

1-й ход. Игрок A выбирает число x из множества двух чисел $\{1, 2\}$.

2-й ход. Игрок B выбирает число y из множества двух чисел $\{1, 2\}$, зная выбор числа x игроком A. Функции выплат игрокам A и B — $W_A(x, y)$ и $W_B(x, y)$ соответственно — задаются так:

$$W_A(1,1) = 1, \quad W_A(1,2) = -1, \quad W_A(2,1) = -2, \quad W_A(2,2) = 2,$$
$$W_B(1,1) = 2, \quad W_B(1,2) = 1, \quad W_B(2,1) = 1, \quad W_B(2,2) = 2.$$

Дерево игры показано на рис. 12. Исход игры зависит от того, каковы намерения игрока B — *максимизировать свой выигрыш*:

$$W_B(x, y) \to \max,$$

или *максимизировать свой относительный выигрыш*:

$$W_B(x, y) - W_A(x, y) \to \max.$$

В первом случае это достигается так.

при $x = 1$ $\quad y = 1$ и $W_B(1,1) = 2$ $(W_A(1,1) = 1)$;

при $x = 2$ $\quad y = 2$ и $W_B(2,2) = 2$ $(W_A(2,2) = 2)$.

Во втором случае:

при $x = 1$ $\quad y = 2$ и $W_B(1,2) - W_A(1,2) = 1 - (-1) = 2)$;

при $x = 2$ $\quad y = 1$ и $W_B(2,1) - W_A(2,1) = 1 - (-2) = 3)$.

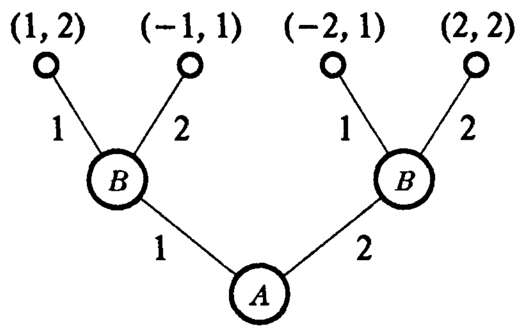


Рис. 12

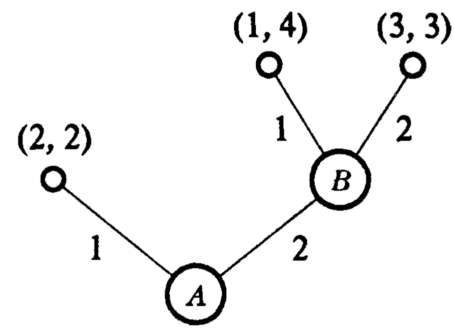


Рис. 13

Пример Б. Игра задается деревом (см. рис. 13).

1-й ход. Игрок A выбирает число x из множества двух чисел $\{1, 2\}$.

Если $x = 1$, то каждый из игроков получает свой выигрыш, равный 2.

Если $x = 2$, то право 2-го хода получает игрок B, где он и выбирает выбирает число y из множества двух чисел $\{1, 2\}$.

При $y = 1$ выигрыш игрока A равен 1, а игрока B — 4. При $y = 2$ оба игрока получают поровну — по 3.

В случае, когда каждый из игроков стремится к получению максимального выигрыша и любые виды кооперации запрещены, исход игры ясен — игрок A выбирает $x = 1$, и игра заканчивается. Но при $x = 2$ и $y = 2$ каждый из игроков получает по 3 (такой исход предпочтительнее простейшего $(1, 1)$), и, если допустить соглашение между игроками, это обстоятельство вполне может изменить исход игры.

Глава XLVIII

БИМАТРИЧНЫЕ ИГРЫ

Предыдущие рассмотрения касались игр двух лиц, в которых интересы игроков были прямо противоположны (антагонистические, или матричные игры), а также позиционных игр, сводимых к матричным. Однако ситуации, в которых интересы игроков хотя и не совпадают, но уже не обязательно являются противоположными, встречаются значительно чаще.

Рассмотрим, например, конфликтную ситуацию, в которой каждый из двух участников имеет следующие возможности для выбора своей линии поведения:

игрок A — может выбрать любую из стратегий $A_1, \ldots, A_m$,
игрок B — любую из стратегий $B_1, \ldots, B_n$.

При этом всякий раз их совместный выбор оценивается вполне определенно:

если игрок A выбрал i-ю стратегию A_i, а игрок B — k-ю стратегию B_k, то в итоге выигрыш игрока A будет равен некоторому числу a_{ik}, а выигрыш игрока B некоторому, вообще говоря, другому числу b_{ik}.

Иными словами, всякий раз каждый из игроков получает свой приз.

Последовательно перебирая все стратегии игрока A и все стратегии игрока B, мы сможем заполнить их выигрышами две таблицы

	B_1	$\ldots$	B_k	$\ldots$	B_n
A_1	a_{11}	$\ldots$	a_{1k}	$\ldots$	a_{1n}
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
A_i	a_{i1}	$\ldots$	a_{ik}	$\ldots$	a_{in}
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
A_m	a_{m1}	$\ldots$	a_{mk}	$\ldots$	a_{mn}

	B_1	$\ldots$	B_k	$\ldots$	B_n
A_1	b_{11}	$\ldots$	b_{1k}	$\ldots$	b_{1n}
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
A_i	b_{i1}	$\ldots$	b_{ik}	$\ldots$	b_{in}
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
A_m	b_{m1}	$\ldots$	b_{mk}	$\ldots$	b_{mn}

Первая из таблиц описывает выигрыши игрока A, а вторая — выигрыши игрока B. Обычно эти таблицы записывают в виде матриц

$$\mathsf{A} = \begin{pmatrix} a_{11} & \ldots & a_{1k} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{i1} & \ldots & a_{ik} & \ldots & a_{in} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & \ldots & a_{mk} & \ldots & a_{mn} \end{pmatrix}, \qquad \mathsf{B} = \begin{pmatrix} b_{11} & \ldots & b_{1k} & \ldots & b_{1n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ b_{i1} & \ldots & b_{ik} & \ldots & b_{in} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ b_{m1} & \ldots & b_{mk} & \ldots & b_{mn} \end{pmatrix}.$$

Здесь A — *платежная матрица игрока A*, а B — *платежная матрица игрока B*.

При выборе игроком A i-й стратегии, а игроком B — k-й стратегии их выигрыши находятся в матрицах выплат на пересечении i-х строк и k-х столбцов: в матрице A это элемент a_{ik}, а в матрице B — элемент b_{ik}.

Таким образом, в случае, когда интересы игроков различны (но не обязательно противоположны), получаются две платежные матрицы: одна — матрица выплат игроку A, другая — матрица выплат игроку B. Поэтому совершенно естественно звучит название, которое обычно присваивается подобной игре — *биматричная*.

Замечание. Рассматриваемые ранее матричные игры, разумеется, можно рассматривать и как биматричные, где матрица выплат игроку B *противоположна* матрице выплат игроку A:

$$b_{ik} = -a_{ik}$$

или

$$\mathsf{A} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad \mathsf{B} = -\mathsf{A} = \begin{pmatrix} -a_{11} & \dots & -a_{1n} \\ \dots & \dots & \dots \\ -a_{m1} & \dots & -a_{mn} \end{pmatrix}.$$

Тем не менее, в общем случае биматричная игра — это игра с ненулевой суммой.

Нам кажется вполне естественным время от времени сопоставлять наши рассмотрения с рассуждениями, проведенными ранее для матричных игр (особенно при попытках разрешения схожих проблем). Подобные сопоставления часто оказываются одновременно и удобными и полезными. Конечно, класс биматричных игр значительно шире класса матричных (разнообразие новых моделируемых конфликтных ситуаций весьма заметно), а, значит, неизбежно увеличиваются и трудности, встающие на пути их успешного разрешения. Впрочем, мы надеемся, что часть из этих трудностей мы сумеем преодолеть уже в настоящем издании.

§ 1. Примеры биматричных игр

Примеры этого раздела описывают некоторые типические конфликтные ситуации, приводящие к биматричным играм. Сначала мы обсудим вопросы, связанные с формализацией рассматриваемых конфликтов (построение платежных матриц), а позднее связанные с рекомендациями по их разрешению.

Пример 20. «Борьба за рынки». Небольшая фирма (игрок A) намерена сбыть партию товара на одном из двух рынков, контролируемых другой, более крупной фирмой (игрок B). Для этого фирма A готова предпринять на одном из рынков соответствующие приготовления (например, развернуть рекламную компанию). Господствующая на рынках фирма B может попытаться воспрепятствовать этому, приняв на одном из рынков предупредительные меры (разумеется, в ремках закона). Не встречая противодействия на рынке, фирма A захватывает его; при наличии препятствий — терпит поражение,

Будем считать для определенности, что проникновение фирмы A на первый рынок более выгодно для нее, нежели проникновение на второй. Естественно также считать, что и борьба за первый рынок потребует вложения бо́льших средств. Например, победа фирмы A на первом рынке принесет ей вдвое больший выигрыш, чем победа на втором, но зато и поражение при попытке освоиться на первом рынке полностью ее разорит, а фирму B избавит от конкурента.

Что же касается второго рынка, то при поражении фирмы A ее потери будут не столь разорительны, но и победа принесет немного.

Таким образом, у фирмы A две стратегии:

A_1 — выбор первого рынка, A_2 — выбор второго рынка.

Такие же стратегии и у фирмы B:

B_1 — выбор первого рынка, B_2 — выбор второго рынка.

Для того, чтобы составить платежные матрицы игроков, нужны расчетные количественные показатели, которые мы приведем здесь в условных единицах:

$$\mathsf{A} = \begin{pmatrix} -10 & 2 \\ 1 & -1 \end{pmatrix}, \quad \mathsf{B} = \begin{pmatrix} 5 & -2 \\ -1 & 1 \end{pmatrix}.$$

Взглянем на выписанные матрицы выплат. Из сказанного выше ясно, что если оба игрока выберут один и тот же рынок, то победа останется за более сильной фирмой B.

То, что в ситуации (A_1, B_1) выигрыш игрока B равен 5, а в ситуации (A_2, B_2) — 1, подчеркивает, что первый рынок более выгоден (удобно расположен, хорошо посещаем и т. п.), чем второй. Выигрыш (-10) игрока A в ситуации (A_1, B_1) (а точнее, проигрыш) в сопоставлении с его выигрышем (-1) в ситуации (A_2, B_2) выглядит, разумеется, вполне сокрушительно. Что же касается ситуации, когда фирмы уделяют основное внимание разным рынкам (A_1, B_2) и (A_2, B_1), то здесь фирму A ждет настоящий выигрыш, бо́льший на более выгодном рынке. Потери, которые при этом несет фирма B, оказываются прямо противоположными.

Замечание. Ясно, что точно рассчитать выгоду и ущерб сторон в этом конфликте заранее довольно трудно. А вот в следующей конфликтной ситуации размеры выигрышей игроков известны со всей определенностью.

Пример 21. «Дилемма узников». Игроками являются два узника, находящихся в предварительном заключении по подозрению в совершении преступления. При отсутствии прямых улик возможность их осуждения в большой степени зависит от того, заговорят они или будут молчать.

Если оба будут молчать, то наказанием будет лишь срок предварительного заключения (потери каждого из узников составят (-1)). Если сознаются, то получат срок, учитывающий признание как смягчающее обстоятельство (потери каждого из узников составят в этом случае (-6)). Если же заговорит только один из узников, а другой будет молчать, то в этом случае заговоривший будет выпущен на свободу (его потери равны 0), а сохраняющий молчание получит максимально возможное наказание (его потери будут равны (-9)).

Эта конфликтная ситуация приводит к биматричной игре, в которой каждый из игроков имеет по две стратегии — молчать (М) или говорить (Г).

Выигрыши игроков A и B соответственно описываются так:

	(М)	(Г)
(М)	−1	−9
(Г)	0	−6

	(М)	(Г)
(М)	−1	−9
(Г)	0	−6

Пример 22. «Семейный спор». Два партнера договариваются о совместном проведении одного из двух действий, (1) и (2), каждое из которых требует их совместного участия.

В случае осуществления первого из этих двух действий выигрыш первого партнера (игрок A) будет вдвое выше выигрыша второго партнера (игрок B). Напротив, в случае осуществления второго из этих двух действий выигрыш игрока A будет вдвое меньше выигрыша игрока B. Если же партнеры выполнят различные действия, то выигрыш каждого из них будет равен нулю.

Эта конфликтная ситуация приводит к биматричной игре, в которой каждый из игроков имеет по две стратегии. Выигрыши игроков A и B соответственно описываются таблицами следующего вида:

	(1)	(2)
(1)	2	0
(2)	0	1

	(1)	(2)
(1)	1	0
(2)	0	2

Пояснение. Довольно понятно, что различные конфликтные ситуации могут иметь одну и ту же формализацию. В частности, рассмотренная биматричная игра частно интерпретируется, как одновременный выбор супругами совместного развлечения: посещение оперного спектакля или хоккейного матча. При этом в посещении оперного театра жена заинтересована в большей степени, чем муж, а при посещении стадиона наблюдается обратная картина. В случае же непреодоления разногласий, возникших при выборе, день оказывается вообще испорченным.

Отсюда и название, вынесенное в заголовок.

Пример 23. «Студент — Преподаватель». Рассмотрим следующую ситуацию. Студент (игрок A) готовится к зачету, который принимает Преподаватель (игрок B). Можно считать, что у Студента две стратегии — подготовиться к сдаче зачета $(+)$ и не подготовиться $(-)$. У Преподавателя также две стратегии — поставить зачет $[+]$ и не поставить зачета $[-]$.

В основу значений функций выигрыша игроков положим следующие соображения:

Выигрыш Студента

	[+]	[−]
(+)	оценка заслуженв	очень обидно
(−)	удалось обмануть	оценка заслужена

Выигрыш Преподавателя

	[+]	[−]
(+)	все нормально	был неправ
(−)	дал себя обмануть	опять придет

Количественно это можно выразить, например, так

	[+]	[−]
(+)	2	−1
(−)	1	0

	[+]	[−]
(+)	1	−3
(−)	−2	−1

§ 2. Смешанные стратегии

В приведенных примерах (позже мы вернемся к подробному рассмотрению каждого) описаны ситуации, в которых интересы игроков не совпадают. Естественно встает вопрос о том, какие рекомендации необходимо дать игрокам для того, чтобы моделируемая конфликтная ситуация разрешилась. Иными словами, что мы будем понимать под решением биматричной игры?

Попробуем ответить на это вопрос так:

вследствие того, что интересы игроков не совпадают, нам нужно построить такое (компромиссное) решение, которое бы в том или ином, но в одинаковом смысле удовлетворяло обоих игроков.

Не пытаясь сразу выражать эту мысль совсем точно, скажем — попробуем найти некую равновесную ситуацию, явное отклонение от которой одного из игроков уменьшало бы его выигрыш.

Подобный вопрос мы ставили и при рассмотрении матричных игр. Напомним, что возникающее при разработке минимаксного подхода понятие равновесной ситуации приводило нас к поиску седловой точки, которая, как оказалось, существует далеко не всегда — конечно, если ограничиваться только чистыми стратегиями игроков A и B, т. е. стратегиями

$$A_1, \ldots, A_m, B_1, \ldots, B_n.$$

Естественно ожидать, что в более сложном случае биматричной игры дело вряд ли обстоит проще.

Однако при расширении матричной игры путем перехода к смешанным стратегиям, т. е. к такому поведению игроков, при котором они чередуют (чистые) стратегии с определенными частотами: игрок A — стратегии $A_1, \ldots, A_m$ с частотами $p_1, \ldots, p_m$, где

$$p_1 \geqslant 0, \ \ldots, \ p_m \geqslant 0, \quad \sum_{i=1}^{m} p_i = 1,$$

а игрок B — стратегии $B_1, \ldots, B_n$ с частотами $q_1, \ldots, q_n$, где

$$q_1 \geqslant 0, \ \ldots, \ q_n \geqslant 0, \quad \sum_{k=1}^{n} q_k = 1,$$

выяснилось, что в смешанных стратегиях равновесная ситуация всегда существует. Иными словами, любая матричная игра в смешанных стратегиях разрешима.

Поэтому, рассматривая здесь биматричные игры, разумно попробовать сразу же перейти к смешанным стратегиям игроков (тем самым, мы предполагаем, что каждая игра может быть многократно повторена в неизменных обстоятельствах).

В матричном случае смешивание стратегий приводило к расширению возможности выплат в том смысле, что расчет строился из вычисления средних выигрышей игроков A и B, которые определялись по элементам платежной матрицы A и вероятностям p_i и q_k:

$$H_A = \sum_{i,k} a_{ik} p_i q_k, \qquad H_B = -\sum_{i,k} a_{ik} p_i q_k.$$

При смешанных стратегиях в биматричных играх также естественно возникают средние выигрыши игроков A и B, определяемые по правилам, в которых уже нет никакой дискриминации игрока B:

$$\boxed{H_A = \sum_{i,k} a_{ik} p_i q_k, \quad H_B = \sum_{i,k} b_{ik} p_i q_k.}$$

§3. 2×2 биматричные игры. Ситуация равновесия

Мы предполагаем уделить основное внимание случаю, когда у каждого из игроков имеется ровно две стратегии, т. е. случаю $m = n = 2$. Поэтому нам кажется уместным выписать приведенные выше формулы именно для такого случая.

В 2×2 биматричной игре платежные матрицы игроков имеют следующий вид

$$\boxed{\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \mathsf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},}$$

вероятности

$$\boxed{p_1 = p, \quad p_2 = 1 - p, \quad q_1 = q, \quad q_2 = 1 - q,}$$

а средние выигрыши вычисляются по формулам

$$\boxed{\begin{aligned} H_A(p, q) &= a_{11}pq + a_{12}p(1-q) + a_{21}(1-p)q + a_{22}(1-p)(1-q), \\ H_B(p, q) &= b_{11}pq + b_{12}p(1-q) + b_{21}(1-p)q + b_{22}(1-p)(1-q), \end{aligned}}$$

где

$$0 \leqslant p \leqslant 1, \quad 0 \leqslant q \leqslant 1.$$

Сформулируем основное определение.

Определение. Будем говорить, что пара чисел

$$(p^*, q^*), \quad 0 \leqslant p^* \leqslant 1, \ 0 \leqslant q^* \leqslant 1,$$

определяет *равновесную ситуацию*, если для любых p и q, подчиненных условиям $0 \leqslant p \leqslant 1$, $0 \leqslant q \leqslant 1$, одновременно выполнены следующие неравенства

$$H_A(p, q^*) \leqslant H_A(p^*, q^*), \quad H_B(p^*, q) \leqslant H_B(p^*, q^*). \qquad (\star)$$

Пояснение. Выписанные неравенства ($\star$) означают следующее: ситуация, определяемая смешанной стратегией (p^*, q^*), является равновесной, если отклонение от нее одного из игроков при условии, что другой сохраняет свой выбор, приводит к тому, что выигрыш отклонившегося игрока может только уменьшиться. Тем самым, получается, что если равновесная ситуация существует, то отклонение от нее невыгодно самому игроку.

Но может ли быть подобная ситуация равновесия в биматричной игре? Ответ на поставленный вопрос дает следующее утверждение.

Теорема 4 (Дж. Нэш). *Всякая биматричная игра имеет хотя бы одну равновесную ситуацию (точку равновесия) в смешанных стратегиях.*

Заметим, что весьма похожее утверждение мы уже встречали при изучении вопроса разрешимости матричных игр. Напомним также, что в соответствующем разделе после формулировки аналогичного утверждения о существовании равновесной ситуации описывался и способ отыскания точки равновесия. Та же проблема встает перед нами и здесь.

Итак, равновесная ситуация существует. Но как ее найти?

Если некоторая пара чисел (p^*, q^*) претендует на то, чтобы определять ситуацию равновесия, то для того, чтобы убедиться в обоснованности этих претензий, или, наоборот, доказать их необоснованность, необходимо проверить справедливость неравенств $(\star)$ для любого p в пределах от 0 до 1 и для любого q в пределах от 0 до 1.

В общем случае число таких проверок бесконечно. И, следовательно, действенный способ определения равновесной ситуации нужно искать где-то в ином месте.

Для этого мы обопремся на следующий теоретический результат.

Теорема 5. *Выполнение неравенств*

$$H_A(p, q^*) \leqslant H_A(p^*, q^*), \qquad H_B(p^*, q) \leqslant H_B(p^*, q^*) \tag{$\star$}$$

равносильно выполнению неравенств

$$\begin{aligned} H_A(0, q^*) &\leqslant H_A(p^*, q^*), & H_B(p^*, 0) &\leqslant H_B(p^*, q^*),\\ H_A(1, q^*) &\leqslant H_A(p^*, q^*), & H_B(p^*, 1) &\leqslant H_B(p^*, q^*). \end{aligned} \tag{$\star\star$}$$

Иными словами, для того, чтобы убедиться в обоснованности претензий пары (p^*, q^*) на то, чтобы определять равновесную ситуацию, достаточно проверить справедливость неравенства

$$H_A(p, q^*) \leqslant H_A(p^*, q^*)$$

только для двух чистых стратегий игрока A ($p = 0$ и $p = 1$) и неравенства

$$H_B(p^*, q) \leqslant H_B(p^*, q^*)$$

только для двух чистых стратегий игрока B ($q = 0$ и $q = 1$).

Четыре неравенства $(\star\star)$ позволяют провести поиск точки равновесия уже вполне конструктивно.

Запишем средние выигрыши игроков A и B в более удобной форме. Имеем

$$\begin{aligned} H_A(p, q) &= (a_{11} - a_{12} - a_{21} + a_{22})pq + (a_{12} - a_{22})p + (a_{21} - a_{22})q + a_{22},\\ H_B(p, q) &= (b_{11} - b_{12} - b_{21} + b_{22})pq + (b_{12} - b_{22})p + (b_{21} - b_{22})q + b_{22}. \end{aligned}$$

Обратимся к первой из полученных формул. Полагая в ней сначала $p = 1$, а потом $p = 0$, получаем, что

$$\begin{aligned} H_A(1, q) &= (a_{11} - a_{12} - a_{21} + a_{22})q + a_{12} + (a_{21} - a_{22})q,\\ H_A(0, q) &= (a_{21} - a_{22})q + a_{22}. \end{aligned}$$

Рассмотрим разности

$$\begin{aligned} H_A(p, q) - H_A(1, q) &= (a_{11} - a_{12} - a_{21} + a_{22})pq + (a_{12} - a_{22})p -\\ &\quad - (a_{11} - a_{12} - a_{21} + a_{22})q + a_{22} - a_{12},\\ H_A(p, q) - H_A(0, q) &= (a_{11} - a_{12} - a_{21} + a_{22})pq + (a_{12} - a_{22})p. \end{aligned}$$

Полагая

$$C = a_{11} - a_{12} - a_{21} + a_{22}, \qquad \alpha = a_{22} - a_{12},$$

получим для них следующие выражения

$$H_A(p,q) - H_A(1,q) = Cpq - \alpha p - Cq + \alpha = Cq(p-1) - \alpha(p-1) = (p-1)(Cq-\alpha),$$

$$H_A(p,q) - H_A(0,q) = Cpq - \alpha p = p(Cq-\alpha).$$

В случае, если пара (p, q) определяет точку равновесия, эти разности неотрицательны

$$H_A(p, q) - H_A(1, q) \geqslant 0, \quad H_A(p, q) - H_A(0, q) \geqslant 0.$$

Поэтому окончательно получаем

$$(p-1)(Cq-\alpha) \geqslant 0,$$
$$p(Cq-\alpha) \geqslant 0.$$

Из формул для функции $H_B(p, q)$ при $q = 1$ и $q = 0$ соответственно имеем

$$H_B(p, 1) = (b_{11} - b_{12} - b_{21} + b_{22})p + (b_{12} - b_{22})p + b_{21},$$
$$H_B(p, 0) = (b_{12} - b_{22})p + b_{22}.$$

Разности

$$H_B(p, q) - H_B(p, 1) \quad \text{и} \quad H_B(p, q) - H_B(p, 0)$$

с учетом обозначений

$$D = b_{11} - b_{12} - b_{21} + b_{22}, \quad \beta = b_{22} - b_{21}$$

приводятся к виду

$$H_B(p, q) = H_B(p, 1) = (q-1)(Dp-\beta),$$
$$H_B(p, q) = H_B(p, 0) = q(Dq-\beta)$$

совершенно так же, как соответствующие разности для функции H_A.

Если пара (p, q) определяет точку равновесия, то эти разности неотрицательны

$$H_B(p, q) - H_B(p, 1) \geqslant 0, \quad H_B(p, q) - H_B(p, 0) \geqslant 0.$$

Поэтому

$$(q-1)(Dp-\beta) \geqslant 0,$$
$$q(Dp-\beta) \geqslant 0.$$

Прежде чем приступать к последующим шагам, подведем некоторые итоги.

Для того, чтобы в биматричной игре

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \mathsf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

пара (p, q) определяла равновесную ситуацию, необходимо и достаточно одновременное выполнение следующих неравенств

$$\boxed{\begin{array}{rc} (p-1)(Cq-\alpha) \geqslant 0, & \multirow{2}{*}{} \\ p(Cq-\alpha) \geqslant 0, & 0 \leqslant p \leqslant 1, \\ (q-1)(Dp-\beta) \geqslant 0, & \\ q(Dp-\beta) \geqslant 0, & 0 \leqslant q \leqslant 1, \end{array}} \qquad (*)$$

где

$$\boxed{\begin{array}{ll} C = a_{11} - a_{12} - a_{21} + a_{22}, & \alpha = a_{22} - a_{12}, \\ D = b_{11} - b_{12} - b_{21} + b_{22}, & \beta = b_{22} - b_{21}. \end{array}} \qquad (**)$$

§ 4. Поиск равновесных ситуаций

Геометрический смысл условий (*) рассмотрим на примерах описанных выше биматричных игр.

Пример 20. «Борьба за рынки» (продолжение). Напомним, что ситуация, сложившаяся в этой задаче, задается платежными матрицами следующего вида

$$A = \begin{pmatrix} -10 & 2 \\ 1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & -2 \\ -1 & 1 \end{pmatrix}.$$

Заменяя в неравенстве (*) величины C, α, D и β их конкретными значениями

$$C = -10 - 2 - 1 - 1 = -14, \qquad \alpha = -1 - 2 = -3,$$
$$D = \quad 5 + 2 + 1 + 1 = \quad 9, \qquad \beta = \quad 1 + 1 = \quad 2,$$

получаем

$$(l) \quad \begin{aligned} (p-1)(-14q-(-3)) &\geqslant 0, \\ p(-14q-(-3)) &\geqslant 0, \end{aligned} \qquad (r) \quad \begin{aligned} (q-1)(9p-2) &\geqslant 0, \\ q(9p-2) &\geqslant 0. \end{aligned}$$

Рассмотрим сначала левую пару неравенств (l)

$$\begin{aligned} (p-1)(-14q+3) &\geqslant 0, \\ p(-14q+3) &\geqslant 0. \end{aligned}$$

Возможны следующие три случая

$$1)\ p = 1, \quad 2)\ p = 0, \quad 3)\ 0 < p < 1.$$

Разберем каждый из этих случаев подробно.

1. Полагая $p = 1$, получаем

$$0 \geqslant 0, \quad -14q + 3 \geqslant 0.$$

Откуда

$$14q - 3 \leqslant 0$$

и, значит,

$$q \geqslant \tfrac{3}{14}.$$

2. Полагая $p = 0$, получаем

$$0 \geqslant 0, \quad -(-14q + 3) \geqslant 0, \quad 0 \geqslant 0.$$

Откуда

$$14q - 3 \geqslant 0$$

и, значит,

$$q \leqslant \tfrac{3}{14}.$$

3. Наконец, положив $0 < p < 1$, получим

$$\begin{aligned} -14q + 3 &\geqslant 0, \\ -14q + 3 &\leqslant 0, \end{aligned}$$

что возможно лишь в случае, если

$$-14q + 3 = 0,$$

т. е.

$$q = \tfrac{3}{14}.$$

Сформулируем результат наших рассмотрений:

$$\begin{aligned} &1^\circ.\ p = 1, && q \leqslant \tfrac{3}{14}, \\ &2^\circ.\ p = 0, && q \geqslant \tfrac{3}{14}, \\ &3^\circ.\ 0 < p < 1, && q = \tfrac{3}{14}. \end{aligned}$$

Перенесем теперь полученные сведения на чертеж.

Введем на плоскости прямоугольную систему координат (p, q) и выделим на ней единичный квадрат, соответствующий неравенствам

$$0 \leqslant p \leqslant 1, \quad 0 \leqslant q \leqslant 1$$

(рис. 1).

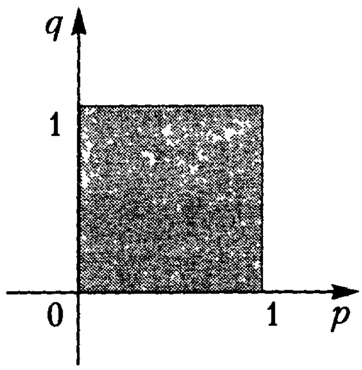


Рис. 1

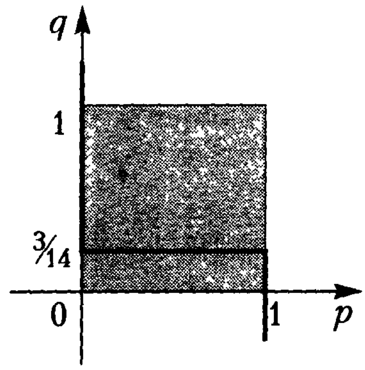


Рис. 2

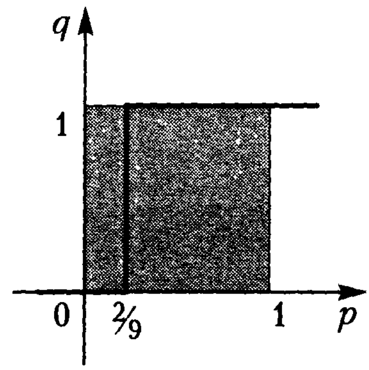


Рис. 3

Нанесем на этот чертеж то множество точек, которое описывается условиями 1°, 2° и 3°. Это множество (на рис. 2 его точки выделены жирной линией) состоит из трех прямолинейных участков — двух вертикальных лучей и одного горизонтального отрезка — и представляет собой «зигзаг». Нас будет интересовать только та его часть, которая попала в заштрихованный на рис. 2 единичный квадрат.

Оставив на время полученные результаты в покое, обратимся к правой части неравенств (r):

$$(q-1)(9p-2) \geqslant 0,$$
$$q(9p-2) \geqslant 0.$$

Три интересных для нас случая

$$1)\ q=1, \quad 2)\ q=0, \quad 3)\ 0<q<1,$$

приводят нас к следующему результату

$$1^\circ.\ q=1, \qquad p \geqslant \tfrac{2}{9},$$
$$2^\circ.\ q=0, \qquad p \leqslant \tfrac{2}{9},$$
$$3^\circ.\ 0<q<1, \quad p = \tfrac{2}{9}.$$

Перенося его на чертеж, получим второй «зигзаг», но уже горизонтальный (рис. 3).

Теперь остается только объединить полученное на рис. 4.

Общая точка построенных зигзагов — точка равновесия — имеет координаты

$$\left(\tfrac{2}{9}, \tfrac{3}{14}\right).$$

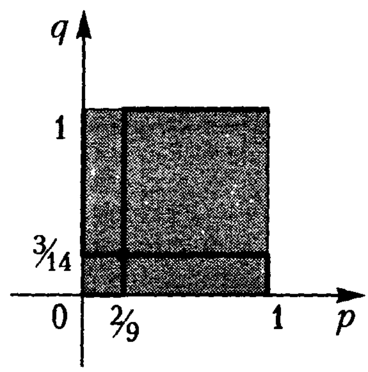


Рис. 4

Соответствующие смешанные стратегии игроков имеют следующий вид

$$P = \left\{\tfrac{2}{9}, \tfrac{7}{9}\right\}, \qquad Q = \left\{\tfrac{3}{14}, \tfrac{11}{14}\right\},$$

а средние выигрыши игроков таковы

$$H_A\left(\tfrac{2}{9}, \tfrac{3}{14}\right) = -\tfrac{4}{7}, \qquad H_B\left(\tfrac{2}{9}, \tfrac{3}{14}\right) = \tfrac{1}{3}.$$

Замечание. Попробуем разбить рассмотренную биматричную игру на две матричные игры с нулевой суммой.

1. *Игра с матрицей* A.

$$\mathrm{A} = \begin{pmatrix} -10 & 2 \\ 1 & -1 \end{pmatrix}.$$

Решая эту игру графическим методом, найдем оптимальную смешанную стратегию для игрока A

$$\left\{\tfrac{1}{7}, \tfrac{6}{7}\right\},$$

цену этой игры

$$\nu_A^0 = -\tfrac{4}{7},$$

а затем и оптимальную смешанную стратегию для игрока B

$$\left\{\tfrac{3}{14}, \tfrac{11}{14}\right\}.$$

2. *Игра с матрицей* B

$$\mathrm{B} = \begin{pmatrix} 5 & -2 \\ -1 & 1 \end{pmatrix}.$$

Решая эту игру графическим методом, найдем оптимальную смешанную стратегию для игрока B

$$\left\{\tfrac{1}{3}, \tfrac{2}{3}\right\},$$

цену этой игры

$$\nu_B^0 = \frac{1}{3},$$

а затем и оптимальную смешанную стратегию для игрока A

$$\left\{\frac{2}{9}, \frac{7}{9}\right\}.$$

Сравнивая полученные результаты с решением биматричной игры, можно заметить следующее: *если каждый игрок будет применять свои стратегии в этой игре, исходя только из матрицы своих выигрышей, то его оптимальный средний выигрыш совпадет с его выигрышем при равновесной ситуации*; кстати, по своей матрице игрок может найти и оптимальную смешанную стратегию другого игрока (но не свою!).

Пример 21. «Дилемма узников» (продолжение). Выигрыши игроков A и B описываются соответствующими матрицами выплат:

$$\mathrm{A} = \begin{pmatrix} -1 & -9 \\ 0 & -6 \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} -1 & 0 \\ -9 & -6 \end{pmatrix}.$$

Проведем необходимые вычисления. Имеем

$$C = -1 - (-9) - 0 + (-6) = 2, \qquad \alpha = -6 - (-9) = 3,$$
$$D = -1 - 0 - (-9) + (-6) = 2, \qquad \beta = -6 - (-9) = 3.$$

Отсюда

$$(l)\quad \begin{aligned} (p-1)(2q-3) &\geqslant 0, \\ p(2q-3) &\geqslant 0, \end{aligned} \qquad (r)\quad \begin{aligned} (q-1)(2p-3) &\geqslant 0, \\ q(2p-3) &\geqslant 0, \end{aligned}$$

получаем, что

$$1^\circ.\ p = 1,\ q \geqslant \tfrac{3}{2}, \qquad 2^\circ.\ p = 0,\ q \leqslant \tfrac{3}{2}, \qquad 3^\circ.\ 0 < p < 1,\ q = \tfrac{3}{2}.$$
$$1^\circ.\ q = 1,\ p \geqslant \tfrac{3}{2}, \qquad 2^\circ.\ q = 0,\ p \leqslant \tfrac{3}{2}, \qquad 3^\circ.\ 0 < q < 1,\ p = \tfrac{3}{2}.$$

Полученные зигзаги изображены на рис. 5.

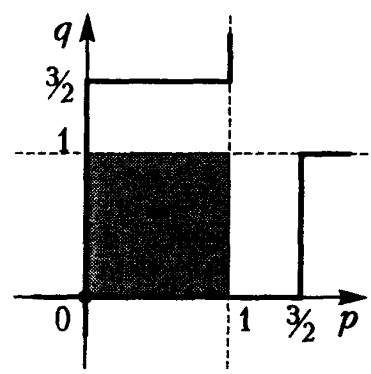


Рис. 5

Единственная равновесная ситуация — $(0, 0)$. Это ситуация, в которой каждый из игроков выбирает вторую чистую стратегию — сознаться — и теряет 6.

Как мы уже отмечали ранее, отклонение от ситуации равновесия одного из игроков не дает ему никаких преимуществ. Однако при одновременном отклонении обоих каждый из них может получить больший выигрыш, нежели в равновесной ситуации. Например, в ситуации $(1, 1)$, когда оба игрока выбирают первую чистую стратегию — молчать, — каждый из них теряет лишь 1.

Напомним, что по условию задачи сговор (создание коалиции) между игроками недопустим.

Совершенно ясно однако, что в рассматриваемых обстоятельствах ситуация $(0, 0)$ неустойчива — любой из узников, изменяя свою стратегию, увеличивает свой выигрыш (избегает наказания).

Пример 22. «Семейный спор» (продолжение). Выигрыши игроков A и B в этой биматричной игре задаются так:

$$\mathrm{A} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Проводя необходимые вычисления

$$C = 2 - 0 - 0 + 1 = 3, \qquad \alpha = 1 - 0 = 1,$$
$$D = 1 - 0 - 0 + 2 = 3, \qquad \beta = 2 - 0 = 2,$$

и рассуждения

$$(l)\quad \begin{aligned} (p-1)(3q-1) &\geqslant 0, \\ p(3q-1) &\geqslant 0, \end{aligned} \qquad (r)\quad \begin{aligned} (q-1)(3p-2) &\geqslant 0, \\ q(3p-2) &\geqslant 0, \end{aligned}$$

получаем, что

$$1^\circ.\ p = 1,\ q \geqslant \tfrac{1}{3}, \qquad 2^\circ.\ p = 0,\ q \leqslant \tfrac{1}{3}, \qquad 3^\circ.\ 0 < p < 1,\ q = \tfrac{1}{3}.$$
$$1^\circ.\ q = 1,\ p \geqslant \tfrac{2}{3}, \qquad 2^\circ.\ q = 0,\ p \leqslant \tfrac{2}{3}, \qquad 3^\circ.\ 0 < q < 1,\ p = \tfrac{2}{3}.$$

Геометрически полученный результат выглядит так (рис. 6).

Данная игра имеет три точки равновесия. Две из них отвечают чистым стратегиям игроков,

$$p = 1,\ q = 1: \quad H_A(1, 1) = 2,\ H_B(1, 1) = 1,$$
$$p = 0,\ q = 0: \quad H_A(0, 0) = 1,\ H_B(0, 0) = 2,$$

одна — смешанная,

$$p = \tfrac{2}{3},\ q = \tfrac{1}{3}: \quad H_A\left(\tfrac{2}{3}, \tfrac{1}{3}\right) = \tfrac{2}{3},\ H_B\left(\tfrac{2}{3}, \tfrac{1}{3}\right) = \tfrac{2}{3}.$$

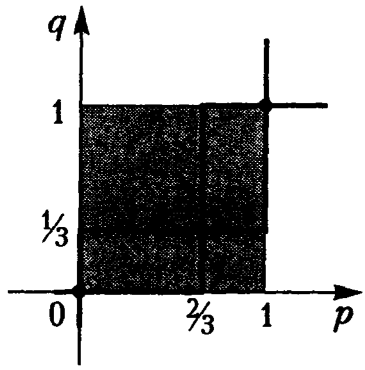


Рис. 6

Признаться, полученные результаты ставят больше вопросов, чем дают ответов.

Ситуации $(1, 1)$ и $(0, 0)$ соответствуют одноаременному выбору игроками своих первых или, соответственно, вторых стратегий, то есть определенной договоренности о совместных действиях.

Однако в данном случае есть еще одна ситуация равновесия, состоящая в выборе игроками вполне определенных смешанных стратегий. В ней оба игрока получают одинаковые выигрыши, правда, меньшие тех, которые даaали две другие равновесные ситуации.

Какой же из этих трех ситуаций равновесия следует отдать предпочтение? Какую выбрать игрокам? Если бы игроки договорились играть оба, скажем, первую чистую стратегию, причем игрок A за получение большего выигрыша, чем игрок B, заплатил бы ему $1/2$, то выигрыш каждым полутора единиц можно было бы считать и выгодным и справедливым. Однако в рамках теории бескоалиционных игр такого рода дележи не рассматриваются.

Вполне ясно, что для выбора каждым из игроков своей линии поведения (напомним, что подобная ситуация может повторяться, и повторяется, многократно) необходимы либо расширение возможностей, имеющихся у игроков, либо иные, измененные критерии.

Наконец, обратимся к последнему из приведенных выше примеров биматричных игр — «Студент — Преподаватель».

Пример 23. «Студент — Преподаватель» (продолжение). Впечатления у каждого из них относительно результатов общения в матричном виде выглядят следующим образом

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -3 \\ -2 & -1 \end{pmatrix}.$$

Проводя необходимые вычисления

$$C = 2 + 1 - 1 + 0 = 2, \qquad \alpha = 0 + 1 = 1,$$
$$D = 1 + 3 + 2 - 1 = 5, \qquad \beta = -1 + 2 = 1,$$

и рассуждения

$$(l)\ \begin{array}{r} (p-1)(2q-1) \geqslant 0, \\ p(2q-1) \geqslant 0, \end{array} \qquad (r)\ \begin{array}{r} (q-1)(5p-1) \geqslant 0, \\ q(5p-1) \geqslant 0, \end{array}$$

получаем, что

$$1^\circ.\ p = 1,\ q \geqslant \tfrac{1}{2}, \qquad 2^\circ.\ p = 0,\ q \leqslant \tfrac{1}{2}, \qquad 3^\circ.\ 0 < p < 1,\ q = \tfrac{1}{2},$$
$$1^\circ.\ q = 1,\ p \geqslant \tfrac{1}{5}, \qquad 2^\circ.\ q = 0,\ p \leqslant \tfrac{1}{5}, \qquad 3^\circ.\ 0 < q < 1,\ p = \tfrac{1}{5}$$

(рис. 7).

Число точек пересечения у зигзагов (равновесных ситуаций) равно трем. Две из них отвечают чистым стратегиям игроков,

$$p = 1,\ q = 1: \quad H_A(1, 1) = 2,\ H_B(1, 1) = 1,$$
$$p = 0,\ q = 0: \quad H_A(0, 0) = 0,\ H_B(0, 0) = -1,$$

одна — смешанная,

$$p = \tfrac{1}{5},\ q = \tfrac{1}{2}: \quad H_A\left(\tfrac{1}{5}, \tfrac{1}{2}\right) = \tfrac{1}{2},\ H_B\left(\tfrac{1}{5}, \tfrac{1}{2}\right) = -\tfrac{7}{5}.$$

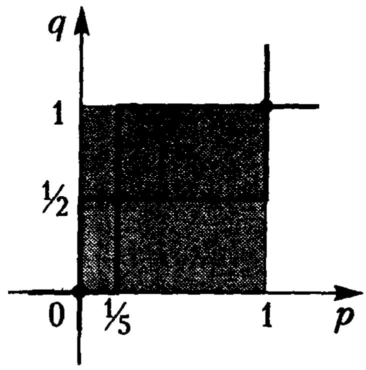


Рис. 7

В данной задаче, а отличие от предыдущей, все довольно ясно: наилучшим является выбор каждым из игроков первой чистой стратегии — хорошо подготовиться к зачету и поставить зачет.

Как нетрудно заметить, тем самым в этой задаче реализуется весьма редкая возможность, когда функции выигрыша каждого из игроков достигают своих максимумов одновременно. Выгодность такой ситуации совершенно ясна. Ее устойчивость также вполне очевидна: любое отклонение от ситуации $(1, 1)$ одним из игроков или обоими игроками может привести разве что к уменьшению их выигрышей.

§ 5. Оптимальность по Парето

Содержательные представления об устойчивости, выгодности и справедливости многообразны. Выше мы рассматривали проявление устойчивости через равновесие. Существует и иной вариант устойчивости ситуации, в большей степени, чем равновесность, отражающий черты ее выгодности. Это оптимальность по Парето.

5.1. Множество Парето

Рассмотрим на плоскости (U, V) множество Ω (рис. 8). Каждая его точка обладает одним из следующих свойств: либо все точки, ближайшие к ней, принадлежат множеству Ω (такая точка называется *внутренней* точкой множества Ω), либо сколь угодно близко от нее расположены как точки множества Ω, так и точки, множеству Ω не принадлежащие (такие точки называются *граничными* точками множества Ω). Граничная точка может как принадлежать множеству Ω, так и не принадлежать. Здесь мы будем рассматривать только такие множества, которым принадлежат все точки границы. Множество всех граничных точек множества называется его *границей*. Обозначение: $\partial\Omega$.

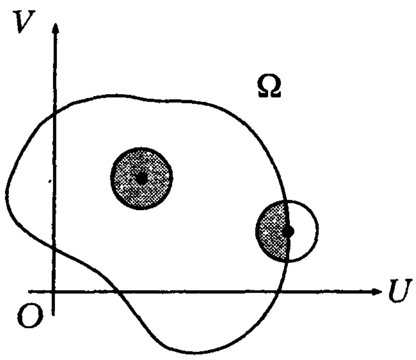


Рис. 8

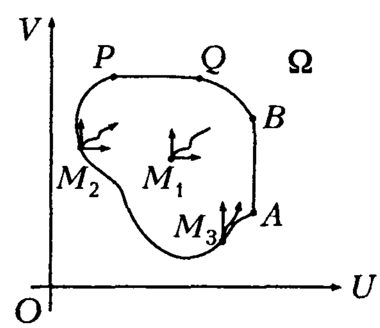


Рис. 9

Пусть M — произвольная точка множества Ω, внутренняя или граничная, и (U, V) — ее координаты. Поставим следующий вопрос: можно ли, оставаясь во множестве Ω, переместиться из точки M в близкую точку так, чтобы при этом увеличились обе ее координаты. Если M — внутренняя точка, то это бесспорно возможно. Если же M — граничная точка, то такое возможно не всегда (рис. 9). Из точек M_1, M_2, M_3 это сделать можно, но уже из точек вертикального отрезка AB можно переместиться, увеличивая лишь координату V (координата U при этом остается неизменной). Перемещая точку горизонтального отрезка PQ вправо, мы увеличиваем координату U (при этом координата V сохраняет свое значение). Что же касается дуги BQ, то перемещение вдоль нее способно увеличить лишь одну из координат при одновременном уменьшении другой.

Тем самым, точки множества Ω можно разбить на три класса:

- в первый класс относятся точки, которые, оставаясь во множестве Ω, можно сдвинуть так, чтобы одновременно увеличились обе координаты (в этот класс попадают все внутренние точки множества Ω и часть его граничных точек),
- второй класс образуют точки, перемещением которых по множеству Ω можно увеличить только одну из координат при сохранении значения второй (вертикальный отрезок AB и горизонтальный отрезок PQ на границе множества Ω),

- в третий класс попадут точки, перемещение которых по множеству Ω способно лишь уменьшить либо одну из координат, либо обе (дуга BQ границы $\partial\Omega$) (рис. 10).

Множество точек третьего класса называется *множеством Парето*, или *границей Парето* данного множества Ω (выделено на рис. 10).

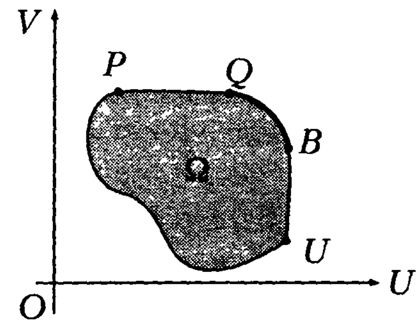


Рис. 10

5.2. Метод идеальной точки

Пусть на плоскости (x, y) задано множество ω (рис. 11) и в каждой точке этого множества определены две непрерывные функции

$$U = \Phi(x, y) \quad \text{и} \quad V = \Psi(x, y).$$

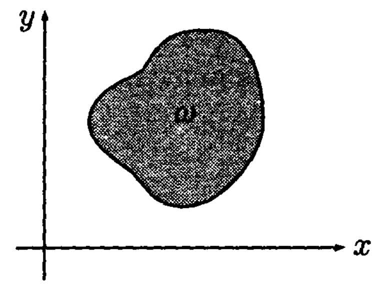


Рис. 11

Рассмотрим следующую задачу.

Во множестве ω найти точку (x_*, y_*), в которой

$$\Phi(x_*, y_*) = \max_{(x, y)\in\omega} \Phi(x, y) \quad \text{и} \quad \Psi(x_*, y*) = \max_{(x, y)\in\omega} \Psi(x, y).$$

Обычно это записывается так

$$\Phi(x, y) \to \max \quad \text{и} \quad \Psi(x, y) \to \max$$
$$(x, y) \in \omega.$$

Сразу же отметим, что в общем случае поставленная задача решения не имеет.

В самом деле, нарисуем на плоскости (U, V) все точки, координаты которых вычисляются по формулам

$$U = \Phi(x, y), \quad V = \Psi(x, y), \quad (x, y) \in \omega.$$

Из рис. 12 видно, что наибольшее значение $U - U_{\max}$ — и наибольше значение $V - V_{\max}$ — достигаются в разных точках, а точка с координатами

$$(U_{\max}, V_{\max})$$

лежит вне множества Ω.

Тем самым, в исходной постановке задача, вообще говоря, неразрешима — удовлетворить обоим требованиями одновременно невозможно. И, следовательно, нужно искать какое-то компромиссное решение.

Опишем один из путей, использующий множество Парето.

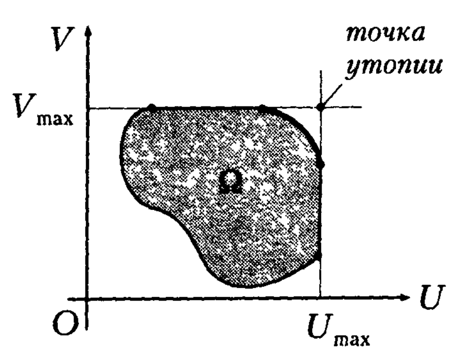


Рис 12

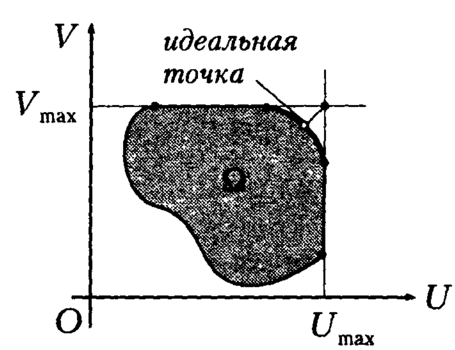


Рис. 13

Сначала на плоскости (U, V) задается целевая точка, в качеств координат которой часто выбирается сочетание наилучших значений обоих критериев U и V.

В данном случае это точка $(U_{\max}, V_{\max})$.

Вследствие того, что обычно такая точка при заданных ограничениях не реализуется, ее называют *точкой утопии*.

Затем строится множество Парето и на нем ищется точка, ближайшая к точке утопии — *идеальная точка* (рис. 13).

5.3. Оптимальность по Парето в биматричной игре

Рассмотрим биматричную игру с 2×2-матрицами

$$\mathsf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{и} \quad \mathsf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}.$$

Пусть $H_A(p, q)$ и $H_B(p, q)$ — средние выигрыши игроков A и B.

Ситуация (p_*, q_*) в биматричной игре А и В наказывается *оптимальной по Парето*, если из того, что

$$H_A(p_*, q_*) \leqslant H_A(p, q) \quad \text{и} \quad H_B(p_*, q_*) \leqslant H_B(p, q),$$

вытекают равенства

$$p = p_*, \quad q = q_*.$$

Иными словами, в оптимальной по Парето ситуации игроки не могут совместными усилиями увеличить выигрыш одного из игроков, не уменьшив при этом выигрыш другого.

Различие ситуации равновесия от ситуации, оптимальной по Парето, состоит в следующем:

- в ситуации равновесия ни один из игроков, действуя в одиночку, не может увеличить своего собственного выигрыша;
- в ситуации, оптимальной по Парето, игроки, действуя совместно, не могут (даже нестрого) увеличить выигрыш каждого.

Обратившись к игре «Дилемма узников», покажем, как практически отыскиваются оптимальные по Парето ситуации.

Напомним, что соответствующие платежные матрицы в этой игре имели следующий вид

$$\mathsf{A} = \begin{bmatrix} -1 & -9 \\ 0 & -6 \end{bmatrix}, \quad \mathsf{B} = \begin{bmatrix} -1 & 0 \\ -9 & -6 \end{bmatrix}.$$

Тем самым, на единичном квадрате

$$0 \leqslant p \leqslant 1, \quad 0 \leqslant q \leqslant 1$$

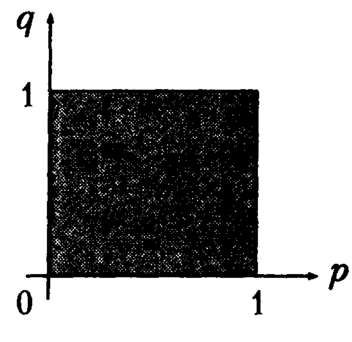


Рис. 14

(рис. 14) возможных значений вероятностей p и q заданы две функции

$$U = H_A(p, q) = -pq - 9p(1-q) - 6(1-p)(1-q),$$
$$V = H_B(p, q) = -pq - 9(1-p)q - 6(1-p)(1-q).$$

Точки с координатами (U, V), вычисленными по приведенным формулам, на плоскости (U, V) заполняют четырехугольник с вершинами $K(-1, -1)$, $L(-9, 0)$, $M(-6, -6)$ и $N(0, -9)$ (рис. 15). Граница Парето этого множества — ломаная NKL.

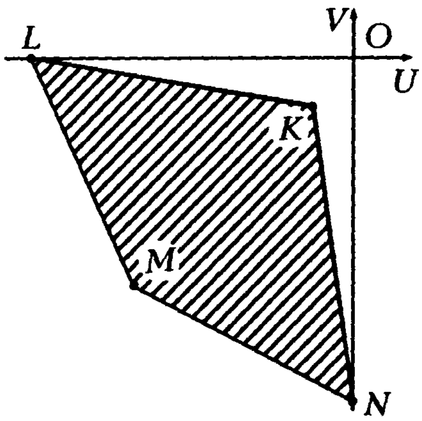


Рис. 15

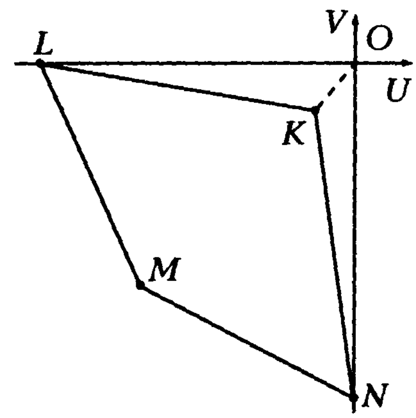


Рис. 16

Каждый из игроков заинтересован в наибольшем значении своего среднего выигрыша

$$U = H_A(p, q) \to \max \quad \text{и} \quad V = H_A(p, q) \to \max.$$

Нетрудно заметить, что в данном случае

$$U_{\max} = 0 \quad \text{и} \quad V_{\max} = 0.$$

Тем самым, точкой утопии в этой задаче является начальная точка $O(0, 0)$. Ближайшая к ней точка множества Парето — $K(-1, -1)$ (рис. 16).

Идеальная точка $K(-1, -1)$ — точка с наибольшими выигрышами для каждого из игроков — оказывается лучше, чем равновесная точка $M(-6, -6)$, и ей соответствуют чистые стратегии обоих игроков

$$p = 1, \quad q = 1.$$

§ 6. Несколько слов в заключение

На анализе полученных результатов стоит остановиться чуть подробнее.

Из приведенных примеров видно, что числа C и D из соотношений (**) на с. 273 могут быть как положительными, так и отрицательными. Они могут, в частности, даже обращаться в нуль.

Рассмотрим однако наиболее интересный в приложениях случай, когда ни C ни D нулю не равны, т. е.

$$CD \neq 0.$$

Тогда, как нетрудно видеть, точка равновесия определяется парой

$$p = \frac{\beta}{D}, \quad q = \frac{\alpha}{C}.$$

Эти формулы являются весьма примечательными: в равновесной ситуации выбор игрока A полностью определяется элементами платежной матрицы игрока B,

$$p = \frac{b_{22} - b_{21}}{b_{11} - b_{12} - b_{21} + b_{22}}$$

(и не зависит от элементов его собственной платежной матрицы), а выбор игрока B в равновесной ситуации полностью определяется элементами платежной матрицы игрока A,

$$q = \frac{a_{22} - a_{12}}{a_{11} - a_{12} - a_{21} + a_{22}}$$

(и не зависит от элементов его собственной платежной матрицы).

Иными словами, равновесная ситуация обоих игроков определяется не столько стремлением увеличить собственный выигрыш, сколько желанием держать под контролем выигрыш другого игрока (минимизировать этот выигрыш). И если, например, заменить в биматричной игре матрицу выплат игроку A, а матрицу выплат игроку B оставить прежней, то игрок A никак не изменит своего «равновесного» поведения (просто не обратит внимания на эту замену), в то время как игрок B изменит свою стратегию на новую, равновесную.

Таким образом, в биматричной (неантагонистической) игре мы вновь встречаемся с антагонизмом. Правда, теперь это уже не *антагонизм интересов* (как это было в антагонистической, матричной игре), а *антагонизм поведения*.

Отметим, что в биматричными играх (в отличие от матричных) при наличии нескольких ситуаций равновесия средний выигрыш игрока в разных равновесных ситуациях различен (напомним, что в матричной игре выигрыш игрока один и тот же вне зависимости от количества точек равновесия).

Но если средние выигрыши разнятся, то какую равновесную ситуацию следует считать оптимальной?

Наконец, еще одно, не менее интересное. Вспомним, с какими трудностями мы столкнулись, пытаясь перевести эмоциональные оценки результатов общения студент-преподаватель в количественные показатели. В целом сохраняя основные соотношения, эти количественные оценки могут, конечно, изменяться как от студента к студенту, так и от преподавателя к преподавателю. Однако, если эти изменения будут не слишком значительными — элементы платежной матрицы пошевельнутся «слегка» — то слегка пошевелятся и зигзаги, не изменяя ни своей общей формы, ни взаимного расположения, а, значит, число равновесных ситуаций не изменится. Впрочем, сказанное относится лишь к случае, когда множество ситуаций равновесия конечно и состоит из нечетного числа точек (одной или трех). Как принято говорить в подобных случаях, это число *устойчиво относительно малых шевелений*.

Конечно, в некоторых биматричных играх равновесные ситуации случаются и в чистых стратегиях (в последнем из разобранных примеров таких ситуаций даже две). И (в принципе это совсем нетрудно) можно дать определение ситуации равновесия в чистых стратегиях. Найти ее (если она, конечно, существует) — дело довольно простое. Но, как показывают приведенные примеры, во-первых, чистой ситуации равновесия может вовсе не быть, а, во-вторых, даже при ее наличии не исключено существование равновесных ситуаций в смешанных стратегиях. И желая найти их все, неизбежно приходится обращаться к описанному выше подходу.

Заключение

Реальные конфликтные ситуации приводят к разным видам игр. Различны и способы их анализа и разрешения.

Мы остановились лишь на трех видах игр — матричных, позиционных и биматричных. Сделанный выбор обусловлен тем, что уже здесь можно наглядно показать, какой смысл вкладывается в термин *игра* и чем именно занимается *теория игр*, а также познакомить с относительно несложным математическим *инструментарием*, опирающимся на ключевые понятия вероятности, матрицы и координаты и позволяющим разрешать простейшие из этих видов игр.

Вместе с тем, нам не хотелось бы, чтобы у читателя сложилось впечатление, что доступными анализу могут быть только игры, описанные выше. Существует обширный, содержательный и интересный, привлекающий неослабевающее внимание исследователей класс игр, в которых хотя бы один из игроков имеет бесконечное множество возможных стратегий — *бесконечные* игры.

В этом — заключительном — разделе мы приведем три примера бесконечных игр: непрерывной игры на единичном квадрате (*непрерывными* называются бесконечные игры, в которых функции выигрышей непрерывно зависят от стратегий, выбираемых игроками), дуэли (играми *с выбором момента времени*, или играми *типа дуэли*, называются игры, характеризующиеся моментом выбора хода и вероятностями получения выигрыша в зависимости от времени, прошедшего от начала игры до момента выбора) и дифференциальной игры поиска (в *дифференциальных* играх допускается делать ходы непрерывно и связывать поведение игроков условиями, описываемыми дифференциальными уравнениями). При этом мы ограничимся, как и ранее, играми двух лиц и проведем лишь постановку задач (описание возможностей в поведении игроков и построение функций выигрышей), хотя для каждой из приводимых игр разработаны достаточно эффективные подходы к построению их решения.

Борьба за рынки (игра на единичном квадрате)

Одна из конкурирующих фирм (игрок A) пытается вытеснить другую фирму (игрока B) с одного из двух рынков сбыта. Предположим, что общая сумма средств, выделенная на это игроком A, равна 1. Типичной стратегией игрока A является разделение выделенной суммы на две части: x $(0 \leqslant x \leqslant 1)$ для первого рынка и $1 - x$ для второго. Подобным образом выглядят и стратегии игрока B: выделение им части y $(0 \leqslant y \leqslant 1)$ своей суммы на первый рынок и $1 - y$ на второй.

Будем считать, что если игрок A добился превосходства на одном из рынков (на другом превосходства автоматически добивается игрок B), то он вытесняет противника с этого рынка и получает выигрыш, пропорциональный избытку вложенных средств с коэффициентом, характеризующим важность рынка (этот коэффициент равен k_1 для первого рынка и k_2 для второго). Тогда функция выигрыша $H(x, y)$ игрока A определяется формулой

$$H(x, y) = \begin{cases} k_1(x - y), & \text{если} \quad 0 \leqslant y < x \leqslant 1, \\ 0, & \text{если} \quad x = y, \\ k_2(y - x), & \text{если} \quad 0 \leqslant x < y \leqslant 1. \end{cases}$$

Ясно, что функция выигрыша игрока B равна $-H(x, y)$.

Дуэль

Два дуэлянта (игроки A и B) начинают сходиться в момент времени $t = 0$. У каждого пистолет заряжен одной пулей. Они встретятся в момент времени $t = 1$ (если только ни один из них не застрелит другого раньше). Каждый из дуэлянтов может выстрелить, когда пожелает. Если при этом одному из них удастся поразить противника, а самому остаться невредимым, то он становится победителем (его выигрыш равен 1) и дуэль тут же прекращается. Если оба промахнутся, дуэль закончится вничью (выигрыш каждого из игроков равен 0). Если оба выстрелят одновременно и каждый поразит противника, то дуэль также считается окончившейся вничью.

Если игрок A произведет выстрел в момент времени x $(0 \leqslant x \leqslant 1)$, то его выстрел будет успешным с вероятностью $p(x)$. Подобным же образом, выстрел игрока B в момент времени y $(0 \leqslant y \leqslant 1)$ будет успешным с вероятностью $q(y)$. При условии $x < y$ игрок A выиграет с вероятностью $p(x)$, а проиграет с вероятностью $(1-p(x))q(y)$. Тем самым, его средний выигрыш при $x < y$ будет равен

$$H(x, y) = 1 \cdot p(x) + (-1) \cdot (1 - p(x))q(y) = p(x) - q(y) + p(x)q(y).$$

С другой стороны, если $x > y$, его средний выигрыш будет равен

$$H(x, y) = 1 \cdot (1 - p(x))q(y) + (-1) \cdot p(x) = p(x) - q(y) - p(x)q(y).$$

При $x = y$ средний выигрыш

$$H(x, x) = 1 \cdot p(x) + (-1) \cdot q(y) = p(x) - q(x).$$

Таким образом, функция выигрыша $H(x, y)$ игрока A имеет вид

$$H(x, y) = \begin{cases} p(x) - q(y) + p(x)q(y), & \text{если} \quad 0 \leqslant x < y \leqslant 1, \\ p(x) - q(x), & \text{если} \quad x = y, \\ p(x) - q(y) - p(x)q(y), & \text{если} \quad 0 \leqslant y < x \leqslant 1, \end{cases}$$

и антагонистическая игра задана. В частности, если игроки стреляют без промаха, $p(x) = q(y) \equiv 1$,

$$H(x, y) = \begin{cases} 1, & \text{если} \quad 0 \leqslant x < y \leqslant 1, \\ 0, & \text{если} \quad x = y, \\ -1, & \text{если} \quad 0 \leqslant y < x \leqslant 1. \end{cases}$$

Дифференциальная игра поиска

Ищущий (игрок A) стремится обнаружить уклоняющегося (игрока B). Оба игрока перемещаются с постоянными скалярными скоростями (α и β соответственно) по плоскости внутри некоторой поисковой области Ω. В любой момент времени каждый из игроков управляет своим перемещением, задавая направление вектора скорости. Пусть (x_A, y_A) и (x_B, y_B) — координаты игроков. Тогда имеем

$$\frac{dx_A}{dt} = \alpha \cos \varphi, \qquad \frac{dx_B}{dt} = \beta \cos \psi,$$

$$\frac{dy_A}{dt} = \alpha \sin \varphi, \qquad \frac{dy_B}{dt} = \beta \sin \psi.$$

Игра поиска заканчивается в тот момент, когда игроки сблизятся на расстояние $l > 0$, иными словами, когда будет выполнено неравенство

$$(x_A - x_B)^2 + (y_A - y_B)^2 \leqslant l^2.$$

В случае успешного обнаружения выигрыш игрока A считается равным 1.

Построение решения в этой игре существенно зависит от характера и степени информированности игроков.

* * *

Все, о чем говорилось в этой книге, — примеры *бескоалиционных* игр, когда любые соглашения, обмен информацией, побочные платежи, совместный выбор стратегий запрещены.

Другой важный класс составляют *кооперативные* игры, в которых разрешены самые разнообразные формы сотрудничества. Возможность соглашений между игроками оказывает существенное влияние на исход игры. Если допустить, например, в игре «Дилемма узников» совместный выбор стратегий, то исход игры может оказаться совсем иным. При наличии побочных платежей по-иному окончится и «Семейный спор».

Вне поля наших рассмотрений остались игры с одним участником, а также игры с участием трех и более игроков; последние особенно интересны, но они и трудны.

Задачи к разделу

1. Найдите нижнюю цену игры, верхнюю цену игры, определите седловые точки, оптимальные чистые стратегии и цену игры (если они существуют):

$$\textbf{а)}\ \begin{pmatrix} -1 & 3 & -2 & 1 \\ 2 & 3 & 4 & -1 \\ 7 & 1 & -5 & 2 \\ -8 & 4 & 3 & -1 \end{pmatrix}, \quad \textbf{б)}\ \begin{pmatrix} 1 & 2 & 1 & 3 & 7 \\ 3 & -5 & -1 & 6 & -2 \\ 3 & 2 & 1 & 4 & 2 \\ -1 & 7 & -2 & -3 & 4 \end{pmatrix}.$$

2. Найдите решения следующих матричных игр:

$$\textbf{а)}\ \begin{pmatrix} -5 & 8 \\ 4 & -7 \end{pmatrix}, \quad \textbf{б)}\ \begin{pmatrix} -1 & 1 & -1 & 2 \\ 0 & -1 & 2 & -2 \end{pmatrix}, \quad \textbf{в)}\ \begin{pmatrix} 4 & 2 & 3 & -1 \\ -4 & 0 & -2 & 2 \end{pmatrix},$$

$$\textbf{г)}\ \begin{pmatrix} 1 & 4 \\ 3 & -2 \\ 0 & 5 \end{pmatrix}, \quad \textbf{д)}\ \begin{pmatrix} 4 & 2 & 0 & 1 & 0 & 2 \\ 4 & 2 & 0 & 2 & 1 & 1 \\ 4 & 3 & 1 & 3 & 1 & 2 \\ 4 & 3 & 4 & -1 & 2 & 2 \\ 4 & 3 & 3 & -2 & 2 & 2 \end{pmatrix}.$$

3. Найдите точные и приближенные решения следующих матричных игр:

$$\textbf{а)}\ \begin{pmatrix} 4 & -2 & 0 \\ 0 & 1 & 2 \\ -3 & 3 & -1 \end{pmatrix}, \quad \textbf{б)}\ \begin{pmatrix} 1 & 0 & -1 \\ 1 & -1 & 2 \\ 0 & 2 & -4 \end{pmatrix}.$$

4. Дайте графическое представление, приведите к нормальной форме и найдите точное решение позиционной игры со следующей функцией выигрышей $W(x, y, z)$:

$$\begin{aligned} W(1,1,1) &= 2, & W(2,1,1) &= -1, \\ W(1,1,2) &= -2, & W(2,1,2) &= 3. \\ W(1,2,1) &= 1, & W(2,2,1) &= 0, \\ W(1,2,2) &= 0, & W(2,2,2) &= -3. \end{aligned}$$

а) *1-й ход* делает игрок A: он выбирает число x из множества двух чисел $\{1, 2\}$. *2-й ход* делает игрок B: не зная о выборе игрока A на 1-м ходе, он выбирает число y из множества двух чисел $\{1, 2\}$. *3-й ход* делает игрок A: он выбирает число z из множества двух чисел $\{1, 2\}$, зная значение y, выбранное игроком B на 2-м ходе, но не помня собственного выбора x на 1-м ходе.

б) *1-й ход* делает игрок A: он выбирает число x из множества двух чисел $\{1, 2\}$. *2-й ход* делает игрок B: зная выбор игрока A на 1-м ходе, он выбирает число y из множества двух чисел $\{1, 2\}$. *3-й ход* делает игрок A: он выбирает число z из множества двух чисел $\{1, 2\}$, не зная значения y, выбранного игроком B на 2-м ходе, но помня собственный выбор x на 1-м ходе.

в) *1-й ход* производится случайно: игрок O выбирает число x, равное 1 с вероятностью 0,3 и равное 2 с вероятностью 0,7. *2-й ход* делает игрок A: он выбирает число y из множества двух чисел $\{1, 2\}$, зная результат случайного выбора на 1-м ходе. *3-й ход* делает игрок B: он выбирает число z из множества двух чисел $\{1, 2\}$, зная выбор y игрока A на 2-м ходе, но не зная случайного выбора x на 1-м ходе.

5. Найдите решение биматричной игры:

а) $A = \begin{pmatrix} 6 & 2 \\ 2 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 6 \\ 8 & 2 \end{pmatrix}$.

б) $A = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$.

в) $A = \begin{pmatrix} 5 & 0 \\ 10 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 5 & 10 \\ 0 & 1 \end{pmatrix}$.

г) «ястреб—голубь» $A = \begin{pmatrix} \dfrac{a-b}{2} & 0 \\ a & \dfrac{a}{2} \end{pmatrix}$, $B = \begin{pmatrix} \dfrac{a-b}{2} & a \\ 0 & \dfrac{a}{2} \end{pmatrix}$.

Найдите ситуации равновесия в каждом из двух случаев: 1) $a > b$, 2) $a < b$.

Ответы

1. а) $\alpha = -2$, $\beta = 2$; **б)** $\nu = 1$, $\{A_1, B_3\}$ и $\{A_3, B_3\}$. **2. а)** $P = \left\{\frac{11}{24}, \frac{13}{24}\right\}$, $Q = \left\{\frac{5}{8}, \frac{3}{8}\right\}$, $\nu = -\frac{1}{8}$; **б)** $P = \left\{\frac{2}{5}, \frac{3}{5}\right\}$, $Q = \left\{\frac{4}{5}, 0, 0, \frac{1}{5}\right\}$, $\nu = -\frac{2}{5}$; **в)** $P = \left\{\frac{6}{11}, \frac{5}{11}\right\}$, $Q = \left\{\frac{3}{11}, 0, 0, \frac{8}{11}\right\}$, $\nu = \frac{4}{11}$; **г)** $P = \left\{\frac{5}{8}, \frac{3}{8}, 0\right\}$, $Q = \left\{\frac{3}{4}, \frac{1}{4}\right\}$, $\nu = \frac{7}{4}$; **д)** $P = \left\{0, 0, \frac{3}{5}, \frac{2}{5}, 0\right\}$, $Q = \left\{0, 0, 0, \frac{1}{5}, \frac{4}{5}, 0\right\}$, $\nu = \frac{7}{5}$. **3. а)** $P \approx \{0{,}14; 0{,}86; 0\}$, $Q \approx \{0{,}43; 0{,}57; 0\}$, $\nu \approx 0{,}46$; **б)** $P \approx \{0{,}05; 0{,}62; 0{,}33\}$, $Q \approx \{0; 0{,}67; 0{,}33\}$, $\nu \approx 1{,}00$. **5. а)** $p = \frac{3}{5}$, $q = \frac{1}{3}$: $H_A\left(\frac{3}{5}, \frac{1}{3}\right) = \frac{10}{3}$, $H_B\left(\frac{3}{5}, \frac{1}{3}\right) = \frac{22}{5}$; **б)** $p = 1$, $q = 1$: $H_A(1, 1) = 4$, $H_B(1, 1) = 1$; $p = 0$, $q = 0$: $H_A(0, 0) = 1$, $H_B(0, 0) = 4$; $p = \frac{1}{5}$, $q = \frac{4}{5}$: $H_A\left(\frac{1}{5}, \frac{4}{5}\right) = \frac{4}{5}$, $H_B\left(\frac{1}{5}, \frac{4}{5}\right) = -\frac{4}{5}$; **в)** $p = 0$, $q = 0$: $H_A(0, 0) = 1$, $H_B(0, 0) = 1$; **г)** 1) $a > b$: $p = 1$, $q = 1$, 2) $a < b$: $p = 1$, $q = 0$; $p = 0$, $q = 1$; $p = \frac{a}{b}$, $q = \frac{a}{b}$.

ПОДСЧЕТ ЧИСЛЕННОСТЕЙ ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ. ПРИНЦИП УМНОЖЕНИЯ. КОМБИНАТОРИКА

В теории вероятностей часто возникает задача о подсчете количества некоторых комбинаций. Наиболее естественно к этой задаче приводит случай конечных или счетных пространств элементарных событий. Рассмотрим постановку, которая охватывает наиболее важные с точки зрения приложений случаи.

Имеется некоторое конечное множество объектов произвольной природы

$$a_1, a_2, \dots, a_N.$$

Производится выбор $n \leqslant N$ объектов по заранее фиксированному правилу. Требуется определить количество различных способов, которыми этот выбор можно осуществить. Для решения поставленной задачи применим так называемый принцип умножения. Если выбираемая совокупность

$$a_{i_1}, a_{i_2}, \dots, a_{i_n} \tag{1}$$

такова, что символ, стоящий на первом месте, может быть выбран m_1 способами, после его выбора символ, стоящий на втором месте — m_2 способами и, вообще,

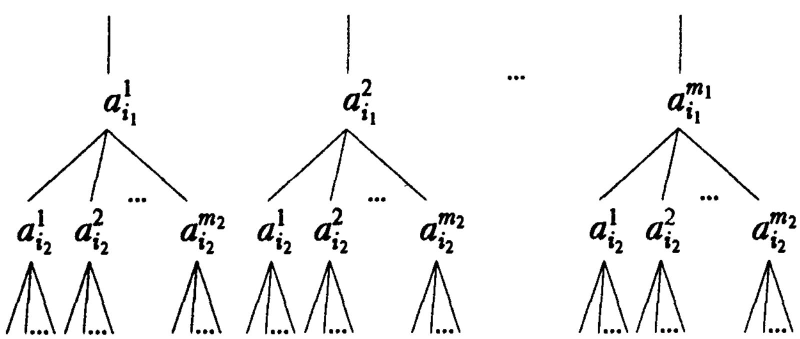


Рис. 1

символ, стоящий на j-м месте может быть выбран m_j способами, после выбора предыдущих символов, то общее количество способов, которыми может быть выбрана последовательность (1), равно

$$k = m_1 \cdot m_2 \cdot \ldots \cdot m_n. \tag{2}$$

Справедливость этого принципа легко усмотреть из рис. 1.

Задача 1. Перестановки. Пусть элементы $a_{i_1}, a_{i_2}, \dots, a_{i_n}$ различны. *Перестановкой* из этих элементов называется их любой упорядоченный набор. Требуется определить количество всех различных перестановок P_N из данных N элементов.

Поскольку все элементы в перестановке различны, то первый элемент может быть выбран N разными способами. После этого элемент, стоящий на втором месте, может быть выбран $N-1$ различными способами и т. д. Из принципа умножения немедленно следует соотношение

$$P_N = N \cdot (N-1) \cdot (N-2) \dots 2 \cdot 1 = N! \tag{3}$$

Задача 2. Размещения. *Размещением* длины n из N элементов a_j называется упорядоченная совокупность n различных элементов a_j. Заметим, что два размещения считаются разными, если они отличаются либо входящими в них элементами, либо порядком элементов, либо их количеством. Требуется определить количество размещений длины n, которые можно образовать из данных N элементов a_j.

Рассуждаем аналогично решению задачи 1. Первый элемент может быть выбран N различными способами, второй — $(N-1)$ способами и т. д. Получаем формулу

$$A_N^n = N(N-1) \cdot \ldots \cdot (N-n+1). \tag{4}$$

Заметим, что формула (3) является частным случаем формулы (4) при $n = N$.

Задача 3. Сочетания. *Сочетанием* длины n из N элементов называется неупорядоченная совокупность различных n элементов из данных N. Сочетания одинаковой длины отличаются друг от друга только входящими в них элементами (в то время как размещения отличаются еще и их порядком). Требуется определить количество различных сочетаний C_N^n длины n, которые можно образовать из данных N элементов a_j.

Легко видеть, что каждому сочетанию отвечает $n!$ различных размещений, составленных из те же элементов, но отличающихся друг от друга порядком элементов. Отсюда

$$C_N^n = \binom{N}{n} = \frac{A_N^n}{n!} = \frac{N(N-1) \cdot \ldots \cdot (N-n+1)}{n!} = \frac{N!}{n!(N-n)!}. \tag{5}$$

Замечание. В различных задачах оказываются полезными следующие соотношения для чисел C_N^n:

1. $C_N^n = C_N^{N-n}$. 2. $\sum\limits_{j=0}^{N} C_N^j = 2^N$. 3. $\sum\limits_{j=0}^{N} (-1)^j C_N^j = 0$.

Задача 4. Найти количество различных «слов», которые можно образовать из букв п, е, р, е, с, т, а, н, о, в, к, а.

«Словом» будем называть последовательность всех перечисленных выше букв, взятую в произвольном порядке. Считая все буквы различными, получим $12!$ «слов». Однако не все из этих «слов» различны — буквы «е» и «я» повторяются и от их перестановки друг с другом построенное слово не меняется. Букву «е» можно переставить в уже сформированном слове $2!$ способами, аналогично дело обстоит и с буквой «а» — слова при этом не меняются. Поэтому количество различных «слов» K дается соотношением

$$K = \frac{12!}{2!2!}.$$

Задача 5. Выбор без возвращения. В урне содержится $(m+n)$ шаров: m — красных и n — черных. Случайным образом выбирают k шаров. Сколькими различными способами можно осуществить этот выбор? Сколько различных наборов содержит ровно $r \leqslant k$ красных шаров? ($r \leqslant m$, порядок шаров в наборе несуществен).

Для первой части задачи цвет шаров безразличен: у нас есть $m+n$ шаров и требуется определить, сколькими способами из них можно выбрать k шаров. Используя результат задачи 3, получаем, что нужное число способов равно C_{m+n}^k.

Для того, чтобы получить совокупность шаров, содержащую r красных шаров, поступим следующим образом: выберем произвольным образом r шаров из данных m красных — этот выбор можно осуществить C_m^r способами, аналогично $k-r$ черных шаров из n черных можно выбрать C_n^{k-r} способами. Поскольку любой набор из r красных шаров можно объединить с любым набором $k-r$ черных шаров, то из принципа умножения следует, что количество различных наборов k шаров, содержащих ровно r красных дается формулой $C_m^r \cdot C_n^{k-r}$.

Задача 6. Выбор с возвращением. В урне содержится $m+n$ шаров: m — красных и n — черных. Производится следующий выбор k шаров: случайным образом вынимают шар, отмечают его цвет и возвращают в урну. Сколько различных k-наборов можно получить таким образом? Сколько k-наборов содержат ровно $r \leqslant k$ красных шаров? (Порядок шаров существен.)

Заметим, что извлекаемый шар может оказаться любым из $m+n$ шаров, содержащихся в урне. Поэтому, в соответствии с принципом умножения, искомое количество k-наборов дается соотношением $(m+n)^k$. Далее, рассуждая аналогично задаче 5, получаем, что r красных шаров можно выбрать m^r различными способами, $k-r$ черных — n^{k-r} способами. Теперь, так как здесь учитывается порядок шаров, надо $m^r n^{k-r}$ умножить на число выборов r мест в k-элементном множестве, т. е. на C_k^r. В итоге получим $C_k^r m^r n^{k-r}$.

Замечание. Следует обратить внимание на различные вероятностные ситуации, описываемые выбором с возвращением и выбором без возвращения. В случае выбора с возвращением, вероятность извлечения шара остается неизменной в течение всего эксперимента, в то время как при выборе без возвращения эта вероятность изменяется с каждым извлеченным шаром.

Задача 7. В ящике находится $2n$ шаров, помеченных номерами от 1 до n так, что один и тот же номер имеют два шара. Из ящика извлекают случайным образом r шаров, $r \leqslant n$. Найти количество различных наборов из r шаров с разными номерами.

Осуществляется выбор r элементов из данных $2n$. Заметим, что если выбран шар с номером j, то прочие шары в наборе уже не могут иметь этого номера и потому выбор последующих шаров должен осуществляться из набора $2n-2$ шаров. Таким образом, первый шар может быть выбран $2n$ способами, второй — $(2n-2)$ способами, третий — $(2n-4)$ способами и, вообще, k-й шар может быть выбран $2n-2(k-1)$ способами. Всего упорядоченных наборов из r шаров будет

$$S = 2n(2n-2)\cdot\ldots\cdot(2n-2(r-1)).$$

Но с точки зрения условия задачи не все эти наборы различны, так как они отличаются не только входящими в них элементами (т. е. номерами шаров), но и порядком, в котором элементы стоят в выбранной совокупности. Поскольку порядок шаров нам безразличен, то число S следует уменьшить в $r!$ раз — ровно столько различающихся порядком наборов можно образовать из r шаров с фиксированными номерами. В результате получим

$$\frac{2n(2n-2)\cdot\ldots\cdot(2n-2(r-1))}{r!} = \frac{2^r n!}{r!(n-r)!}$$

различных наборов.

Задача 8. В условиях предыдущей задачи случайным образом извлекают $2r < 2n$ шаров. Найти количество различных наборов, содержащих ровно $m < r$ парных шаров $(2r \leqslant m+n)$.

Очевидно, что набор из $2r$ шаров, содержащий ровно m пар шаров, содержит еще $2r-2m$ различных шаров (т. е. непарных) из оставшихся $2n-2m$ шаров. Заметим, что всего существует C_n^m различных m пар шаров из данных n пар. Оставшиеся $2r-2m$ шаров можно выбрать (см. задачу 7)

$$\frac{(2n-2m)\cdot(2n-2m-2)\cdot\ldots\cdot(2n+2m-4r+2)}{(2r-2m)!} = \frac{2^{2r-2m}(n-m)!}{(2r-2m)!(m+n-2r)!}$$

способами. Следовательно, искомое количество способов дается формулой

$$S = \frac{2^{2r-2m} C_n^m (n-m)!}{(2r-2m)!(m+n-2r)!}.$$

Предметный указатель

Л

М

Н

О

П

Р

С

Т

У

Ф

Х

Ц

Ч

Э

Оглавление