МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ
БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«ВЯТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

Факультет экономики и менеджмента
Кафедра математического моделирования в экономике

А.В. Ряттель

Теория вероятностей и математическая статистика

Учебное пособие (лекционный материал)

Киров 2011

УДК 330

Допущено к изданию методическим советом факультета
экономики и менеджмента ФГБОУ ВПО «ВятГУ»
в качестве учебного пособия
для студентов направления подготовки
080500.62 «Бизнес-информатика»
всех профилей подготовки и форм обучения

Рецензент:
кандидат физико-математических наук, доцент кафедры прикладной математики и информатики ФГБОУ ВПО «ВятГУ» М.Ю. Здоровенко

Р Ряттель А.В.

Теория вероятностей и математическая статистика: учебное пособие для студентов направления подготовки 080500.62 «Бизнес-информатика» всех профилей подготовки и форм обучения/ А.В. Ряттель.- Киров: ПРИП ФГБОУ ВПО «ВятГУ», 2011.- 178 с.

УДК 330

В издании излагается лекционный материал по дисциплине «Теория вероятностей и математическая статистика», предусмотренных учебным планом дисциплины.

Редактор Е.В. Кайгородцева

Содержание лекционного курса «Теория вероятностей и математическая статистика»

§1. Случайные события (4 часа)

Теория вероятностей – наука о случайных явлениях (событиях). Какие явления можно назвать случайными? Ответ, который можно дать сразу, – это события, не поддающиеся объяснению. А если их объяснить, то перестанут ли события быть случайными? Приведем несколько примеров.

Пример 1. Саша Иванов - средний студент и обычно дает правильные ответы лишь на половину экзаменационных билетов. На очередном экзамене Саша на билет ответил и получил положительную оценку. Какие события можно считать случайными:

а) Саше попался «хороший» билет – событие A;

б) Саша ответил на билет – событие B;

в) Саша сдал экзамен – событие C.

Событие A – случайное, так как Саша мог взять и «плохой» билет, но почему ему попался «хороший» - это объяснить трудно. Событие B - не случайно, так как Саша может ответить только на «хороший» билет. Событие C – случайное, так как состоит из нескольких событий и, по меньшей мере, одно из них случайное (событие A).

Пример 2. Саша и Маша разыгрывают билет на концерт. Какие из следующих событий можно считать случайными?

а) Только Саша выиграл билет – событие A;

б) Только Маша выиграла билет – событие B;

в) Саша или Маша выиграли билет – событие C;

г) Оба выиграли билет – событие D.

События A и B – случайные; событие C – не случайное, так как оно обязательно произойдет. Событие D – не случайное, так как оно никогда, при данных условиях, произойти не может.

Тем не менее, все эти события имеют смысл и изучаются в теории вероятностей (при этом событие C называется **достоверным**, а событие D – **невозможным**).

Пример 3. Рассмотрим работу столовой, в плане обслуживания клиентов. Моменты прихода посетителей (событие A) заранее предсказать невозможно, более того, время, затрачиваемое клиентами на обед (событие B), для разных клиентов - различное. Следовательно, события A и B можно считать случайными, а процесс обслуживания клиентов – случайным процессом (или случайным явлением).

Пример 4. Английский ботаник Браун (Brown), изучая под микроскопом пыльцу хвойных растений в воде, открыл, что взвешенные частицы двигаются беспорядочно под действием толчков со стороны молекул окружающей среды.

Это беспорядочное движение частиц А. Эйнштейн назвал (1905-1906) броуновским (от имени Brown), а позднее Н. Винер создал теорию винеровских процессов (1920-1930), являющихся непрерывным аналогом броуновского движения. Выяснилось, что частица размером в один микрон (10^{-4} см) испытывает за секунду со стороны молекул более 10^{15} ударов. Чтобы определить траекторию частицы, нужно за секунду измерить параметры 10^{15} ударов. Это практически невозможно. Таким образом, мы вправе броуновское движение считать **случайным**. Поступив так, Эйнштейн открыл новые возможности изучения броуновского движения, а заодно, и тайн микромира.

Здесь случайность проявляется как незнание или неумение получить достоверную информацию о движении частиц.

Из примеров следует, что случайные события не существуют в единственном числе, у каждого из них должно быть, по меньшей мере, альтернативное событие.

Таким образом, под **случайными** будем понимать наблюдаемые **события, каждое из которых обладает возможностью реализоваться в данном наблюдении, но реализуется лишь одно из них**.

Кроме того, мы предполагаем, что любое случайное событие «за бесконечное время реализуется бесконечное число раз».

Это условие хотя и образное, но достаточно точно отражает суть понятия случайного события в теории вероятностей.

В самом деле, изучая случайное событие, нам важно знать не только факт его появления, но и то, как часто случайное событие появляется в сравнении с другими, то есть знать его вероятность. Для этого необходимо иметь достаточный набор статистических данных, но это уже предмет математической статистики.

Итак, можно утверждать, что в природе нет ни одного физического явления, которое бы не содержало элемент случайности, а это означает, что, изучая случайность, мы познаем закономерности окружающего нас мира. Современная теория вероятностей редко применяется для изучения отдельного явления, состоящего из небольшого числа факторов. Основной ее задачей является выявление **закономерностей** в массовых случайных явлениях и их изучение.

Вероятностный (статистический) метод изучает явления с общих позиций, помогает специалистам познать их суть, не останавливаясь на несущественных деталях. Это является большим преимуществом по сравнению с точными методами других наук. Не следует думать, что теория вероятностей противопоставляет себя другим наукам, наоборот, она их дополняет и развивает.

Например, вводя в детерминированную модель случайную составляющую, часто получают более точные и глубокие результаты изучаемого физического процесса. Эффективным оказывается и вероятностный подход для явлений, которые декларируются случайными, независимо от того, являются они таковыми или нет.

В теории вероятностей такой подход называется **рандомизацией** (random – случайный).

Исторические сведения

Принято считать, что теория вероятностей своему возникновению обязана азартным играм, однако аналогичные права на нее может предъявить, например, и страхование. В любом случае, теория вероятностей и математическая статистика появились благодаря потребностям практики.

Первые серьезные работы по теории вероятностей возникли в середине XVII века из переписки Паскаля (1623 – 1662) и Ферма (1601 – 1665) при изучении азартных игр. Одним из основателей современной теории вероятностей является Яков Бернулли (1654 – 1705). Изложение основ теории вероятностей принадлежит Муавру (1667 – 1754) и Лапласу (1749 – 1827).

С именем Гаусса (1777 – 1855) связан один из самых фундаментальных законов теории вероятностей – нормальный закон, а с именем Пуассона (1781 – 1840) – закон Пуассона. Кроме того, Пуассону принадлежит теорема закона больших чисел, обобщающая теорему Бернулли.

Большой вклад в развитие теории вероятностей и математической статистики внесли русские и советские математики.

Автором первого учебника по теории вероятностей на русском языке и учителем П.Л. Чебышева (1821 – 1894) был В.Я. Буняковский (1804 – 1889). П.Л. Чебышеву принадлежат фундаментальные работы по закону больших чисел, А.А. Маркову (1856 – 1922) – авторство создания теории стохастических процессов (марковских процессов). Его ученик А.М. Ляпунов (1857 – 1918) доказал центральную предельную теорему при достаточно общих условиях, разработал метод характеристических функций.

Среди советских математиков, сформировавших теорию вероятностей как математическую науку, следует отметить С.Н. Бернштейна (1880 – 1968), А. Я. Хинчина (1894 – 1959) (стационарные случайные процессы, теория массового обслуживания), А.Н. Колмогорова (1903 – 1987) (автора аксиоматического построения теории вероятностей; ему принадлежат фундаментальные работы по теории стохастических процессов), Б. В. Гнеденко (р.1911) (теория массового обслуживания, стохастические процессы), А.А. Боровкова (р. 1931) (теория массового обслуживания).

П.1. Понятие случайного события

Как уже отмечалось в предисловии, теория вероятностей изучает массовые случайные явления. А что же такое случай? Как к нему относиться? Если нам повезло, говорим о счастливом случае, если нет, то это – несчастливый случай.

Однако, в целом, к случайностям мы относимся отрицательно, поскольку заранее не знаем, как себя эта случайность проявит. Конечно, случайность портила и портит жизнь человека, но она ему и помогает. Для борьбы со случайностью разработаны эффективные методы. Выясняется, что описание и формализация случайности является одним из самых мощных инструментов научного описания мира.

Под случаем мы обычно понимаем либо ограниченность необходимой информации, либо неумение её использовать, либо полное отсутствие информации (за исключением той информации, что она отсутствует). Итак, будем считать, что случай, случайность - понятия для нас интуитивно ясные.

Разобьем случайность на два класса: «хорошая» случайность – когда можно выявить какие-то закономерности её проявления (то есть имеет смысл говорить о её количественной оценке), и «дурная» случайность – закономерностей никаких нет (мистика, колдовство, прилёт инопланетян и др.)

«Хорошую» случайность, в отличие от «дурной», можно формализовать. Изучением именно «хорошей» случайности, и только ею, занимается современная **теория вероятностей** – математическая наука, которая по известным вероятностям одних случайных событий позволяет находить вероятности других случайных событий.

Случайные события будем называть просто событиями, а их количественную оценку - вероятностью события, которая является числом из промежутка [0;1]. Прежде всего, мы научимся получать комбинации событий и вычислять соответствующие им вероятности. Это позволит нам адекватно оценить действительность, прогнозировать результаты, вырабатывать оптимальную стратегию поведения.

П.2 Операции над событиями

Первоначальным и, тем самым, математически неопределяемым понятием для нас, является пространство Ω случайных событий*. Оно состоит из

---

* Мы рассматриваем дискретное пространство, являющееся частным случаем более общего пространства элементарных событий. Для изучаемого курса теории вероятностей этого достаточно. Более общие случаи будем оговаривать особо.

элементарных событий (точек) ω_1, ω_2, ..., ω_n,… представляющих неразложимый исход теоретического эксперимента. Количество точек из Ω может быть **конечно** или **счетно**. Стандартная запись: $\Omega=\{\omega_1, \omega_2, ..., \omega_n, ...\}$. Любой конечный (или даже счетный) набор элементарных событий, например, $\{\omega_{i_1}, \omega_{i_2}, ..., \omega_{i_m}\} \subset \Omega$, назовем случайным событием. Случайные события обозначают буквами: A, B, ….

Пусть $A = \{\omega_{i_1}, \omega_{i_2}, ..., \omega_{i_m}\} \subset \Omega$. Будем говорить, что событие, A произошло, если наступило одно из элементарных событий, $\{\omega_{i_1}, \omega_{i_2}, ..., \omega_{i_m}\}$.

Объединением (суммой) двух событий A и B называется событие $A\cup B$, состоящее из элементарных событий, принадлежащих хотя бы одному из событий A или B.

Пересечением (произведением) событий A и B называется событие $A\cap B$, состоящее из элементарных событий, содержащихся одновременно в событиях A и B.

Дополнением (разностью) событий A и B называется событие $A\backslash B$, состоящее из элементарных событий события A, не содержащихся в событии B.

Пусть $A \subset \Omega$, тогда **противоположным** событию A называется событие $\overline{A} \subset \Omega$, состоящее из элементарных событий пространства Ω, не содержащихся в событии A, то есть $\overline{A} = \Omega \setminus A$.

Пусть A, $B \subset \Omega$. Они образуют **алгебру** событий, если:

1) $A \cup B \subset \Omega$,

2) $A \cap B \subset \Omega$,

3) $\overline{A} \subset \Omega$.

Кроме того, если выполнено условие

4) $A \setminus B \subset \Omega$,

то имеем **поле** событий.

Очевидно обобщение на любое конечное число событий $A_1, A_2, \ldots, A_n \subset \Omega$.

Событие, которое никогда не происходит (то есть не содержит ни одной точки), называется **невозможным**, обозначается символом $\varnothing$ и

$$\forall\ (A \subset \Omega) \Rightarrow (\varnothing \subset A).$$

Событие $A = \Omega$ всегда происходит и называется **достоверным**, при этом полагаем $\overline{\Omega} = \varnothing$.

События $A_1, A_2 \subset \Omega$ **несовместны**, если $A_1 \cap A_2 = \varnothing$ (то есть события A_1 и A_2 не имеют общих точек).

События $A_1, A_2, \ldots, A_n$ образуют **полную группу**, если $\Omega \subseteq \bigcup_{i=1}^{n} A_i$, а если A_i, $i = 1, 2, \ldots, n$, попарно несовместные, то есть $\forall i \neq j$, $j = 1, 2, \ldots, n$, $A_i \cap A_j = \varnothing$, тогда $\bigcup_{i=1}^{n} A_i = \Omega$.

Если каждое появление события A влечет за собой появление события B, то говорят, что A есть часть B, то есть $A \subset B$.

Многие задачи теории вероятностей содержат бесконечное число исходов (например, точки на отрезке прямой, поверхности и др.), и мы можем столкнуться с трудностями теоретического характера, если любое подмножество отрезка или поверхности будем считать событием. Чтобы их избежать, мы вводим специальный класс $\mathcal{F}$ подмножеств, состоящий из **несчетных** множеств, где любое его подмножество есть событие. Формально это выглядит следующим образом.

Пусть пространство Ω - произвольное множество (в том числе, несчетное), а $\mathcal{F}$ класс подмножеств из множества Ω.

$\mathcal{F}$ называется **σ- алгеброй**, если

1) $\Omega \subset \mathcal{F}$,

2) $\forall A_i \subset \mathcal{F}, i \in N \Rightarrow \left(\bigcup\limits_{i=1}^{\infty} A_i \subset F \quad \text{и} \quad \bigcap\limits_{i=1}^{\infty} A_i \subset F \right)$,

3) $A \subset \mathcal{F} \Rightarrow \overline{A} \subset \mathcal{F}$.

Таким образом, алгебра событий замкнута относительно конечного числа теоретико-множественных операций (объединения, пересечения, отрицания), а σ- алгебра замкнута относительно бесконечного числа этих операций.

Замечание. По условию, класс подмножеств $\mathcal{F}$ содержится в пространстве Ω и одновременно, сам содержит это пространство. Возможность такой формализации становится понятной, если считать $\mathcal{F}$ оператором «наведения порядка» в Ω. Тогда, если, например, Ω интерпретировать как единичный объем жидкости, а $\mathcal{F}$ - как губку, то, если вся жидкость находится в губке всюду плотно, получается, что с одной стороны, губка находится в жидкости, а с другой стороны, вся жидкость находится в губке.

Мерой или количественной оценкой случайных событий из Ω служит **вероятность** p – число, удовлетворяющее следующим аксиомам.

Аксиома 1. Любому событию $A \subset \Omega$, удовлетворяющему условиям 1) – 3), поставлено в соответствие неотрицательное число $p = P\{A\}$, называемое его вероятностью.

Аксиома 2. $P\{\Omega\} = 1$.

Аксиома 3. Если события $A_1, A_2, ..., A_n, ...$ попарно несовместны, то

$$P\left\{\bigcup_{i=1}^{\infty} A_i\right\} = \sum_{i=1}^{\infty} P\{A_i\}.$$

Пространство Ω, с заданной на нем алгеброй (σ - алгеброй) событий и определенной для каждого события вероятностью, которая удовлетворяет аксиомам 1-3, является центральным понятием, определяющим аксиоматический подход к построению теории вероятностей, введенный А.Н. Колмогоровым в 30-х годах прошлого века.

Определение. Тройку $(\Omega,\mathcal{F},P)$ будем называть вероятностным пространством.

Замечание. В данном курсе теории вероятностей мы обсуждаем только такие случаи, для которых любое подмножество Ω есть событие, а потому введение σ- алгебры $\mathcal{F}$ излишне. Однако в целях конструктивности изложения мы будем писать $(\Omega,\mathcal{F},P)$, подразумевая под вероятностным пространством (Ω, P).

Следствия из аксиом

Следствие 1.

$P\{\varnothing\} = 0$.

В самом деле, имеем $\Omega = \Omega \cup \varnothing$ и $\Omega \cap \varnothing = \varnothing$, то есть Ω и $\varnothing$ несовместны. Следовательно, $1 = P\{\Omega\} = P\{\Omega \cup \varnothing\} =$ { по аксиоме 3 } $= P\{\Omega\} + P\{\varnothing\} = 1 + P\{\varnothing\}$. Отсюда $P\{\varnothing\} = 0$.▼

Следствие 2.

Если $A \subset \Omega$, то $P\{\overline{A}\} = 1 - P\{A\}$.

Доказательство сразу следует из условия $A \cup \overline{A} = \Omega$, $A \cap \overline{A} = \varnothing$.▼

Следствие 3.

Если $A \subset \Omega$, то $0 \le P\{A\} \le 1$.

В самом деле, так как $\varnothing \subset A \subset \Omega$, то $P\{\varnothing\} < P\{A\} < P\{\Omega\}$, тогда $0 < P\{A\} < 1$. Знак равенства возможен тогда, когда $A = \varnothing$ или $A = \Omega$, или $P\{A\} = 0$ и $A \ne \varnothing$.▼

Следствие 4 (Теорема сложения).

Для любых $A, B \subset \Omega$ имеет место

$\boldsymbol{P\{A \cup B\} = P\{A\} + P\{B\} - P\{A \cap B\}}$.

В самом деле, имеем

$A \cup B = A \cup (B \setminus (A \cap B))$ и $B = (A \cap B) \cup (B \setminus (A \cap B))$.

События правой части несовместные, отсюда

$$P\{A \cup B\} = P\{A\} + P\{B \setminus (A \cap B)\},$$

$$P\{B\} = P\{A \cap B\} + P\{B \setminus (A \cap B)\}.$$

Вычитая из первого равенства второе, получаем

$$P\{A \cup B\} - P\{B\} = P\{A\} - P\{A \cap B\}.\blacktriangledown$$

Следствие 5.

Для любых $A, B \subset \Omega$,

$P\{A \cup B\} \le P\{A\} + P\{B\}$.

Доказательство следует из условия $P\{A \cap B\} \ge 0$ и следствия 4. ▼

Очевидны обобщения на произвольное число событий.

Определение. События A, B из вероятностного пространства $(\Omega, \mathcal{F}, P)$ называются **независимыми**, если вероятность их совместного осуществления равна произведению вероятностей этих событий, то есть

$\mathbf{P\{A \cap B\} = P\{A\} \cdot P\{B\}}$.

Из определения сразу следует, что

1) Для любого $A \subset \Omega$ события A и Ω независимы.

2) Если $P\{B\} = 0$ и событие $A \subset \Omega$ произвольно, то B и A независимы.

3) Если события A и B_i независимы, $i = 1, 2$ и $B_1 \supset B_2$, то A и $(B_1 \setminus B_2)$ независимы.

4) Если события A и B_i независимы и B_i попарно несовместны, то есть $\forall i \ne j$ $B_i \cap B_j = \varnothing$, то A и $\bigcup_{i=1}^{n} B_i$ также независимы.

5) Событие A не зависит от самого себя тогда и только тогда, когда либо $P\{A\} = 0$, либо $P\{A\} = 1$.

6) Если события A и B несовместны, то есть $A \cap B = \varnothing$ и $P\{A\} > 0$, $P\{B\} > 0$, то A и B зависимы.

В самом деле, пусть события A и B независимы, тогда $P\{A \bigcap B\} = P\{A\} \cdot P\{B\} > 0$, но по условию $P\{A \bigcap B\} = P\{\varnothing\} = 0$. Получили противоречие, то есть A и B - зависимы.

Замечание 1. Понятие независимости в теории вероятностей имеет более глубокий смысл, чем независимость обычная. Принято считать события независимыми, если они не связаны причинно. На практике, понятие зависимости и независимости случайных событий относительно. Если события слабо связаны, и эта связь несущественно влияет на конечный результат, то такие события считают независимыми, поскольку в этом случае построение математических моделей реальных ситуаций становится много проще. Наиболее глубоко в теории вероятностей изучены именно независимые события.

Замечание 2. Из аксиоматического построения вероятности события следует, что событие случайно, если оно не достоверно и не невозможно. Это определение через отрицание и из него следует, что имеет смысл говорить о вероятности как о некотором определенном, но неизвестном нам числе. Утверждение, что вероятность события A существует, нуждается в обосновании, а если оно принято в качестве гипотезы, то в последующей проверке. Это следует учитывать при построении математических моделей реальных ситуаций.

Рассматривая вероятность события как число из промежутка [0,1], мы обычно предполагаем в какой его части это число будет находиться. И чем больше мы имеем информации о случайном событии, тем точнее предположение. Это позволяет нам определить вероятность как меру возможности (уверенности) появления случайного события.

Именно так Блез Паскаль в письме Пьеру Ферма в 1654 году написал: «Я считаю более простым и естественным принять степень уверенности в появлении достоверного события равной единице. Тем самым, возможность наступления случайных событий соизмеряется с тем, какую часть единицы они составляют».

Так впервые была формализована связь между случайным событием и числом, его измеряющим, – вероятностью.

П.3. Элементы комбинаторики

Комбинаторика – раздел математики, занимающийся решением задач, связанных с **выбором и расположением элементов из некоторой совокупности**.

В классической теории вероятностей комбинаторика, в основном, используется для выбора и подсчета числа комбинаций событий с идентичными свойствами. Кроме того, первоначально комбинаторика применялась для нахождения вероятностей событий, обладающих различного вида симметриями.

Пример 1. Сколько существует различных k - мерных векторов, координаты которых составлены из чисел множества $A = \{ 1, 2, ..., n \}$.

Решение. Будем исходить из того, что два вектора считаются равными, если соответствующие координаты представлены одинаковыми цифрами, иначе – различные.

Число различных k -мерных векторов находим следующим образом.

Первой координатой может являться любое из n чисел множества A, второй - также любое из n чисел, то есть, для каждого фиксированного числа $r = 1, 2, ..., n$ первой координаты имеем n вариантов для второй. Таким образом, всего имеем $n \times n = n^2$ двумерных различных векторов, далее по индукции получаем, что всего различных k -мерных векторов будет n^k.

Пример 2. Сколько существует различных трехзначных чисел?

Решение. Всего цифр десять: 1, 2, 3, 4, 5, 6, 7, 8, 9, 0. На первом месте может быть любая цифра, кроме нуля, на втором и третьем месте любая из десяти цифр. Следовательно, всего различных чисел $9 \cdot 10^2 = 900$.

Пример 3. Сколько существует различных k - мерных векторов, у которых числовые значения координат , взятых из множества $A = \{1, 2, ..., n\}$, не повторяются.

Решение. Аналогично примеру 1, первой координатой может являться любая из n цифр множества A, второй - любая из оставшихся $(n - 1)$ цифр, не совпадающей с первой, и т.д. Таким образом, получаем всего

$$A_n^k = n \cdot (n-1) \cdot ... \cdot (n-k+1)$$

различных векторов.

В частном случае, при $k = n$ имеем $A_n^n = n \cdot (n-1) \cdot ... \cdot 2 \cdot 1$ различных векторов. Это число обозначается $n! = 1 \cdot 2 \cdot 3 \cdot ... \cdot n$ – (эн - факториал).

Замечание. Часто $n!$ называют перестановками, так как $n!$ количественно определяет число различных перестановок элементов, из которых они состоят. Например, число перестановок трехтомного собрания сочинений равно шести: (1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1), где цифра означает номер тома.

Пример 4. Сколько существует различных k - мерных векторов, у которых числовые значения координат, взятых из множества $A = \{1, 2, ..., n\}$, не только не повторяются, но и их координаты принадлежат различным подмножествам множества A. Напомним, что два множества считаются различными, если они отличаются хотя бы одним элементом.

Решение. Пусть x – число таких k - мерных векторов. Возьмем один из них. Всего существует $k!$ перестановок координат этого вектора. Умножая $k!$ на x, получим число A_n^k векторов, удовлетворяющих условию примера 3, но тогда

$$x \cdot k! = A_n^k.$$

Отсюда искомое число векторов равно

$$x = A_n^k / k!,$$

или

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

Если $k > n$, то $x = 0$.

Каждый из примеров представляет собой достаточно распространенный способ выбора в комбинаторике.

Мы будем придерживаться «урновой» схемы: имеется сосуд, в котором находятся n тщательно перемешанных шаров различающихся только своими порядковыми номерами. Если из урны извлечено k шаров, то будем говорить, что имеем выборку объема k. Шары из урны извлекаются случайным образом, подобно лототрону, при этом извлечение шаров может осуществляться с возвращением или без возвращения.

При выборе с возвращением фиксируется номер шара, а сам он возвращается в урну; при выборе без возвращения - шар в урну не возвращается, то есть выборка не содержит повторяющихся шаров.

Итак, из урны последовательно извлекается k шаров. Сколько различных вариантов выборки объема k можно получить, если выбор осуществляется:

а) с возвращением, и порядок следования шаров в выборке важен. Число вариантов равно n^k. Этот способ называется **простым случайным выбором,** и соответствует примеру 1;

б) без возвращения, и порядок следования элементов в выборке важен. Число вариантов равно A_n^k. Способ выбора называется **размещениями**. В соответствие с примером 3, имеем

$$A_n^k = n \cdot (n-1) \cdot \ldots \cdot (n-k+1),$$

при $k > n$, $A_n^k = 0$;

в) без возвращения, порядок следования элементов в выборке не важен. Способ выбора называется **сочетаниями**, число вариантов равно C_n^k и, в соответствие с примером 4, вычисляется по формуле:

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

при $k > n \Rightarrow C_n^k = 0$.

Рассмотрим несколько частных случаев, имеющих самостоятельное значение.

Определение. Выборкой объема k из n элементов с повторениями называется такая выборка, в которой любой из k ее элементов может повториться более одного раза.

Пример 5. Сколько существует размещений с повторениями при выборке k шаров из n?

Решение. Так как по определению любой из k шаров в размещениях может быть повторен от 1 до k раз, то всего вариантов выбора есть n^k, то есть имеет место простой случайный выбор (см. пример 1).

Пример 6. Сколько существует сочетаний с повторениями при выборе k шаров из n?

Решение. Расположим n шаров на прямой и ограничим их слева и справа вертикальными черточками | 00 ... 0| (шару соответствует 0).

Возьмем еще (k-1) черточку и произвольно распределим черточки между шарами, причем, между соседними шарами может находиться одна или более черточек. Интерпретируя две соседние черточки как ящик, получим, что число шаров между соседними черточками – это число повторных шаров в ящике. Перечисляя возможные расположения (k- 1) черточек между шарами, получим число сочетаний с повторениями.

Итак, задача свелась к следующей: имеется (n + k - 1) – мерный вектор, координаты которого состоят из n шаров и (k - 1) черточек. Так как число способов расположения (k - 1) черточек по (n + k - 1) месту равно C_{n+k-1}^{k-1} (см. пример 4), то это и есть искомое число вариантов выбора k шаров из n с повторениями.

Замечание. Формула сочетаний с повторениями используется, например, при подсчете числа решений (в целых числах, включая ноль) диофантова уравнения

$$\sum_{i=1}^{\kappa} x_i = n.$$

Число m частных производных порядка k от функции n переменных также вычисляется по формуле $m = C_{n+k-1}^{k-1}$.

Приведем некоторые свойства сочетаний.

Рассмотрим бином Ньютона

$$(a+b)^n = \sum_{k=0}^{n} C_n^k \cdot a^k \cdot b^{n-k}, \qquad (1)$$

где $C_n^0 = C_n^n = 1$, 0! = 1.

Благодаря формуле бинома Ньютона, сочетания иногда называют биномиальными коэффициентами.

Если в (1) $a = b = 1$, то получаем

$$2^n = \sum_{k=0}^{n} C_n^k,$$

если $a = -b$, будем иметь

$$0 = \sum_{k=0}^{n} (-1)^k \cdot C_n^k.$$

Если $k \le n$, то для вычисления сочетаний имеем формулу

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

В самом деле,

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

Отсюда следует, что

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

Для любого целого k и n имеем

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

В самом деле,

$$C_n^{k-1}+C_n^k=\frac{n!}{(k-1)!\cdot(n-k+1)!}+\frac{n!}{k!\cdot(n-k)!}=\frac{n!}{(k-1)!\cdot(n-k)!}\cdot\left(\frac{1}{n-k+1}+\frac{1}{k}\right)=$$
$$=\frac{n!\cdot(n+1)}{(k-1)!\cdot(n-k)!\cdot(n-k+1)\cdot k}=\frac{(n+1)!}{k!\cdot(n-k+1)!}=C_{n+1}^k.\blacktriangledown$$

Пример 7. В урне находятся n пронумерованных шаров, из которых k красные и (n - k) черные. Наудачу выбираем без возвращения r шаров. Сколько различных выборок объема r можно получить, если среди выбранных r шаров s – красных?

Решение. Разделим урну условно на две половины так, что в одной находятся k красных шаров, а в другой (n - k) черных. Среди k красных шаров s шаров можно выбрать C_k^s способами, а среди (n - k) черных шаров (r – s) шаров можно выбрать C_{n-k}^{r-s} способами. Поскольку на каждую фиксированную выборку красных шаров приходится C_{n-k}^{r-s} выборок черных, то всего выборок объема r будет $C_k^s\cdot C_{n-k}^{r-s},\ s=0,1,\dots,r$.

Замечание. Если в предыдущей задаче мы выбирали бы r шаров из n без учета их цвета, то всего различных выборок было бы C_n^r. С другой стороны, если учесть все возможные варианты выбора красных шаров s, то получаем, что всего их будет $\sum\limits_{s=0}^{r}C_k^sC_{n-k}^{r-s}$.

Таким образом, имеем формулу

$$\sum_{s=0}^{r}C_k^sC_{n-k}^{r-s}=C_n^r.$$

П.4. Вычисление вероятностей событий

Для вычисления вероятности $P\,\{A\}$ события $A\subset\Omega$ необходимо построить математическую модель изучаемого объекта, которая содержит событие A. Основой модели является вероятностное пространство ($\Omega,\mathcal{F},P$), где Ω - пространство элементарных событий ω, $\mathcal{F}$ – класс событий с введенными над ними операциями композиции, $p = P\{A\}$ – вероятность любого события A, имеющего смысл в Ω и входящего в класс событий $\mathcal{F}$. Если, например, $A = \left\{\omega_{i_1},\ \omega_{i_2}, \ldots, \omega_{i_m}\right\}$, то из аксиомы 3, вероятностей, следует, что $p = \sum_{j=1}^{m} P\left\{\omega_{i_j}\right\}$.

Таким образом, вычисление вероятности события A, сведено к вычислению вероятностей элементарных событий, его составляющих, а так как они являются «базовыми», то методы их вычисления не обязаны зависить от аксиоматики теории вероятностей.

Здесь рассмотрены три подхода к вычислению вероятностей элементарных событий:

1) классический;
2) геометрический;
3) статистический или частотный.

Классический метод вычисления вероятностей

Из аксиоматического определения вероятности следует, что вероятность существует для любого события $A \subset \Omega$, но как ее вычислить, об этом ничего не говорится, хотя известно, что для каждого элементарного события ω_i существует вероятность p_i, такая, что сумма вероятностей всех элементарных событий пространства Ω равна единице, то есть

$$\sum_{i=1} P\left\{\omega_i\right\} = \sum_{i=1} p_i = 1.$$

На использовании этого факта основан **классический** метод вычисления вероятностей случайных событий, который в силу своей специфичности, дает способ нахождения вероятностей этих событий непосредственно из аксиом.

Пусть дано фиксированное вероятностное пространство $(\Omega,\mathcal{F},P)$, в котором:

а) Ω состоит из **конечного** числа n элементарных событий,

б) каждому элементарному событию ω_i поставлена в соответствие вероятность

$$P\{\omega_i\} = \frac{1}{n},\ i = 1, 2, \ldots,\ n.$$

Рассмотрим событие $A \subset \Omega$, которое состоит из m элементарных событий: $A = \{\omega_{i_1},\ \omega_{i_2}, \ldots, \omega_{i_m}\}$, тогда из аксиомы 3 вероятностей, в силу несовместности элементарных событий, следует, что

$$P\ \{A\} = \sum_{j=1}^{m} P\ \{\omega_{i_j}\} = \frac{m}{n}.$$

Тем самым имеем формулу

$$P\{A\} = \frac{m}{n}, \tag{2}$$

которую можно интерпретировать следующим образом: вероятность событию A произойти равна отношению числа элементарных событий, благоприятствующих появлению событию A, к числу всех элементарных событий из Ω.

В этом суть классического метода вычисления вероятностей событий.

Замечание. Приписав одинаковую вероятность каждому из элементарных событий пространства Ω, мы, с одной стороны, имея вероятностное пространство и опираясь на аксиомы теории вероятностей, получили правило вычисления вероятностей любых случайных событий из пространства Ω по формуле (2), с другой стороны, это дает нам основание считать все элементарные события **равновозможными** и вычисление вероятностей любых случайных событий из Ω свести к «урновой» схеме независимо от аксиом.

Из формулы (2) следует, что вероятность события A зависит только от числа элементарных событий, из которых оно состоит и не зависит от их конкретного содержания. Таким образом, чтобы воспользоваться формулой (2), необходимо найти число точек пространства Ω и число точек, из которых состоит событие $A \subset \Omega$, но тогда это уже задача комбинаторного анализа.

Рассмотрим несколько примеров.

Пример 8. В урне из n шаров - k красных и (n - k) черных. Наудачу извлекаем без возвращения r шаров. Какова вероятность того, что в выборке из r шаров s шаров – красных?

Решение. Пусть событие $\{A\} \sim \{$в выборке из r шаров s - красных$\}$. Искомая вероятность находится по классической схеме, формула (2):

$$P\{A\} = \frac{m}{\tilde{n}},$$

где $\tilde{n}$ - число возможных выборок объема r, которые различаются хотя бы одним номером шара, а m – число выборок объема r, в которых s шаров красных. Для $\tilde{n}$, очевидно, число возможных вариантов выборки равно C_n^r, а m, как следует из примера 7, равно $C_k^s \cdot C_{n-k}^{r-s}$. Таким образом, искомая вероятность равна

$$P\{A\} = C_k^s \cdot C_{n-k}^{r-s} / C_n^r .$$

Пусть дан набор попарно несовместных событий A_s, $s = 0, 1, ..., r$, образующих полную группу, тогда

$$\sum_{s=0}^{r} P\{A_s\} = 1.$$

В этом случае говорят, что имеем **распределение** вероятностей событий A_s.

Распределения вероятностей является одним из фундаментальных понятий современной теории вероятностей и составляет основу аксиомами Колмагорова.

Определение. Распределение вероятностей

$$P\{A\} = C_k^s \cdot C_{n-k}^{r-s} / C_n^r, \ s = 0, 1, ..., r, \tag{3}$$

определяется **гипергеометрическое распределение**.

Боровков А.А. в своей книге на примере формулы (3) поясняет природу задач теории вероятностей и математической статистики следующим образом: зная состав генеральной совокупности, мы с помощью гипергеометрического распределения можем выяснить, каким может быть состав выборки – это типичная задача теории вероятностей **(прямая задача)**. В естественных науках решают обратную задачу: по составу выборок, определяют природу генеральных совокупностей – это **обратная задача,** и она, образно говоря, составляет содержание математической статистики.

Обобщением биномиальных коэффициентов (сочетаний) являются **полиномиальные** коэффициенты, которые своим названием обязаны разложению полинома вида

$$\left(\sum_{i=1}^{k} a_i\right)^n = \sum_{r_1+r_2+\ldots+r_k=n} A(r_1, r_2, \ldots, r_k)\cdot a_1^{r_1}\cdot a_2^{r_2}\cdot\ldots\cdot a_k^{r_k},$$

где $$A(r_1, r_2, \ldots, r_k) = \frac{n!}{r_1!\cdot r_2!\cdot\ldots\cdot r_k!}, \qquad (4)$$

по степеням слагаемых.

Полиномиальные коэффициенты (4) часто применяются при решении комбинаторных задач.

Теорема. Пусть имеется k различных ящиков, по которым раскладываются пронумерованные шары. Тогда число размещений шаров по ящикам так, чтобы в ящике с номером r находилось r_i шаров, $i = 1,2,\ldots k$, $\left(\sum_{i=1}^{k} r_i = n\right)$, определяется полиномиальными коэффициентами (4).

Доказательство. Поскольку порядок расположения ящиков важен, а шаров в ящиках - не важен, то для подсчета размещений шаров в любом ящике можно воспользоваться сочетаниями.

В первом ящике r_1 шаров из n можно выбрать $C_n^{r_1}$ способами, во втором ящике r_2 шаров, из оставшихся (n - r_1) можно выбрать $C_{n-r_1}^{r_2}$ способами и так далее, в (k – 1) ящик r_{k-1} шаров выбираем $C_{n-(r_1+r_2+\ldots+r_{k-2})}^{r_{k-1}}$ способами; в ящик k – оставшиеся $r_k = n - \sum_{i=1}^{k-1} r_i$ шаров попадают автоматически, одним способом.

Таким образом, всего размещений будет

$$C_n^{r_1}\cdot C_{n-r_1}^{r_2}\cdot\ldots\cdot C_{n-r_1-\ldots-r_{k-2}}^{r_{k-1}}=\frac{n!\cdot(n-r_1)!\cdot\ldots\cdot(n-r_1-\ldots-r_{k-2})!}{r_1!\cdot(n-r_1)!\cdot r_2!\cdot(n-r_2)!\cdot\ldots\cdot r_{k-1}!\cdot(n-r_1-\ldots-r_{k-1})}=$$

$$=\frac{n!}{r_1!\cdot r_2!\cdot\ldots\cdot r_k!}.\blacktriangledown$$

Пример. По n ящикам случайно распределяются n шаров. Считая, что ящики и шары различимы, найти вероятности следующих событий:

а) {все ящики не пустые} = A_0;

б) {один ящик пуст} = A_1;

в) {два ящика пустых} = A_2;

г) {три ящика пустых } = A_3;

д) {(n-1) – ящик пуст} = A_4.

Решить задачу для случая n = 5.

Решение. Из условия следует, что распределение шаров по ящикам есть простой случайный выбор, следовательно, всех вариантов n^n.

Прежде, чем считать благоприятные варианты, опишем общий подход к их нахождению. Расположим (в порядке возрастания номеров) ящики, в которых находятся неразличимые шары, например,

333221…1.

Эта последовательность означает, что в первом, втором и третьем ящиках по три шара, в четвертом и пятом по два шара, в остальных (n – 5) ящиках по одному шару. Всего таких размещений шаров по ящикам будет $n!\ /\ (3!\cdot 2!\cdot(n-5)!)$. Так как шары на самом деле различимы, то на каждую такую комбинацию будем иметь $n!/(3!\cdot 3!\cdot 3!\cdot 2!\cdot 2!)$ размещений шаров. Таким образом, всего вариантов будет $(n!\cdot n!)/(3!\cdot 2!\cdot(n-5)!\cdot 3!\cdot 3!\cdot 3!\cdot 2!\cdot 2!)$.

Переходим к решению по пунктам примера:

а) так как в каждом ящике находится по одному шару, то имеем последовательность 111…11, для которой число размещений равно $n!/\ n!$ = 1.

Если шары различимы, то имеем $n!/\ 1!$ размещений, следовательно, всего вариантов $m = 1 \cdot n! = n!$, отсюда

$$P\{A_0\} = \frac{n!}{n^n}.$$

б) если один ящик пуст, то какой-то ящик содержит два шара, тогда имеем последовательность 211…10, для которой число размещений равно $n! \cdot (n-2)!$. Так как шары различимы, то для каждой такой комбинации имеем $n!/\ 2!$ размещений. Всего вариантов

$$m = \frac{n! \cdot n!}{2! \cdot (n-2)!} = n! \cdot C_n^2,$$

тогда

$$P\{A_1\} = \frac{n! \cdot C_n^2}{n^n}.$$

в) если два ящика пусты, то имеем две последовательности: 311…100 и 221…100. Для первой число размещений равно $n!/\ (2! \cdot (n-3)!)$. На каждую такую комбинацию имеем $n!/\ 3!$ размещений шаров. Итак, для первой последовательности, число вариантов равно

$$k_1 = \frac{n!}{2!(n-3)!} \cdot \frac{n!}{3!} = \frac{n!}{2!} \cdot C_n^3.$$

Для второй последовательности всего вариантов будет

$$k_2 = \frac{n!}{2! \cdot 2! \cdot (n-4)!} \cdot \frac{n!}{2! \cdot 2!} = \frac{n! \cdot 4! \cdot C_n^4}{2! \cdot 2! \cdot 2! \cdot 2!}.$$

Окончательно имеем

$$m = k_1 + k_2 = \frac{n!}{2!} \cdot C_n^3 + 3 \cdot \frac{n! \cdot C_n^4}{2!}.$$

Отсюда

$$P\{A_2\} = \frac{n!}{2!\,n^n}\left(C_n^3 + 3C_n^4\right).$$

г) для трех пустых ящиков будет три последовательности: 411…1000, либо 3211…1000, либо 22211…1000.

Для первой последовательности имеем

$$k_1 = n!/(3!\cdot(n-4)!)\cdot n!/4! = \frac{n!}{3!}C_n^4 .$$

Для второй последовательности

$$k_2 = n!/(3!\cdot(n-5)!\cdot n!/3!\cdot 2! = \frac{n!}{3!}\cdot C_n^5\cdot\frac{5!}{3!\cdot 2!} .$$

Для третьей последовательности получаем

$$k_3 = n!/(3!\cdot 3!\cdot(n-6)!)\cdot n!/2!\cdot 2!\cdot 2! = \frac{n!\cdot C_n^6\cdot 6!}{3!\cdot 3!\cdot 2!\cdot 2!\cdot 2!} .$$

Всего вариантов $m = k_1 + k_2 + k_3$,

или

$$m = \frac{n!}{3!}\left(C_n^4 + C_n^5\cdot C_5^3 + \frac{C_n^6\cdot 6!}{3!\cdot 2!\cdot 2!\cdot 2!}\right).$$

Искомая вероятность равна

$$P\{A_3\} = \frac{n!(C_n^4 + C_n^5\cdot C_5^3 + 15C_n^6)}{3!\cdot n^n} .$$

д) если (n -1) ящик пуст, то все шары должны находиться в одном из ящиков. Очевидно, что число комбинаций равно

$$m = \frac{n!}{(n-1)!}\cdot\frac{n!}{n!} = n .$$

Соответствующая этому событию вероятность равна

$$P\{A_4\} = \frac{n}{n^n} = \frac{1}{n^{n-1}} .$$

При n = 5, имеем

$$P\ \{A_0\} = \frac{5!}{5^5} = \frac{24}{625},$$

$$P\ \{A_1\} = \frac{5!\cdot C_5^2}{5^5} = \frac{48}{125},$$

$$P\{A_2\} = \frac{5!}{2!\cdot 5^5}\left(C_5^3 + 3C_5^4\right) = \frac{12}{25},$$

$$P\{A_3\} = \frac{5!}{3!\cdot 5^5}\left(C_5^4 + C_5^5 \cdot C_5^3\right) = \frac{12}{125},$$

$$P\{A_4\} = \frac{5}{5^5} = \frac{1}{625}.$$

Заметим, что при $n = 5$ события A_i должны образовывать полную группу, что соответствует действительности. В самом деле

$$\sum_{i=1}^{4} P\{A_i\} = \frac{24}{625} + \frac{48}{125} + \frac{12}{25} + \frac{12}{125} + \frac{1}{625} = 1.$$

Геометрический метод вычисления вероятностей

Недостаток классического метода вычисления вероятностей в том, что он рассматривает конечное число равновозможных событий.

И если можно еще этот метод расширить на счетное число событий, то на большее его возможностей недостаточно. Однако идеи классического метода можно использовать на геометрических образах и, тем самым, рассматривать несчетные множества событий.

Пусть дана область D^n из пространства R^n, $n = 1, 2, 3$, с определенной на ней мере – $mes\ D^n$ (мера прямой – длина, мера плоскости – площадь, мера пространства – объем).

В области D^n выделяется часть A^n (вообще говоря, неодносвязная) с мерой $mes\ A^n$.

В область D^n наудачу бросается точка. Какова вероятность того, что эта точка попадет в область A^n? Считая, как и при классическом подходе, попадание точки в область, пропорциональной только ее мере, будем иметь

$$P\{A\} = \frac{mes\ A^n}{mes\ D^n},$$

где область D^n соответствует пространству элементарных событий Ω, с той разницей, что D^n не нормирована.

Пример 1. Вычислить вероятность того, что для наудачу взятого значения $x \in [0,2\pi)$, значение $y = \sqrt{0,5 - \sin^2 x}$ существует.

Решение. Обозначим через A искомое событие, а его геометрический образ через $\widetilde{A}$. Значение y существует, если $0,5 - \sin^2 x \ge 0$, то есть $|\sin x| \le \frac{1}{\sqrt{2}}$, для $x \in [\,0,\, 2\pi\,)$.

В силу симметрии, в качестве области D достаточно взять промежуток $[0,\pi]$, тогда $mes\, D = \pi - 0 = \pi$.

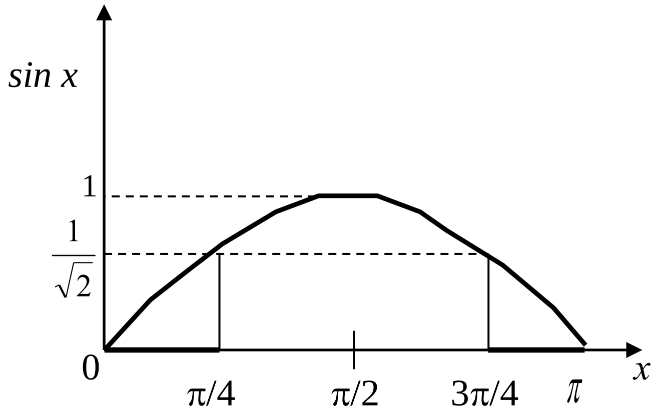
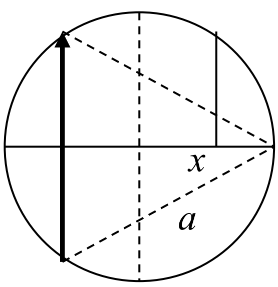


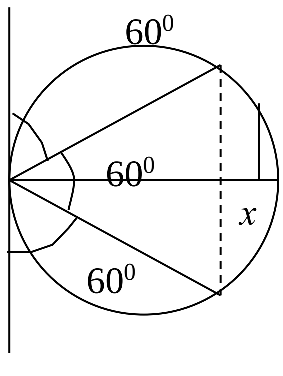
Рис. 1

Из рис. 1 видно, что область $\widetilde{A} = \left\{ \left[0; \frac{\pi}{4}\right] \cup \left[\frac{3\pi}{4}; \pi\right] \right\}$, тогда $mes\,\widetilde{A} = \pi - 2\frac{\pi}{4} = \frac{\pi}{2}$. Окончательно, $P\{A\} = \frac{\pi}{2\pi} = 0,5$.

Пример 2 (Парадокс Бертрана). Наудачу берется хорда в круге. Чему равна вероятность того, что ее длина ℓ превосходит длину стороны вписанного равностороннего треугольника (событие A)?

Решение 1. Из соображений симметрии, не нарушая общности, зададим направление хорды (рис. 2а).





а) *б*)

Рис. 2

Проведем диаметр длиной *d*, перпендикулярный этому направлению. Очевидно, что эти и только эти хорды, пересекающие диаметр в промежутке $\left[\frac{d}{4};\frac{3d}{4}\right]$, будут превосходить стороны правильного треугольника. В самом деле, сторона правильного треугольника $a = d\sqrt{3}/2$, длина хорды находится из пропорции: $\left(\frac{\ell}{2}\right)\Big/(3d/4)=(d/4)/(\ell/2)$. Таким образом $P\{A\}=\left(\frac{3d}{4}-\frac{d}{4}\right)\Big/d=0{,}5$.

Решение 2. Из соображений симметрии, закрепим один конец хорды на окружности. Касательная к окружности в этой точке и две стороны правильного треугольника образуют углы по 60^0 каждый. Задаче удовлетворяют только хорды, попадающие в средний угол, а это третья часть окружности. Отсюда $P\{A\}=1/3$.

Решения задачи дают разные ответы, хотя логических противоречий нет. Суть в том, что в задаче не определено понятие проведения хорды наудачу. Так какой ответ верный?

Очевидно тот, который учитывает все возможные ситуации, то есть имеющий наибольшую вероятность. Ясно, что если будет построен геометрический аналог, с вероятностью превосходящей 0,5, то и ответ будет другой. Ответ $P\{A\}=0{,}5$.

Второе решение, с точки зрения теории вероятностей, дает результат для более частной задачи.

Статистическое определение вероятности

Классический и геометрический методы вычисления вероятностей событий представляют собой теоретическую схему, которая основывается на аксиомах теории вероятностей, и, тем самым, не зависит от реального объекта исследования.

Для применения этих методов необходимо владеть всей информацией о возможных исходах эксперимента (пространство Ω). На практике мы далеко не всегда можем описать пространство Ω, даже в случае равновозможности элементарных событий.

Например, вычислить вероятность всхожести семян практически невозможно, если использовать классический подход, поскольку трудно пересчитать количество зерен для посадки, да и размеры зерен влияют на их всхожесть. Можно говорить лишь о приближенных значениях вероятности всхода семян, определяя приближенно их среднее количество на единичном участке поля.

Рассмотрим вновь пример с подбрасыванием монеты. Пусть у нас есть основание считать монету несимметричной. Тогда, никакие соображения относительно вероятности выпадения герба не будут иметь решающего значения, кроме как проведение испытаний. Естественно возникает вопрос: чему равна вероятность выпадения герба для этой монеты? Пусть при $n = 1000$ подбрасываний, герб выпал 450 раз, тогда доля выпадений герба составила 0,45. Отклонение от 0,5 всего 5%. Много это или мало? Можно ли считать монету симметричной?

Ответ на эти вопросы может дать статистический метод вычисления вероятностей событий.

Пусть проведено n испытаний, в которых событие A появилось m раз.

Определение. Доля числа случаев, в которых событие A появилось, называется **частотой** появления события A и вычисляется по формуле

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

(5)

Говоря о частоте, прежде всего, считают, что результат любого испытания заранее **не предсказуем**; учитываются только те результаты, которые мы ожидали получить. Если появился новый результат, то мы должны предполагать, что он возник из равноценных начальных условий и одних и тех же начальных знаний.

Испытания должны быть **независимыми**, в том смысле, что, во-первых, каждое повторное испытание проводится при одном и том же комплексе начальных условий (строго говоря, испытания не могут быть повторены в точности, поэтому мы должны так ставить эксперименты, чтобы они казались нам одинаковыми), во-вторых, результатом эксперимента являются два исхода: событие A появилось и событие A не появилось.

Частота должна быть **устойчива**, то есть, при достаточно большом числе испытаний, значения частоты подвержены малым колебаниям, которые тем меньше, чем больше число испытаний.

Определение. Число p, к которому сходится частота, при неограниченном увеличении числа испытаний, называется **статистической** вероятностью, то есть

$$p = \lim_{n \to \infty} W_n\left(A\right),$$

где $p = P\{A\}$ – вероятность события A.

Данное определение требует комментариев. Обычно, еще до проведения испытаний, в зависимости от глубины наших знаний об объекте, мы ожидаем получить конкретный результат. Поскольку число испытаний всегда конечно, то за вероятность события A мы принимаем либо значение частоты, либо число близкое к нему (в частности, то которое мы ожидали получить). Таким образом, значение вероятности события A, полученное статистическим методом, зависит от двух факторов: частоты и субъективных знаний об объекте исследования.

Например, при достаточно большом числе испытаний с подбрасыванием монеты, незначительными отклонениями значений частоты от 0,5 можно пренебречь, если нет оснований, считать монету несимметричной, либо это отклонение не может существенно повлиять на конечный результат.

Недостаток статистического определения вероятности в том, что алгоритм ее вычисления не дает ответа на основной вопрос: «Является ли принимаемое нами значение вероятности события A ее истинным значением?». Поэтому возникает ощущение того, что вероятность не является объективной характеристикой случайного события.

Тем не менее, статистический метод является наиболее общим и универсальным подходом к вычислению вероятностей случайных событий. Например, для подтверждения симметричности монеты, Д' Бюффон подбросил ее 4040 раз (2048 раз выпал герб), Пирсон провел 24000 испытаний (12012 раз выпал герб).

П.5. Условная вероятность

Пусть имеем вероятностное пространство – $(\Omega,\mathcal{F},P)$ и события A, B $\subset \Omega$, произвольны, причем P $\{B\} > 0$.

Определение. Условной вероятностью $P\{A/B\}$ называется число, определяемое формулой:

$$P\{A/B\} = \frac{P\{A \cap B\}}{P\{B\}}, \; P\{B\} > 0. \tag{6}$$

Следует читать: $\{A/B\} \sim \left\{ \begin{matrix} \text{событие } A \text{ произойдет, при условии,} \\ \text{что } B \text{ произошло.} \end{matrix} \right\}$

Формула (6) считается определением, ниоткуда не выводится и является отражением здравого смысла. Поясним это на примере геометрического изображения событий (рис. 3).

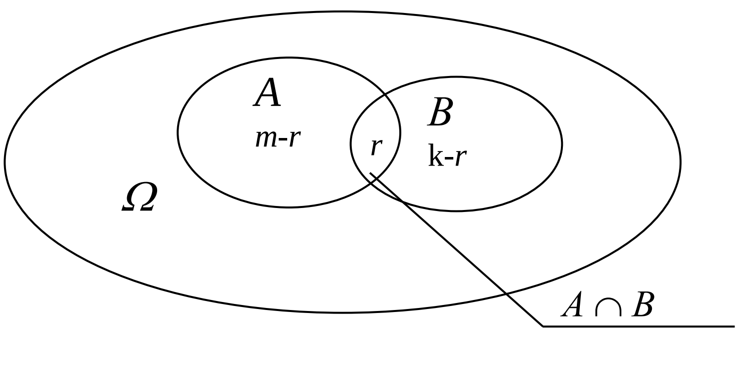


Рис. 3

Пусть пространство Ω состоит из n ($\geq m+k$) точек, равноправных между собой. Событие A насчитывает m точек, событие B - k точек и событие $A \cap B - r$ точек. По определению, событие происходит, если в результате эксперимента реализовалась какая - либо из точек, составляющих это событие. Для условной вероятности (6), фраза: «Событие A произойдет при условии, что B произошло»,- означает, что должна реализоваться одна из точек события $A \cap B$, где событие B играет роль вероятностного пространства. Следовательно, $P\{A/B\}$ – есть оценка доли участия события A в реализации события B, то есть

$$P\{A/B\} = \frac{r}{k}.$$

С другой стороны, если рассматривать все пространство Ω, то

$$P\{A \bigcap B\} = \frac{r}{n},\ P\{B\} = \frac{k}{n}.$$

По формуле (6) получаем

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

Теорема умножения. Пусть $A, B \subset \Omega$, тогда

$$P\{A \cap B\} = P\{B\} \cdot P\{A/B\} = P\{A\} \cdot P\{B/A\}. \qquad (7)$$

Доказательство. Если $P\{B\} \neq 0$, то (7) сразу следует из (6). Если же $P\{B\}=0$, то $P\{A \cap B\} = 0$ и, следовательно, (7) тривиально.▼

Определение. События $A, B \subset \Omega$ независимы, если

$$P\{A/B\} = P\{A\}. \qquad (8)$$

В самом деле, для независимых событий, по определению, имеем

$$P\{A \cap B\} = P\{A\} \cdot P\{B\}.$$

Делая в (7) замену по формуле (8), получаем эквивалентность определений независимости событий.

Определение. События независимы в совокупности, если

$$P\left\{\bigcap_{k=1}^{n} B_k\right\} = \prod_{k=1}^{n} P\left\{B_k\right\}.$$

Замечание. Попарной независимости событий (см. аксиому 3 вероятности) недостаточно для независимости их в совокупности (пример Бернштейна [2]).

П.6. Формула полной вероятности и формула Байеса (Bayes)

Формула полной вероятности

Пусть $(\Omega, \mathcal{F}, P)$ произвольное вероятностное пространство, в котором события $A, B_1, B_2, ..., B_n \subset \Omega$, удовлетворяют условиям:

1) события B_k, $k = 1, 2, \ldots, n$, попарно несовместны, то есть

$$B_i \cap B_j = \varnothing, \ \forall\ i \neq j,\ i, j = 1, 2, \ldots, n;$$

2) событие A происходит с одним, и только одним, из событий B_k, то есть

$$A = \bigcup_{k=1}^{n} (A \cap B_k);$$

тогда имеет место **формула полной вероятности**

$$P\{A\} = \sum_{k=1}^{n} P\{B_k\} \cdot P\{A / B_k\}. \tag{9}$$

Доказательство. Имеем $A = \bigcup_{k=1}^{n} (A \cap B_k)$, так как события $(A \cap B_1), \ldots, (A \cap B_n)$ попарно несовместны, то по аксиоме 3:

$$P\{A\} = P\left\{\bigcup_{k=1}^{n} (\mathrm{A} \cap B_k)\right\} = \sum_{k=1}^{n} P\{A \cap B_k\}.$$

Применяя теорему умножения получим

$$P\{A\} = \sum_{k=1}^{n} P\left\{B_k\right\} \cdot P\{A / B_k\}. \blacktriangledown$$

Замечание. События B_1, B_2, ..., B_n называют априорными гипотезами (apriory), и обычно, в литературе, на них накладывают еще дополнительное условие - они образуют полную группу событий. Это условие не является обязательным, хотя и методически оправдано, в том смысле, что при решении задач, в целях проверки правильности выбора гипотез, должно выполняться

$$\sum_{k=1}^{n} P\{B_k\} = 1.$$

На самом деле для гипотез B_k выполняется неравенство $P\{A\} \le \sum_{k=1}^{n} P\{B_k\} \le 1$.

Если заранее о вероятностях гипотез B_k,, ничего неизвестно, то каждой из n гипотез B_k приписывается одинаковая вероятность n^{-1}.

Вышесказанное в замечании проиллюстрируем рисунком.

Пусть событие A область, представляющая собой круг малого диаметра (рис. 4) пространства Ω.

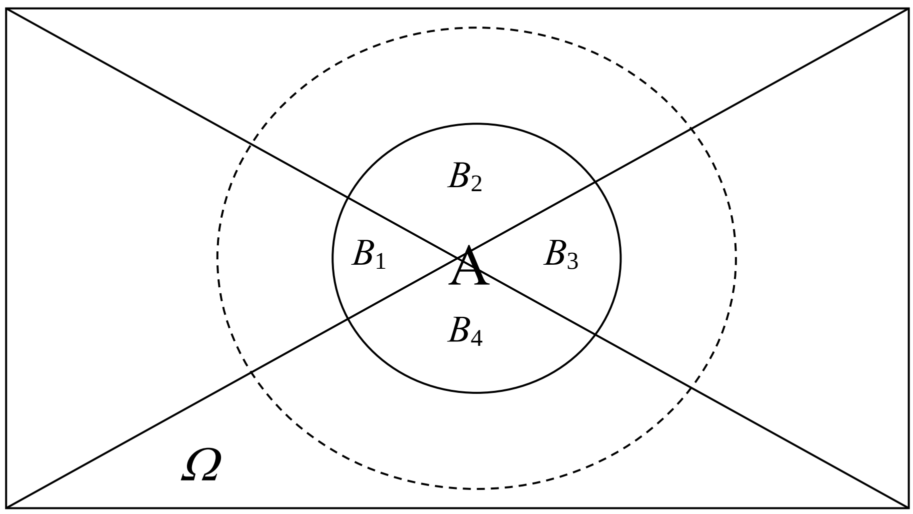


Рис. 4

Под гипотезами B_k, k =1,2,3,4, можно считать области

а) из которых состоит событие A (как на рис. 4),

тогда

$$P\left\{\bigcup_{k=1}^{4} B_k\right\} = P\{A\},$$

б) являющиеся секторами большого круга, граница которого помечена пунктиром, тогда

$$P\{A\} < P\left\{\bigcup_{k=1}^{4} B_k\right\} < 1,$$

в) являющиеся треугольниками, из которых состоит пространство Ω, тогда

$$P\left\{\bigcup_{k=1}^{4} B_k\right\} = \sum_{k=1}^{4} P\{B_k\} = 1.$$

В последнем случае несовместные события $B_к$ образуют полную группу.

Пример. Применяя формулу полной вероятности, вычислить вероятность того, что при подбрасывании симметричного кубика выпадет четная грань.

Решение 1. Вероятностное пространство $(\Omega,\mathcal{F},P)$ дискретное,

$\Omega = \{\omega_1, \omega_2, ..., \omega_6\}$; $\mathcal{F}$ - множество всех подмножеств пространства Ω,

$P\{\omega_i\} = 1/6$**, i = 1, 2, …, 6.**

Пусть $A = \{\omega_2, \omega_4, \omega_6\}$ - выпадение четной грани, $A \subset \mathcal{F}$. $B_{2\cdot i}$ $(\subset \mathcal{F})$ – выпадение грани с цифрой $2\cdot i$, $i = 1, 2, 3$, $P\{B_{2\cdot i}\} = P\{\omega_{2\cdot i}\}$.

Заметим, что здесь $\sum_{i=1}^{3} P\{B_{2\cdot i}\} = 1/2 < 1$. Далее, $P\{A / B_{2\cdot i}\} = 1$, $i = 1, 2, 3$. Применяя (9), получаем

$$P\{A\} = \frac{1}{6}\cdot 1 + \frac{1}{6}\cdot 1 + \frac{1}{6}\cdot 1 = \frac{1}{2}.$$

Решение 2. Положим $B_2 = \{\omega_2\}$, $B_4 = \{\omega_4\}$, $B_6 = \{\omega_1, \omega_3, \omega_5, \omega_6\}$.

Тогда $P\{B_2\} = \frac{1}{6}$, $P\{B_4\} = \frac{1}{6}$, $P\{B_6\} = \frac{2}{3}$, (здесь $\sum_{i=1}^{3} B_{2\cdot i} = 1$),

$$P\{A / B_2\} = P\{A / B_4\} = 1,\ P\{A / B_6\} = \frac{1}{4}.$$

Применяя (9), получим

$$P\{A\} = \frac{1}{6}\cdot 1 + \frac{1}{6}\cdot 1 + \frac{2}{3}\cdot\frac{1}{4} = \frac{1}{2}.$$

Пример показывает, что для гипотез достаточно, что их объединение содержит хотя бы те точки, из которых состоит событие A, то есть $A \subseteq \bigcup_{k=1}^{n} B_k$.

Формула Байеса

Пусть события A, B_1, ... , $B_n \subset \Omega$, удовлетворяют условиям, необходимым для получения формулы (9), тогда имеет место формула Байеса

$$P\{B_k / A\} = \frac{P\{B_k\}\cdot P\{A/B_k\}}{\sum_{k=1}^{n} P\{B_k\}\cdot P\{A/B_k\}},\ k = 1,2,...,n. \qquad (10)$$

Доказательство. Рассмотрим правую часть формулы (7) теоремы умножения вероятностей, предварительно положив $B = B_k$, $k = 1, 2, ..., n$:

$$P\{B_k\}\cdot P\{A/B_k\} = P\{A\}\cdot P\{B_k / A\}.$$

Отсюда

$$P\{B_k / A\} = \frac{P\{B_k\}\cdot P\{A/B_k\}}{P\{A\}}.$$

Учитывая, что $P\{A\} = \sum_{k=1}^{n} P\{B_k\}\cdot P\{A/B_k\}$, получаем (10). ▼

Вероятности $P\{B_k / A\}$, $k = 1,2,...,n$, называют апостериорными вероятностями гипотез B_k, поскольку оценка происходит после того, как событие A произошло.

Пример. Студенту предложили карточку с пятью вариантами ответов, причем лишь один правильный. Пусть студент правильно решит задачу с вероятностью p и неверно с вероятностью $1 - p = q$. Будем считать, что в этом случае в ответе студент напишет любой из пяти вариантов с вероятностью k =

$= 5^{-1}$. Известно, что студент получил верный ответ. Какова вероятность того, что он его угадал (событие A)[2].

Решение. Пусть $\{B_1\}$~{студент правильно решает задачу}, $\{B_2\}$~{неверно}. Требуется найти $P\{B_2/A\}$.

Имеем $P\{B_1\} = p$, $P\{B_2\} = 1 - p = q$. Далее

$$P\{A/B_1\} = 1,\ P\{A/B_2\} = k^{-1} = 5^{-1}.$$

Используя формулу (10), получаем

$$P\{B_2/A\} = \frac{P\{B_2\}\cdot P\{A/B_2\}}{P\{B_1\}\cdot P\{A/B_1\} + P\{B_2\}\cdot P\{A/B_2\}} = \frac{(1-p)\cdot k^{-1}}{p\cdot 1 + (1-p)\cdot k^{-1}} = \frac{1-p}{(k-1)\cdot p + 1} =$$

$$= \frac{1-p}{4p+1}.$$

§2. Повторение независимых испытаний (2 часа)

П.1. Формула Бернулли

Независимые испытания называются испытаниями **Бернулли**, если их можно повторить любое число раз при одних и тех же условиях, причем каждый раз возможно лишь два исхода: появление события A или события $\overline{A}$ и вероятности исходов испытаний не изменяются. Испытания Бернулли – схема теоретическая, и поэтому ее пригодность к описанию опыта должна быть обоснована.

Пусть известна вероятность появления события A (при соблюдении комплекса заданных условий), то есть $P\{A\} = p$. Положим $P\{\overline{A}\} = q$, $q = 1 - p$. Провели n независимых испытаний. Какова вероятность того, что событие A появилось ровно k раз, $k = 0, 1, 2, \ldots, n$?

Построим вероятностное пространство $(\Omega, \mathcal{F}, P)$. Любая точка (элементарное событие) пространства элементарных событий Ω представляет собой n – мерный вектор, каждая координата которого есть 1 или 0 (1- появилось событие A, 0 – событие $\overline{A}$). Очевидно, что число точек пространства Ω равно 2^n. Класс $\mathcal{F}$ - множество всех подмножеств пространства Ω.

Фиксируем k. Нас интересуют только те точки пространства, которые состоят из векторов, содержащих k единиц и $n - k$ нулей. По теореме умножения вероятностей каждая такая точка (вектор) имеет вероятность $p^k \cdot q^{n-k}$. Число точек, очевидно, равно числу способов, которыми можно расположить k единиц по n местам. Как известно, это число равно C_n^k. По теореме сложения вероятностей для несовместных событий, получаем **формулу Бернулли**:

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k},\ q = 1 - p,\ k = 0, 1, \ldots, n. \qquad (11)$$

Каждой точке пространства Ω соответствует вероятность, вычисляемая по формуле $p^k \cdot q^{n-k}$, $k = 0, 1, \ldots, n$. Для фиксированного k имеем C_n^k точек пространства Ω. Следовательно, всего точек в Ω будет $C_n^0 + C_n^1 + \ldots + C_n^n = 2^n$.

Наконец, $P\{\Omega\} = 1$, что следует из равенства

$$\sum_{k=0}^{n} C_n^k \cdot p^k \cdot q^{n-k} = (p+q)^n = 1. \qquad (12)$$

Здесь использован бином Ньютона $(a+b)^n = \sum_{k=0}^{n} C_n^k \cdot a^{n-k} \cdot b^k$, поэтому формулу (11) часто называют **биномиальным распределением**.

Таким образом, мы не только вывели формулу Бернулли (11), но и построили события, являющиеся элементарными для нового вероятностного пространства, удовлетворяющего аксиомам 1-3 вероятности.

Из формулы (11), в частности, следует, что вероятность того, что A не появится ни в одном из n испытаний, равна q^n, а вероятность того, что A появится хотя бы раз, равна $1 - q^n$.

В самом деле, получаем из (12), с учетом (11):

$$\sum_{k=1}^{n} P_n(k) = 1 - P_n(0) = 1 - q^n. \blacktriangledown$$

Покажем, что при $n \to \infty$, для любого фиксированного k, $C_n^k \cdot p^k \cdot q^{n-k} \to 0$. В самом деле, при каждом фиксированном k

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

Разделим числитель и знаменатель на n^k, тогда

$$C_n^k \cdot p^k \cdot q^{n-k} = 1\cdot\left(1-\frac{1}{n}\right)\cdot\ldots\cdot\left(1-\frac{k-1}{n}\right)\cdot\frac{p^k\cdot n^k\cdot(1-p)^{n-k}}{k!}.$$

Введем обозначения

$$b=\frac{p^k\cdot(1-p)^{-k}}{k!},\quad a^{-1}=(1-p)\quad \text{, где}\quad a>1.$$

Тогда имеем

$$C_n^k \cdot p^k \cdot q^{n-k} = 1\cdot\left(1-\frac{1}{n}\right)\cdot\ldots\cdot\left(1-\frac{k-1}{n}\right)\cdot b\cdot\frac{n^k}{a^n} \le b\cdot\frac{n^k}{a^n}\to 0,$$

так как $\lim\limits_{n\to\infty} n^k/a^n = 0$, а b - постоянная.▼

Учитывая формулу (12) и доказанное утверждение, замечаем, что каждый член суммы

$$\sum_{k=0}^{n} C_n^k \cdot p^k \cdot q^{n-k}$$

убывает, при $n\to\infty$, хотя сама сумма всегда равна единице. Так как слагаемые суммы имеют разные значения, то интерес представляет тот индекс $k = k_0$, для которого $C_n^k \cdot p^k \cdot q^{n-k}$ имеет максимальное значение.

Легко показать, что, функция $P_n(k) = C_n^k \cdot p^k \cdot q^{n-k}$ аргумента k, имеет один максимум. Тогда для нахождения k_0 можно рассмотреть отношение:

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

Далее

$$p\cdot\frac{n-k+1}{kq} = \frac{(n+1)p-kp}{kq} = \frac{(n+1)p-k+kq}{kq} = 1+\frac{(n+1)p-k}{kq}.$$

Возможны ситуации: $P_n(k) > P_n(k-1)$, тогда $(n+1)p > k$;

$P_n(k) < P_n(k-1)$, тогда $(n+1)p < k$;

$$P_n(k) = P_n(k\text{-}1), \text{ тогда } (n+1)p = k.$$

Последнее выполняется, если $(n+1)p$ – целое.

Таким образом, имеем $(n+1)p - 1 < k_0 \le (n+1)p$,

или

$$k_0 = \begin{cases} (n+1)p, & \text{если } (n+1)p \ - \text{целое}, \\ [(n+1)p], & \text{если } (n+1)p - \text{не целое}, \end{cases} \tag{13}$$

где $[x]$ – целая часть числа x

Число k_0 - называется **наивероятнейшим** числом.

Если $(n+1)p$ - целое, то имеем второе значение $k_0 = [(n+1)\cdot p - 1] = [np]$.

Пример 1. Вероятность того, что база уложится в данный день недели в норму расходов, равна ¾. Какова вероятность того, что она уложится в норму расходов в каждый из пяти дней недели?

Решение. Считая, что расходы базы практически не зависят от выбранного дня недели, воспользуемся формулой Бернулли (11). Имеем $n = 7$, $p = ¾$, $k = 5$. Тогда искомую вероятность можно обозначить как $P_7(5)$. Имеем

$$P_7(5) = C_7^5 \left(\frac{3}{4}\right)^5 \left(\frac{1}{4}\right)^2 = 21 \cdot 3^5 / 4^7 \approx 0{,}311.$$

Пример 2. Вероятность изготовления бракованной детали на станке равна 0,01. Найти вероятность того, что из 5000 деталей, изготовленных на станке

а) ровно 50 деталей бракованных,

б) бракованных деталей не более 50.

Решение. Из смысла задачи можно считать, что детали изготовлены независимо друг от друга. Воспользуемся формулой Бернулли, где $n = 5000$, $p = 0{,}01$, $q = 0{,}99$.

Имеем для а) $P_{5000}(50) = C_{5000}^{50}(0{,}01)^{50}(0{,}99)^{4950} \approx 0{,}0572$.

б) $P_{5000}\{0 \le k \le 50\} = \sum_{k=0}^{50} C_{5000}^{k}(0{,}01)^{k}(0{,}99)^{5000-k} \approx 0{,}5033$.

Чтобы получить числовые значения искомых вероятностей, требуется применение технических средств. Однако, если вычислять вероятности $P_n(k)$

непосредственно, то при больших k (или близких к нулю) их значенния будут ничтожно малы. Поэтому удобно пользоваться рекуррентными формулами.

Для этого находим наивероятнейшее число $k_0 = 51$, при котором значение вероятности $P_{5000}(51)$ максимальное. Затем, используя рекуррентные формулы: $P_{5000}(k-1) = ((n+k-1)/k \cdot q) \cdot P_{5000}(k)$, k = 51, 50, …, 1, получаем требуемые значения вероятностей.

Видно, что вычисление вероятностей, непосредственно по формулам, при больших n, k, задача трудновыполнимая, если не пользоваться техническими средствами. Числовые значения вероятностей можно получить легче, если воспользоваться приближенными методами.

П.2. Локальная теорема Муавра – Лапласа

Теорема. Пусть в n независимых испытаниях, вероятность появления события A постоянна и равна p ($0 < p < 1$), тогда имеет место асимптотическая оценка

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k} \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{npq}}\varphi(x), \tag{14}$$

где

$$x = \frac{k-np}{\sqrt{npq}},\ \varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-x^2\big/2\right)$$

Справедливость формулы (14) проиллюстрирована на рис. 5.

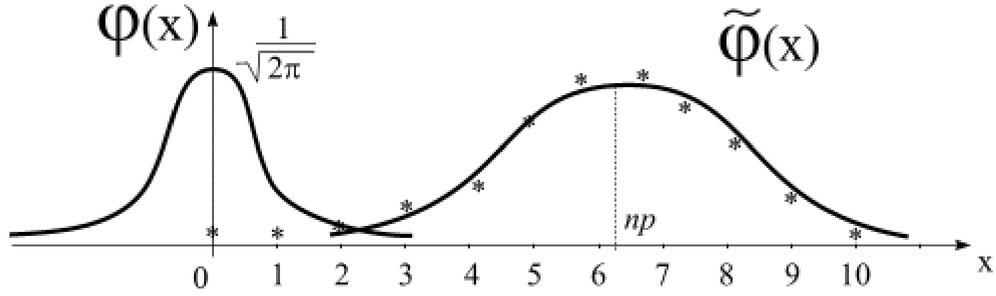


Рис. 5

Изобразим координаты (k, $P_n(k)$) звездочками. Функцию $P_n(k)$ аргумента k, можно приблизить, в соответствии с формулой (14):

$$\tilde{\varphi}(x) = \frac{1}{\sqrt{2\pi npq}} \exp\left(-(k-np)^2/(2npq)\right),$$

где np – координата центра тяжести (среднее значение), а $\sqrt{npq}$ характеризует меру «сжатости» около центра np.

Делая замену $x = (k-np)\big/\sqrt{npq}$, мы преобразуем произвольную функцию $\tilde{\varphi}(x)$ к стандартной $\varphi(x)$, у которой координата центра тяжести $np = 0$, а $\sqrt{npq} = 1$. Из рисунка видно, что при $n \to \infty$, $P_n(k) \to 0$ (при этом всегда $\sum_{k=0}^{n} P_n(k) = 1$) ошибка уменьшается. Для удобства вычислений, функция $\varphi(x)$ табулирована (см. приложение, табл. 3). Сама функция называется **кривой Гаусса**. Функция $\varphi(x)$ – четная, $\varphi(x) \to 0$ при $|x| \to \infty$, $\varphi(x) < 10^{-4}$, при $|x| > 5$, $\max\limits_{x}\{\varphi(x)\} = 1/\sqrt{2\pi} \approx 0{,}3989$.

Для практических приложений (при $n > 10$, $p \to \frac{1}{2}$) используют формулу

$$C_n^k \cdot p^k \cdot q^{n-k} \approx \frac{1}{\sqrt{npq}} \cdot \varphi\left(\frac{k-np}{\sqrt{npq}}\right). \qquad (15)$$

Пример. Решить пример п 1.5, а).

Решение. Имеем

$$P_{5000}(50) = C_{5000}^{50} \cdot (0{,}01)^{50} \cdot (0{,}99)^{4950} \approx \frac{1}{\sqrt{npq}} \cdot \varphi\left(\frac{k-np}{\sqrt{npq}}\right),$$

$k = 50$, $np = 50$, $\sqrt{npq} = \sqrt{50 \cdot 0{,}99} \approx 7{,}04$.

Итак,

$$P_{5000}(50) \approx \frac{1}{7{,}04}\varphi(0) = \left\{ \begin{matrix} \text{по табл.} \\ \text{для } \varphi(x) \end{matrix} \right\} = \frac{0{,}3989}{7{,}04} \approx 0{,}057.$$

П.3. Интегральная теорема Муавра-Лапласа

Теорема. Пусть в n независимых испытаниях вероятность появления события A в каждом испытании постоянна и равна p, $0< p < 1$, тогда, для любых $-\infty < a < b < \infty$, равномерно относительно a, b, при $n \to \infty$, имеет место асимптотическая оценка

$$P\{k_1 \le k \le k_2\} = \sum_{k=k_1}^{k_2} C_n^k \cdot p^k \cdot q^{n-k} \xrightarrow[n\to\infty]{} \int_{\grave{a}}^{b} \varphi(x)dx, \qquad (16)$$

где φ(x) - кривая Гаусса, $a = \dfrac{k_1 - np}{\sqrt{npq}}$, $b = \dfrac{k_2 - np}{\sqrt{npq}}$.

Функция $\varPhi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$ называется функцией Лапласа.

Так как $P_n\{0 \le k \le n\} = 1$ для любого n, то из (16) должно следовать, что

$$\varPhi(+\infty) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = 1.$$

В самом деле, положим $\mathcal{J} = \int\limits_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt$, тогда

$$\mathcal{J}^2 = \left(\int\limits_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt \right)^2 = 4\int\limits_{0}^{\infty} e^{-\frac{t^2}{2}} dt \cdot \int\limits_{0}^{\infty} e^{-\frac{x^2}{2}} dx = 4\int\limits_{0}^{\infty}\int\limits_{0}^{\infty} e^{-\frac{t^2+x^2}{2}} dtdx.$$

Введем полярные координаты:

$t = r\cos\varphi$, $x = r\sin\varphi$, $\varphi \in [0, \pi/2]$, $r \ge 0$, $dtdx = rdrd\varphi$.

Отсюда

$\mathcal{J}^2 = 4\int\limits_{0}^{\frac{\pi}{2}}\int\limits_{0}^{\infty} re^{-\frac{r^2}{2}} drd\varphi = 4\int\limits_{0}^{\frac{\pi}{2}} d\varphi \cdot \int\limits_{0}^{\infty} e^{-\frac{r^2}{2}} rdr = 2\pi \cdot 1 = 2\pi$, $\mathcal{J} = \sqrt{2\pi}$ - интеграл Пуассона.

Следовательно, $\varPhi(+\infty) = \dfrac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = 1$.▼

Для практических приложений вместо (16) используют формулу:

$$P\,\{k_1 \le k \le k_2\} \approx \Phi\,(в\,) - \Phi\,(a\,), \qquad (17)$$

где

$$a = \frac{k_1 - np}{\sqrt{npq}}\,,\ b = \frac{k_2 - np}{\sqrt{npq}}\,.$$

Учитывая, что $\Phi\,(+\infty) = 1$, легко получить $\Phi\,(x) + \Phi\,(-x) = 1$.

В самом деле, пусть $x > 0$, тогда $\Phi(x) = \int\limits_{-\infty}^{x} \varphi(t)dt$, а $\Phi(-x) = \int\limits_{-\infty}^{-x} \varphi(t)dt =$

$$= \left\{ \begin{matrix} \text{замена} \\ y = -x, \quad dx = -dy \end{matrix} \right\} = -\int\limits_{+\infty}^{x} = \int\limits_{x}^{\infty}\ .$$

Отсюда

$$\Phi\,(x) + \Phi\,(-x) = \int\limits_{-\infty}^{x} + \int\limits_{x}^{\infty} = 1.$$

Функция $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$ - табулирована, ее значения приведены в табл. 4 приложения.

Таблица составлена для $x < 0$, а для $x > 0$, значения находятся по формуле

$$\Phi\,(x) + \Phi\,(-x) = 1.$$

Пример. Решить пример п 1.5, б).

Решение. Имеем

$$P_{5000}\,\{0 \le k \le 50\} = \sum_{k=0}^{50} C_{5000}^{k} (0,01)^k \cdot (0,99)^{5000-k} \approx \Phi(b) - \Phi(a)\,,$$

$a = \dfrac{0-50}{7,07} \approx -7$, $b = \dfrac{50-50}{7,07} = 0$.

По табл. 5 приложения находим

$$\Phi(0) - \Phi(-7) = 0,5 - 0,0 = 0,5\,.$$

Отсюда $P_{5000}\,\{0 \le k \le 50\} \approx 0,5$.

Сравнивая решение задачи п.1.5. а), б), можно предположить, что, так как $k = 50$ – наивероятнейшее число, с большой вероятностью реализуется событие $\{\ 40 \le \text{k} \le 60\}$, с центром в точке k_0:

$$P_{5000}\{40 \le k \le 60\} \approx 0{,}037 \cdot 20 \approx 0{,}7\,.$$

Заметим, что $\sqrt{npq}$ характеризует средние отклонения от среднего значения *np* (чем меньше $\sqrt{npq}$, тем «круче» кривая Гаусса в точке симметрии).

П.4. Формула Пуассона

Приближенные формулы Муавра-Лапласа перестают быть эффективными при больших отклонениях вероятности *p* или q от 0,5 и бессмысленны при $p \to 0$, поскольку в этом случае, для разумного приближения, требуется проведение очень большого числа независимых испытаний.

Однако, во многих задачах пищевой промышленности, биологии, сельского хозяйства, в технике и электронике, возникают именно такие задачи, то есть приходится рассматривать объекты, состоящие из очень большого числа однородных элементов, каждый из которых имеет малую реализацию целевой функции (например, всхожесть зерна, выход из строя транзистора и др.).

Возникает задача оценки, например, вероятности всхожести семян, именно для таких случаев. Соответствующая оценка предложена Пуассоном.

Пусть в *n* независимых испытаниях вероятность появления события *A*, в каждом из испытаний, равна *p* (причем *p* близко к нулю), тогда имеет место оценка Пуассона:

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k} \approx \frac{\lambda^k}{k!} \cdot e^{-\lambda}, \text{ где } \lambda = np,\ k << n,$$

(где символ «<< » читается: «много меньше»).

В самом деле, при $k = 0$, имеем

$$P_n(0) = q^n = (1-p)^n = \left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda}\,.$$

Рассмотрим отношение

$$\forall\ k \ge 1\ \ \frac{P_n(k+1)}{P_n(k)} = \frac{n\cdot(n-1)\cdot\ldots\cdot(n-k)\cdot k!\cdot p^{k+1}\cdot q^{n-k-1}}{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)\cdot(k+1)!\cdot p^{k}\cdot q^{n-k}}.$$

После упрощений, получаем

$\frac{P_n(k+1)}{P_n(k)} = \frac{(n-k)p}{(k+1)q} = \frac{np}{(k+1)q} - \frac{kp}{(k+1)q} \approx \frac{\lambda}{k+1}$, так как $k << n$, $q \to 1$ и $np = \lambda$.

Таким образом, имеем

$$P_n(0) = e^{-\lambda},\ P_n(k+1) \approx P_n(k)\cdot\frac{\lambda}{k+1},\ k = 1,2,\ldots,n.$$

Окончательно, получим

$$P_n(k) \approx \frac{\lambda^k}{k!}e^{-\lambda},\ k = 0,1,\ldots,n. \blacktriangledown \tag{18}$$

Формула $V_k = \frac{\lambda^k}{k!}e^{-\lambda},\ k \in \{0\} \cup N$, называется формулой Пуассона. Очевидно, что $\sum_{k=0}^{\infty} V_k = 1$.

Значения V_k формулы Пуассона для различных k и λ представлены в приложении (табл. 2).

Пример. В книге на 1000 страниц 100 опечаток. Какова вероятность обнаружить, в наудачу взятой странице, хотя бы одну опечатку?

Решение. Имеем n = 100, p = 0,001, np = 0,1. В силу независимости выбора страниц искомая вероятность находится по формуле:

$$P_{100}\{1 \le k \le 100\} = \sum_{k=1}^{100} C_{100}^{k}\cdot(0{,}001)^k\cdot(0{,}999)^{100-k} = 1 - P_{100}(0).$$

Из формулы (18) получаем $P_{100}(0) \approx V_0 = e^{-0{,}1} \approx 0{,}9048$.

Таким образом, $P\{1 \le k \le 100\} \approx 1 - 0{,}9048 = 0{,}0952$.

Полученное значение вероятности согласуется и с интуитивным смыслом, так как в среднем одна опечатка приходится на 10 страниц.

Рассмотренные нами приближенные формулы для формулы Бернулли имеют важное самостоятельное значение. В качестве приложения оценим событие $\{\ |\mu / n - p|\ \ < \varepsilon\}$, где μ / n- частота, ε > 0.

Прежде всего, формулу (17), в интегральной теореме Муавра- Лапласа, преобразуем к виду:

$$P\{k_1 \le k \le k_2\} = P\left\{\frac{k_1 - np}{\sqrt{npq}} \le \frac{k - np}{\sqrt{npq}} \le \frac{k_2 - np}{\sqrt{npq}}\right\} = P\left\{a \le \frac{k - np}{\sqrt{npq}} \le b\right\} \approx \Phi(b) - \Phi(a).$$

Отсюда

$$P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = P\left\{-\varepsilon < \frac{\mu}{n} - p < \varepsilon\right\} = P\{-n\varepsilon < \mu - np < n\varepsilon\} =$$

$$= P\left\{-\varepsilon\sqrt{\frac{n}{pq}} < \frac{\mu - np}{\sqrt{npq}} < \varepsilon\sqrt{\frac{n}{pq}}\right\} \xrightarrow[n\to\infty]{} \left[\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right) - \Phi\left(-\varepsilon\sqrt{\frac{n}{pq}}\right)\right] = 1 - 2\Phi\left(-\varepsilon\sqrt{\frac{n}{pq}}\right).$$

Таким образом:

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = 1 - 2\Phi(-\infty) = 1. \qquad (19)$$

Асимптотическая формула (19) является одной из теорем закона больших чисел (теорема Бернулли п. 3.1); и обосновывает определение статистической вероятности (см. формулу 4, п.1.3.2.). Для практических приложений, вместо (19), обычно пользуются приближенной формулой:

$$P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} \approx 1 - 2\Phi\left(-\varepsilon\sqrt{\frac{n}{pq}}\right). \qquad (20)$$

Это трансцендентное уравнение всегда имеет решение, если неизвестное только одно.

Пример. Сколько повторных испытаний симметричной монеты нужно провести, чтобы с вероятностью не меньшей 0,98, частота появления герба отклонилась от его вероятности не более чем на 0,01.

Решение. Из (20), при ε = 0,01, p = 0,5, имеем

$$P\left\{\left|\frac{\mu}{n}-\frac{1}{2}\right|<0,01\right\}\approx 1-2\Phi\left(-0,01\sqrt{\frac{n}{\frac{1}{2}\cdot\frac{1}{2}}}\right);$$

$$1-2\cdot\Phi(-0,02\cdot\sqrt{n})=0,98\,,\Phi(-0,02\cdot\sqrt{n})=0,01.$$

По табл. 4 приложения значение аргумента находим из равенств

$$\Phi\ (x) = 0,01 \Rightarrow -0,02\cdot\sqrt{n} = -2,3$$

или

$$\sqrt{n}\approx 115\ \Rightarrow\ n\geq 13225.$$

§3. Случайные величины (4 часа)

П.1. Понятие случайной величины

Пусть дано произвольное вероятностное пространство ($\Omega,\mathcal{F},P$).

Введем одно из основных понятий теории вероятностей – **случайную величину**. Интуитивно, случайная величина – это переменная (функция), которая в результате эксперимента принимает одно из множества своих возможных значений.

Определение: Измеримая функция ξ, определенная на пространстве элементарных событий Ω и принимающая значения из области действительных чисел, называется случайной величиной

$$\xi=\xi(\omega),\ \omega\in\Omega,\ \xi\in R\,. \qquad (21)$$

Поскольку для любого элементарного события $\omega\in\Omega$ определена вероятность его реализации, то очевидно, что каждое значение случайной величины так же имеет свою вероятность. Таким образом, с каждой случайной величиной связано **распределение вероятностей** (в определении это измеримость).

Различают два основных вида случайных величин: **дискретные** и **непрерывные.**

Случайная величина, которая принимает **конечное** или **счетное** число значений, называется дискретной.

Например, к дискретным случайным величинам относятся:

а) число отказов технического устройства за определенное время;

б) количество посетителей столовой в каждый рабочий день за месяц;

в) число появлений гербов при подбрасывании монеты в серии из n испытаний.

Случайная величина называется **непрерывной**, если она принимает значения из интервала или, может быть, всей действительной оси.

Например, к непрерывным случайным величинам относятся:

а) время работы технического устройства до первого отказа;

б) отсутствие посетителей в столовой в течение не более чем один час;

в) величина ошибки измерения физических величин.

Дискретная случайная величина полностью определяется своим законом распределения, который можно представить в виде табл. 1:

Таблица 1

ξ	x_1	x_2	…	x_n	…
p	p_1	p_2	…	p_n	…

$(\Sigma p_i=1)$

где x_i, $i\in N$ возможные значения случайной величины, а p_i – соответствующие им вероятности. При этом сумма вероятностей всех значений случайной величины всегда равна единице.

П.2. Функция распределения вероятностей случайной величины

В общем случае, случайная величина полностью определяется своей функцией распределения.

Определение. **Функцией распределения** $F(x)$ случайной величины ξ называется вероятность события $\{\xi < x\},\ x \in R$, то есть

$$F(x) = P\ \{\xi < x\}, x \in R\ . \qquad (22)$$

Функция распределения существует для любой случайной величины.

Свойства функции распределения

1) Монотонность: $\forall(x_1 \le x_2) \Rightarrow (\ F\ (x_1) \le F\ (x_2)$;

2) непрерывность слева: $\lim\limits_{x \to x_0 - 0} F(x) = F(x_0)$

3) число разрывов 1-го рода не более чем счетно (ступенчатая функция);

4) $F(-\infty) = 0$;

5) $F(+\infty) = 1$.

Любая функция, удовлетворяющая свойствам 1) – 5), является функцией распределения и обратно.

Для любой дискретной случайной величины можно построить ее функцию распределения. Более того, можно сказать, что величина $\xi = \xi\ (\omega)$ называется случайной, если она имеет функцию распределения.

Определение. Индикатором события A называется случайная величина:

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

Пример 1. Построить функцию распределения индикатора события A, если известно, что $P\ \{\omega \in A\} = p$.

Решение. Случайная величина $I_A(\omega)$ – дискретная. Закон ее распределения имеет вид:

I_A	0	1
p	1-p	p

.

Функцию распределения определим формулой:

$$F(x) = \begin{cases} 0, & x \le 0; \\ 1 - p, & 0 < x \le 1; \\ 1, & x > 1, \end{cases}$$

график которой представлен на рис. 6.

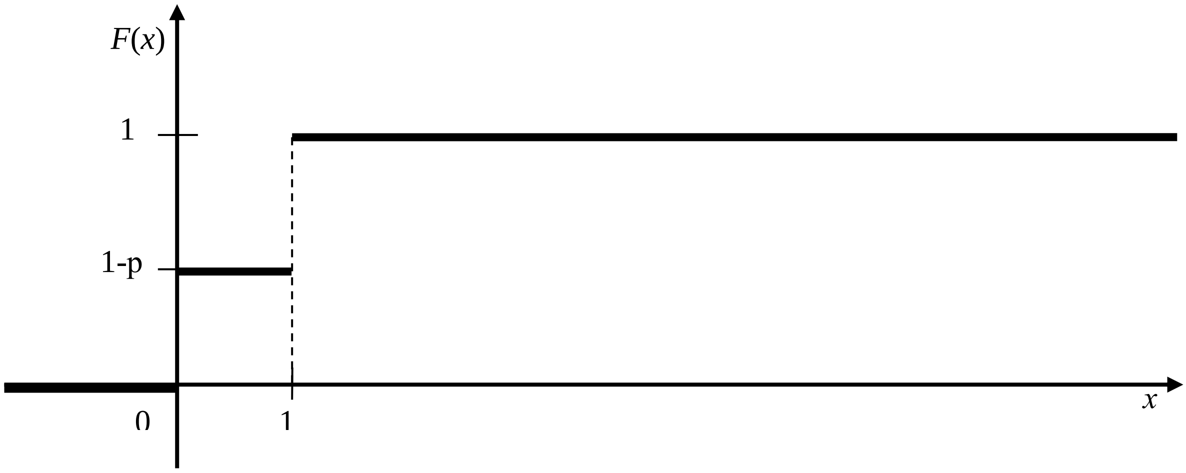


Рис. 6

П.3. Плотность вероятности случайной величины

Очевидно, что вероятностная характеристика случайной величины, полученная из функции распределения с помощью формальных математических операций и определенная на всей действительной оси, несет столько же информации о случайной величине, что и сама функция распределения. Исключение могут составлять некоторые точки действительной оси, число которых не более чем счетно.

К такой характеристике можно отнести плотность.

Определение. Функция $\rho(x)$, удовлетворяющая условиям:

а) $\rho(x) \geq 0, \; \forall\, x \in R$,

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

называется **плотностью** случайной величины ξ.

Из определения видно, что плотность играет ту же роль, что и закон распределения дискретной случайной величины.

Физически, плотность характеризует распределение единичной массы на действительной оси. Ёе изменение на участке длиной Δx, примыкающего к точке x, оценивается интегралом

$$\int\limits_{x}^{x+\Delta x} \rho(t)dt.$$

Свойства плотности

Свойство 1. $$F(x) = \int\limits_{-\infty}^{x} \rho(t)\,dt\,. \qquad (23)$$

Доказательство. Проверим свойства 1)-5) функции распределения.

Пусть $\rho\,(x)$ - плотность, тогда $\rho\,(x) \geq 0$, отсюда

1) $\forall(x_1 \geq x_2) \Rightarrow \int\limits_{-\infty}^{x_1} \rho(x)dx \geq \int\limits_{-\infty}^{x_2} \rho(x)dx;$

2) непрерывность слева следует либо из непрерывности $\rho(x)$, либо из ее кусочной непрерывности с разрывами первого рода;

3) следует из существования интеграла на действительной оси;

4) $F(-\infty) = \int\limits_{-\infty}^{-\infty} \rho(x)dx = 0;$

5) $F(+\infty) = 1$, из определения.

Свойство 2. $$F'(x) = \rho(x)\,. \qquad (24)$$

Для доказательства достаточно продифференцировать (23) по переменному верхнему пределу.▼

Учитывая (23) и (24), функцию распределения называют **интегральной**, а плотность **дифференциальной** характеристикой случайной величины.

Плотность имеет смысл для такой случайной величины, функция распределения которой дифференцируема; обычно, это непрерывная случайная величина. Функция распределения - это вероятность, и по определению безразмерна. Для плотности, как следует из формулы (24), размерность обратна размерности случайной величины. Физически, плотность характеризует мгновенное изменение случайной величины в точке x.

Для дискретной случайной величины понятие плотности лишено смысла, поскольку, как видно из примера для индикатора, она либо равно нулю, либо имеет бесконечное изменение в точке разрыва.

Отметим некоторые полезные свойства функции распределения и плотности:

1) $P\{a \le \xi < b\} = F(b) - F(a)$,

2) $P\{a \le \xi < b\} = \int\limits_a^b f(x)dx$,

3) $P\ \{x \le \xi \le x{+}dx\} = \rho\,(x)dx$,

4) $P\ \{\ \xi = a\} = F\ (a{+}0) - F(a{-}0)$,

5) $P\ \{\ \xi \le a\ \} = F\ (a\ {+}0\ )$.

Упражнение. Доказать свойства 1 – 5.

П.4. Основные распределения случайных величин

Пример 1. Пусть случайная величина ξ есть число появлений события A в n независимых испытаниях (вероятность появления события А в любом испытании равна p). Построить функцию распределения.

Решение. Рассмотрим событие $\{\xi < x\} \sim \left\{\begin{matrix} \text{число появлений события } A \\ \text{меньших} \quad x \end{matrix}\right\}$, $x \in R$.

По условию, если $P\{\xi = k\} = C_n^k \cdot p^k \cdot q^{n-k}$, то полагаем, $F(x)=0$ для $x \le 0$. $F(x) = P\{\xi < x\} = \sum\limits_{k<x} P\{\xi = k\}$, для $0 < x \le n$, и $F(x) = 1$, для $x > n$. Таким образом,

$$F(x) = \begin{cases} 0, & x \le 0; \\ \sum\limits_{\hat{e}<\tilde{o}} C_n^k \cdot p^k \cdot q^{n-k}, & 0 < x \le n; \\ 1, & x > n. \end{cases} \qquad (25)$$

График функции имеет ступенчатый вид (рис.7):

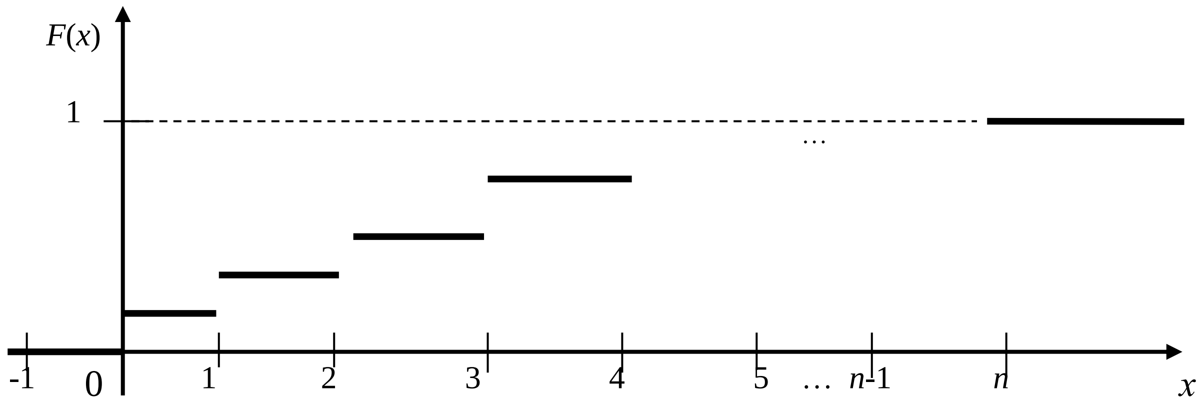


Рис. 7

Из графика видно, что свойства 1) – 5) выполняются. Величину скачка функции в точке $x = k$ находим из равенства

$$P\{\xi = k\} = F(k+0) - F(k-0) = C_n^k \cdot p^k \cdot q^{n-k}.$$

Пример 2. Будем говорить, что случайная величина ξ имеет распределение Пуассона, если ее функция распределения имеет вид:

$$F(x) = \begin{cases} 0, & x \le 0; \\ \sum\limits_{k<x} \dfrac{\lambda^k}{k!} \cdot e^{-\lambda}, & x > 0. \end{cases} \qquad (26)$$

Свойства 1)- 4) очевидны. Проверим 5):

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

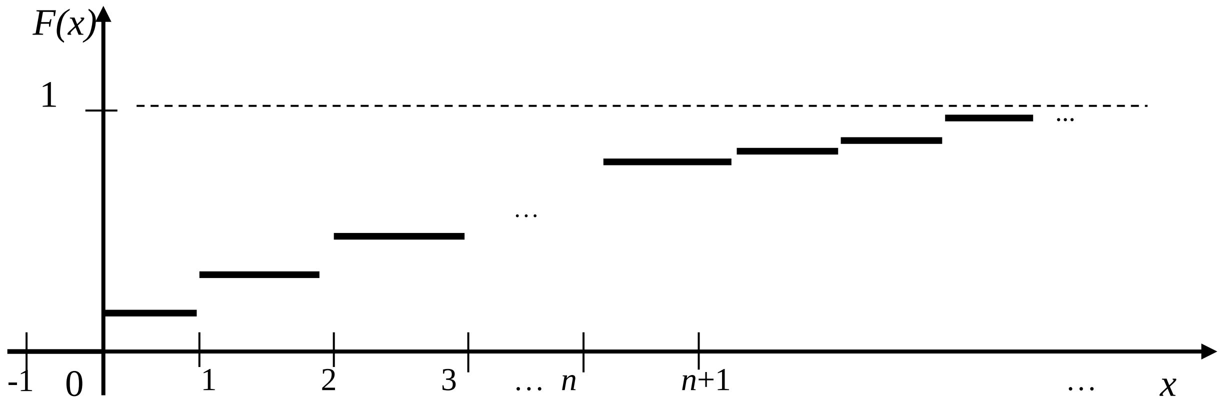


Рис. 8

Величина скачка в точке $x = k$ равна $V_k = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$, $k = 0,1,2,\ldots$. Число разрывов счетно. График функции представлен на рис. 8.

Пример 3. Будем говорить, что случайная величина ξ равномерно распределена на (a, $в$], если ее функция распределения имеет вид:

$$F(x) = \begin{cases} 0, & x \le a; \\ \frac{x-a}{b-a}, & a < x \le b; \\ 1, & x > b. \end{cases} \tag{27}$$

Плотность равномерного распределения

$$\rho(x) = F'(x) = \begin{cases} 0, & x \le a; \\ \frac{1}{b-a}, & a < x \le b; \\ 0, & x > b. \end{cases}$$

(28)

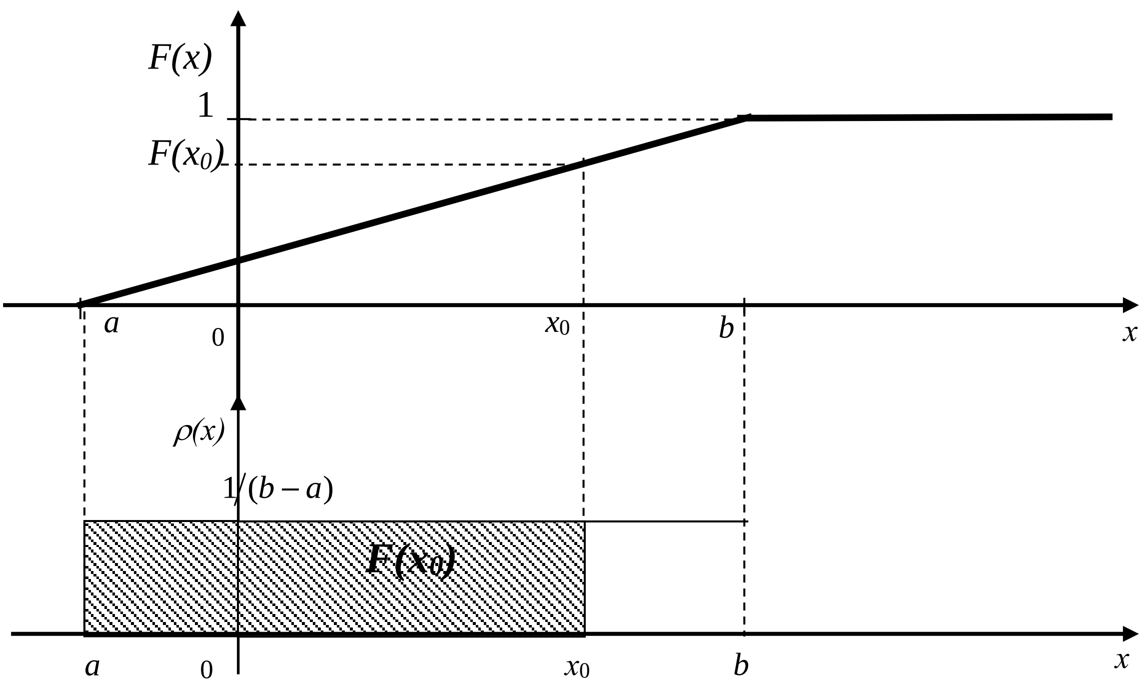


Рис. 9

Из графиков (рис.9) видно, что значение $F(x_0) = \frac{x_0 - a}{b-a}$ есть площадь (интеграл) области, ограниченной справа прямой $x = x_0$.

Пример 4. Случайная величина ξ распределена нормально, если ее функция распределения имеет вид:

$$F(x)=\frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\int\limits_{-\infty}^{x}\exp\left(-\left(y-a\right)^2/2\sigma^2\right)\,dy,$$

а плотность

$$\rho\left(x\right)=\frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\exp\left(-(x-a)^2/2\sigma^2\right),\ \sigma>0,\ a-const.$$

Свойства функции распределения 1 - 4 очевидны. Проверим свойство 5.

$$F(+\infty)=\frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\int\limits_{-\infty}^{\infty}e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\left\{\begin{matrix}\dfrac{x-a}{\sigma}=y\\ dx=\sigma dy\end{matrix}\right\}=\frac{1}{\sqrt{2\pi}}\cdot\int\limits_{-\infty}^{\infty}e^{-\frac{y^2}{2}}dy=$$

={интеграл Пуассона $\mathcal{I}=\sqrt{2\pi}$ }$=\dfrac{1}{\sqrt{2\pi}}\cdot\sqrt{2\pi}=1$.

Схематично график плотности (рис. 10) имеет вид:

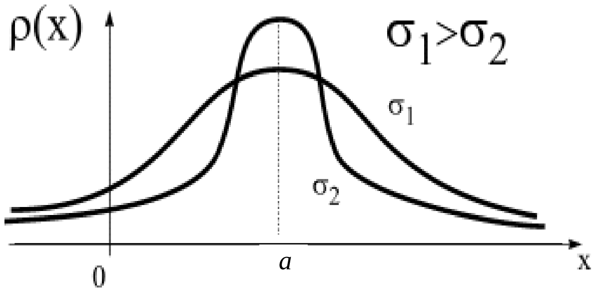


Рис. 10

Постоянная a характеризует сдвиг функции $\rho(x)$ по оси OX относительно начала координат, а σ - меру «сжатости» кривой около центра в точке $x = a$.

Пример 5. Случайная величина ξ имеет показательное (экспоненциальное) распределение, если ее функция распределения определяется по формулой

$$F(x)=\begin{cases}0,\ x\le 0,\\ 1\text{-}e^{-x\lambda},\ x>0,\ (\lambda>0\text{ - const}).\end{cases}$$

(29)

Если x – интерпретировать как время, то функция распределения будет иметь вид (рис. 11):

$$F(t) = 1 - e^{-\lambda t}, t \geq 0.$$

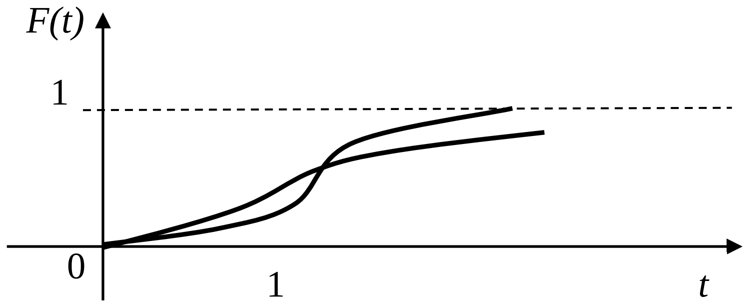


Рис. 11

Это распределение играет важную роль в технике и носит название функции надежности, λ - интенсивность с размерностью обратной времени [1].

Плотность $\rho(t) = \lambda \cdot e^{-\lambda t}$, ее график функции имеет вид (рис. 12):

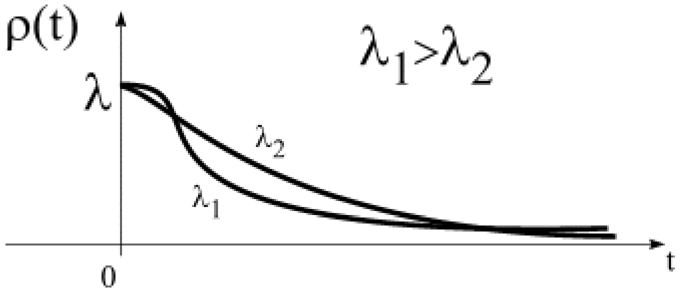


Рис. 12

П.5. Числовые характеристики случайных величин

Случайная величина полностью определяется своей функцией распределения (или плотностью, если она существует). Однако, чтобы эту функцию найти, требуется иметь не только большой объем статистических данных, но и быть уверенным в том, что они отражают все существенные свойства случайной величины. К сожалению, это бывает редко, а во многих случаях в этом нет необходимости. Достаточно бывает проанализировать часть свойств случайной величины.

Рассмотрим некоторые типичные плотности, и определим по ним числовые характеристики, знание которых поможет получить информацию о случайной величине, без знания вида самой плотности.

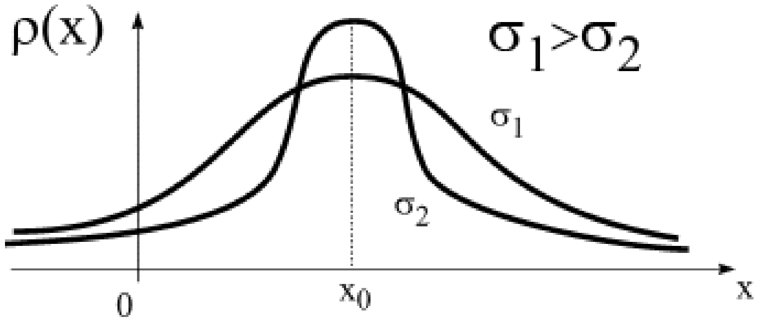


Рис. 13

Из графиков плотностей (рис. 13, 14) видно, что желательно знать абсциссу центра тяжести x_0, сгруппированность большей части площади около центра σ_1, σ_2, асимметричность μ_3, крутость μ_4, число максимумов x_1, x_2, вероятность максимального значения плотности x_0 (рис.13), x_1, и другие.

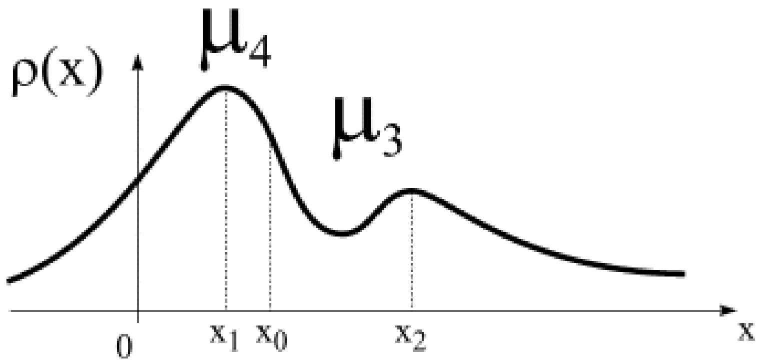


Рис. 14

Знание хотя бы части этих характеристик позволяет достичь желаемой цели без знания плотности. Наконец, при исследовании какой-либо проблемы, мы начинаем ее изучение с общих позиций, оцениваем ее в среднем. Именно для изучения этих сторон, в первую очередь, и предназначены числовые характеристики случайных величин. Мы рассмотрим здесь лишь некоторые из них.

Математическое ожидание, мода, медиана

Пусть имеем произвольное вероятностное пространство $(\Omega,\mathcal{F},P)$, на котором определена случайная величина ξ.

Определение. **Математическим ожиданием** или **средним значением** случайной величины ξ называется число $M\xi$, которое находится по формуле:

а) если случайная величина ξ дискретна, то есть задана табл. 2,

Таблица 2

ξ	x_1	x_2	…	x_n	…
p	p_1	p_2	…	p_n	…

$(\Sigma p_i=1)$, $i\in N$,

то

$$M\xi = \sum_i x_i p_i \,, \tag{30}$$

и существует, при условии, что ряд в правой части (30) сходится;

б) если случайная величина ξ непрерывна с плотностью $\rho(x)$, то

$$M\xi = \int_{-\infty}^{\infty} x \cdot \rho(x)dx, \tag{31}$$

и существует, при условии, что несобственный интеграл в правой части (31) сходится.

Математическое ожидание аналогично понятию средневзвешенного и интерпретируется как абсцисса центра тяжести распределения массы на прямой.

Свойства.

1). Если $\xi = a$ – const, то $Ma = a$.

В самом деле, рассматривая a как дискретную случайную величину с законом распределения

$$P\{\xi = a\} = 1,\; P\{\xi \neq a\} = 0,$$

получаем по формуле (30):

$$M\xi = 0\cdot P\{\xi \neq a\} + a\cdot P\{\xi = a\} = a\cdot 1 = a. \blacktriangledown$$

2). Постоянную можно выносить за знак математического ожидания

$$M(a\xi) = aM\xi.$$

В самом деле, если ξ, например, непрерывная случайная величина, то

$$M(a\xi) = \int\limits_{-\infty}^{\infty} a \cdot x \cdot \rho(x)dx = a \cdot \int\limits_{-\infty}^{\infty} x \cdot \rho(x)dx = a \cdot M\xi \text{ .}\blacktriangledown$$

3). Для любых случайных величин ξ, η

$$M\,(\xi+\eta) = M\xi + M\eta$$

4). Если случайные величины ξ, η независимы, то

$$M\,(\xi\cdot\eta) = M\xi\cdot M\eta.$$

В самом деле, если случайные величины независимы, то их совместная плотность, равна произведению плотностей случайных величин*, то есть

$$\rho_{\xi\cdot\eta}(x,y) = \rho_{\xi}(x)\cdot\rho_{\eta}(y),$$

тогда

$$M(\xi\cdot\eta) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} x\cdot y\cdot\rho(x,y)dxdy = \int\limits_{-\infty}^{\infty} x\cdot\rho(x)\cdot y\rho(y)dxdy = \int\limits_{-\infty}^{\infty} x\rho(x)dx\cdot\int\limits_{-\infty}^{\infty} y\rho(y)dy =$$

$= M\xi\cdot M\eta.\blacktriangledown$

5). Всегда $|M\xi| \le M|\xi|$.

В самом деле, имеем

$$|M\xi| = \left|\int\limits_{-\infty}^{\infty} x\cdot\rho(x)dx\right| \le \int\limits_{-\infty}^{\infty}|x|\rho(x)dx = M|\xi| \text{ .}\blacktriangledown$$

Пример. Найти математическое ожидание индикатора события $A \subset \Omega$.

Решение. По определению $I(\omega)=I_A(\omega) = \begin{cases}1, \text{если } \omega\in A,\\ 0, \text{если } \omega\notin A.\end{cases}$, тогда для любого $\omega \in \Omega$

$$M\,(I(\omega)) = P(\omega)\cdot 1 - (1 - P\,(\omega))\cdot 0 = P(\omega).$$

Так как $A = \bigcup\omega$, то $M(I(A)) = M(\sum(I(\omega)) = \sum M(I(\omega)) = \sum P(\omega) = P\{A\}$.

Таким образом, вероятность события $A \subset \Omega$ можно записать через математическое ожидание индикатора события A.

* О независимости случайных величин смотри ниже

Математическое ожидание случайной величины ξ является важнейшей, среди ее «линейных характеристик». На практике, в качестве характеристик, дополняющих математическое ожидание, используют **моду** и **медиану**.

Определение. **Модой** M_0 дискретной случайной величины ξ называется ее наивероятнейшее значение κ_0.

Модой непрерывной случайной величины ξ называется любое из значений x, в котором плотность имеет максимум.

Графическая интерпретация моды приведена на рис. 14.

Определение. **Медианой** непрерывной случайной величины ξ называется ее значение M_e, для которого

$$P\{\xi < M_e\}=P\{\xi > M_e\}.$$

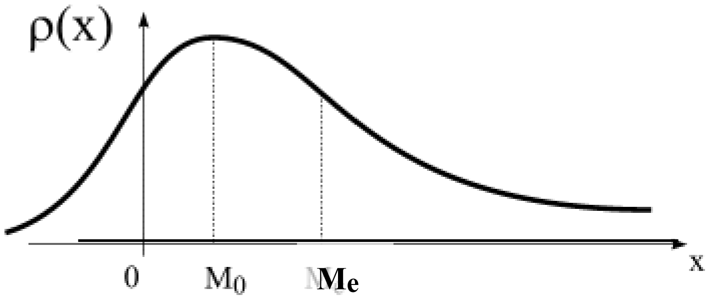


Рис. 15

На рис. 15 изображена плотность вероятности, где медиана есть абсцисса $M_e = x$, для которой

$$\int_{-\infty}^{x} \rho(t)dt = \int_{x}^{+\infty} \rho(t)dt.$$

Можно определить медиану и для дискретной случайной величины, например, как среднее арифметическое наименьшего и наибольшего ее значений [1], однако обычно медиана используется при изучении непрерывных случайных величин.

Моменты

Пусть случайная величина ξ имеет математическое ожидание $M\xi = a$. Введем новую случайную величину $\eta = \xi - a$. Случайная величина η называется **отклонением** случайной величины ξ.

Математическое ожидание отклонения равно 0.

В самом деле, имеем

$$M\eta = M(\xi - a) = M\xi - Ma = a - a = 0\ .\blacktriangledown$$

Геометрически это означает, что среднее значение отклонения всегда находится в начале координат.

Определение. **Начальным** моментом $\nu_к$ порядка $к$ случайной величины ξ называется математическое ожидание случайной величины $\xi^к$: $\nu_k = M(\xi^k)$, $к \in N$, и вычисляется по формуле:

а) $\nu_k = \sum_i (x_i)^k \cdot p_i$, $i \in N$, если ξ - дискретная; (32)

б) $\nu_k = \int_{-\infty}^{\infty} x^k \cdot \rho(x)dx$, если ξ - непрерывная. (33)

Начальные моменты порядка $к$ существуют, если их правые части в (32) и (33) имеют смысл.

Математическое ожидание есть начальный момент первого порядка $M\xi = \nu_1$.

Определение. **Центральным** моментом $\mu_к$ порядка $к$, случайной величины ξ, называется математическое ожидание $к$-ой степени отклонения

$\mu_к = M(\xi - M\xi)^к$, $к \in N$.

и вычисляется по формуле:

а) $\mu_k = \sum_i (x_i - M\xi)^k \cdot p_i$, $i \in N$, если ξ - дискретная; (34)

б) $\mu_k = \int_{-\infty}^{\infty} (x - M\xi)^k \cdot \rho(x)dx$, если ξ - непрерывная. (35)

Очевидно, что если существует момент порядка $к$, то существуют все моменты низшего порядка.

Определение. **Дисперсией** случайной величины ξ называется центральный момент второго порядка. Дисперсия обозначается символом $D\xi$:

$$D\xi = M(\xi - M\xi)^2 . \qquad (36)$$

Дисперсия число неотрицательное, и характеризует средние отклонения случайной величины от ее среднего значения.

Свойства дисперсии

Пусть ξ, η - случайные величины и $\alpha, \beta \in R$, тогда

1) дисперсия постоянной равна 0, то есть $D\alpha = 0$.

В самом деле, при $\xi = \alpha$

$$D\alpha = M(\alpha - M\alpha)^2 = M(\alpha - \alpha)^2 = M0 = 0. \blacktriangledown$$

2) для любой случайной величины ξ и $\alpha \in R$

$$D(\alpha\xi) = \alpha^2 \cdot D\xi.$$

В самом деле,

$D(\alpha\xi) = M(\alpha\xi - M(\alpha\xi))^2 = M(\alpha(\xi - M\xi))^2 = M(\alpha^2 \cdot (\xi - M\xi)^2) = \alpha^2 \cdot M(\xi - M\xi)^2 = =\alpha^2 \cdot D\xi . \blacktriangledown$

3) если ξ и η независимы, то

$D(\xi \pm \eta) = D\xi + D\eta.$

В самом деле, имеем:

$D(\xi \pm \eta) = M(\xi \pm \eta - M(\xi \pm \eta))^2 = M((\xi - M\xi) \pm (\eta - M\eta))^2 = M(\xi - M\xi)^2 +$
$+M(\eta - M\eta)^2 \pm 2M((\xi - M\xi) \cdot (\eta - M\eta)) = D\xi + D\eta \pm M(\xi - M\xi) \cdot M(\eta - M\eta) =$
$=D\xi + D\eta,$

так как из независимости ξ и η следует независимость их отклонений. $\blacktriangledown$

Часто вместо формулы (36) используют эквивалентную ей формулу

$$D\xi = M(\xi^2) - (M\xi)^2. \qquad (37)$$

В самом деле,

$D\xi = M(\xi - M\xi)^2 = M(\xi^2 - 2\xi \cdot M\xi + (M\xi)^2) = M\xi^2 - 2M\xi \cdot M\xi + (M\xi)^2 = M\xi^2 - (M\xi)^2. \blacktriangledown$

Для практических приложений более удобной характеристикой случайной величины является **среднеквадратичное** (**стандартное**) отклонение σ, вычисляемое по формуле:

$$\sigma\xi = \sqrt{D\xi} . \tag{38}$$

Среднеквадратичное отклонение имеет ту же размерность, что и сама случайная величина, а дисперсия - квадратичную. Основной недостаток стандартного отклонения, в отличие от дисперсии, в том, что оно не обладает свойством аддитивности. Это означает, что, если для независимых случайных величин ξ и η

$D(\xi + \eta) = D\xi + D\eta,$

то для стандартного отклонения

$$\sigma(\xi+\eta) = \sqrt{D(\xi+\eta)} = \sqrt{D\xi + D\eta} .$$

Математическое ожидание и дисперсия наиболее популярные числовые характеристики случайных величин, поскольку они отражают наиболее важные свойства распределения. Для детального изучения случайных величин применяются моменты высших порядков. Мы рассмотрим здесь коэффициент **асимметрии** и **эксцесс**.

Определение. **Асимметрией** распределения называется свойство кривой распределения, указывающее на отличие от симметричности распределения случайной величины.

Мерой асимметрии распределения является **коэффициент** асимметрии $S_к$, определяемый равенством

$$S_k = \frac{\mu_3}{\sigma^3} ,$$

где μ_3- третий центральный момент распределения вероятностей случайной величины ξ.

Асимметрия положительна, если $S_к > 0$, отрицательна, если $S_к < 0$ и равна нулю, если распределение симметрично.

При положительной асимметрии более «длинная» часть плотности распределения лежит правее моды и отрицательна, если левее моды.

Замечание. Для распределений симметричных относительно математического ожидания все моменты нечетного порядка, если они существуют, равны нулю.

В самом деле, например, если случайная величина имеет плотность $\rho(x)$, то

$$\mu_{2\kappa+1} = \int_{-a}^{+a} (x - \mu_{\xi})^{2\kappa+1} \rho(x)dx = 0,\ \kappa = 0,\ 1,\ ...,$$

что сразу следует из свойств интеграла от нечетной функции с симметричными пределами.

Таким образом, любой центральный момент нечетного порядка может быть использован для характеристики асимметрии.

Определение. **Коэффициентом эксцесса** (эксцессом) распределения вероятности случайной величины ξ называется числовая характеристика E_k, определяющая «островершинность» плотности распределения, и вычисляется по формуле:

$$E_k = \frac{\mu_4}{\sigma^4} - 3,$$

где μ_4 – четвертый центральный момент вероятностного распределения. Число 3 связано с эксцессом нормального распределения, так как для него $\mu_4 / \sigma^4 = 3$. В силу исключительной важности нормального распределения в теории вероятностей с ним сравниваются распределения вероятностей отличных от нормального. Таким образом, для нормального распределения $E_k = 0$. Если вершина распределения более «остра» чем нормальное, то эксцесс положителен, если более «плоска», то эксцесс отрицателен. Геометрическая интерпретация этого факта представлена на рис. 16.

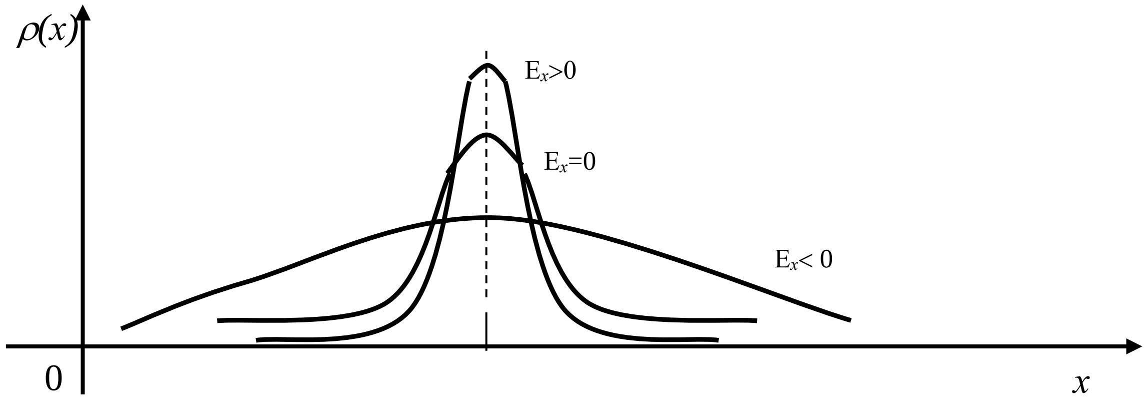


Рис. 16

Рассмотренные числовые характеристики случайных величин являются наиболее употребительными на практике. Достаточно часто ими пользуются для приближенной замены одного распределения другим, более подходящим.

П.6. Вычисление числовых характеристик стандартных распределений

Вычислим математическое ожидание и дисперсию основных распределений.

1. **Биномиальное распределение.**

Имеем $P\{\xi = k\} = C_n^k \cdot p^k \cdot q^{n-k}$, $q = 1 - p$, $k = 0, 1, ..., n$ тогда

$$M\xi = \sum_k x_k p_k = \sum_{k=1}^{n} k \cdot C_n^k \cdot p^k \cdot q^{n-k} = \sum_{k=1}^{n} \frac{k \cdot n \cdot (n-1) \cdot ... \cdot (n-k+1)}{k!} \cdot p^k \cdot q^{n-k} =$$

$$= \{m = k-1\} = np \cdot \sum_{m=0}^{n-1} \frac{(n-1) \cdot ... \cdot (n-1-m+1)}{m!} \cdot p^m \cdot q^{n-1-m} = np \cdot \sum_{m=0}^{n-1} C_{n-1}^m \cdot p^m \cdot q^{n-1+m} =$$

$$= np \cdot 1 = np.$$

Для вычисления дисперсии воспользуемся следующим приемом [5]. Введем независимые случайные величины ξ_i:

$$\forall_i,\ \xi_i = \begin{cases} 1, & \text{если событие } A \text{ появилось в испытании } i, \\ 0, & \text{в противном случае, } i = 1, 2, ..., n. \end{cases}$$

Тогда

$$\xi=\sum_{i=1}^{n}\xi_i,\ \forall\ i \Rightarrow\ D\xi_i=\sum_i x_i^2 p_i-(M\xi_i)^2=1^2\cdot p+0\cdot q^2-(1\cdot p+0\cdot q)^2=$$

$$=p-p^2=p\cdot q.$$

В силу независимости случайных величин ξ_i:

$$D\xi=D\left(\sum_{i=1}^{n}\xi_i\right)=\sum_{i=1}^{n}D\xi_i=\sum_{i=1}^{n}(p\cdot q)=n\cdot p\cdot q.$$

Итак, для биноминального распределения

$M\xi=np,\qquad D\xi=npq.$

2. Распределение Пуассона.

Имеем $P\{\xi=k\}=\dfrac{\lambda^k}{k!}\cdot e^{-\lambda}$, κ=0, 1, 2, … (см стр. 50), тогда

$$M\xi=\sum_{\kappa=1}^{\infty}x_k\cdot p_k=\sum_{k=1}^{\infty}k\cdot V_k=\sum_{k=1}^{\infty}k\cdot\frac{\lambda^k}{k!}\cdot e^{-\lambda}=\lambda\cdot e^{-\lambda}\cdot\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}=\{k-1=i\}=$$

$$=\lambda\cdot e^{-\lambda}\cdot\sum_{i=0}^{\infty}\frac{\lambda^i}{i!}=\lambda\cdot e^{-\lambda}\cdot e^{\lambda}=\lambda.$$

Далее,

$$D\xi=M(\xi^2)-\left(M\xi\right)^2=\sum_{\hat{e}=1}^{\infty}k^2\cdot\frac{\lambda^k}{k!}\cdot e^{-\lambda}-\lambda^2=\lambda\cdot e^{-\lambda}\cdot\sum_{\hat{e}=1}^{\infty}\kappa\cdot\frac{\lambda^{\kappa-1}}{(\kappa-1)\ !}-\lambda^2=$$

$$=\lambda\cdot e^{-\lambda}\cdot\sum_{k=1}^{\infty}\frac{\lambda^{k-1}\cdot(k-1+i)}{(k-1)\ !}-\lambda^2=\lambda\cdot e^{-\lambda}\cdot\left(\sum_{k=1}^{\infty}\frac{(k-1)\cdot\lambda^{k-1}}{(k-1)\ !}+\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)\ !}\right)-\lambda^2=$$

$$=\{k-1=i\}=\lambda\cdot e^{-\lambda}\cdot\left(\sum_{i=0}^{\infty}\frac{i\cdot\lambda^i}{i!}+\sum_{i=0}^{\infty}\frac{\lambda^i}{i!}\right)-\lambda^2=\lambda\cdot e^{-\lambda}\cdot\left(\lambda\cdot e^{\lambda}+e^{\lambda}\right)-\lambda^2=\lambda^2+\lambda-$$

$$-\lambda^2=\lambda.$$

Здесь мы воспользовались рядом Маклорена $e^{\lambda}=\sum_{k=0}^{\infty}\dfrac{\lambda^k}{k!}$. Таким образом, для распределения Пуассона

$M\xi=\lambda,\ D\xi=\lambda.$

3. Равномерное распределение.

Используя формулы (28) и (31), имеем

$$M\xi=\int\limits_{-\infty}^{\infty}x\cdot\rho(x)dx=\int\limits_{-\infty}^{a}x\cdot 0dx+\int\limits_{a}^{b}x\cdot\frac{1}{b-a}dx+\int\limits_{b}^{\infty}x\cdot 0dx=\int\limits_{a}^{b}\frac{x}{b-a}dx=\frac{x^2}{2(b-a)}\Bigg|_{a}^{b}=$$

$$=\frac{b+a}{2}.$$

$$D\xi=\int\limits_{-\infty}^{\infty}(x-M\xi)^2\cdot\rho(x)dx=\int\limits_{a}^{b}\left(x-\frac{b+a}{2}\right)^2\cdot\frac{1}{b-a}dx=\frac{(b-a)^2}{12}.$$

Таким образом, для равномерного распределения

$$M\xi=\frac{b+a}{2},\ D\xi=\frac{(b-a)^2}{12}.$$

4. Экспоненциальное распределение.

Учитывая, что $\rho(x)=\lambda\cdot e^{-\lambda x}$, $x\geq 0$, получаем

$$M\xi=\lambda^{-1},\ D\xi=\lambda^{-2}.$$

5. Нормальное распределение.

$$M\xi=\frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\int\limits_{-\infty}^{\infty}x\cdot\exp\left(-(x-a)^2/2\sigma^2\right)dx=\left\{\begin{matrix}\dfrac{x-a}{\sigma}=t\\ dx=\sigma\cdot dt\end{matrix}\right\}=\frac{1}{\sqrt{2\pi}}\cdot\int\limits_{-\infty}^{\infty}(A+\sigma t)\cdot e^{-t^2/2}dt=$$

$$=\frac{a}{\sqrt{2\pi}}\cdot\int\limits_{-\infty}^{\infty}e^{-t^2/2}dt-\frac{\sigma}{\sqrt{2\pi}}\cdot\int\limits_{-\infty}^{\infty}e^{-t^2/2}d\left(-\frac{t^2}{2}\right)=\frac{a}{\sqrt{2\pi}}\cdot\sqrt{2\pi}+0=a,$$

так как $\int\limits_{-\infty}^{\infty}e^{-t^2/2}dt=\sqrt{2\pi}$.

$$D\xi=\frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\int\limits_{-\infty}^{\infty}(x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\left\{\begin{matrix}\dfrac{x-a}{\sigma}=t\\ dx=\sigma\cdot dt\end{matrix}\right\}=\frac{\sigma^2}{\sqrt{2\pi}}\cdot\int\limits_{-\infty}^{\infty}t^2\cdot e^{-t^2/2}dt=$$

$$=-\frac{\sigma^2}{\sqrt{2\pi}}\cdot t\cdot e^{-t^2/2}\Bigg|_{-\infty}^{\infty}+\frac{\sigma^2}{\sqrt{2\pi}}\cdot\int\limits_{-\infty}^{\infty}\exp\left(-\frac{t^2}{2}\right)dt=0+\sigma^2=\sigma^2.$$

Итак, для нормального распределения

$$M\xi=a,\ D\xi=\sigma^2.$$

Вычислим эксцесс E_k.

$$\mu_4 = \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\int_{-\infty}^{\infty}(x-a)^4\cdot\exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)dx = \left\{\begin{matrix}\dfrac{x-a}{\sigma}=t\\ dx=\sigma\cdot dt\end{matrix}\right\} = \frac{\sigma^4}{\sqrt{2\pi}}\cdot\int_{-\infty}^{\infty}t^4\cdot e^{-t^2/2}dt =$$

$$= \frac{\sigma^4}{\sqrt{2\pi}}\cdot\int_{-\infty}^{\infty}t^4\cdot e^{-t^2/2}dt = \frac{\sigma^4}{\sqrt{2\pi}}\cdot\int_{-\infty}^{\infty}t^3\cdot d(e^{-t^2/2}) = -\frac{\sigma^4}{\sqrt{2\pi}}\cdot t^3\cdot e^{-t^2/2}\Bigg|_{-\infty}^{\infty} +$$

$$+\frac{3\sigma^4}{\sqrt{2\pi}}\cdot\int_{-\infty}^{\infty}t^2\cdot e^{-t^2/2}dt = \text{\{учитывая вычисление дисперсии\}}$$

$$= \frac{3\sigma^4}{\sqrt{2\pi}}\cdot\sqrt{2\pi} = 3\sigma^4.$$

Отсюда $E_k = \dfrac{\mu_4}{\sigma^4} - 3 = \dfrac{3\sigma^4}{\sigma^4} - 3 = 0$.

Итак, для нормального распределения

$$M\xi = a,\ D\xi = \sigma^2,\ \ \mu_4 = 3,\ E_k = 0.$$

П.7. Приложения нормального распределения

Пусть случайная величина ξ имеет нормальное распределение. Вычислим вероятность события $\{\alpha < \xi < \beta\}$.

Из свойств функции распределения имеем:

$$P\{\alpha \le \xi < \beta\} = F(\beta) - F(\alpha) = \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\int_{-\infty}^{\beta}\exp(-(y-a)^2/2\sigma^2)\,dy -$$

$$-\frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\int_{-\infty}^{\alpha}\exp\left(-(y-a)^2/2\sigma^2\right)dy.$$

Введя новую переменную $x = (y-a)/\sigma$, получаем

$$P\{\alpha \le \xi < \beta\} = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right), \text{ где } \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-t^2/2}dt.$$

В частности

$$P\{\ |\xi - a|\ < \delta\} = 1 - 2\Phi\left(-\frac{\delta}{\sigma}\right), \tag{39}$$

(или $P\{\ |\xi - a|\ < \delta\} = 2\Phi\left(\dfrac{\delta}{\sigma}\right)$, если $\Phi(x) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{0}^{x}e^{-t^2/2}dt$).

Так как на вероятность события не влияют неслучайные преобразования над случайной величиной, то, полагая $a = 0$, получим:

$$P\left\{\ |\xi|\ < \delta\right\} = 1 - 2\Phi\left(-\frac{\delta}{\sigma}\right). \qquad (40)$$

Пример (правило трех σ). Пусть случайная величина ξ, характеризующая ошибки измерений, подчинена нормальному закону. Найти вероятность того, что ошибки измерения ξ не превзойдут 3σ.

Решение. Ошибкой измерения называется отклонение реального результата от истиного. Ошибка измерения является случайной величиной, если она есть результат действия только случайных факторов (в отличие от систематических ошибок, которые изменяются по определенному закону или постоянны во всей серии испытаний).

Если ошибки измерения случайны, то они симметричны относительно нуля, поскольку в силу их случайности, разумно предположить, что ошибки, равные по величине и противоположные по знаку, должны встречаться одинаково часто.

С учетом сказанного, воспользуемся формулой (40). Имеем

$P\{|\ \xi\ | < 3\sigma\}$= 1 - 2Φ(-3) = 1 - 0,00135 = 0,9973.

Таблица 3

$n\sigma$	σ	2σ	3σ	4σ
$P\{\mid \xi \mid < n\sigma\}$	0,6826	0,9544	0,9973	0,9997
%	32%	5%	0,7%	0,07%

Важность «правила трех σ» в том, что ошибки измерения для нормально распределенных величин, превышающих 3σ практически невозможны, менее 3σ существенны (см. табл. 3).

П.8. Функции от случайных величин

Рассмотренные нами основные законы распределения случайных величин, в чистом виде, встречаются не так уж и часто. Область их применения можно значительно расширить, если случайные величины, описывающие случайные явления, выражать через функцию от других случайных величин или хотя бы, через неслучайную функцию одной случайной величины.

Пусть, например, нас интересует распределение случайной величины η, которая связана функционально со случайной величиной ξ по формуле $\eta=\varphi(\xi)$, для которой функция распределения $F_{\xi}(x)$ известна. Задача состоит в нахождении функции распределения случайной величины η, где

$$F_{\eta}(y)=P\{\eta < y\},\ y\in R.$$

В некоторых случаях решение задачи может быть получено из здравого смысла, которое должно быть проверено формально.

Пример. Дискретная случайная величина ξ задана законом распределения

ξ	-1	0	1
p	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

Построить закон распределения случайной величины $\eta=\xi^2$.

Решение. Случайная величина η неотрицательна и, очевидно, принимает два значения. Закон распределения имеет вид:

η	0	1
p	$\frac{2}{3}$	$\frac{1}{3}$

Замечание. Законы распределения для случайных величин ξ^2 и $\xi\cdot\xi$ различны. Это означает, что $\xi^2\neq\xi\cdot\xi$. В самом деле, слева мы над случайной величиной произвели неслучайную математическую операцию: возведение в квадрат и получили два значения η: 0, 1, а справа стоит произведение двух случайных величин. Эта случайная величина η, для которой имеем три различных значения. Закон распределения имеет вид:

$\eta=\xi\cdot\xi$	-1	0	1
p	$\frac{2}{9}$	$\frac{5}{9}$	$\frac{2}{9}$

Рассмотренный пример демонстрирует подход построения закона распределения для функции одного случайного аргумента дискретной случайной величины.

Проведем построение функции распределения случайной величины η, являющейся функцией случайной величины ξ, с заданным распределением $F_\xi(x)$.

Определение. Пусть функция $y = \varphi(x)$ строго монотонна в области $(x\in X)\&(y\in У)$, где символ «&» читается как «и».

Функция φ^{-1} называется обратной к φ, если она определена на множестве $У$ и

$$\varphi^{-1}(\varphi(x)) = \varphi(\varphi^{-1}(x)) \equiv x\,.$$

Если $\varphi(x)$ возрастающая и $\varphi(x)< y$, то $x< \varphi^{-1}(y)$, если $\varphi(x)$ убывающая, то $x > \varphi^{-1}(y)$.

Рассмотрим случайную величину ξ с функцией распределения $F_\xi(x)$ и плотностью $\rho_\xi(x)$. Пусть $y = \varphi(x)$ строго монотонная и дифференцируемая, вместе со своей обратной, функция. Пусть случайная величина $\eta=\varphi(\xi)$. Требуется найти $F_\eta(y)$ и $\rho_\eta(y)$, где $y\in У$ - значения случайной величины η.

По определению, имеем

$F_\eta(y) = P\{\eta< y\}$, тогда, если $\varphi(x)$ возрастающая, то

$P\{\varphi(\xi) < y\} = P\{\xi < \varphi^{-1}(y)\}$.

Отсюда

$F_\eta(y) = F_\xi(\varphi^{-1}(y))$,

или

$$F_\eta(y) = \int\limits_{-\infty}^{\varphi^{-1}(y)} \rho_\xi(x)dx.$$

Для плотности получаем:

$$\rho_\eta(y) = \frac{d}{dy}\,F_\eta(y) = \rho_\xi\left(\varphi^{-1}(y)\right)\cdot\frac{d\left(\varphi^{-1}(y)\right)}{dy}. \qquad (42)$$

Если $\varphi(x)$ - убывающая, то

$P\left\{\varphi(\xi) < y\right\} = P\left\{\xi \geq \varphi^{-1}(y)\right\}$, но тогда

$F_\eta(y) = 1 - P\left\{\xi < \varphi^{-1}(y)\right\}$.

Отсюда

$$F_\eta(y) = 1 - \int\limits_{-\infty}^{\varphi^{-1}(y)} \rho_\xi(x)dx.$$

Для плотности, аналогично, получаем

$$\rho_\eta(y) = \frac{d}{dy}F_\eta(y) = -\rho_\xi\left(\varphi^{-1}(y)\right)\cdot\frac{d}{dy}\varphi^{-1}(y) \qquad (43)$$

Учитывая, что производная убывающей функции отрицательна, правая часть (43) положительна.

Объединяя (42) и (43), получаем:

$$\rho_\eta(y) = \rho_\xi\left(\varphi^{-1}(y)\right)\cdot\left|\frac{d}{dy}\varphi^{-1}(y)\right| \qquad (44)$$

Пример. Пусть случайная величина ξ имеет функцию распределения $F_\xi(x)$, а случайная величина $\eta = \alpha\xi + \beta$, $\alpha, \beta \in R$.

Найти $F_\eta(y)$ и $\rho_\eta(y)$, где y – значение случайной величины η из области У.

Решение. Если $\alpha > 0$, функция $y = \alpha x + \beta$ возрастающая, тогда

$$F_\eta(y) = F_\xi\left(\frac{y-\beta}{\alpha}\right),$$

отсюда

$$\rho_\eta(y) = \frac{d}{dy}F_\xi\left(\frac{y-\beta}{\alpha}\right) = \frac{1}{\alpha}\rho_\xi\left(\frac{y-\beta}{\alpha}\right)$$

Если $\alpha < 0$, то $y = \alpha x + \beta$ - убывающая, тогда

$$F_\eta(y) = 1 - F_\xi\left(\frac{y-\beta}{\alpha}\right),$$

отсюда

$$\rho_\eta(y) = -\frac{1}{\alpha}\rho_\xi\left(\frac{y-\beta}{\alpha}\right).$$

Учитывая (44), для плотности получаекм

$$\rho_\eta(y) = \frac{1}{|\alpha|}\rho_\xi\left(\frac{y-\beta}{\alpha}\right).$$

Пример. Пусть случайная величина ξ имеет экспоненциальное распределение $F(x) = 1 - \exp(-\lambda x)$, $x > 0$. Найти распределение случайной величины $\eta = e^{-\beta\xi}$, $\beta > 0$.

Решение. Имеем

$$F_\eta(y) = P\{\eta < y\} = P\{e^{-\beta\xi} < y\}.$$

Если $y < 0$, то событие $\{e^{-\beta\cdot\xi} < y\}$ невозможно, тогда $F_\eta(y) = 0$. Если $y > 1$, то событие $\{e^{-\beta\cdot\xi} < y\}$ - достоверное, тогда $F_\eta(y) = 1$.

Пусть теперь $y \in (0,1)$, тогда

$$P\{e^{-\beta\cdot\xi} < y\} = P\{-\beta\xi < \ln y\} = P\left\{\xi > -\frac{1}{\beta}\ln y\right\} = 1 - P\left\{\xi < -\frac{1}{\beta}\ln y\right\} =$$

$$= \exp\left(\frac{\lambda}{\beta}\ln y\right) = y^{\frac{\lambda}{\beta}}.$$

Таким образом, с учетом того, что $P\{\eta=0\}=P\{\eta=1\}=0$ имеем,

$$F_\eta(y)=\begin{cases}0, \text{ для } y\le 0,\\ y^{\frac{\lambda}{\beta}}, \text{ для } 0<y\le 1,\\ 1, \text{ для } y>1, \qquad \lambda,\beta>0.\end{cases}$$

П.9. Многомерные случайные величины

Развитие аппарата многомерных случайных величин в теории вероятностей так же важно, как и развитие функций нескольких переменных в математическом анализе.

Пусть имеем вероятностное пространство (Ω, $\mathcal{F}$,P), (Ω совпадает с R^n) с заданными на нем случайными величинами $\xi_1=\xi_1(\omega)$, $\xi_2=\xi_2(\omega)$,…, $\xi_n=\xi_n(\omega)$, $\omega\in\Omega$.

Определение. Случайную величину $\bar{\xi}(\omega)=(\xi_1, \xi_2, \dots, \xi_n)$ назовем n – мерным случайным вектором, являющимся отображением $\Omega \to R^n$. Отображение измеримо, в том смысле, что для любого множества из класса $\mathcal{F}$ определена функция распределения.

$$F_{\xi_1,\dots,\xi_n}(x_1,\dots,x_n)=F_{\bar{\xi}(\omega)}(\bar{x}), \text{ где } \bar{x}=(x_1,\dots,x_n)\,. \tag{45}$$

Функция распределения (45) однозначно определяет распределение вероятностей $P\{\xi_1< x_1, \xi_2< x_2,\dots,\xi_n< x_n\}$ и обладает свойствами, вполне аналогичными свойствам функции распределения одной переменной, а именно, $F_{\bar{\xi}(\omega)}(\bar{x})$:

1) неубывающая по каждому аргументу вектора $\bar{x}$;
2) $\lim\limits_{x_i\to+\infty} F(x_1,\dots,x_n)=F(x_1,x_2,\dots,x_{i-1},x_{i+1},\dots x_n)$, $i=1,2,\dots,n$;
3) непрерывна слева по каждому аргументу;
4) $\lim\limits_{x_\kappa\to-\infty} F(x_1,\dots,x_n)=0$, $\kappa=1,2,\dots,n$;

5) $$\lim_{\substack{x_1 \to \infty \\ \cdots\cdots\cdots \\ x_n \to \infty}} F(x_1, ..., x_n) = 1.$$

Дальнейшее построение теории многомерных случайных величин приведем для двухмерного случайного вектора (ξ, η).

Пусть Ω не обязательно совпадает с R^2. Рассмотрим случайный вектор $(\xi, \eta) \in \Omega_{\xi,\eta} \subset R^2$, где $\Omega_{\xi,\eta} = \Omega_\xi \times \Omega_\eta$. Любое подмножество $A \subset \Omega_{\xi,\eta}$ назовем событием. Класс $\mathcal{F}$ определим как алгебру событий, каждое из которых можно получить из множеств

$$((\xi < x) \& (\eta < y)), \text{ где } x \in \Omega_\xi,\ y \in \Omega_\eta.$$

Вероятность события A определим как

$$P_{\xi,\eta}\{A\} = P\{\omega / (\xi(\omega) \ \& \ \eta(\omega)) \in A\}, \forall A \subset \Omega_{\xi,\eta}.$$

Тем самым построено вероятностное пространство $(\Omega_{\xi,\eta}, P_{\xi,\eta})$, как частный случай рассмотренного ранее.

Из свойств функции распределения легко получить:

а) $P\{a_1 \le \xi < в_1,\ \eta < y\} = F(в_1, y) - F(a_1, y)$

Графическая иллюстрация представлена на рис. 17 (включение границы в допустимую область обозначено жирной чертой).

б) $P\{a_1 \le \xi < в_1,\ a_2 \le \eta < в_2\} = F(в_1, в_2) - F(a_1, в_2) - F(в_1, a_2) + F(a_1, a_2)$

Графическая иллюстрация представлена на рис. 18.

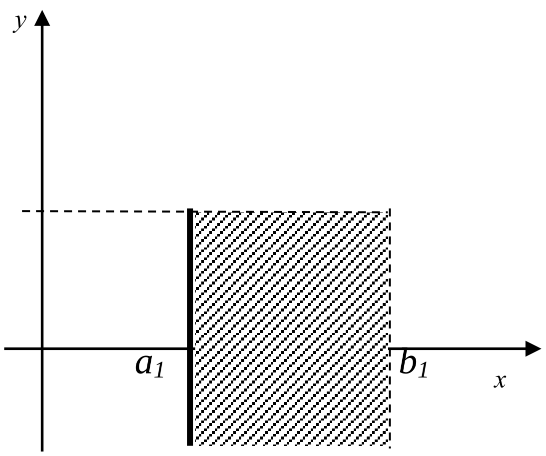


Рис. 17

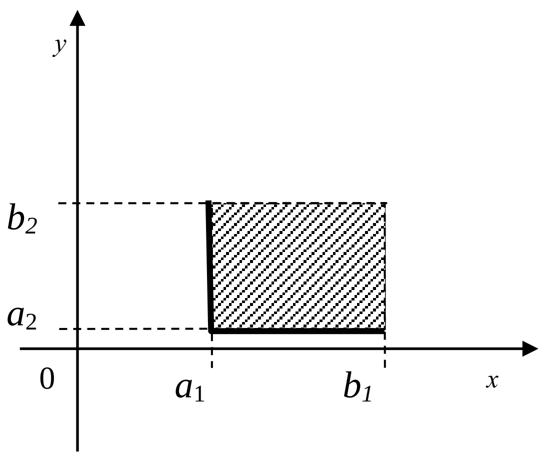


Рис. 18

Распределение вероятностей случайного вектора назовем дискретным, если он принимает не более, чем счетное число значений.

Распределение случайного вектора назовем абсолютно непрерывным, если для любого подмножества A, со значениями из R^2,

$$P\{(\xi,\eta)\in A\}=\iint_{A}\rho(x,y)\,dxdy$$

или в эквивалентной форме

$$F_{\xi,\eta}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}\rho(z,t)\,dzdt. \qquad (46)$$

Функция $\rho(x,y)$ называется плотностью распределения случайного вектора (ξ,η). Из (46) следует, что

$$\frac{\partial^2 F(x,y)}{\partial x\partial y}=\rho(x,y).$$

Если существует плотность $\rho(x,y)$, то существуют и плотности $\rho_\xi(x)$, $\rho_\eta(y)$.

Рассмотрим два примера, часто используемые в приложениях.

Пример (полиномиальное распределение).

Пусть (ξ,η) - целочисленный случайный вектор. Распределение (ξ,η) зададим формулой:

$$P_n\{\ \xi=k_1,\ \eta=k_2\}=\frac{n!}{k_1!\cdot k_2!\cdot(n-k_1-k_2)!}\cdot p_1^{k_1}\cdot p_2^{k_2}\cdot(1-p_1-p_2)^{n-k_1-k_2},$$

где $0\le \kappa_1+\kappa_2\le n$, $p_1=P\{A_1\}$, $p_2=P\{A_2\}$, $1-p_1-p_2=P\{\Omega\setminus(A_1\cup A_2)\}$, $A_1, A_2\subset\Omega$, $A_1\cap A_2=\varnothing$.

Применим описанное распределение при построении математической модели процесса разделения сыпучих материалов.

При разделении сыпучего материала на три группы по среднему диаметру частиц (события A_i) установлено, что вероятность $P\{A_i\}$ частицы принадлежать группе i равна p_i, i=1,2,3. Распределение вероятностей, того, что среди n частиц κ_i частиц принадлежит группе i ($\kappa_1+\kappa_2+\kappa_3=n$) определяется по формуле полиномиального распределения, где $\kappa_3=n-\kappa_1-\kappa_2$.

Пример (Двумерное нормальное распределение).

Пусть дана двумерная случайная величина (ξ,η). Будем говорить, что она нормально распределена, если ее функция распределения имеет вид:

$$\Phi_{\xi,\eta}(x,y)=\frac{1}{2\pi\cdot\sigma_\xi\cdot\sigma_\eta\cdot\sqrt{1-r^2}}\cdot\int_{-\infty}^{x}\int_{-\infty}^{y}\exp\left(-\frac{1}{2\left(1-r^2\right)}\cdot k^2\left(t,z\right)\right)dtdz$$

где $\kappa^2\left(x,y\right)=\frac{\left(x-M\xi\right)^2}{\sigma_\xi^2}-\frac{2r\left(x-M\xi\right)\cdot\left(y-M\eta\right)}{\sigma_\xi\cdot\sigma_\eta}+\frac{\left(y-M\eta\right)}{\sigma_\eta^2}$ - называется эллипсом рассеивания, а $r=\frac{M\left(\left(\xi-M\xi\right)\cdot\left(\eta-M\eta\right)\right)}{\sigma_\xi\cdot\sigma_\eta}$ - называется коэффициентом корреляции, о котором будем говорить ниже.

Из свойств функций распределения многомерных случайных величин следует, что, зная закон вектора (ξ,η), можно определить законы распределения отдельных случайных величин ξ, η. В самом деле, имеем

$$F_\xi\left(x\right)=F_{\xi,\eta}\left(x,\infty\right)=\int_{-\infty}^{x}\int_{-\infty}^{\infty}\rho\left(t,y\right)\,dtdy\,. \tag{47}$$

Дифференцируя (47) по x, получаем, для плотности распределения ξ, формулу:

$$\rho_\xi(x)=F_\xi^{'}\left(x\right)=\int_{-\infty}^{\infty}\rho\left(x,y\right)dy\,.$$

Аналогично, получаем выражение плотности распределения η:

$$\rho_\eta(y)=F_\eta^{'}\left(y\right)=\int_{-\infty}^{\infty}\rho\left(x,y\right)\,dx\,.$$

Естественно, возникает обратная задача: зная законы распределения отдельных случайных величин, найти закон распределения случайного вектора?

Для независимых случайных величин это, очевидно, можно, но для зависимых, в общем случае, нельзя, если зависимость неизвестна.

Обратная задача решается положительно, если ввести **условные законы** распределения.

Определение. Условным законом распределения случайной величины ξ случайного вектора (ξ,η), называется функция распределения $F_\xi(x/y)$, полученная при условии, что случайная величина $\eta = y$.

Аналогично определяется условная плотность распределения $\rho_\xi(x/y)$.

Зная закон распределения одной из случайных величин и условный закон распределения другой, можно построить закон распределения двумерного вектора.

Из определения следует, что

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

отсюда, по определению условной вероятности,

$$F_\xi(x/y) = P\{(\xi < x)/(\eta = y)\} = \frac{1}{\rho_\eta(y)}\int_{-\infty}^{x}\rho(t,y)dt.$$

Дифференцируя это равенство по x, получаем условную плотность

$$\rho_\xi(x/y) = \frac{1}{\rho_\eta(y)}\cdot\rho(x,y). \qquad (48)$$

Из (48) следует, что

$$\rho(x,y) = \rho_\eta(y)\cdot\rho(x/y) = \rho_\xi(x)\rho(y/x). \qquad (49)$$

Формула (49) называется **теоремой умножения законов распределения** случайных величин.

Аналогичные формулы можно получить для условного закона распределения случайной величины η вектора (ξ,η).

Определение. Случайные величины ξ, η называются **независимыми**, если закон каждой из них не зависит от того, какое значение приняла другая случайная величина, иначе они зависимы.

Независимость означает, что для случайной величины (ξ,η), $\forall(x,y) \in R^2$ выполняется

$$P\{(\xi < x),(\eta < y)\} = P\{\xi < x\}\cdot P\{\eta < y\} \tag{50}$$

или

$$F_{\xi,\eta}(x, y) = F_{\xi}(x)\cdot F_{\eta}(y). \tag{51}$$

Если случайная величина (ξ,η) имеет плотность, то

$$\rho_{\xi,\eta}(x, y) = \rho_{\xi}(x)\cdot\rho_{\eta}(y). \tag{52}$$

Любое из этих равенств является необходимым и достаточным условием независимости случайных величин ξ,η.

Пример. Случайная величина (ξ,η) задана законом распределения (табл.4).

Таблица 4

$\eta \backslash \xi$	0	1
0	$\frac{1}{3}$	$\frac{1}{3}$
1	$\frac{1}{6}$	$\frac{1}{6}$

Проверить, являются ли случайные величины ξ,η независимыми, если их законы распределения имеют вид:

ξ	0	1
p_{ξ}	$\frac{1}{2}$	$\frac{1}{2}$

η	0	1
p_{η}	$\frac{2}{3}$	$\frac{1}{3}$

Решение. Для проверки, достаточно показать справедливость формулы (50) или (51).

Положим $p_{1\bullet} = P\{\xi = 0\}$, $p_{2\bullet} = P\{\xi = 1\}$, а $p_{\bullet 1} = P\{\eta = 0\}$, $p_{\bullet 2} = P\{\eta = 1\}$.

Из свойства 2), функции распределения многомерных случайных величин, следует, что

а) для случайной величины ξ:

$$p_{1\bullet} = p_{11} + p_{12}, \; p_{2\bullet} = p_{21} + p_{22},$$

б) для случайной величины η:

$$p_{\bullet 1} = p_{11} + p_{12}, \; p_{\bullet 2} = p_{21} + p_{22},$$

где $\begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$.

Представим табл. 4 в виде:

$\eta \backslash \xi$	0	1	$\sum$
0	$\frac{1}{2}\cdot\frac{2}{3}$	$\frac{1}{2}\cdot\frac{2}{3}$	$\frac{2}{3}$
1	$\frac{1}{2}\cdot\frac{1}{3}$	$\frac{1}{2}\cdot\frac{1}{3}$	$\frac{1}{3}$
Σ	$\frac{1}{2}$	$\frac{1}{2}$	$\sum = 1$

Отсюда видно, что условие независимости (50) выполняется. Для проверки достаточно рассмотреть всевозможные пары с повторениями из выборки $(-0,5;\ 0,5;\ 1,5)$ для пары (x,y). Например, из $(x, y) = (0,5; 0,5)$ следует, что $P\{\xi < 0,5;\ \eta < 0,5\} = 1/3$. С другой стороны $P\{\xi < 0,5\} \cdot P\{\eta < 0,5\} = 1/2 \cdot 2/3 = 1/3$.

Важное значение в теории вероятностей и ее приложениях имеют суммы случайных величин. Нахождение функции распределения суммы по известным распределениям отдельных случайных величин является первостепенным.

Теорема. Пусть ξ_1 и ξ_2 положительно определенные независимые случайные величины с распределениями F_1 и F_2. Если $\rho_2(x)$ – плотность случайной величины ξ_2, то

$$P\{\xi_1+\xi_2<x\}=\int_0^x F_1(x-t)\cdot\rho_2(t)\,dt,\ x\in[0,\infty).$$

Доказательство. Рассмотрим промежуток $[0, x]$ и разобьем его на n частей $x_0 = 0 < x_1 < \ldots < x_{n-1} < x_n = x$. Пусть $\Delta x_i = x_i - x_{i-1}$ - длина интервала i, $i = 1, 2,\ldots, n$. Рассмотрим события $\{A_i, B_i\} = \{\xi_1 < x - x_i,\ \xi_2 \in (x_{i-1}, x_i]\}$. Реализация любого из них приводит к появлению события $\{\xi_1 + \xi_2 < x\}$.

Так как события $\{A_i, B_i\}$ несовместны, а случайные величины ξ_1 и ξ_2 независимы, то $P\{\xi_1+\xi_2<x\}\geq\sum\limits_{i=1}^{n}P\{A_i\}\cdot P\{B_i\}$.

Учитывая, что $P\{B_i\}=P\{x_{i-1}<\xi_2\leq x_{i-1}+\Delta x_i\}=\rho(\tilde{x}_i)\Delta x_i$, где $\tilde{x}_i$ некоторое число из промежутка $(x_{i-1}, x_i]$, получаем

$$P\{\xi_1+\xi_2<x\}=\sum_{i=1}^{n}F(x-x_i)\rho_2(\tilde{x}_i)\Delta x_i.$$

Рассматривая правую часть как интегральную сумму, при $n\to\infty$ и $\max\Delta x_i\to 0$, будем иметь

$$P\{\xi_1+\xi_2<x\}=\int_0^x F(x-t)\cdot\rho_2(t)\,dt. \blacktriangledown \qquad (53)$$

Если $\rho_1(x)$ – плотность случайной величины ξ_1, то аналогично

$$P\{\xi_1+\xi_2<x\}=\int_0^x F_2(x-t)\cdot\rho_1(t)\,dt. \qquad (53')$$

Формулы (53) и (53') называются **сверткой** распределений F_1 и F_2 и обозначаются символом

$$F_1 * F_2(x) = \int_0^x F_1(x-t)\rho_2(t)dt.$$

Дифференцируя формулы (53) и (53') по аргументу *x*, получаем аналогичную формулу для плотности суммы ($\xi_1 + \xi_2$):

$$\rho_{\xi_1+\xi_2}(x) = \int_0^x \rho_1(x-t)\cdot\rho_2(t)dt = \int_0^x \rho_2(x-t)\cdot\rho_1(t)dt.$$

Если случайные величины ξ_1, ξ_2 определены на всей действительной оси, то для свертки имеет место

$$\rho_{\xi_1+\xi_2}(x) = \int_{-\infty}^{\infty} \rho_1(x-t)\cdot\rho_2(t)dt = \int_{-\infty}^{\infty} \rho_2(x-t)\cdot\rho_1(t)dt. \quad (54)$$

Доказательство можно найти, например, в [5].

В классе распределений операция свертки обладает свойствами:

1) коммутативности: $F_1 * F_2 = F_2 * F_1$;

2) ассоциативности: $F_1 * (F_2 * F_3) = (F_1 * F_2) * F_3$;

3) дистрибутивности:

$(\alpha_1 F_1 * \alpha_2 F_2) * F_3 = \alpha_1(F_1 * F_2) + \alpha_2(F_2 * F_3),\ \alpha_1, \alpha_2 \in R.$.

Пример. Пусть даны случайные величины ξ_i, $i = 1,2,3,4$, равномерно распределенные на [0,1). Требуется вычислить плотность их суммы для различных *i*.

Решение. Имеем $\rho_{\xi_i}(x) = \begin{cases} 0, & x \notin [0,1), \\ 1, & x \in [0,1), \end{cases} \quad i = 1,2,3,4.$

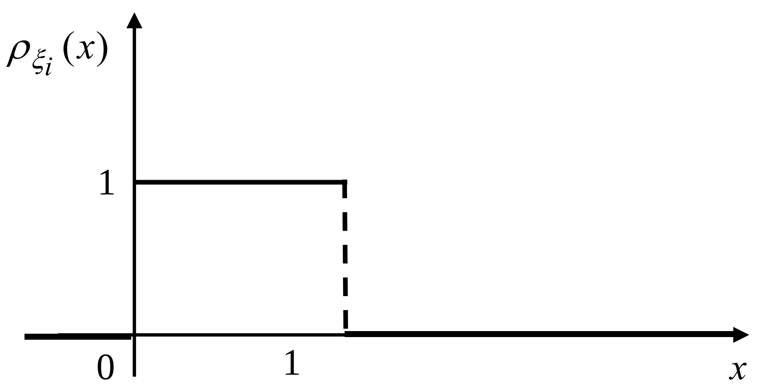
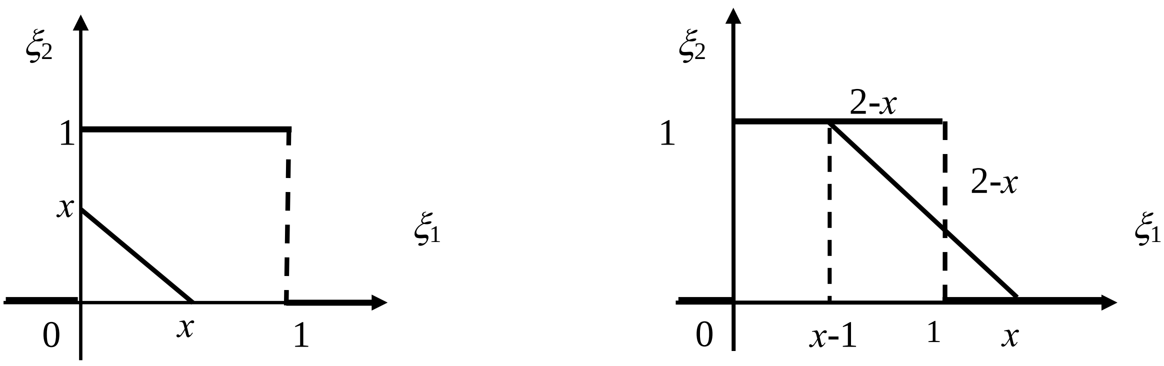


Рис. 19

Для вычисления плотности воспользуемся геометрическим подходом. Изобразим ξ_1 и ξ_2 на плоскости в декартовой системе координат

Вычислим вероятности событий $\xi_1 + \xi_2 < x$ для разных $x \in R$. Очевидно, что

$$P\{(\xi_1 + \xi_2 < x)/(x \in [-\infty, 0))\} = 0, \text{ а } P\{(\xi_1 + \xi_2 < x)/(x \in [2, \infty))\} = 1.$$



а) б)

Рис. 20

Пусть $x \in [0,1]$ (рис.20, а)), тогда $P\{(\xi_1 + \xi_2 < x)/(x \in [0,1))\} = \dfrac{x^2}{2}$,

соответствующая плотность $\rho_{\xi_1+\xi_2}(x) = x, \quad x \in [0,1]$.

Для $x \in [1,2]$ (рис. 20, б)), имеем

$P\left\{(\xi_1 + \xi_2 < x)/(x \in [1,2])\right\} = 1 - \dfrac{(2-x)^2}{2}$, отсюда $\rho_{\xi_1+\xi_2}(x) = 2 - x$, $\forall\, x \in [1,2]$.

Таким образом, получаем

$$\rho_{\xi_1+\xi_2}(x) = \begin{cases} 0, \ x \notin [0,2]; \\ x, \ x \in [0,1]; \\ 2 - x, \ x \in (1,2]. \end{cases} \qquad (55)$$

Графическая иллюстрация плотности представлена на рис. 21.

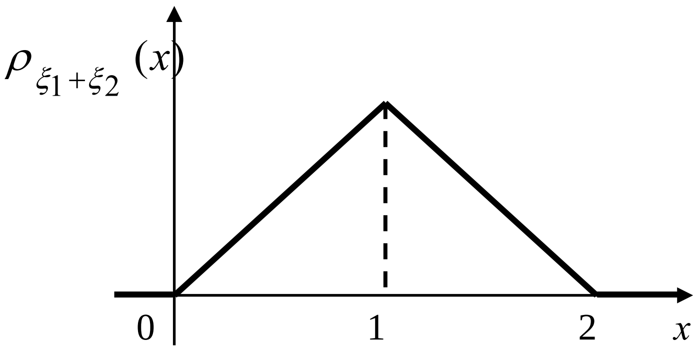


Рис. 21

Для более общего случая геометрический образ трудно воспринимается, поэтому воспользуемся следующими рассуждениями.

Для $x \notin [0,2]$, очевидно, $\rho_{\xi_1+\xi_2}(x)=0$.

Пусть $x \in [0,1]$, тогда по формуле свертки, имеем (рис. 22, а)) $\rho_{\xi_1+\xi_2}(x)=\int\limits_0^x dy = x$.

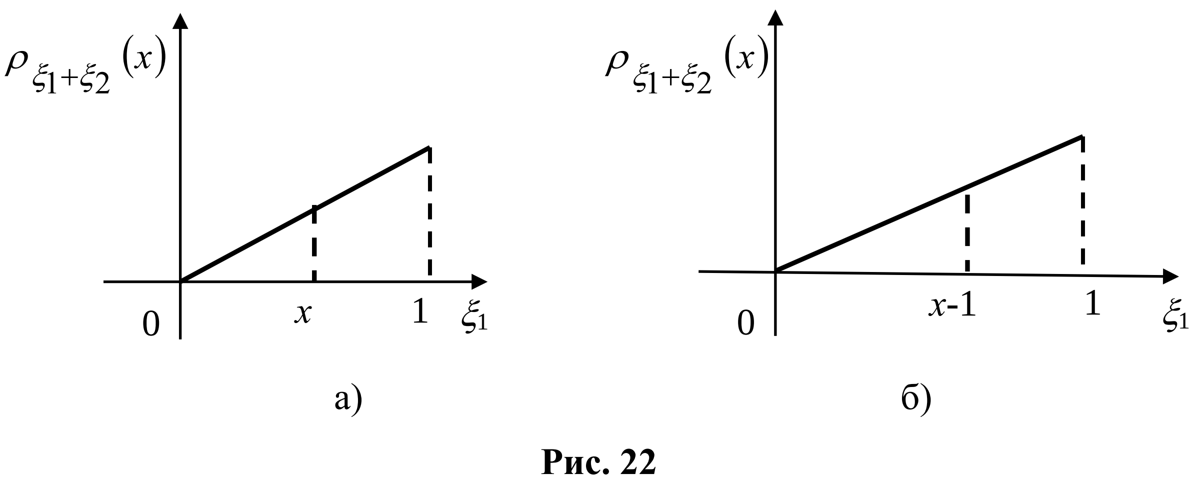


Рис. 22

Для $x \in [1,2]$, получаем (рис. 22, б)) $\rho_{\xi_1+\xi_2}(x)=\int\limits_{x-1}^{1} dy = 2-x$.

По аналогии с предыдущим вычислим плотность:

$$\rho_{\xi_1+\xi_2+\xi_3}(x)=\begin{cases} 0, \text{ для } x\notin[0,3), \\ \int\limits_0^x (x-y)\,dy=\dfrac{x^2}{2}, \text{ для } x\in[0,1], \\ \int\limits_{x-1}^1 (x-y)\,dx+\int\limits_0^{x-1} 2-(x-y)\,dy= \\ =1-\dfrac{(x-1)^2}{2}-\dfrac{(2-x)^2}{2}, \text{ для } x\in(1,2], \\ \int\limits_{x-2}^1 2-(x-y)\,dy=\dfrac{(3-x)^2}{2}, \text{ для } \; x\in(2,3]. \end{cases}$$

Плотность суммы $\xi_1+\xi_2+\xi_3$ «склеена» из трех кусков различных парабол.

Соответствующая кривая напоминает плотность нормального распределения (рис. 23), что является не случайным. Объяснение этому факту дает центральная предельная теорема, о которой мы будем говорить ниже.

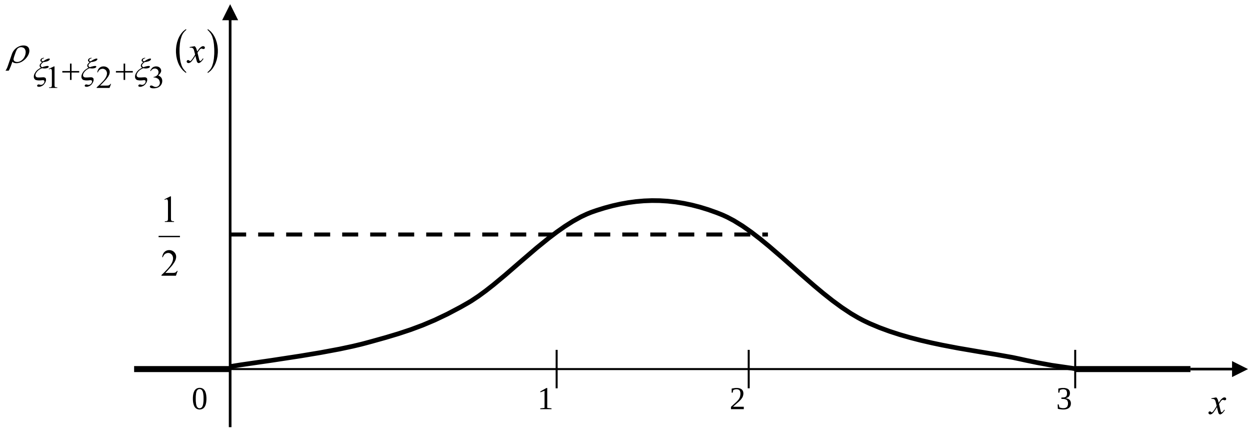


Рис. 23

Наконец, для суммы $\xi_1+\xi_2+\xi_3+\xi_4$, имеем

$$\rho_{\xi_1+\xi_2+\xi_3+\xi_4}(x)=\begin{cases} 0,\ \text{для } x\notin[0,4), \\ \int\limits_0^x \dfrac{(x-y)^2}{2}\,dx,\ \text{для } x\in[0,1], \\ \int\limits_{x-1}^{1} \dfrac{(x-y)^2}{2}\,dy+\int\limits_0^{x-1} 1-\dfrac{(2-(x-y))^2}{2}- \\ -\dfrac{(x-y-1)^2}{2}\,dy,\ \text{для } x\in(1,2], \\ \int\limits_{x-2}^{1} 1-\dfrac{(2-(x-y))^2}{2}-\dfrac{(x-y-1)^2}{2}\,dy+ \\ +\int\limits_0^{x-2}(3-(x-y))^2\,dy,\ \ \text{для } \ x\in(2,3], \\ \int\limits_{x-3}^{1}\dfrac{(3-(x-y))^2}{2}\,dy,\ \text{для } \ x\in(3,4). \end{cases}$$

Или, окончательно,

$$\rho_{\xi_1+\xi_2+\xi_3+\xi_4}(x)=\begin{cases} 0,\ \text{для } x\notin[0,4), \\ \dfrac{x^3}{6},\ \text{для } x\in[0,1], \\ x-1-\dfrac{(x-3)^3}{3}+\dfrac{(2-x)^3}{6},\ \text{для } \ x\in(1,2], \\ 3-x+\dfrac{(x-2)^3}{6}-\dfrac{(3-x)^3}{6},\ \text{для } \ x\in(2,3], \\ \dfrac{(4-x)^3}{6},\ \ \text{для } \ \ x\in(3,4]. \end{cases}$$

Искомая плотность «склеена» из четырех кусков кривых третьего порядка, которая еще ближе приближена к нормальной кривой.

Задача. Прибор испытывается на экстремальную нагрузку. Установлено, что время x, до которого он выдерживает нагрузку, является случайной величиной ξ с равномерным распределением на единичном отрезке [0,1).
Требуется найти $P\{\xi+\xi< t\}$ - вероятность того, что за время t, $t\in[0,2]$, прибор откажет два раза. Решить задачу для $t \div 0{,}5;\ 1;\ 1{,}5$ (временем восстановления прибора пренебречь).

Решение. Пусть x момент первого выхода прибора из строя, $x \in [0,1]$, тогда за оставшееся время $t - x$ прибор должен выйти из строя еще раз. Иллюстрация этого факта приведена на рис. 24.

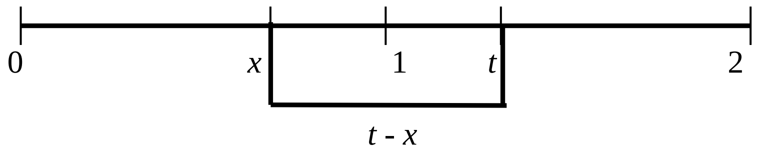


Рис. 24

Пусть $\rho_\xi(x)$ - плотность распределения времени работы прибора до первого отказа, тогда $\rho_{\xi+\xi}(t)$ - плотность распределения времени двукратного выхода прибора из строя за время t, которая вычисляется по формуле:

$$\rho_{\xi+\xi}(t) = \int_0^1 \rho_\xi(t-x)\rho_\xi(x)dx /$$

или, учитывая результаты предыдущего примера (формула (55)), получаем

$$\rho_{\xi+\xi}(t) = \begin{cases} 0, & t \notin [0,2]; \\ t, & t \in [0,1]; \\ 2-t, & t \in (1,2]. \end{cases}$$

Таким образом

$$P\{\xi + \xi < t\} = \int_0^t \rho_{\xi+\xi}(x)dx$$

или

$$P\{\xi + \xi < t\} = \begin{cases} 0, & t \notin [0,2]; \\ \dfrac{t^2}{2}, & t \in [0,1]; \\ 2t - \dfrac{t^2}{2} - 1, & t \in (1,2]. \end{cases}$$

Вероятности для конкретных значений t приведены в табл. 5

Таблица 5

t	0,5	1	1,5
$P\{\xi+\xi<t\}$	1/8	1/2	7/8

Как видно из примеров, изучение суммы случайных величин имеет важное значение (например, в теории надежности, теории восстановления, химии, технологических процессах и др. [1]).

Если для математического ожидания суммы случайных величин задача решается просто, то для нахождения их дисперсии мы сталкиваемся с трудностями, если случайные величины зависимы.

П.10. Моменты многомерных случайных величин

Определим, по аналогии, начальные и центральные моменты системы двух случайных величин в предположении существования их ряда распределения (или плотности).

Определение. **Начальным** моментом порядка $(\kappa + r)$ системы случайных величин (ξ,η) называется число

$$\nu_{k,r} = M(\xi^k\cdot\eta^r) = \begin{cases} \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}(x_i)^k\cdot(y_j)^r p_{ij}, & \text{если } (\xi,\ \eta) \text{ - дискретная} \\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} x^k\cdot y^r \rho(x,y)dxdy, & \text{если } (\xi,\ \eta) \text{ - непрерывная} \end{cases}$$

В частности, $\nu_{1,0} = M\xi$, $\nu_{0,1} = M\eta$, $n\in N$.

Определение. **Центральным** моментом порядка $(\kappa+r)$ системы случайных величин $(\xi,\ \eta)$ называется число

$$\mu_{k,r} = M\Big(\big(\xi - M\xi\big)^k\cdot\big(\eta - M\eta\big)^r\Big), \tag{57}$$

в частности $\mu_{2,0} = D\xi$, $\mu_{0,2} = D\eta$.

Выясним что представляют собой начальный момент $\nu_{1,1}$ и центральный момент $\mu_{1,1}$ системы (ξ,η).

Рассмотрим более общую задачу. Пусть $\xi_1, \xi_2, \ldots, \xi_n$ - произвольные случайные величины. Вычислим дисперсию их суммы:

$$D\left(\sum_{i=1}^{n}\xi_i\right)=M\left(\sum_{i=1}^{n}\xi_i-M\left(\sum_{i=1}^{n}\xi_i\right)\right)^2=M\left(\sum_{i=1}^{n}\left(\xi_i-M\xi_i\right)\right)^2=M\left(\sum_{i=1}^{n}(\xi_i-M\xi_i)^2+\right.$$

$$\left.+2\sum_{i<j}\left(\xi_i-M\xi_i\right)\cdot\left(\xi_j-M\xi_j\right)\right)=\sum_{i=1}^{n}M\left(\xi_i-M\xi_i\right)^2+2\sum_{i<j}M\left(\left(\xi_i-M\xi_i\right)\cdot\left(\xi_j-M\xi_j\right)\right)$$

или

$$D\left(\sum_{i=1}^{n}\xi_i\right)=\sum_{i=1}^{n}D\xi_i+2\sum_{i<j}M\left(\left(\xi_i-M\xi_i\right)\cdot\left(\xi_j-M\xi_j\right)\right). \tag{58}$$

Число слагаемых во второй сумме, правой части (58), равно C_n^2.

Если бы случайные величины были независимы, то, в силу свойств математического ожидания,

$$\forall i\neq j,\ M\left(\left(\xi_i-M\xi_i\right)\cdot\left(\xi_j-M\xi_j\right)\right)=M\left(\xi_i-M\xi_i\right)\cdot M\left(\xi_j-M\xi_j\right)=0.$$

Число $M\left(\left(\xi_i-M\xi_i\right)\cdot\left(\xi_j-M\xi_j\right)\right)$, в формуле (58), является вторым смешанным центральным моментом $\mu_{1,1}$ пары случайных величин ξ_i, ξ_j.

Его можно рассматривать как меру зависимости случайных величин.

Определение. **Ковариацией** случайных величин ξ, η называется математическое ожидание произведения их отклонений

$$Cov\left(\xi,\eta\right)=M\left(\left(\xi-M\xi\right)\cdot\left(\eta-M\eta\right)\right). \tag{59}$$

Ковариация существует, если существуют дисперсии каждой из случайных величин. Очевидно, что дисперсия есть частный случай ковариации, так как при $\xi=\eta$ имеем $Cov(\xi,\xi)=M\left(\xi-M\xi\right)^2=D\xi$.

Если случайные величины независимы, то ковариация равна 0. Утверждение сразу следует из свойств математического ожидания отклонений этих величин. Обратное неверно.

Вместо формулы (59) часто используется формула

$$Cov(\xi,\eta)=M(\xi\cdot\eta)-M\xi\cdot M\eta. \tag{60}$$

В самом деле, имеем

$$M\left(\left(\xi - M\xi\right)\cdot\left(\eta - M\eta\right)\right) = M\left(\xi\cdot\eta - \xi\cdot M\eta - \eta\cdot M\xi + M\xi\cdot M\eta\right) = M(\xi\eta) -$$

$$-2M\xi\cdot M\eta + M\xi\cdot M\eta \;=\; M(\xi\eta) - M\xi\cdot M\eta\,. \blacktriangledown$$

Пример. (Игра в лотерею). У каждого играющего в лотерею свой номер. Карточки с номерами собирают и тщательно тусуют. Затем по очереди, в соответствие с номером, игроки подходят и берут карточки. Получает приз тот, кто взял свой номер. Оценить, сколько призов в среднем следует приготовить.

Решение. Определим сл. в. $\xi_\kappa = \begin{cases} 1, & \text{произошлособытие}\, A_\kappa, \\ 0, & \text{произошлособытие}\, \overline{A_\kappa}. \end{cases}$

A_κ~{игрок с номером κ вытащил карточку с номером κ}. Пусть $\xi = \sum\limits_{\kappa=1}^{n}\xi_\kappa$ - случайная величина, характеризующая число призеров. Ясно, что свою карточку игрок берет с вероятностью n^{-1}:

$$P\{\xi_k = 1\} = \frac{1}{n},\; P\{\xi_k = 0\} = \frac{n-1}{n},\; \forall\, k\,,$$

но тогда

$$M\xi_k = 1\cdot\frac{1}{n} + 0\cdot\frac{n-1}{n} \;=\; \frac{1}{n},$$

то есть, имеем одно совпадение при любом n.

Найдем дисперсию:

$$D\xi = D\sum\xi_k = \sum D\xi_k + 2\sum_{i<j} Cov\left(\xi_i,\xi_j\right),\; \forall\, k\,,$$

$$D\xi_k = M\left(\xi_k\right)^2 - \left(M\xi_k\right)^2 = \frac{1}{n} - \frac{1}{n^2} = \frac{n-1}{n^2},$$

$$Cov\left(\xi_i,\xi_j\right) = M\left(\left(\xi_i - M\xi_i\right)\cdot\left(\xi_j - M\xi_j\right)\right) = M(\xi_i\cdot\xi_j) - M\xi_i\cdot M\xi_j\,.$$

Из определения ξ_k следует, что $\xi_i \cdot \xi_j$ может быть равно 1 или 0, причем $\xi_i \cdot \xi_j = 1$, если обе карты на своем месте, то есть $\forall\ i, j$,

$$M\left(\xi_i \cdot \xi_j\right) = \frac{1}{n} \cdot \frac{1}{n-1} = \frac{1}{n(n-1)}.$$

Итак, $Cov\left(\xi_i, \xi_j\right) = \dfrac{1}{n(n-1)} - \dfrac{1}{n^2} = \dfrac{1}{n^2(n-1)}$, а так как число слагаемых у второй суммы есть C_n^2, то

$$D\xi = n \cdot \frac{n-1}{n^2} + 2 \cdot C_n^2 \cdot \frac{1}{n^2 \cdot (n-1)} = \frac{n-1}{n} + \frac{n \cdot (n-1)}{n^2 \cdot (n-1)} = 1.$$

Таким образом, с учетом средних отклонений, получается, что независимо от числа игроков следует приготовить $M\xi + \sqrt{D\xi} = 1 + 1 = 2$ приза в среднем.

Пример. Пусть случайная величина (ξ, η) распределена равномерно в круге *D* радиуса *R* (для простоты центр круга поместим в начало координат). Определим плотность

$$\rho(x, y) = \begin{cases} \left(\pi R^2\right)^{-1}, \text{если } x^2 + y^2 \le R^2, \\ 0, \quad \text{если } x^2 + y^2 > R. \end{cases}$$

Вычислим $Cov(\xi, \eta)$. В силу симметрии, $M\xi = 0$, $M\eta = 0$. Далее,

$$Cov(\xi, \eta) = \iint_D x \cdot y \cdot \rho(x, y)\,dxdy = \frac{1}{\pi R^2} \cdot \int_{-R}^{R} dx \int_{-\sqrt{R^2 - x^2}}^{\sqrt{R^2 - x^2}} dy = 0,$$

то есть ξ, η некоррелированы, хотя и зависимы. В самом деле, имеем

$$\rho_\xi(x) = \int_{-\infty}^{\infty} \rho(x, y)\,dy = \int_{-R}^{R} \frac{1}{\pi R^2}\,dy = \frac{2R}{\pi R^2} = \frac{2}{\pi R},\quad \rho_\eta(y) = \int_{-\infty}^{\infty} \frac{1}{\pi R^2}\,dx = \frac{2}{\pi R},$$

тогда

$$\rho_\xi(x) \cdot \rho_\eta(y) = \frac{4}{\pi R^2} \ne \frac{1}{\pi R^2} = \rho(x, y).$$

Задача. Доказать, что для нормального распределения случайных величин ξ, η из некоррелированности вытекает их независимость.

Из примера видно, что величина ковариации зависит от размерности случайных величин ξ, η. Целесообразно ввести безразмерную характеристику, которая будет являться мерой зависимости случайных величин.

Определение. Случайная величина ξ^* называется **нормированной**, если $M\xi^*=0$, $D\xi^*=1$. Любую случайную величину ξ можно нормировать заменой

$$\xi^*=\frac{\xi-M\xi}{\sqrt{D\xi}}.$$

Определение. **Коэффициентом корреляции** $r(\xi,\eta)$ случайных величин ξ, η, входящих в двумерную случайную величину (ξ,η), назовем нормированную ковариацию:

$$r(\xi,\eta)=\frac{Cov(\xi,\eta)}{\sqrt{D\xi}\cdot\sqrt{D\eta}}.$$

Если $r(\xi,\eta)=0$, то говорят, что случайные величины ξ,η некоррелированы. Независимые случайные величины всегда некоррелированы.

Если рассматривать геометрический образ случайных величин в декартовых координатах, то для их некоррелированности достаточно, чтобы их совместное распределение было симметрично относительно прямой параллельной любой из осей координат.

Свойства коэффициента корреляции

1) $|r(\xi,\eta)|\le 1$.

В самом деле, имеем

$$D(\xi^*\pm\eta^*)=D\xi^*+D\eta^*+2r(\xi,\eta)=2\pm 2r(\xi,\eta)\ge 0,$$

или

$1\pm r(\xi,\eta)\ge 0,$ что эквивалентно неравенству $-1\le r(\xi,\eta)\le 1$. ▼

2) если ξ и η независимы, то $r(\xi,\eta)=0$. Доказательство следует из независимости ξ^* и η^*.▼

3) $r(\xi,\eta)=\pm 1$ тогда и только тогда, когда случайные величины линейно зависимы, то есть $\eta=\alpha\xi+\beta, \quad \alpha,\beta\in R.$

Если $\alpha<0$, то $r(\xi,\eta)=-1$, если $\alpha>0$, то $r(\xi,\eta)=1$.

Пусть $\eta=\alpha\xi+\beta$, тогда

$$r(\xi,\eta)=\frac{Cov(\xi,\alpha\xi+\beta)}{\sqrt{D\xi}\cdot\sqrt{D\xi}\cdot|\alpha|}=\frac{\alpha}{|\alpha|}M\left(\xi^*\right)^2=\frac{\alpha}{|\alpha|}\cdot 1=\frac{\alpha}{|\alpha|}.\blacktriangledown$$

Обратно, пусть $|r(\xi,\eta)|=1$. Возьмем $r(\xi,\eta)=-1$,

тогда

$$D\left(\xi^*+\eta^*\right)=2(1+r(\xi,\eta))=0,\ \text{следовательно}\ \xi^*+\eta^*=c.$$

Возвращаясь к ξ, η, получаем

$$\frac{\xi-M\xi}{\sqrt{D\xi}}+\frac{\eta-M\eta}{\sqrt{D\eta}}=c.$$

Для $r(\xi,\eta)=1$, возьмем $D(\xi^*-\eta^*)$. $\blacktriangledown$

4) Если $r(\xi,\eta)\neq 0$, то случайные величины зависимы.

Из свойств коэффициента корреляции следует, что он характеризует тесноту линейной связи в том смысле, что с возрастанием одной случайной величины другая линейно увеличивается (если $r(\xi,\eta)>0$) или линейно уменьшается (если $r(\xi,\eta)<0$), в среднем.

Пусть имеем набор случайных величин ξ_1, ξ_2, …, ξ_n, представляющих некоторый объект для исследования. Математический анализ этого объекта можно проводить, если известны:

а) математические ожидания $M\xi_i$,

б) дисперсии случайных величин $D\xi_i$,

в) парные коэффициенты корреляции r_{ij}, число которых равно C_n^2,

где $i, j = 1, 2, \ldots, n$.

Для оценки совокупного поведения системы случайных величин, часто рассматривают **корреляционную** матрицу

$$
(r_{ij}) = \begin{pmatrix} 1 & r_{12} & \dots & r_{1n} \\ r_{21} & 1 & \dots & r_{2n} \\ . & . & . & . \\ r_{n1} & r_{n2} & \dots & 1 \end{pmatrix},
$$

где $r_{ij} = r_{ji}$ – коэффициенты корреляции случайных величин ξ_i, ξ_j.

Изучение этой матрицы задача достаточно сложная и выходит за рамки данного курса [1].

Замечание. Можно рассматривать корреляционную матрицу, элементами которой являются корреляции случайных величин, а по главной диагонали идут их дисперсии. По такой матрице можно судить о величине рассеяния системы случайных величин относительно их среднего значения.

Еще одним видом зависимости случайных величин являются линии **регрессии**.

Пусть имеем пару $(\xi,\eta) = (\xi(x),\eta(y))$.

Определение. Условным математическим ожиданием случайной величины η (или ξ), при фиксированном значении случайной величины $\xi = x$, (или $\eta = y$) назовем число

$$
M\left(\eta / \xi = x \right) = \begin{cases} \sum\limits_j y_j \cdot p_{ij}, \text{если } \xi,\eta \text{ - дискретные}, \\ \int\limits_{-\infty}^{\infty} y \cdot \rho\left(y / x \right) dy, \text{если } \xi,\eta \text{ - непрерывные}, \end{cases} \tag{61}
$$

$$
\left(M\left(\xi / \eta = y \right) = \begin{cases} \sum\limits_i x_i \cdot p_{ij}, \text{если } \xi,\eta \text{ - дискретные}, \\ \int\limits_{-\infty}^{\infty} x \cdot \rho\left(x / y \right) dx, \text{если } \xi,\eta \text{ - непрерывные}. \end{cases} \right) \tag{62}
$$

Из формул (61), (62) видно, что при изменении значений x (значений y) изменяется и величина $M\left(\eta / \xi = x \right)$ $\left(\text{величина } M\left(\xi / \eta = y \right) \right)$.

Определение. Функция $\varphi(x) = M\left(\eta/_{\xi = x}\right)$ называется **функцией регрессии** случайной величины ξ на η. Аналогично, $\psi(y) = M\left(\xi/_{\eta = y}\right)$- функция регрессии случайной величины η на ξ.

Кривые, которые заданы функциями $y = \varphi(x)$ и $x = \psi(y)$, называются линиями регрессии. Ясно, что функции φ, ψ не являются, вообще говоря, обратными, поскольку взаимное влияние случайных величин друг на друга, как правило, различно. Это становится очевидным, если заметить, что при фиксированном значении $\xi = x$, значение $y = \varphi\ (x)$ определяется как условное среднее значений $y \in У$, которое может даже и не принадлежать множеству $У$.

Разумно считать, что линии регрессии $y = \varphi\ (x)$ и $x = \psi(y)$ различны и имеют, как правило, сложную функциональную зависимость.

Нахождение истинной линии регрессии задача трудновыполнимая. На практике, как правило, поступают следующим образом.

Если, из каких-либо соображений, общий вид линии регрессии известен, например, $y = \varphi\ (\alpha, \beta, \gamma)$, то задача состоит в нахождении числовых параметров α, β, γ, которые обычно являются композицией из моментов случайных величин. Эта задача решаемая.

Другим, наиболее часто используемым в инженерных расчетах подходом, является приближение линии регрессии многочленом степени n. Приближение тем точнее, чем выше его степень. Обычно выбирают метод наименьших квадратов (М.Н.К.). Пусть задан класс многочленов, накладывающий на выборку одинаковое число связей, которое равно числу неопределенных коэффициентов многочлена. Наилучшее уравнение приближенной регрессии дает тот многочлен степени n, для которого наименьшее значение имеет функция $\sum_{i=1}^{n} \left(y_i - f(x_i)\right)^2$, где (x_i, y_i) –совокупность опытных данных.

Пример. Пусть линия регрессии приближается линейной функцией, то есть

$$\hat{y} = \alpha x + \beta .$$

Здесь $\hat{y}$ означает, что истинная линия регрессии $y = \varphi(x)$ заменяется на линейную функцию.

Требуется подобрать коэффициенты α, β так, чтобы эта линейная функция была наилучшей.

Решение. Рассмотрим коэффициент корреляции, имеем

$$\sigma_\xi \cdot \sigma_\eta \cdot r(\xi,\eta) = Cov(\xi,\eta) = Cov(\xi,\alpha\xi+\beta)$$

где $\eta = \alpha\xi + \beta$, а $\sigma_\xi = \sqrt{D\xi}$, $\sigma_\eta = \sqrt{D\eta}$.

Учитывая (60),

$$Cov(\xi,\alpha\xi+\beta) = M\left(\xi \cdot (\alpha\xi+\beta)\right) - M\xi \cdot M(\alpha\xi+\beta),$$

или

$$Cov(\xi,\alpha\xi+\beta) = \alpha\left(M\xi^2 - (M\xi)^2\right) + \beta \cdot M\xi - \beta M\xi = \alpha\sigma_\xi^2 .$$

Отсюда

$$\alpha = \frac{Cov(\xi,\alpha\xi+\beta)}{\sigma_\xi^2} = \frac{\sigma_\eta}{\sigma_\xi} \cdot r(\xi,\eta).$$

Для нахождения β заметим, что, если линия регрессии линейна, то $\forall(x \in X)$ справедливо $M\left(\eta / \xi = x\right) = \alpha x + \beta$, то есть

$$M\eta = \alpha M\xi + \beta$$

или

$$\beta = M\eta - \alpha \cdot M\xi = M\eta - \frac{\sigma_\eta}{\sigma_\xi} \cdot r(\xi,\eta) \cdot M\xi .$$

Таким образом,

$$\hat{y} = r(\xi,\eta) \cdot \frac{\sigma_\eta}{\sigma_\xi} \cdot x + M\eta - \frac{\sigma_\eta}{\sigma_\xi} \cdot r(\xi,\eta) \cdot M\xi . \qquad (63)$$

Аналогично получаем

$$\hat{x} = r(\xi,\eta)\cdot\frac{\sigma_\xi}{\sigma_\eta}\cdot y + M\xi - \frac{\sigma_\xi}{\sigma_\eta}\cdot r(\xi,\eta)\cdot M\eta. \qquad (64)$$

Упражнение. Показать, что те же самые результаты дает метод наименьших квадратов.

Если $r(\xi,\eta)\neq 0$, то случайные величины зависимы и оцениваются острым углом между прямыми (63) и (64).

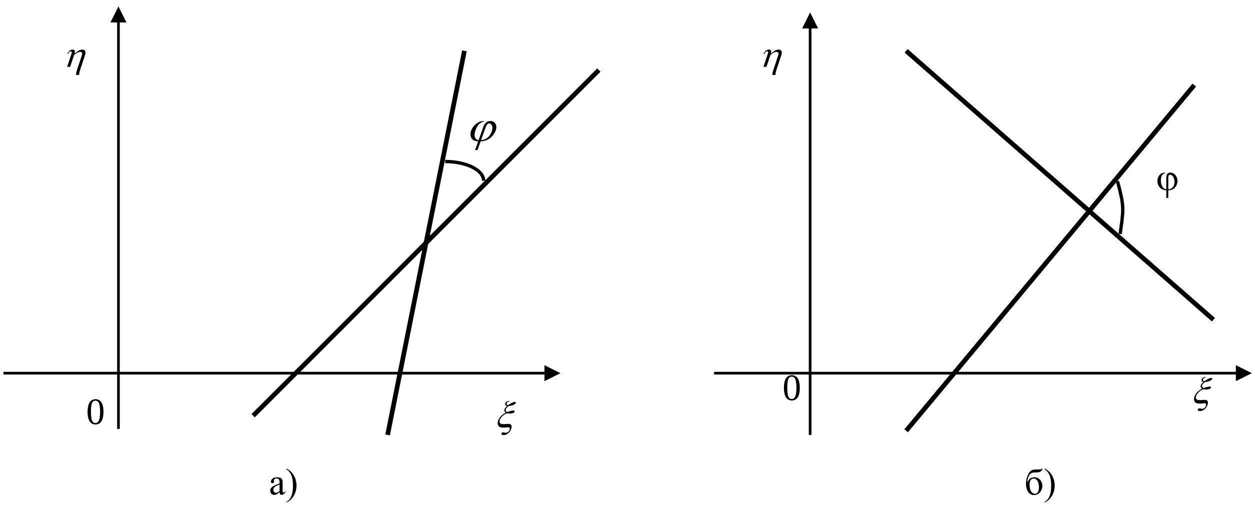


Рис. 25

На рис. 25 а), коэффициент корреляции $r(\xi,\eta) > 0$, так как с увеличением значений ξ, значения η увеличиваются. На рис. 25 б), коэффициент корреляции $r(\xi,\eta) < 0$, так как с увеличением значений ξ, значения η убывают. Если прямые перпендикулярны, то $r(\xi,\eta) = 0$. Это означает, что случайные величины некоррелированы.

П.11. Случайные процессы

Пусть имеем произвольное вероятностное пространство $(\Omega,\mathcal{F},P)$, на котором задана случайная величина $\xi = \xi(\omega)$, $\omega \in \Omega$. Рассмотрим функцию $\varphi(t)=\xi(\omega,t)$, где $t \in T$, T – множество, интерпретируемое нами как время.

Определение. Семейство случайных величин $\xi(t)$ параметра $t\in T$, определенное для каждого $\omega\in\Omega$, называется **случайным процессом**.

На вероятностном пространстве можно задать любое конечное, или даже бесконечное, семейство случайных величин. Однако, если они независимы, то такой процесс трудно назвать случайным. Отличительной особенностью случайного процесса является, именно, зависимость случайных величин, которая осуществляется через связующий параметр t. Задание вида зависимости приводит к возникновению классов случайных процессов (например, процессы с независимыми приращениями [2], марковские процессы [5], пуассоновские процессы [8] и др.), каждый из которых имеет свою, наиболее эффективную область применения. Способы описания случайных процессов, в основном, индивидуальны и приспособлены под тот или иной класс. Существует развитая общая теория случайных процессов [6], которая является в свою очередь, частью общей теории ***случайных функций***.

Пример 1. Последовательность случайных величин ξ_1, ξ_2, …, ξ_n, …, являющаяся результатом работы датчика случайных чисел, образующего двухзначные целые числа в моменты времени $t\in\{1,2,\ldots\}$, образует случайный процесс. В этом случае говорят о случайном процессе с **дискретным временем** и **дискретными состояниями**.

Пример 2. Срок службы фильтра, модулей по очистке воды, можно рассматривать как случайный процесс со временем $t\in T=[0,T_1]$, где пространство элементарных событий Ω = {хлор, нефтепродукты, фенол, пестициды, тяжелые металлы и др.}, T_1 – срок службы фильтра. Здесь случайный процесс с непрерывным временем и дискретным числом состояний.

Пример 3. Пусть время работы технического устройства является случайной величиной ξ с экспоненциальной функцией распределения

$$F(t) = 1 - \exp(-\lambda t), \quad \lambda = 1/t_{cp}, \tag{63}$$

t_{cp} – среднее время наработки прибора на отказ.
Число отказов $к = 0,1,2,\ldots n$ прибора за время T образует случайный процесс с непрерывным временем и дискретными состояниями.

Пусть имеем случайный процесс $\xi(t) = \varphi(\omega,t)$. При любом допустимом значении $t = t_1$ случайный процесс становится случайной величиной $\xi(t_1)=\varphi(\omega,t_1)$, для которой имеем свою функцию распределения $F_1(x) = F(x, t_1)$.

Будем говорить, что в этом случае имеем **сечение случайного процесса**. Если фиксировать случайное событие $\omega_1 \in \Omega$, то имеем реализацию случайного процесса, то есть неслучайную функцию $\xi_1 = \xi(\omega_1, t)$, являющейся функцией параметра t. Итак, при одном сечении процесса и одной реализации, вся информация о нем содержится в системе

$$\begin{cases} F_1(x) = F(x, t_1), \\ \xi_1 = \xi(\omega_1, t). \end{cases}$$

Интуитивно ясно, что этой информации явно недостаточно для изучения случайного процесса.

Рассмотрим пример 3. Так как имеем случайный процесс с непрерывным временем, то при $t = t_1$, случайная величина $\xi_1 = \varphi(\omega, t_1)$ имеет закон распределения, представленный табл. 6.

Таблица 6

$\xi(t_1)$	$x_0(t_1)$	$x_1(t_1)$	…	$x_n(t_1)$	…
$P(t_1)$	$P_0(t_1)$	$P_1(t_1)$	…	$P_n(t_1)$	…

$\left(\sum P_n(t_1) = 1\right)$.

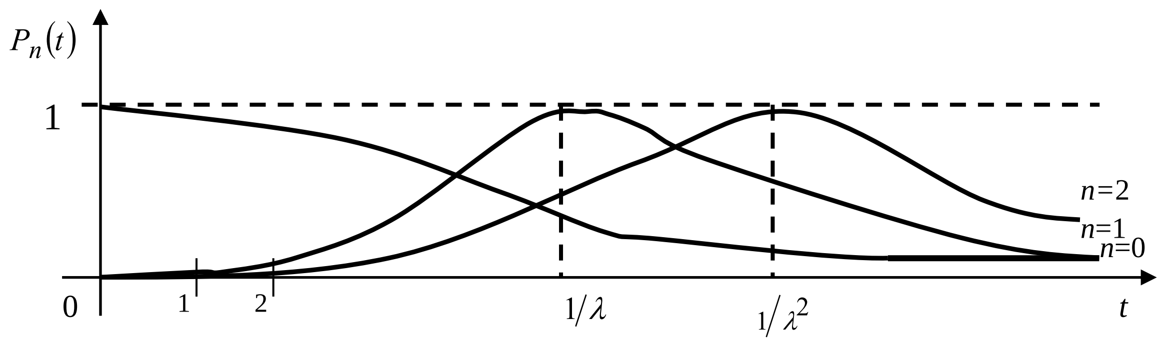


Рис. 26

Значения случайной величины $\xi(t_1) : x_n(t_1) = n,\ n = 0, 1, \dots$.

Зафиксируем элементарное событие $\omega = \omega_n$, $\omega_n \in \Omega$, тогда имеем реализацию случайного процесса, или функцию одной переменной $\xi_1 = \xi_1(t)$, $t \in [0,\infty)$.

Рис. 26 иллюстрирует ситуацию, когда в начальный момент времени t=0 прибор находится в рабочем состоянии. При n = 1 максимальное значение $\xi_1(t)=P_1(t)$ равно $1/\lambda$, при n = 2 максимальное значение $\xi_2(t) = P_2(t)$ равно $1/\lambda^2$.

Разобранный нами случайный процесс называется **пуассоновским**.

Из примера видно, что изучение случайных процессов тесно связано с распределением вероятностей. Однако далеко не всегда реализация случайного процесса имеет траекторию, представляющую «вероятностную кривую» (то есть без пересечений и со значениями из [0,1]).

Пример 4. Частица совершает движение «скачком», двигаясь по перпендикулярным прямым, на бумаге в клетку, случайным образом, фиксируя свое положение в точке пересечения прямых в моменты t_i, i = 1, 2, …, n. В результате имеем достаточно сложную кривую (рис. 27). Реализация случайного процесса – траектория, описываемая параметрически парой $(\xi_1(t), \xi_2(t))$, $t \in [0,T)$.

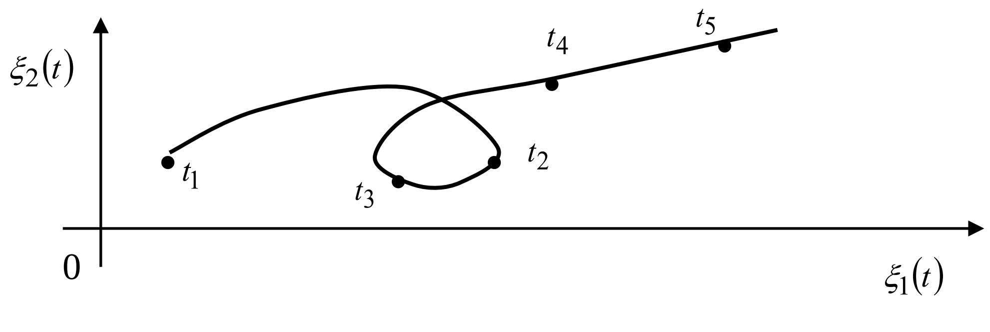


Рис. 27

Пусть имеем случайный процесс $\xi(t)=\varphi(\omega,t)$, тогда, при любом фиксированном t=t_1 (по аналогии со случайной величиной), имеем закон распределения

$$F(t,x) = P\{\xi(t) < x\}. \tag{64}$$

Функция (64), двух переменных, $F(t, x)$, называется одномерным законом распределения случайного процесса $\xi(t)$. Очевидно, что одномерный закон распределения не может являться достаточной характеристикой случайного процесса, поскольку он представляет только одно сечение, по которому невозможно делать выводы о всем процессе.

Ясно, что чем больше сечений, тем точнее задан случайный процесс, однако, и сложность изучения его резко возрастает. Для n сечений мы имеем функцию $2n$ переменных. На практике, более чем двумерные законы используются редко.

Рассмотрим двумерный закон распределения

$$F(t_1, t_2, x_1, x_2) = P\{\xi(t_1) < x_1, \xi(t_2) < x_2\}, \tag{65}$$

который составлен по двум сечениям процесса. Функция четырех переменных (65), оказывается, является исчерпывающей характеристикой для специального типа случайных процессов – процессов **без последействия** или **марковских** процессов (пример 3).

Марковские процессы широко используются в инженерной практике, поскольку основаны на независимости специального класса случайных событий.

П.12. Марковские процессы

Понятие марковкой цепи принадлежит русскому математику А.А. Маркову (в статьях 1906-1908 гг., где он использовал новое понятие для статистического анализа распределения букв в поэме А.С. Пушкина «Евгений Онегин»). Само понятие «Цепь Маркова» было предложено русским математиком А.Я. Хинчиным.

Пусть имеем некоторую систему S, которая может находиться в одном из конечного или счетного множества несовместных состояний C_i, $i \in N$. Переход системы от состояния к состоянию, вообще говоря, случаен и возможен только в фиксированные моменты времени t_n, $n = 0,1,2,\ldots$. Опишем функционирование системы в терминах случайных процессов.

Пусть в момент времени t_n, система S перешла из состояния C_j в состояние C_i. Для ее описания зададим дискретный случайный процесс функцией $\xi(t_n) = \varphi(\omega_i, t_n)$, $i = 1, 2, \ldots$; $n = 0, 1, \ldots$. Элементарное событие ω_i отражает пребывание системы S в состоянии C_i. Кроме того, нам необходимо задать начальное распределение вероятностей для момента времени $t = t_0$ и, в общем случае, задать все сечения процесса и возможность его реализации.

Получить такую информацию о случайном процессе задача трудновыполнимая, да и в ряде случаев не нужная, если использовать понятие **цепей Маркова**.

В самом деле, пусть имеем последовательность (цепь) зависимых целочисленных случайных величин $\xi_n = \xi(t_n)$, $n = 0,1,\ldots$. Если в момент t_n система пришла в состояние C_i, то будем считать, что $\xi_n = i$.

Определение. Последовательность случайных величин $\{\xi_n\}$, $n = 0,1,\ldots$ образует **цепь Маркова**, если

$$P\{\xi_n = i / \xi_0 = i_0,\ \xi_1 = i_1,\ \ldots,\ \xi_{n-2} = i_{n-2},\ \xi_{n-1} = j\} = P\{\xi_n = i / \xi_{n-1} = j\} = p_{ij}^{(n)}, \quad (66)$$

с начальными условиями

$$P\{\xi_0 = i\} = p_i^{(0)},\ \sum_i p_i^{(0)} = 1. \quad (67)$$

Вероятности $p_{ij}^{(n)}$ - называются вероятностями перехода. Свойство (66), цепи Маркова, называется **свойством отсутствия последействия**, которое интерпретируется так: **поведение процесса в будущем зависит только от фиксированного настоящего и не зависит от его прошлого**.

Определение. Цепь Маркова $\{\xi_n\}$, $n = 0, 1, \ldots$, называется **однородной,** если вероятности перехода $p_{ij}^{(n)}$ не зависят от времени, то есть

$$\forall n,\ p_{ij}^{(n)} = p_{ij}. \quad (68)$$

Определение. Цепь Маркова называется **неприводимой**, если каждое ее состояние может быть достигнуто из любого другого, то есть для любых двух состояний системы S C_i, C_j, существует целое число κ, такое, что $p_{ij}^{(\kappa)} > 0$.

Для однородной цепи имеем $p_{ij} > 0$.

Пусть $\pi_j^{(n)} = P\{\xi_n = j\}$ - вероятность того, что в момент времени t_n система находится в состоянии C_j. Интерес представляет существование предела

$$\pi_j = \lim_{n\to\infty} \pi_j^{(n)}, \ j \in N . \qquad (69)$$

Нахождение распределения $\{\pi_j\}$ является основной задачей цепей Маркова. Если предел существует, то говорят, что система S имеет **стационарный** режим функционирования, если $\forall j \quad \pi_j > 0$. Предельные вероятности $\{\pi_j\}$ не зависят от начальных условий, $\forall j$, π_j, означают долю времени, в течение которого система находится в состоянии C_j, $j \in N$, и однозначно определяются равенствами:

$$\sum_i \pi_i = 1, \qquad (70)$$

$$\pi_j = \sum_i \pi_i \cdot p_{ij}, \ \forall j . \qquad (71)$$

Формула (70) называется условием нормировки.

Система алгебраических уравнений (71) является однородной, и для ее однозначного решения необходимо использовать (70), при этом, любое одно уравнение из системы (71) можно исключить.

Матрица Π, составленная из элементов p_{ij}, называется матрицей вероятностей перехода:

$$\Pi = \left(p_{ij} \right). \qquad (72)$$

Зададим вектор вероятностей состояний системы

$\overline{\pi} = \left(p_1, p_2, ..., p_n, ... \right)$.

тогда система (71) записывается в виде

$$\overline{\pi} = \overline{\pi} \cdot \Pi . \qquad (73)$$

Часто представляют интерес переходы системы из состояния в состояние в произвольный момент времени (переходный режим).

Для этого нужно определить распределение вероятностей $\left\{\pi_j^{(n)}\right\}$ пребывания системы в состоянии C_j в момент t_n. Зададим вектор вероятностей $\overline{\pi}^{(n)}$ в момент t_n равенством

$$\overline{\pi}^{(n)} = \left(p_1^{(n)}, p_2^{(n)}, \ldots, p_n^{(n)}, \ldots\right).$$

Используя (71) и определение вероятностей переходов (66), имеем

$$\overline{\pi}^{(1)} = \overline{\pi}^{(0)} \cdot \Pi ,$$

где $\overline{\pi}^{0} = \left(p_1^0, p_2^0, \ldots\right)$ - начальное состояние системы (67). Отсюда для любого n, по рекуррентной формуле, получаем

$$\overline{\pi}^{(n)} = \overline{\pi}^{(n-1)} \cdot \Pi = \overline{\pi}^{0} \cdot \Pi^n, \quad n \in N . \tag{74}$$

Уравнение (74) дает общий метод вычисления вероятностей на n-м шаге процесса по заданной матрице переходов Π и начальном распределении $\overline{\pi}^{0}$.

Если стационарный режим существует, то

$$\overline{\pi} = \lim_{n\to\infty} \overline{\pi}^{(n)}. \tag{75}$$

Пример. Рассмотрим систему S, которая находится, в любой момент времени t, в одном из трех состояний C_1, C_2, C_3. Переход системы от состояния к состоянию происходит мгновенно в фиксированные моменты времени $t_к = к$, $к \in N$, в соответствии с размеченным графом [3] состояний рисунка 28.

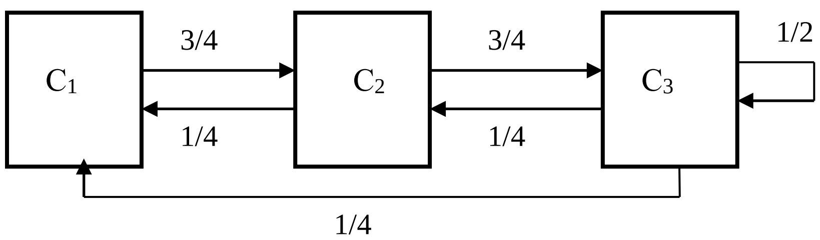


Рис. 28

Требуется оценить скорость сходимости к стационарному режиму и вычислить стационарное распределение вероятностей.

Решение. Вычислим стационарное распределение вероятностей, то есть найдем собственный вектор $\pi = (p_1, p_2, p_3)$, где $p_i = P\{C_i\}$, i=1, 2, 3.

Имеем $\pi = \pi \cdot \Pi$, где

$$\Pi = \begin{pmatrix} 0 & 3/4 & 1/4 \\ 1/4 & 0 & 3/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}.$$

С учетом условия нормировки $p_1 + p_2 + p_3 = 1$ имеем систему

$$\begin{cases} p_1 = 0 \cdot p_1 + (1/4) \cdot p_2 + (1/4) \cdot p_3, \\ p_2 = (3/4) \cdot p_1 + 0 \cdot p_2 + (1/4) \cdot p_3, \\ p_3 = (1/4) \cdot p_1 + (3/4) \cdot p_2 + (1/2) \cdot p_3, \qquad (*) \\ 1 = p_1 + p_2 + p_3. \end{cases}$$

Решая ее (например, без уравнения помеченного (*)), получаем стационарное распределение вероятностей:

$$\bar{\pi} = (0,2;\ 0,28;\ 0,52).$$

Оценим скорость сходимости. Для этого вычислим вероятности перехода $p_{ij}^{(n)}$ по формуле (74) при различных начальных условиях:

а) $p^{(0)}$ = (1,0,0),

результаты представлены в виде табл. 7.

Таблица 7

n	0	1	2	3	4	…	∞
$p_1^{(n)}$	1	0	0,250	0,178	0,203	…	0,2
$p_2^{(n)}$	0	0,75	0,062	0,359	0,254	…	0,28
$p_3^{(n)}$	0	0,25	0,688	0,454	0,543	…	0,52

б) $p^{(0)}$ = (0,1,0),

соответствующие результаты отражены в табл. 8.

Таблица 8

n	0	1	2	3	4	…	∞
$p_1^{(n)}$	0	0	0,187	0,203	0,199	…	0,2
$p_2^{(n)}$	1	0,75	0,375	0,250	0,289	…	0,28
$p_3^{(n)}$	0	0,25	0,438	0,547	0,512	…	0,52

в) $p^{(0)} = (0,0,1)$,

в итоге получаем табл. 9:

Таблица 9

n	0	1	2	3	4	…	∞
$p_1^{(n)}$	0	0	0,187	0,203	0,199	…	0,2
$p_2^{(n)}$	0	0,75	0,313	0,266	0,285	…	0,28
$p_3^{(n)}$	1	0,25	0,500	0,531	0,516	…	0,52

Из таблиц видно, что вхождение системы в стационарный режим происходит достаточно быстро, так как, уже после четырех шагов, вероятности мало отличаются от предельных, независимо от начальных условий.

Замечание. Оценка скорости сходимости переходных вероятностей к стационарным зависит от **собственных** значений матрицы П и иллюстрируется **барицентрической системой координат** [8].

П.13. Непрерывные цепи Маркова

Если система с конечным или счетным числом состояний может переходить из одного состояния в другое в любой момент времени t, то будем говорить, что задана **цепь Маркова** с **непрерывным** временем.

Определение. Случайный процесс $\xi(t)$ образует **непрерывную цепь Маркова**, если для произвольной последовательности $\{t_i\}$, $i = 1, 2, \ldots, n$, такой, что $t_0 < t_1 < \ldots < t_n$, выполняется

$$P\left\{\xi\left(t_n\right)=j/\xi\left(t_0\right)=i_0,\ \ \xi\left(t_1\right)=i_1,\ \ldots,\ \xi\left(t_{n-1}\right)=i_{n-1}\right\}=P\left\{\xi\left(t_n\right)=j/\xi\left(t_{n-1}\right)=i_{n-1}\right\}.$$
(76)

Это определение является непрерывным аналогом определения (66). Интерпретация та же самая: состояние системы *S* в будущем зависит только от

текущего ее состояния (настоящего) и не зависит от того, как и когда система попала в это состояние.

Определение. Цепи Маркова, сформулированные в терминах случайных процессов, называются марковскими процессами.

Рассмотренные здесь цепи Маркова - суть марковские процессы.

Непрерывные марковские процессы отличаются от дискретных тем, что случайные изменения состояний системы зависят от непрерывно изменяющихся параметров.

Пусть задан процесс $\xi(t)$, определяющий состяние ситемы в момент времени $t \in T$. Зададим процесс ее развития: если, в данный момент времени τ ($\tau<t$), система находится в состоянии i, то в последующий момент времени t она будет находится в состоянии j с вероятностью $P_{ij}(\tau,t)$, независимо от поведения системы до момента τ. Для такой системы вероятности

$$P_{ij}(\tau,t)=P\left\{\xi(t)=j/\xi(\tau)=i\right\},\ i,j\in N,$$

называются **переходными вероятностями** марковского процесса.

Определение. Марковский процесс $\xi(t)$ называется **однородным**, если переходные вероятности $P_{ij}(\tau,t)$ зависят только от разности (t-τ), то есть

$$P_{ij}(\tau,t)=P_{ij}(t-\tau).$$

В общем случае вместо переходных вероятностей можно рассматривать соответствующие плотности вероятностей. В качестве примера, можно привести броуновское движение, распространенное на непрерывный случай (Винеровский процесс).

§4. Закон больших чисел и предельные теоремы (2 часа)

П.1. Понятие закона больших чисел

Реальности нашей действительности подтверждают принцип, согласно которому определенного вида действия случайных событий, при весьма широких допущениях, приводят к результату, не являющемуся случайным. Благодаря этому принципу, законы теории вероятностей можно получать из закономерностей, присущих массовым случайным явлениям (событиям).

Некоторые законы можно сформулировать в виде предельных теорем, основывающихся на устойчивости среднего результата случайных факторов (например, оценка вероятности, числовых характеристик случайных величин, их распределений и др.).

Мы рассмотрим две группы предельных теорем.

К первой группе относятся теоремы, в которых, при определенным образом организованных условиях, доказывается сходимость случайных величин или функций от них к некоторым постоянным. Эта группа теорем носит название **закона больших чисел**. Примером может являться теорема Бернулли.

К другой группе предельных теорем относятся теоремы, использующие предельные свойства сумм случайных величин, где пределом последовательности частичных сумм являются не постоянные, а неслучайные функции (например, нормальное распределение или распределение Пуассона). В частности, в **центральных** предельных теоремах формулируются условия, при которых последовательности частичных сумм случайных величин сходятся к **нормальному** распределению

Замечание. Предельные теоремы и приближенные формулы (например, формула Пуассона) справедливы тогда, когда общее число испытаний заранее фиксировано. Если допустить, что, например, при подбрасывании монеты, игрок может закончить игру в выгодный для себя момент, то в целом результат игры не может быть оценен нормальным распределением, поскольку считается, что за достаточно длительное время произойдет любое, пусть и маловероятное, но мыслимое событие, то есть при $n \to \infty$ вероятность любому событию произойти близка к единице.

Докажем некоторые неравенства, которые, может быть, малопригодны на практике, в силу своей общности, но очень важны и эффективны в теоретических исследованиях.

Теорема. Для любой случайной величины ξ, с заданной функцией распределения $F(x)$, имеет место неравенство

$$\forall \varepsilon>0,\ P\{|\xi|\geq\varepsilon\}\leq\frac{M|\xi|^{\kappa}}{\varepsilon^{\kappa}},\ \kappa\in N. \tag{110}$$

Доказательство. Пусть $\rho(x)$ - плотность случайной величины ξ, тогда имеем цепочку неравенств

$$P\{\ |\xi|\geq\varepsilon\}=P\{\ (\xi\leq-\varepsilon)\,\&\,(\xi\geq\varepsilon)\}=\{A=(-\infty,-\varepsilon]\cup[\varepsilon,\infty)\}=P\{\ \xi\in A\}=$$

$$=\int_A\rho(x)\ dx\leq\int_A\frac{|x|^k}{\varepsilon^k}\rho(x)\ dx\leq\frac{1}{\varepsilon^k}\int_{-\infty}^{\infty}|x|^k\cdot\rho(x)\ dx=\frac{M|\xi|^k}{\varepsilon^k}.\blacktriangledown$$

Следствие 1. Пусть случайная величина ξ положительно определена, тогда

$$P\{\xi\geq\varepsilon\}\leq\frac{M\xi}{\varepsilon}.$$

Следствие 2. (Неравенство **Чебышева**). Если дисперсия случайной величины ξ существует, то

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

или

$$P\{|\xi-M\xi|\geq\varepsilon\}\leq\frac{D\xi}{\varepsilon^2}. \tag{111}$$

Вместо (111), часто используют неравенство

$$P\{|\xi-M\xi|<\varepsilon\}\geq1-\frac{D\xi}{\varepsilon^2}. \tag{112}$$

Пример. Оценить вероятность того, что произвольная случайная величина, с конечной дисперсией, отклонится от своего математического ожидания более чем на 3σ.

Решение. Из формулы (111), имеем

$$P\left\{\left|\xi - M\xi\right| \geq 3\sigma\right\} \leq \frac{\sigma^2}{(3\sigma)^2} = 1/9 = 0,(1).$$

Для сравнения, если случайная величина распределена нормально, то

$$P\left\{\left|\xi - M\xi\right| \geq 3\sigma\right\} \leq 0{,}00135,$$

а если имеет распределение Пуассона, то

$$P\left\{\left|\xi - M\xi\right| \geq 3\sigma\right\} \leq 1/9.$$

Полученная оценка для распределения Пуассона

$$V_k = \frac{\lambda^k}{k!}e^{-\lambda},\ k = 0,\ 1,\ \ldots\ ,$$

отражает факт универсальности его применения, в том смысле, что распределение устойчиво к некоторым ослаблениям условий необходимых при его выводе.

П.2. Закон больших чисел

Теорема (Чебышева). Для последовательности независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_n, \ldots$, с дисперсиями ограниченными в совокупности (то есть $\forall n,\ D\xi_n < C$), имеет место асимптотическая оценка*

$$\forall \varepsilon > 0,\ P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}M\xi_i\right| < \varepsilon\right\} \to 1,\ n \to \infty. \qquad (113)$$

Доказательство. Положим $\bar{\xi} = \frac{1}{n}\sum_{i=1}^{n}\xi_i$, тогда $M\bar{\xi} = \frac{1}{n}\sum_{i=1}^{n}M\xi_i$. Для каждого фиксированного *n*, в силу (112), имеем

$$P\left\{\ \left|\bar{\xi} - M\bar{\xi}\right| < \varepsilon\right\} \geq 1 - \frac{D\bar{\xi}}{\varepsilon^2}. \qquad (114)$$

Из свойств дисперсии следует, что

$$D\bar{\xi} = D\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right) = \frac{1}{n^2}D\sum_{i=1}^{n}\xi_i = \frac{1}{n^2}\sum_{i=1}^{n}D\xi_i \leq \frac{nC}{n^2} = \frac{C}{n}.$$

---

* Последовательность $\{\xi_n\}$ сходится по вероятности к ξ, если для всякого $\varepsilon > 0$, $P\{|\xi_n - \xi| > \varepsilon\} \to 0$ при $n \to \infty$. Формула (113) определяется как сходимость по вероятности дополнительной вероятности к событию $\{|\xi_n - \xi| \leq \varepsilon\}$.

Усилив (114), получим

$$P\left\{\left|\bar{\xi}-M\bar{\xi}\right|<\varepsilon\right\}\geq 1-\frac{C}{n}.$$

Переходя к пределу при $n\to\infty$, и, учитывая, что вероятность больше единицы не бывает, получаем требуемое. ▼

Замечание. Теорема Чебышева справедлива и для случайных величин, у которых функции распределения, вообще говоря, различны.

Из теоремы Чебышева можно получить важные частные случаи.

Теорема. (Хинчина). Дана последовательность независимых случайных величин $\xi_1,\xi_2,...,\xi_n,...$ с одним и тем же распределением и ограниченной дисперсией, тогда

$$\forall\varepsilon>0,\ P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i-a\right|<\varepsilon\right\}\to 1 \text{ при } n\to\infty,$$

где

$$a=M\xi_i,\ \forall i.$$

Пример. Дана последовательность независимых случайных величин $\xi_1,\xi_2,...,\xi_n,...$, заданных на вероятностном пространстве $(\Omega,\mathcal{F},P)$ с функцией распределения Коши:

$$\forall i,\ F_{\xi_i}(x)=\frac{1}{2}+\frac{1}{\pi}arctgx,\ i\in N.$$

Можно ли применить к этой последовательности теорему Хинчина?

Решение. По условию теоремы Хинчина, дисперсии случайных величин ограничены. Проверим это условие. Найдем сначала математическое ожидание. Имеем

$$\forall i,\ M\xi_i=\int_{-\infty}^{\infty}xF'_{\xi_i}(x)dx=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x}{1+x^2}dx.$$

Так как математическое ожидание не существует, то не существует и дисперсия, следовательно, теорема Хинчина для этой последовательности неприменима по двум ее условиям.

Замечание. Для теоремы Хинчина, вообще говоря, кроме независимости случайных величин достаточно существование конечного математического ожидания.

Из теоремы Хинчина следует, что, если при многочисленных измерениях некоторой величины, допускаются случайные ошибки, то их среднее арифметическое дает измерение, наиболее близкое к истинному.

Теорема (Бернулли). Пусть μ - число появлений события A в n независимых испытаниях, а p– вероятность появления события A в каждом испытании, тогда

$$\forall \varepsilon > 0,\ P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} \to 1 \text{ при } n \to \infty.$$

Доказательство. Рассмотрим независимые случайные величины ξ_i, где

$$\forall i,\ \xi_i = \begin{cases} 1, \text{если в испытании } i \text{ событие } A \text{ появилось,} \\ 0, \text{в противном случае,} \end{cases}$$

тогда $\mu = \sum\limits_{i=1}^{n} \xi_i$. Далее $\forall i,\ M\xi_i = 1 \cdot p + 0 \cdot q = p,\ D\xi_i = p \cdot q < 1/4$.

Ограниченность дисперсии следует из того, что, взяв производную от выражения $(p \cdot q)'_p = (p(1-p))' = 1 - 2p$,

получаем, что максимальное значение p = 1/2,

тогда $q = 1 - p = 1/2$.

Все условия теоремы Чебышева выполнены. Учитывая, что $\mu = \sum\limits_{i=1}^{n} \xi_i$, получаем

$$\forall \varepsilon > 0,\ P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n} p\right| < \varepsilon\right\} \to 1 \text{ при } n \to \infty. \blacktriangledown$$

Теорема Бернулли утверждает, что чем больше мы будем проводить независимых испытаний, тем точнее будет оценка вероятности события A в среднем.

Теорема (Пуассона). Если для последовательности независимых испытаний вероятность появления события A в испытании κ равна $p_\kappa,\ \kappa \in N$, то

$$\forall \varepsilon > 0\,,\ P\left\{\left|\frac{\mu}{n} - \frac{1}{n}\sum p_i\right| < \varepsilon\right\} \to 1 \text{ при } n \to \infty.$$

Доказательство. Достаточно заметить, что

$$\forall i,\ D\xi_i = p_i \cdot q_i < 1/4,$$

то есть дисперсии ограничены в совокупности. Все остальные условия теоремы Чебышева, очевидно, выполняются. ▼

Из теоремы Пуассона следует, что, если при проведении независимых испытаний, вероятность появления события A меняется незначительно за счет случайных причин, то при достаточном числе испытаний мы получим значение близкое к истинному значению вероятности события A.

Для **произвольной** последовательности случайных величин закон больших чисел может быть сформулирован следующим образом.

Теорема (Маркова). Дана последовательность произвольных случайных величин $\xi_1, \xi_2, ..., \xi_n, \ldots$ и для каждого фиксированного n

$$\frac{D\left(\sum\limits_{i=1}^{n} \xi_i\right)}{n^2} \xrightarrow[n\to\infty]{} 0, \tag{115}$$

тогда

$$\forall \varepsilon > 0\,,\ P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n} M\xi_i\right| < \varepsilon\right\} \to 1 \text{ при } n \to \infty.$$

Условие (115) означает, что для любого конечного n, среди случайных величин $\xi_1, \xi_2, ..., \xi_n$ нет таких, которые существенно влияли бы на их сумму.

Закон больших чисел фактически обосновывает статистическую вероятность [4], устойчивую к ослаблениям условий ее получения, если число испытаний достаточно велико.

П.3. Центральные предельные теоремы

Пусть дана последовательность случайных величин $\xi_1, \xi_2, ..., \xi_n, \ \dots$. Пусть $\zeta_n = \sum_{\kappa=1}^{n} \xi_\kappa$ - частичная сумма последовательности.

В классической постановке центральные предельные теоремы описывают поведение частичных сумм ζ_n, когда *n* достаточно большое.

В большинстве случаев, будем говорить, что к последовательности случайных величин применима центральная предельная теорема [5], если для любых действительных чисел $\alpha < \beta$.

$$P\{\alpha B_n < \zeta_n - A_n < \beta B_n\} \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} exp\left(-\frac{x^2}{2}\right) dx, \qquad (116)$$

где $B_n^2 = \sum_{\kappa=1}^{n} D\xi_\kappa$, а $A_n = \sum_{\kappa=1}^{n} M\xi_\kappa$.

Практически, формула (116) дает оценку вероятности того, что сумма нескольких случайных величин примет значение из заданного интервала. В современных центральных предельных теоремах в качестве предельных распределений рассматриваются распределения, отличные от нормального, обычно это локальные предельные теоремы (распределение Пуассона, Коши распределения, распределение Стьюдента, χ^2 – распределение и др).

Теорема (Линдеберга – Леви). Пусть $\xi_1, \xi_2, ..., \xi_n, \dots$- последовательность независимых случайных величин, заданных на одном вероятностном пространстве. Пусть существуют

$$M\xi_n = a, \text{ и } D\xi_n = \sigma^2, \ \forall n,$$

тогда для любых $\alpha < \beta$ имеет место

$$P\left\{\alpha < \frac{\sum_{\kappa=1}^{n} \xi_\kappa - na}{\sigma\sqrt{n}} < \beta \right\} \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} exp\left(-\frac{x^2}{2}\right) dx.$$

Теорема (Ляпунова). Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ - последовательность независимых случайных величин, у которых существуют конечное математическое ожидание - $M\xi_\kappa$, дисперсия - $D\xi_\kappa$ и третий центральный момент $\mu_3^\kappa = M\left(\xi_\kappa - M\xi_\kappa\right)^3$, κ, $n \in N$.

Если
$$\lim_{n\to\infty} \sum_{k=1}^{n} \left|\mu_3^k\right| \Bigg/ \left(\sum_{k=1}^{n} (\sigma_k)^{3/2} \right) = 0, \tag{117}$$

то закон распределения нормированной суммы

$$\bar{\xi}_n = \frac{\xi_n - A_n}{B_n}$$

сходится по вероятности к нормальному распределению, то есть

$$P\left\{ \bar{\xi}_n < x \right\} \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2} \right) dt .$$

Замечание. Условие (117) означает, что влияние любой из случайных величин ξ_κ последовательности на их сумму несущественно, иначе сходимость определялась бы распределением той случайной величины, у которой дисперсия намного больше, чем у других случайных величин.

Следствие. Если дана последовательность $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ одинаково распределенных случайных величин с конечной дисперсией, то имеет место сходимость равномерно по x

$$P\left\{ \frac{1}{n\sqrt{D\xi}} \sum_{k=1}^{n} \left(\xi_k - M\xi_k \right) < x \right\} \xrightarrow[n\to\infty]{} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2} \right) dt .$$

Пример 1. В n испытаниях Бернулли с вероятностью успеха p в каждом из испытаний, число успехов можно представить в виде суммы

$$\xi = \sum_{i=1}^{n} \xi_i ,$$

где ξ_i равно единице или нулю, в зависимости от того был ли успех в испытании i или нет (см. теорему Бернулли). Так как испытания независимые, то к сумме применимо следствие из центральной предельной теоремы. Поэтому распределение для ξ приблизительно нормально, то есть

$$M\xi = np,\ \sigma = \sqrt{npq}.$$

В сущности, это есть теорема Муавра-Лапласа.

Пример 2. На странице 88 рассмотрен пример со случайными величинами ξ_i, $i = 1,2,3,4$, одинаково и равномерно распределенными на [0,1). При рассмотрении сумм ξ_1, $\xi_1 + \xi_2$, $\xi_1 + \xi_2 + \xi_3$, $\xi_1 + \xi_2 + \xi_3 + \xi_4$, было отмечено, что их плотности тем больше напоминают нормальное, чем больше слагаемых в сумме.

Теперь в силу теоремы Линдеберга-Леви мы можем утверждать, что

$$\rho\left(\sum_{i=1}^{n}\xi_i, x\right)\xrightarrow[n\to\infty]{}\frac{1}{\sqrt{2\pi}}\exp\left(-x^2/2\right),$$

где $\rho\left(\sum\limits_{i=1}^{n}\xi_i, x\right)$ - плотность случайной величины $\sum\limits_{i=1}^{n}\xi_i$.

Задача. Доказать, что

$$\lim_{n\to\infty}\sum_{\kappa=0}^{n}\frac{n^{\kappa}}{\kappa!}e^{-n} = \frac{1}{2}.$$

Решение. Воспользуемся следствием к теореме Ляпунова. Предположим, что сл. в. $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ распределены в соответствии с законом Пуассона, с параметром λ=1. В силу следствия, имеем

$$P\left\{\frac{1}{n\cdot 1}\sum_{\kappa=1}^{n}(\xi_\kappa - 1) < x\right\}\xrightarrow[n\to\infty]{}\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)dt,$$

или

$$P\left\{\sum_{1}^{n}\xi_\kappa < nx + n\right\} = \{\text{при}\, x = 0\} = P\left\{\sum_{1}^{n}\xi_\kappa < n\right\}\xrightarrow[\text{n}\to\infty]{}\Phi(0) = \frac{1}{2},$$

так как

$$\Phi(0) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0}e^{-\frac{t^2}{2}}dt = \frac{1}{2}.$$

Покажем, что

$$P\left\{\sum_{k=1}^{n}\xi_k < n\right\} = \sum_{k=0}^{n}\frac{n^k}{k!}e^{-n} .$$

Предварительно докажем лемму.

Лемма. Пусть даны случайные величины ξ_1, ξ_2, распределенные по закону Пуассона с параметрами λ_1, λ_2, тогда

$$P\{\xi_1 + \xi_2 = \kappa\} = \frac{(\lambda_1 + \lambda_2)^\kappa}{\kappa!}e^{-(\lambda_1+\lambda_2)}.$$

Доказательство. Пусть

$$P\{\xi_1 = k\} = \frac{\lambda_1^k}{k!}\cdot e^{-\lambda_1},\ P\{\xi_2 = k\} = \frac{\lambda_2^k}{k!}\cdot e^{-\lambda_2},$$

тогда $P\{\xi_1 + \xi_2 = \kappa\} = P\{\xi_1 = \kappa - \xi_2\},\ \kappa\in N$. Отсюда

$$P\{\xi_1 + \xi_2 = k\} = \sum_{i=0}^{k}P\{\xi_2 = i\}\cdot P\{\xi_1 = k - i\} = \sum_{i=0}^{\hat{e}}\left(\frac{\lambda_2^i}{i!}e^{-\lambda_2}\cdot\frac{\lambda_1^{k-i}\cdot e^{-\lambda_1}}{(k-i\ )!}\right) =$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{k!}\cdot\sum_{i=0}^{k}\frac{k!\cdot\lambda_2^i\cdot\lambda_1^{k-i}}{i!\cdot(k-i)!} = \frac{e^{-(\lambda_1+\lambda_2)}}{k!}\sum_{i=0}^{k}C_k^i\cdot\lambda_2^i\cdot\lambda_1^{k-i} = \frac{(\lambda_1+\lambda_2)^k}{k!}\cdot e^{-(\lambda_1+\lambda_2)}.\blacktriangledown$$

Далее, по индукции, легко показать, что, если $\xi_1, \xi_2, \ldots, \xi_n$ распределены по закону Пуассона, то

$$P\left\{\sum_{i=1}^{n}\xi_i = r\right\} = \frac{\left(\sum_{i=1}^{n}\lambda_i\right)^r}{r!}e^{-\sum_{i=1}^{n}\lambda_i}.\blacktriangledown$$

Вернемся к задаче.

Рассмотрим событие $\sum_{i=1}^{n}\xi_i < r$. Его можно представить в виде объединения несовместных событий:

$$\sum_{i=1}^{n}\xi_i < r = \bigcup_{j=0}^{n}\left(\sum_{i=1}^{n}\xi_i = j\right),\ \forall r.$$

Следовательно, для $r = n$

$$P\left\{\sum_{i=1}^{n}\xi_i < n\right\} = P\left\{\bigcup_{j=0}^{n-1}\left(\sum_{i=1}^{n}\xi_i = j\right)\right\} = \sum_{i=0}^{n} P\left\{\sum_{i=1}^{n}\xi_i = j\right\} = \sum_{i=0}^{n}\frac{n^i}{i!}e^{-n} \xrightarrow[n\to\infty]{} \frac{1}{2}.$$

Замечание. Лемма представляет собой пример вычисления свертки дискретных неотрицательных целочисленных случайных величин, распределенных по закону Пуассона.

Практически, центральными предельными теоремами, можно пользоваться и тогда, когда имеется сумма небольшого числа случайных величин (n>10).

В частности, широко применяется приближенная замена одних плотностей на другие, более удобные.

Не следует думать, что все центральные предельные теоремы используют нормальное распределение как предельное. Например, в случайных процессах в качестве предельных рассматриваются χ^2-распределение, распределение Пуассона, Коши распределение, гамма распределение и др. Эти распределения относятся к классу безгранично делимых распределений (то есть таких, которые представимы как n–кратная свертка ($n\in N$) , одинаковых распределений вероятностей). Доказано, что они и только они могут быть предельными для сумм независимых случайных величин. Рассматриваются также предельные теоремы перехода от дискретных случайных процессов к непрерывным.

§5. Выборочный метод.
Статистические оценки параметров распределения (2 часа)

П.1. Введение в математическую статистику

Основной задачей математической статистики является разработка методов получения научно обоснованных выводов о массовых явлениях и процессах из данных наблюдений и экспериментов. Эти выводы и заключения относятся не к отдельным испытаниям, из повторения которых складывается данное массовое явление, а представляют собой утверждения об общих вероятностных характеристиках данного процесса, то есть о вероятностях, законах распределения, математических ожиданиях, дисперсиях и т. д. Такое использование фактических данных как раз и является отличительной чертой статистического метода.

Пусть мы располагаем сведениями (обычно довольно ограниченными), например, о числе дефектных изделий в изготовленной в определенных условиях продукции или о результатах испытаний материалов на разрушение и т. п. Собранные нами данные могут представлять непосредственный интерес в смысле информации о качестве той или иной партии продукции. Статистические же проблемы возникают тогда, когда мы на основе той же информации начинаем делать выводы относительно более широкого круга явлений. Так, например, нас может интересовать качество технологического процесса, для чего мы оцениваем вероятность получения в нем дефектного изделия или среднюю долговечность изделия. В этом случае мы рассматриваем собранный материал не ради его самого, а лишь как некую пробную группу или выборку, представляющую только серии из возможных результатов, которые мы могли бы встретить при продолжении наблюдений массового процесса в данной обстановке. Выводы и оценки, основанные на материале наблюдений, отражают случайный состав пробной группы и поэтому считаются приблизительными оценками вероятностного характера. Во многих случаях теория указывает, как наилучшим способом использовать имеющуюся информацию для получения по возможности более точных и надежных характеристик, указывая при этом степень надежности выводов, объясняющуюся ограниченностью запаса сведений.

В математической статистике рассматриваются две основные категории задач: оценивание и статистическая проверка гипотез. Первая задача разделяется на точечное оценивание и интервальное оценивание параметров распределения. Например, может возникнуть необходимость по наблюдениям получить точечные оценки параметров $M\xi$ и $D\xi$. Если мы хотим получить некоторый интервал, с той или иной степенью достоверности содержащий истинное значение параметра, то это задача интервального оценивания.

Вторая задача – проверка гипотез – заключается в том, что мы делаем предположение о распределении вероятностей случайной величины (например, о значении одного или нескольких параметров функции распределения) и решаем, согласуются ли в некотором смысле эти значения параметров с полученными результатами наблюдений.

П.2. Понятие выборочного метода

Пусть нам нужно обследовать количественный признак в партии экземпляров некоторого товара. Проверку партии можно проводить двумя способами:

1) провести сплошной контроль всей партии;

2) провести контроль только части партии.

Первый способ не всегда осуществим, например, из–за большого числа экземпляров в партии, из–за дороговизны проведения операции контроля, из–за того, что контроль связан с разрушением экземпляра (проверка электролампы на долговечность ее работы).

При втором способе множество случайным образом отобранных объектов называется **выборочной совокупностью** или **выборкой**. Все множество объектов, из которого производится выборка, называется **генеральной совокупностью**. Число объектов в выборке называется **объемом выборки**. Обычно будем считать, что объем генеральной совокупности бесконечен.

Выборки разделяются на **повторные** (с возвращением) и **бесповторные** (без возвращения).

Обычно осуществляются бесповторные выборки, но благодаря большому (бесконечному) объему генеральной совокупности ведутся расчеты и делаются выводы, справедливые лишь для повторных выборок.

Выборка должна достаточно полно отражать особенности всех объектов генеральной совокупности, иначе говоря, выборка должна быть **репрезентативной** (представительной).

Выборки различаются по способу отбора.

1. Простой случайный отбор.

Все элементы генеральной совокупности нумеруются и из таблицы случайных чисел берут, например, последовательность любых 30-ти идущих подряд чисел. Элементы с выпавшими номерами и входят в выборку.

2. Типический отбор.

Такой отбор производится в том случае, если генеральную совокупность можно представить в виде объединения подмножеств, объекты которых однородны по какому–то признаку, хотя вся совокупность такой однородности не имеет (партия товара состоит из нескольких групп, произведенных на разных предприятиях). Тогда по каждому подмножеству проводят простой случайный отбор, и в выборку объединяются все полученные объекты.

3. Механический отбор.

Отбирают каждый двадцатый (сотый) экземпляр.

4. Серийный отбор.

В выборку подбираются экземпляры, произведенные на каком–то производстве в определенный промежуток времени.

В дальнейшем под генеральной совокупностью мы будем подразумевать не само множество объектов, а множество значений случайной величины, принимающей числовое значение на каждом из объектов. В действительности генеральной совокупности как множества объектов может и не существовать. Например, имеет смысл говорить о множестве деталей, которые **можно произвести**, используя данный технологический процесс. Используя какие–то известные нам характеристики данного процесса, мы можем оценивать параметры этого несуществующего множества деталей. Размер детали – это случайная величина, значение которой определяется воздействием множества факторов, составляющих технологический процесс. Нас, например, может интересовать вероятность, с которой эта случайная величина принимает значение, принадлежащее некоторому интервалу. На этот вопрос можно ответить, зная закон распределения этой случайной величины, а также ее параметры, такие как $M\xi$ и $D\xi$.

Итак, отвлекаясь от понятия генеральной совокупности как множества объектов, обладающих некоторым признаком, будем рассматривать генеральную совокупность как случайную величину ξ, закон распределения и параметры которой определяются с помощью выборочного метода.

Рассмотрим выборку объема n, представляющую данную генеральную совокупность. Первое выборочное значение x_1 будем рассматривать как реализацию, как одно из возможных значений случайной величины ξ_1, имеющей тот же закон распределения с теми же параметрами, что и случайная величина ξ. Второе выборочное значение x_2 – одно из возможных значений случайной величины ξ_2 с тем же законом распределения, что и случайна величина ξ. То же самое можно сказать о значениях x_3, x_4,..., x_n .

Таким образом на выборку будем смотреть как на совокупность независимых случайных величин ξ_1, ξ_2, ..., ξ_n, распределенных так же, как и случайная величина ξ, представляющая генеральную совокупность. Выборочные значения x_1, x_2, ..., x_n – это значения, которые приняли эти случайные величины в результате 1-го, 2-го, ..., n-го эксперимента.

П.3. Вариационный ряд

Пусть для объектов генеральной совокупности определен некоторый признак или числовая характеристика, которую можно замерить (размер детали, удельное количество нитратов в дыне, шум работы двигателя). Эта характеристика – случайная величина ξ, принимающая на каждом объекте определенное числовое значение. Из выборки объема n получаем значения этой случайной величины в виде ряда из n чисел:

$$x_1, x_2,..., x_n. \qquad (*)$$

Эти числа называются значениями признака.

Среди чисел ряда (*) могут быть одинаковые числа. Если значения признака упорядочить, то есть расположить в порядке возрастания или убывания, написав каждое значение лишь один раз, а затем под каждым значением x_i признака написать число m_i, показывающее сколько раз данное значение встречается в ряду (*):

x_1	x_2	x_3	...	x_k
m_1	m_2	m_3	...	m_k

то получится таблица, называемая **дискретным вариационным рядом**. Число m_i называется частотой i-го значения признака.

Очевидно, что x_i в ряду (*) может не совпадать с x_i в вариационном ряду. Очевидна также справедливость равенства

$$\sum_{i=1}^{k} m_i = n .$$

Если промежуток между наименьшим и наибольшим значениями признака в выборке разбить на несколько интервалов одинаковой длины, каждому интервалу поставить в соответствие число выборочных значений признака, попавших в этот интервал, то получим **интервальный вариационный ряд**. Если признак может принимать любые значения из некоторого промежутка, то есть является непрерывной случайной величиной, приходится выборку представлять именно таким рядом. Если в вариационном интервальном ряду каждый интервал $[\alpha_i; \alpha_{i+1})$ заменить лежащим в его середине числом $(\alpha_i+\alpha_{i+1})/2$, то получим дискретный вариационный ряд. Такая замена вполне естественна, так как, например, при измерении размера детали с точностью до одного миллиметра всем размерам из промежутка [49,5; 50,5), будет соответствовать одно число, равное 50.

П.4. Точечные оценки параметров генеральной совокупности

Во многих случаях мы располагаем информацией о виде закона распределения случайной величины (нормальный, бернуллиевский, равномерный и т. п.), но не знаем параметров этого распределения, таких как $M\xi$, $D\xi$. Для определения этих параметров применяется выборочный метод.

Пусть выборка объема n представлена в виде вариационного ряда. Назовем **выборочной средней** величину

$$\bar{x} = \frac{x_1 m_1 + x_2 m_2 + \ldots + x_k m_k}{n} = x_1 \frac{m_1}{n} + x_2 \frac{m_2}{n} + \ldots + \frac{m_k}{n}$$

Величина $\omega_i = \dfrac{m_i}{n}$ называется **относительной частотой** значения признака x_i. Если значения признака, полученные из выборки не группировать и не представлять в виде вариационного ряда, то для вычисления выборочной средней нужно пользоваться формулой

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \ .$$

Естественно считать величину $\bar{x}$ выборочной оценкой параметра $M\xi$. Выборочная оценка параметра, представляющая собой число, называется **точечной оценкой**.

Выборочную дисперсию

$$\sigma^2 = \sum_{i=1}^{k} (x_i - \bar{x})^2 \omega_i = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

можно считать точечной оценкой дисперсии $D\xi$ генеральной совокупности.

Приведем еще один пример точечной оценки. Пусть каждый объект генеральной совокупности характеризуется двумя количественными признаками x и y. Например, деталь может иметь два размера – длину и ширину. Можно в различных районах измерять концентрацию вредных веществ в воздухе и фиксировать количество легочных заболеваний населения в месяц. Можно через равные промежутки времени сопоставлять доходность акций данной корпорации с каким-либо индексом, характеризующим среднюю доходность всего рынка акций. В этом случае генеральная совокупность представляет собой двумерную случайную величину ξ, η. Эта случайная величина принимает значения x, y на множестве объектов генеральной совокупности. Не зная закона совместного распределения случайных величин ξ и η, мы не можем говорить о наличии или глубине корреляционной связи между ними, однако некоторые выводы можно сделать, используя выборочный метод.

Выборку объема n в этом случае представим в виде таблицы, где i-тый отобранный объект (i= 1,2,...n) представлен парой чисел x_i, y_i :

x_1	x_2	...	x_n
y_1	y_2	...	y_n

Выборочный коэффициент корреляции рассчитывается по формуле

$$r_{xy} = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y}$$

Здесь

$$\overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i \,,\ \sigma_x = \sqrt{\sigma_x^{\ 2}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}\,,$$

$$\sigma_y = \sqrt{\sigma_y^{\ 2}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}\,.$$

Выборочный коэффициент корреляции можно рассматривать как точечную оценку коэффициента корреляции $\rho_{\xi\eta}$, характеризующего генеральную совокупность.

Выборочные параметры $\bar{x}, s_x, r_{xy}$ или любые другие зависят от того, какие объекты генеральной совокупности попали в выборку и различаются от выборки к выборке. Поэтому они сами являются случайными величинами.

Пусть выборочный параметр δ рассматривается как выборочная оценка параметра Δ генеральной совокупности и при этом выполняется равенство

$$M\delta = \Delta.$$

Такая выборочная оценка называется **несмещенной.**

Для доказательства несмещённости некоторых точечных оценок будем рассматривать выборку объема n как систему n независимых случайных величин ξ_1, ξ_2,... ξ_n , каждая из которых имеет тот же закон распределения с теми же параметрами, что и случайная величина ξ, представляющая генеральную совокупность. При таком подходе становятся очевидными равенства: $Mx_i = M\xi_i = M\xi$; $Dx_i = D\xi_i = D\xi$ для всех k = 1,2,...n.

Теперь можно показать, что выборочная средняя $\bar{x}$ есть несмещенная оценка средней генеральной совокупности или , что то же самое, математического ожидания интересующей нас случайной величины ξ:

$$M\bar{x} = M\,\frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n}\left(M\xi_1 + M\xi_2 + \ldots + M\xi_n\right) = \frac{1}{n}nM\xi = M\xi.$$

Выведем формулу для дисперсии выборочной средней:

$$D\bar{x} = D\,\frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n^2}\left(D\xi_1 + D\xi_2 + \ldots + D\xi_{n1}\right) = \frac{1}{n^2}nD\xi = \frac{D\xi}{n}.$$

Найдем теперь, чему равно математическое ожидание выборочной дисперсии σ^2. Сначала преобразуем σ^2 следующим образом:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - M\xi + M\xi - \bar{x}\right)^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(\left(x_i - M\xi\right)^2 - 2\left(x_i - M\xi\right)\left(\bar{x} - M\xi\right) + \left(\bar{x} - M\xi\right)^2\right) =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(x_i - M\xi\right)^2 - \left(\bar{x} - M\xi\right)^2$$

Здесь использовано преобразование:

$$\sum_{i=1}^{n}2\left(x_i - M\xi\right)\left(\bar{x} - M\xi\right) = 2\left(\bar{x} - M\xi\right)\sum_{i=1}^{n}\left(x_i - M\xi\right) =$$

$$= 2\left(\bar{x} - M\xi\right)\left(\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}M\xi\right) = 2\left(\bar{x} - M\xi\right)\left(n\bar{x} - nM\xi\right) = 2n\left(\bar{x} - M\xi\right)^2$$

Теперь, используя полученное выше выражение для величины σ^2, найдем ее математическое ожидание.

$$M\sigma^2 = M\left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - M\xi\right)^2 - \left(\bar{x} - M\xi\right)^2\right) =$$

$$= \frac{1}{n}\sum_{i=1}^{n}M\left(x_i - M\xi\right)^2 - M\left(\bar{x} - M\xi\right)^2 = \frac{1}{n}nD\xi - D\bar{x} =$$

$$= D\xi - \frac{D\xi}{n} = \frac{n-1}{n}D\xi.$$

Так как $M\sigma^2 \neq D\xi$, **выборочная дисперсия не является несмещенной оценкой дисперсии генеральной совокупности**.

Чтобы получить несмещенную оценку дисперсии генеральной совокупности, нужно умножить выборочную дисперсию на $\frac{n}{n-1}$. Тогда получится величина $s^2 = \frac{n}{n-1}\sigma^2$, называемая **исправленной выборочной дисперсией**.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

Пусть имеется ряд несмещенных точечных оценок одного и того же параметра генеральной совокупности. Та оценка, которая имеет наименьшую дисперсию, называется **эффективной**.

Полученная из выборки объема n точечная оценка δ_n параметра Δ генеральной совокупности называется **состоятельной**, если она сходится по вероятности к Δ. Это означает, что для любых положительных чисел ε и γ найдется такое число $n_{\varepsilon\gamma}$, что для всех чисел n, удовлетворяющих неравенству $n > n_{\varepsilon\gamma}$ выполняется условие $P\left(\left|\delta_n - \Delta\right| < \varepsilon\right) > 1 - \gamma$. $\bar{x}$ и s^2 являются несмещёнными, состоятельными и эффективными оценками величин $M\xi$ и $D\xi$.

П.5. Интервальные оценки

Точечные оценки параметров генеральной совокупности могут быть приняты в качестве ориентировочных, первоначальных результатов обработки выборочных данных. Их недостаток заключается в том, что неизвестно, с какой точностью оценивается параметр. Если для выборок большого объема точность обычно бывает достаточной (при условии несмещенности, эффективности и состоятельности оценок), то для выборок небольшого объема вопрос точности оценок становится очень важным.

Введем понятие интервальной оценки неизвестного параметра генеральной совокупности (или случайной величины ξ, определенной на множестве объектов этой генеральной совокупности). Обозначим этот параметр через Δ. По сделанной выборке по определенным правилам найдем числа Δ_1 и Δ_2, так чтобы выполнялось условие:

$$P(\Delta_1 < \Delta < \Delta_2) = P(\Delta \in (\Delta_1; \Delta_2)) = \gamma$$

Числа Δ_1 и Δ_2 называются **доверительными границами**, интервал (Δ_1, Δ_2) — **доверительным интервалом** для параметра Δ. Число γ называется **доверительной вероятностью** или **надежностью** сделанной оценки.

Сначала задается надежность. Обычно ее выбирают равной 0.95, 0.99 или 0.999. Тогда вероятность того, что интересующий нас параметр попал в интервал (Δ_1, Δ_2) достаточно высока. Число $(\Delta_1 + \Delta_2) / 2$ – середина доверительного интервала – будет давать значение параметра Δ с **точностью** $(\Delta_2 - \Delta_1) / 2$, которая представляет собой половину длины доверительного интервала.

Границы Δ_1 и Δ_2 определяются из выборочных данных и являются функциями от случайных величин $x_1, x_2, \ldots, x_n$, а следовательно – сами случайные величины. Отсюда – доверительный интервал (Δ_1, Δ_2) тоже случаен. Он может покрывать параметр Δ или нет. Именно в таком смысле нужно понимать случайное событие, заключающееся в том, что доверительный интервал покрывает число Δ.

Доверительный интервал для математического ожидания нормального распределения при известной дисперсии

Пусть случайная величина ξ (можно говорить о генеральной совокупности) распределена по нормальному закону, для которого известна дисперсия $D\xi = \sigma^2$ $(\sigma > 0)$. Из генеральной совокупности (на множестве объектов которой определена случайная величина) делается выборка объема n. Выборка $x_1, x_2, \ldots, x_n$ рассматривается как совокупность n независимых случайных величин, распределенных так же как ξ (подход, которому дано объяснение выше по тексту).

Ранее также обсуждались и доказаны следующие равенства:

$Mx_1 = Mx_2 = \ldots = Mx_n = M\xi$;

$Dx_1 = Dx_2 = ... = Dx_n = D\xi$;

$M\,\bar{x} = M\xi$;

$D\,\bar{x} = D\xi/n$;

Достаточно просто доказать (мы доказательство опускаем), что случайная величина $\bar{x}$ в данном случае также распределена по нормальному закону.

Обозначим неизвестную величину $M\xi$ через a и подберем по заданной надежности γ число $d > 0$ так, чтобы выполнялось условие:

$$P(|\bar{x} - a| < d) = \gamma \qquad (1)$$

Так как случайная величина $\bar{x}$ распределена по нормальному закону с математическим ожиданием $M\bar{x} = M\xi = a$ и дисперсией $D\bar{x} = D\xi/n = \sigma^2/n$, получаем:

$$P(|\bar{x} - a| < d) = P(a - d < \bar{x} < a + d) =$$

$$= \Phi\left(\frac{a + d - a}{\frac{\sigma}{\sqrt{n}}}\right) - \Phi\left(\frac{a - d - a}{\frac{\sigma}{\sqrt{n}}}\right) = 2\Phi\left(\frac{d\sqrt{n}}{\sigma}\right)$$

Осталось подобрать d таким, чтобы выполнялось равенство $2\Phi\left(\frac{d\sqrt{n}}{\sigma}\right) = \gamma$ или $\Phi\left(\frac{d\sqrt{n}}{\sigma}\right) = \frac{\gamma}{2}$.

Для любого $\gamma \in [0;1]$ можно по таблице найти такое число t, что $\Phi(t) = \gamma/2$. Это число t иногда называют **квантилем**.

Теперь из равенства

$$\frac{d\sqrt{n}}{\sigma} = t$$

определим значение d: $d = \frac{\sigma t}{\sqrt{n}}$.

Окончательный результат получим, представив формулу (1) в виде:

$$P\left(\bar{x} - \frac{\sigma t}{\sqrt{n}} < a < \bar{x} + \frac{\sigma t}{\sqrt{n}}\right) = \gamma .$$

Смысл последней формулы состоит в следующем: с надежностью γ доверительный интервал

$$\left(\bar{x} - \frac{\sigma t}{\sqrt{n}}; \bar{x} + \frac{\sigma t}{\sqrt{n}}\right)$$

покрывает неизвестный параметр $a = M\xi$ генеральной совокупности. Можно сказать иначе: точечная оценка $\bar{x}$ определяет значение параметра $M\xi$ с точностью $d = \sigma t / \sqrt{n}$ и надежностью γ.

Задача. Пусть имеется генеральная совокупность с некоторой характеристикой, распределенной по нормальному закону с дисперсией, равной 6,25. Произведена выборка объема $n = 27$ и получено средневыборочное значение характеристики $\bar{x} = 12$. Найти доверительный интервал, покрывающий неизвестное математическое ожидание исследуемой характеристики генеральной совокупности с надежностью $\gamma = 0{,}99$.

Решение. Сначала по таблице для функции Лапласа найдем значение t из равенства $\Phi(t) = \gamma / 2 = 0{,}495$. По полученному значению $t = 2{,}58$ определим точность оценки (или половину длины доверительного интервала) d: $d = 2{,}5 \times 2{,}58 / \sqrt{27} \approx 1{,}24$. Отсюда получаем искомый доверительный интервал: (10,76; 13,24).

Доверительный интервал для математического ожидания нормального распределения при неизвестной дисперсии

Пусть ξ – случайная величина, распределенная по нормальному закону с неизвестным математическим ожиданием $M\xi$, которое обозначим буквой a. Произведем выборку объема n. Определим среднюю выборочную $\bar{x}$ и исправленную выборочную дисперсию s^2 по известным формулам.

Случайная величина

$$t = \frac{(\bar{x} - a)\sqrt{n}}{s}$$

распределена по закону Стьюдента с $n - 1$ степенями свободы.

Задача заключается в том, чтобы по заданной надежности γ и по числу степеней свободы n – 1 найти такое число t_γ, чтобы выполнялось равенство

$$P\left(\left|\frac{(\bar{x}-a)\sqrt{n}}{s}\right| < t_\gamma\right) = \gamma \qquad (2)$$

или эквивалентное равенство

$$P\left(\bar{x} - t_\gamma \frac{s}{\sqrt{n}} < a < \bar{x} + t_\gamma \frac{s}{\sqrt{n}}\right) = \gamma\ . \qquad (3)$$

Здесь в скобках написано условие того, что значение неизвестного параметра a принадлежит некоторому промежутку, который и является доверительным интервалом. Его границы зависят от надежности γ , а также от параметров выборки $\bar{x}$ и s.

Чтобы определить значение t_γ по величине γ, равенство (2) преобразуем к виду:

$$P\left(\left|\frac{(\bar{x}-a)\sqrt{n}}{s}\right| \geq t_\gamma\right) = 1-\gamma$$

Теперь по таблице для случайной величины t, распределенной по закону Стьюдента, по вероятности 1 – γ и числу степеней свободы n – 1 находим t_γ . Формула (3) дает ответ поставленной задачи.

Задача. На контрольных испытаниях 20-ти электроламп средняя продолжительность их работы оказалась равной 2000 часов при среднем квадратическом отклонении (рассчитанном как корень квадратный из исправленной выборочной дисперсии), равном 11-ти часам. Известно, что продолжительность работы лампы является нормально распределенной случайной величиной. Определить с надежностью 0,95 доверительный интервал для математического ожидания этой случайной величины.

Решение. Величина 1 – γ в данном случае равна 0,05. По таблице распределения Стьюдента, при числе степеней свободы, равном 19, находим: t_γ = 2,093. Вычислим теперь точность оценки: 2,093×121/$\sqrt{20}$ = 56,6. Отсюда получаем искомый доверительный интервал:

(1943,4; 2056,6).

Доверительный интервал для дисперсии нормального распределения

Пусть случайная величина ξ распределена по нормальному закону, для которого дисперсия $D\xi$ неизвестна. Делается выборка объема n . Из нее определяется исправленная выборочная дисперсия s^2. Случайная величина

$$\chi^2 = \frac{(n-1)s^2}{D\xi}$$

распределена по закону χ^2 с $n-1$ степенями свободы. По заданной надежности γ можно найти сколько угодно границ χ_1^2 и χ_2^2 интервалов, таких, что

$$\text{Ð}\left(\chi_1^2 < \chi^2 < \chi_2^2\right) = \gamma \qquad (*)$$

Найдем χ_1^2 и χ_2^2 из следующих условий:

$$P(\chi^2 \le \chi_1^2) = (1-\gamma)/\,2 \qquad (**)$$

$$P(\chi^2 \ge \chi_2^2) = (1-\gamma)/\,2 \qquad (***)$$

Очевидно, что при выполнении двух последних условий справедливо равенство (*).

В таблицах для случайной величины χ^2 обычно дается решение уравнения $P(\chi^2 \ge \chi_q^2) = q$. Из такой таблицы по заданной величине q и по числу степеней свободы $n-1$ можно определить значение χ_q^2. Таким образом, сразу находится значение χ_2^2 в формуле (***).

Для определения χ_1^2 преобразуем (**):

$$P(\chi^2 \ge \chi_1^2) = 1 - (1-\gamma)/\,2 = (1+\gamma)/\,2$$

Полученное равенство позволяет определить по таблице значение χ_1^2.

Теперь, когда найдены значения χ_1^2 и χ_2^2, представим равенство (*) в виде

$$P\left(\chi_1^2 < \frac{(n-1)s^2}{D\xi} < \chi_2^2\right) = \gamma\,.$$

Последнее равенство перепишем в такой форме, чтобы были определены границы доверительного интервала для неизвестной величины $D\xi$:

$$P\left(\frac{(n-1)s^2}{\chi_2^{\ 2}} < D\xi < \frac{(n-1)s^2}{\chi_1^{\ 2}}\right) = \gamma .$$

Отсюда легко получить формулу, по которой находится доверительный интервал для стандартного отклонения:

$$P\left(\frac{\sqrt{(n-1)}\,s}{\sqrt{\chi_2^{\ 2}}} < \sqrt{D\xi} < \frac{\sqrt{(n-1)}s}{\sqrt{\chi_1^{\ 2}}}\right) = \gamma \qquad (****)$$

Задача. Будем считать, что шум в кабинах вертолетов одного и того же типа при работающих в определенном режиме двигателях — случайная величина, распределенная по нормальному закону. Было случайным образом выбрано 20 вертолетов, и произведены замеры уровня шума (в децибелах) в каждом из них. Исправленная выборочная дисперсия измерений оказалась равной 22,5. Найти доверительный интервал, накрывающий неизвестное стандартное отклонение величины шума в кабинах вертолетов данного типа с надежностью 98%.

Решение. По числу степеней свободы, равному 19, и по вероятности (1 – 0,98)/2 = 0,01 находим из таблицы распределения χ^2 величину χ_2^2 = 36,2. Аналогичным образом при вероятности (1 + 0,98)/2 = 0,99 получаем χ_1^2 = 7,63. Используя формулу (****), получаем искомый доверительный интервал: (3,44; 7,49).

§6. Проверка статистических гипотез (2 часа)

П.1. Задачи статистической проверки гипотез

Статистическая проверка гипотез является вторым после статистического оценивания параметров распределения и в то же время важнейшим разделом математической статистики.

Методы математической статистики позволяют проверить предположения о законе распределения некоторой случайной величины (генеральной совокупности), о значениях параметров этого закона (например $M\xi$, $D\xi$), о наличии корреляционной зависимости между случайными величинами, определенными на множестве объектов одной и той же генеральной совокупности.

Пусть по некоторым данным имеются основания выдвинуть предположения о законе распределения или о параметре закона распределения случайной величины (или генеральной совокупности, на множестве объектов которой определена эта случайная величина). Задача заключается в том, чтобы подтвердить или опровергнуть это предположение, используя выборочные (экспериментальные) данные.

Гипотезы о значениях параметров распределения или о сравнительной величине параметров двух распределений называются **параметрическими гипотезами**.

Гипотезы о виде распределения называются **непараметрическими гипотезами**.

Проверить статистическую гипотезу – это значит проверить, согласуются ли данные, полученные из выборки с этой гипотезой. Проверка осуществляется с помощью **статистического критерия**. **Статистический критерий – это случайная величина, закон распределения которой (вместе со значениями параметров) известен в случае, если принятая гипотеза справедлива**. Этот критерий называют еще **критерием согласия** (имеется в виду согласие принятой гипотезы с результатами, полученными из выборки).

Гипотезу, выдвинутую для проверки ее согласия с выборочными данными, называют **нулевой гипотезой** и обозначают H_0. Вместе с гипотезой H_0 выдвигается **альтернативная** или **конкурирующая** гипотеза, которая обозначается H_1. Например:

1) H_0: $M\xi = 0$ 2) H_0: $M\xi = 0$ 3) H_0: $M\xi = 0$

H_1: $M\xi \neq 0$ H_1: $M\xi > 0$ H_1: $M\xi = 2$

Пусть случайная величина K – статистический критерий проверки некоторой гипотезы H_0. При справедливости гипотезы H_0 закон распределения случайной величины K характеризуется некоторой известной нам плотностью распределения $p_K(x)$.

Выберем некоторую малую вероятность α, равную 0,05 , 0,01 или еще меньшую. Определим **критическое значение критерия** $K_{кр}$ как решение одного из трех уравнений, в зависимости от вида нулевой и конкурирующей гипотез:

$$P(K > K_{кр}) = \alpha \qquad (1)$$

$$P(K < K_{кр}) = \alpha \qquad (2)$$

$$P((K < K_{кр1}) \cap (K > K_{кр2})) = \alpha \qquad (3)$$

Возможны и другие уравнения, но они встречаются значительно реже, чем приведенные.

Решение уравнения (1) (то же самое для уравнений (2) и (3)) заключается в следующем: по вероятности α, зная функцию $p_K(x)$, заданную как правило таблицей, нужно определить $K_{кр}$.

Что означает условие (1)?

Если гипотеза H_0 справедлива, то вероятность того, что критерий K превзойдет некоторое значение $K_{кр}$ очень мала – 0,05 , 0,01 или еще меньше, в зависимости от нашего выбора. Если $K_в$ – значение критерия K, рассчитанное по выборочным данным, превзошло значение $K_{кр}$, это означает, что выборочные данные не дают основания для принятия нулевой гипотезы H_0 (например, если α=0,01 , то можно сказать, что произошло событие, которое при справедливости гипотезы H_0 встречается в среднем не чаще, чем в одной из ста выборок). В этом случае говорят, что **гипотеза H_0 не согласуется с выборочными данными и должна быть отвергнута**. Если $K_в$ не превосходит $K_{кр}$, то говорят, что **выборочные данные не противоречат гипотезе H_0**, и нет оснований отвергать эту гипотезу.

Для уравнения (1) область $K > K_{кр}$ называется **критической** областью. **Если значение $K_в$ попадает в критическую область, то гипотеза H_0 отвергается**.

Для уравнения (1) область $K < K_{кр}$ называется **областью принятия гипотезы**. **Если значение $K_в$ попадает в область принятия гипотезы, то гипотеза H_0 принимается**.

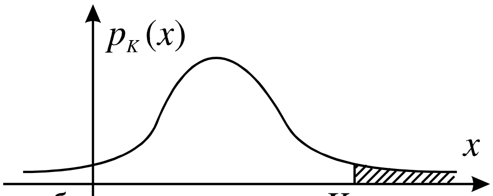


Рис. 1.

Рисунок 1. иллюстрирует решение уравнения (1). Здесь $p_K(x)$ – известная плотность распределения случайной величины K при условии справедливости гипотезы H_0.

Пусть выбрано некоторое малое значение вероятности α, по нему определено значение $K_{кр}$ и по выборочным данным определено значение $K_в$, которое попало в критическую область. В этом случае гипотеза H_0 отвергается, но она может оказаться справедливой, просто случайно произошло событие, которое имеет очень малую вероятность α. В этом смысле α есть вероятность отвержения правильной гипотезы H_0.

Отвержение правильной гипотезы называется **ошибкой первого рода**. Вероятность α называется уровнем значимости. Таким образом **уровень значимости – это вероятность совершения ошибки первого рода**.

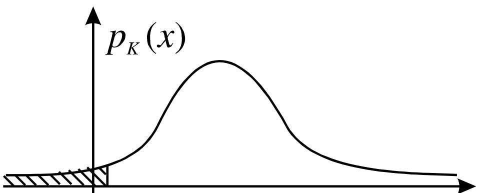


Рис. 2.

Критическая область, полученная для уравнения (1) и приведенная на рисунке 1., называется **правосторонней**.

Уравнение (2) определяет **левосторонюю критическую область**. Ее изображение приводится на рисунке 2.

Отметим, что каждая из заштрихованных фигур на рисунках 1. и 2. имеет площадь, равную α.

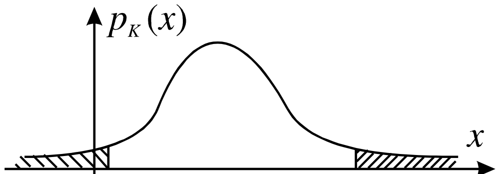


Рис. 3.

Уравнение (3) определяет **двусторонюю критическую область**. Такая область изображена на рисунке 3. Здесь критическая область состоит из двух частей. В случае двусторонней критической области границы ее частей $K_{кр1}$ и $K_{кр2}$ определяются таким образом, чтобы выполнялось условие:

$$P(K \le K_{кр}) = P(K \ge K_{кр}) = \alpha / 2.$$

На рисунке 3. площадь каждой из заштрихованных фигур равна $\alpha / 2$.

Вид критической области зависит от того, какая гипотеза выдвинута в качестве конкурирующей.

Чем меньше уровень значимости, тем меньше вероятность отвергнуть проверяемую гипотезу H_0, когда она верна, то есть совершить ошибку первого рода. Но с уменьшением уровня значимости расширяется область принятия гипотезы H_0 и увеличивается вероятность принятия проверяемой гипотезы, когда она неверна, то есть когда предпочтение должно быть отдано конкурирующей гипотезе.

Пусть при справедливости гипотезы H_0 статистический критерий K имеет плотность распределения $p_0(x)$, а при справедливости конкурирующей гипотезы H_1 – плотность распределения $p_1(x)$. Графики этих функций приведены на рисунке 4. При некотором уровне значимости находится критическое значение $K_{кр}$ и правостороняя критическая область. Если значение $K_в$, определенное по выборочным данным, оказывается меньше, чем $K_{кр}$, то гипотеза H_0 принимается. Предположим, что справедлива на самом деле конкурирующая гипотеза H_1. Тогда вероятность попадания критерия в область принятия гипотезы H_0 есть некоторое число β, равное площади фигуры, образованной графиком функции $p_1(x)$ и полубесконечной частью горизонтальной координатной оси, лежащей слева от точки $K_{кр}$. Очевидно, что β – это вероятность того, что будет принята неверная гипотеза H_0.

Принятие неверной гипотезы называется ошибкой второго рода. В рассмотренном случае число β – это вероятность ошибки второго рода. **Число** 1 – β, **равное вероятности того, что не совершается ошибка второго рода, называется мощностью критерия**. На рисунке 4 мощность критерия равна площади фигуры, образованной графиком функции $p_1(x)$.и полубесконечной частью горизонтальной координатной оси, лежащей справа от точки $K_{кр}$.

Выбор статистического критерия и вида критической области осуществляется таким образом, чтобы мощность критерия была максимальной.

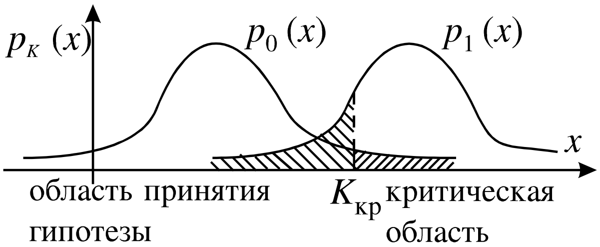


Рис. 4.

П.2. Проверка статистической гипотезы о математическом ожидании нормального распределения при известной дисперсии

Пусть имеется нормально распределенная случайная величина ξ,, определенная на множестве объектов некоторой генеральной совокупности. Известно, что $D\xi = \sigma^2$. Математическое ожидание $M\xi$ неизвестно. Допустим, что имеются основания предполагать, что $M\xi = a$, где a – некоторое число (такими основаниями могут быть ограниченные сведения об объектах генеральной совокупности, опыт исследования подобных совокупностей и т. д.). Будем считать также, что имеется другая информация, указывающая на то, что $M\xi = a_1$, где $a_1 > a$.

I. Выдвигаем нулевую гипотезу H_0: $M\xi = a$;

при конкурирующей гипотезе H_1: $M\xi = a_1$.

Делаем выборку объема n: $x_1, x_2,..., x_n$. В основе проверки лежит тот факт, что случайная величина $\bar{x}$ (выборочная средняя) распределена по нормальному закону с дисперсией σ^2/n и математическим ожиданием, равным a в случае справедливости H_0, и равным a_1 в случае справедливости H_1.

Очевидно, что если величина $\bar{x}$ оказывается достаточно малой, то это дает основание предпочесть гипотезу H_0 гипотезе H_1. При достаточно большом значении $\bar{x}$ более вероятна справедливость гипотезы H_1. Задачу можно было бы поставить так: требуется найти некоторое критическое число, которое разбивало бы все возможные значения выборочной средней (в условиях данной задачи это все действительные числа) на два полубесконечных промежутка. При попадании $\bar{x}$ в левый промежуток следовало бы принимать гипотезу H_0, а при попадании $\bar{x}$ в правый промежуток предпочтение следовало бы оказать гипотезе H_1. Однако на самом деле поступают несколько иначе.

В качестве статистического критерия выбирается случайная величина

$$z = \frac{(\bar{x} - a)\sqrt{n}}{\sigma},$$

распределенная по нормальному закону , причем $Mz = 0$ и $Dz = 1$ (это следует из свойств математического ожидания и дисперсии) в случае справедливости гипотезы H_0. Если справедлива гипотеза H_1, то $Mz = a^* = (a_1 - a)\sqrt{n}/\sigma$, $Dz = 1$.

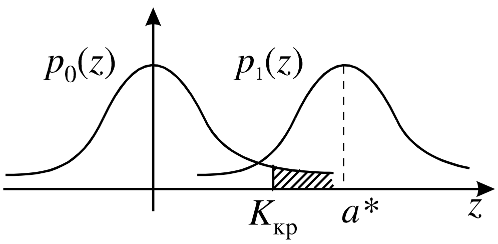


Рис.1.

На рисунке 1. изображены графики $p_0(z)$ и $p_1(z)$ – функций плотности распределения случайной величины z при справедливости гипотез H_0 и H_1, соответственно.

Если величина $\bar{x}$, полученная из выборочных данных, относительно велика, то и величина z велика, что является свидетельством в пользу гипотезы H_1. Относительно малые значения $\bar{x}$ приводят к малым значениям z, что свидетельствует в пользу гипотезы H_0. Отсюда следует, что должна быть выбрана правосторонняя критическая область. По принятому уровню значимости α (например $\alpha = 0{,}05$), используя то, что случайная величина z распределена по нормальному закону, определим значение $K_{кр}$ из формулы

$$\alpha = P(K_{кр} < z < \infty) = \Phi(\infty) - \Phi(K_{кр}) = 0{,}5 - \Phi(K_{кр}).$$

Отсюда $\Phi(K_{êð}) = \dfrac{1 - 2\alpha}{2}$, и осталось воспользоваться таблицей функции Лапласа для нахождения числа $K_{кр}$.

Если величина z, полученная при выборочном значении $\bar{x}$, попадает в область принятия гипотезы ($z < K_{кр}$), то гипотеза H_0 принимается (делается вывод, что выборочные данные не противоречат гипотезе H_0). Если величина z попадает в критическую область, то гипотеза H_0 отвергается.

В данной задаче может быть подсчитана мощность критерия:

$$1-\beta = \Phi(\infty) - \Phi\left(K_{\hat{e}ð} - \frac{(a_1 - a)}{\sigma}\sqrt{n}\right)$$

Мощность критерия тем больше, чем больше разность $a_1 - a$.

II. Если в предыдущей задаче поставить другое условие:

H_0: $M\xi = a$;

H_1: $M\xi = a_1$, $a_1 < a$,

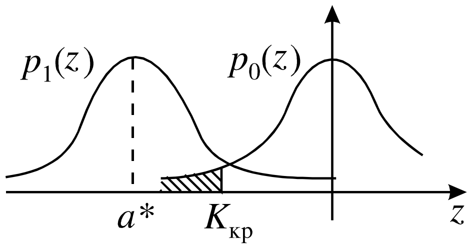


Рис. 2.

то сохранив смысл всех рассуждений, здесь придется рассматривать левостороннюю критическую область, как изображено на рисунке 2. Здесь, как и в предыдущем случае, $a^* = (a_1 - a)\sqrt{n}/\sigma$, а величина $K_{кр}$ определяется из формулы

$$\alpha = P(-\infty < z < K_{кр}) = \Phi(K_{кр}) - \Phi(-\infty) = \Phi(K_{кр}) + \frac{1}{2}.$$

Используя формулу $-\Phi(K_{кр}) = \Phi(-K_{кр})$, получаем:

$$\Phi(-K_{кр}) = \frac{1-2\alpha}{2}.$$

Отметим, что по смыслу задачи здесь $K_{кр}$ – отрицательное число.

Значения z, вычисленные по выборочным данным, превышающие $K_{кр}$, согласуются с гипотезой H_0. Если величина z попадает в критическую область ($z < K_{кр}$), то гипотезу H_0 следует отвергнуть, считая предпочтительной гипотезу H_1.

III. Рассмотрим теперь такую задачу:

H_0: $M\xi = a$;

H_1: $M\xi \neq a$.

В данном случае большие отклонения величины z от нуля в положительную или отрицательную сторону должны приводить к заключению о ложности гипотезы H_0, то есть здесь следует рассматривать двустороннюю

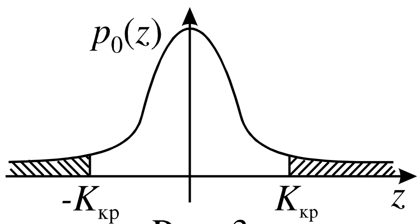


Рис. 3.

критическую область, как изображено на рисунке 3.

Критическое значение $K_{кр}$ определяется с помощью соотношения

$$P(-K_{кр} < z < K_{кр}) = 1 - \alpha = \Phi(K_{кр}) - \Phi(-K_{кр}) = 2\Phi(K_{кр}).$$

Из этого соотношения следует:

$$\Phi(K_{кр}) = \frac{1-\alpha}{2}.$$

П.3. Проверка гипотезы о равенстве дисперсий

Гипотезы о дисперсии играют очень важную роль в экономико–математическом моделировании, так как величина рассеяния экспериментальных выборочных данных относительно рассчитанных теоретических значений соответствующих параметров, характеризующаяся дисперсией, дает возможность судить о пригодности (адекватности) теории или модели, на основании которой строится теория.

Пусть нормально распределенная случайная величина ξ определена на некотором множестве, образующем генеральную совокупность, а нормально распределенная случайная величина η определена на другом множестве, которое тоже составляет генеральную совокупность. Из обеих совокупностей делаются выборки: из первой – объема n_1, а из второй – объема n_2 (отметим, что объем выборки не всегда можно определить заранее, как например в случае, если он равен количеству рыб, попавших в сеть). По каждой выборке рассчитывается исправленная выборочная дисперсия: s_1^2 для выборки из первой совокупности и s_2^2 для выборки из второй совокупности.

Поставим задачу: с помощью выборочных данных проверить статистическую гипотезу H_0: $D\xi = D\eta$. В качестве конкурирующей гипотезы будем рассматривать идею, заключающуюся в том, что дисперсия той совокупности, для которой исправленная выборочная дисперсия оказалась наибольшей, больше дисперсии другой совокупности. Критерий берется в следующем виде:

$$F = \frac{S^{**}}{S^{*}}.$$

Здесь S^{**}– наибольшая из двух оценок s_1^2 и s_2^2, а S^*– наименьшая из тех же двух оценок.

Критерий F распределен по закону Фишера с k_1 и k_2 степенями свободы. Здесь

$$k_1 = n_1 - 1,\ k_2 = n_2 - 1, \text{ если } S^{**} = s_1^2;$$

$$k_1 = n_2 - 1,\ k_2 = n_1 - 1, \text{ если } S^{**} = s_2^2.$$

В этой задаче естественно рассматривать правостороннюю критическую область, так как достаточно большие выборочные значения критерия F свидетельствуют в пользу конкурирующей гипотезы.

При заданном уровне значимости q (обычно q =0,05 или q =0,01) критическое значение $F_{кр}$ определяется из таблицы распределения Фишера. В случае $F > F_{кр}$ гипотеза H_0 отвергается, а в случае $F < F_{кр}$ – принимается.

Пусть два множества некоторых объектов, обладающих количественным признаком, подвергнуты выборочному контролю. Значения количественного признака есть распределенные по нормальному закону случайные величины, которые мы обозначим ξ_1 и ξ_2, соответственно, для первого и для второго множеств. Из первого множества сделана выборка объема n_1=21 и подсчитана исправленная выборочная дисперсия, оказавшаяся равной 0,75. Из второго множества сделана выборка объема n_2=11. Эта выборка дала значение исправленной выборочной дисперсии, равное 0,25. Выдвигаем гипотезу H_0: $D\xi_1 = D\xi_2$. Конкурирующая гипотеза H_1 заключается в том, что $D\xi_1 > D\xi_2$. В данном случае выборочное значение $F_в$ критерия Фишера равно 3. При выбранном уровне значимости $q = 0,05$ по числам степеней свободы k_1=20, k_2=10 находим по таблице распределения Фишера $F_{кр}$=2,77. Так как $F_в > F_{кр}$, гипотеза о равенстве дисперсий должна быть отвергнута.

§7. Элементы корреляционного и регрессионного анализа (2 часа)

П.1. Понятие регрессии и корреляции

Диалектический подход к изучению природы и общества требует рассмотрения явлений в их взаимосвязи и непрестанном изменении.

Понятия *корреляции* и *регрессии* появились в середине XIX в. благодаря работам английских статистиков Ф. Гальтона и К. Пирсона. Первый термин произошел от латинского «correlatio» — соотношение, взаимосвязь. Второй термин (от лат. «regressio» — движение назад) введен Ф. Гальтоном, который, изучая зависимость между ростом родителей и их детей, обнаружил явление «регрессии к среднему» — у детей, родившихся у очень высоких родителей, рост имел тенденцию быть ближе к средней величине.

В естественных науках часто речь идет о функциональной зависимости (связи), когда каждому значению одной переменной соответствует вполне *определенное значение другой* (например, скорость свободного падения тела в вакууме в зависимости от времени и т.п.).

В экономике в большинстве случаев между переменными величинами существуют зависимости, когда каждому значению одной переменной соответствует не какое-то определенное, а **множество** возможных значений другой переменной. Иначе говоря, каждому значению одной переменной соответствует *определенное (условное) распределение другой переменной.* Такая зависимость получила название ***статистической*** (или с***тохастической, вероятностной***).

Возникновение понятия статистической связи обусловливается тем, что зависимая переменная подвержена влиянию ряда неконтролируемых или неучтенных факторов, а также тем, что измерение значений переменных неизбежно сопровождается некоторыми случайными ошибками. Примером статистической связи является зависимость урожайности от количества внесенных удобрений, производительности труда на предприятии от его энерговооруженности и т.п.

В силу неоднозначности статистической зависимости между Y и X для исследователя, в частности, представляет интерес усредненная по *x* схема зависимости, т.е. закономерность в изменении условного математического ожидания $M_X(Y)$ (математического ожидания случайной переменной Y, вычисленного в предположении, что переменная X приняла значение x в зависимости от x.

Определение. *Корреляционной зависимостью между двумя переменными величинами называется функциональная зависимость между значениями одной из них и условным математическим ожиданием другой.*

Корреляционная зависимость может быть представлена в виде:

$$M_x(Y)=\varphi(x) \quad (1) \quad \text{или} \quad M_Y(X)=\varphi(y) \qquad (2)$$

Уравнения (1) и (2) называются *модельными уравнениями регрессии* (или просто уравнениями *регрессии*) соответственно Y по X и X по Y, функции φ(x) и ψ(y) - *модельными функциями регрессии* (или *функциями регрессии*), а их графики — *модельными линиями регрессии* (или *линиями регрессии*).

Для отыскания модельных уравнений регрессии, вообще говоря, необходимо знать *закон распределения двумерной случайной величины (X,Y)*. На практике исследователь, как правило, располагает лишь *выборкой* пар значений (x_i, y_i) ограниченного объема. В этом случае речь может идти об оценке (приближенном выражении) по выборке функции регрессии. Такой наилучшей (в смысле метода наименьших квадратов) оценкой является *выборочная линия (кривая) регрессии Y по X:*

$$y_x = \hat{\varphi}\left(x, b_0, b_1, \ldots, b_p\right), \qquad (3)$$

где y_x — *условная (групповая) средняя* переменной Y при фиксированном значении переменной X= x; $b_0, b_1 \ldots b_p$ — параметры кривой.

Аналогично определяется *выборочная линия (кривая) регрессии* X по Y:

$$x_y = \hat{\psi}\left(y, c_0, c_1, \ldots, c_p\right), \qquad (4)$$

где x_y — *условная (групповая) средняя* переменной *X* при фиксированном значении переменной Y = y; $c_0, c_1, \ldots, c_p$ — параметры кривой.

Уравнения (3), (4) называют также выборочными уравнениями регрессии соответственно Y по X и X по Y.

Статистические связи между переменными можно изучать методами корреляционного и регрессионного анализа.

Основной задачей регрессионного анализа является установление

формы и изучение зависимости между переменными. Основной задачей корреляционного анализа — выявление связи между случайными переменными и оценка ее тесноты.

П.2. Линейная парная регрессия

Данные о статистической зависимости удобно задавать в виде корреляционной таблицы.

Рассмотрим в качестве примера зависимость между суточной выработкой продукции Y (т) и величиной основных производственных фондов X (млн руб.) для совокупности 50 однотипных предприятий (табл. 1).

В дальнейшем для краткости там, где это очевидно по смыслу, мы часто и выборочные уравнения (линии) регрессии будем называть просто уравнениями (линиями) регрессии.

(В таблице через x_i и y_j обозначены середины соответствующих интервалов, а n_i и n_j — соответственно их частоты).

Изобразим полученную зависимость графически точками координатной плоскости (рис. 1). Такое изображение статистической зависимости называется полем корреляции.

Для каждого значения x_i ($i = 1,2,...,l$), т.е. для каждой строки корреляционной таблицы вычислим групповые средние

$$\bar{y}_i = \frac{\sum_{j=1}^{m} y_j n_{ij}}{n_i}, \quad (5)$$

где n_{ij} — частоты пар (x_i, y_j) и $n_i = \sum_{i=1}^{m} n_{ij}$, m — число интервалов по переменной Y.

Таблица 1

Величина ОПФ, млн. руб.(X)	Середины интервалов	Суточная выработка продукции, т(Y)					Всего n_i	Групповая средняя, т $\bar{y}_i$
		7–11	11–15	15–19	19–23	23–27		
	x_i \ y_j	9	13	17	21	25		
20–25	22,5	2	1	–	–	–	3	10,3
25–30	27,5	3	6	4	–	–	13	13,3
30–35	32,5	–	3	11	7	–	21	17,8
35–40	37,5	–	1	2	6	2	11	20,3
40–45	42,5	–	—	–	1	1	2	23,0
Всего n_j		5	11	17	14	3	50	—
Групповая средняя $\bar{x}_j$, млн руб.		25,5	29,3	31,9	35,4	39,2	–	–

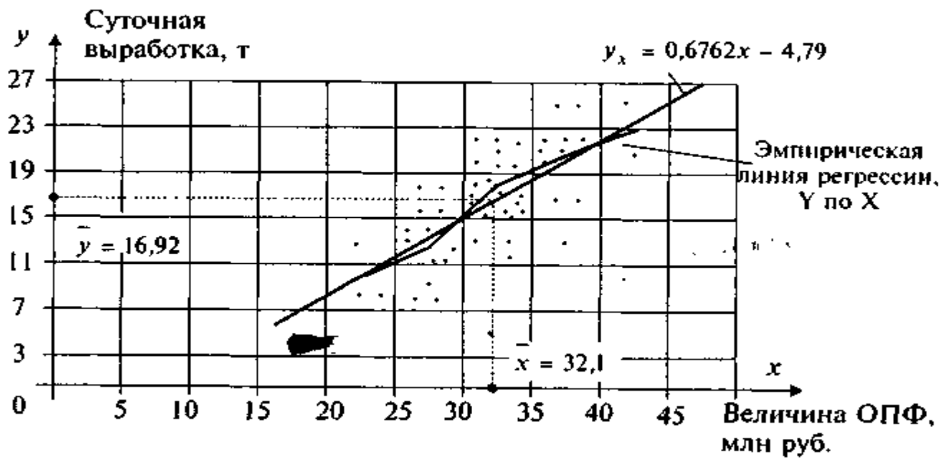


Рис. 1

Вычисленные групповые средние $\overline{y_i}$ поместим в последнем столбце корреляционной таблицы и изобразим графически в виде ломаной, называемой *эмпирической линией регрессии* Y по X (рис. 1).

Аналогично для каждого значения y_j (j = 1,2,...,m) по формуле

$$\bar{x}_j = \frac{\sum_{i=1}^{l} x_i n_{ij}}{n_j} \qquad (6)$$

вычислим групповые средние x, (см. нижнюю строку корреляционной таблицы), где $n_j = \sum_{i=1}^{l} n_{ij}$, l - число интервалов по переменной X.

По виду ломаной можно предположить наличие линейной корреляционной зависимости Y по X между двумя рассматриваемыми переменными, которая графически выражается тем точнее, чем больше объем выборки (число рассматриваемых предприятий) *n*:

$$n = \sum_{i=1}^{l} n_i = \sum_{j=1}^{m} n_j = \sum_{i=1}^{l}\sum_{j=1}^{m} n_{ij} \,. \qquad (7)$$

Поэтому уравнение регрессии (3) будем искать в виде:

$$y_x = b_0 + b_1 x \,. \qquad (8)$$

Найдем формулы расчета неизвестных параметров уравнения линейной регрессии. С этой целью применим метод наименьших квадратов, согласно которому неизвестные параметры $Ь_0$ и $Ь_1$ выбираются таким образом, чтобы сумма квадратов отклонений эмпирических групповых средних $\overline{y}_i$ вычисленных по формуле (5), от значений y_{x_i}, найденных по уравнению регрессии (8), была минимальной:

$$S = \sum_{i=1}^{l}\left(y_{x_i} - \overline{y}_i\right)^2 n_i = \sum_{i=1}^{l}\left(b_0 + b_1 x_i - \overline{y}_i\right)^2 n_i \to \min \,. \qquad (9)$$

На основании необходимого условия экстремума функции двух переменных S = S(Ь$_0$, b$_1$,) приравниваем нулю ее частные производные, т.е.

$$\begin{cases} \dfrac{dS}{db_0} = 2\sum\limits_{i=1}^{l}\left(b_0 + b_1 x_i - \overline{y}_i\right) n_i = 0, \\ \dfrac{dS}{db_1} = 2\sum\limits_{i=1}^{l}\left(b_0 + b_1 x_i - \overline{y}_i\right) x_i n_i = 0, \end{cases}$$

откуда после преобразований получим систему нормальных уравнений для определения параметров линейной регрессии:

$$\begin{cases} b_0 \sum\limits_{i=1}^{l} n_i + b_1 \sum\limits_{i=1}^{l} x_i n_i = \sum\limits_{i=1}^{l} \overline{y}_i n_i \,, \\ b_0 \sum\limits_{i=1}^{l} x_i n_i + b_1 \sum\limits_{i=1}^{l} x_i^2 n_i = \sum\limits_{i=1}^{l} x_i \overline{y}_i n_i \,. \end{cases} \qquad (10)$$

Учитывая (5), преобразуем выражения:

$$\sum_{i=1}^{l}\bar{y}_i n_i = \sum_{i=1}^{l}\left(\frac{\sum_{j=1}^{m} y_j n_{ij}}{n_i}\right) n_i = \sum_{i=1}^{l}\sum_{j=1}^{m} y_j n_{ij} = \sum_{j=1}^{m} y_j \sum_{i=1}^{l} n_{ij} = \sum_{j=1}^{m} y_j n_j,$$

$$\sum_{i=1}^{l} x_i \bar{y}_i n_i = \sum_{i=1}^{l} x_i\left(\frac{\sum_{j=1}^{m} y_j n_{ij}}{n_i}\right) n_i = \sum_{i=1}^{l}\sum_{j=1}^{m} x_i y_j n_{ij}.$$

Теперь с учетом (7), разделив обе части уравнений (10) на *n*, получим систему нормальных уравнений в виде:

$$\begin{cases} b_0 + b_1\bar{x} = \bar{y}, \\ b_0\bar{x} + b_1\overline{x^2} = \overline{xy}, \end{cases} \quad (11$$

где соответствующие средние определяются по формулам:

$$\bar{x} = \frac{\sum_{i=1}^{l} x_i n_i}{n}, \quad \bar{y} = \frac{\sum_{j=1}^{m} y_j n_j}{n}, \qquad \textbf{(12.12)}$$

$$\overline{xy} = \frac{\sum_{i=1}^{l}\sum_{j=1}^{m} x_i y_j n_{ij}}{n}, \qquad \textbf{(12.13)}$$

$$\overline{x^2} = \frac{\sum_{i=1}^{l} x_i^2 n_i}{n}. \qquad \textbf{(12.14)}$$

Подставляя значение $Ь_0$ = $\bar{y}$ - $Ь_x$ из первого уравнения системы (11) в уравнение регрессии (8), получим $\bar{y}_x - \bar{y} = b_1(\tilde{o} - \tilde{o})$ (15)

Коэффициент $Ь_1$ в уравнении регрессии, называемый выборочным коэффициентом регрессии (или просто коэффициентом регрессии) У по Х, будем обозначать символом $Ь_{ух}$. Теперь уравнение регрессии Y по X запишется так:

$$y_x - \bar{y} = b_{yx}(x - \bar{x}). \qquad \textbf{(12.16)}$$

Коэффициент регрессии У по Х показывает, на сколько единиц в среднем изменяется переменная Y при увеличении переменной X на одну единицу.

Решая систему (12.11), найдем

$$b_{yx} = b_1 = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{\overline{xy} - \bar{x}\,\bar{y}}{s_x^2} = \frac{\mu}{s_x^2}, \qquad (12.17)$$

где S_x^2 — выборочная дисперсия переменной X:

$$s_x^2 = \overline{x^2} - \bar{x}^2 = \frac{\sum_{i=1}^{l} x_i^2 n_i}{n} - \left(\bar{x}\right)^2; \qquad (12.18)$$

μ — выборочный корреляционный момент или выборочная ковариация:

$$\mu = \overline{xy} - \bar{x}\,\bar{y} = \frac{\sum_{i=1}^{l}\sum_{j=1}^{m} x_i y_j n_{ij}}{n} - \bar{x}\,\bar{y}. \qquad (12.19)$$

Рассуждая аналогично и полагая уравнение регрессии (4) линейным, можно привести его к виду:

$$x_y - \bar{x} = b_{xy}\left(y - \bar{y}\right), \qquad (12.20)$$

где

$$b_{xy} = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\overline{y^2} - \bar{y}^2} = \frac{\overline{xy} - \bar{x}\,\bar{y}}{s_y^2} = \frac{\mu}{s_y^2} \qquad (12.21)$$

— выборочный коэффициент регрессии (или просто коэффициент регрессии) X по Y, показывающий, на сколько единиц в среднем изменяется переменная X при увеличении переменной У на одну единицу,

$$s_y^2 = \overline{y^2} - \bar{y}^2 = \frac{\sum_{j=1}^{m} y_j^2 n_j}{n} - \left(\bar{y}\right)^2 \qquad (12.22)$$

—выборочная дисперсия переменной Y.

Так как числители в формулах (17) и (21) для $Ь_{ух}$ и $Ь_{ху}$ совпадают, а знаменатели — положительные величины, то коэффициенты регрессии $Ь_{ух}$ и $Ь_{ху}$, имеют одинаковые знаки, определяемые знаком μ. Из уравнений регрессии (16) и (20) следует, что коэффициенты $Ь_{ух}$ и $1/Ь_{ху}$ определяют угловые коэффициенты (тангенсы углов наклона) к оси ох соответствующих линий регрессии, пересекающихся в точке ($\bar{y}$,$\bar{x}$) (см. рис. 3).

П.3. Коэффициент корреляции

Перейдем к оценке тесноты корреляционной зависимости. Рассмотрим наиболее важный для практики и теории случай линейной зависимости вида (16).

На первый взгляд подходящим измерителем тесноты связи Y от X является коэффициент регрессии Ь$_{ух}$ ибо, как уже отмечено, он показывает, на сколько единиц в среднем изменяется Y, когда X увеличивается на одну единицу. Однако Ь$_{ух}$ зависит от единиц измерения переменных. Например, в полученной ранее зависимости он увеличится в 1000 раз, если величину основных производственных фондов X выразить не в млн руб., а в тыс. руб. Очевидно, что для «исправления» Ь$_{ух}$ как показателя тесноты связи нужна такая стандартная система единиц измерения, в которой данные по различным характеристикам оказались бы сравнимы между собой. Статистика знает такую систему единиц. Эта система использует в качестве единицы измерения переменной ее среднее квадратическое отклонение S.

Представим уравнение (16) в эквивалентном виде:

$$\frac{y_x - \bar{y}}{s_y} = \left(b_{yx}\frac{s_x}{s_y}\right)\frac{x - \bar{x}}{s_x}. \qquad (28)$$

В этой системе величина

$$r = b_{yx}\frac{s_x}{s_y} \qquad (29)$$

показывает, на сколько величин S_y изменится в среднем Y, когда X увеличится на одно S_x Величина ***r*** является показателем тесноты связи и называется выборочным коэффициентом корреляции (или просто коэффициентом корреляции).

На рис. 2 приведены две корреляционные зависимости переменной Y по X. Очевидно, что в случае а) зависимость между переменными менее тесная и коэффициент корреляции должен быть меньше, чем в случае б), так как точки корреляционного поля а) дальше отстоят от линии регрессии, чем точки поля б). Нетрудно видеть, что $\boldsymbol{r}$ совпадает по знаку с Ь$_{yx}$ (а значит, и с Ь$_{xy}$).

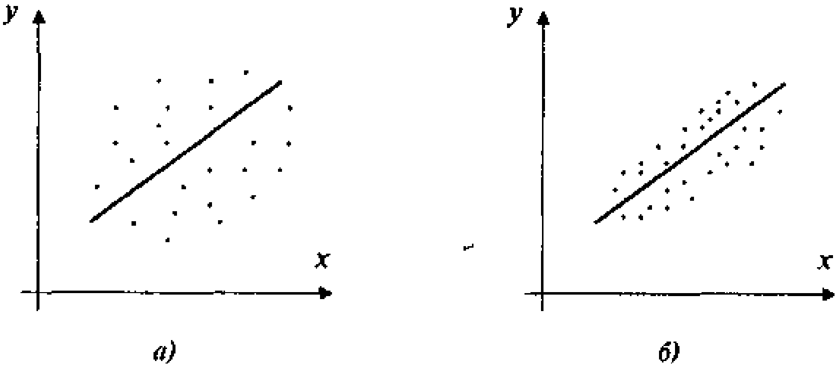


Рис. 2

Если $\boldsymbol{r} > 0$ (Ь$_{yx}$>0, Ь$_{xy}$>0), то корреляционная связь между переменными называется прямой, если $\boldsymbol{r}$< О (Ь$_{yx}$ <0, Ь$_{xy}$<0) — обратной. При прямой (обратной) связи увеличение одной из переменных ведет к увеличению (уменьшению) условной (групповой) средней другой.

Учитывая (17), формулу для $\boldsymbol{r}$ представим в виде:

$$r = \frac{\overline{xy} - \bar{x}\,\bar{y}}{s_x s_y}. \qquad \textbf{(12.30)}$$

Отсюда видно, что формула для $\boldsymbol{r}$ симметрична относительно двух переменных, т.е. переменные X и Y можно менять местами. Тогда аналогично (24) можно записать:

$$r = b_{xy}\frac{s_y}{s_x}. \qquad \textbf{(12.31)}$$

Найдя произведение обеих частей равенств (29) и (31), получим

$$r^2 = b_{yx}b_{xy} \qquad \textbf{(12.32)}$$

$$r = \pm\sqrt{b_{yx}b_{xy}}, \qquad \textbf{(12.33)}$$

т.е. коэффициент корреляции $\boldsymbol{r}$ переменных X и Y есть средняя геометрическая коэффициентов регрессии, имеющая их знак.

П.4. Проверка значимости и интервальная оценка параметров связи

В практических исследованиях о тесноте корреляционной зависимости между рассматриваемыми переменными судят фактически не по величине генерального коэффициента корреляции ρ (который обычно неизвестен), а по величине его выборочного аналога ***r***. Так как ***r*** вычисляется по значениям переменных, случайно попавшим в выборку из генеральной совокупности, то в отличие от параметра ρ параметр ***r*** — величина случайная.

Пусть вычисленное значение ***r*** = 0. Возникает вопрос, объясняется ли это действительно существующей линейной корреляционной связью между переменными X и Y в генеральной совокупности или является следствием случайности отбора переменных в выборку (т.е. при другом отборе возможно, например, ***r*** = 0 или изменение знака ***r***).

Обычно в этих случаях проверяется гипотеза H_0: об отсутствии линейной корреляционной связи между переменными в генеральной совокупности, т.е. H_0: **ρ** = 0. При справедливости этой гипотезы статистика

$$t=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

имеет t-распределение Стьюдента с k = n—2 степенями свободы. Поэтому гипотеза H_0 отвергается, т.е. выборочный коэффициент корреляции ***r*** значимо (существенно) отличается от нуля, если

$$|t|=\frac{|r|\sqrt{n-2}}{\sqrt{1-r^2}}\rangle t_{1-a:k},$$

где $t_{1-a;k}$ — табличное значение t-критерия Стьюдента, определенное на уровне значимости **α** при числе степеней свободы k = n-2.

Для значимого коэффициента корреляции ***r*** целесообразно найти доверительный интервал (интервальную оценку), который с заданной надежностью **γ** = 1 - α содержит (точнее, «накрывает») неизвестный генеральный коэффициент корреляции **ρ**. Для построения такого интервала необходимо знать выборочное распределение коэффициента корреляции ***r***, которое при **ρ** = 0 несимметрично и очень медленно (с ростом ***n***) сходится к нормальному распределению. Поэтому прибегают к специально подобранным функциям от ***r***, которые сходятся к хорошо изученным распределениям. Чаще всего для подбора функции применяют Z-преобразование Фишера.

$$Z = \frac{1}{2}\ln\frac{1+r}{1-r}. \qquad (45)$$

Распределение уже при небольших ***n*** является приближенно нормальным с математическим ожиданием

$$M(Z) = \frac{1}{2}\ln\frac{1+\rho}{1-\rho} + \frac{\rho}{2(n-1)} \qquad (46)$$

дисперсией $\sigma_2^2 = \frac{1}{n-3}$ (47)

Поэтому вначале строят доверительный интервал для M{z}:

$$z - t_{1-a}\frac{1}{n-3} \le M\left(z\right) \le z + t_{1-a}\frac{1}{\sqrt{n-3}}$$

где $t_{1-\alpha}$ — нормированное отклонение z, определяемое с помощью функции Лапласа:

$$\hat{O}\left(t_{1-a}\right) = \gamma = 1-\alpha$$

При определении границ доверительного интервала для **ρ**, т.е. для перехода от z к **ρ**, существует специальная таблица. При ее отсутствии переход может быть осуществлен по формуле:

$$r = thz = \frac{e^2 - e^{-2}}{e^2 + e^{-2}}$$

где th z— гиперболический тангенс z.

Если коэффициент корреляции значим, то коэффициенты регрессии b_{yx} и b_{xy} также значимо отличаются от нуля, а интервальные оценки для соответствующих генеральных коэффициентов регрессии β_{yx} и β_{xy} могут быть получены по формулам, основанным на том, что статистики $(b_{yx}-\beta_{xy})/S_{byx}$, $(b_{xy}-\beta_{yx})\backslash S_{bxy}$ имеют t-распределение Стьюдента с (n—2) степенями свободы:

$$b_{yx}-t_{1-a:n-2}\frac{s_y\sqrt{1-r^2}}{s_x\sqrt{n-2}}\le\beta_{yx}\le b_{yx}+t_{1-a:n-2}\cdot\frac{s_y\sqrt{1-r^2}}{s_x\sqrt{n-2}} \quad (51)$$

$$b_{xy}-t_{1-a:n-2}\frac{s_x\sqrt{1-r^2}}{s_y\sqrt{n-2}}\le\beta_{xy}\le b_{xy}+t_{1-a:n-2}\cdot\frac{s_x\sqrt{1-r^2}}{s_y\sqrt{n-2}} \quad (52)$$

П.5. Корреляционное отношение и индекс корреляции

Введенный выше коэффициент корреляции, как уже отмечено, является полноценным показателем тесноты связи лишь в случае линейной зависимости между переменными. Однако часто возникает необходимость в достоверном показателе интенсивности связи при любой форме зависимости.

Для получения такого показателя вспомним правило сложения дисперсий:

$$s_y^2=s_{iy}^{i2}+\delta_{iy}^2$$

$$s_y^2=\frac{\sum_{j=1}^{m}\left(y_j-y\right)^{-2}n_i}{n} \quad (53, 54)$$

S^2_{iy} — средняя групповых дисперсий S^2_{iy} , или остаточная диспепсия —

$$s_{iy}^{i2}=\frac{\sum_{i-1}^{i}s_{iy}^2n_i}{n} \qquad s_{iy}^2=\frac{\sum_{j-1}^{m}\left(y_i-y_j\right)^{-2}}{n}$$

$$\delta_{iy}^2=\frac{\sum_{i-1}^{i}\left(y_i-y\right)n_i^{-2}}{n}$$

(55, 56, 57) межгрупповая дисперсия

Остаточной дисперсией измеряют ту часть колеблемости Y, которая возникает из-за изменчивости неучтенных факторов, не зависящих от X. Межгрупповая дисперсия выражает ту часть вариации У, которая обусловлена изменчивостью X. Величина

$$\eta_{yx} = \sqrt{\frac{\delta_{iy}^2}{s_y^2}} \quad (58)$$

получила название **эмпирического корреляционного отношения У по X**. Чем теснее связь, тем большее влияние на вариацию переменной доказывает изменчивость X.по сравнению с неучтенными факторами, тем выше n_{yx}. Величина n_{yx}. , называемая эмпирическим коэффициентом детерминации, показывает, какая часть общей вариации ***У*** обусловлена вариацией X. Аналогично вводится эмпирическое корреляционное отношение X по У:

$$\eta_{yx} = \sqrt{\frac{\delta_{ix}^2}{s_x^2}} \quad (59)$$

Отметим основные **свойства корреляционных отношений** (при достаточно большом объеме выборки *n*):

1.Корреляционное отношение есть неотрицательная величина, не превосходящая 1: $0 < \eta < 1$.

2. Если $\eta = 0$, то корреляционная связь отсутствует. Если $\eta = 1$, то между переменными существует функциональная зависимость.

3. $\eta_{yx} \neq \eta_{xy}$, т.е. в отличие от коэффициента корреляции $\boldsymbol{r}$ (для которого $\boldsymbol{r}_{xy} = \boldsymbol{r}_{yx} = \boldsymbol{r}$) при вычислении корреляционного отношения существенно, какую переменную считать независимой, а какую — зависимой.

Эти свойства справедливы как для эмпирических корреляционных отношений *n*, так и для теоретических — R .

Эмпирическое корреляционное отношение η_{yx} является показателем рассеяния точек корреляционного поля относительно эмпирической линии регрессии, выражаемой ломаной, соединяющей значения $\overline{y}_i$. Однако в связи с тем, что закономерное изменение $\overline{y}_i$ нарушается случайными зигзагами ломаной, возникающими вследствие остаточного действия неучтенных факторов, η_{yx} преувеличивает тесноту связи. Поэтому наряду с η_{yx}. рассматривается показатель тесноты связи R_{yx}, характеризующий рассеяние точек корреляционного поля относительно линии регрессии y_x (12.3).

Показатель R_{yx} получил название **теоретического корреляционного отношения** или индекса корреляции Y по X.

$$R_{yx} = \sqrt{\frac{\delta_r^{i2}}{s_y^2}} = \sqrt{1 - \frac{s_r^{i2}}{s_y^2}} \quad (60)$$

где дисперсии δ_y^2 и S_y^2 определяются по формулам (54) —(56), в которых групповые средние $\bar{y}_i$, заменены условными средними y_{xi}, вычисленными по уравнению регрессии (16). Подобно R_{yx} вводится и индекс корреляции X по Y

$$R_{xy} = \sqrt{\frac{\delta_x^2}{s_x^2}} = \sqrt{1 - \frac{s_x^{i2}}{s_x^2}} \quad (61)$$

Достоинством рассмотренных показателей η и R является то, что они могут быть вычислены при любой форме связи между переменными. Хотя η и завышает тесноту связи по сравнению с R, но для его вычисления не нужно знать уравнение регрессии. Корреляционные отношения η и R связаны с коэффициентом корреляции ***r*** следующим образом:

$$0 \le |r| \le R \le \eta \le 1. \quad (62)$$

Можно показать, что в случае линейной модели (3), т.е. зависимости, $y_x - \bar{y} = b_{yx}(x - \bar{x})$ индекс корреляции R_{yx} равен коэффициенту корреляции ***r*** (по абсолютной ветчине): $R_{yx} = |\,r\,|$

$$R_{yx} = \frac{|xy - xy|}{s_x^2}\sqrt{\frac{s_x^2}{s_y^2}} = \frac{|xy - xy|}{s_x s_y} = |r|.$$

Коэффициент детерминации R^2, равный квадрату индекса корреляции (для парной линейной модели — r^2), показывает долю общей вариации зависимой переменной, обусловленной регрессией или изменчивостью объясняющей переменной. Чем ближе R^2 к 1, тем теснее наблюдения примыкают к линии регрессии, тем лучше регрессия описывает зависимость переменных.

Расхождение между η^2 и R^2 (или r^2) может быть использовано для проверки линейности корреляционной зависимости.

Проверка значимости корреляционного отношения η основана на том, что статистика

$$F = \frac{\eta^2 (n-m)}{(1-\eta^2)(m-1)} \quad (63)$$

(где m — число интервалов по группировочному признаку) имеет F-распределение Фишера—Снедекора с $k_1 = m - 1$ и $k_2 = n - m$ степенями свободы. Поэтому η значимо отличается от нуля, если $F > F_{\alpha,, k1, k2}$, где $F_{\alpha,, k1, k2}$— табличное значение F-критерия на уровне значимости α при числе степеней свободы k_1 и k_2.

Индекс корреляции R двух переменных значим, если значение статистики

$$F = \frac{R^2 (n-2)}{1-R^2} \quad (64)$$

больше табличного $F_{\alpha,, k1, k2}$, где $к_1 = 1$ и $к._2 = n - 2$.

ГЛОССАРИЙ

Аддитивная функция – функция $\mu(A)$ множеств- элементов алгебры **A**, для которой из условия $A \cap B =$ ш следует, что $\mu(A \cup B) = \mu(A) + \mu(B)$.

Алгебра множеств – система подмножеств **A** множества **Ω**, элементы которой удовлетворяют следующим требованиям:

а) $\Omega \in$ **A**; б) для любых **A** и **B**, принадлежащих **A**, следует, что $A \cup B \in$ **A** и $A \cap B \in$ **A**; в) если $A \in$ **A**, то $\overline{A} = \Omega \setminus A \in$ **A**.

Борелевская алгебра множеств B(R) – система подмножеств множества действительных чисел **R**, получающаяся путём применения операций объединения, пересечения и дополнения к элементам системы $I = \{[a, b)\}$, где **a** и **b** – произвольные действительные числа.

Вероятностное пространство <Ω,A, P> - тройка объектов, где

Ω - множество элементарных исходов;

A - σ-алгебра случайных событий;

P – вероятностная функция.

Дискретная случайная величина – случайная величина, областью возможных значений которой является не более чем счётное множество D действительных чисел $\{x_k\}, \quad k = 1, 2, \ldots, n, \infty$. Закон распределения вероятностей дискретной случайной величины задаётся путём определения набора положительных чисел $\{p_k\}, \quad k = 1, 2, \ldots, n, \infty$, таких, что $\sum_{k=1}^{n,\infty} p_k = 1$. Здесь: $p_k = P(\xi(\omega) = x_k)$.

Дисперсия случайной величины $D\xi = M[\xi - M\xi]^2$ - мера разброса значений случайной величины около её математического ожидания.

Доверительный интервал $I_\gamma = (\alpha, \beta)$ - интервал, в котором с вероятностью, не меньшей чем γ, находится значение неизвестной числовой характеристики ϑ, то есть интервал, для которого справедливо: $P(\alpha < \vartheta < \beta) \geq \gamma$.

Закон больших чисел (ЗБЧ) – совокупность теорем, в которых на последовательность случайных величин $\{\xi_k\}$, $k = 1,2,...$, налагаются условия, при которых их среднее арифметическое $\frac{1}{n}\sum_{k=1}^{n}\xi_k$ сходится по вероятности к постоянной величине – среднему арифметическому их математических ожиданий: $\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}M\xi_k\right| < \varepsilon\right) = 1$.

Измеримое пространство <Ω,A> - пара объектов, где **Ω** - множество элементарных исходов, **A -** алгебра случайных событий, на которой вводится числовая функция множеств $\mu(A)$, которая при выполнении условий нормированности и аддитивности, называется вероятностной мерой множества **A**.

Классическое определение вероятности – определение вероятности наступления случайного события, основанное на равновозможности реализации элементарных исходов конечного множества элементарных исходов **Ω**. Если мощность множества **Ω** равна $|\Omega| = n$, а мощность подмножества **A**, являющегося случайным событием, равна $|A| = m$, то по классическому определению вероятности вероятность наступления случайного события **A** будет равна $P(A) = \frac{m}{n}$.

Ковариационный момент – смешанный центральный момент второго порядка μ_{11} двумерной случайной величины:

$$cov(\xi_1, \xi_2) = \mu_{11} = M[(\xi_1 - M\xi_1)\cdot(\xi_2 - M\xi_2)].$$

Компонента случайного вектора – скалярная случайная величина $\xi_k(\omega)$, являющаяся проекцией случайного вектора $\xi(\omega)$ на **k**-тую координатную ось R_k. То есть, если $\xi(\omega): \Omega \to R^n$ и π_k- проектор, отображающий R^n в R_k, то $\xi_k(\omega)$ является композицией отображений:

$$\xi_k(\omega) = \pi_k \circ \xi(\omega): \Omega \to R^n \to R_k.$$

Коэффициент линейной корреляции – мера статистической силы связи между случайными величинами. Вычисляется по формуле $\rho = \dfrac{\alpha_{11} - m_1 \cdot m_2}{\sigma_1 \cdot \sigma_2}$. Применяется в тех случаях, когда статистическая связь имеет линейный характер.

Критерий проверки основной гипотезы – случайная величина, статистика элементов выборки, закон распределения вероятностей которой зависит от предполагаемой гипотезы.

Математическое ожидание – числовая характеристика случайной величины, $M\xi = \int\limits_R x dF(x)$. Математическое ожидание есть среднее значение случайной величины ξ. Интерпретируется как координата центра тяжести единичной массы распределённой на числовой оси.

Множество элементарных исходов – множество, элементами, которого является все возможные элементарные исходы. В результате проведения испытания всегда реализуется один, и только один элементарный исход.

Начальный момент k-того порядка – числовая характеристика случайной величины, являющаяся значением абсолютно сходящегося несобственного интеграла от функции $g(\xi) = \xi^k$ по функции распределения случайной величины, то есть: $\alpha_k = M[\xi^k] = \int\limits_R x^k dF(x)$.

Независимость случайных величин. Случайные величины $\xi_1(\omega)$ и $\xi_2(\omega)$ называются независимыми, если закон распределения вероятностей одной из них не зависит от другой случайной величины.

Точнее: пусть случайные величины $\xi_1(\omega)$ и $\xi_2(\omega)$ являются компонентами двумерной случайной величины $\xi(\omega) = (\xi_1(\omega), \xi_2(\omega))$, принимающей значения в R^2. Эти компоненты называются независимыми, если для любого множества $\mathbf{B}, B \in \mathbf{B}(R^2)$, представимого как декартово произведение $B = B_1 \times B_2$, $B_1 \subset R_1$ и $B_2 \subset R_2$, будет справедливо:

$$P(\xi(\omega) \in B) = P_1(\xi_1(\omega) \in B_1) \cdot P_2(\xi_2(\omega) \in B_2),$$

Где P_1 и P_2 - частные вероятностные функции компонент.

Независимость случайных величин непрерывного типа – Случайные величины непрерывного типа $\xi_1(\omega)$ и $\xi_2(\omega)$ (компоненты двумерного случайного вектора) будут независимыми тогда, только тогда, когда для любой пары $(x,y)\in R^2$ выполняется равенство $p(x,y)=p_1(x)\cdot p_2(y)$, где $p(x,y)$ - плотность вероятности двумерного случайного вектора $\xi(\omega)=(\xi_1(\omega),\xi_2(\omega))$, а $p_1(x)$ и $p_2(y)$ - плотности вероятностей его компонент $\xi_1(\omega)$ и $\xi_2(\omega)$.

Независимость случайных величин дискретного типа – Случайные величины дискретного типа $\xi_1(\omega)$ и $\xi_2(\omega)$ (компоненты двумерного случайного вектора) будут независимыми тогда, только тогда, когда для любой пары $(x_i,y_k)\in R^2$ выполняется равенство $p_{i,k}=p_{i*}\cdot p_{*k}$, где $p_{i,k}=P(\xi(\omega)=(x_i,y_k))$, а $p_{i*}=P(\xi_1(\omega)=x_i)$ и $p_{*k}=P(\xi_2(\omega)=y_k)$.

Независимость случайных событий. Случайные события называются независимыми, если условная вероятность наступления любого из них равна его безусловной вероятности: $P(A/B)=P(A)$ или $P(B/A)=P(B)$.

Непрерывная случайная величина – случайная величина, областью возможных значений которой является множество **D** мощности континуум и положительной меры Лебега. Закон распределения вероятностей непрерывной случайной величины задаётся путём определения на этом множестве плотности вероятности $p(x)$ - кусочно-непрерывной, неотрицательной функции, такой что $\int\limits_D p(x)dx=1$.

Несмещённость точечной оценки. Точечная оценка ϑ^* числовой характеристики ϑ называется несмещённой, если $M\vartheta^*=\vartheta$.

Остаточная дисперсия – мера разброса значений одной из компонент (например ξ_2) двумерной случайной величины $\xi=(\xi_1;\xi_2)$около её математического ожидания, вызванного внутренними свойствами этой компоненты. При линейном виде статистической связи между компонентами величина остаточной дисперсии компоненты ξ_2 равна $D\xi_{ост} = \left(1-\rho^2\right)\cdot D\xi_2$, где ρ - коэффициент линейной корреляции между компонентами ξ_1 и ξ_2.

Ошибка I рода – отклонение верной гипотезы H_0. Возникает в том случае, когда при справедливости в реальности гипотезы H_0 наблюдаемое значение критерия $T_{набл}$ попадает в критическую область $S_{кр}$. Вероятность ошибки **I** рода равна $\alpha = P\left(T_{набл} \in S_{кр} \middle| H_0\right)$.

Ошибка II рода – принятие неверной гипотезы H_0. Возникает в том случае, когда при справедливости в реальности гипотезы H_1 наблюдаемое значение критерия $T_{набл}$ попадает в область допустимых значений $Q_{доп}$. Вероятность ошибки **II** рода равна $\beta = P\left(T_{набл} \in Q_{доп} \middle| H_1\right)$.

Повторные независимые испытания – серия одинаковых испытаний, в каждом из которых с постоянными вероятностями **p** и **q** может произойти только одно из взаимно противоположных событий **A** или $\overline{A}$.

Плотность вероятности – неотрицательная, кусочно-непрерывная функция, удовлетворяющая условию: $\int\limits_R p(x)dx = 1$. Плотность вероятности описывает распределение вероятностей случайной величины $\xi(\omega)$ непрерывного типа.

Распределение χ^2 - (распределение Пирсона) распределение вероятностей случайной величины $\chi_n^2 = \sum\limits_{k=1}^{n} \xi_k^{\ 2}$, где все ξ_k, $k = 1,\ldots,n,$ независимые случайные величины, имеющие нормальное распределение вероятностей N(0;1).

Распределение Стьюдента – (**t**-распределение) распределение вероятностей случайной величины $t_n = \dfrac{\xi_0}{\sqrt{\dfrac{1}{n}\sum\limits_{k=1}^{n}\xi_k^{\ 2}}}$, где все ξ_k, $k = 0,1,\ldots,n,$ независимые случайные величины, имеющие нормальное распределение вероятностей N(0;1).

Распределение Фишера-Снедекора – (**F**-распределение) распределение вероятностей случайной величины $f_{m,n} = \dfrac{\frac{1}{m}\chi_m^2}{\frac{1}{n}\chi_n^2}$.

Ряд распределения – таблица, состоящая из двух строк, с помощью которой задаётся закон распределения дискретной случайной величины:

$\xi(\omega)$	x_1	x_2	...	x_k	...	x_n	...
p_k	p_1	p_2	...	p_k	...	p_n	...

.

Где $\{x_k\},\quad k = 1,2,...,n$ или ∞; $p_k = P(\xi(\omega) = x_k)$. Всегда $\sum_{k=1}^{n,\infty} p_k = 1$.

Свёртка функций распределения – несобственный интеграл, определяющий функцию распределения случайной величины, являющейся суммой независимых случайных величин. Если $\xi = \xi_1 + \xi_2$, то функция распределения ξ будет равна: $F(x) = \int_{-\infty}^{+\infty} F_2(x - x_1) dF_1(x_1)$, где $F_1(x_1)$ и $F_2(x_2)$ - функции распределения случайных величин-слагаемых.

Состоятельность точечной оценки. Точечная оценка ϑ^* числовой характеристики ϑ называется состоятельной, если она сходится по вероятности к этой точечной оценке, то есть: $\lim_{n\to\infty} P(|\vartheta^* - \vartheta| < \varepsilon) = 1$.

Статистика – любая функция элементов выборки $\{x_i\},\quad i = 1,...,n$: $\vartheta^* = g(x_1, x_2, ..., x_n)$.

Сходимость по вероятности. Последовательность случайных величин $\{\xi_n\},\quad n = 1,2,3,...$ сходится по вероятности к случайной величине ξ (обозначение: $\xi_n \xrightarrow{P} \xi$), если выполняется условие $\lim_{n\to\infty} P(|\xi_n - \xi| < \varepsilon) = 1$.

Сходимость по распределению. Последовательность случайных величин $\{\xi_n\},\quad n = 1,2,3,...$ сходится по распределению к случайной величине ξ (обозначение: $\xi_n \xrightarrow{w} \xi$), если соответствующая последовательность функций распределения $\{F_n(x)\}$ слабо сходится к функции распределения $F(x)$ случайной величины ξ ($F_n(x) \Rightarrow F(x)$).

Условная вероятность $P(A/B)$ - вероятность наступления случайного события **A**, вычисленная при предположении, что случайное событие **B** произошло. Определяется по формуле: $P(A/B) = \dfrac{P(A \cap B)}{P(B)}$.

Условная плотность вероятности $p(y/\xi_1 = x_0)$ - плотность вероятности условной случайной величины $\xi_2/\xi_1 = x_0$, является законом распределения вероятностей второй компоненты при любом фиксированном значении первой компоненты. Определяется по формуле: $p(y/\xi_1 = x_0) = \dfrac{p(x_0, y)}{p_1(x_0)}$, где $p(x, y)$ - плотность вероятности двумерной случайной величины $\xi = (\xi_1, \xi_2)$, $p_1(x) = \int\limits_{R_2} p(x, y)dy$ - частная плотность вероятности первой компоненты ξ_1.

Функция распределения – функция $F(x)$, описывающая изменение вероятности случайного события $\{\xi \in (-\infty; x)\}$ при изменении **x**, то есть $F(x) = P(-\infty < \xi < x)$. Определяя функцию распределения $F(x)$, мы задаём закон распределения вероятностей случайной величины ξ.

Функция распределения вектора - функция $F(x) = F(x_1, x_2, \ldots, x_n)$, описывающая изменение вероятности случайного события $\{\xi = \{\xi_1, \xi_2, \ldots, \xi_n\} \in S_x\}$, где $S_x = (-\infty; x_1) \times (-\infty; x_2) \times \ldots \times (-\infty; x_n)$, при изменении $x = (x_1, x_2, \ldots, x_n) \in R^n = R_1 \times R_2 \times \ldots \times R_n$, то есть $F(x) = P(\xi \in S_x)$. Определяя функцию распределения $F(x)$, мы задаём закон распределения вероятностей случайного вектора ξ.

Функция регрессии – функция, описывающая зависимость значений условных математических ожиданий одной из компонент двумерной случайной величины от другой компоненты. Функция $f(x) = M[\eta/\xi = x]$ - функция регрессии компоненты η на изменение компоненты ξ.

Функция $\varphi(y) = M[\xi/\eta = y]$ - функция регрессии компоненты ξ на изменение компоненты η.

Характеристическая функция – комплексно-значная функция действительного аргумента, являющаяся математическим ожиданием функции $g(\xi) = e^{it\xi}$ случайной величины ξ, где $t \in R$, то есть: $\varphi_\xi(t) = M\left[e^{it\xi}\right]$.

Частная функция распределения – функция распределения любой **k**-той компоненты ξ_k вектора $\xi = \{\xi_1, \xi_2, ..., \xi_n\}$. Определение частной функции распределения основано на свойстве согласованности функции распределения многомерной случайной величины, например, если **n**=2, то $F_1(x) = \lim\limits_{y\to\infty} F(x,y)$ и $F_2(y) = \lim\limits_{x\to\infty} F(x,y)$.

Частные распределения компонент случайного вектора - распределения вероятностей компонент вектора, являющихся скалярными случайными величинами. Частное распределение каждой компоненты получается как проекция вероятностной функции вектора на соответствующую координатную ось. Если $\xi = \{\xi_1, \xi_2\}$ и **P** вероятностная функция вектора, то частное распределение P_1 компоненты ξ_1 определяется равенством: $P_1 = P(\xi \in B_1 \times R_2)$, где $B_1 \in$ **B**(R_1). Аналогично, частное распределение P_2 компоненты ξ_2 определяется равенством: $P_2 = P(\xi \in R_1 \times B_2)$, где $B_2 \in$ **B**(R_2).

Центральная предельная теорема (ЦПТ) – совокупность теорем, в которых на последовательность случайных величин $\{\xi_k\}$, $k = 1, 2, ...$, налагаются условия, при которых их центрированная и нормированная сумма $\dfrac{\sum\limits_{k=1}^{n} \xi_k - \sum\limits_{k=1}^{n} M\xi_k}{\sqrt{\sum\limits_{k=1}^{n} D\xi_k}}$ сходится по распределению к нормальному закону **N**(0;1).

Эффективная оценка – точечная оценка числовой характеристики, имеющая наименьшую дисперсию.

ОБОЗНАЧЕНИЯ

Ω	- пространство элементарных случайных событий ω_i,
Γ	- генеральная совокупность,
$\varnothing$	- невозможное случайное событие,
$\mathcal{F}$	- σ-алгебра событий,
$A^k{}_n$	- число размещений,
$C^k{}_n$	- число сочетаний,
$p = P\ \{A\}$	- вероятность случайного события,
N	- множество целых чисел,
R	- множество действительных чисел,
$\wedge$, &	- «и»,
$\vee$	- «или»,
$\forall$	- квантор всеобщности
$\exists$	- квантор существования
$[x]$	- целая часть числа x,
exp(x)	- e^x,
$\xi = \xi\,(\omega)$	- случайная величина (сл.в.),
$F(x)=P\ \{\xi < X\}$	- функция распределения сл.в. ξ,
$\rho_\xi(x)$	- плотность сл.в. ξ,
$\varphi(x)$	- функция Гаусса (нормированная плотность нормального распределения),
$\Phi(x)$	- функция Лапласа (нормированное нормальное распределение),
$M\xi$	- математическое ожидание сл.в. ξ,
$D\xi$	- дисперсия сл.в. ξ,
σ_ξ	- средне - квадратическое отклонение сл.в. ξ,
$Cov(\xi,\eta)$	- ковариация сл.в. ξ, η,

$r_{\xi,\eta}$	- коэффициент корреляции сл.в. ξ, η,
$o(\Delta t)$	- бесконечно малая более высокого порядка, чем Δt
$\xi = \xi(\omega, t)$	- случайный процесс,
$V_k(t)$	- пуассоновский процесс,
m_r	- абсолютная частота,
W_r	- относительная частота,
$\widetilde{F}_n$	- эмпирическая функция распределения,
$k(z)$	- функция Колмогорова,
χ^2	- распределение Пирсона,
$\overline{\xi}$	- средне арифметическая,
s^2	- выборочная дисперсия,
$\hat{s}^2$	- исправленная дисперсия,
γ	- доверительная вероятность,
$S_{n-1}(x)$	- плотность распределения Стьюдента с (n-1) степенью свободы,
$v_{n-1}(x)$	- плотность распределения χ^2 с (n-1) степенью свободы,
$\overset{P}{\overline{\lim\limits_{n\to\infty}}}$	- сходимость по вероятности,
$\lambda(t)$	- интенсивность (плотность) потока случайных событий

ПРИЛОЖЕНИЯ

Таблица 1

Значения некоторых числовых величин

$\sqrt{2}$=1,41421356237309504880168872420969807856970…

π=3,14159265358979323846264338327950288419720…

e=2,71828182845904523536028747135266249775720…

Значения функции $f(x)=e^{-x}$.

x	*exp(-x)*	*x*	*exp(-x)*	*x*	*exp(-x)*	*x*	*exp(-x)*
0	1,000000	0,4	0,67032	0,8	0,449329	3	0,049787
0,02	0,980199	0,42	0,657047	0,82	0,440432	3,2	0,040762
0,04	0,960789	0,44	0,644036	0,84	0,431711	3,4	0,033373
0,06	0,941765	0,46	0,631284	0,86	0,423162	3,6	0,027324
0,08	0,923116	0,48	0,618783	0,88	0,414783	3,8	0,022371
0,1	0,904837	0,5	0,606531	0,9	0,40657	4	0,018316
0,12	0,88692	0,52	0,594521	0,92	0,398519	4,2	0,014996
0,14	0,869358	0,54	0,582748	0,94	0,390628	4,4	0,012277
0,16	0,852144	0,56	0,571209	0,96	0,382893	4,6	0,010052
0,18	0,83527	0,58	0,559898	0,98	0,375311	4,8	0,00823
0,2	0,818731	0,6	0,548812	1	0,367879	5	0,006738
0,22	0,802519	0,62	0,537944	1,02	0,360595	5,2	0,005517
0,24	0,786628	0,64	0,527292	1,04	0,353455	5,4	0,004517
0,26	0,771052	0,66	0,516851	1,06	0,346456	5,6	0,003698
0,28	0,755784	0,68	0,506617	1,08	0,339596	5,8	0,003028
0,3	0,740818	0,7	0,496585	1,1	0,332871	6	0,002479
0,32	0,726149	0,72	0,486752	1,12	0,32628	6,2	0,002029
0,34	0,71177	0,74	0,477114	1,14	0,319819	6,4	0,001662
0,36	0,697676	0,76	0,467666	1,16	0,313486	6,6	0,00136
0,38	0,683861	0,78	0,458406	1,18	0,307279	6,8	0,001114
						7	0,000912

Таблица 2

Значения функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$.

x	0	1	2	3	4	5	6	7	8	9
0,1	0,3968	0,3965	0,3961	0,3956	0,3951	0,3945	0,3939	0,3932	0,3925	0,3918
0,2	0,3907	0,3902	0,3894	0,3885	0,3876	0,3867	0,3857	0,3847	0,3836	0,3825
0,3	0,3809	0,3802	0,3790	0,3778	0,3765	0,3752	0,3739	0,3725	0,3712	0,3697
0,4	0,3677	0,3668	0,3653	0,3637	0,3621	0,3605	0,3589	0,3572	0,3555	0,3538
0,5	0,3514	0,3503	0,3485	0,3467	0,3448	0,3429	0,3410	0,3391	0,3372	0,3352
0,6	0,3325	0,3312	0,3292	0,3271	0,3251	0,3230	0,3209	0,3187	0,3166	0,3144
0,7	0,3115	0,3101	0,3079	0,3056	0,3034	0,3011	0,2989	0,2966	0,2943	0,2920
0,8	0,2890	0,2874	0,2850	0,2827	0,2803	0,2780	0,2756	0,2732	0,2709	0,2685
0,9	0,2654	0,2637	0,2613	0,2589	0,2565	0,2541	0,2516	0,2492	0,2468	0,2444
1	0,2413	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	0,2173	0,2155	0,2131	0,2107	0,2083	0,2059	0,2036	0,2012	0,1989	0,1965
1,2	0,1937	0,1919	0,1895	0,1872	0,1849	0,1826	0,1804	0,1781	0,1758	0,1736
1,3	0,1709	0,1691	0,1669	0,1647	0,1626	0,1604	0,1582	0,1561	0,1539	0,1518
1,4	0,1494	0,1476	0,1456	0,1435	0,1415	0,1394	0,1374	0,1354	0,1334	0,1315
1,5	0,1292	0,1276	0,1257	0,1238	0,1219	0,1200	0,1182	0,1163	0,1145	0,1127
1,6	0,1107	0,1092	0,1074	0,1057	0,1040	0,1023	0,1006	0,0989	0,0973	0,0957
1,7	0,0939	0,0925	0,0909	0,0893	0,0878	0,0863	0,0848	0,0833	0,0818	0,0804
1,8	0,0788	0,0775	0,0761	0,0748	0,0734	0,0721	0,0707	0,0694	0,0681	0,0669
1,9	0,0655	0,0644	0,0632	0,0620	0,0608	0,0596	0,0584	0,0573	0,0562	0,0551
2	0,0539	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0,0439	0,0431	0,0422	0,0413	0,0404	0,0396	0,0387	0,0379	0,0371	0,0363
2,2	0,0354	0,0347	0,0339	0,0332	0,0325	0,0317	0,0310	0,0303	0,0297	0,0290
2,3	0,0283	0,0277	0,0270	0,0264	0,0258	0,0252	0,0246	0,0241	0,0235	0,0229
2,4	0,0224	0,0219	0,0213	0,0208	0,0203	0,0198	0,0194	0,0189	0,0184	0,0180
2,5	0,0175	0,0171	0,0167	0,0163	0,0158	0,0154	0,0151	0,0147	0,0143	0,0139
2,6	0,0136	0,0132	0,0129	0,0126	0,0122	0,0119	0,0116	0,0113	0,0110	0,0107
2,7	0,0104	0,0101	0,0099	0,0096	0,0093	0,0091	0,0088	0,0086	0,0084	0,0081
2,8	0,0079	0,0077	0,0075	0,0073	0,0071	0,0069	0,0067	0,0065	0,0063	0,0061
2,9	0,0060	0,0058	0,0056	0,0055	0,0053	0,0051	0,0050	0,0048	0,0047	0,0046
3	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026	0,0025	0,0025
3,2	0,0024	0,0023	0,0022	0,0022	0,0021	0,0020	0,0020	0,0019	0,0018	0,0018
3,3	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014	0,0013	0,0013
3,4	0,0012	0,0012	0,0012	0,0011	0,0011	0,0010	0,0010	0,0010	0,0009	0,0009
3,5	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007	0,0007	0,0007	0,0006
3,6	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005	0,0004
3,7	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003	0,0003	0,0003	0,0003
3,8	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002	0,0002
3,9	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0001	0,0001

Таблица значений функции $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{\frac{-t^2}{2}}dt$

x	*0,00*	*0,01*	*0,02*	*0,03*	*0,04*	*0,05*	*0,06*	*0,07*	*0,08*	*0,09*
0,0	0,00000	0,00399	0,00798	0,01197	0,01595	0,01994	0,02392	0,02790	0,03188	0,03586
0,1	0,03983	0,04380	0,04776	0,05172	0,05567	0,05962	0,06356	0,06749	0,07142	0,07535
0,2	0,07962	0,08317	0,08706	0,09095	0,09483	0,09871	0,10257	0,10642	0,11026	0,11409
0,3	0,11791	0,12172	0,12552	0,12930	0,13307	0,13683	0,14058	0,14431	0,14803	0,15173
0,4	0,15542	0,15910	0,16276	0,16640	0,17003	0,17364	0,17724	0,18082	0,18439	0,18793
0,5	0,19146	0,19497	0,19847	0,20194	0,23891	0,20884	0,21226	0,21566	0,21904	0,22240
0,6	0,22575	0,22907	0,23237	0,23565	0,20540	0,24215	0,24537	0,24857	0,25175	0,25490
0,7	0,25804	0,26115	0,26424	0,26730	0,27035	0,27337	0,27637	0,27935	0,28230	0,28524
0,8	0,28814	0,29103	0,29389	0,29673	0,29955	0,30234	0,30511	0,30785	0,31057	0,31327
0,9	0,31594	0,31859	0,32121	0,32381	0,32639	0,32894	0,33147	0,33398	0,33646	0,33891
1,0	0,34134	0,34375	0,34614	0,34850	0,35083	0,35314	0,35543	0,35769	0,35993	0,36214
1,1	0,36433	0,36650	0,36864	0,37076	0,37286	0,37493	0,37698	0,37900	0,38100	0,38298
1,2	0,38493	0,38686	0,38877	0,39065	0,39251	0,39435	0,39607	0,39796	0,39973	0,40147
1,3	0,40320	0,40490	0,40658	0,40824	0,40988	0,41149	0,41309	0,41466	0,41621	0,41774
1,4	0,41924	0,42073	0,42220	0,42364	0,42507	0,42647	0,42786	0,42922	0,43056	0,43189
1,5	0,43319	0,43448	0,43574	0,43699	0,43822	0,43943	0,44062	0,44179	0,44295	0,44408
1,6	0,44520	0,44630	0,44738	0,44845	0,44950	0,45053	0,45154	0,45254	0,45352	0,45449
1,7	0,45543	0,45637	0,45728	0,45818	0,45907	0,45994	0,46080	0,46164	0,46246	0,46327
1,8	0,46407	0,46485	0,46562	0,46638	0,46712	0,46784	0,46856	0,46926	0,46995	0,47062
1,9	0,47128	0,47193	0,47257	0,47320	0,77381	0,47441	0,47500	0,47558	0,47615	0,47670
2,0	0,47725	0,47778	0,47831	0,47882	0,47932	0,47982	0,48030	0,48077	0,48124	0,48169
2,1	0,482214	0,48257	0,48300	0,48341	0,48382	0,48422	0,48461	0,48500	0,48537	0,48574
2,2	0,48610	0,48645	0,48679	0,48713	0,48745	0,48778	0,48809	0,48840	0,48870	0,48899
2,3	0,48928	0,48956	0,48983	0,49010	0,49036	0,49061	0,49086	0,49111	0,49131	0,49158
2,4	0,49180	0,49202	0,49224	0,49245	0,49266	0,49286	0,49305	0,49324	0,49343	0,49361
2,5	0,49379	0,49396	0,49413	0,49430	0,49446	0,49461	0,49477	0,49492	0,49506	0,49520
2,6	0,49534	0,49547	0,49560	0,49573	0,49585	0,49598	0,49609	0,49621	0,49632	0,49643
2,7	0,49653	0,49664	0,49674	0,49683	0,49693	0,49702	0,49711	0,49720	0,49728	0,49736
2,8	0,49744	0,49752	0,49760	0,49767	0,49774	0,47781	0,49788	0,49795	0,49801	0,49807
2,9	0,49813	0,49819	0,49825	0,49831	0,49836	0,49841	0,49846	0,49851	0,49856	0,49861

3,0	0,49865	*3,1*	0,49903	*3,2*	0,49931	*3,3*	0,49952	*3,4*	0,49966
3,5	0,49977	*3,6*	0,49984	*3,7*	0,49989	*3,8*	0,49993	*3,9*	0,49995
4,0	0,499968	*4,5*	0,499997	*5,0*	0,49999997				

Таблица 4

Критические точки распределения χ^2

Число степеней свободы k	Уровень значимости α					
	0,01	0,025	0,05	0,95	0,975	0,89
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	25,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	0,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

Таблица 5

Критические точки распределения Стьюдента

Число степеней свободы k	Уровень значимости α (двусторонняя критическая область)					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,3	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,96
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29
	0,05	0,025	0,01	0,005	0,001	0,0005
	Уровень значимости α (односторонняя критическая область)					

Критические точки распределения Фишера-Снедекора

									Уровень значимости α=0,05										
$k_2 \backslash k_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	240	242	244	246	248	249	250	251	252	253	254
2	18,5	19,0	19,2	19,2	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	23,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	211	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	182	1,77	1,71	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,93	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	3,84	3,00	3,60	2,37	2,21	2,10	2,01	1,94	1,83	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00
									Уровень значимости α=0,01										
1	4052	4999,5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,42	99,43	99,45	99,46	99,47	99,47	,99,48	99,49	99,50
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,05	26,87	26,69	26,60	26,50	26,41	26,32	26,22	26,13
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,37	14,20	14,02	13,93	13,84	13,75	13,65	13,56	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11	9,02
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,97	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,95	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,40	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4,00	3,91
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,40	4,25	4,10	4,02	3,94	3,86	3,78	3,69	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,78	3,70	3,62	3,54	3,45	3,36
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,82	3,66	3,59	3,51	3,43	3,34	3,25	3,17
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,80	3,66	3,51	3,43	3,35	3,27	3,18	3,09	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,55	3,41	3,26	3,18	3,10	3,02	2,93	2,84	2,75
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,46	3,31	3,16	3,08	3,00	2,92	2,83	2,75	2,62
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,37	3,23	3,08	3,00	2,92	2,84	2,75	2,66	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,30	3,15	3,00	2,92	2,84	2,76	2,67	2,58	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,52	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,17	3,03	2,88	2,80	2,72	2,64	2,55	2,46	2,36
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,98	2,83	2,75	2,67	2,58	2,50	2,40	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,93	2,78	2,70	2,62	2,54	2,45	2,35	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,31	2,21
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,99	2,85	2,70	2,62	2,54	2,456	2,36	2,27	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,96	2,81	2,66	2,58	2,50	2,42	2,33	2,23	2,13
27	7,68	5,49	4,60	4,11	3,789	3,56	3,39	3,26	3,15	3,06	2,93	2,789	2,63	2,55	2,47	2,38	2,29	2,20	2,10
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,90	2,75	2,60	2,52	2,44	2,35	2,26	2,17	2,06
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,87	2,73	2,57	2,49	2,41	2,33	2,23	2,14	2,03
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,84	2,70	2,55	2,47	2,39	2,30	2,21	2,11	2,01
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,66	2,52	2,37	2,29	2,20	2,11	2,02	1,92	1,80
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,73	1,60
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,34	2,19	2,03	1,95	1,86	1,76	1,66	1,53	1,38
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,32	1,00

Список литературы

1. Основная литература

1) Гмурман, В.Е. Теория вероятностей и математическая статистика: учеб. пособие / В.Е. Гмурман. – 12-е издание., перераб. – М.: Высш. шк., 2008. – 479 с.: ил.
2) Кремер, Н.Ш. Теория вероятностей и математическая статистика : учебник / Н.Ш. Кремер. – 2-е изд., перераб. И доп. – М.: ЮНИТИ, 2006. – 576 с.
3) Бочаров, П.П. Теория вероятностей и математическая статистика : учебник / П.П. Бочаров, А.В. Печенкин. – М.: Физматлит, 2005. – 295 с.

2. Дополнительная литература

1) Вентцель, Е.С. Теория вероятностей и ее инженерные приложения: учебное пособие / Е.С. Вентцель, Л.А. Овчаров. – 3-е изд., перераб. И доп. – М.: Академия, 2003. – 464 с.: ил.
2) Белько, И.В. Теория вероятностей и математическая статистика. Примеры и задачи : учеб. пособие / И.В. Белько, Г.П. Свирид. – Минск: ООО «Новое знание», 2002. – 250 с.
3) Балдин, К.В. Теория вероятностей и математическая статистика : учебник [Электронный ресурс] / Колемаев В.А. – М.: Юнити-Дана, 2010. – 353 с.
Теория вероятностей и математическая статистика : учеб. пособие: для студентов высших учебных заведений, обучающихся по направлению «Экономика» и другим экономическим специальностям / М-во образования Рос. Федерации, Рос. экон. акад. им. Г.В. Плеханова; [Л.Г. Бирюкова и др.], под ред. В.И. Ермакова. – М.: ИНФРА-М, 2008. – 286 с. : ил.

Учебное издание

Ряттель Александра Владимировна

Теория вероятностей и математическая статистика

Учебное пособие (лекционный материал)

Подписано в печать . Печать цифровая. Бумага для офисной техники.

Усл. Печ. л. 11,13. Тираж 100 экз. Заказ

Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Вятский государственный университет»

610000, Киров, ул. Московская, 36, тел.: (8332) 64-23-56, http://vytsu.ru