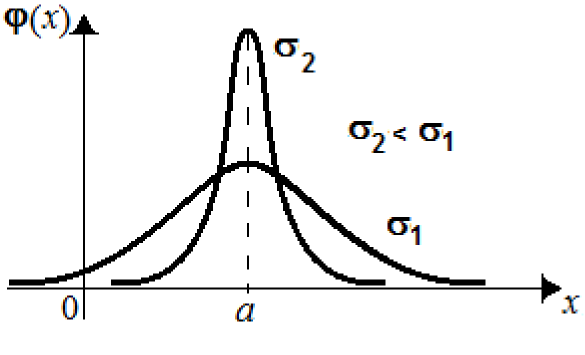
С.П. Горбиков, Л.В. Филатов

ЛЕКЦИИ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ



Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Нижегородский государственный архитектурно-строительный университет»

С.П. Горбиков, Л.В. Филатов

ЛЕКЦИИ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

Курс лекций

Утверждено редакционно-издательским советом университета
в качестве учебного пособия

Нижний Новгород
ННГАСУ
2011

УДК 519.2 (076.5)

Рецензенты:

Сморкалова В.М. - к.т.н., доцент кафедры Прикладной теории вероятностей факультета ВМК ННГУ им. Н.И. Лобачевского

Стребуляев С.Н. - к. т. н., доцент, с.н.с. НИИ ПМК

Горбиков С.П., Филатов Л.В. Лекции по теории вероятностей и математической статистике. [Текст]: учебное пособие для вузов./ Горбиков С.П., Филатов Л.В.; Нижегор. Гос. Архитектур.- строит. ун-т – Н.Новгород: ННГАСУ, 2011.-105с.

Рассматриваются основные положения теории вероятностей и математической статистики. Приводится большое количество примеров и иллюстраций, поясняющих теоретический материал. Пособие может использоваться как преподавателями соответствующего курса, так и студентами для самостоятельной работы.

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Введение

Теория вероятностей, как и любая другая наука, возникла из потребностей практики. Интуитивно понятие вероятного и случайного всегда связываются с неоднозначностью и непредсказуемостью наблюдаемых явлений. Так, например, при бросании монеты невозможно предсказать упадет она орлом или решеткой. Легче рассчитать движение светил небесных, чем ответить на этот вопрос! Такая непредсказуемость явления определяется тем, что имеется множество объективных и субъективных причин, учесть которые при исследовании не представляется возможным. Однако когда подобные явления наблюдаются в массовом порядке, оказывается, что они часто подчиняются определенным закономерностям, называемым статистическими. Изучение закономерностей мира случайных явлений и составляет предмет теории вероятностей.

Теория вероятностей строгая математическая наука, отказываясь от детерминистических математических моделей, свойственных предсказуемым явлениям, она строит и использует при изучении случайных явлений свои специфические вероятностные модели. Обширной частью современной теории вероятностей является математическая статистика, наука о методах наблюдения и обработки результатов массовых явлений, в которых фактор случайности имеет немаловажное значение.

Становление теории вероятностей связано с трудами Б. Паскаля, П.Ферма, Я. Бернулли в XVII в. и их попытками проведения расчетов в азартных играх, поэтому игровые модели чрезвычайно популярны при изложении теории. Дальнейшее развитие теория вероятностей получила в XVIII века в трудах К. Гаусса, П. Лапласа, С. Пуассона в связи с широким применением математических методов анализа. В XIX веке русские математики П.Л. Чебышев, А.А. Марков, А.М. Ляпунов провели обоснование вероятностного метода, доказав ряд предельных теорем. В дальнейшем теория вероятностей получила развитие в работах Н. Винера, Р. Фишера, А.Н. Колмогорова и ряда других ученых XX века.

Лекция № 1

Предмет теории вероятностей

1. События, частота и вероятность

На практике часто встречаются ситуации, результат которых трудно спрогнозировать.

Пример. То, что застрахованный дом пострадает или будет уничтожен в течение некоторого периода времени – дело случая. Но страховой орган должен рассчитывать сумму страхового взноса за этот период.

Пример. Сколько времени будет идти маршрутка с площади Комсомольской до площади Горького – дело случая. Но Вы должны рассчитать время своего приезда.

Впервые такими ситуациями занялись математики в середине XVII века. Возникновение теории вероятностей связано с именами Гюйгенса (1629-1695), Паскаля (1623-1662), Ферма (1601-1665) и Якоба Бернулли (1654-1705), одного представителя из многочисленного математического клана Бернулли (31 математик). При этом лишь азартные игры были главным побудительным моментом в их деятельности.

Далее мы будем иметь дело со случайными событиями, поэтому нужно определить это понятие.

Определение. Событие *– исход некоторого опыта.*

Определение. Случайное событие *– то, что может произойти либо не произойти.*

Естественнее всего случайные события характеризовать следующим понятием.

Определение. Относительной частотой p^* *случайного события* A *называется отношение числа* m^* *появлений данного события к общему числу* n^* *проведённых испытаний, в каждом из которых может появиться или нет данное событие:*

$$p^* = p^*(A) = \frac{m^*}{n^*}\ .$$

Чаще всего оказывается (по крайней мере, теория вероятностей имеет дело именно с такими частотами, а иные ситуации в ней **не рассматриваются**), что:

$$p^* = p^*(A) = \frac{m^*}{n^*} \underset{n^*\to+\infty}{\longrightarrow} p = p(A)\,,$$

где p - некоторое число.

Определение. *Такое число* p *называется* ***вероятностью*** *появления случайного события* A.

Пример. Наблюдение броуновского движения (хаотического движения мельчайших частиц вещества, взвешенных в жидкости). Хаос здесь объясняется ударами молекул жидкости. Кинетическая теория газов даёт возможность подсчитать вероятность того, что в данном объёме жидкости не будет ни одной частицы, будет $1,2,3,\ldots$ частицы.

Для проверки теории проводились эксперименты. Шведский учёный Сведберг провёл 518 экспериментов. В подвергшейся наблюдению части пространства: 112 раз не было частиц, 168 раз была одна частица, 130 раз было две частицы, 69 раз было три частицы, 32 раза было четыре частицы, 5 раз было пять частицы, 1 раз было шесть частиц, 1 раз было семь частиц. Таким образом, он составил таблицу относительных частот:

$$p^*(0) = \frac{m^*(0)}{n^*} = \frac{112}{518} \approx 0{,}216,$$

$$p^*(1) = \frac{168}{518} \approx 0{,}324,$$

$$p^*(2) = \frac{130}{518} \approx 0{,}251,$$

$$p^*(3) = \frac{69}{518} \approx 0{,}133,$$

$$p^*(4) = \frac{32}{518} \approx 0{,}062,$$

$$p^*(5) = \frac{5}{518} \approx 0{,}0104,$$

$$p^*(6) = p^*(7) = \frac{1}{518} \approx 0{,}002.$$

Результаты наблюдений показали хорошее совпадение с теоретически предсказанными вероятностями.

Очень часто необходимо предсказывать характер протекания многих процессов, т.е. находить вероятности некоторых сложных событий, хотя мы можем определить частоты (а в пределе - вероятности) некоторых простых событий. Например, необходимо определить (с высокой степенью достоверности) поражение мишени хотя бы одним выстрелом из трёх произведённых, хотя мы легко можем определить вероятность попадания в мишень при одном выстреле.

Тогда строят модель таким образом. Полагают известной вероятность события A:

$$p(A) \approx p^*(A) = \frac{m^*}{n^*}$$

и с помощью определённых процедур находят вероятности нужных случайных событий.

Определение вероятности появления события по вероятностям элементарных событий, изучение вероятностных закономерностей (различных случайных событий) и является предметом теории вероятностей.

Укажем на одно важное свойство вероятности случайных событий. Поскольку $0 \le m^* \le n^*$, а $p = \lim (m^* / n^*)$, то *вероятность случайного события измеряется в долях единицы, т.е.* $0 \le p \le 1$,

2. Классификация событий

Рассмотрим простейший пример, который мы будем изучать с разных сторон в следующих двух параграфах первой лекции.

Пример № 1. Бросили игральную (шестигранную) кость (один раз). Найти вероятность того, что выпадет: 1) «6»; 2) чётное число; 3) нечётное число; 4) число, меньшее «5».

Прелюдия к решению. Рассмотрим следующие элементарные события (возможно, на их основе представим нужные нам события):

A_1 - бросили игральную кость и выпала «1»;

A_2 - бросили игральную кость и выпала «2»;

A_3 - бросили игральную кость и выпала «3»;

A_4 - бросили игральную кость и выпала «4»;

A_5 - бросили игральную кость и выпала «5»;

A_6 - бросили игральную кость и выпала «6».

Теперь легко представить, что:

1) событие A, состоящее в том, что бросили игральную кость и выпала «6», есть событие A_6, т.е.

$$A = A_6;$$

2) событие B, состоящее в том, что бросили игральную кость, а выпало чётное число, представляет собой множество, состоящее из трёх событий,

$$B = \{A_2, A_4, A_6\};$$

3) событие C, состоящее в том, что бросили игральную кость, а выпало нечётное число, представляет собой множество, состоящее из трёх событий,

$$C = \{A_1, A_3, A_5\};$$

4) событие D, состоящее в том, что бросили игральную кость, а выпало число, меньшее «5», представляет собой множество, состоящее из четырёх событий,

$$D = \{A_1, A_2, A_3, A_4\}.$$

Чтобы научиться находить вероятности сложных событий, нужно провести их классификацию и научиться проводить операции над ними.

Определение. Сумма $A_1 + A_2 + A_3 + \ldots + A_n$ *конечного числа событий* $A_1, A_2, A_3, \ldots, A_n$ *– событие, состоящее в наступлении хотя бы одного из них.*

Пример. В примере № 1 событие C равно сумме событий $C = A_1 + A_3 + A_5$.

Определение. Произведение $A_1 \cdot A_2 \cdot A_3 \cdot \ldots \cdot A_n$ *конечного числа событий* $A_1, A_2, A_3, \ldots, A_n$ *– событие, состоящее в наступлении всех этих событий.*

Пример. В примере № 1 событие $A_2 + A_4$ есть произведение событий B и D: $A_2 + A_4 = B \cdot D = BD$ (математики экономят на знаке произведения).

Определение. Противоположным *событием* $\overline{A}$ *называется событие, состоящее в не появлении события* A.

Пример. В примере № 1 событие B есть противоположное к событию C: $B = \overline{C}$.

Рассмотрим важные для дальнейшего понятия.

Определение. *Два события называются* ***несовместными****, если наступление одного из них исключает возможность наступления другого. В противоположном случае события называются* ***совместными.***

Пример. В примере № 1 события B и C - несовместные, а события A и B - совместные.

Определение. *События называются* ***равновозможными (равновероятными)****, если вероятность наступления каждого из них одна и та же.*

Пример. В примере № 1 события B и C являются равновозможными, если кость сделана без изъянов. Также следует признать равновозможными и события $A_1, A_2, A_3, \ldots, A_6$.

Определение. *События называются* ***элементарными,*** *если их наступление нельзя связать с наступлением других событий в этом опыте.*

Пример. Извлечение карты «Дама пик» из перемешанной колоды карт – событие элементарное.

Определение. *События называются* ***сложными,*** *если их наступление в опыте можно связать с наступлением других событий в этом опыте.*

Пример. Извлечение «пиковой карты» из перемешанной колоды карт – событие сложное, так как его наступление связано с рядом событий в этом опыте, а именно, извлечение «Туз пик», «Король пик», …

Определение. *События образуют* ***полную группу****, если при каждом испытании может появиться любое из них и не может появиться какое-либо иное (отличное от входящих в группу) событие.*

Пример. В примере № 1 события B и C образуют такую полную группу, если не учитывать, что кость при бросании может встать на ребро, исчезнуть (провалиться под пол), …

Определение. *Событие называется* ***достоверным****, если оно не может не произойти в условиях данного опыта.*

Вероятность достоверного события равна 1, т.к. для этого события $m^* = n^*$ (напомним, что $p = \lim \frac{m^*}{n^*}$).

Пример. В примере № 1 событие $A_1 + A_2 + A_3 + \ldots + A_6$ есть как раз такое достоверное событие.

Определение. *Событие, которое не может произойти в условиях данного опыта, называется* ***невозможным*** *событием.*

Вероятность невозможного события равна 0, т.к. для этого события $m^* = 0$ (а $p = \lim \frac{m^*}{n^*}$).

Пример. В примере № 1 событие, равное произведению двух событий BC, является как раз невозможным событием. Невозможное событие представляет собой и событие, состоящее в выпадении $\frac{1}{3}$.

3. Классический способ нахождения вероятности

Пусть мы имеем **полную группу равновозможных, несовместных,** случайных событий.

Определение. *Событие (из такой группы) называется* ***благоприятствующим*** *появлению события A, если появление этого события (из такой группы) влечёт за собой появление события A.*

Пример. В примере № 1 событие B, выпало чётное число при бросании один раз игральной кости, имеет в качестве благоприятствующих событий, следующие события: A_2, A_4, A_6.

Собственно сам классический способ нахождения вероятности заключается в следующем (а как может быть по-другому?) простом соображении.

Вероятность события A равна отношению числа m благоприятствующих случайных событий к числу всех возможных случайных событий n, образующих полную группу равновозможных несовместных событий:

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

Исходя из приведённого правила, можно опять установить, что для событий, входящих в состав полной группы равновозможных несовместных событий, имеет место два свойства вероятности:

$$p(\textit{достоверное событие}) = \frac{n}{n} = 1,$$

$$p(\textit{невозможное событие}) = \frac{0}{n} = 0.$$

Теперь можно вернуться к примеру № 1 и предложить окончательное его решение.

Пример. В примере № 1 полную группу равновозможных несовместных событий составляют события $A_1, A_2, A_3, \ldots, A_6$, т.к.:

$A_1, A_2, A_3, \ldots, A_6$ образуют полную группу (об этом мы уже говорили),

понятно, что все эти события равновероятны (если кость сделана без изъянов),

понятно, что все эти события несовместны (если кость при бросании не упадёт на ребро). Поэтому $n = 6$.

Тогда событие A, состоящее в том, что бросили игральную кость и выпала «6», имеет вероятность:

$$p(A) = \frac{m}{n} = \frac{1}{6},$$

т.к. благоприятствующим событием является лишь событие A_6, т.е. $m = 1$.

Событие B, состоящее в том, что бросили игральную кость, а выпало чётное число, имеет вероятность:

$$p(B) = \frac{m}{n} = \frac{3}{6} = 0{,}5,$$

т.к. благоприятствующими событиями являются события A_2, A_4, A_6, т.е. $m = 3$.

Событие C, состоящее в том, что бросили игральную кость, а выпало нечётное число, имеет вероятность:

$$p(C) = p(B) = 0{,}5,$$

о чём мы уже говорили.

Событие D, состоящее в том, что бросили игральную кость, а выпало число, меньшее «5», имеет вероятность:

$$p(D) = \frac{m}{n} = \frac{4}{6} = \frac{2}{3},$$

т.к. благоприятствующими событиями являются события A_1, A_2, A_3, A_4, т.е. $m = 4$.

4. Графическая интерпретация событий

Пусть в опыте мы имеем полную группу равновозможных случайных событий $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$. Если при этом события ω_i являются элементарными, то такую группу событий называют *пространством элементарных событий*. Любое сложное событие A в опыте всегда связано с рядом элементарных событий из Ω, которые являются благоприятствующими для его наступления $A=\{\omega_{A1},\omega_{A2},\ldots,\omega_{Ak}\}\subset\Omega$. Тогда элементарное событие может быть изображено графически точкой в пространстве Ω, а сложное событие множеством A в пространстве Ω [1-4]. Такое изображение событий представляется диаграммой Эйлера-Вена на рис.1.1.

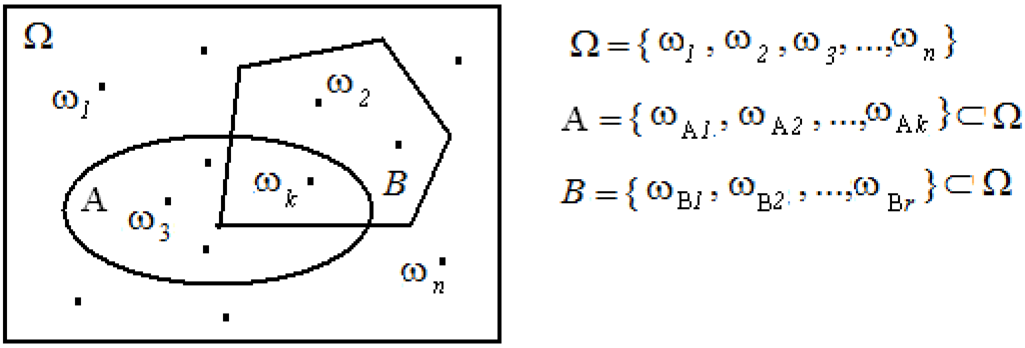


Рис.1.1. Диаграмма Эйлера-Вена для интерпретации событий в пространстве элементарных событий и операций над событиями

Такая интерпретация события, как множества в пространстве событий, позволяет легко и наглядно изображать события, операции над событиями как операции над множествами, понять соотношения алгебры событий:

$$A+\overline{A}=\Omega,\qquad A\overline{A}=\otimes,\qquad A(B+C)=AB+AC,\qquad \overline{AB}=\overline{A}+\overline{B},\qquad \overline{A+B}=\overline{A}\,\overline{B}.$$

Если под массой (модулем) события $|A|$ понимать число, характеризующее количество благоприятствующих элементарных исходов, то классический способ вычисления вероятности интерпретируется как отношение массы события к массе пространства:

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

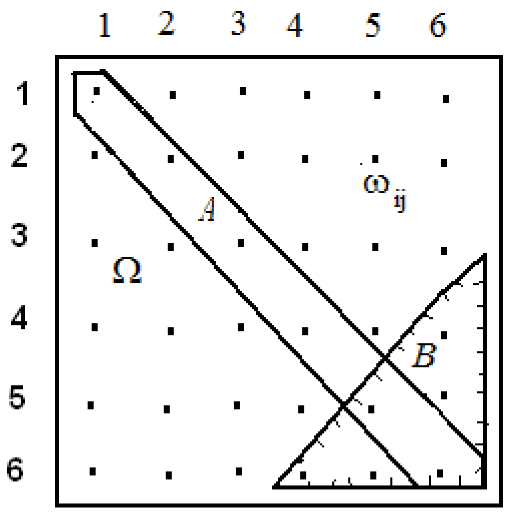


Рис.1.2. Изображение событий A и B в пространстве Ω

Пример. Пусть в опыте бросаются две игральные кости. Событие A состоит в выпадении дубля, а событие B - в выпадении суммы очков на обеих костях не менее 10. Эти события изображены на рис. 1.2. множествами, где точками изображаются элементарные исходы $\{\omega_{ij}\}$, тогда:

$$\Omega=\{\omega_{ij}\},\qquad |\Omega|=36,$$

$$A=\{\omega_{11},\omega_{22},\omega_{33},\omega_{44},\omega_{55},\omega_{66}\},\ \ |A|=6,$$

$$B=\{\omega_{66},\omega_{65},\omega_{56},\omega_{55},\omega_{64},\omega_{46}\},\ \ |B|=6,$$

$$P(A)=|A|/|\Omega|=1/6,\ \ P(B)=|B|/|\Omega|=1/6.$$

Лекция № 2

Вычисление вероятности событий

1. Элементы комбинаторики и вычисление вероятности событий

Необходимые сведения из комбинаторики [5,6] изучим на простейших примерах.

Пример 1. При игре в русское лото из мешка поочерёдно извлекают все 90 бочонков (с различной нумерацией). Найти вероятность того, что бочонки извлекут в порядке убывания нумерации.

Решение. Здесь Ω, множество всех равновозможных несовместных событий, образующих полную группу, представляет собой:

$$\Omega = \{(1,2,3,4,\ldots,90), (2,1,3,4,\ldots,90), (1,3,2,4,\ldots,90), \ldots, (90,89,88,\ldots,2,1)\},$$

где $(i_1, i_2, \ldots, i_{90})$ обозначает комбинацию чисел $1,2,\ldots,90$, указанных на бочонках, извлечённых из мешка один за другим в результате какого-то опыта. При этом порядок, в котором следуют числа $i_1, i_2, \ldots, i_{90}$, имеет существенное значение!

Договоримся, что эти комбинации отличаются друг от друга хотя бы одним числом, стоящим на соответствующем месте. Понятно, что Ω - полная группа (т.к. все возможные комбинации чисел от 1 до 90 здесь поименованы) равновозможных (т.к. нет предпочтения ни одной комбинации перед другими) несовместных (т.к. одновременно обе различные комбинации появиться не могут) событий.

Тогда, если мы найдём число n всех комбинаций во множестве Ω, то нужная нам вероятность есть:

$$p = \frac{m}{n} = \frac{1}{n},$$

т.к. число благоприятствующих комбинаций равно единице (комбинация $(90,89,\ldots,2,1)$, и только она, ибо порядок чисел имеет значение).

А число n найти просто. Поскольку порядок чисел имеет значение, постольку при первом извлечении бочонка у нас всего 90 возможностей, при втором – 89 (т.к. один бочонок уже извлечён из мешка). При подсчёте числа n между этими числами нужно поставить знак умножить, т.к. на всякое i_1 найдётся 89 возможностей i_2. И так далее до предпоследнего извлечения бочонка, когда останется только 2 возможности. Поэтому:

$$n = 90 \cdot 89 \cdot 88 \cdot \ldots \cdot 3 \cdot 2 = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot 88 \cdot 89 \cdot 90.$$

В этом последнем виде и определено в комбинаторике число

$$N! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (N-1) \cdot N,$$

носящее название « N – факториал». Оно представляет собой число всех возможных комбинаций из N чисел, расставленных по N местам, при этом порядок, занимаемый числами, имеет существенное значение.

Итак, искомая вероятность равна:

$$p = \frac{1}{90!},$$

т.к. порядок, в котором следуют числа во всевозможных комбинациях, имеет существенное значение.

Величина этого числа, стоящего в знаменателе, огромна: $90! > 10^{81}$, т.к. $90 > 10$, $89 > 10$, $88 > 10, \ldots, 10 \geq 10$. Поэтому встретиться на практике с такой комбинацией невероятно (вероятность такой встречи практически равна нулю)!

Пример 2. При игре в русское лото из мешка поочерёдно извлекают (на сей раз) 86 бочонков (с различной нумерацией). Найти вероятность того, что бочонки появятся в строго убывающем порядке, начиная с бочонка под номером 90 (точнее появятся в таком порядке: 90, 89, 88, …, 6, 5).

Решение. Здесь Ω - множество всех равновозможных несовместных событий, образующих полную группу - представляет собой:

$$\Omega = \{(1,2,3,4,\ldots,85,86), (2,1,3,4,\ldots,85,86), (1,3,2,4,\ldots,85,86), \ldots, (90,89,88,\ldots,6,5)\},$$

где $(i_1, i_2, \ldots, i_{86})$ обозначает комбинацию чисел $i_1, i_2, \ldots, i_{86}$, указанных на бочонках, извлечённых из мешка один за другим в результате какого-то опыта. При этом порядок, в котором следуют числа $i_1, i_2, \ldots, i_{90}$, опять имеет существенное значение!

Снова договоримся, что эти комбинации отличаются друг от друга хотя бы одним числом, стоящим на соответствующем месте. Понятно, что Ω - полная группа (т.к. все возможные комбинации чисел от 1 до 90 здесь поименованы) равновозможных (т.к. нет предпочтения ни одной комбинации перед другими) несовместных (т.к. одновременно обе различные комбинации появиться не могут) событий.

Тогда, если мы найдём число n всех комбинаций во множестве Ω, то нужная нам вероятность есть:

$$p = \frac{m}{n} = \frac{1}{n},$$

т.к. число благоприятствующих комбинаций равно единице (комбинация $(90,89,\ldots,6,5)$, и только она). Порядок чисел фиксирован!

А число n найти по-прежнему просто. При первом извлечении бочонка у нас всего 90 возможностей, при втором – 89 (т.к. один бочонок уже извлечён из мешка). При подсчёте числа n между этими числами нужно поставить знак умножить, т.к. на всякое i_1 найдётся 89 возможностей i_2. И так далее до последнего извлечения бочонка, когда останется только 5 возможностей. Поэтому:

$$n = 90 \cdot 89 \cdot 88 \cdot \ldots \cdot 6 \cdot 5 = \frac{90!}{4!} = \frac{90!}{(90-86)!} = A_{90}^{86},$$

где A_{90}^{86} называется числом размещений.

Итак, мы познакомились с ещё одним числом, имеющим большое значение для комбинаторики (да и для нас тоже)!

Числом размещений A_k^l *называется частное от деления:*

$$A_k^l = \frac{k!}{(k-l)!} \quad (k \ge l).$$

Оно представляет собой число всех возможных комбинаций из k чисел, расставленных по l местам, при этом порядок, занимаемый числами, имеет существенное значение.

Поэтому искомая вероятность равна:

$$p = \frac{1}{A_{90}^{86}} = \frac{1}{5 \cdot 6 \cdot \ldots \cdot 89 \cdot 90}.$$

Величина этого числа, стоящего в знаменателе, по-прежнему огромна. Поэтому встретиться на практике с такой комбинацией невероятно!

Пример 3. Найти вероятность угадать в лотерее «6 из 49» (когда извлекают 6 чисел из различных (!) 49 чисел) при заполнении одного варианта:

1) все шесть номеров;
2) три номера.

Решение. Займёмся сначала решением первой задачи. При заполнении одного варианта выбирают 6 чисел из чисел, следующих друг за другом, от 1 до 49. При этом *порядок, в котором указаны числа в выбранном для игры варианте, не имеет значения*!

Поэтому Ω состоит из групп комбинаций, а каждая группа составлена из комбинаций всевозможных наборов заранее определённых 6 чисел $(i_1, i_2, \ldots, i_6)$. Проще говоря, набор всех возможных комбинаций (а всего их $6!$, как следует из только что разобранного примера)

$$\{(i_1, i_2, i_3, \ldots, i_6), (i_2, i_1, i_3, \ldots, i_6), (i_1, i_3, i_2, \ldots, i_6), \ldots, (i_6, i_5, i_3, \ldots, i_1)\}$$

и составляет одну такую группу. А из этих групп и составлено в свою очередь множество Ω.

Но как подсчитать число n всевозможных групп? Понятно, что это n есть частное от деления числа A_{49}^6 (т.е. числа всех возможных комбинаций из 49 чисел, расставленных по 6 местам, при этом порядок имеет существенное значение) на число $6!$ (т.е. число всех возможных комбинаций из 6 чисел, расставленных по 6 местам, при этом порядок имеет существенное значение):

$$n = \frac{A_{49}^6}{6!} = \frac{\frac{49!}{(49-6)!}}{6!} = \frac{49!}{6!\,(49-6)!} = C_{49}^6,$$

которое носит название «число сочетаний из 49 по 6 местам».

Итак, мы пришли к понятию ещё одного важного числа для комбинаторики.

Числом сочетаний из k *элементов по* l *элементам называется число:*

$$C_k^l = \frac{k!}{l!(k-l)!} \qquad (k \ge l),$$

обозначающее число способов, которыми можно расположить k чисел по l местам (при этом порядок, занимаемый числами, не имеет значения).

Итак, чтобы найти искомую вероятность, нужно $m=1$ (т.к. число благоприятствующих событий равно единице) поделить на только что найденное n. Поэтому искомая вероятность равна:

$$p=\frac{1}{C_{49}^{6}}=\frac{6!\cdot 43!}{49!}=\frac{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 1\cdot 2\cdot\ldots\cdot 43}{1\cdot 2\cdot\ldots\cdot 43\cdot 44\cdot 45\cdot 46\cdot 47\cdot 48\cdot 49}=\frac{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6}{44\cdot 45\cdot 46\cdot 47\cdot 48\cdot 49}=\frac{1}{13983816}.$$

Это значит, что просто так, без каких-то ухищрений, выиграть в эту игру нельзя: «выигрывает одна из 14 миллионов попыток».

Перейдём теперь к решению второй задачи. Для этого осталось подсчитать число m (ибо число n только что подсчитано). Но что значит угадать «три номера из шести»? Это означает «три угадали, а три в указанном варианте не угадали». А такая комбинация означает, что в ней три номера из шести указаны правильно (порядок чисел в указанном варианте не имеет значения, чему соответствует число C_6^3), а три неправильно (порядок чисел в указанном варианте по-прежнему не имеет значения, чему соответствует число $C_{49-6}^3=C_{43}^3$). А между этими числами нужно поставить знак умножения, т.к. на всякое i_1 из возможностей C_6^3 найдётся одна из возможностей C_{43}^3. Поэтому

$$m=C_6^3C_{43}^3.$$

Отсюда, искомая (во второй раз) вероятность равна:

$$p=\frac{m}{n}=\frac{C_6^3C_{43}^3}{C_{49}^6}=\frac{\frac{6!}{3!\,3!}\,\frac{43!}{3!\,40!}}{\frac{49!}{6!\,43!}}=\frac{\frac{2\cdot 3\cdot 4\cdot 5\cdot 6}{2\cdot 3\cdot 2\cdot 3}\,\frac{41\cdot 42\cdot 43}{2\cdot 3}}{\frac{44\cdot 45\cdot 46\cdot 47\cdot 48\cdot 49}{2\cdot 3\cdot 4\cdot 5\cdot 6}}=\frac{\frac{4\cdot 5}{1}\,\frac{41\cdot 42\cdot 43}{2\cdot 3}}{\frac{44\cdot 45\cdot 46\cdot 47\cdot 48\cdot 49}{2\cdot 3\cdot 4\cdot 5\cdot 6}}=$$

$$=\frac{4\cdot 5\cdot 41\cdot 42\cdot 43}{\frac{44\cdot 45\cdot 46\cdot 47\cdot 48\cdot 49}{4\cdot 5\cdot 6}}=\frac{4\cdot 5\cdot 41\cdot 42\cdot 43}{11\cdot 9\cdot 46\cdot 47\cdot 8\cdot 49}\approx 0{,}01765.$$

Проверим практикой полученный результат (ибо «практика – критерий истины»). Возьмём наугад результат какого-нибудь тиража лотереи «6 из 49». В 406 тираже, состоявшемся в 2004 году, всего было сыграно 46283 вариантов ($n^*=46283$). Из них было угадано «три номера из шести» в 685 вариантах ($m^*=685$). Частота этого события равна:

$$p^*=\frac{m^*}{n^*}=\frac{685}{46283}\approx 0{,}01480.$$

О лучшем (совпадении) трудно было бы и мечтать: вероятность почти одинакова с частотой!

Схема урн. Отметим, что рассмотренная выше задача описывает так называемую «схему урн» (рис 2.1), состоящую в следующем.

Пусть в урне тщательно перемешаны шары, отличающиеся только цветом и пусть, например, белых там N_1, а черных N_2. Наугад из урны извлекаются n шаров. Какова вероятность события A, состоящего в том, что среди извлеченных будет n_1 белых и n_2 черных? Схема изображена ниже:

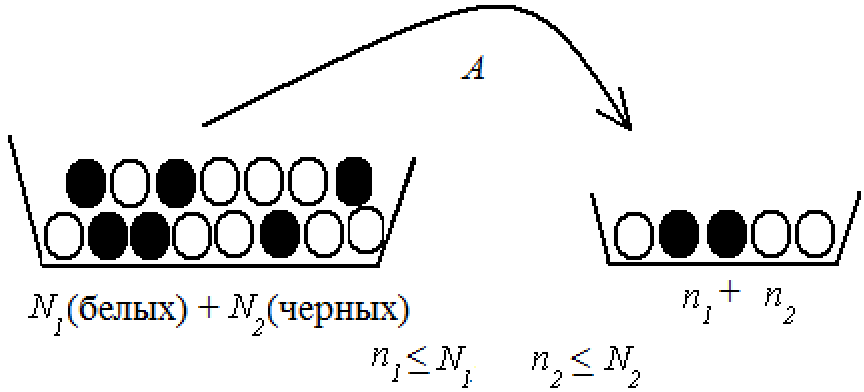


Рис. 2.1. Схема урн с белыми и черными шарами

Из вышеприведенной задачи понятными становится следующие формулы вероятности событий:

$$p(A)=\frac{C_{N_1}^{n_1}\cdot C_{N_2}^{n_2}}{C_{N_1+N_2}^{n_1+n_2}},\qquad p(A)=\frac{C_{N_1}^{n_1}\cdot C_{N_2}^{n_2}\cdot\ldots\cdot C_{N_m}^{n_m}}{C_{N_1+N_2+\cdots+N_m}^{n_1+n_2+\cdots+n_m}}.$$

Вторая из них для случая многоцветных шаров (белые, черные, синие и др.).

2. Геометрические вероятности

Это понятие касается следующего класса задач. Представим себе, что на плоскости расположены две области M и m, причем область m целиком распложена в области M. Их площади, соответственно, равны S_m и S_M. В область M наудачу бросают точку. Какова вероятность того, что точка попадёт также и в область m?

Если предположить, что точка может попасть в любую часть области M, а вероятность попадания в область m пропорциональна лишь её площади и не зависит ни от расположения m, ни от её формы, то искомая вероятность:

$$p=\frac{S_m}{S_M}.$$

Это и есть так называемое *«правило нахождения геометрической вероятности»* [7].

Аналогично могут быть определены вероятности попадания точки:

1) в объёмную область v величиной V_v, содержащуюся в объёмной области V величиной V_V, если точка брошена наугад в объём V:

$$p=\frac{V_v}{V_V};$$

2) на отрезок l величиной L_l, расположенный на отрезке L величиной L_L, если точка брошена наугад на отрезок L:

$$p=\frac{L_l}{L_L}.$$

Пример. Круглый диск радиуса R разбит на два сектора. Длина дуги одного из них (заштрихованного) равна радиусу R (рис. 2.2). По быстро вращающемуся диску произведён выстрел. Цель поражена. Найти вероятность того, что попали в заштрихованную часть.

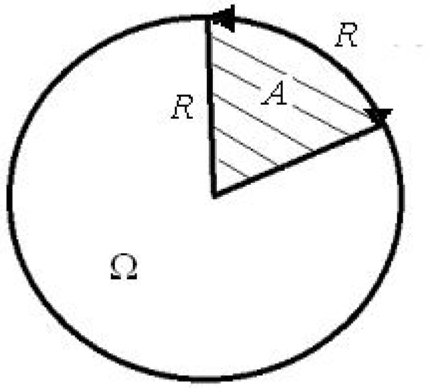


Рис. 2.2. Иллюстрация к задаче о попадании в сектор диска

Решение. Идеология решения задачи проста. Пусть событие A есть событие, состоящее в том, что попали именно в заштрихованную часть. Тогда искомая вероятность равна $P(A)=\dfrac{S_m}{S_M}$, где S_m - площадь заштрихованной части, S_M - площадь круга ($S_M=\pi R^2$).

Проблема лишь в том, как найти площадь заштрихованной части. Но S_m относится к S_M также, как длина дуги заштрихованной части ($L_m=R$) относится к длине круга ($L_M=2\pi R$): $p=\dfrac{S_m}{S_M}=\dfrac{L_m}{L_M}=\dfrac{R}{2\pi R}=\dfrac{1}{2\pi}$, что и требовалось найти.

Пример. Задача Бюффона (или задача об игле) [7]. Пусть на плоскость, разлинованную параллельными линиями с расстоянием $2a$, наудачу брошен отрезок (игла) длиной $2l<2a$. Какова вероятность пересечения линии иглой?

Событие A состоит в пересечении линии на плоскости. Игла пересекает только одну линию в силу ограничения $2l<2a$, или не пересекает ни одной. Пусть u - расстояние от центра иглы до ближайшей линии, а φ - угол наклона иглы к линиям. Тогда множество всех равновозможных событий $\Omega=\{0\le u\le a;\,0\le\varphi\le\pi\}$, а множество всех благоприятствующих исходов для события $A=\{0\le u\le l\sin\varphi;\,0\le\varphi\le\pi\}$ и оба эти множества изображены ниже на рисунке. Вероятность события A вычисляется как геометрическая:

$$S_\Omega=\pi a,\quad S_A=\int_0^\pi l\sin\varphi\cdot d\varphi=2l,\quad P(A)=\frac{S_A}{S_\Omega}=\frac{2l}{\pi a}\ .$$

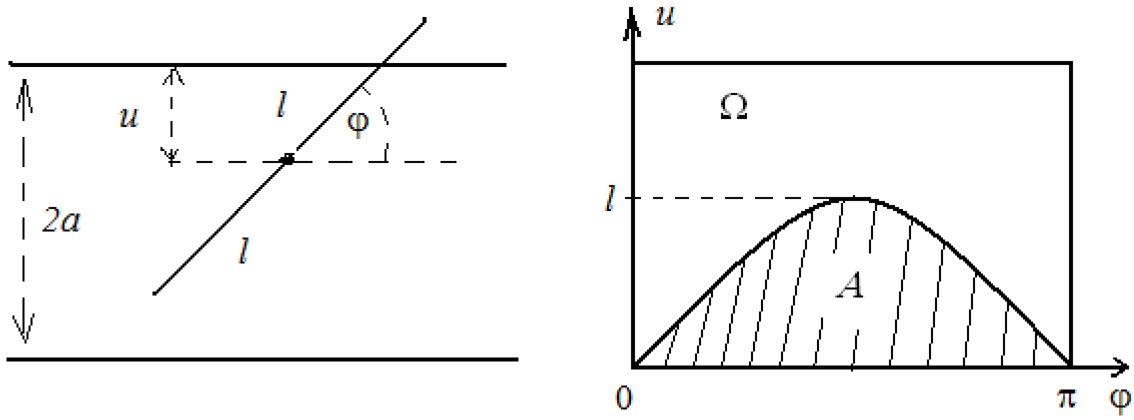


Рис 2.3. Иллюстрация к задаче Бюфона

Лекция № 3

Вероятности сложных событий

Часто возникает ситуация, когда вероятность искомого события может быть вычислена через известные вероятности ряда более простых событий, наступление или отсутствие которых приводит к искомому событию.

1. Определение условной вероятности

Начнем с определения.

Определение. *Если* $P(A) > 0$, *то частное* $\dfrac{P(AB)}{P(A)}$ *называется* ***условной вероятностью*** *события* B *при условии* A *(или* ***условной вероятностью*** *события* B *при условии, что событие* A *произошло).*

Оно обозначается:

$$P_A(B) = P(B/A) = \frac{P(AB)}{P(A)}.$$

Смысл условной вероятности открывается из следующего рассуждения. Пусть рассматриваются геометрические вероятности. Событие A состоит в том, что бросаем точку на часть плоскости Ω и попадаем в фигуру A, а событие B - попадаем в фигуру B (см. рис. 3.1). Событие AB состоит в том, что бросаем точку и она попадает в общую часть фигур A и B (на рис. 3.1 эта часть забита точками). Тогда $\dfrac{P(AB)}{P(A)}$ характеризует, какую часть по отношению к части A (событию A) составляет часть AB (событие AB).

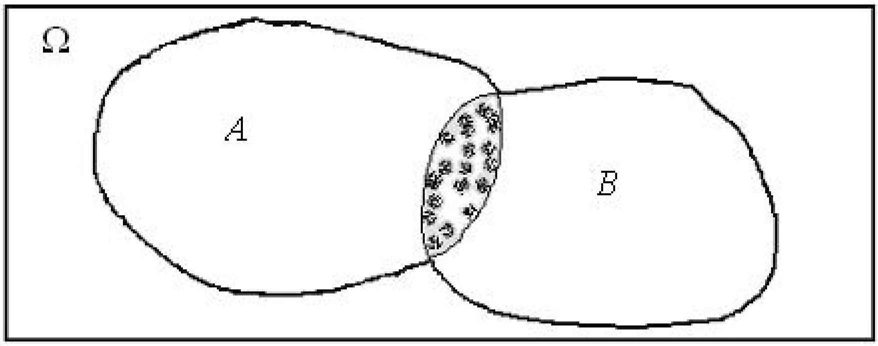


Рис 3.1. Иллюстрация понятия условной вероятности

Иными словами,

$$\frac{P(AB)}{P(A)} = P_A(B) = \frac{\textit{вероятность попасть в } AB}{\textit{вероятность попасть в } A} =$$

$$= \textit{вероятность того, что попали в } B \textit{ при условии, что находимся в } A.$$

Вывод из сказанного получается следующий: $P_A(B)$ действительно обозначает вероятность того, что B произойдёт при условии, что A произошло.

2. Независимость событий

Понятие «независимости» играет ключевую роль в теории вероятностей: оно выделило теорию вероятностей из теории меры (ибо в теории вероятностей находятся вероятности различных событий – суть измеряется мера определенного множества по сравнению с множеством единичной меры).

Однако перейдём к понятию независимости. Если A и B два события, то естественно сказать, что событие B не зависит от события A, если знание того, что свершилось событие A, никак не влияет на вероятность события B. Иначе говоря (при условии $P(A) > 0$),

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

По определению условной вероятности:

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

Поэтому

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

откуда

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

Последнее равенство и принято в теории вероятностей за определение независимости двух событий.

Итак, *два события A и B называются* ***независимыми****, если*

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

Прелесть этого определения ещё и в том, что оно годится и для случая, когда $P(A) = 0$ (в отличие от рассуждений в начале этого пункта).

Пример. Безотказная работа прибора определяется работой двух узлов, соединённых последовательно. Вероятность безотказной работы i-ого узла равна:

$$p_1 = 0{,}9,\ p_2 = 0{,}8.$$

Узлы работают независимо друг от друга. Какова вероятность безотказной работы всего прибора.

Решение. Введём следующие обозначения:

A - событие, состоящее в безотказной работе всего прибора;

A_i - событие, состоящее в безотказной работе i-ого узла прибора ($i = 1, 2$).

Тогда в силу «последовательности» соединения

$$A = A_1 A_2\ .$$

Поэтому

$$P(A) = P(A_1 A_2)\,,$$

а в силу независимости работы узлов прибора (вероятность произведения равна произведению вероятностей):

$$P(A) = P(A_1 A_2) = P(A_1)P(A_2) = p_1 p_2 = 0{,}9 \cdot 0{,}8 = 0{,}72.$$

Всякое последовательное соединение приводит к потере устойчивости в работе прибора!

3. Вероятность произведения событий

Это очень просто. Из определения условной вероятности (напишем определение наоборот):

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

следует, что вероятность произведения событий (в общем случае) равна

$$P(AB) = P(A)P(B/A)\,.$$

И всё! Новая формула готова!

Аналогично (от перемены букв в определении само определение не изменится!):

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

поэтому

$$P(AB) = P(B)P(A/B)\,.$$

Окончательно получается следующее утверждение.

Теорема умножения вероятностей. *Вероятность произведения двух событий равна произведению вероятности одного события на произведение условной вероятности другого события при условии, что первое событие произошло:*

$$P(AB) = P(A)P(B/A) = P(B)P(A/B)\,.$$

Приведём получаемую по индукции теорему об умножении конечного числа событий:

$$P(A_1 A_2 A_3 \dots A_{n-1} A_n) = P(A_1)P(A_2 / A_1)P\big(A_3/(A_1 A_2)\big)\dots P\big(A_n/(A_1 A_2 A_3 \dots A_{n-1})\big).$$

4. Теорема сложения вероятностей событий

Начнем с геометрической иллюстрации. Пусть рассматривается геометрическая вероятность в случае $n = 2$ (плоский случай). Событие A состоит в том, что бросаем точку на часть плоскости Ω и попадаем в фигуру A, а событие B - попадаем в фигуру B (см. рис. 3.2). Найдем вероятность того, что бросаем точку в область Ω и попадаем в фигуру $A \cup B$, т.е. забитую точками на рис. 3.2 фигуру. Эта фигура $A \cup B$ соответствует событию, состоящему в наступлении или события A или события B, т.е. события $A + B$.

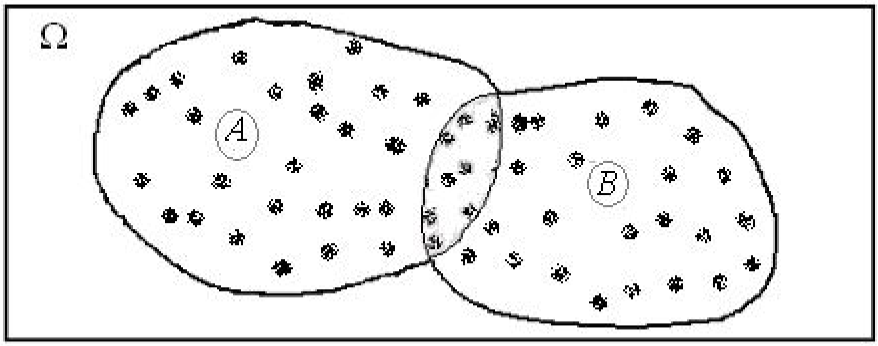


Рис. 3.2. Иллюстрация к теореме сложения вероятностей

В силу геометрической вероятности эта вероятность $P(A + B)$ равна:

$$P(A + B) = \frac{S_{A+B}}{S_\Omega},$$

где S_{A+B} - площадь фигуры $A \cup B$, а S_Ω - площадь области Ω. Осталось найти площадь S_{A+B}. Она равна:

$$S_{A+B} = S_A + S_B - S_{AB} \;,$$

где S_A - площадь фигуры A, S_B - площадь фигуры B, S_{AB} - площадь общей части фигур A и B, «забитой» на рис. 3.2 пятнами. Тогда:

$$P(A + B) = \frac{S_{A+B}}{S_\Omega} = \frac{S_A + S_B - S_{AB}}{S_\Omega} = P(A) + P(B) - P(AB),$$

где по определению геометрической вероятности:

$$S_A / S_\Omega = P(A)$$

вероятность события A,

$$S_B / S_\Omega = P(B)$$

вероятность события B,

$$S_{AB} / S_\Omega = P(AB)$$

вероятность события AB.

Тем самым, мы приходим к равенству

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

которое и составляет содержание *теоремы о сложении вероятностей совместных событий*, но доказательство её в общем случае гораздо сложнее и его мы оставляем без внимания.

Теорема о сложении вероятностей совместных событий. *Вероятность суммы совместных событий A и B равна:*

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

Здесь слова «вероятность совместных событий» имеют принципиальное значение, т.к. для несовместных событий получается несколько иная теорема. Разберёмся в этом. Для несовместных событий A и B основным свойством является равенство (они вместе произойти не могут):

$$P(AB) = 0.$$

Поэтому теорема переписывается в следующем виде.

Теорема о сложении вероятностей несовместных событий. *Вероятность суммы несовместных событий A и B равна:*

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

Пример. «Не кладите все яйца в одну корзину». В два банка положены деньги (слава Богу, что некто догадался положить их именно в два банка). Банки работают независимо друг от друга (часто встречающаяся ситуация). Вероятность разорения первого банка равна $0{,}1$, а второго - $0{,}2$. Какова вероятность того, что деньги сохранятся хотя бы в одном из банков.

Решение. Чтобы решить вероятностную задачу, главное, ввести правильные обозначения. Попробуем ввести следующие события.

A_1 - деньги взяты из первого банка,

A_2 - деньги взяты из второго банка.

Тогда событие $A_1 + A_2$ означает, что деньги взяты либо из первого, либо из второго банка, либо из обоих банков сразу (вам очень повезло). А найти нужно именно вероятность этого события $P(A_1 + A_2)$. По формуле сложения вероятностей совместных событий получаем:

$$P(A_1 + A_2) = P(A_1) + P(A_2) - P(A_1 A_2).$$

Вероятность $P(A_1)$ того, что первый банк останется «на плаву», составляет с вероятностью $P(\overline{A_1})$ того, что первый банк разорится, в сумме 1 (т.к. событие $A_1 + \overline{A_1}$ есть достоверное событие). Поэтому:

$$P(A_1) = 1 - P(\overline{A_1}) = 1 - 0{,}1 = 0{,}9.$$

Аналогично найдем

$$P(A_2) = 1 - P(\overline{A_2}) = 1 - 0{,}2 = 0{,}8.$$

А вероятность произведения двух событий $P(A_1 A_2)$ равна произведению вероятностей $P(A_1)P(A_2)$, как произведение независимых событий. Поэтому:

$$P(A_1 + A_2) = P(A_1) + P(A_2) - P(A_1)P(A_2) = 0{,}9 + 0{,}8 - 0{,}9 \cdot 0{,}8 = 1{,}7 - 0{,}72 = 0{,}98.$$

То есть искомая вероятность получается больше вероятностей $P(A_1)$ и $P(A_2)$, а, значит, права пословица!

5. Формула полной вероятности

Эта формула работает в том случае, когда событие происходит в опыте вместе с рядом других событий, составляющих полную группу. Такая ситуация иллюстрируется на рис. 3.3.

Допустим, что события $B_1, B_2, \ldots, B_n$ образуют полную группу (т.е. какое-то из них непременно происходит) несовместных (т.е. два разных события одновременно произойти не могут) событий. Тогда верна следующая теорема.

Теорема. *Если событие A может осуществляться только при выполнении одного из событий $B_1, B_2, \ldots, B_n$, которые образуют полную группу несовместных событий, то:*

$$p(A) = p(B_1)p(A/B_1) + p(B_2)p(A/B_2) + \ldots + p(B_n)p(A/B_n) = \sum_{i=1}^{n} p(B_i)p(A/B_i).$$

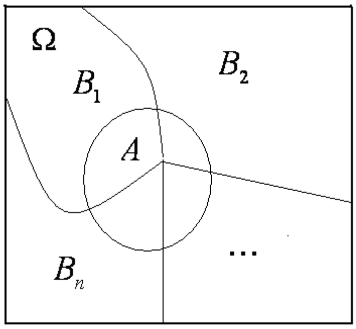


Рис. 3.3. Иллюстрация к формуле полной вероятности

Доказательство. Представим событие A как событие, умноженное на достоверное событие (от этого результат не изменится), а достоверное событие как сумму всех событий B_i:

$$A = A(B_1 + B_2 + \ldots + B_n) = AB_1 + AB_2 + \ldots + AB_n.$$

Тогда по формуле сложения вероятностей (события $B_1, B_2, \ldots, B_n$ - несовместные, а значит и события $AB_1, AB_2, \ldots, AB_n$ совместно произойти не могут) вероятность суммы равна сумме вероятностей:

$$p(A) = p(AB_1 + AB_2 + \ldots + AB_n) = p(AB_1) + p(AB_2) + \ldots + p(AB_n).$$

Далее по формуле умножения вероятностей получаем искомое соотношение:

$$p(A) = p(B_1)p(A/B_1) + p(B_2)p(A/B_2) + \ldots + p(B_n)p(A/B_n).$$

Что и требовалось доказать.

Отметим, что события $B_1, B_2, \ldots, B_n$ часто называют *гипотезами*.

Пример. В городе 5 банков, три из них («хорошие» банки) разорятся с вероятностью $p_1 = 0{,}05$, два («плохие» банки) – с $p_2 = 0{,}3$. Найти вероятность сохранения вклада, если деньги доверены наудачу одному из банков. (Общая ситуация. Вы никогда не знаете наверняка вероятность сохранения Вашего вклада).

Решение. Введем соответствующие обозначения:

событие B_1 - деньги доверены «хорошему» банку;

событие B_2 - деньги доверены «плохому» банку;

событие A - деньги сохранены.

Тогда можно найти вероятность того, что деньги сохранятся в «хорошем» банке (деньги будут сохранены, при условии того, что они будут доверены «хорошему» банку):

$$p(A/B_1) = 1 - 0{,}05 = 0{,}95,$$

т.к. эта вероятность в совокупности с вероятностью разорения «хорошего» банка даст единицу (достоверное событие).

Аналогично находится вероятность того, что деньги сохранятся в «плохом» банке (деньги будут сохранены при условии того, что они будут доверены «плохому» банку):

$$p(A/B_2) = 1 - 0{,}3 = 0{,}7.$$

Теперь можно найти вероятность того, что деньги доверены «хорошему» банку (по классическому способу нахождения вероятности событий, а число 3 равно числу благоприятствующих случаев, а всего - 5 банков):

$$p(B_1) = 3/5 = 0{,}6$$

и «плохому» банку (по классическому способу нахождения вероятности событий, а число 2 равно числу благоприятствующих случаев):

$$p(B_2) = 2/5 = 0{,}4.$$

Тогда по формуле полной вероятности:

$$p(A) = p(B_1)p(A/B_1) + p(B_2)p(A/B_2) = 0{,}6 \cdot 0{,}95 + 0{,}4 \cdot 0{,}7 = 0{,}57 + 0{,}28 = 0{,}85.$$

Т.е. вероятность оказалась «размытой по середине» между вероятностью сохранить деньги в «хорошем» банке и вероятностью сохранить в «плохом» банке.

6. Формула Байеса

Она нужна для того, чтобы переоценить вероятность $P(B_i/A)$ события B_i при условии того, что событие A произошло. Формула работает в том случае, когда события $B_1, B_2, \ldots, B_n$ образуют полную группу (т.е. какое-то из них непременно происходит) несовместных (т.е. два разных события одновременно произойти не могут) событий.

Пусть событие A произошло. Тогда переоценим вероятность события $P(B_i/A)$, исходя из формулы $P(AB_i)$. По формуле умножения вероятностей:

$$P(AB_i) = P(A)P(B_i/A),$$

где вероятность события A будем считать известной, а вероятность события B_i/A надо определить. Но по формуле умножения вероятностей, с другой стороны:

$$P(AB_i) = P(B_i)P(A/B_i),$$

где обе вероятности, участвующие в правой части формулы, также будем считать известными. Тогда, сравнивая правые части обеих формул, приходим к следующему:

$$P(A)P(B_i / A) = P(B_i)P(A / B_i) .$$

Отсюда находим неизвестную величину $P(B_i / A)$:

$$P(B_i / A) = \frac{P(B_i)P(A / B_i)}{P(A)} ,$$

или, пользуясь формулой полной вероятности, можно расписать знаменатель и получить окончательный вид *формулы Байеса*:

$$P(B_i / A) = \frac{P(B_i)P(A / B_i)}{\sum_{k=1}^{n} P(B_k)P(A / B_k)} .$$

Пример. В городе находится 10 банков. Вероятность того, что деньги сохранятся в 4-х банках («хороших») равна 0,95, а в остальных банках («плохих») равна 0,8. Вкладчик сохранил деньги в наудачу взятом банке. (Общая ситуация: Вы никогда не знаете наверняка, в каком именно банке храните деньги: в «хорошем» или в «плохом».)

Что вероятнее: вкладчик держал деньги в «хорошем» банке или в «плохом»?

Решение. Введем обозначения:

событие A - деньги сохранены;

событие B_1 - выбран «хороший» банк;

событие B_2 - выбран «плохой» банк.

Причем события B_1 и B_2 как раз и образуют полную группу (т.к. в какой-то банк положены деньги) несовместных (т.к. деньги хранились лишь в одном из банков) событий.

Тогда можно найти следующие вероятности:

$P(A / B_1) = 0{,}95$ - вероятность того, что деньги сохранятся, если выбран «хороший» банк;

$P(A / B_2) = 0{,}8$ - вероятность того, что деньги сохранятся, если выбран «плохой» банк;

$P(B_1) = \frac{4}{10}$ - вероятность того, что выбран «хороший» банк (определяется по классическому способу нахождения вероятности событий, а число 4 равно числу благоприятствующих случаев);

$P(B_2) = \frac{6}{10}$ - вероятность того, что выбран «плохой» банк (определяется по классическому способу нахождения вероятности событий, а число 6 равно числу благоприятствующих случаев).

Тогда $P(B_1 / A)$ (вероятность того, что был выбран «хороший» банк при условии, что деньги были сохранены) равна, по формуле Байеса:

$$P(B_1 / A) = \frac{P(B_1)P(A / B_1)}{P(B_1)P(A / B_1) + P(B_2)P(A / B_2)} =$$

$$= \frac{0{,}4 \cdot 0{,}95}{0{,}4 \cdot 0{,}95 + 0{,}6 \cdot 0{,}8} = \frac{0{,}38}{0{,}38 + 0{,}48} < \frac{1}{2} .$$

То есть вероятнее, что был выбран «плохой» банк при условии, что деньги были сохранены.

Так и в жизни. Никогда нельзя доверять тому, что говорят. Даже, если это говорит честный человек. Например, он купил автомобиль китайского производства, и автомобиль оказался хорошим. Человек утверждает, раз мой оказался хорошим, значит можно покупать китайские автомобили. Это неправда, просто этих автомобилей на рынке больше (по количеству) и хороший автомобиль вполне мог попасться этому человеку!

Лекция № 4

Схема независимых испытаний

Схема независимых испытаний представляет собой сочетание следующих факторов. Пусть производится серия из n независимых испытаний (что такое «независимость» мы говорили на прошлой лекции). В каждом испытании может возникнуть событие A с одной и той же вероятностью.

1. Формула Бернулли

Имеет место следующая теорема.

Теорема (формула Бернулли) *Вероятность того, в n независимых испытаниях событие A наступит ровно m раз, равна:*

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

где $C_n^m = \dfrac{n!}{m!\,(n-m)!}$ *- число сочетаний из n по m, p - вероятность события A, $q = 1-p$ - вероятность события $\overline{A}$ (т.е. вероятность ненаступления события A).*

Доказательство. Начнём с малого. Пусть B обозначает исходное событие, т.е. появление ровно m раз события A в n независимых испытаниях (вероятность появления A в одном испытании, напомню, равняется p).

Событие A может, например, появиться (событие B_1) следующим образом: вначале испытаний событие A наступает ровно m раз, а затем оно $n-m$ раз не наступает (значит, наступит противоположное событие $\overline{A}$):

$$B_1 = \underbrace{A \cdot A \cdot \ldots \cdot A}_{m \text{ раз}} \cdot \underbrace{\overline{A} \cdot \overline{A} \cdot \ldots \cdot \overline{A}}_{n-m \text{ раз}}.$$

Найдём вероятность этого события. Поскольку все события независимы («вероятность произведения равна произведению вероятностей»), то:

$$p(B_1) = p(A)^m p(\overline{A})^{n-m}.$$

А вероятность $p(\overline{A})$ найдём, исхитрившись. События A и $\overline{A}$ образуют полную группу, т.е. $p(A + \overline{A}) = 1$. Кроме того, они несовместны (т.к. вместе произойти не могут).

Поэтому («вероятность суммы равна сумме вероятностей»):

$$p(A+\overline{A}) = p(A) + p(\overline{A}) = 1,$$

Откуда:

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

Поэтому

$$p(B_1) = p(A)^m p(\overline{A})^{n-m} = p^m q^{n-m}.$$

Но событие A может появиться и другим образом. Например,

$$B_2 = \underbrace{A \cdot A \cdot \ldots \cdot A}_{m-1 \ \textit{раз}} \cdot \underbrace{\overline{A} \cdot \overline{A} \cdot \ldots \cdot \overline{A}}_{n-m \ \textit{раз}} \cdot A.$$

Нетрудно убедиться в том, что вероятность $p(B_2)$ по-прежнему равна:

$$p(B_2) = p(A)^m p(\overline{A})^{n-m} = p^m q^{n-m}.$$

Но как пересчитать все эти возможности (ясно, что они все являются несовместными, а поэтому $p(B)$ будет равно числу (сумме) всех этих возможностей умноженной на $p(B_i)$)? Число всех возможных таких вариантов событий B_i равно C_n^m, числу способов, которыми можно расположить n чисел по m местам (при этом порядок, занимаемый числами, не имеет значение):

числа $1,2,3,\ldots,m$ располагаются по m местам (событие B_1),

числа $1,2,3,\ldots,m-1,n$ располагаются по m местам (событие B_2),

и т.д.

Поэтому

$$p(B) = C_n^m p^m q^{n-m} = p_n(m).$$

Что и требовалось доказать.

Пример. Для нормальной работы автобазы на линии должно быть не менее восьми автомашин, на автобазе всего десять машин. Вероятность невыхода каждой автомашины на линию равна $0{,}1$. Найти вероятность нормальной работы автобазы.

Решение. Прежде всего поймём, что значит вероятность нормальной работы автобазы:

$$p\begin{pmatrix}\textit{нормальной}\\ \textit{работы}\\ \textit{автобазы}\end{pmatrix} = p\begin{pmatrix}\textit{на линии}\\ 8,9,10\ \textit{машин}\end{pmatrix} = p\begin{pmatrix}\textit{на линии}\\ 8\ \textit{машин}\end{pmatrix} + p\begin{pmatrix}\textit{на линии}\\ 9\ \textit{машин}\end{pmatrix} + p\begin{pmatrix}\textit{на линии}\\ 10\ \textit{машин}\end{pmatrix}.$$

Причём последнее равенство справедливо, т.к. несовместными являются события «8 машин на линии», «9 машин на линии» и «10 машин на линии».

Вероятность того, что 8 автомашин на линии, равна «вероятности того, что в 10 независимых испытаниях событие A (выход одной машины на линию) наступит ровно 8 раз»:

$$p_{10}(8) = C_{10}^8 p^8 q^2,$$

где p - вероятность выхода одной машины на линию, а q - вероятность невыхода одной машины на линию. Поскольку по условию задачи $q = 0{,}1$, постольку $p = 1 - q = 0{,}9$.

Окончательно,

$$p\begin{pmatrix}\textit{на линии}\\ 8\ \textit{машин}\end{pmatrix}=C_{10}^{8}p^{8}q^{2}=C_{10}^{8}0{,}9^{8}0{,}1^{2}.$$

Аналогично:

$$p\begin{pmatrix}\textit{на линии}\\ 9\ \textit{машин}\end{pmatrix}=C_{10}^{9}p^{9}q^{1}=C_{10}^{9}0{,}9^{9}0{,}1 \quad \text{и} \quad p\begin{pmatrix}\textit{на линии}\\ 10\ \textit{машин}\end{pmatrix}=C_{10}^{10}p^{10}q^{0}=C_{10}^{10}0{,}9^{10}.$$

Поэтому:

$$p\begin{pmatrix}\textit{нормальной}\\ \textit{работы}\\ \textit{автобазы}\end{pmatrix}=C_{10}^{8}0{,}9^{8}0{,}1^{2}+C_{10}^{9}0{,}9^{9}0{,}1+C_{10}^{10}0{,}9^{10}=$$

$$=\frac{10!}{2!\,8!}0{,}9^{8}0{,}01+\frac{10!}{1!\,9!}0{,}9^{9}0{,}1+\frac{10!}{0!\,10!}0{,}9^{10}=0{,}01\cdot\frac{9\cdot 10}{2}\cdot 0{,}9^{8}+0{,}1\cdot\frac{10}{1}\cdot 0{,}9^{9}+0{,}9^{10}=$$

$$=0{,}45\cdot 0{,}43046721+0{,}38742049+0{,}34867844\approx 0{,}9298.$$

Сделаем вывод. Поскольку в статистике считается, что событие, вероятность которого «больше $0{,}95\approx$ достоверное событие», постольку базу, иногда, будет «лихорадить». Полностью нормальной её работу считать нельзя! А для исправления ситуации следует прикупить автомашины или поработать над уменьшением вероятности *невыхода каждой автомашины на линию*.

2. Полиноминальная формула Бернулли

Пусть в каждом из n независимых испытаниях происходит одно из событий $\{A_1, A_2, \ldots A_m\}$, образующих полную группу. Вероятности наступления каждого из событий не изменяется в серии испытаний и равны $P(A_k)=p_k$, а само событие A_k наступает в серии испытаний ровно n_k. Причем $\sum p_k = 1$, а $\sum n_k = n$. Тогда верна следующая полиноминальная формула Бернулли:

$$P(n_1, n_2, \ldots n_m)=\frac{n!}{n_1!n_2!\ldots n_m!}p_1^{n_1}p_2^{n_2}\ldots p_m^{n_m},$$

которая дает вероятность получения различных комбинаций количеств событий в серии испытаний. При $m=2,\ A_1=A,\ A_2=\overline{A},\ p_1=p,\ p_2=q$ полиноминальная формула превращается в обычную биноминальную формулу Бернулли.

Пример. Вычислить вероятность того, что в шахматном турнире из 5 партий с соперником, у которого выиграть партию вы сможете с вероятностью 0,3, а проиграть 0,5, вы две партии выиграете, а проиграете только одну.

Решение. В каждой из независимых партий (испытаний) возможны три исхода. При этом $m=3, p_1=0{,}3, p_2=0{,}2, p_3=0{,}5$. Тогда ответ на вопрос задачи будет следующим:

$$P(2,2,1)=\frac{5!}{2!2!1!}0{,}3^{2}0{,}2^{2}0{,}5^{1}=0{,}054.$$

3. Теоремы Муавра-Лапласа

В условиях действия схемы Бернулли (производится n независимых испытаний, событие A наступает ровно m раз, вероятность наступления A в одном испытании равна p) *при большом* n подсчитать $p_n(m)$ по формуле Бернулли затруднительно (нужно подсчитать большие факториалы, большие степени и т.п.).

Для упрощения расчётов придумали формулу, но приближённую (пришлось «заплатить точностью»). Приведем без доказательства соответствующую теорему.

Локальная теорема Муавра-Лапласа. *Если вероятность* p *наступления события* A *в каждом из* n *независимых испытаний постоянна,* $0 < p < 1$*, число* n *испытаний велико, то вероятность* $p_n(m)$ *того, что в* n *независимых испытаниях событие* A *наступит ровно* m *раз, приближённо равно:*

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

где $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, $x = \frac{m-np}{\sqrt{npq}}$, $q = 1-p$.

Функция $\varphi(x)$ затабулирована (в [2] значения этой функции даются в таблице приложения 1).

При использовании этой таблицы полезно иметь в виду, что: функция $\varphi(x)$ чётная (т.е. $\varphi(x) = \varphi(-x)$); функция $\varphi(x)$ убывает при $|x| \to +\infty$; функция $\varphi(x) \geq 0$.

Пример. Найти вероятность того, что при 600 выстрелах мишень будет поражена ровно 250 раз, если вероятность поражения мишени при одном выстреле равна $0{,}4$.

Решение. Понятно, что искомую вероятность $p_n(m)$ можно найти по формуле Бернулли, но кто возьмется подсчитать C_{600}^{250} (т.е. $600!, 250!, \ldots$)? А с помощью локальной теоремы Муавра-Лапласа - запросто! Здесь $m = 250,\ n = 600,\ p = 0{,}4,\ q = 0{,}6$:

$$p_{600}(250) \approx \frac{\varphi(x)}{\sqrt{600 \cdot 0{,}4 \cdot 0{,}6}} = \frac{\varphi(x)}{\sqrt{6 \cdot 4 \cdot 6}} = \frac{\varphi(x)}{12},$$

$$x = \frac{(m-np)}{\sqrt{npq}} = \frac{250 - 600 \cdot 0{,}4}{12} = 10/12 \approx 0{,}833.$$

Поэтому по таблице приложения 1 находим $\varphi(0{,}833) \approx 0{,}2820$, откуда:

$$p_{600}(250) \approx \frac{0{,}2820}{12} \approx 0{,}0235.$$

Теперь пусть перед нами поставлена следующая задача. Найти вероятность того, что из достаточно большого числа объектов от a до b объектов имеют определённое свойство: $p_n(a \le m \le b)$.

Например, нужно найти вероятность того, что из 400 семей от 300 до 360 семей имеют автомобиль:

$$p_{400}(300 \le m \le 360).$$

В условиях действия схемы Бернулли (производится n независимых испытаний, событие A наступает ровно m раз, вероятность наступления A в одном испытании равна p) *при большом* n подсчитать $p_n(a \le m \le b)$ по формуле Бернулли затруднительно (нужно подсчитать большие факториалы, большие степени и т.п.). Подсчитать по локальной теореме Муавра-Лапласа? Но она приближённая, а поэтому мы сложим большое число ошибок и в итоге получим пшик!

На помощь приходит интегральная теорема Муавра-Лапласа, которую также приведём без доказательства.

Интегральная теорема Муавра-Лапласа. *Если вероятность p наступления события A в каждом из n независимых испытаний постоянна, $0 < p < 1$, то вероятность того, что число m наступления события A в n независимых испытаниях заключена в пределах от «a» до «b» (включительно) при достаточно большом числе n приближённо равна*

$$p_n(a \le m \le b) \approx \Phi(x_2) - \Phi(x_1),$$

где *функция* $\Phi(x)$ *равна* $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$ *(функция Лапласа),* $x_1 = \frac{a - np}{\sqrt{npq}}$, $x_2 = \frac{b - np}{\sqrt{npq}}$, $q = 1 - p$.

Функция $\Phi(x)$ затабулирована (в [2] значения этой функции даются в таблице приложения 2).

При использовании этой таблицы полезно иметь в виду, что: функция $\Phi(x)$ нечётная ($\Phi(x) = -\Phi(-x)$); функция $\Phi(x)$ возрастает при увеличении положительного значения x; функция $\Phi(x) \le 0{,}5$.

Пример. Вероятность появления события в каждом из 2100 независимых испытаний равна 0,7. Найти вероятность того, что событие появится не менее 1470 раз.

Решение. Нас интересует вероятность

$$p_{2100}(1470 \le m \le 2100) \approx \Phi(x_2) - \Phi(x_1),$$

где

$$x_2 = \frac{2100 - 2100 \cdot 0{,}7}{\sqrt{2100 \cdot 0{,}7 \cdot 0{,}3}} = \frac{2100 \cdot (1 - 0{,}7)}{\sqrt{21 \cdot 7 \cdot 3}} = \frac{21 \cdot 100 \cdot 0{,}3}{21} = 30,$$

следовательно, надо отыскать $\Phi(30)$, но в табл. приложения 2 из [2] даётся только значение $\Phi(5) = 0{,}499997$ (для больших значений аргумента значений не приводится). Но так как $\Phi(x)$ возрастает при значениях $x \geq 0$ и $\Phi(x) \leq 0{,}5$, то заключаем, что $\Phi(30) \approx 0{,}5$.

Кроме того $x_1 = \dfrac{1470 - 2100 \cdot 0{,}7}{21} = \dfrac{1470 - 1470}{21} = 0$, поэтому $\Phi(x_1) = \Phi(0) =$ $= \dfrac{1}{\sqrt{2\pi}} \int\limits_0^0 e^{-z^2/2} dz = 0$ (по свойству определенного интеграла). Отсюда:

$$p_{2100}(1470 \leq m \leq 2100) \approx \Phi(x_2) - \Phi(x_1) = 0{,}5 .$$

4. О границах применимости схемы Бернулли

Схема повторных испытаний Бернулли в силу своей простоты и наглядности имеет большое распространение не только в учебном процессе, но и практических приложениях. Однако надо помнить, что в ее основе лежат три основополагающих принципа:

- однородность (одинаковость, повторяемость) испытаний,
- независимость испытаний между собой,
- стационарность испытания (постоянство вероятностей его исходов).

Пример. На складе находится N деталей, причем из них M бракованных. Рабочий взял наудачу n деталей, какова вероятность того, что у него окажется m бракованных деталей.

Решение. Часто студенты, не владеющие полностью материалом курса, начинают рассуждать так: «Неважно, как рабочий взял эти детали, то ли он их получил у кладовщицы все сразу, то ли он сам набирал эти детали в тару по очереди (что верно в силу однородности и симметрии опыта). Но раз есть повторяемость, то почему бы не применить формулу Бернулли для повторных испытаний с вероятностью взять бракованную деталь равной $p = M / N$»?

Например, для $N = 10$, $M = 3$, $n = 5$, $m = 2$ эти студенты получают $P_5(2) = C_5^2 0{,}3^2 0{,}7^3 = 0{,}3087$. Однако они забывают, что вероятность взять бракованную деталь $p = M / N$ верна только для первого испытания. В последующих испытаниях она изменяется (становится условной), что нарушает стационарность и независимость испытаний. Если бы они знали «схему урн» или вспомнили о ней, то правильное решение находится тривиально:

$$P_5(2) = \frac{C_3^2 \cdot C_7^3}{C_{10}^5} = 0{,}417 \cdot$$

Заметим, что при условии «большого» склада $n << N$ $m << M$ разница в решении будет несущественной, так как вероятность взять бракованную деталь изменяется медленно. Так для значений $N = 100$, $M = 30$ $n = 5$, $m = 2$ решение по «Бернулли» будет тем же $P_5(2) = C_5^2 0{,}3^2 0{,}7^3 = 0{,}3087$, а вот правильное:

$$P_5(2) = \frac{C_{30}^2 \cdot C_{70}^3}{C_{100}^5} = 0{,}3163 \cdot$$

Разница менее существенна, но студенту остается уповать на «милость» преподавателя.

Лекция № 5

Дискретны случайные величины

1. Закон распределения дискретной случайной величины

Начнём с примера. Если Вы сдаёте экзамен, то результатом сдачи может явиться:

а) сам факт сдачи экзамена, это известное нам *случайное событие*;

б) оценка, полученная на экзамене. А когда речь идёт *о количестве*, то это уже новая характеристика, ранее нами не исследованная. Это уже *величина*. А поскольку она носит случайный характер (заранее никак не угадаешь, сдадите ли экзамен на «4,5» или на «4,7»), то это величина - *случайная*. Так появляется необходимость рассмотрения *случайных величин*.

Определение. *Переменная X, принимающая в результате испытаний случайным образом одно и только одно из конечной или бесконечной последовательности значений $x_1, x_1, \ldots, x_k, \ldots$, называется **дискретной случайной величиной**, если каждому значению x_k соответствует определённая вероятность p_k того, что переменная X примет значение x_k ($p_k = P(X = x_k)$).*

Обозначение. Обозначать дискретные случайные величины будем латинскими буквами X, Y, Z,..., а их возможные значения $x, y, z, \ldots$.

Определение. *Зависимость вероятности p_k от значения x_k называется **законом распределения** вероятностей дискретной случайной величины.*

Закон распределения удобно задавать в виде таблицы:

Значения случайной величины (X)	x_1	x_2	x_3	$\ldots$	x_n	$\ldots$
Вероятности значений (P)	p_1	p_2	p_3	$\ldots$	p_n	$\ldots$

Получается так называемый ***ряд распределения*** дискретной случайной величины.

Причём (что очень важно!) события $X = x_k$ - несовместные и единственно возможные (переменная X принимает одно и только одно значение), поэтому (по теореме сложения вероятностей и т.к. в итоге получается достоверное событие):

$$p_1 + p_2 + p_3 + \ldots + p_n = \sum_{i=1}^{n} p_i = 1,$$

если X принимает конечное n число значений, и

$$p_1 + p_2 + p_3 + \ldots + p_n + \ldots = \sum_{i=1}^{\infty} p_i = 1,$$

если X принимает бесконечное число значений. То есть получается *необходимое условие существования закона распределения*!

Закон распределения дискретной случайной величины можно задавать графически (рис. 5.1).

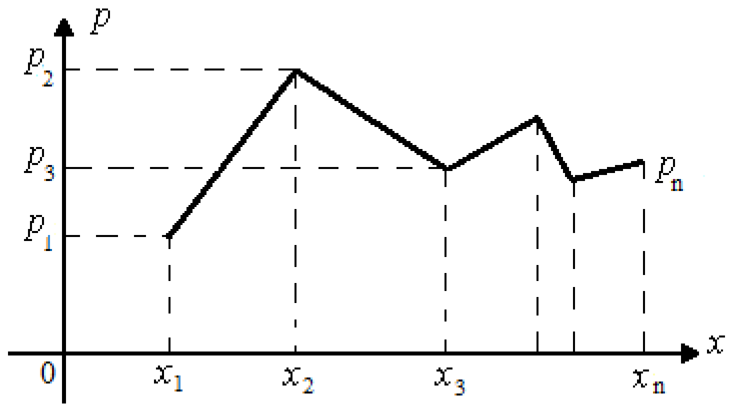


Рис. 5.1. Многоугольник распределения

Получается так называемый ***многоугольник распределения вероятностей*** дискретной случайной величины (на рис. 5.1 этот многоугольник изображается ломаной жирной линией).

Пример. Составить закон распределения числа попаданий в цель при четырех выстрелах, если вероятность попадания при одном выстреле равна $0{,}1$.

Решение. Прежде всего, разберём, сколько значений может принимать случайная величина X в данном случае. Можем совсем промахнуться (X будет равно 0), попасть один раз из четырёх ($X=1$), попасть два раза из четырёх ($X=2$), попасть три раза из четырёх ($X=3$), попасть четыре раз из четырёх ($X=4$).

Теперь о вероятностях. Нас интересует вероятность: 0 раз попасть при четырёх выстрелах. Иными словами, вероятность того, что в 4 независимых испытаниях событие A наступит ровно 0 раз. В схеме Бернулли мы её обозначали как $p_4(0)$, а находили по формуле $(p=0{,}1,\ q=1-0{,}1=0{,}9,\ 0!=1)$:

$$p_4(0)=C_4^0p^0q^4=\frac{4!}{0!\ 4!}0{,}9^4=0{,}9^4=0{,}6561=p_0.$$

Аналогично

$$p_4(1)=C_4^1p^1q^3=\frac{4!}{1!\ 3!}0{,}1^10{,}9^3=0{,}2916=p_1,\ p_4(2)=C_4^2p^2q^2=\frac{4!}{2!\ 2!}0{,}1^20{,}9^2=0{,}0{,}0486=p_2$$

$$p_4(3)=C_4^3p^3q^1=\frac{4!}{3!\ 1!}0{,}1^30{,}9^1=0{,}0036=p_3\quad p_4(4)=C_4^4p^4q^0=\frac{4!}{4!\ 0!}0{,}1^40{,}9^0=0{,}0001=p_4.$$

Причём легко подсчитать, что:

$$p_0+p_1+p_2+p_3+p_4=\sum_{i=0}^{4}p_i=1.$$

Поэтому закон распределения числа попаданий в цель при четырех выстрелах имеет вид

X	0	1	2	3	4
P	0,6561	0,2916	0,0486	0,0036	0,0001

2. Пример распределения дискретной случайной величины. Биномиальное распределение

Пусть мы по-прежнему находимся в рамках действия схемы независимых испытаний: проводится n независимых испытаний, вероятность появления события A во всех испытаниях одинакова и равна p.

Можно рассмотреть случайную величину X, число появлений события A в этих испытаниях. Понятно, что мы находимся в области действия схемы Бернулли, и значение $X = m$ соответствует вероятности

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

Поэтому появляется закон распределения, который получил название - *биномиальное распределение* (рис. 5.2):

X	0	1	…	m	…	n
P	q^n	npq^{n-1}	…	$C_n^m p^m q^{n-m}$	…	p^n

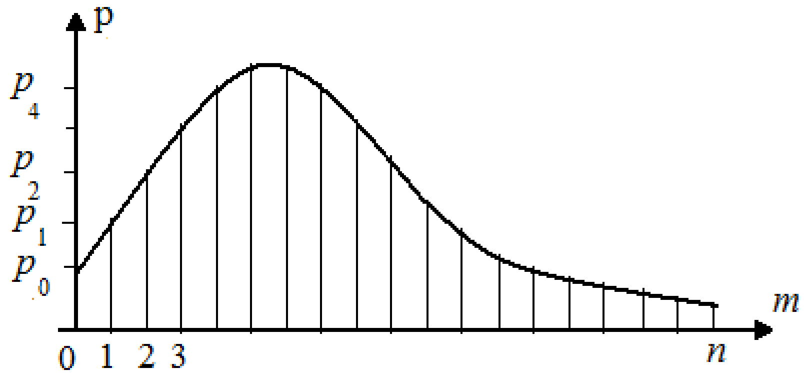


Рис. 5.2. Многоугольник биноминального распределения

Название закона происходит от разложения суммы двух чисел в n – ой степени, предложенного великим Ньютоном (*бином Ньютона*):

$$(p+q)^n = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = q^n + npq^{n-1} + \ldots + C_n^m p^m q^{n-m} + \ldots + p^n.$$

Кстати, последняя запись (формула) означает, что просуммированы все используемые в биномиальном распределении вероятности, а т.к. $p + q = 1$, то сумма всех вероятностей равна 1:

$$\sum_{k=0}^{n} p_k = 1,$$

т.е. у нас действительно выполняется необходимое условие существования закона распределения!

3. Пример распределения дискретной случайной величины. Распределение Пуассона

Пусть мы по-прежнему находимся в рамках действия схемы независимых испытаний: проводится n независимых испытаний, вероятность появления события A во всех испытаниях одинакова и равна p. Пусть к тому же n - велико, а p - мало. Тогда, применяя теорему Муавра-Лапласа $(n >> 1)$ для нахождения $p_n(m)$, будет получать значительные ошибки ($p \approx 0$).

Выход для частного, но практически важного, случая стремления $n \to \infty$, $p \to 0$ видится в следующем: предположим, что $p \cdot n = \lambda = \text{const}$ (т.е. в каждой серии из n независимых испытаний $p = p(n)$ зависит от n, но произведение $p \cdot n = \lambda = \text{const}$).

Тогда справедлива следующая теорема.

Теорема Пуассона. *При стремлении* $n \to \infty$ *так, что* $p \cdot n = \lambda = \text{const}$, *имеет место следующее:*

$$p(X = m) \to \frac{\lambda^m}{m!} e^{-\lambda}.$$

Поэтому появляется смысл во введении следующего распределения, которое получило название (по названию теоремы) – ***распределения Пуассона:***

X	0	1	2	3	…	m	…
P	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2}{2} e^{-\lambda}$	$\frac{\lambda^3}{3!} e^{-\lambda}$	…	$\frac{\lambda^m}{m!} e^{-\lambda}$	…

Проверим, выполняется ли необходимое условие существования закона распределения:

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

Но по свойству степенных рядов:

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda},$$

отсюда

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

Итак, если $n \to \infty$, но $p \cdot n = \lambda = \text{const}$ ($p \to 0$ при $n \to \infty$,), то биномиальное распределение превращается в распределение Пуассона.

4. Пример распределения дискретной случайной величины. Геометрическое распределение

Пусть мы по-прежнему находимся в рамках действия схемы независимых испытаний: проводится n независимых испытаний, вероятность появления события A во всех испытаниях одинакова и равна p. Но теперь испытания заканчиваются, если наступает событие A.

Поэтому случайная величина X - число испытаний, проведенных до *первого появления* события A. Найдём соответствующую вероятность:

$$p(X=k)=p(\underbrace{\overline{A}\cdot\overline{A}\cdot\ldots\cdot\overline{A}}_{k-1\ \text{раз}}\cdot A).$$

По определению независимости испытаний:

$$p(\underbrace{\overline{A}\cdot\overline{A}\cdot\ldots\cdot\overline{A}}_{k-1\ \text{раз}}\cdot A)=\underbrace{p(\overline{A})\cdot p(\overline{A})\cdot\ldots\cdot p(\overline{A})}_{k-1\ \text{раз}}\cdot p(A).$$

Отсюда:

$$p(X=k)=p(\overline{A})^{k-1}p(A)=q^{k-1}p.$$

Поэтому получается следующий закон распределения:

X	1	2	3	…	k	…
P	p	$q\,p$	$q^2\,p$	…	$q^{k-1}p$	…

Он носит название ***геометрический закон распределения*** (т.к. вероятности $p, qp, q^2p, \ldots$ суть члены геометрической прогрессии со знаменателем q).

Можно доказать, что при $0<q=1-p<1$ имеет место равенство:

$$\sum_{k=1}^{\infty}p_i=\sum_{k=1}^{\infty}q^{k-1}p=1,$$

и тем самым проверить необходимое условие существования закона распределения.

Действительно:

$$\sum_{k=1}^{\infty}q^{k-1}p=p\sum_{k=1}^{\infty}q^{k-1}=p(1+q+q^2+\ldots).$$

Чтобы найти сумму бесконечного ряда, найдём его частичную сумму:

$$1+q+q^2+\ldots+q^{n-1}=\frac{1-q^n}{1-q},$$

что, в свою очередь, легко проверить непосредственно перемножая:

$$(1-q)(1+q+q^2+\ldots+q^{n-1})=1+q+q^2+\ldots+q^{n-1}-q-q^2-q^3-\ldots-q^{n-1}-q^n=$$
$$=1+(q-q)+(q^2-q^2)+\ldots+(q^{n-1}-q^{n-1})-q^n=1-q^n.$$

Откуда

$$\sum_{k=1}^{\infty}q^{k-1}=\lim_{n\to\infty}\sum_{k=1}^{n}q^{k-1}=\lim_{n\to\infty}(1+q+q^2+\ldots q^{n-1})=\lim_{n\to\infty}\frac{1-q^n}{1-q}=\frac{1-\lim\limits_{n\to\infty}q^n}{1-q}=\frac{1}{1-q},$$

т.к. $q=1-p>0$.

Поэтому

$$\sum_{k=1}^{\infty}q^{k-1}p=p\sum_{k=1}^{\infty}q^{k-1}=p\frac{1}{1-q}=p\cdot\frac{1}{p}=1,$$

что и требовалось доказать.

Лекция № 6

Непрерывные случайные величины

1. Функция распределения непрерывной и дискретной случайной величины

Начнём с определения.

Определение. Непрерывной случайной величиной *называется переменная, которая может принимать случайным образом любые значения в некотором интервале числовой оси.*

Обозначение. Обозначать непрерывные случайные величины будем латинскими буквами X, Y, Z,...

Пример. Пусть между двумя населёнными пунктами A и B протянута телефонная линия. Расстояние между ними равно S_{AB}. Точку возможного обрыва линии будем характеризовать случайной величиной X, которая принимает значения на интервале от нуля до S_{AB}.

Тогда точка обрыва, точка x (то есть случайная величина X примет значение $X = x$), не может являться вероятностной характеристикой произошедшего обрыва: вероятность $P(X = x) = 0$.

На самом деле, по геометрической вероятности:

$$P(X = x) = \frac{S_x}{S_{AB}},$$

где S_x - длина точки x, S_{AB} - расстояние между пунктами A и B, но $S_x = 0$!

Как охарактеризовать с вероятностной точки зрения линию обрыва?

Непрерывную случайную величину X характеризуют с помощью функции распределения.

Функцией распределения *случайной величины X называется функция $F(x)$, выражающая для каждого числа x вероятность того, что случайная величина X примет какое-либо значение, меньшее числа x:*

$$F(x) = P(X < x).$$

Функция распределения $F(x)$ определена для всех x: $-\infty < x < +\infty$, а значения принимает на отрезке $[0,1]$, т.к. вероятность любого события находится именно в этих пределах.

Функцией распределения можно характеризовать и дискретные случайные величины.

Пример. Пусть X - число попаданий в цель при четырех выстрелах, если вероятность попадания при одном выстреле равна $0{,}1$ (пример из предыдущей лекции). Найти и изобразить функцию распределения этой случайной величины X.

Решение. В предыдущей лекции мы нашли, что закон распределения X имеет вид:

X	0	1	2	3	4
P	0,6561	0,2916	0,0486	0,0036	0,0001

Найдём функцию распределения $F(x) = P(X < x)$. При значении $x = 5$ функция $F(x = 5) = 1$, т.к. событие $X < 5$ можно составить из пяти несовместных событий:

$$X = 0,\ X = 1,\ X = 2,\ X = 3,\ X = 4,$$

вероятности которых в сумме дают 1. И это будет справедливо для тех значений x, которые удовлетворяют неравенству $x > 4$. Поэтому $F(x) = 1$ при значениях $x > 4$.

Но как только x принимает значение $x = 4$, сразу из перечисленного выше множества событий исключается событие $X = 4$ (т.к. $F(4) = P(X < 4)$). Поэтому:

$$P(X < 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 1 - 0{,}0001 = 0{,}9999.$$

И так будет справедливо только для тех значений x, которые удовлетворяют неравенствам $3 < x \le 4$. Поэтому $F(x) = 0{,}9999$ для тех значений x, которые удовлетворяют неравенствам $3 < x \le 4$. И так далее.

Последним в этом списке будет значение $x = 0$, которому не соответствует ни одно из событий $X = i,\ i = \overline{0{,}4}$ (т.е. переменная i принимает значения от 0 до 4 с шагом равным 1). Поэтому для значений $x \le 0$ значение функции распределения равно $F(x) = 0$.

Итак, для рассматриваемой здесь случайной величины функция распределения имеет вид:

$$F(x) = \begin{cases} 0 & \textit{при} \quad x \le 0, \\ 0{,}6561 & \textit{при} \quad 0 < x \le 1, \\ 0{,}9477 & \textit{при} \quad 1 < x \le 2, \\ 0{,}9963 & \textit{при} \quad 2 < x \le 3, \\ 0{,}9999 & \textit{при} \quad 3 < x \le 4, \\ 1 & \textit{при} \quad x > 4. \end{cases}$$

Графиком функции распределения $F(x)$ является «набор из горизонтальной линии $y = 0$ $(x \le 0)$ и горизонтальных стрелок» (рис. 6.1), которые говорят о том, что предел справа у функции не достигается в пяти случаях:

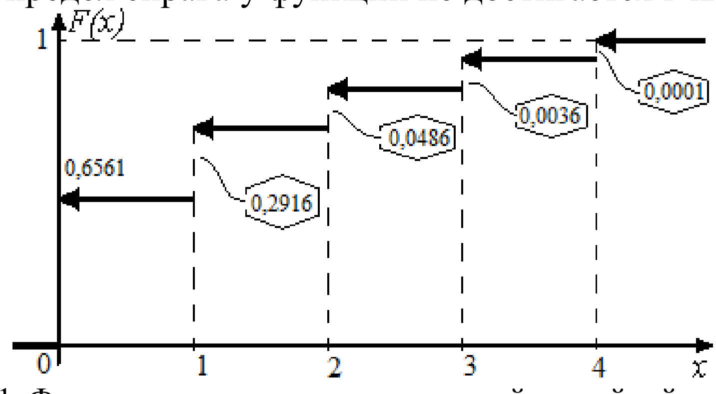


Рис. 6.1. Функция распределения дискретной случайной величины

2. Свойства функции распределения

Свойство 1^o. *Вероятность того, что случайная величина X примет какое-либо значение x, удовлетворяющее неравенству $x_1 \le X < x_2$, равна приращению функции распределения $F(x)$ на этом интервале:*

$$P(x_1 \le X < x_2) = F(x_2) - F(x_1).$$

Доказательство. Разобьём событие $X < x_2$ на два несовместных события: $X < x_1$ и $x_1 \le X < x_2$. Тогда получим (по теореме о сложении вероятностей несовместных событий):

$$P(X < x_2) = P(X < x_1) + P(x_1 \le X < x_2).$$

Поскольку первые две вероятности, участвующие в последнем равенстве, суть функции распределения, постольку получается такое равенство:

$$F(x_2) = F(x_1) + P(x_1 \le X < x_2),$$

откуда и получается:

$$P(x_1 \le X < x_2) = F(x_2) - F(x_1).$$

Что и требовалось доказать.

Свойство 2^o. *Функция распределения равна: от минус бесконечности - нулю, а от плюс бесконечности - единице. Иными словами:*

$$F(-\infty) = 0, \quad F(+\infty) = 1.$$

Доказательство. Поскольку $F(-\infty) = P(X < -\infty)$ - есть вероятность пустого множества, постольку $F(-\infty) = 0$, а т.к. $F(+\infty) = P(X < +\infty)$ - есть вероятность достоверного события, то $F(+\infty) = 1$.

Что и требовалось доказать.

Свойство 3^o. *Функция распределения (любой случайной величины) - неубывающая функция.*

Доказательство. Поскольку (свойство 1^o)

$$P(x_1 \le X < x_2) = F(x_2) - F(x_1)$$

при $x_2 > x_1$, а вероятность всегда

$$P(x_1 \le X < x_2) \ge 0,$$

постольку

$$F(x_2) - F(x_1) \ge 0, \text{ т.е. } F(x_2) \ge F(x_1) \text{ при } x_2 > x_1.$$

Что и требовалось доказать.

Таким образом, функция распределения $F(x)$ не убывает, её значения расположены на отрезке $[0,1]$. При стремлении $x \to -\infty$ функция распределения обращается в ноль, а при стремлении $x \to +\infty$ функция распределения обращается в единицу. Примерный график функции распределения $F(x)$ приведён на рис. 6.2:

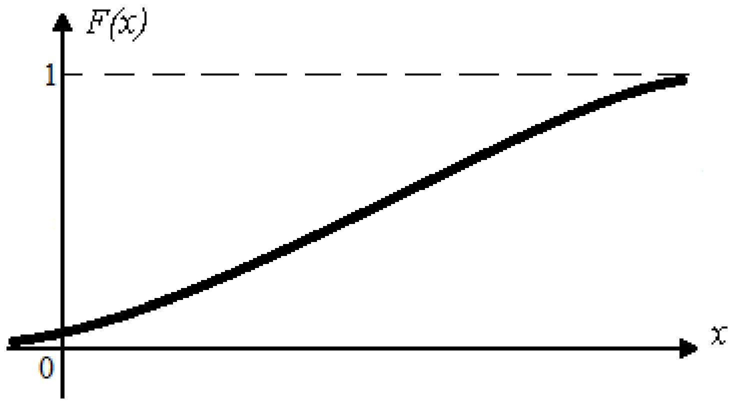


Рис. 6.2. Функция распределения непрерывной случайной величины

Свойство 4^o**.** *Если функция распределения* $F(x)$ *непрерывна в точке* $x = C$*, то вероятность того, что случайная величина* X *принимает значение* $X = C$*, равна нулю:*

$$P(X = C) = 0.$$

Доказательство. Оценим вероятность $P(X = C)$:

$$P(X = C) \le P(C - \varepsilon \le X < C + \varepsilon),$$

причём это верно для любого ε. Но по свойству 1^o

$$P(X = C) \le P(C - \varepsilon \le X < C + \varepsilon) = F(C + \varepsilon) - F(C - \varepsilon).$$

Теперь перейдём к пределу (т.к. ε - любое) в этом неравенстве (неравенство сохранится):

$$\lim_{\varepsilon \to 0} P(X = C) \le \lim_{\varepsilon \to 0}\big(F(C + \varepsilon) - F(C - \varepsilon)\big) = \lim_{\varepsilon \to 0} F(C + \varepsilon) - \lim_{\varepsilon \to 0} F(C - \varepsilon).$$

Предел слева равен самому выражению, а справа запишем выражения (обозначения) для пределов:

$$P(X = C) \le F(C + 0) - F(C - 0).$$

Поскольку предел справа $F(C + 0)$ равен (для непрерывной функции)

$$F(C + 0) = F(C),$$

а предел слева $F(C - 0)$ также равен (для непрерывной функции)

$$F(C - 0) = F(C),$$

постольку

$$P(X = C) = F(C) - F(C) = 0.$$

Что и требовалось доказать.

Следствие из свойства 4^o**.** *Для непрерывной функции распределения* $F(x)$ *справедливо следующее:*

$$P(x_1 < X < x_2) = P(x_1 < X \le x_2) = P(x_1 \le X \le x_2) = P(x_1 \le X < x_2) = F(x_2) - F(x_1).$$

Доказательство. Поскольку по свойству 1^o:

$$F(x_2) - F(x_1) = P(x_1 \le X < x_2),$$

а $P(x_1 \le X < x_2)$ отличается от вероятности $P(x_1 \le X \le x_2)$ только на величину $P(X = x_2)$, т.е. вероятностью, которая равна 0 (по свойству 4^o), постольку:

$$F(x_2) - F(x_1) = P(x_1 \le X < x_2) = P(x_1 \le X < x_2) + P(X = x_2) = P(x_1 \le X \le x_2).$$

Последнее равенство справедливо по теореме о сложении вероятностей несовместных событий. Аналогично доказываются два других равенства.

3. Плотность распределения вероятностей непрерывной случайной величины

Определение. Плотностью распределения вероятностей (или сокращённо ***плотностью вероятности***) $\varphi(x)$ *непрерывной случайной величины называется производная от её функции распределения* $F(x)$, *если только существует эта производная:*

$$\varphi(x) = F'(x).$$

Пример. Найти плотность вероятности случайной величины X (величины Релея), которая принимает неотрицательные значения, а её функция распределения равна $F(x) = 1 - e^{-k^2x^2}$.

Решение. Т.к. $F(x=0)=0$ и $F(x)$ не убывает ($1-e^{-k^2x^2} > 0$ при $x<0$), то на самом деле:

$$F(x) = \begin{cases} 1-e^{-k^2x^2} & при\ x \ge 0 \\ 0 & при\ x < 0 \end{cases}.$$

Поэтому:

$$\varphi(x) = F'(x) = \begin{cases} (1-e^{-k^2x^2})'_x & при\ x \ge 0 \\ 0 & при\ x < 0 \end{cases} = \begin{cases} -\left(e^{-k^2x^2}\right)(-k^2 \cdot 2x) & при\ x \ge 0 \\ 0 & при\ x < 0 \end{cases} =$$

$$= \begin{cases} 2k^2 x e^{-k^2x^2} & при\ x \ge 0 \\ 0 & при\ x < 0 \end{cases}.$$

4. Свойства плотности вероятности

Свойство 1^o. *Вероятность того, что случайная величина* X *примет какое-либо значение* x *из замкнутого интервала* $[a,b]$, *равна*

$$P(a \le X \le b) = \int_a^b \varphi(x)\,dx.$$

Доказательство. Функция распределения $F(x)$ - непрерывна, т.к. существует производная $\varphi(x) = F'(x)$. Поэтому по следствию из свойства 4^o для непрерывной функции распределения:

$$P(a \le X \le b) = F(b) - F(a),$$

а по формуле Ньютона-Лейбница:

$$F(b) - F(a) = \int_a^b F'(x)\,dx = \int_a^b \varphi(x)\,dx.$$

Поэтому $P(a \le X \le b) = \int_a^b \varphi(x)\,dx$.

Что и требовалось доказать.

Пример. Плотность вероятности случайной величины X задана:

$$\varphi(x) = \frac{1}{\pi(1+x^2)}.$$

Найти вероятность того, что случайная величина X примет значение на интервале $[0;5]$.

Решение. По только что доказанному свойству:

$$P(0 \le X \le 5) = \int_0^5 \varphi(x)\,dx = \int_0^5 \frac{1}{\pi(1+x^2)}\,dx = \frac{1}{\pi}\int_0^5 \frac{dx}{(1+x^2)} = \frac{1}{\pi}\operatorname{arctg} x \Big|_0^5 = \frac{1}{\pi}(\operatorname{arctg} 5 - 0) \approx 0{,}435.$$

Свойство 2^o. *Функция* $\varphi(x)$, *плотность распределения вероятностей, всегда неотрицательна, т.е.*

$$\varphi(x) \ge 0.$$

Доказательство. Поскольку $F(x_2) \ge F(x_1)$ при $x_2 > x_1$ (по свойству 3^o для функции распределения), то:

$$\varphi(x = x_1) = F'(x = x_1) = \lim_{x_2 \to x_1} \frac{F(x_2) - F(x_1)}{x_2 - x_1} \ge 0,$$

как отношение двух неотрицательных величин.

Что и требовалось доказать.

Свойство 3^o. $\int_{-\infty}^{+\infty} \varphi(x)\,dx = 1.$

Доказательство. По только что установленному свойству 2^o ($\varphi(x) \ge 0$) плотности вероятности:

$$\int_{-\infty}^{+\infty} \varphi(x)\,dx \ge \int_{-A}^{+A} \varphi(x)\,dx$$

при любом достаточно большом A. Но по свойству 1^o для плотности вероятности:

$$\int_{-A}^{+A} \varphi(x)\,dx = P(-A \le X \le A)$$

при любом достаточно большом A. Следовательно:

$$\int_{-\infty}^{+\infty} \varphi(x)\,dx \ge \int_{-A}^{+A} \varphi(x)\,dx = P(-A \le X \le A) = F(A) - F(-A)$$

по следствию для непрерывной функции распределения. Откуда, переходя к пределу при $A \to +\infty$ (неравенство сохранится), получаем:

$$\int_{-\infty}^{+\infty} \varphi(x)\,dx \ge \lim_{A \to +\infty} \big(F(A) - F(-A)\big) = \lim_{A \to +\infty} F(A) - \lim_{A \to +\infty} F(-A) = F(+\infty) - F(-\infty).$$

Откуда по свойству 2^o для функции распределения:

$$\int_{-\infty}^{+\infty} \varphi(x)\,dx \ge F(+\infty) - F(-\infty) = 1 - 0 = 1.$$

Поскольку вероятность события не может быть больше 1, постольку

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

В силу доказанных сейчас свойств, функция $\varphi(x)$ плотности распределения вероятностей всегда неотрицательна (по свойству 2^o). Она стремится к нулю при стремлении $x \to -\infty$ и $x \to +\infty$ (т.к. по свойству 3^o площадь между графиком функции $\varphi(x)$ и осью абсцисс равна единице). Примерный график функции $\varphi(x)$ плотности распределения вероятностей изображён на следующем рис. 6.3.

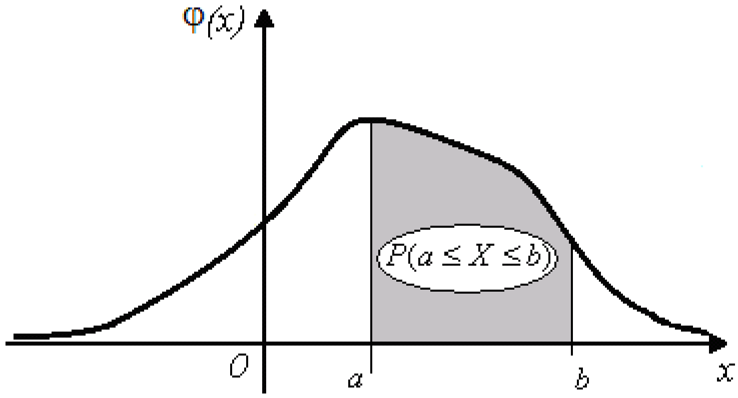


Рис. 6.3. Иллюстрация свойств 1-3 функции плотности распределения

Свойство 4^o. *Функция распределения* $F(x)$ *равна*

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

Доказательство. Для несобственного интеграла

$$\int_{-\infty}^{x} \varphi(t)\, dt$$

справедливо:

$$\int_{-\infty}^{x} \varphi(t)\, dt = \lim_{A \to -\infty} \int_{A}^{x} \varphi(t)\, dt ,$$

а по свойству 1^o для плотности распределения вероятностей:

$$\int_{A}^{x} \varphi(t)\, dt = P(A \le X \le x) .$$

По следствию из свойства 4^o для непрерывной функции распределения:

$$\int_{A}^{x} \varphi(t)\, dt = P(A \le X \le x) = F(x) - F(A) .$$

Поэтому, переходя к пределу, получим:

$$\int_{-\infty}^{x} \varphi(t)\, dt = \lim_{A \to -\infty} \int_{A}^{x} \varphi(t)\, dt = \lim_{A \to -\infty} \big(F(x) - F(A)\big) = F(x) - \lim_{A \to -\infty} F(A) = F(x) - F(-\infty) .$$

По свойству 2^o для функции распределения $F(-\infty) = 0$, т.е.

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

Что и требовалось доказать.

Итак, для полной характеристики случайной величины достаточно знать или функцию распределения, или плотность распределения вероятностей (т.к. одну из них можно выразить через другую):

$$F(x)=\int_{-\infty}^{x}\varphi(t)\,dt \quad \text{или} \quad \varphi(x)=F'(x).$$

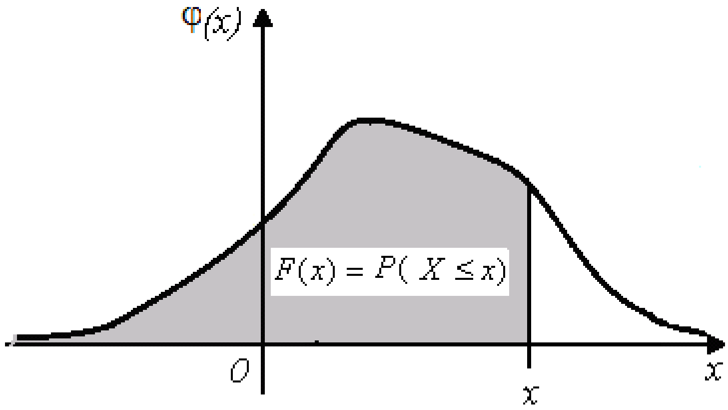


Рис. 6.4. Иллюстрация свойства 4 функции плотности распределения

Пример. Найти функцию распределения случайной величины, плотность вероятности которой:

$$\varphi(x)=\frac{1}{\pi(1+x^2)}.$$

Решение. По только что доказанному свойству:

$$F(x)=\int_{-\infty}^{x}\varphi(t)\,dt=\int_{-\infty}^{x}\frac{1}{\pi(1+t^2)}dt=$$

$$=\frac{1}{\pi}\operatorname{arctg}\ t\Big|_{-\infty}^{x}=\frac{1}{\pi}(\operatorname{arctg}\ x-\operatorname{arctg}(-\infty))=\frac{1}{\pi}(\operatorname{arctg}\ x+\frac{\pi}{2})=\frac{1}{2}+\frac{1}{\pi}\operatorname{arctg}x\ .$$

Лекция № 7

Примеры распределения непрерывных случайных величин

1. Равномерное распределение

Определение. *Случайная величина с плотностью вероятности*

$$\varphi(x)=\begin{cases}0 & \text{при } x<a \quad \text{и} \quad x>b\\ C & \text{при } a<x<b\end{cases},\quad \text{где } C=\text{const},$$

называется ***равномерно распределённой величиной***.

Равномерный закон распределения используется: при анализе ошибок измерения, когда проводятся численные расчёты; в ряде задач массового обслуживания.

Найдём величину C из условия $\int\limits_{-\infty}^{+\infty}\varphi(x)\,dx = 1$ (свойство 3^o плотности вероятности):

$$\int\limits_{-\infty}^{+\infty}\varphi(x)\,dx = \int\limits_{-\infty}^{a}0\cdot dx + \int\limits_{a}^{b}C\cdot dx + \int\limits_{b}^{+\infty}0\cdot dx = C\int\limits_{a}^{b}dx = C\,x\Big|_a^b = C(b-a) \underset{\substack{должно\\быть}}{=} 1.$$

Поэтому $C = \dfrac{1}{b-a}$, а плотность вероятности равномерно распределённой величины имеет вид:

$$\varphi(x) = \begin{cases} 0 & при\ x < a \quad и \quad x > b \\ \dfrac{1}{b-a} & при\ a < x < b \end{cases}.$$

Найдём также функцию распределения равномерно распределённой величины. По свойству 4^o для плотности вероятности:

$$F(x) = \int\limits_{-\infty}^{x}\varphi(t)\,dt = \begin{cases} \int\limits_{-\infty}^{x}0\cdot dt & при\ x < a \\ \int\limits_{-\infty}^{a}0\cdot dt + \int\limits_{a}^{x}\dfrac{1}{b-a}dt & при\ a < x < b \\ \int\limits_{-\infty}^{a}0\cdot dt + \int\limits_{a}^{b}\dfrac{1}{b-a}dt + \int\limits_{b}^{x}0\cdot dt & при\ b < x \end{cases} =$$

$$= \begin{cases} 0 & при\ x < a \\ \int\limits_{a}^{x}\dfrac{1}{b-a}dt & при\ a < x < b \\ \int\limits_{a}^{b}\dfrac{1}{b-a}dt & при\ b < x \end{cases} = \begin{cases} 0 & при\ x < a \\ \dfrac{1}{b-a}(x-a) & при\ a < x < b \\ \dfrac{1}{b-a}(b-a) & при\ b < x \end{cases} = \begin{cases} 0 & при\ x < a \\ \dfrac{x-a}{b-a} & при\ a < x < b \\ 1 & при\ b < x \end{cases}.$$

Графики функций $F(x)$ и $\varphi(x)$ приведены ниже на рис. 7.1. На графике для функции $\varphi(x)$ четыре стрелки означают, что левый или правый пределы не достижимы функцией в соответствующей точке.

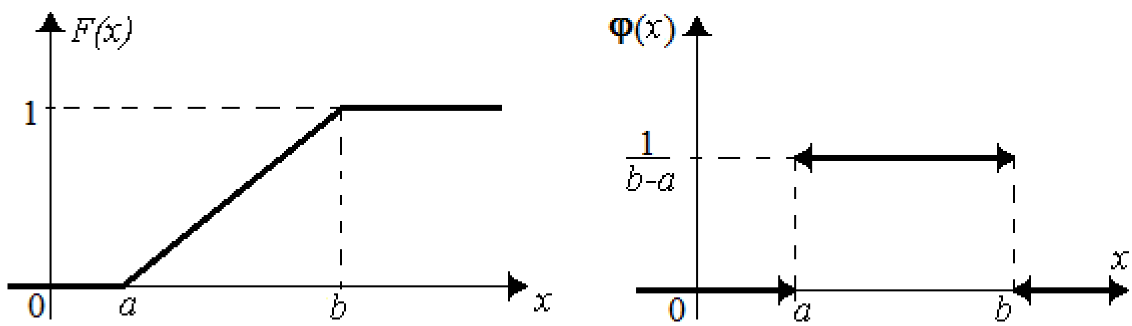


Рис. 7.1. Равномерное распределение

Пример. Поезда метрополитена идут регулярно с интервалом 2 минуты. Пассажир выходит на платформу в случайный момент времени. Какова вероятность того, что ждать пассажиру придётся не более полминуты.

Решение. Пусть случайная величина X - время ожидания пассажира. Тогда её плотность вероятности равна:

$$\varphi(x)=\begin{cases}0 & при\ \ x<0\ \ и\ \ x>2\\ \dfrac{1}{2} & при\ \ 0<x<2\end{cases}.$$

Поэтому по свойству 1^o для плотности вероятности получим:

$$P(0\le X\le \frac{1}{2})=\int_0^{\frac{1}{2}}\frac{1}{2}dx=\frac{1}{2}\int_0^{\frac{1}{2}}dx=\frac{1}{2}x\Big|_0^{1/2}=\frac{1}{4}.$$

2. Нормальное распределение

Определение. *Случайная величина имеет* ***нормальный закон распределения*** *(закон* ***Гаусса****), если её плотность распределения вероятностей имеет вид:*

$$\varphi(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}},$$

где a и σ - параметры распределения $(\sigma>0,\ -\infty<a<+\infty)$.

Нормальный закон распределения наиболее часто встречается на практике. Главная его особенность – он является предельным законом, к которому приближаются другие законы распределения (при типичных условиях).

Плотность вероятности $\varphi(x)$ - функция, похожая на колокол. Зависимость от параметров такова (рис. 7.2). При уменьшении только параметра σ, график функции вытягивается и поднимается вверх по оси ординат. А при увеличении только параметра a, график симметрично передвигается вправо вдоль оси абсцисс:

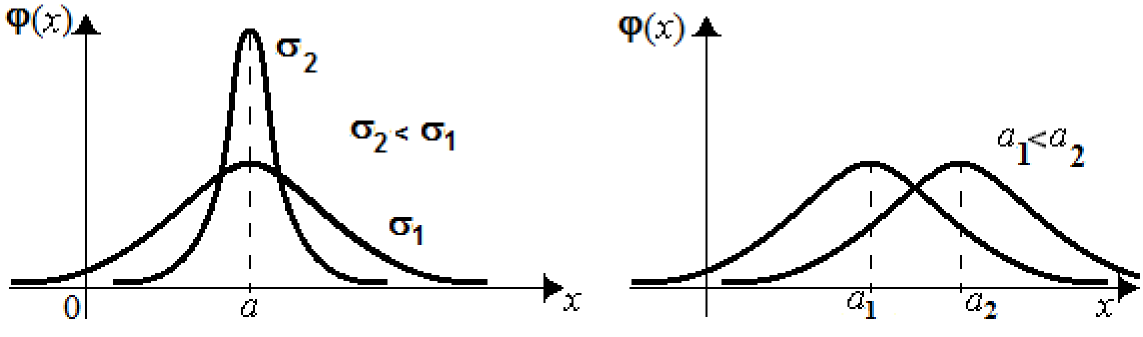


Рис. 7.2. Функция плотности распределения нормальной величины

Функция распределения $F(x)$ нормального распределения

$$F(x)=\int_{-\infty}^{x}\varphi(t)\,dt=\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{x}e^{-\frac{(t-a)^2}{2\sigma^2}}\,dt$$

имеет вид, изображенный на рис. 7.3:

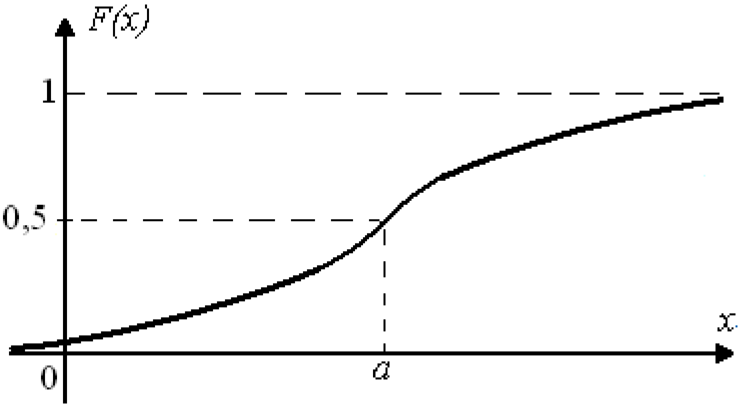


Рис. 7.3. Функция распределения нормальной величины

а) Правило «трёх сигм»

Найдём вероятность того, что изучаемая случайная величина (распределённая нормально) примет значение в пределах от α до β:

$$P(\alpha \le X \le \beta) = \int_{\alpha}^{\beta} \varphi(x)\, dx = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}}\, dx .$$

Для этого воспользуемся известным из математического анализа свойством определённого интеграла:

$$\int_{a}^{b} f(x)\, dx = \int_{a}^{c} f(x)\, dx + \int_{c}^{b} f(x)\, dx$$

и, используя ещё одно свойство:

$$\int_{a}^{c} f(x)\, dx = -\int_{c}^{a} f(x)\, dx ,$$

окончательно получим:

$$\int_{a}^{b} f(x)\, dx = \int_{c}^{b} f(x)\, dx - \int_{c}^{a} f(x)\, dx .$$

Этим равенством и воспользуемся (при условии, что роль c играет параметр a из нормального закона)

$$P(\alpha \le X \le \beta) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}}\, dx = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{a}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}}\, dx \; - \; \frac{1}{\sqrt{2\pi}\,\sigma} \int_{a}^{\alpha} e^{-\frac{(x-a)^2}{2\sigma^2}}\, dx .$$

Далее сделаем замену $t = \dfrac{x-a}{\sigma}$ в определённых интегралах (тогда $dt = d(\dfrac{x-a}{\sigma})$ или $dx = \sigma\, dt$):

$$P(\alpha \le X \le \beta) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{a}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}}\, dx \; - \; \frac{1}{\sqrt{2\pi}\,\sigma} \int_{a}^{\alpha} e^{-\frac{(x-a)^2}{2\sigma^2}}\, dx =$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} \int\limits_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} \, \sigma\, dt \; - \; \frac{1}{\sqrt{2\pi}\,\sigma} \int\limits_{\frac{a-a}{\sigma}}^{\frac{\alpha-a}{\sigma}} e^{-\frac{t^2}{2}} \, \sigma\, dt =$$

$$= \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} \, dt \; - \; \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{\frac{\alpha-a}{\sigma}} e^{-\frac{t^2}{2}} \, dt = \Phi(\frac{\beta-a}{\sigma}) - \Phi(\frac{\alpha-a}{\sigma}),$$

где функция Лапласа $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{z^2}{2}} dz$ затабулирована и приводится в приложении 2 из [2]. В частном случае, когда интервал симметричен относительно точки a, эта формула выглядит так:

$$P(a-\varepsilon \le X \le a+\varepsilon) = \Phi(\frac{a+\varepsilon-a}{\sigma}) - \Phi(\frac{a-\varepsilon-a}{\sigma}) = \Phi(\frac{\varepsilon}{\sigma}) - \Phi(\frac{-\varepsilon}{\sigma}) = 2\Phi(\frac{\varepsilon}{\sigma})$$

или так:

$$P(|X-a| \le \varepsilon) = 2\Phi(\frac{\varepsilon}{\sigma}).$$

Отсюда правило «трёх сигм» выводится следующим образом. Рассмотрим вероятность того, что изучаемая случайная величина (распределённая нормально) примет значение в пределах от $a-3\sigma$ до $a+3\sigma$:

$$P(a-3\sigma \le X \le a+3\sigma) = 2\Phi(3\sigma/\sigma) = 2\Phi(3).$$

Из таблицы для функции Лапласа находим, что $\Phi(x) = 0{,}49865$, поэтому

$$P(a-3\sigma \le X \le a+3\sigma) = 2 \cdot 0{,}49865 \approx 0{,}9973,$$

т.е. *вероятность встретить значение изучаемой случайной величины именно на интервале* $[a-3\sigma, a+3\sigma]$ *велика* - $0{,}9973$!!!

3. Показательный (экспоненциальный) закон распределения

Определение. *Случайная величина X имеет* ***показательный (экспоненциальный) закон распределения,*** *если её плотность распределения вероятностей имеет вид:*

$$\varphi(x) = \begin{cases} 0 & при\ x < 0 \\ \lambda e^{-\lambda x} & при\ x \ge 0 \end{cases},$$

где λ – параметр распределения ($\lambda > 0$).

Он возникает в теории массового обслуживания, теории надёжности. Например, интервал времени T между двумя соседними событиями (заявками) в потоке поступающих заявок на обслуживание (ремонт телевизоров, автомобилей, …) имеет показательный закон распределения (с интенсивностью λ).

Примерный график плотности распределения вероятностей $\varphi(x)$ приводится на рис. 7.4.

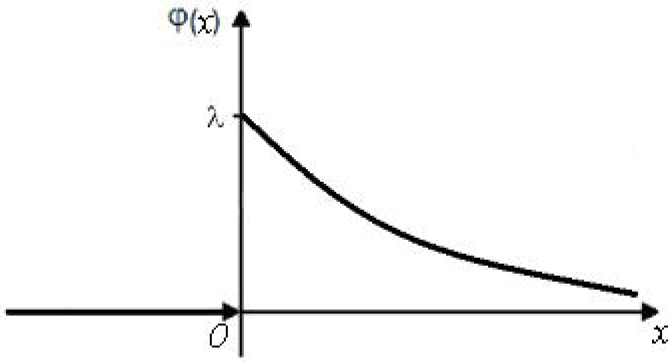


Рис. 7.4. Функция плотности распределения показательной величины

Определим вид функции распределения для показательного закона:

$$F(x)=\int_{-\infty}^{x}\varphi(t)\,dt=\begin{cases}\int_{-\infty}^{x}0\cdot dt=0 & \text{при } x<0\\ \int_{-\infty}^{x}\lambda e^{-\lambda t}dt=\int_{-\infty}^{0}0\cdot dt+\int_{0}^{x}\lambda e^{-\lambda t}dt=-\int_{0}^{x}e^{-\lambda t}d(-\lambda t) & \text{при } x\geq 0\end{cases}=$$

$$=\begin{cases}0 & \text{при } x<0\\ -e^{-\lambda t}\Big|_{0}^{x} & \text{при } x\geq 0\end{cases}=\begin{cases}0 & \text{при } x<0\\ 1-e^{-\lambda x} & \text{при } x\geq 0\end{cases}.$$

Примерный график функции распределения $F(x)$ приводится на рис.7.5.

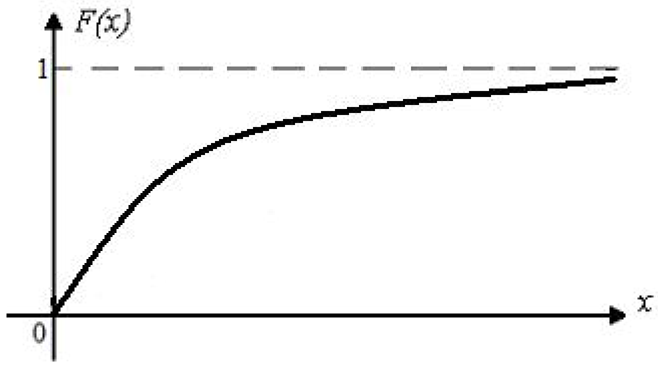


Рис. 7.5. Функция распределения показательной величины

4. Логарифмически-нормальное распределение

Определение. *Случайная величина* $X(X>0)$ *имеет* ***логарифмически-нормальное (логнормальное) распределение,*** *если её натуральный логарифм* $\ln X$ *подчинён нормальному закону:*

$$F(x)=P(\ln X<\ln x)=\begin{cases}0 & \text{при } x<0\\ \dfrac{1}{\sigma\sqrt{2\pi}}\int_{0}^{\ln x}e^{-\frac{(t-a)^2}{2\sigma^2}}dt & \text{при } x\geq 0\end{cases}.$$

Отсюда, функция $\varphi(x)$ плотность распределения вероятностей логнормального распределения имеет вид (по правилу дифференцирования интеграла, зависящего от параметра)

$$\varphi(x)=\begin{cases}0 & при\ \ x<0\\ \dfrac{1}{\sigma\sqrt{2\pi}\,x}e^{-\frac{(\ln x-a)^2}{2\sigma^2}} & при\ \ x\ge 0\end{cases}.$$

Примерный вид графика функции $\varphi(x)$ приведён на рис. 7.6.

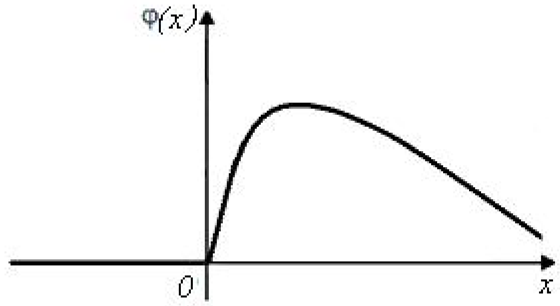


Рис. 7.6. Функция плотности распределения логнормальной величины

Логнормальное распределение встречается при описании распределения доходов, банковских вкладов, долговечности изделий в режиме износа – старения, месячной зарплаты, посевных площадей под различные культуры и т.п.

5. Вейбуловское распределение

В инженерной практике часто используется распределение Вейбулла-Гнеденко:

$$F(x)=\begin{cases}0 & при\ x<0\\ 1-e^{-(x/\alpha)^{\beta}} & при\ x\ge 0\end{cases},\quad \varphi(x)=\begin{cases}0 & при\ x<0\\ (\beta/\alpha)\left(x/\alpha\right)^{\beta-1}e^{-(x/\alpha)^{\beta}} & при\ x\ge 0\end{cases}$$

с параметрами $\alpha>0$ и $\beta>0$. Данное распределение часто используется для описания распределения экстремальных значений системы случайных величин:

$$X_{\max}(n)=\max(X_1,X_2,...X_n),\ \ X_{\min}(n)=\min(X_1,X_2,...X_n),$$

кроме того оно используется для описания времени и интенсивности отказов в теории надежности сложных систем.

Частными случаями распределения Вейбулла-Гнеденко являются следующие распределения :

- Показательное распределение $\alpha=1/\lambda>0$, $\beta=1$,
- Релеевское распределение $\alpha>0$, $\beta=2$.

Лекция № 8

Числовые характеристики случайных величин

1. Математическое ожидание. Дискретные случайные величины

Пусть задана случайная величина X своим рядом распределений

Значения случайной величины (X)	x_1	x_2	x_3	$\ldots$	x_n	$\ldots$
Вероятности значений (P)	p_1	p_2	p_3	$\ldots$	p_n	$\ldots$

Определение. Математическим ожиданием *дискретной случайной величины X называется сумма произведений всех возможных значений x_i случайной величины на вероятности p_i этих значений:*

$$\sum_{i=1}^{n(+\infty)} x_i p_i = M[X] = M(X) = m_X .$$

Примечания. 1. Все три приведённые обозначения математического ожидания равноправны.

2. Верхний индекс у знака суммы (n или $+\infty$) соответствует тому, что X принимает конечное число значений (n) или бесконечное число значений.

3. Сумма всех участвующих (в определении) вероятностей равна единице (попробуйте это доказать):

$$\sum_{i=1}^{n(+\infty)} p_i = 1 .$$

Пример. Два стрелка стреляют по мишени. Число очков, которые выбивает каждый стрелок, заданы следующими законами распределений:

X_1	1	2	3	4	5
$P(X_1 = x_k)$	0	0	0,3	0,4	0,3

X_2	1	2	3	4	5
$P(X_2 = x_k)$	0,1	0,1	0,1	0,2	0,5

Какой из стрелков опытнее, кого нужно взять на соревнование (между вузами), кому нужно отдать предпочтение?

Решение. Задача па первый взгляд непростая. Но если воспользоваться понятием математического ожидания (найти среднее число очков, выбиваемых каждым стрелком при одном выстреле), задача переходит в разряд решаемых.

Действительно, для первого стрелка математическое ожидание равно:

$$M(X_1) = 1\cdot 0 + 2\cdot 0 + 3\cdot 0{,}3 + 4\cdot 0{,}4 + 5\cdot 0{,}3 = 0{,}9 + 1{,}6 + 1{,}5 = 4{,}0 ,$$

(т.е. он выбивает в среднем $4{,}0$ очка за один выстрел), а для второго стрелка

$$M(X_2)=1\cdot 0{,}1+2\cdot 0{,}1+3\cdot 0{,}1+4\cdot 0{,}2+5\cdot 0{,}5=0{,}1+0{,}2+0{,}3+0{,}8+2{,}5=3{,}9.$$

То есть пока опытнее (точнее) первый стрелок. Но если второй «уберёт» шлейф плохих выстрелов, то он станет вне конкуренции!

Свойства математического ожидания

Свойство 1^o. *Математическое ожидание постоянной величины равно этой постоянной:*

$$X=C=\text{const}\Rightarrow M(X)=M(C)=C$$

Доказательство. Действительно, пусть случайная величина X равна $C=$ const с вероятность $p=1$. Тогда по определению математического ожидания:

$$M(X)=C\cdot 1=C.$$

Что и требовалось доказать.

Свойство 2^o. *Постоянный множитель можно выносить за знак математического ожидания:*

$$M(k\cdot X)=k\cdot M(X),\ \text{где } k=\text{const}.$$

Доказательство. Докажем это для случайной величины X, которая принимает конечное число n значений $x_i, i=1,\dots,n$. По определению математического ожидания:

$$M(k\cdot X)=\sum_{i=1}^{n}k\cdot x_i\cdot p_i.$$

Отсюда, вынося константу за знак суммы, получим:

$$\sum_{i=1}^{n}k\cdot x_i\cdot p_i=k\cdot\sum_{i=1}^{n}x_i\cdot p_i=k\cdot M(X).$$

Что и требовалось доказать.

Свойство 3^o. *Математическое ожидание суммы конечного числа случайных величин равно сумме математических ожиданий этих величин:*

$$M(X_1+X_2+\ldots+X_k)=M(X_1)+M(X_2)+\ldots+M(X_k).$$

Доказательство. Докажем это для математического ожидания величины, составленной из суммы двух случайных величин X и Y, причем случайная величина X, принимает конечное число n значений $x_i, i=1,\dots,n$, а случайная величина Y, принимает конечное число m значений $y_j, j=1,\dots,m$. Тогда математическое ожидание суммы двух величин X и Y равно:

$$M(X+Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}(x_i+y_j)P(X=x_i,Y=y_j)=$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m}x_iP(X=x_i,Y=y_j)+\sum_{i=1}^{n}\sum_{j=1}^{m}y_jP(X=x_i,Y=y_j).$$

Попробуем разобраться с первой двойной суммой

$$\sum_{i=1}^{n}\sum_{j=1}^{m}x_iP(X=x_i,Y=y_j).$$

В ней от значения x_i не зависит внутренняя сумма, поэтому вынесем x_i за знак этой суммы:

$$\sum_{i=1}^{n}\ \sum_{j=1}^{m} x_i\, P(X=x_i, Y=y_j) = \sum_{i=1}^{n} x_i \sum_{j=1}^{m} P(X=x_i, Y=y_j).$$

Событие, состоящее в том, что X примет значение x_i (вероятность этого события равна $P(X=x_i)$), влечёт за собой событие, которое состоит в том, что $X+Y$ примет значения:

$$X=x_i, Y=y_1 \text{ или } X=x_i, Y=y_2, \quad \ldots \quad, \text{ или } X=x_i, Y=y_m,$$

а вероятности этих несовместных событий равны соответственно:

$$P(X=x_i, Y=y_1) \text{ или } P(X=x_i, Y=y_2), \quad \ldots \quad, \text{ или } P(X=x_i, Y=y_m)\ .$$

Тогда вероятность $P(X=x_i)$ первоначального события $X=x_i$ равна (по теореме о сложении вероятностей несовместных событий):

$$P(X=x_i) = \sum_{j=1}^{m} P(X=x_i, Y=y_j).$$

Поэтому первая двойная сумма равна:

$$\sum_{i=1}^{n} x_i \sum_{j=1}^{m} P(X=x_i, Y=y_j) = \sum_{i=1}^{n} x_i\, P(X=x_i) = M(X).$$

Последнее же равенство следует из определения математического ожидания.

Со второй двойной суммой поступим аналогично, но прежде заметим, что она не изменится, если поменять порядок суммирования:

$$\sum_{i=1}^{n}\ \sum_{j=1}^{m}\ y_j P(X=x_i, Y=y_j) = \sum_{j=1}^{m}\ \sum_{i=1}^{n}\ y_j P(X=x_i, Y=y_j)$$

(Это известное свойство можно проверить, расписав его для случаев $i=2, j=2$ или $i=3, j=2$). А далее:

$$\sum_{j=1}^{m}\ \sum_{i=1}^{n}\ y_j P(X=x_i, Y=y_j) = \sum_{j=1}^{m} y_j \sum_{i=1}^{n} P(X=x_i, Y=y_j) = \sum_{j=1}^{m} y_j\, P(Y=y_j) = M(Y).$$

Поэтому окончательно получаем:

$$M(X+Y) = \sum_{i=1}^{n}\ \sum_{j=1}^{m}\ x_i P(X=x_i, Y=y_j) + \sum_{i=1}^{n}\ \sum_{j=1}^{m}\ y_j P(X=x_i, Y=y_j) =$$

$$= M(X) + M(Y).$$

Что и требовалось доказать.

Свойство 4^o. *Математическое ожидание произведения* ***независимых*** *случайных величин равно произведению их математических ожиданий:*

$$M(XY) = M(X)\cdot M(Y).$$

Без доказательства (для заинтересовавшихся студентов это доказательство – повод повысить итоговую оценку).

Пример. Вероятность попадания в мишень при одном выстреле равна $p=0{,}2$. Найти число независимых выстрелов, обеспечивающее математическое ожидание равное 5 попаданиям в мишень (например, при стрельбе по кораблю).

Решение. Пусть случайные величины $X_1, X_2, \ldots, X_n$ - есть число попаданий в мишень при каждом из n выстрелов. Тогда, согласно условию задачи, все эти случайные величины $X_i, i = 1,2,\ldots,n,$ имеют один и тот же закон распределения

X_i	0	1
P	0,8	0,2

Математическое ожидание этих случайных величин равно:

$$M(X_i) = 0 \cdot 0{,}8 + 1 \cdot 0{,}2 = 0{,}2 .$$

Математическое ожидание числа попаданий в мишень при n выстрелах равно $M(X_1 + X_2 + \ldots + X_n)$. По свойству 3^o (математического ожидания):

$$M(X_1 + X_2 + \ldots + X_n) = M(X_1) + M(X_2) + \ldots + M(X_n) = n \cdot M(X_i) = n \cdot 0{,}2 = 5 ,$$

где последнее равенство записано в силу условия задачи. Отсюда

$$n = 5 : 0{,}2 = 10 .$$

2. Математическое ожидание. Непрерывные случайные величины

Определение. Математическим ожиданием непрерывной с*лучайной величины* X *с плотностью распределения вероятностей* $\varphi(x)$ *называется следующий интеграл:*

$$M(X) = \int_{-\infty}^{+\infty} x\,\varphi(x)\,dx .$$

Свойства математического ожидания непрерывной случайной величины точно такие же, как и у $M(X)$ дискретной случайной величины.

Пример. Случайная величина X распределена равномерно на отрезке $[0;10]$. Найти её математическое ожидание.

Решение. По свойству равномерно распределённой на отрезке $[0;10]$ случайной величины её плотность распределения вероятностей равна:

$$\varphi(x) = \begin{cases} 0 & \textit{при } x < 0 \quad \textit{и} \quad x > 10 \\ C & \textit{при } 0 < x < 10 \end{cases},$$

где $C = 1/10$ (попробуйте обосновать это).

Тогда по определению математического ожидания непрерывной случайной величины:

$$M(X) = \int_{-\infty}^{+\infty} x\,\varphi(x)\,dx = \int_{-\infty}^{0} x \cdot 0\,dx + \int_{0}^{10} x \cdot \frac{1}{10} dx + \int_{10}^{+\infty} x \cdot 0\,dx =$$

$$= \int_{0}^{10} x \cdot \frac{1}{10}\,dx = \frac{1}{10}\int_{0}^{10} x\,dx = \frac{1}{10}\frac{x^2}{2}\Big|_0^{10} = \frac{1}{10}\left(\frac{10^2}{2} - 0\right) = 5 .$$

3. Дисперсия и среднее квадратическое отклонение случайной величины

Начнём сразу с двух определений.

Определение. Дисперсией *случайной величины* X *называется величина:*

$$D(X) = M\left[(X - M(X))^2\right].$$

Дисперсия говорит о среднем квадрате отклонения от среднего (математического ожидания).

Определение. Средним квадратическим отклонением *случайной величины* X *называется величина:*

$$\sigma(X) = \sqrt{D(X)}.$$

Среднее квадратическое отклонение говорит о среднем отклонении от математического ожидания.

Для дискретной случайной величины дисперсия имеет вид:

$$D(X) = \sum_{i=1}^{n(+\infty)} (x_i - M(X))^2.$$

Для непрерывной случайной величины с плотностью распределения вероятностей $\varphi(x)$ дисперсия имеет вид:

$$D(X) = \int_{-\infty}^{+\infty} (x - M(X))^2 \varphi(x)\, dx.$$

Примечание. Полезно знать, что для нормально распределенной случайной величины X (напомним, что её плотность распределения вероятностей имеет вид $\varphi(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$) математическое ожидание X равно "a", а среднее квадратическое отклонение X равно "σ", т.е. величинам, входящим в определение самого закона.

Пример. Найти дисперсию и среднее квадратическое отклонение случайной величины X, заданной следующим законом распределения:

X	$-0,1$	$-0,01$	0	$0,01$	$0,1$
P	0,1	0,2	0,4	0,2	0,1

Решение. Сначала найдём математическое ожидание случайной величины X:

$$M(X) = (-0{,}1)\cdot 0{,}1 + (-0{,}01)\cdot 0{,}2 + 0\cdot 0{,}4 + 0{,}01\cdot 0{,}2 + 0{,}1\cdot 0{,}1 = 0.$$

Теперь настала очередь дисперсии:

$$D(X) = (-0{,}1-0)^2\cdot 0{,}1 + (-0{,}01-0)^2\cdot 0{,}2 + (0-0)^2\cdot 0{,}4 + (0{,}01-0)^2\cdot 0{,}2 + (0{,}1-0)^2\cdot 0{,}1 = \\ = 0{,}00204$$

и среднего квадратического отклонения:

$$\sigma(X) = \sqrt{D(X)} = \sqrt{0{,}00204} \approx 0{,}04517.$$

Пример. Найти дисперсию и среднее квадратическое отклонение равномерно распределённой на отрезке $[0;10]$ случайной величины X.

Решение. Поскольку математическое ожидание этой случайной величины X мы нашли ранее $(M(X)=5)$, а её плотность распределения вероятностей имеет вид:

$$\varphi(x)=\begin{cases}0 & при\ x<0 \quad и \quad x>10\\ 1/10 & при\ 0<x<10\end{cases},$$

то дисперсия её считается следующим образом:

$$D(X)=\int_{-\infty}^{+\infty}(x-M(X))^2\,\varphi(x)\,dx=\int_{-\infty}^{0}(x-5)^2\cdot 0\,dx+\int_{0}^{10}(x-5)^2\cdot 0{,}1\,dx+\int_{10}^{+\infty}(x-5)^2\cdot 0\,dx=$$

$$=0{,}1\int_{0}^{10}(x-5)^2\,dx=0{,}1\int_{0}^{10}(x-5)^2\,d(x-5)=0{,}1\,\frac{(x-5)^3}{3}\,\Big|_0^{10}=0{,}1\frac{5^3-(-5)^3}{3}=\frac{25}{3}\approx 8{,}333.$$

Отсюда среднее квадратическое отклонение для случайной величины X равно:

$$\sigma(X)=\sqrt{D(X)}\approx\sqrt{8{,}333}\approx 2{,}89.$$

4. Свойства дисперсии

Свойство 1^o. *Дисперсия постоянной величины равна нулю:*

$$X=\text{const}=C\Rightarrow D(X)=D(C)=0.$$

Доказательство. Действительно, пусть случайная величина X равна $C=$ const с вероятность $p=1$. Поскольку тогда $M(C)=C$, то по определению дисперсии:

$$D(X)=M\left[(X-M(X))^2\right]=M\left[(C-M(C))^2\right]=M\left[(C-C)^2\right]=M[0^2]=M[0]=0.$$

Что и требовалось доказать.

Свойство 2^o. *Постоянный множитель можно вносить за знак математического ожидания, но* ***в квадрате:***

$$D(k\cdot X)=k^2\cdot D(X),\ \text{где}\ k=\text{const}.$$

Доказательство. Используя определение дисперсии и свойство 2^o математического ожидания, получим:

$$D(k\cdot X)=M\left[(kX-M(kX))^2\right]=M\left[(k\cdot X-k\cdot M(X))^2\right]=M\left[k^2\cdot(X-M(X))^2\right]=$$
$$=k^2M\left[(X-M(X))^2\right]=k^2D(X).$$

Что и требовалось доказать.

Свойство 3^o. *Дисперсия равна математическому ожиданию квадрата минус квадрат математического ожидания:*

$$D(X)=M\left[(X)^2\right]-(M(X))^2=M\left[X^2\right]-M^2(X).$$

Доказательство. Используя определение дисперсии и свойство 3^o математического ожидания, получим:

$$D(X)=M\left[(X-M(X))^2\right]=M\left[(X^2-2\cdot X\cdot M(X)+(M(X))^2)\right]=$$

$$=M\left[X^2\right]+M\left[-2\cdot X\cdot M(X)\right]+M\left[(M(X))^2\right].$$

Поскольку математическое ожидание – суть константа, то по свойству 2^o, а затем по свойству 1^o математического ожидания, приходим к следующему:

$D(X) = M\left[X^2\right] + M\left[-2M(X)\cdot X\right] + M\left[(M(X))^2\right] = M\left[X^2\right] - 2M(X)\cdot M\left[X\right] + \left(M(X)\right)^2$.
Теперь, приводя подобные, получаем:

$$D(X) = M\left[X^2\right] - 2M(X)\cdot M\left[X\right] + \left(M(X)\right)^2 = M\left[X^2\right] - \left(M(X)\right)^2$$

Что и требовалось доказать.

Свойство 4^o. *Дисперсия суммы двух **независимых** случайных величин равна сумме дисперсий этих величин:*

$$D(X+Y) = D(X) + D(Y).$$

Без доказательства (для заинтересовавшихся студентов это доказательство – повод повысить итоговую оценку).

Свойство 5^o. *Дисперсия случайной величины ограничивает вероятность её отклонение от своего математического ожидания (неравенство Чебышева П.Л.):*

$$P(|X - M(X)| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2}.$$

Без доказательства.

Пример. При ракетной стрельбе в «заданный район» среднеквадратическое отклонение от цели имеет значение $\sigma = 20 м$. Оценить радиус круга безопасности, где с вероятностью не мене $0{,}99$ ракеты не ложатся.

Решение. Пусть R - координата точки падения по дальности. Тогда вероятность выхода за ε-зону ограничена (по неравенству Чебышева) следующим:

$$P(|R - M(R)| \geq \varepsilon) \leq \frac{D(R)}{\varepsilon^2},$$

но она должна быть не больше $0{,}99$. Это будет выполнено, если

$$\frac{D(R)}{\varepsilon^2} = \frac{400}{\varepsilon^2} \leq 0{,}01,$$

т. е. при $\varepsilon \geq 200\ м$.

Заметим: если предположить, что дальность распределена по нормальному закону, то:

$$P(|R - M(R)| \geq \varepsilon) = 1 - 2\Phi(\frac{\sigma}{\varepsilon}) \leq 0{,}01,$$

а значит $\Phi(\frac{\sigma}{\varepsilon}) \geq 0{,}495$, или при $\varepsilon \geq 51{,}6\ м$. Как видим, знание закона распределения существенно уточняет круг безопасности!

На рис. 8.1 показан геометрический смысл основных числовых характеристик случайной величины.

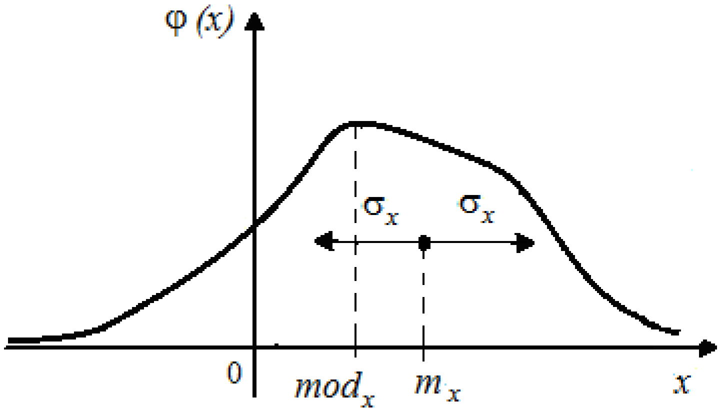


Рис. 8.1. Геометрическая иллюстрация понятий математического ожидания $M(X) = m_x$ моды mod_X и дисперсии $D(X) = \sigma_x^2$ случайной величины

Так, математическое ожидание $M(X) = m_x$ характеризует центр распределения или среднее ожидаемое значение величины и геометрически оно изображается как координата центра тяжести фигуры, образованной осью x и линией функции $\varphi(x)$. Дисперсия $D(X) = \sigma_x^2$ и среднее квадратическое отклонение $\sigma(X) = \sigma_x$ характеризуют средний ожидаемый разброс (широту, изменчивость) значений величины возле математического ожидания. Наиболее вероятное значение случайной величины является ее мода mod_X, оно соответствует максимуму функции $\varphi(x)$.

5. Моменты распределения случайной величины

Для описания распределения случайной величины иногда недостаточно только знания математического ожидания и дисперсии. Для более полного описания необходимо ввести еще ряд числовых характеристик распределения, и такими характеристиками могут быть моменты высших порядков.

А именно, начальный теоретический момент порядка k

$$A_k(X) = \int_{-\infty}^{+\infty} x^k \cdot \varphi(x)\, dx$$

и центральный теоретический момент порядка k

$$B_k(X) = \int_{-\infty}^{+\infty} (x - M(X))^k \cdot \varphi(x)\, dx.$$

Можно заметить, что $A_0(X) = 1$, $A_1(X) = M(X)$, $A_2(X) = M(X^2)$, а так же $B_0(X) = 1$, $B_1(X) = 0$, $B_2(X) = D(X)$.

В технических приложениях часто используются моменты 3-го и 4-го порядка, где они (безразмерные) имеют специальные названия:

$A_s = \dfrac{B_3(X)}{\sigma^3(X)}$ - асимметрия случайной величины,

$E_s = -3 + \dfrac{B_4(X)}{\sigma^4(X)}$ - эксцесс случайной величины.

Асимметрия случайной величины равна нулю у случайной величины симметричной относительно своего математического ожидания, а ее значение характеризует степень асиметрии ее распределения. Эксцесс равен нулю у нормальной случайной величины, а его значение характеризует степень отклонения от нормального закона распределения. Смысловое значение асиметрии («скошенности») и эксцесса («островершинности») иллюстрируется на рис.8.2.

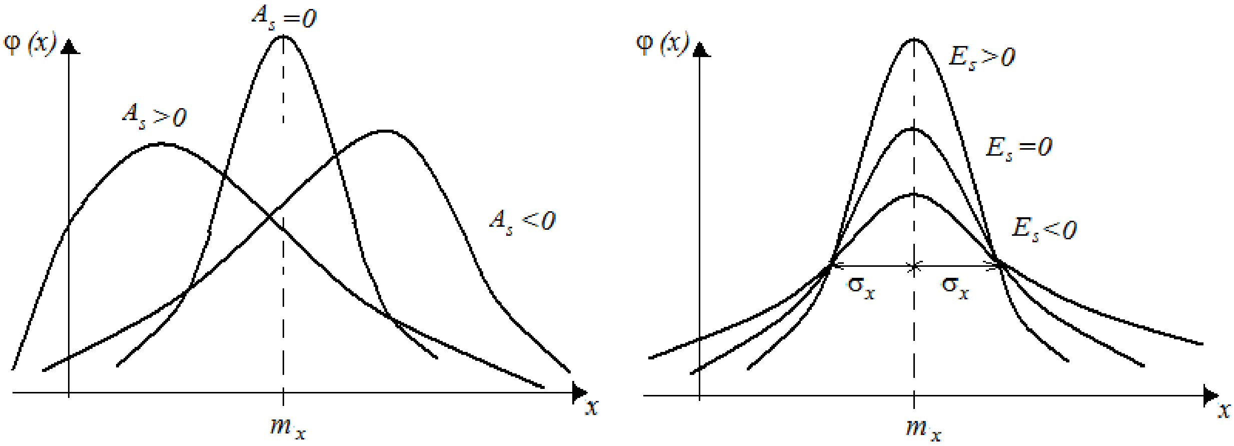


Рис. 8.2. Геометрическая иллюстрация понятий асимметрии A_s и эксцесса E_s

Лекция № 9

Закон больших чисел

Закон больших чисел (ЗБЧ) представляет собой ситуацию, когда совокупное действие большого числа случайных факторов приводит к результату, независящему от случая. Все основные формулы и выводы следующего раздела этих лекций (математической статистики) основаны на результатах ЗБЧ

Мы рассмотрим здесь только некоторые важные теоремы, являющиеся яркими представителями в соответствующих областях и применяемые ниже, хотя число похожих теорем больше сотни (при самых различных предположениях, подходящих для разнообразных жизненных ситуаций).

1. Теорема Чебышева

Прежде всего, посмотрим (приводим без доказательства) на теорему П.Л. Чебышева (1821 – 1894). Он унаследовал крупное, процветающее поместье. Но, чтобы составить состояние своим сёстрам, необходимое как приданое (надо сказать, что в то время состояние переходило лишь по мужской линии), он играл на бирже. И очень успешно (и сумел-таки составить выгодную компанию своим сестрам и выдать их замуж), а в игре ему помогали его научные результаты!

Теорема Чебышева. *Пусть:*

а) $X_1, X_2, \ldots, X_n$ *- независимые случайные величины;*

б) существуют $M(X_i)$ *и* $D(X_i)$ *для всех* $i = 1,2,\ldots,n$*;*

в) $D(X_i) \le C$ *(при некотором положительном* C*) для всех* $i = 1,2,\ldots,n$.

Тогда:

$$\lim_{n\to+\infty} P\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{M(X_1) + M(X_2) + \ldots + M(X_n)}{n}\right| \le \varepsilon\right) = 1$$

при любом $\varepsilon > 0$.

Следствие. *Если независимые случайные величины* $X_1, X_2, \ldots, X_n$ *имеют одинаковые математические ожидания*

$$M(X_i) = a,\ i = 1,2,\ldots,n,$$

и

$$D(X_i) \le C,\ i = 1,2,\ldots,n,$$

то

$$\lim_{n\to+\infty} P\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - a\right| \le \varepsilon\right) = 1.$$

То есть (в общем случае) среднее арифметическое в пределе не отличается от математического ожидания (с вероятностью 1)!

Пример. Посмотрим на ситуацию в страховом бизнесе. Пусть X_i - убыток какого-то страхователя (того, кто страхуется) при наступлении страхового случая. Понятно, что все эти убытки имеют примерно одно и же математическое ожидание:

$$M(X_i) = a.$$

Тогда (по следствию из теоремы Чебышева) средний убыток всех страхователей:

$$\frac{X_1 + X_2 + \ldots + X_n}{n}$$

есть величина постоянная!

2. Центральная предельная теорема

Это на самом деле группа теорем, устанавливающих связь с нормальным законом распределения величины X с функцией плотности распределения вероятности (рис. 9.1):

$$\varphi(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}},$$

где $a = M(X),\ \sigma^2 = D(X)$ параметры распределения.

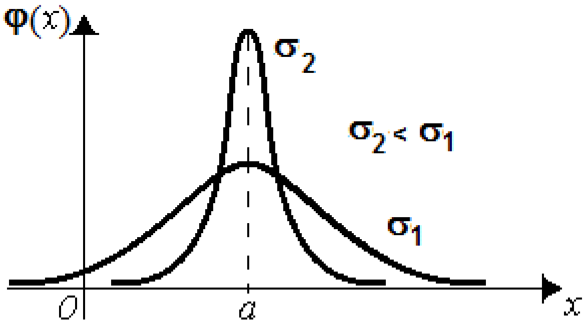


Рис. 9.1. Плотность распределения нормальной случайной величины

Приведём формулировку одной из таких теорем (приводим без доказательства).

Теорема Ляпунова. *Если:*

а) $X_1, X_2, \ldots, X_n$ *- независимые случайные величины;*

б) существуют $M(X_i) = a_i$ *и* $D(X_i) = \sigma_i^2$ *для всех* $i = 1,2,\ldots,n$;

в) существуют величины $M\left(|X_i - a_i|^3\right) = m_i$ *и*

$\lim\limits_{n\to+\infty} \dfrac{\sum_{i=1}^{n} m_i}{\left(\sum_{i=1}^{n}\sigma_i^2\right)^{3/2}} = 0$ *то закон распределения величины* $Y_n = \sum_{i=1}^{n} X_i$ *(при* $n \to +\infty$*) неограниченно приближается к нормальному закону с математическим ожиданием* $\sum_{i=1}^{n} a_i$ *и дисперсией* $\sum_{i=1}^{n} \sigma_i^2$, *т.е.:*

$$\lim_{n\to+\infty} P\left(\left|\frac{Y_n-\sum_{i=1}^{n}a_i}{\sqrt{\sum_{i=1}^{n}\sigma_i^2}}\right|\le z\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-\frac{t^2}{2}}dt=\frac{1}{2}+\Phi(z),$$

где $\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-\frac{t^2}{2}}dt$ *есть известная нам функция Лапласа.*

Смысл теоремы состоит в том, что чем сложней случайная величина, чем больше факторов, влияющих на ее значение, тем ближе она к нормально распределенной случайной величине.

Следствие. *Если независимые случайные величины* $X_1, X_2, \ldots, X_n$ *имеют одинаковые математические ожидания и дисперсии*

$$M(X_i)=a,\ D(X_i)=\sigma^2,\quad i=1,2,\ldots,n$$

и существуют величины $M\left(\left|X_i-a_i\right|^3\right)=m$*, то закон распределения величины*

$$Y_n=\frac{1}{n}\sum_{i=1}^{n}X_i$$

при $n\to+\infty$ *неограниченно приближается к нормальному закону с теми же параметрами* a *и* σ.

Пример. Пусть X_i - потребление электроэнергии жильцами квартиры номер i в многоквартирном, многоэтажном доме. Тогда по теореме Чебышева среднее потребление:

$$\frac{\sum_{i=1}^{n}X_i}{n}\approx a,$$

а по теореме Ляпунова величина:

$$Y_n=\frac{1}{n}\sum_{i=1}^{n}X_i$$

является случайной величиной, имеющей нормальный закон распределения (т.е. будет отличаться от величины a, как нормально распределённая случайная величина).

Пример. Представим величину Бернулли Y_n (количество наступления события A в серии из n испытаний) в виде суммы независимых величин, так называемых «индикаторов» каждого из испытаний:

$$Y_n=\sum_{k=1}^{n}X_i.$$

Здесь X_i - случайные величины - «индикаторы испытания»:

X_i	1	0
p_i	p	$q=1-p$

$$M(X_i)=p\cdot 1+q\cdot 0=p,\quad D(X_i)=p\cdot 1^2+q\cdot 0^2-p^2=pq.$$

Тогда по свойствам математического ожидания и дисперсии случайная величина Бернулли X_n будет иметь следующие параметры:

$$M(Y_n)=\sum_{i=1}^{n}M(X_i)=np,\quad D(Y_n)=\sum_{i=1}^{n}D(X_i)=npq,\ \sigma(Y_n)=\sqrt{npq},$$

а в соответствии с центральной предельной теоремой при большом количестве испытаний ($n\to+\infty$), она будет иметь распределение, близкое к нормальному закону, с параметрами $a=np$ и $\sigma=\sqrt{npq}$:

$$F(Y_n)\approx\frac{1}{2}+\Phi\left(\frac{Y_n-np}{\sqrt{npq}}\right).$$

3. Теорема Бернулли

Важнейшее методологическое значение для теории вероятностей и математической статистики имеет следующая теорема о частоте события. В серии испытаний Бернулли частоту события определим как:

$$\nu_n(A)=\frac{Y_n}{n}.$$

Теорема Бернулли.

Если количество испытаний велико, то частота события в испытании является нормальной случайной величиной с математическим ожиданием, равным вероятности события.

Действительно, поскольку частота события $\nu_n(A)$ в силу центральной теоремы при $n\to+\infty$ является величиной нормальной, а в силу основных свойств математического ожидания и дисперсии имеет математическое ожидание $M(\nu_n)=p$ и дисперсию $D(\nu_n)=pq/n$.

В соответствии с формулами Муавра – Лапласа, величина отклонения частоты и вероятности события имеет следующую вероятность:

$$P(|\nu_n-p|\le\varepsilon)=2\Phi\left(\frac{\varepsilon}{\sigma}\right)=2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)\xrightarrow[n\to\infty]{}1$$

для любого $\varepsilon>0$.

Таким образом, с ростом количества испытаний частота события стремится к его вероятности.

МАТЕМАТИЧЕЧСКАЯ СТАТИСТИКА

Лекция № 10

Выборочный метод

Для установления закономерностей, которым подчинены случайные события и случайные величины, теория вероятности, как и любая другая наука, обращается к опыту – наблюдениям, измерениям, экспериментам. Результаты наблюдений за случайными величинами объединяются в наборы статистических данных. Задачей математической статистики, раздела современной теории вероятностей, является разработка методов сбора и обработки статистических данных, а также их анализа с целью установления законов распределения наблюдаемых случайных величин [8, 9].

1. Генеральная и выборочная совокупность данных

Генеральной совокупностью является набор всех мыслимых статистических данных, при наблюдениях случайной величины:

$$x_Г = \{x_1, x_2, x_3, ..., x_N\} = \{x_i; i = 1, N\}.$$

Наблюдаемая случайная величина X называется признаком или фактором выборки. Генеральная совокупность есть статистический аналог случайной величины, ее объем N обычно велик, поэтому из нее выбирается часть данных, называемая выборочной совокупностью или просто выборкой

$$x_B = \{x_1, x_2, x_3, ..., x_n\} = \{x_i; i = 1, n\}, \quad x_B \subset x_Г, n \leq N.$$

Использование выборки для построения закономерностей, которым подчинена наблюдаемая случайная величина, позволяет избежать ее сплошного (массового) наблюдения, что часто бывает ресурсоемким процессом, а то и просто невозможным. Однако выборка должна удовлетворять следующим основным требованиям:

- выборка должна быть представительной, т.е. сохранять в себе пропорции генеральной совокупности,
- объем выборки должен быть небольшим, но достаточным для того, чтобы полученные результаты ее анализа обладали необходимой степенью надежности. В табл. 1 приводятся примеры генеральных и выборочных совокупностей.

Таблица 1

Генеральная совокупность	*Выборочная совокупность*
Данные переписи населения страны по разным признакам	Данные опроса случайных прохожих по тем же признакам
Времена работы электроламп, выпущенных заводом	Лабораторные данные о времени работы испытанных электроламп

Отметим, что в более строгом смысле выборку можно представить как многомерную случайную величину $\vec{X}_B = \{X_1, X_2, X_3, ..., X_n\} = \{X_i; i = 1, n\}$, у которой все компоненты X_i распределены одинаково и по закону распределения наблюдаемой случайной величины. В этом смысле выборочные значения x_B есть одна из реализаций величины $\vec{X}_B$.

2. Статистическое распределение выборки. Выборочный ряд, полигон, гистограмма и комулянта выборки

Возможные значения элементов выборки $x_B = \{x_i; i = 1, n\}$, называются вариантами x_j выборки, причем число вариант m меньше чем объем выборки n. Варианта может повторяться в выборке несколько раз, число повторения варианты x_j в выборке называется частотой варианты n_j. Причем $n_1 + n_2 + ... + n_m = n$. Величина $w_j = n_j / n$ называется относительной частотой варианты x_j.

Упорядоченный по возрастанию значений набор вариант совместно с соответствующими им частотами называется вариационно-частотным рядом выборки:

$$V_{xn} = \{x_j, n_j; j = 1, m\}; \quad V_{xw} = \{x_j, \nu_j; j = 1, m\}.$$

Ломаная линия, соединяющая точки вариационно-частотного ряда на плоскости (x, n) или (x, ν) называется полигоном частот.

Пример 1. Пусть дана выборка полуденных температур месяца мая своим вариационно-частотным рядом, приведенным в табл. 2.

Таблица 2

x_j	0	2	3	7	8	12	14	16	19	23	25	27	30
n_j	2	1	1	2	3	4	2	3	6	2	1	3	1

На рис.10.1 приводится полигон частот рассматриваемой выборки.
Вариационно-частотный ряд имеет существенный недостаток, а именно, ненаглядность полигона в случае малой повторяемости вариант, например, при наблюдении непрерывного признака его повторяемость в выборке маловероятна. Более общей формой описания элементов выборки, является гистограмма выборки.

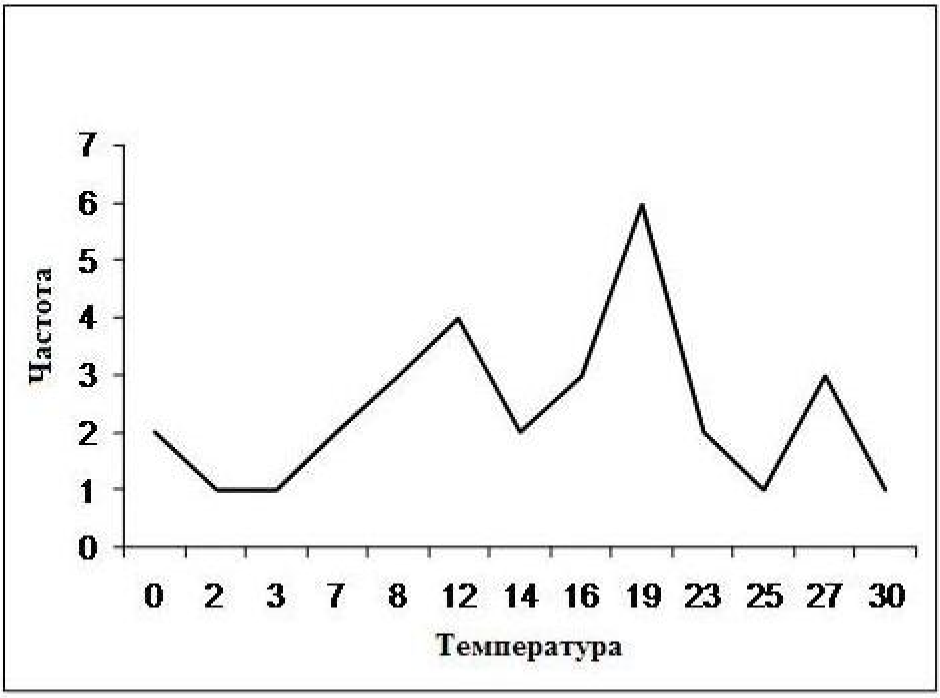


Рис.10.1 Полигон частот

Для построения гистограммы, разобьем интервал значений выборки $R = x_{\max} - x_{\min}$ на m интервалов $h_j = (x_j, x_{j+1})$ длины $h = R/m$ с границами $x_j = x_{\min} + h \cdot (j-1)$. Число элементов выборки x_B, попадающих в интервал, h_j называется частотой n_j интервала, кроме того вводятся следующие величины:

$\nu_j = n_j / n$ ~ относительная частота интервала,

$w_j = \nu_j / h_j$ ~ плотность относительной частоты интервала.

Совокупность интервалов, наблюдаемой в выборке случайной величины и соответствующих им частот, называется гистограммой выборки.

$$H_{xn} = \{h_j, n_j; j = 1, m\}, \quad H_{x\nu} = \{h_j, \nu_j; j = 1, m\}, \quad H_{xw} = \{h_j, w_j; j = 1, m\}$$

Для частот гистограммы выполнены следующие условия нормировки:

$$\sum_{j=1}^{m} n_j = n, \quad \sum_{j=1}^{m} \nu_j = 1, \quad \sum_{j=1}^{m} w_j h = 1$$

Число интервалов гистограммы m должно быть оптимальным, чтобы, с одной стороны, была достаточной повторяемость интервалов, а с другой стороны не должны сглаживаться особенности выборочной статистики. Рекомендуется значение $m \cong 1 + 3{,}2\lg(n)$. На плоскости (x, n) гистограмма представляется ступенчатой фигурой.

Пример 2. Наблюдаемые значения полуденной температуры месяца мая разбиты на 6 интервалов, соответствующая гистограмма задана следующей табл. 3:

Таблица 3

h_j	0-5	5-10	10-15	15-20	20-25	25-30
n_j	4	5	6	9	3	4

Гистограмма наблюдаемых температур приводится на рис. 10.2.

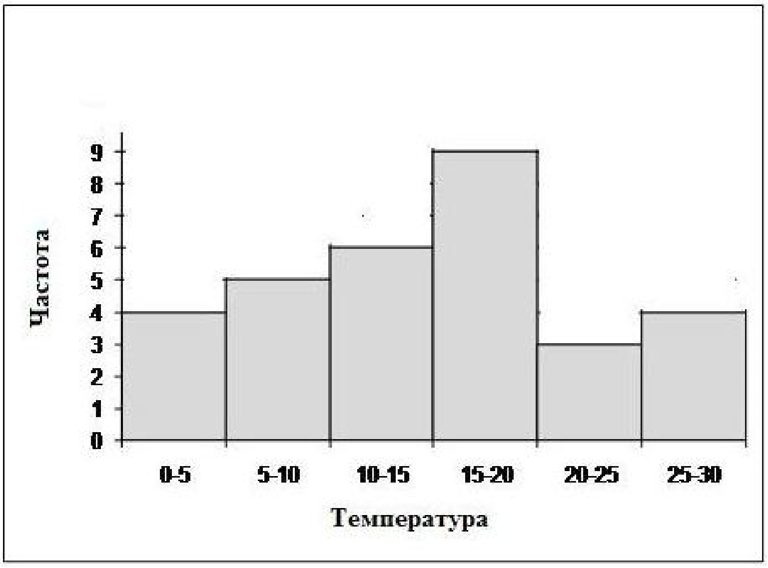


Рис. 10.2 Гистограмма частот

Выборочной или эмпирической функцией распределения называется функция $F_n(x)$, определяющая для каждого значения x относительную частоту события $\{X<x\}$ в выборке, которая вычисляется через сумму соответствующих частот:

$$F_n(x) = \frac{1}{n}\sum_{x_j<x} n_j \,.$$

В нашем примере выборочная функция распределения (иногда называемая комулянтой) приводится на рис.10.3.

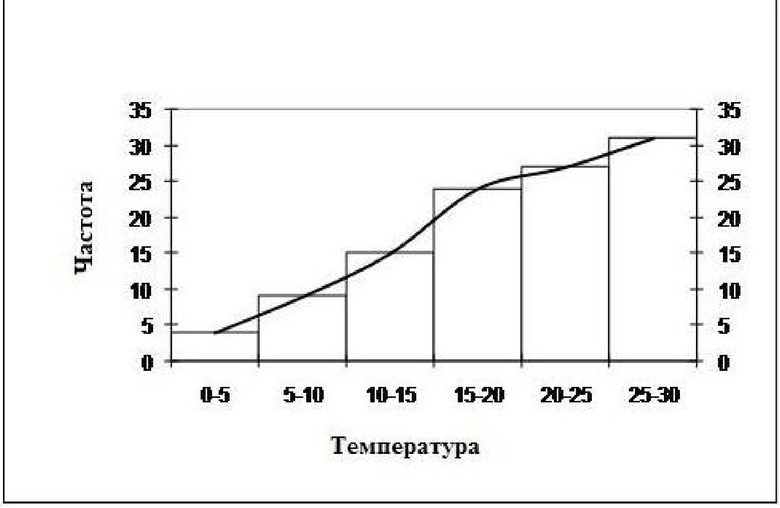


Рис. 10.3 Комулянта частот

При увеличении объема выборки относительная частота события приближается к вероятности этого события (теорема Бернулли), поэтому выборочная функция распределения $F_n(x)$ является оценкой теоретической функции распределения $F(x)$ для случайной величины X.

$$\lim_{n\to\infty} P\{\left|F_n\ (x) - F(x)\right| < \varepsilon\} = 1 \quad \text{для любого } x \text{ и } \varepsilon > 0.$$

Это утверждение строго доказано и носит форму теоремы Гливенко [7].

3. Выборочные характеристики

Помимо полигона и гистограммы выборка характеризуется следующими числовыми величинами:

Основные характеристики

$\bar{x}_B = \dfrac{1}{n}\sum_{i=1}^{n} x_i$ ~ выборочное среднее;

$D_B = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_B)^2$ ~ выборочная дисперсия;

$\sigma_B = \sqrt{D_B}$ ~ выборочное среднеквадратическое отклонение;

$S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_B)^2$ ~ исправленная выборочная дисперсия;

$S = \sqrt{S^2}$ ~ исправленное выборочное среднеквадратическое отклонение (выборочный стандарт).

Дополнительные характеристики

$a_k = \dfrac{1}{n}\sum_{i=1}^{n} x_i^{\kappa}$ ~ выборочный начальный момент порядка *k*;

$b_k = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_B)^{\kappa}$ ~ выборочный центральный момент порядка *k*;

Часто используются моменты 3-го и 4-го порядков в следующей форме:

$A_B = \dfrac{b_3}{\sigma_B^3}$ ~ выборочная асимметрия;

$E_B = -3 + \dfrac{b_4}{\sigma_B^4}$ ~ выборочный эксцесс.

В статистической практике рассматриваются так же групповые характеристики, например, в интервальных группах гистограммы выборки вычисляются средние интервальные значения и дисперсии.

Пример 3. Рассмотрим вычисление выборочных характеристик для выборки, представленной в примере 1. У этой выборки объема $n = 31$ имеется m=13 вариант x_j и столько же соответствующих им частот n_j, которые расположены в первых двух столбцах табл. 4.

Таблица 4

x_j	n_j	$n_j x_j$	$(x_j - \bar{x}_B)$	$n_j(x_j-\bar{x}_B)^2$	$n_j(x_j-\bar{x}_B)^3$	$n_j(x_j-\bar{x}_B)^4$
0	1	0,00	-14,87	221,15	-3288,65	48905,41
2	1	2,00	-12,87	165,66	-2132,23	27443,84
3	1	3,00	-11,87	140,92	-1672,86	19858,41
7	1	7,00	-7,87	61,95	-487,62	3838,07
8	5	40,00	-6,87	236,05	-1621,90	11144,01
12	6	72,00	-2,87	49,45	-141,98	407,63
14	2	28,00	-0,87	1,52	-1,32	1,15
16	2	32,00	1,13	2,55	2,88	3,25
19	5	95,00	4,13	85,24	351,98	1453,33
23	2	46,00	8,13	132,16	1074,35	8733,44
25	1	25,00	10,13	102,60	1039,21	10526,20
27	3	81,00	12,13	441,34	5353,03	64927,08
30	1	30,00	15,13	228,89	3462,85	52389,54
$\sum$	31	461,00		1869,48	1937,74	249631,36
		$\bar{x}_B$ = 14,87		60,31	62,51	8052,62
				7,77	0,13	-0,79

В последующих столбцах табл. 4, в соответствие с методом сводных таблиц, приводится расчет выборочных моментов и выборочных характеристик через варианты и частоты выборки:

$$\bar{x}_B = \frac{1}{n}\sum_{j=1}^{m} n_j x_{ji} = 14{,}87\ ;$$

$$D_B = \frac{1}{n}\sum_{j=1}^{m} n_j (x_j - \bar{x}_B)^2 = 60{,}31\,;\sigma_B = \sqrt{D_B} = 7.77\ ;$$

$$S^2 = \frac{n}{n-1} D_B = \frac{31}{30} 60{,}31 = 62.32\ ;\ S = \sqrt{62.32} = 7.89$$

Причем выполняется $a_0 = 1,\ \ a_1 = \bar{x}_B, a_2 = \overline{x_B^2},\ \ b_0 = 1,\ \ b_1 = 0,\ \ b_2 = D_B$.

$$b_3 = \frac{1}{n}\sum_{j=1}^{m} n_j (x_i - \bar{x}_B)^3 = 62.51\ ;\ \ A_B = \frac{b_3}{\sigma_B^3} = 0.13\ ;$$

$$b_4 = \frac{1}{n}\sum_{j=1}^{m} n_j (x_i - \bar{x}_B)^4 = 8052.62\ ;\ E_B = -3 + \frac{b_4}{\sigma_B^4} = -0.79\ .$$

Отметим, что все приведенные числовые характеристики являются случайными величинами, поскольку получены по случайно взятой выборке. На элементах другой выборки наблюдений над той же случайной величиной X числовые характеристики в общем случае изменят свое значение

Лекция № 11

Выборочные распределения

Если наблюдаемая случайная величина X является нормальной, т.е $X = N(a,\sigma)$, где a - математическое ожидание, σ - среднеквадратическое отклонение, то случайная величина среднего выборочного $\overline{X}_B = \frac{1}{n}\sum_{i=1}^{n} X_i$ так же является нормальной $X_B = N(a,\sigma/\sqrt{n})$. Здесь $X_i = N(a,\sigma)$ нормальные случайные величины, совпадающие с наблюдаемой величиной. Рассмотрим стандартные нормальные величины $\xi = N(0;1)$ в виде:

$$\xi_0 = \frac{X_B - a}{\sigma/\sqrt{n}}, \quad \xi_i = \frac{X_i - a}{\sigma}$$

и построим из них случайные величины Пирсона χ_n^2 и Стьюдента t_n.
Тогда получим [9,10]:

$$\chi_{n-1}^2 = \sum_{i=1}^{n} \xi_i^2 = \frac{1}{\sigma^2}\sum_{i}^{n} (X_i - a)^2 = \frac{nD_B}{\sigma^2} = \frac{n-1}{\sigma^2} S^2,$$

$$t_{n-1} = \frac{\xi_0}{\sqrt{\chi_n^2/(n-1)}} = \frac{X_B - a}{\sigma_B/\sqrt{n-1}} = \frac{X_B - a}{S/\sqrt{n}}.$$

Отсюда видно, что случайная величина выборочной дисперсии D_B распределена пропорционально «Хи-квадрат» случайной величине с *n-1* степенью свободы, а отклонение выборочного среднего от математического ожидания распределено пропорционально *t*-величине Стьюдента с *n-1* степенью свободы.

При сравнении двух выборок объемов n_1 и n_2 часто используется случайная величина Фишера со степенями свободы n_1 и n_2:

$$F_{n_1,n_2} = \frac{\chi_{n_1}^2/n_1}{\chi_{n_2}^2/n_2}.$$

1. Распределения Стьюдента и Пирсона

Распределения величин χ_n^2 и t_n известны аналитически в виде функции плотности распределения вероятностей

$$f_{\chi_n^2}(x) = \frac{1}{2^{n/2}\,\Gamma(n/2)} x^{0,5n-1} e^{-0,5x}, \qquad f_{T_m}(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(n/2)\sqrt{\pi n}}\left(1+\frac{x^2}{n}\right)^{-0,5(n+1)},$$

здесь $\Gamma(y) = \int_0^{\infty} e^{-t} t^{y-1} dt$ - функция Эйлера, обладающая свойством $\Gamma(y) = (y-1)\Gamma(y-1)$, в силу которого при целом положительном $y = k$ имеет место $\Gamma(\kappa) = (\kappa-1)\cdot\Gamma(\kappa-1) = (\kappa-1)\cdot(\kappa-2)...3\cdot 2\cdot 1 = (\kappa-1)!$

Графический вид функций плотности представлен ниже на рис. 11.1, 11.2 для различного количества степеней свободы.

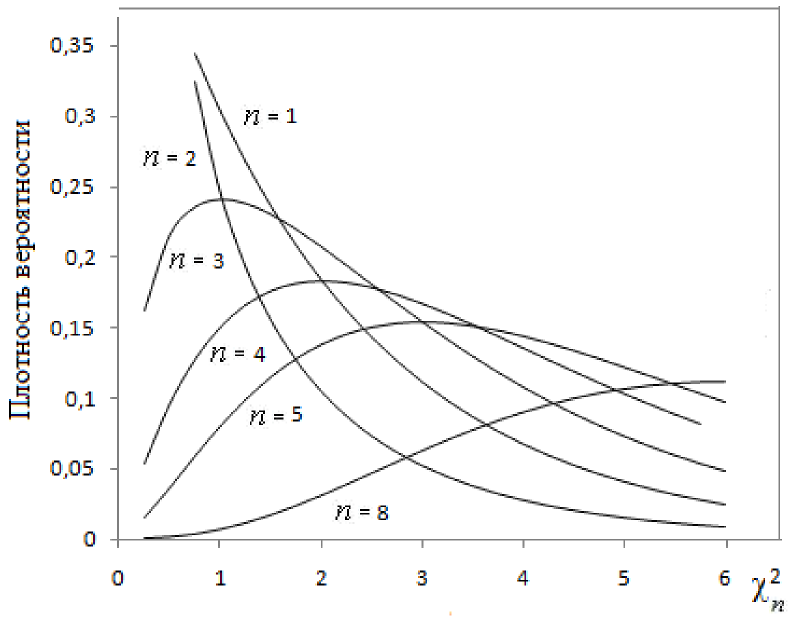


Рис.11.1 Кривые «Хи-квадрат» распределения

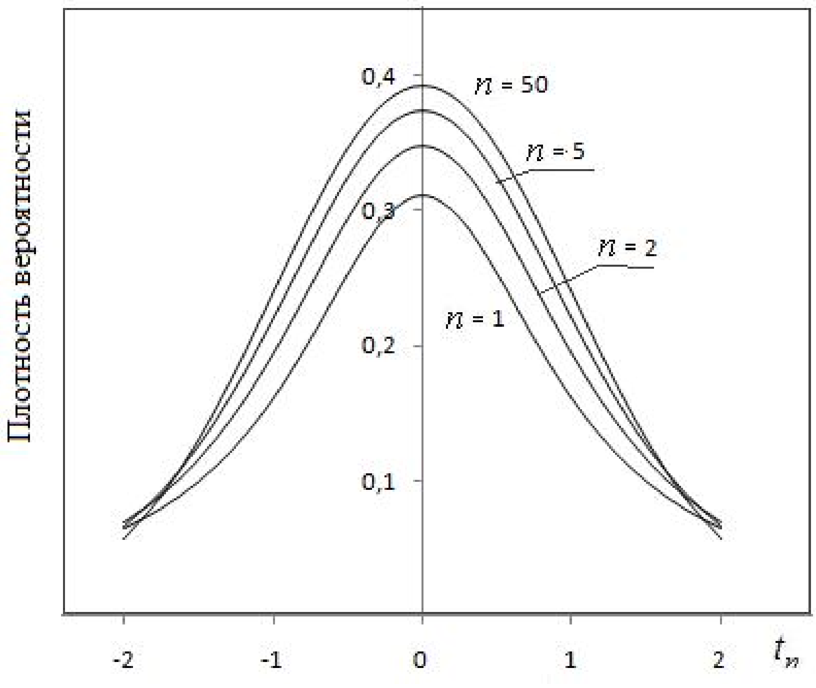


Рис.11.2 Кривые распределения Стьюдента

Числовые характеристики распределений «Хи-квадрат» и Стьюдента следующие:

$$M[\chi_n^2] = n, \quad D[\chi_n^2] = 2n, \qquad M[t_n] = 0, \; D[t_n] = \frac{n}{n-2}.$$

Можно заметить, что с ростом числа степеней свободы, указанные распределения будут приближаться к нормальному распределению, что соответствует центральной предельной теореме теории вероятностей.

2. Таблицы распределения выборочных величин

Обычно выборочные распределения задаются таблично в виде левосторонних функций распределения $F(x,n)$ и/или обратных к ним правосторонних квантилей $x_{кр} = x_{кр}(\alpha, n)$, графический смысл которых изображен на рис.11.3. Таблица значений этих величин известна [10] и они приводятся в приложениях 2-5.

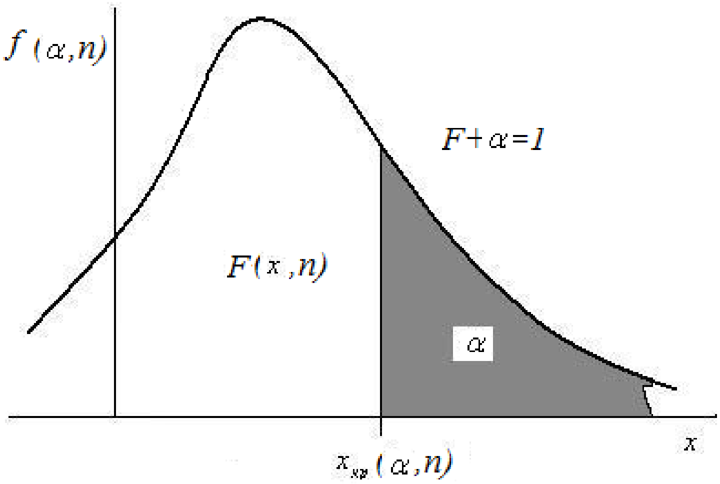


Рис.11.3 Правосторонняя квантиль $x_{кр} = x_{кр}(\alpha, n)$

В статистическом комплексе программ MS Excel-2007 эти распределения представлены следующими функциями:

$\alpha_{\chi 2}(x,n) = ХИ2РАСП(x,n)$ - правостороннее χ^2 распределение Пирсона,

$\chi^2_{кр}(\alpha,n) = ХИ2РАСПОБР(\alpha,n)$ - правосторонняя χ^2 квантиль Пирсона,

$\alpha_T(x,n) = СТЬЮДРАСП(x,n,1)$ - правостороннее t-распредел. Стьюдента,

$2\alpha_T(x,n) = СТЬЮДРАСП(x,n,2)$ - двухстороннее t –распределение,

$T_{кр}(\alpha/2,n) = СТЬЮДРАСПОБР(\alpha,n)$ - двухсторонняя t –квантиль,

$\alpha_F(x,n_1,n_2) = FРАСП(x,n_1,n_2)$ - правостороннее F-распределение

$F_{кр}(\alpha,n_1,n_2) = FРАСПОБР(\alpha,n_1,n_2)$ - правосторонняя квантиль Фишера.

Для работы с нормальной случайной величиной имеются следующие полезные функции:

$f(x) = НОРМРАСП(x,a,\sigma,л)$ - весовая функция

$F(x) = НОРМРАСП(x,a,\sigma,и)$ - интегральная функция

$x_{кр} = НОРМОБР(F,a,\sigma)$ - обратная интегральная функция;

$\varphi(x) = НОРМСТРАСП(x)$ - весовая функция со стандартными параметрами $(a=0, \sigma=1)$

$x_{кр} = НОРМСТОБР(F)$ - обратная стандартная интегральная функция;

$\Phi(x) = -0{,}5 + НОРМСТОБР(x)$ - Функция Лапласа.

Лекция № 12

Статистические оценки параметров распределения

Пусть распределение наблюдаемой случайной непрерывной величины X (признак генеральной совокупности), задается функцией плотности вероятности $f_X(x,\theta)$, где θ параметр или параметры распределения. Допустим, что вид функции $f_X(x,\theta)$ известен или ограничен некоторым классом функций, а параметр θ неизвестен и должен быть оценен по выборке $x_B = \{x_i, n\} = \{x_1, x_2, ...x_n\}$, где *n* – объем выборки.

1. Точечные оценки

Точечной статистической оценкой параметров распределения или характеристик наблюдаемой случайной величины X, называется построенная по данным выборки объема *n* величина:

$$\theta_n^* = \theta_n^*(x_1, x_2, ...x_n).$$

Оценка θ_n^* является так же случайной величиной, т.к. зависит от случайной выборки, поэтому ее можно представить как функцию от случайных величин $\Theta_n^* = \Theta_n^*(X_1, X_2, ...X_n)$, где X_i независимые случайные величины, распределенные так же как и сама величина X. Для того, чтобы оценки, получаемые по данным различных выборок соответствовали истинному значению параметра θ, оценка должна удовлетворять следующим требованиям.

Оценка должна быть *несмещенной*, т.е. ее математическое ожидание должно совпадать с истинным значением параметра для любого объема *n*

$$M(\theta_n^*) = \theta$$

или хотя бы асимптотически несмещенной: $M(\theta_n^*) \xrightarrow[n\to\infty]{} \theta$.

Оценка должна быть *состоятельной*, т.е. с ростом объема выборки оценка должна сходится по вероятности к истинному значению параметра:

$$P(\left|\theta_n^* - \theta\right| < \varepsilon) \xrightarrow[n\to\infty]{} 1 \quad \text{для любого } \varepsilon > 0.$$

Для состоятельности оценки достаточно выполнения следующего:

$$D(\theta_n^*) \xrightarrow[n\to\infty]{} 0,$$

действительно, из неравенства Чебышева $P(\left|\theta_n^* - \theta\right| < \varepsilon) \geq 1 - \dfrac{D(\theta_n^*)}{\varepsilon^2}$

для случайной величины θ_n^* следует состоятельность оценки.

Построенная оценка для использования на практике должна быть *эффективной*, т.е. ее дисперсия должна быть минимальной среди всех возможных оценок при фиксированном объеме выборки: $D(\theta^*_{n\,ef}) = \min D(\theta^*_n)$.

Величину дисперсии эффективной оценки можно найти, используя неравенство Рао-Крамера:

$$D(\theta^*_n) \ge \frac{1}{n \cdot I} = D(\theta^*_{n,ef}),$$

где $I(\theta) = M\left[\frac{\partial}{\partial\theta}\ln f(x,\theta)\right]^2 = \int\limits_{-\infty}^{\infty}\left[\frac{f'_\theta(x,\theta)}{f(x,\theta)}\right]^2 f(x,\theta)dx$ - информация Фишера. Коэффициент эффективности оценки $k_{ef} = D(\theta^*_{n\,ef}) / D(\theta^*_n)$ показывает степень эффективности оценки θ^*_n, *если* $k_{ef}(\theta^*_n) \xrightarrow[n\to\infty]{} 1$, то говорят об асимптотической эффективности оценки.

Отметим, что на практике не всегда удается удовлетворить всем перечисленным требованиям к оценке, но введенные свойства оценок всегда позволяют проранжировать имеющиеся оценки по их качеству.

В качестве примера рассмотрим оценки математического ожидания $M(X) = m$ и дисперсии $D(X) = d$ наблюдаемой случайной величины X.

Построим точечные оценки:

$$m^* = \overline{X}_B = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad d^* = D_B = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_B)^2$$

и рассмотрим их свойства. Поскольку $M(X_i) = m$ и $D(X_i) = d$ то можно вычислить, что для оценки m^* справедливо:

$$M(m^*) = m; \qquad D(m^*) = (d/n) \to 0 \quad \text{при } n \to \infty.$$

Из этого следует несмещенность и состоятельность оценки m^*.

Рассматривая же оценку d^* можно получить:

$$M(d^*) = \frac{n-1}{n} d \ne d \; ; \; D(d^*) \approx \frac{1}{n} \xrightarrow[n\to\infty]{} 0.$$

Из чего следует состоятельность, и смещенность оценки d^*. Смещеность оценки здесь легко может быть исправлена. Рассмотрим оценку:

$$d^* = \frac{n}{n-1} D_B = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_B)^2 = S^2.$$

Видим, что оценка $d^* = S^2$ является уже не только состоятельной, но и несмещенной так как $M(d^*) = d$. Величина S^2 называется исправленной (уточненной) выборочной дисперсией, а величина S исправленным среднеквадратическим выборочным отклонением (выборочный стандарт).

В заключении напомним, что относительная частота w_n появления события в независимых испытаниях Бернулли является несмещенной, состоятельной и эффективной оценкой неизвестной вероятности этого события $p^* = w_n$ (теорема Бернулли), а эмпирическая функция выборочного распределения $F_n(x)$ является состоятельной несмещенной оценкой неизвестной функцией распределения $F(x)$ наблюдаемой случайной величины $F^*(x) = F_n(x)$ (теорема Гливенко).

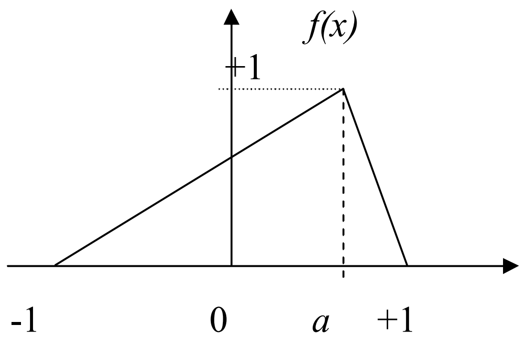
2. Методы построения точечных оценок

Метод моментов для нахождения точечных оценок неизвестных параметров распределения $f(x,\theta_1,\theta_2\cdots\theta_K)$ наблюдаемой в выборке случайной величины X, состоит в приравнивании теоретических моментов к выборочным моментам. Для нахождения r параметров θ_K начальные A_K или центральные B_K моменты до порядка r включительно приравниваются к соответствующим эмпирическим выборочным моментам a_K, b_K, тем самым получим систему r нелинейных уравнений метода моментов.

$$M(X^\kappa) = A_\kappa(\theta_1,\theta_2\cdots\theta_\kappa) = a_\kappa = \frac{1}{n}\sum_{i=1}^{n} x_i^\kappa \quad \text{или}$$

$$M[(X-\overline{X})^\kappa] = B_\kappa(\theta_1,\theta_2\cdots\theta_\kappa) = в_\kappa = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x}_B)^\kappa\,.$$

Например, построим оценку параметра, a случайной величины X, имеющей треугольное распределение (рис.12.1), по заданной выборке $x_B = \{x_i; i=1,n\}$, где n – объем выборки:



$$f(x,a) = \begin{cases} \dfrac{x+1}{a+1}\cdot при\cdot -1 < x < a \\ \dfrac{x-1}{a-1}\cdot при\cdot a < x < +1 \end{cases}$$

Рис. 12.1. Треугольное распределение

Поскольку неизвестный параметр один то, вычисляя и приравнивая только первые начальные теоретические и эмпирические моменты

$$M(X) = A_1(a) = \int_{-1}^{+1} x\cdot f(x,a)dx = \frac{a}{3}, \quad \overline{x}_B = a_1 = \frac{1}{n}\sum_{i=1}^{n} x_i\;,$$

получим оценку $a_n^* = 3\cdot\overline{x}_B$.

Метод моментов достаточно простой в применении и дает состоятельные оценки, однако их эффективность и несмещенность требуют дополнительных исследований.

Метод максимального правдоподобия основан на принципе правдоподобия, состоящем в том, что наблюдаемые в опыте события имеют большую вероятность, а маловероятные события практически не наблюдаемы. Вероятность наблюдения в опыте выборки $x_B = \{x_i; i = 1, n\}$ оценивается функцией правдоподобия

$$L(x_1, x_2, \ldots x_n, \theta) = f_X(x_1, \theta) \cdot f_X(x_2, \theta) \cdot \ldots \cdot f_X(x_n, \theta),$$

поскольку данная нам выборка уже получена в опыте, то она должна обладать максимальным правдоподобием. За оценку θ_n^* неизвестного параметра распределения θ принимается его значение, при котором функция правдоподобия максимальна, поэтому уравнение метода для нахождения оценки θ_n^*:

$$L(x_i, \theta_n^*) = \max_{\theta} L(x_i, \theta) \Rightarrow \frac{\partial}{\partial\theta} L(x_i, \theta) = 0, \text{ при условии } \frac{\partial^2}{\partial\theta^2} L(x_i, \theta)\Big|_{\theta=\theta_m^*} < 0.$$

Для решения этих уравнений чаще используется логарифм функции правдоподобия $l(x_1, x_2, \ldots x_n, \theta) = \ln[L(x_1, x_2, \ldots x_n, \theta)]$, поскольку максимум этих функций достигается при одном значении неизвестного параметра θ.

Например, рассмотрим случайную величину Пуассона $X_n^* > 0$ с плотностью распределения $f(x, \theta) = \frac{\lambda^x}{x!} e^{-\lambda}$, где λ неизвестный параметр распределения. Тогда функция правдоподобия и уравнение метода имеют вид:

$$L(x_i, \theta) = \frac{\lambda^{x_1}}{x_1!} e^{-\lambda} \, \frac{\lambda^{x_2}}{x_2!} e^{-\lambda} \ldots \frac{\lambda^{x_n}}{x_n!} e^{-\lambda} = \frac{\lambda^{x_1 + x_2 + \ldots + x_n}}{x_1! x_2! \ldots x_n!} e^{-n\lambda}$$

$$l(x_i, \theta) = \ln L(x_i, \theta) = (x_1 + x_2 + \ldots + x_n) \ln \lambda - n\lambda - \ln(x_1! \, x_2! \ldots x_n!)$$

$$\frac{\partial}{\partial\theta} l(x_i, \theta) = \frac{1}{\lambda} \sum_{i=1}^{n} x_i - n = 0 \quad \Rightarrow \quad \lambda_n^* = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}_B.$$

Доказано что метод максимального правдоподобия позволяет строить состоятельные и эффективные оценки.

Метод наименьших квадратов основан на идее минимизации суммы квадратов отклонения выборочных данных (или их функции) от строящейся оценки, он не требует знания закона распределения наблюдаемой случайной величины и кратко называется методом МНК.

Например, рассмотрим оценку дисперсии $D(X_i) = \sigma^2$ случайной величины по выборке $x_B = \{x_i; i = 1, n\}$, где n – объем выборки. Построим функцию для квадратов отклонения $R(\sigma^2) = \sum_{i=1}^{n} [(x_i - \bar{x}_B)^2 - \sigma^2]^2 \Rightarrow \min_{\sigma^2}$,

из условия минимума $\frac{dR}{d\sigma^2} = -2 \sum_{i=1}^{n} [(x_i - \bar{x}_B)^2 - \sigma^2] = 0$ и $\frac{d^2R}{d^2\sigma^2} = 2n > 0$ находим

$$(\sigma^2)^* = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}_B)^2.$$

3. Интервальные оценки и алгоритм их построения

В отличие от точечных оценок типа $\theta \approx \theta_n^*$ интервальные оценки задают интервал значений, где оцениваемый параметр находится с заданной вероятностью, т.е. это оценки типа $P(\left|\theta - \theta_n^*\right| \le \varepsilon) = \gamma$.

Надежностью оценки (доверительной вероятностью) называется вероятность γ, с которой оцениваемый параметр находится в интервале:

$$\theta_n^* - \varepsilon_\gamma \le \theta \le \theta_n^* + \varepsilon_\gamma \ .$$

Полуширина доверительного интервала ε_γ называется точностью оценки, соответствующей надежности γ. Для построения доверительного интервала (нахождения по γ величины ε_γ) необходимо знать закон распределения оценки случайной величины θ_n^*.

Пусть в выборке $x_B = \{x_i; i = \overline{1, n}\}$ наблюдается нормальная случайная величина $X = N(a, \sigma)$ с неизвестными параметрами распределения a и σ.

Построим доверительный интервал для математического ожидания a:

$$\bar{x}_B - \varepsilon_\gamma \le a \le \bar{x}_B + \varepsilon_\gamma,$$

принимая за точечную оценку a, величину $a^* = \bar{x}_B$ и учитывая что величина $(\bar{x}_B - a)/(S/\sqrt{n}) = t_{n-1}$ имеет распределение Стьюдента с $n-1$ степенью свободы. Решение уравнения $P(\left|\bar{x}_B - a\right| \le \varepsilon) = \gamma$ относительно ε при заданном значении γ эквивалентно решению уравнения:

$$P\left(\frac{\left|\bar{x}_B - a\right|}{S/\sqrt{n}} < \frac{\varepsilon}{S/\sqrt{n}}\right) = \gamma \text{ или } P(\left|t\right| < t_\gamma) = \gamma .$$

Его решение получим в виде $\varepsilon_\gamma = t_\gamma \cdot S/\sqrt{n}$, где $t_\gamma = t_\gamma(1-\gamma, n-1)$ двухсторонняя квантиль Стьюдента (рис. 12.2).

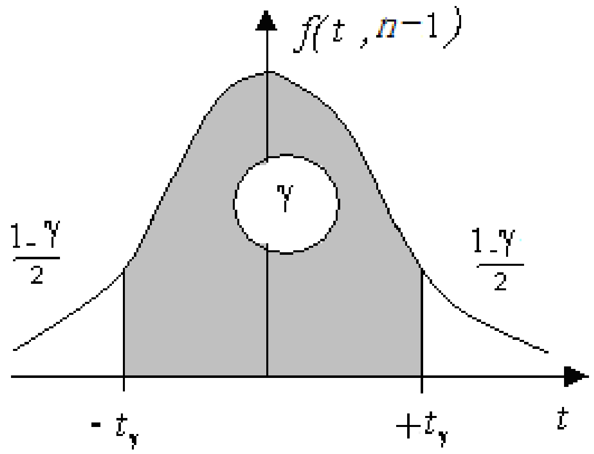


Рис. 12.2 Двухсторонняя квантиль Стьюдента

Построим теперь доверительный интервал для среднеквадратического отклонения σ:

$$S-\varepsilon_\gamma \le \sigma \le S+\varepsilon_\gamma .$$

Принимая за оценку σ величину $\sigma^* = S$ и учитывая, что величина $S^2(n-1)/\sigma^2 = \chi^2_{n-1}$ имеет χ^2-распределение с n - 1 степенью свободы. Решение уравнение $P(|S-\sigma| \le \varepsilon) = \gamma$ относительно ε при заданном параметре γ эквивалентно решению уравнения:

$$P(\chi^2_+ < \frac{S^2(n-1)}{\sigma^2} < \chi^2_-) = \gamma ,$$

тогда получим его решение в виде $S\sqrt{\frac{n-1}{\chi^2_-}} < \sigma < S\sqrt{\frac{n-1}{\chi^2_+}}$, где величины $\chi^2_\pm = F_{\chi^2 кр}(\frac{1\pm\gamma}{2}, n-1)$ являются правосторонними "хи-квадрат" квантилями (рис.12.3).

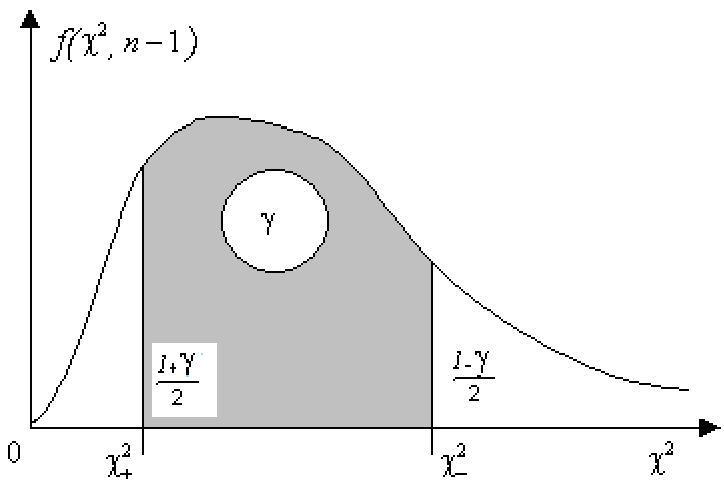


Рис. 12.3 Двухсторонняя "хи-квадрат"квантиль.

Пример: Пусть наблюдается выборка объемом n =16 со средним выборочным значением $\bar{x}_B = 20{,}2$ и выборочной дисперсией $D_B = 0{,}6$. Построить доверительные интервалы для неизвестного математического ожидания a и среднеквадратического отклонения σ для надежности $\gamma = 0{,}95$.

Исправленная дисперсия $S^2 = (16/15)\cdot 0{,}6 = 0{,}64$, а исправленное выборочное среднеквадратическое отклонение $S = 0{,}8$.

По таблице квантилей для распределения Стьюдента в приложении 3 находим $t_\gamma = t_\gamma(1-0{,}95{,}16-1) = 2{,}13$, тогда $\varepsilon_\gamma = 2{,}13\cdot 0{,}8/4 = 0{,}43$ и тогда доверительный интервал для математического ожидания a будет таким:

20,2-0,43< a <20,2+0,43 или 19,77< a <20,63.

По таблице для квантилей χ^2 - распределения в приложении 4 находим $\chi^2_+ = F_{\chi^2}(\frac{1+\gamma}{2}, n-1) = 6{,}26$ $\chi^2_- = F_{\chi^2}(\frac{\gamma-1}{2}, n-1) = 27{,}5$ и тогда $0{,}519 < \sigma < 1{,}238$.

Лекция № 13

Проверка статистических гипотез

Имея дело со случайными величинами, в различных областях человеческой деятельности часто приходится высказывать предположения о виде распределения случайной величины или о значениях ее параметров. Эти предположения строятся с целью прогнозирования поведения случайной величины и принятия решений в условиях неопределенности.

Статистической гипотезой называется любое предположение о виде распределения случайной величины $f_X(x,\theta)$ или/и о значении неизвестных параметров распределения θ.

$H = \{X \sim f_X(x,\theta); \theta = \theta_0\}$ – статистическая гипотеза

Высказанная статистическая гипотеза должна быть проверена по результатам наблюдений (измерений) случайной величины [11], в результате чего, гипотеза принимается или отвергается с определенной степенью риска совершить ошибку.

1. Простые и сложные статистические гипотезы

Статистическая гипотеза *H* называется *простой*, если она однозначно определяет закон распределения случайной величины X, например, для непрерывных величин в виде функции распределения $F_X(x,\theta)$ или функции плотности распределения вероятности $f_X(x,\theta)$ с определенными значениями параметров θ.

Гипотеза является *сложной,* если, в ней неизвестный закон распределения предполагается принадлежащим к некоторому допустимому множеству распределений.

Пример простой статистической гипотезы:

Длина ж/б перекрытия распределена по нормальному закону $N(a,\sigma)$ со следующими параметрами: математическое ожидание *a*=600см, среднеквадратическое отклонение σ=0,75см.

Пример сложной статистической гипотезы:

Толщина ж/б перекрытия распределена по нормальному закону $N(a,\sigma)$ со следующими параметрами: математическое ожидание *a*=20см, среднеквадратическое отклонение 0,5<σ <0,75 см.

2. Проверка статистических гипотез

Выдвинутая статистическая гипотеза *H* должна быть проверена. Как и в любой другой науке, критерием ее проверки является опыт, т.е. наблюдение (измерение) случайной величины. В математической статистике эти наблюдения представляются выборкой

$x_B = \{x_i, n\} = \{x_1, x_2, ... x_n\}$ объема *n*. Критерий проверки должен отвергать или принимать гипотезу по результатам наблюдения. В силу случайной природы наблюдаемых в выборке значений x_j, в результате применения критерия возможны следующие случайные события, их вероятности и совершаемые при этом ошибки представлены в следующей таблице:

Таблица 5

Результат проверки гипотезы	Вероятность	Наличие ошибки
Гипотеза *H* отвергается, когда она верна	α	ошибка I-рода
Гипотеза *H* принимается, когда она верна	$1-\alpha$	нет ошибки
Гипотеза *H* принимается, когда она не верна	β	ошибка II-рода
Гипотеза *H* отвергается, когда она не верна	$1-\beta$	нет ошибки

Из табл. 5 видно, что с вероятностью α при проверке может быть совершена ошибка I рода, когда отвергается верная гипотеза и с вероятностью β ошибка II рода, когда принимается неверная гипотеза. Поэтому первым требованием к критерию проверки является минимизация вероятности ошибок, однако здесь нужно отметить два существенных момента.

Во-первых, ошибки I и II рода могут иметь различную значимость с точки зрения их последствий. Так, например, для гипотезы *H* ={Партия ж/б перекрытий аварийно опасна и не должна поставляться на стройки} ошибка I-го рода приводит к поставке на стройку аварийно опасных изделий, что может повлечь человеческие жертвы. Ошибка же II-го рода здесь приводит к забраковыванию безопасной партии изделий, что влечет к экономическим потерям завода ЖБК. Ясно, что значимость ошибки I рода в приведенном примере выше, чем ошибки II рода, т.к. человеческие жертвы несравнимы с любыми экономическими потерями и значит недопустимы. Поэтому принято считать, что ошибки I рода более значимы чем ошибки II рода, если это не так, то проверяемую гипотезу необходимо переформулировать соответствующим образом (например, перейти к противоположной гипотезе).

Во-вторых, ошибки I и II рода находятся в некотором противоречии друг с другом, поскольку, если при построении критерия уменьшать вероятность одной из них, то вероятность другой будет возрастать. Так, например, при использовании гипотетического критерия “ничему не верю”, отвергающего любую гипотезу, ошибки II рода совершаться не будут ($\beta = 0$, “ложь не пройдет”), но при этом всегда будут совершаться ошибки I- рода ($\alpha = 1$”правда не установится“).

Учитывая сказанное, при построении критерия проверки статистической гипотезы необходимо сначала задаться допустимым уровнем риска на совершение ошибки I рода, как наиболее значимой, а затем минимизировать ошибки II рода.

3. Построение критерия проверки гипотезы

Пусть необходимо проверить простую гипотезу $H_0 = \{X \leftrightarrow f_X(x,\theta)\}$, состоящую в предположении о виде функции плотности распределения случайной величины X с вполне определенными параметрами θ. Построим критерий, однозначно принимающий или отвергающий проверяемую гипотезу по полученной в наблюдении за случайной величиной X выборке $x_B = \{x_1, x_2, ... x_n\}$ объема *n.* Помимо основной гипотезы H_0 (“нулевой”) рассмотрим еще одну или несколько альтернативных гипотез $H_1, H_2, H_3, ..., H_m$ каждая из которых противоречит основной. Альтернативные гипотезы необходимы при построении критерия проверки основной гипотезы для ее сравнения с имеющимися альтернативами (все познается в сравнении с чем-то).

Критерий проверки гипотезы состоит из двух составляющих:

Во-первых, в качестве критерия принимается некоторая случайная величина K, с известным распределением при условии справедливости основной $f_K(k/H_0)$ и хотя бы частично известным для альтернативных гипотез $f_K(k/H_j)$ j=1, ..*m*. Кроме того значения критерия должны быть вычисляемы по наблюдаемой выборке x_B, т.е. $k_{nab} = k(x_i)$.

Во-вторых, строится решающее правило для критерия проверки, согласно которому гипотеза будет приниматься или отвергаться. Для этого, назовем критической областью критерия те значения величины K, при которых гипотеза отвергается. Критическую область будем обозначать K_{kr}. Тогда решающее правило критерия проверки будет следующим:

$k_{nab} \in K_{kr} \Rightarrow H_0$ отвергается (по наблюдаемой выборке),

$k_{nab} \notin K_{kr} \Rightarrow H_0$ принимается (нет оснований отвергать гипотезу).

Точки значения критерия K, где критическая область критерия проверки K_{kr} отделяется от области принятия гипотезы, называются критическими точками критерия k_{kr}. Как построить критическую область критерия или, что равносильно, как найти критические точки критерия? Ниже рассмотрим ответ на этот вопрос.

Зададимся вероятностью α ошибки I-го рода, как наиболее значимой. Исключить такую ошибку при проверке гипотезы невозможно ($\alpha \neq 0$), но в вероятностных задачах это не является трагедией. На практике обычно эту вероятность задают достаточно малой величиной $\alpha = 0{,}05$; $\alpha = 0{,}025$; $\alpha = 0{,}005$ и называют уровнем значимости критерия.

Если из условия

$$P(k \in K_{kr}) = \int\limits_{K_{kr}} f_K(k / H_0)dx = \alpha,$$

можно определить критические точки k_{kr} однозначно, то задача построения критической области критерия решена. В противном случае, когда еще остается свобода выбора критических точек, рассмотрим влияние альтернативных гипотез. Поскольку величина:

$$\int\limits_{K_{kr}} f_K(k / H_j)dx = 1 - \beta_j$$

есть вероятность правильного отбрасывания H_0 при условии справедливости H_j, то ее называют мощностью критерия по отношению к альтернативной гипотезе H_j. Поэтому при заданном уровне значимости α, критическую область критерия нужно строить так, чтобы мощность критерия была максимальной, а именно:

$(1 - \beta_j) \Rightarrow \max$, для наиболее мощного критерия (НМК) относительно гипотезы H_j, максимизация проводится по параметрам сложной гипотезы**;**

$\min\ (1 - \beta_j) \Rightarrow \max$, для равномерно наиболее мощного критерия (РНМК), в случае наличия нескольких сложных гипотез.

Величина β_j - есть вероятность принять неверную гипотезу H_0 при условии справедливости альтернативной гипотезы H_j.

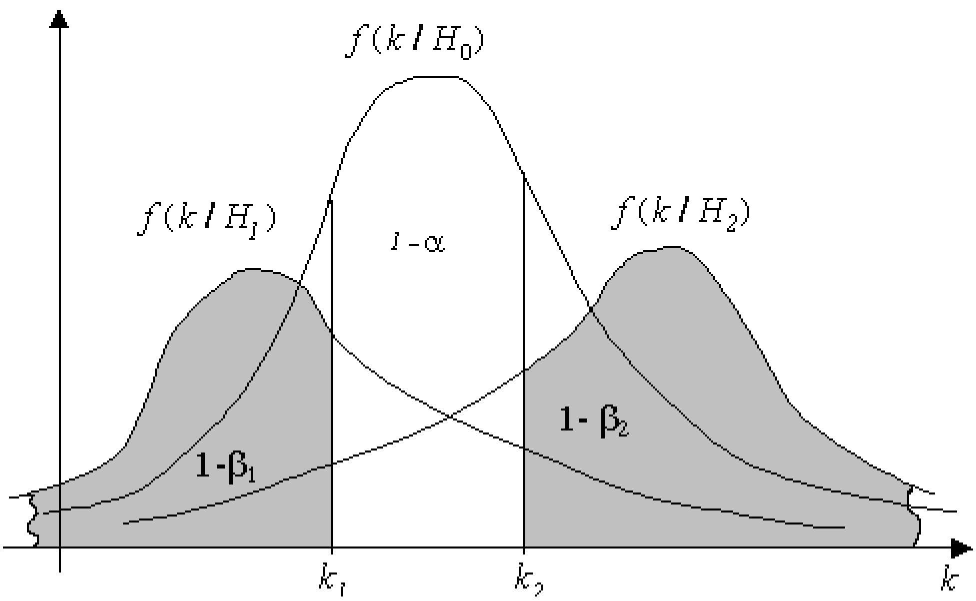


Рис. 13.1. Двухсторонняя критическая область критерия $K_{kr} = \{k > k_2, k < k_1\}$ при наличии двух альтернативных гипотез H_1, H_2.

На рис. 13.1 приведена графическая интерпретация алгоритма построения критической области одномерного критерия. Видим, что структура критической области, зависит от наличия альтернативных гипотез и их “расположения” относительно основной.

Лекция № 14

Примеры построения критериев проверки гипотез

1. Проверка гипотез о значении параметров распределения

Пусть случайная величина X распределена нормально по закону $N(a,\sigma)$ с неизвестными параметрами a,σ и наблюдается в выборке $x_B = \{x_i, n\} = \{x_1, x_2, ...x_n\}$ объема *n*. Нормальный закон распределения $N(a,\sigma)$ задается следующей функцией плотности распределения вероятности:

$$f_X(x,a,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(x-a)^2}{2\sigma^2}), \quad M[X]=a, \quad D[X]=\sigma^2.$$

По данным выборки могут быть получены выборочное среднее $\overline{x}_B$ и выборочный стандарт S **:**

$$\overline{x}_B = \frac{1}{n}\sum_{j=1}^{n} x_j, \qquad S = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(X_{cp} - x_j)^2}\,.$$

Эти величины являются случайными и по ним могут быть построены оценки математического ожидания $a = M[x]$ и дисперсии $\sigma = D[x]$ наблюдаемой в выборке случайной величины *X*.

Ниже проверим ряд простых статистических гипотез об истинных значениях параметров нормальной случайной величины X.

1.1. $H_0 = \{a = a_0\}$. Проверим сначала гипотезу о равенстве значения истинного (гипотетического) математического ожидания a некоторой величине a_0. Основная гипотеза тем самым будет следующей $H_0 = \{a = a_0\}$. В качестве критерия K возьмем случайную величину имеющую, при справедливости основной гипотезы, распределение Стьюдента с *n*-1 степенями свободы:

$$K = \frac{\overline{x}_B - a_0}{S/\sqrt{n}} = t_{n-1}$$

Задаваясь уровнем значимости α для проверяемой гипотезы H_0, будем строить критическую область K_{kr} в зависимости от вида единственной конкурирующей (альтернативной) гипотезы H_1 в следующих случаях:

Случай А: $H_1 = \{a > a_0\}$. В этом случае, при справедливости конкурирующей гипотезы ожидаем сдвиг вероятных значений критерия K в большую сторону (рис.14.1), поэтому критическая область критерия будет правосторонней $K_{kr} = \{k > k_{kr}\}$. Критическая точка k_{kr} однозначно определяется из условия равенства вероятности ошибки I-рода заданному уровню значимости $P(k > k_{kr}) = \alpha$. Решение этого уравнения $k_{kr} = t_{kr}(\alpha; n-1)$ представляет собой правостороннюю квантиль распределения случайной величины Стьюдента и приводится таблицей в приложении 3.

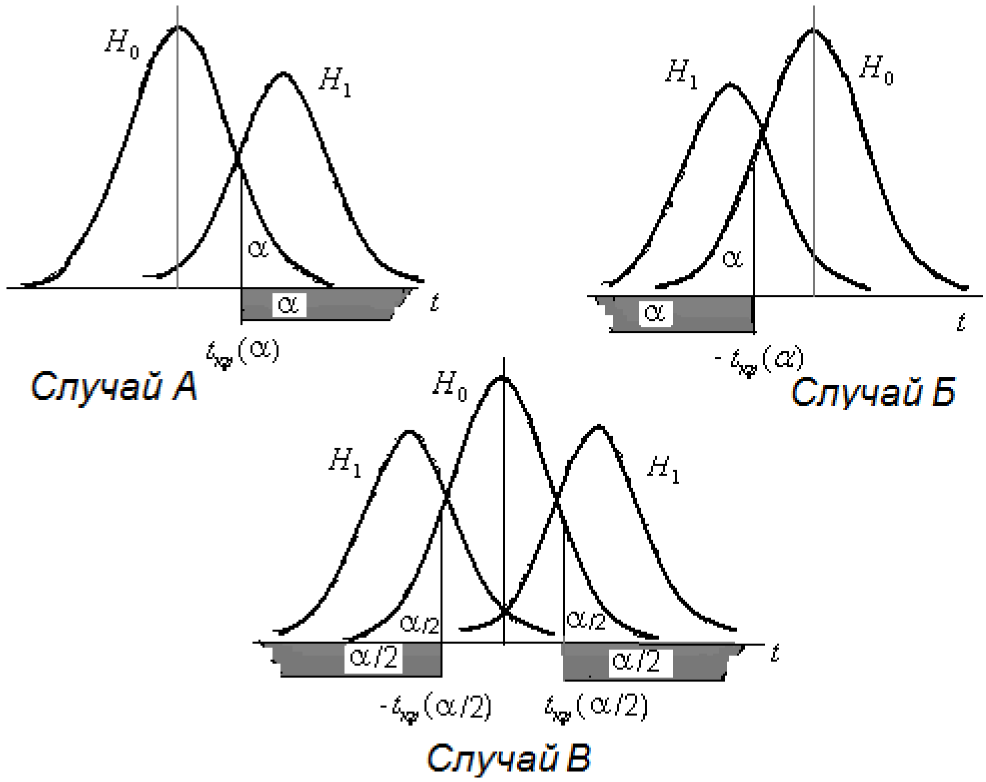


Рис. 14.1 Критические области гипотезы $H_0 = \{a = a_0\}$.

Случай Б: $H_1 = \{a < a_0\}$. В этом случае, критическая область критерия будет левосторонней $K_{kr} = \{k < k_{kr}\}$, а значения критерия отрицательными (рис.14.1). Критическая точка k_{kr} определяется из уравнения $P(k < k_{kr}) = \alpha$, решение которого, в силу симметрии распределения Стьюдента, будет следующим $k_{kr} = -t_{kr}(\alpha; n-1)$.

Случай В: $H_1 = \{a \neq a_0\}$. В этом случае критическая область критерия будет двухсторонней $K_{kr} = \{k < k_{kr1}; k > k_{kr2}\}$. Однако, здесь критические точки k_{kr1}, k_{kr2} не определяются однозначно из уравнения $P(k < k_{kr1}) + P(k > k_{kr2}) = \alpha$. Доказано [9], что при условии $P(k < k_{kr1}) = \alpha/2$ и $P(k > k_{kr2}) = \alpha/2$ мощность критерия $(1-\beta)$ по отношению к конкурирующей гипотезе H_1 будет максимальной. Тогда из этих уравнений критические точки находятся однозначно и представляют собой двухстороннюю квантиль распределения случайной величины Стьюдента:

$$k_{kr1} = -t_{kr}(\alpha/2; n-1), \quad k_{kr2} = t_{kr}(\alpha/2; n-1) .$$

Рассмотрим числовой пример: Пусть по выборке объема n=16 получена оценка математического ожидания наблюдаемой нормальной случайной величины $\overline{x}_B = 10{,}2$ и оценка среднеквадратического отклонения $S = 6{,}5$. Поскольку, каждая оценка есть величина случайная (получена по конкретной случайной выборке), то проверим гипотезу о том, что истинное математическое ожидание наблюдаемой величины равна 15 т.е. $H_0 = \{a = 15\}$. Зададимся уровнем значимости гипотезы $\alpha = 0{,}05$ и альтернативной гипотезой $H_1 = \{a \neq 15\}$. Наблюдаемое в выборке значение критерия $k_{nab} = (10{,}2 - 15) \cdot 4 / 6{,}5 = -2{,}954$. Критическая область K_{kr} двухсторонняя, а критические точки будут:

$$k_{kr1} = -t_{kr}(0{,}025;15) = -2{,}13\,;\; k_{kr2} = +t_{kr}(0{,}025;15) = +2{,}13\,.$$

Видим, что k_{nab} принадлежит критической области и значит, гипотеза *отвергается*, т.е. отличие наблюдаемого значения математического ожидания от гипотетического значительны.

1.2. $H_0 = \{\sigma^2 = \sigma_0^2\}$ Проверим теперь гипотезу о том, что истинная (гипотетическая) дисперсия случайной величины равна σ_0^2. Проверяемая гипотеза $H_0 = \{\sigma^2 = \sigma_0^2\}$ В качестве критерия возьмем одномерную случайную величину K, имеющую распределение «хи-квадрат» с n-1 степенями свободы:

$$K = \frac{S^2}{\sigma_0^2}(n-1) = \chi_{n-1}^2 .$$

Здесь S^2 оценка σ_0^2, полученная по выборке $x_B = \{x_i, i = 1, n\}$.

Задаваясь уровнем значимости α для проверяемой гипотезы H_0, будем строить критическую область K_{kr} в зависимости от вида единственной конкурирующей (альтернативной) гипотезы H_1 в следующих случаях (рис.14.2):

Случай А: $H_1 = \{\sigma^2 > \sigma_0^2\}$. В этом случае, при справедливости конкурирующей гипотезы ожидаем сдвиг наиболее вероятных значений критерия K в большую сторону, поэтому критическая область будет правосторонней.

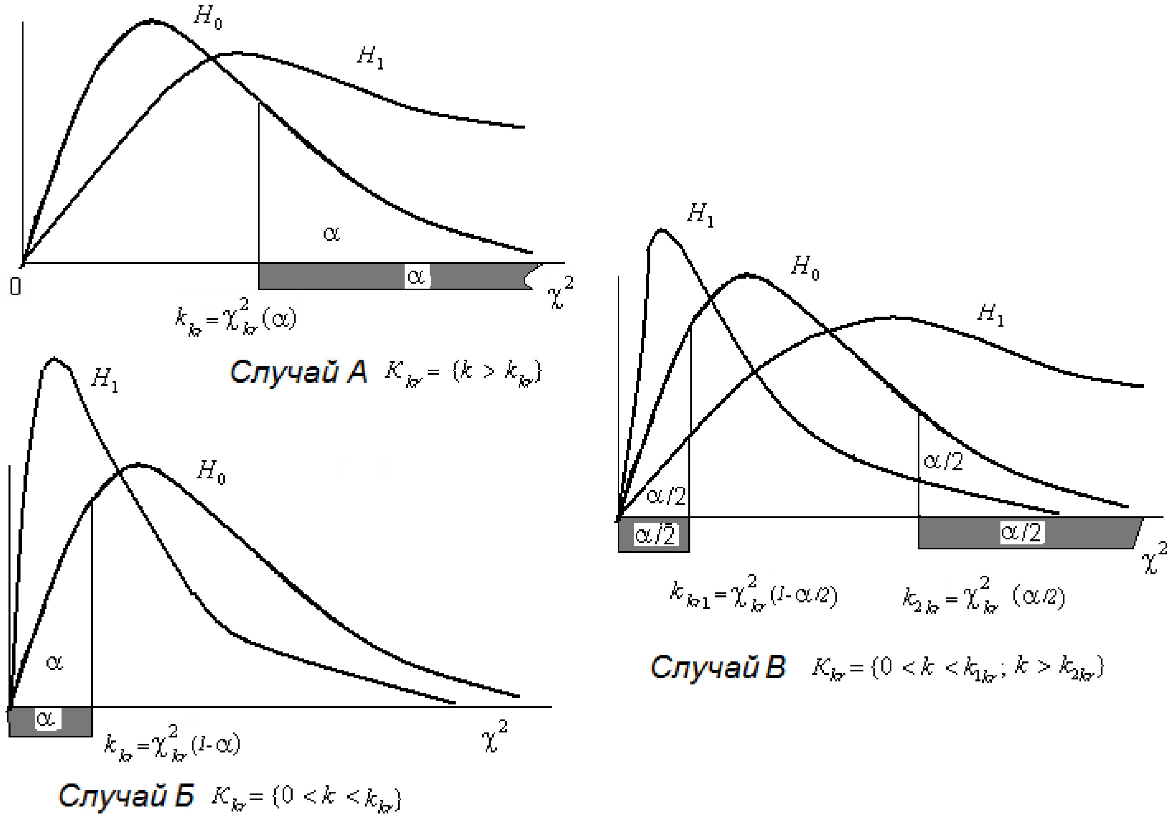


Рис. 14.2 Критические области гипотезы $H_0 = \{\sigma^2 = \sigma_0^2\}$.

Критическая точка k_{kr} здесь однозначно определяется согласно общему подходу к построению критических областей критерия из условия равенства вероятности ошибки I-рода заданному уровню значимости α:

$$P(k > k_{kr}) = \int_{k_{kr}}^{\infty} \chi^2(k; n-1)dk = \alpha .$$

Решение этого уравнения $k_{kr} = \chi^2_{kr}(\alpha; n-1)$ находятся однозначно, и представляет собой правостороннюю квантиль «хи-квадрат» распределения случайной величины и приводится в приложении 4.

Случай Б: $H_1 = \{\sigma^2 < \sigma_0^2\}$. В этом случае критическая область критерия будет левосторонней, а критическая точка однозначно определяется из уравнения :

$$P(k < k_{kr}) = \int_{0}^{k_{kr}} \chi^2(k; n-1)dk = \alpha$$

Левосторонняя критическая точка может быть легко выражена через функцию для правосторонней критической точки. Действительно, т.к. $P(k < k_{kr}) + P(k > k_{kr}) = 1$, то $P(k > k_{kr}) = 1 - \alpha$ и тогда решение для левосторонней точки будет следующим $k_{kr} = \chi^2_{kr}(1-\alpha; n-1)$.

<u>Случай В</u>: $H_1 = \{\sigma^2 \neq \sigma_0^2\}$. В этом случае, объединяющем два предыдущих случая, критическая область критерия будет

двухсторонней $K_{kr}=\{k<k_{kr1};k>k_{kr2}\}$. Однако, здесь критические точки k_{kr1}, k_{kr2} не определяется однозначно из уравнения

$$P(k<k_{kr1})+P(k>k_{kr2})=1-\int_{k_{kr1}}^{k_{kr2}}\chi^2(k,n-1)dk=\alpha.$$

Доказано [9], что при условиях $P(k<k_{kr1})=\alpha/2,\ P(k>k_{kr2})=\alpha/2$ мощность критерия $(1-\beta)$ по отношению к конкурирующей гипотезе H_1 будет максимальной, тогда из этих двух условий критические точки находятся однозначно:

$$k_{kr1}=\chi^2_{kr}(1-\alpha/2;n-1);\qquad k_{kr2}=\chi^2_{kr}(\alpha/2;n-1).$$

Рассмотрим числовой пример: Пусть по выборке объема *n*=15 получена оценка дисперсии наблюдаемой нормальной случайной величины $S^2=40{,}25$ или оценка среднеквадратического отклонения $S=6{,}5$. Поскольку, каждая оценка есть величина случайная (получена по конкретной случайной выборке), то проверим гипотезу о том, что истинная дисперсия наблюдаемой величины равна 36, т.е. $H_0=\{\sigma^2=36\}$. Зададимся уровнем значимости гипотезы H_0 $\alpha=0{,}05$ и альтернативной гипотезой $H_1=\{\sigma^2\neq 36\}$.

Наблюдаемое значение критерия $k_{nab}=(15-1)40{,}25/36=15{,}653$. Критическая область $K_{kr}=\{k<k_{1kr};k>k_{2kr}\}$ двухсторонняя, а критические точки будут:

$$k_{kr1}=\chi^2_{kr}(1-0{,}025;14)=5{,}63;\ \ k_{kr2}=\chi^2_{kr}(0{,}025;14)=26{,}1.$$

Видим, что $k_{nab}=15{,}653$ не принадлежит критической области и значит, гипотеза *принимается*, т.е. отличия наблюдаемого значения дисперсии от гипотетического незначительны. Если бы, такая оценка дисперсии была получена по выборке меньшего объема *n*=7, то

$$k_{kr1}=\chi^2_{kr}(1-0{,}025;6)=14{,}4;\qquad k_{kr2}=\chi^2_{kr}(0{,}025;6)=1{,}24.$$

тогда наблюдаемое значение критерия $k_{nab}=15{,}653$ попадает в критическую область и тогда проверяемая гипотеза *отвергается*.

Отметим, что при проверке гипотез $H_0=\{a=\overline{x}_B\}$ и $H_0=\{\sigma^2=S^2\}$ при уровне значимости α будут построены двухсторонние критические области такими, что область принятия гипотез $\overline{K}_{kr}$ совпадет с доверительными интервалами, построенными с надежностью $\gamma=1-\alpha$.

2. Критерий согласия Пирсона

Критериями согласия называются критерии проверки статистических гипотез о виде распределения случайной величины. Проверяемая гипотеза имеет вид:

$$H_0 = \{X \sim f_X(x, \theta_1, \theta_2, ...\theta_r),$$

где $\theta_1, \theta_2, ...\theta_r$ - принятые в гипотезе параметры распределения. Пирсон предложил и обосновал следующий критерий проверки гипотезы H_0 по отношению к единственной альтернативной противоположной гипотезе $H_1 = \overline{H}_0$.

Пусть по полученной выборке $x_B = \{x_i, i = 1, n\} = \{x_1, x_2, ...x_n\}$ построена гистограмма наблюдаемых частот $H_{Xn} = \{h_j, n_j; j = 1, m\}$. Построим, так же теоретические частоты n_j^T для интервалов h_j при условии справедливости проверяемой гипотезы H_0. Теоретические частоты вычисляются через вероятность P_j нахождения случайной величины X в интервале $h_j = (x_j, x_{j+1})$ по формуле:

$$\frac{n_j^T}{n} \approx P_j = F(x_{j+1}) - F(x_j) = \int_{x_j}^{x_{j+1}} f_X(x, \theta_s) dx \approx hf(x_{j+0.5}, \theta_s),$$

где $F(x_j)$ - функция распределения для случайной величины X, h – шаг интервалов гистограммы, $x_{j+0.5} = 0{,}5 \cdot (x_j + x_{j+1})$ центры интервалов h_j гистограммы. Таким образом, получим теоретические частоты $n_j^T \approx n \cdot P_j$. Показано [9], что величина :

$$\sum_{j=1}^{m} \frac{(n_j - n_j^T)^2}{n_j^T} = \chi^2_{m-r-1},$$

при достаточно большом объеме выборки имеет «хи-квадрат» распределение с $m - r - 1$ степенями свободы и может быть использована в качестве критерия для проверки гипотезы H_0. Задаваясь уровнем значимости α можем однозначно определить правостороннюю критическую область критерия из уравнения

$$P(\chi^2 > \chi^2_{kr}) = \alpha$$

Его решение представляет собой правостороннюю квантиль «хи-квадрат» распределения $\chi^2_{kr} = \chi^2_{kr}(\alpha, m - r - 1)$ и приведено в приложении 4.

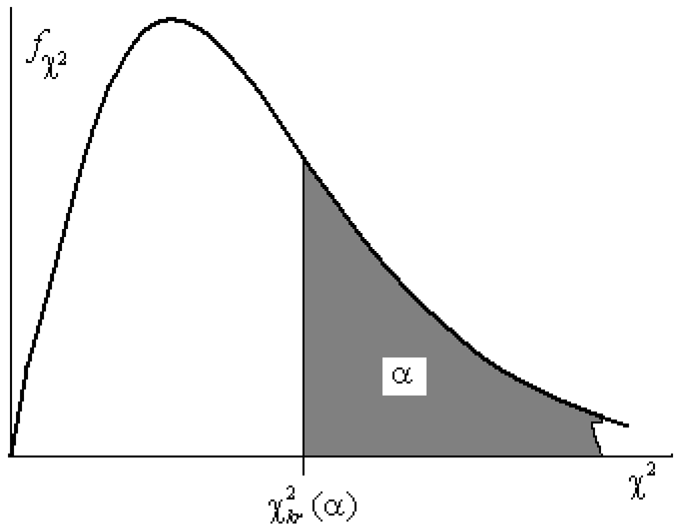


Рис. 14.3. Критическая область критерия Пирсона.

Определив, таким образом, критическую точку χ^2_{kr}, сравним ее с наблюдаемым значением χ^2_{nab} получим правило проверки гипотезы:

- если $\chi^2_{nab} < \chi^2_{kr}$, то гипотеза принимается (отклонения теоретических и наблюдаемых частот незначительны),
- если же $\chi^2_{nab} > \chi^2_{kr}$, то гипотезу необходимо отвергнуть (отклонения частот значительны).

Числовой пример: Проверим гипотезу о нормальном распределении полуденных температур месяца мая для выборки, приведенной в лекции 10, при уровне значимости гипотезы $\alpha = 0{,}05$. Вычислив выборочные характеристики $\overline{x}_B = 14{,}6$ и $S = 7{,}5$, примем их за оценки параметров нормального распределения. Таким образам проверяемая гипотеза такова:

$$H_0 = \{X = N(a, \sigma); a = \overline{x}_B; \sigma = S\}.$$

Учитывая, что для нормальной случайной величины X функция распределения имеет вид $F(x) = \frac{1}{2} + \Phi(\frac{x-a}{\sigma})$, где $\Phi(x)$ - функция Лапласа (приложение 2), то для теоретических частот получим формулу:

$$n_j^T \approx n \cdot [\Phi\left(\frac{x_{j+1} - \overline{x}_B}{S}\right) - \Phi\left(\frac{x_j - \overline{x}_B}{S}\right)],$$

где x_j, x_{j+1} – соответственно левая и правая границы каждого из интервалов h_j разбиения данных в гистограмме. Все результаты приведем в таблице 8 и на рис.14.4.

Таблица 8.

h_j	0-5	5-10	10-15	15-20	20-25	25-30	$\sum$
n_j	3	6	8	7	3	4	31
n_j^T	2,31	5,26	7,79	7,53	4,74	1,95	29,6
χ^2_{nab}	0,205	0,105	0.006	0,037	0,639	2,171	3,162

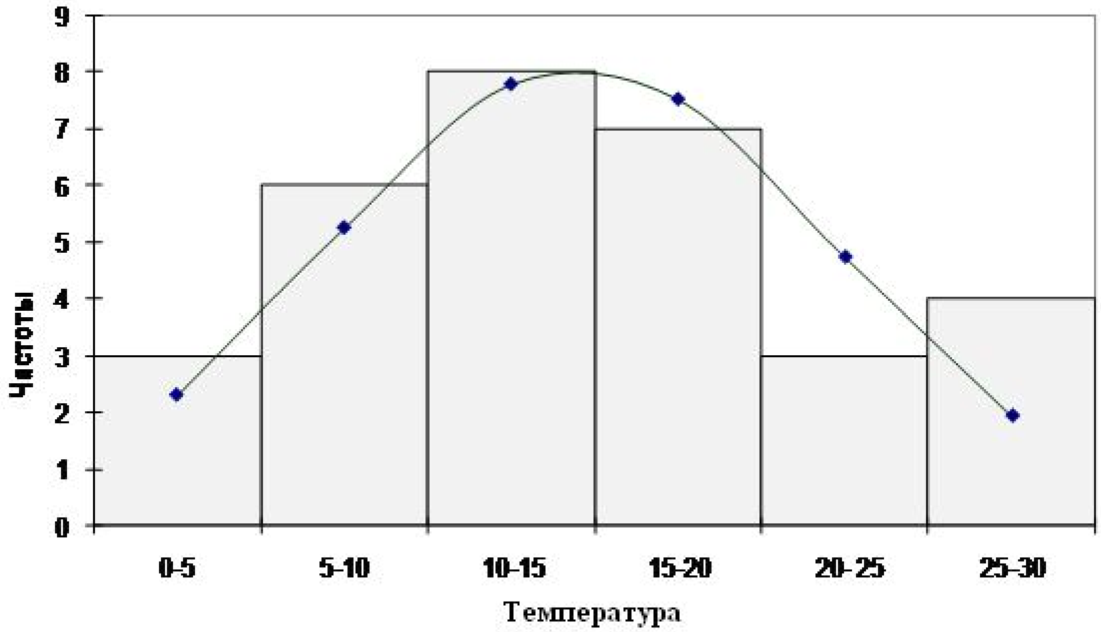


Рис. 14.4. Гистограмма наблюдаемых частот и кривая теоретических частот.

По заданному уровню значимости проверяемой гипотезы H_0 определим критическую точку распределения «хи-квадрат» используя приложение 4. Получим, что $\chi^2_{kr} = \chi^2_{kr}(0{,}05;6-2-1) = 7{,}8$.

Поскольку $\chi^2_{nab} = 3{,}162 < \chi^2_{kr} = 7{,}8$, то гипотеза H_0 принимается (нет оснований ее отвергнуть), т.к. отклонения частот незначительны.

Примеры заданий для проверки различных статистических гипотез для самостоятельной работы студентов приводятся в [12].

Лекция № 15

Элементы корреляционного анализа

Две случайные величины X и Y могут быть независимыми между собой, зависимыми строго функционально $Y = \varphi(X)$ или зависимыми статистически. При статистической зависимости между случайными величинами распределение одной из величин зависит от того, какое значение имеет другая случайная величина. Степень статистической зависимости величин X и Y характеризует теоретический коэффициент корреляции Пирсона

$$\rho_{XY} = \frac{M(X \cdot Y) - M(X) \cdot M(Y)}{\sqrt{D(X) \cdot D(Y)}},$$

обладающий следующими свойствами:

1) его значение по модулю не превышает единицы $-1 \le \rho_{XY} \le 1$.
2) для независимых величин X и Y $\rho_{XY} = 0$,
3) для линейно зависимых величин $\rho_{XY} = \pm 1$.

Сама статистическая зависимость описывается функциями условного распределения, например, для непрерывных случайных величин функциями плотности условного распределения $f_X(x|y)$ или $f_y(y|x)$. Однако нахождение этих функций и их практическое использование обычно затруднено и малоэффективно. Чаще статистическая зависимость рассматривается в более простом виде, в виде функциональной зависимости числовых характеристик одной из величин от значения другой величины. Такая зависимость называется корреляционной и описывается функциями регрессии $\hat{Y}(x)$ или $\hat{X}(y)$. Так например, наиболее часто используется регрессия в форме условного математического ожидания:

$$M(Y|x) = \int_{\Omega_y} y f_y(y|x) dy = \hat{Y}(x).$$

Корреляционная зависимость приближает статистическую зависимость функциональной зависимостью и имеет следующий вид:

$$Y = \hat{Y}(x) + \varepsilon.$$

Здесь Y - объясняемая переменная, x - значение объясняющей переменной X, а ε - случайная величина ошибки (невязки) корреляции с нулевым математическим ожиданием $M(\varepsilon) = 0$ при любом значении x. Дисперсия же ошибки $D(\varepsilon)$ не нулевая, но при «хорошей» функции регрессии она не должна быть большой, и не должна зависеть от переменной x. Построение таких функций регрессии является задачей регрессионного анализа.

Для приближенного построения функции регрессии будем искать наилучшее в определенном, но довольно широком, m-параметрическом классе функций $U_m\{\hat{y}(x,\theta_1,\theta_2,..\theta_m)\}$ таким образом, что бы дисперсия ошибки $D(\varepsilon,\theta_1,\theta_2,..\theta_m)$ как функция от параметров θ_k была минимальной. Такое приближение называется среднеквадратической регрессией в классе U_m. Для приближенного построения функции регрессии можно так же воспользоваться данными наблюдений за величинами X и Y, полученными в выборке (x_i, y_i) объема n. Такие оценки для функции регрессии $\hat{y}(x)$ ищутся так же в кассе U_m, имеют минимальное суммарное отклонение от наблюдаемых значений y_i, строятся методом наименьших квадратов и называются выборочной среднеквадратической регрессией.

1. Эмпирическая линейная среднеквадратическая регрессия

Линейная регрессия является простейшей регрессионной моделью, согласно которой функция регрессии является линейной 2-х параметрической функцией:

$$\hat{y}(x) = a + вx\,,$$

где $a, в$ - неопределенные коэффициенты, которые оценим по наблюдаемым данным. Пусть имеется двухфакторная выборка n наблюдений (x_i, y_i) за величинами X и Y, которую будем называть корреляционным полем. Помимо выборочных средних значений $\bar{x}, \bar{y}$ и выборочных дисперсий $D_x = \sigma_x^2,\ D_y = \sigma_y^2,$ вычислим так же среднее произведение $\overline{xy}$ и выборочный (эмпирический) коэффициент корреляции $r_{xy} = \dfrac{\overline{xy} - \bar{x}\cdot\bar{y}}{\sigma_x \cdot \sigma_y}$, который является выборочным аналогом теоретического коэффициента корреляции Пирсона ρ_{XY}.

Построим коэффициенты $a, в$ методом наименьших квадратов. Для этого найдем такие значения $a, в$, которые минимизируют сумму квадратов отклонения y_i и $\hat{y}_i = \hat{y}(x_i)$, то есть ошибки $e_i = y_i - \hat{y}_i$

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(y_i - a - bx_i)^2 \Rightarrow \min_{a,b}.$$

Из необходимых условий минимума найдем искомые значения $a, в$:

$$\frac{\partial}{\partial a}\sum_{i=1}^{n} e_i^2 = -2\sum_{i=1}^{n}(y_i - a - bx_i) = 0\,; \quad \Rightarrow \bar{y} = a + в\bar{x}\,; \ \Rightarrow a = \bar{y} - в\bar{x}\,,$$

$$\frac{\partial}{\partial b}\sum_{i=1}^{n} e_i^2 = -2\sum_{i=1}^{n}(y_i - a - bx_i)x_i = 0\,; \Rightarrow \overline{xy} = a\bar{x} + в\overline{x^2}\,; \Rightarrow в = \frac{\overline{xy} - \overline{x}\overline{y}}{\overline{x^2} - \bar{x}^2}.$$

Через выборочный коэффициент корреляции r_{xy}, коэффициент $в$ представим в форме $в = r_{xy}\dfrac{\sigma_y}{\sigma_x}$, а уравнение выборочной линейной среднеквадратической регрессии имеет одну из следующих форм:

$$\hat{y}(x) = a + вx; \qquad \hat{y}(x) = \bar{y} + в(x - \bar{x});$$

$$\hat{y}(x) - \bar{y} = r_{xy}\frac{\sigma_y}{\sigma_x}(x - \bar{x}); \qquad \frac{\hat{y}(x) - \bar{y}}{\sigma_y} = r_{xy}\frac{(x - \bar{x})}{\sigma_x}.$$

2. Свойства линейной регрессии и коэффициента корреляции

Построенная выборочная линейная среднеквадратичная регрессия является простейшим приближение корреляционной зависимости, показывает тенденцию (тренд) этой зависимости и изображается прямой на корреляционном поле, наименее уклоняющейся от его точек. Прямая линия регрессии $\hat{y}(x) = a + вx$ проходит через точку $(\bar{x}, \bar{y})$, отсекает от оси x отрезок a, и имеет угол наклона с тангенсом равным $в$, как это изображено на рис. 15.1.

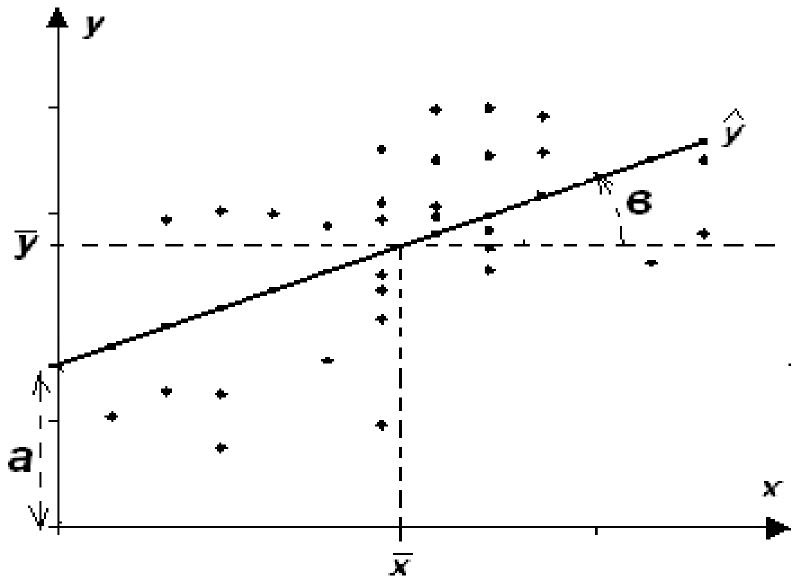


Рис. 15.1 Прямая линейной среднеквадратической регрессии

Выборочный коэффициент корреляции r_{xy} характеризует степень корреляционной зависимости наблюдаемых величин X и $У$ и обладает следующими свойствами:

1) его значения по модулю не превышают единицы ($|r_{xy}| \le 1$),
2) для независимых X и У коэффициент близок к нулю $(r_{xy} \approx 0)$,
3) для линейно зависимых величин он близок к единице $(|r_{xy}| \approx 1)$.

Геометрически он показывает «тесноту» корреляционного поля возле прямой линии регрессии, что иллюстрирует рис. 15.2 для различных значений коэффициента.

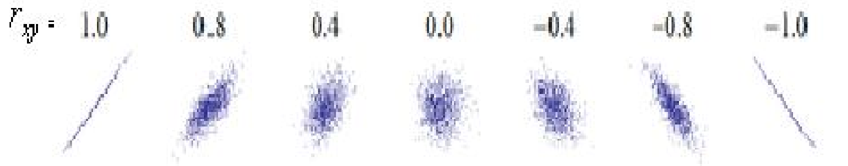


Рис. 15.2 Корреляционное поле для различных уровней корреляции величин

Из рис. 15.2 видно, что некоррелированной выборке $(r_{xy} \approx 0)$ соответствует неориентированное шаровое корреляционное поле, с ростом r_{xy} поле сжимается и ориентируется к прямой линии регрессии. Знак коэффициента говорит о нарастающем или убывающем тренде зависимости.

Ошибки регрессии $e_i = y_i - \hat{y}_i$ имеют нулевое среднее значение $\bar{e} = 0$, так как $\bar{y} = \bar{\hat{y}}$, и минимальную в соответствии с методом наименьших квадратов дисперсию $D_e = \frac{1}{n}\sum_{i=1}^{n} e_i^2 = D_y(1 - r_{xy}^2)$, так называемую остаточную дисперсию, которая тем меньше, чем выше коэффициент корреляции. Величина выборочной дисперсии D_e является статистической оценкой для дисперсии ошибки $D(\varepsilon)$, однако, это смещенная оценка. Несмещенной (исправленной) оценкой является величина $S^2 = \frac{n}{n-2} D_e$, величина $S = (\frac{1}{n-2}\sum_{i=1}^{n} e_i^2)^{1/2}$ называется стандартной ошибкой регрессии. Ошибки для коэффициентов регрессии вычисляются по формулам:

$$S_b^2 = \frac{S^2}{n \cdot D_x}, \; S_a^2 = \frac{\overline{x^2} \cdot S^2}{n \cdot D_x}.$$

В корреляционном анализе также вводится понятие коэффициента детерминации $R^2 = D_{\hat{Y}} / D_Y$, показывающего долю объясненной части дисперсии, объясняемой переменной Y. Поскольку $D_y = D_{\hat{y}} + D_e$, то коэффициент детерминации представим так же в следующем виде:

$$R^2 = 1 - \frac{D_e}{D_y} = r_{xy}^2,$$

показывающем его прямую связь с коэффициентом корреляции.

Известно [9] распределение случайных величин, связанных с введенными выше коэффициентами при условии независимости величин X и Y:

$\frac{r_{xy}\sqrt{n-2}}{\sqrt{1-r_{xy}^2}} = t_{n-2}$ ~ распределение Стьюдента с $\nu = n-2$ степенями свободы,

$\frac{R^2(n-2)}{1-R^2} = F_{1,n-2}$ ~ F-распределение Фишера с $\nu_1 = 1, \nu_2 = n-2$ степенями свободы.

Эти величины используется для построения критериев значимости выборочных коэффициентов r_{xy} и R^2, и их распределение приводится приложениях 3 и 5 соответственно. Действительно, например, задаваясь уровнем значимости α проверяемой гипотезы $H_0 = \{\rho_{XY} = 0\}$, соответствующей независимости величин *X* и *Y*, можно сравнить наблюдаемое значение критерия t_{nab} с критическим значением $t_{kr}(\alpha)$. Если $|t_{nab}| < t_{kr}(\alpha)$, то гипотеза принимается, что говорит о незначимости выборочного коэффициента корреляции, мало отличного от нуля. Если же $|t_{nab}| > t_{kr}(\alpha)$, то гипотеза отвергается, то есть выборочный коэффициент корреляции, а значит и уравнение регрессии, значимы. Значимость коэффициента корреляции говорит о том, что полученный по данной выборке коэффициент неслучайно отличен от нуля, а корреляционная зависимость между наблюдаемыми величинами существенна.

Аналогично строится критерий Фишера для проверки гипотезы $H_0 = \{R^2 = 0\}$ о значимости коэффициента детерминации R^2:

если $F_{nab} < F_{kr}(\alpha)$, то гипотеза H_0 принимается, т.е. R^2 незначим.

Выводы критериев значимости r_{xy} и R^2 идентичны [9].

Значимость коэффициентов регрессии может быть оценена по критериям Стьюдента

$$\frac{a}{S_a} = t_{n-2}, \quad \frac{b}{S_b} = t_{n-2}.$$

3. О множественной регрессии

На практике, объясняемая переменная Y часто зависит не от одной, а нескольких объясняющих переменных $X_к$. Пусть таких переменных будет $m \geq 1$, и они наблюдаются вместе с переменной Y в многофакторной выборке $(y_i x_{1i}, x_{2i}, \ldots, x_{mi})$ объема *n*. Построим выборочную линейную регрессию в форме:

$$\hat{y}(x_1, x_2, \ldots x_m) = b_0 + b_1 x_1 + b_2 x_2 + \ldots.. + b_m x_m + e.$$

Если введем следующие вектора $\vec{x} = (1, x_1, x_2, \ldots x_m)$, $\vec{b} = (b_0, b_1, b_2, \ldots. b_m)^T$, то ее можно записать в векторном виде:

$$\hat{y}(\vec{x}) = \vec{x} \cdot \vec{b}.$$

Введем матрицу измерений X, вектор измерения $\vec{y}$ и переменных $\vec{x}_i = (1, x_{1i}, x_{2i}, \ldots x_{mi})$, а так же вектор регрессии $\hat{\vec{y}}$:

$$X = \begin{pmatrix} 1 & x_{11} & x_{21} & x_{31} & \ldots & x_{m1} \\ 1 & x_{12} & x_{22} & x_{32} & \ldots & x_{m2} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_{1n} & x_{2n} & x_{3n} & \ldots & x_{mn} \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}, \quad \vec{x}_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \ldots \\ x_{in} \end{pmatrix}, \quad \hat{\vec{y}} = \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \ldots \\ \hat{y}_n \end{pmatrix},$$

Тогда вектор регрессии будет $\hat{\vec{y}} = X \cdot \vec{b}$, а ошибки регрессии $\vec{e} = \vec{y} - \hat{\vec{y}}$.

Построим оценки коэффициентов регрессии $\vec{b}$ методом наименьших квадратов, для чего рассмотрим суммарную ошибку регрессии

$$\sum_{i=1}^{n} e_i^2 = \vec{e}^T \cdot \vec{e} = (\vec{y} - \hat{\vec{y}})^T \cdot (\vec{y} - \hat{\vec{y}}) = (\vec{y} - X \cdot \vec{b})^T \cdot (\vec{y} - X \cdot \vec{b}).$$

Подберем такие коэффициенты $\vec{b}$, при которых суммарная ошибка регрессии минимальна, для этого рассмотрим условие минимума:

$$\frac{\partial}{\partial \vec{b}} \sum_{i=1}^{n} e_i^2 = -2 \cdot (X^T \cdot y - X^T \cdot X \cdot \vec{b}) = 0, \quad \Rightarrow \quad \vec{в} = (X^T \cdot X)^{-1} \cdot (X^T \cdot \vec{y}).$$

Таким образом, оценка для коэффициентов регрессии - построена. Матрица, входящая в выражение для коэффициентов имеет вид средних перекрестных произведений:

$$X^T \cdot X = n \cdot \begin{pmatrix} 1 & \overline{x_1} & \overline{x_2} & \overline{x_3} & \ldots & \overline{x_m} \\ \overline{x_1} & \overline{x_1 x_1} & \overline{x_1 x_2} & \overline{x_1 x_3} & \ldots & \overline{x_1 x_m} \\ \overline{x_2} & \overline{x_2 x_1} & \overline{x_2 x_2} & \overline{x_2 x_3} & \ldots & \overline{x_2 x_m} \\ \overline{\ldots} & \overline{\ldots} & \overline{\ldots} & \overline{\ldots} & \ldots & \overline{\ldots} \\ \overline{x_m} & \overline{x_m x_1} & \overline{x_m x_2} & \overline{x_m x_3} & \ldots & \overline{x_m x_m} \end{pmatrix} = W, \quad X^T \cdot \vec{y} = n \cdot \begin{pmatrix} \overline{y} \\ \overline{y x_1} \\ \overline{y x_2} \\ \overline{\ldots} \\ \overline{y x_m} \end{pmatrix}.$$

Значимость построенного уравнения линейной среднеквадратической регрессии $\hat{y}(\vec{x}) = \vec{x} \cdot \vec{b}$ проверяется по значимости коэффициентов регрессии $b_к$ или коэффициента детерминации $R^2 = \frac{D_{\hat{Y}}}{D_Y} = 1 - \frac{D_e}{D_y}$. Для проверки вычисляются: $D_e = \frac{1}{n} \sum_{i=1}^{n} e_i^2$ - дисперсия ошибок регрессии, $S^2 = \frac{1}{n-m-1} \cdot \sum_{i=1}^{n} e_i^2$ - несмещенная стандартная ошибка регрессии, $S_{b_К}^2 = S^2 \cdot W_{кк}^{-1}$ - несмещенные дисперсии коэффициентов регрессии.

Для построения критериев значимости воспользуемся известными статистиками:

$\frac{b_k}{S_{b_k}} = t_{n-m-1}$ ~ распределение Стьюдента с $\nu = n - m - 1$ степенями свободы,

$\frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m} = F_{\nu_1,\nu_2}$ ~ F-распределение Фишера с $\nu_1 = m, \;\; \nu_2 = n - m - 1$ степенями свободы. Задаваясь уровнем значимости α проверяемой гипотезы $H_0 = \{b_k = 0\}$, соответствующей независимости величин X_r и Y, можно сравнить наблюдаемое значение критерия t_{nab} с критическим значением $t_{kr}(\alpha)$. Если $|t_{nab}| < t_{kr}(\alpha)$, то гипотеза принимается, что говорит о незначимости коэффициента $b_к$, мало отличного от нуля, то есть о незначимости переменной X_r в уравнении регрессии, такие переменные желательно исключить из модели регрессии. Аналогично проверяется гипотеза о значимости коэффициента детерминации $H_0 = \{R^2 = 0\}$, соответствующей значимости всего уравнения регрессии в целом. Сравнивая наблюдаемое значение критерия F_{nab}, с критическим значением $F_{kr}(\alpha)$, можно утверждать, что если $F_{nab} < F_{kr}(\alpha)$, то гипотеза принимается, что говорит о незначимости коэффициента R^2, мало отличного от нуля, то есть о не значимости уравнения регрессии в целом.

Помимо значимости построенного уравнения регрессии, его качество оценивается так же отсутствием зависимости между объясняющими переменными X_r (мультиколлинеарности), отсутствием зависимости величины дисперсии ошибок D_e от переменных $X_к$ и Y (гетероскедастичности), отсутствием зависимости ошибок $e_i = y_i - \hat{y}_i$ между собой (например, автокорреляции).

Мультиколлинеарность приводит к неустойчивости обращения матрицы *W*, а ее устранение возможно путем исключения из регрессионной модели малозначимых и сильнозависимых объясняющих переменных (факторов). Для такого исключения построим корреляционную матрицу парных коэффициентов корреляции: $r_{ij} = \frac{\overline{x_i x_j} - \bar{x}_i \cdot \bar{x}_j}{\sigma_{x_i} \cdot \sigma_{x_j}}$. Тогда, если $r_{iк} \approx 1$, а коэффициент $b_к$ незначим или малозначим, то переменную $X_к$ можно исключить из модели регрессии, если коэффициент детерминации при этом значимо не уменьшается.

Гетероскедастичность и автокорреляция могут быть установлены при помощи теста ранговой корреляции Спирмена и теста Дарбина-Уотсона соответственно [1]. Влияние этих нежелательных для качества регрессии факторов может быть ослаблено путем различного рода преобразования переменных регрессионной модели [9].

3. О нелинейной регрессии

Иногда линейная модель регрессии бывает недостаточной, с точки зрения ее качества и значимости, поэтому может быть использованы нелинейные модели. В простейшей форме нелинейность может быть учтена путем введения инструментальных переменных $z_r = \varphi_r(x_k)$, которые входят в модель регрессии обычным линейным образом. При этом часто используются степенная функция $z_r = x^\lambda$, логарифмическая $z_r = \ln x$, показательная $z_r = e^x$ и иногда тригонометрическая $z_r = Sin(\omega x + \varphi_0)$ для выявления циклических факторов в зависимостях. Например, нелинейная модель 2-го порядка может быть построена следующим образом (рис. 15.3):

$$\hat{y}(x) = b_0 + b_1 x + b_2 x^2 + e \quad \Rightarrow \quad \hat{y}(x, z_2) = b_0 + b_1 x + b_2 z_2 + e,$$

где $z_2 = x^2$ - инструментальная переменная. Введение новых членов в модель регрессии, в том числе и инструментальных, оправдано тогда, когда значимо повышается коэффициент детерминации.

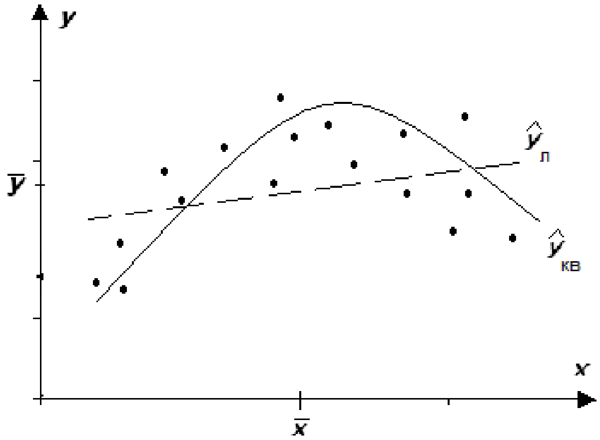
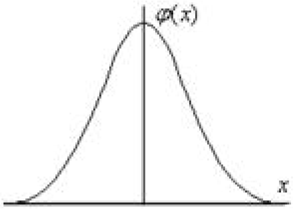


Рис.15.3 Кривая нелинейной среднеквадратической регрессии 2-го порядка.

Иногда строится мультипликативная модель регрессии $\hat{y}(x_1, x_2, ...x_m) = b_0 \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot .... \cdot x_m^{b_m} \cdot e$, которая путем логарифмирования может быть сведена к обычной аддитивной линейной модели для инструментальных переменных $z = \ln y$, $z_к = \ln x_k$.



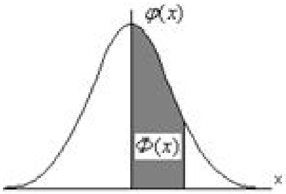
Приложение 1

Весовая функция нормального стандартного распределения (функция Гаусса)

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

x	+0,00	+0,01	+0,02	+0,03	+0,04	+0,05	+0,06	+0,07	+0,08	+0,09
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	0,3970	0,3965	0,3961	0,3956	0,3951	0,3945	0,3939	0,3932	0,3925	0,3918
0,2	0,3910	0,3902	0,3894	0,3885	0,3876	0,3867	0,3857	0,3847	0,3836	0,3825
0,3	0,3814	0,3802	0,3790	0,3778	0,3765	0,3752	0,3739	0,3725	0,3712	0,3697
0,4	0,3683	0,3668	0,3653	0,3637	0,3621	0,3605	0,3589	0,3572	0,3555	0,3538
0,5	0,3521	0,3503	0,3485	0,3467	0,3448	0,3429	0,3410	0,3391	0,3372	0,3352
0,6	0,3332	0,3312	0,3292	0,3271	0,3251	0,3230	0,3209	0,3187	0,3166	0,3144
0,7	0,3123	0,3101	0,3079	0,3056	0,3034	0,3011	0,2989	0,2966	0,2943	0,2920
0,8	0,2897	0,2874	0,2850	0,2827	0,2803	0,2780	0,2756	0,2732	0,2709	0,2685
0,9	0,2661	0,2637	0,2613	0,2589	0,2565	0,2541	0,2516	0,2492	0,2468	0,2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	0,2179	0,2155	0,2131	0,2107	0,2083	0,2059	0,2036	0,2012	0,1989	0,1965
1,2	0,1942	0,1919	0,1895	0,1872	0,1849	0,1826	0,1804	0,1781	0,1758	0,1736
1,3	0,1714	0,1691	0,1669	0,1647	0,1626	0,1604	0,1582	0,1561	0,1539	0,1518
1,4	0,1497	0,1476	0,1456	0,1435	0,1415	0,1394	0,1374	0,1354	0,1334	0,1315
1,5	0,1295	0,1276	0,1257	0,1238	0,1219	0,1200	0,1182	0,1163	0,1145	0,1127
1,6	0,1109	0,1092	0,1074	0,1057	0,1040	0,1023	0,1006	0,0989	0,0973	0,0957
1,7	0,0940	0,0925	0,0909	0,0893	0,0878	0,0863	0,0848	0,0833	0,0818	0,0804
1,8	0,0790	0,0775	0,0761	0,0748	0,0734	0,0721	0,0707	0,0694	0,0681	0,0669
1,9	0,0656	0,0644	0,0632	0,0620	0,0608	0,0596	0,0584	0,0573	0,0562	0,0551
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0,0440	0,0431	0,0422	0,0413	0,0404	0,0396	0,0387	0,0379	0,0371	0,0363
2,2	0,0355	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	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

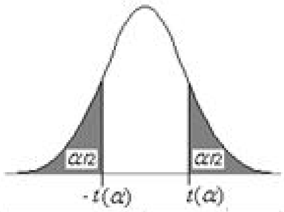
Приложение 2



Интегральная функция нормального стандартного распределения(функция Лапласа)

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$$

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	0,07926	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	0,20540	0,20884	0,21226	0,21566	0,21904	0,22240
0,6	0,22575	0,22907	0,23237	0,23565	0,23891	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	0,34849	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	0,39617	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	0,42785	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	0,47381	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	0,48214	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	0,49134	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	0,49781	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	0,49869	0,49874	0,49878	0,49882	0,49886	0,49889	0,49893	0,49896	0,49900
3,1	0,49903	0,49906	0,49910	0,49913	0,49916	0,49918	0,49921	0,49924	0,49926	0,49929
3,2	0,49931	0,49934	0,49936	0,49938	0,49940	0,49942	0,49944	0,49946	0,49948	0,49950
3,3	0,49952	0,49953	0,49955	0,49957	0,49958	0,49960	0,49961	0,49962	0,49964	0,49965
3,4	0,49966	0,49968	0,49969	0,49970	0,49971	0,49972	0,49973	0,49974	0,49975	0,49976
3,5	0,49977	0,49978	0,49978	0,49979	0,49980	0,49981	0,49981	0,49982	0,49983	0,49983
3,6	0,49984	0,49985	0,49985	0,49986	0,49986	0,49987	0,49987	0,49988	0,49988	0,49989
3,7	0,49989	0,49990	0,49990	0,49990	0,49991	0,49991	0,49992	0,49992	0,49992	0,49992
3,8	0,49993	0,49993	0,49993	0,49994	0,49994	0,49994	0,49994	0,49995	0,49995	0,49995
3,9	0,49995	0,49995	0,49996	0,49996	0,49996	0,49996	0,49996	0,49996	0,49997	0,49997
4,0	0,49997	0,49997	0,49997	0,49997	0,49997	0,49997	0,49998	0,49998	0,49998	0,49998
4,1	0,49998	0,49998	0,49998	0,49998	0,49998	0,49998	0,49998	0,49998	0,49999	0,49999
4,2	0,49999	0,49999	0,49999	0,49999	0,49999	0,49999	0,49999	0,49999	0,49999	0,49999
4,3	0,49999	0,49999	0,49999	0,49999	0,49999	0,49999	0,49999	0,49999	0,49999	0,49999
4,4	0,49999	0,49999	0,50000	0,50000	0,50000	0,50000	0,50000	0,50000	0,50000	0,50000
4,5	0,50000	0,50000	0,50000	0,50000	0,50000	0,50000	0,50000	0,50000	0,50000	0,50000

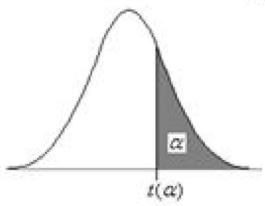


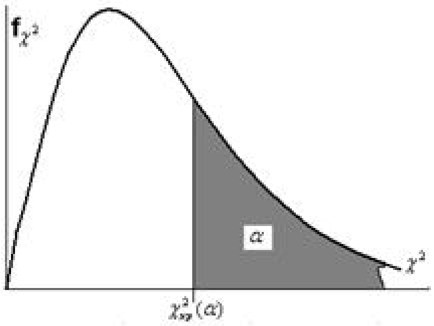
Приложение 3

Обратное t-распределение Стьюдента

$$t_{кр} = t_{кр}(\alpha, m)$$

Число степеней свободы m	Уровни значимости α для двухсторонней критической области					
	0,1	0,05	0,02	0,01	0,002	0,001
1	6,3137515	12,706205	31,820516	63,656741	318,30884	636,61925
2	2,9199856	4,3026527	6,9645567	9,9248432	22,327125	31,599055
3	2,3533634	3,1824463	4,5407029	5,8409093	10,214532	12,923979
4	2,1318468	2,7764451	3,7469474	4,6040949	7,1731822	8,6103016
5	2,0150484	2,5705818	3,36493	4,032143	5,8934295	6,8688266
6	1,9431803	2,4469118	3,1426684	3,707428	5,2076262	5,9588162
7	1,8945786	2,3646243	2,9979516	3,4994833	4,7852896	5,4078825
8	1,859548	2,3060041	2,8964594	3,3553873	4,5007909	5,0413054
9	1,8331129	2,2621572	2,8214379	3,2498355	4,2968057	4,7809126
10	1,8124611	2,2281388	2,7637695	3,1692727	4,1437005	4,5868939
11	1,7958848	2,2009852	2,7180792	3,1058065	4,024701	4,4369793
12	1,7822875	2,1788128	2,680998	3,0545396	3,9296333	4,3177913
13	1,7709334	2,1603687	2,6503088	3,0122758	3,8519824	4,2208317
14	1,7613101	2,1447867	2,6244941	2,9768427	3,7873902	4,1404541
15	1,7530503	2,1314495	2,6024803	2,9467129	3,7328344	4,0727652
16	1,7458837	2,1199053	2,5834872	2,9207816	3,6861548	4,0149963
17	1,7396067	2,1098156	2,566934	2,8982305	3,6457674	3,9651263
18	1,7340636	2,100922	2,5523796	2,8784405	3,6104849	3,9216458
19	1,7291328	2,093024	2,5394832	2,8609346	3,5794001	3,8834059
20	1,7247182	2,0859634	2,527977	2,8453397	3,5518083	3,8495163
21	1,7207429	2,0796138	2,517648	2,8313596	3,5271537	3,8192772
22	1,7171443	2,0738731	2,5083245	2,8187561	3,504992	3,7921307
23	1,7138715	2,0686576	2,4998667	2,8073357	3,4849644	3,7676268
24	1,7108821	2,0638985	2,4921595	2,7969395	3,4667773	3,7453986
25	1,7081407	2,0595385	2,4851072	2,7874358	3,4501887	3,7251439
26	1,7056179	2,0555294	2,4786298	2,7787145	3,4349972	3,7066117
27	1,7032884	2,0518305	2,4726599	2,7706829	3,4210336	3,6895917
28	1,7011309	2,0484071	2,4671401	2,7632624	3,4081552	3,6739064
29	1,699127	2,0452296	2,4620214	2,7563859	3,3962403	3,659405
30	1,6972609	2,0422724	2,4572615	2,7499957	3,3851849	3,6459586
40	1,683851	2,0210754	2,4232568	2,7044593	3,3068777	3,5509658
50	1,675905	2,0085591	2,4032719	2,6777933	3,2614091	3,4960129
60	1,6706489	2,0002978	2,3901195	2,660283	3,2317091	3,4602005
70	1,6669145	1,9944371	2,3808075	2,6479046	3,2107891	3,4350145
80	1,6641246	1,9900634	2,3738682	2,6386906	3,1952577	3,4163375
90	1,6619611	1,9866745	2,3684974	2,6315652	3,1832708	3,4019353
100	1,6602343	1,9839715	2,3642174	2,6258905	3,1737395	3,3904913
	0,05	0,025	0,01	0,005	0,001	0,0005
	Уровни значимости α для односторонней критической области					





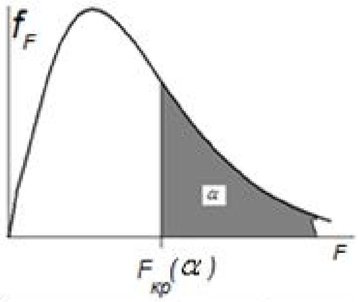
Приложение 4

Обратное "Хи-квадрат" распределение Пирсона

$$\chi^2_{кр} = \chi^2_{кр}(\alpha, m)$$

Число	Уровни значимости α для правосторонней критической области							
степеней	0,001	0,01	0,025	0,05	0,95	0,975	0,99	0,999
1	10,828	6,635	5,024	3,841	0,004	0,001	0,000	0,000
2	13,816	9,210	7,378	5,991	0,103	0,051	0,020	0,002
3	16,266	11,345	9,348	7,815	0,352	0,216	0,115	0,024
4	18,467	13,277	11,143	9,488	0,711	0,484	0,297	0,091
5	20,515	15,086	12,833	11,070	1,145	0,831	0,554	0,210
6	22,458	16,812	14,449	12,592	1,635	1,237	0,872	0,381
7	24,322	18,475	16,013	14,067	2,167	1,690	1,239	0,598
8	26,124	20,090	17,535	15,507	2,733	2,180	1,646	0,857
9	27,877	21,666	19,023	16,919	3,325	2,700	2,088	1,152
10	29,588	23,209	20,483	18,307	3,940	3,247	2,558	1,479
11	31,264	24,725	21,920	19,675	4,575	3,816	3,053	1,834
12	32,909	26,217	23,337	21,026	5,226	4,404	3,571	2,214
13	34,528	27,688	24,736	22,362	5,892	5,009	4,107	2,617
14	36,123	29,141	26,119	23,685	6,571	5,629	4,660	3,041
15	37,697	30,578	27,488	24,996	7,261	6,262	5,229	3,483
16	39,252	32,000	28,845	26,296	7,962	6,908	5,812	3,942
17	40,790	33,409	30,191	27,587	8,672	7,564	6,408	4,416
18	42,312	34,805	31,526	28,869	9,390	8,231	7,015	4,905
19	43,820	36,191	32,852	30,144	10,117	8,907	7,633	5,407
20	45,315	37,566	34,170	31,410	10,851	9,591	8,260	5,921
21	46,797	38,932	35,479	32,671	11,591	10,283	8,897	6,447
22	48,268	40,289	36,781	33,924	12,338	10,982	9,542	6,983
23	49,728	41,638	38,076	35,172	13,091	11,689	10,196	7,529
24	51,179	42,980	39,364	36,415	13,848	12,401	10,856	8,085
25	52,620	44,314	40,646	37,652	14,611	13,120	11,524	8,649
26	54,052	45,642	41,923	38,885	15,379	13,844	12,198	9,222
27	55,476	46,963	43,195	40,113	16,151	14,573	12,879	9,803
28	56,892	48,278	44,461	41,337	16,928	15,308	13,565	10,391
29	58,301	49,588	45,722	42,557	17,708	16,047	14,256	10,986
30	59,703	50,892	46,979	43,773	18,493	16,791	14,953	11,588
40	73,402	63,691	59,342	55,758	26,509	24,433	22,164	17,916
50	86,661	76,154	71,420	67,505	34,764	32,357	29,707	24,674
60	99,607	88,379	83,298	79,082	43,188	40,482	37,485	31,738
70	112,317	100,425	95,023	90,531	51,739	48,758	45,442	39,036
80	124,839	112,329	106,629	101,879	60,391	57,153	53,540	46,520
90	137,208	124,116	118,136	113,145	69,126	65,647	61,754	54,155
100	149,449	135,807	129,561	124,342	77,929	74,222	70,065	61,918
120	173,617	158,950	152,211	146,567	95,705	91,573	86,923	77,755
150	209,265	193,208	185,800	179,581	122,692	117,985	112,668	102,113
200	267,541	249,445	241,058	233,994	168,279	162,728	156,432	143,843
300	381,425	359,906	349,874	341,395	260,878	253,912	245,972	229,963

Приложение 5



Обратное F-распределение Фишера

$$F_{кр} = F_{кр}(\alpha,\ m_1, m_2)$$

Число степ.	Уровень значимости для правосторонней критической области $\alpha = 0{,}05$									
$m2\ /\ m1$	1	2	3	4	5	6	7	8	9	10
1	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54	241,88
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35

Число степ.	Уровень значимости для правосторонней критической области $\alpha = 0{,}01$									
$m2\ /\ m1$	1	2	3	4	5	6	7	8	9	10
1	4052,2	4999,5	5403,4	5624,6	5763,6	5859,0	5928,4	5981,1	6022,5	6055,8
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37

СОДЕРЖАНИЕ

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

ЛИТЕРАТУРА

1. Гмурман В.Е. Теория вероятностей и математическая статистика, – М: Высшая школа, 2001, -408с.

2. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. – М: Высшая школа, 2001, -400с.

3. Письменный Д.Т. Конспект лекций по теории вероятностей и математической статистике. – М: Айрис Пресс, 2004, -252с.

4. Кремер Н.Ш. Теория вероятностей и математическая статистика. – М: ЮНИТИ, 2004, -573с.

5. Виленкин Н.Я. Популярная комбинаторика. –М: Наука, 1975, -208с.

6. Колмогоров А.Н., Журбенко И.Г., Прохоров А.В. Введение в теорию вероятностей. –М: Наука, 1982, -161с.

7. Гнеденко Б.В. Курс теории вероятностей. – М: Наука, 1975, -451с.

8. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и ее инженерные приложения. – М: Наука, 1988, -368с.

9. Смирнов Н.В., Дунин-Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений. М: Наука, 1969, -512с.

10. Корн Г, Корн Т. Справочник по математике. М: Наука, 1977, -832с.

11. Методика статистической обработки эмпирических данных/ РТМ 44-62. -М: Госстандарт, 1966, -101с.

12. Филатов Л.В. Проверка статистических гипотез. Нижний Новгород, ННГАСУ, 2003, -41с.

Горбиков Сергей Павлович
Филатов Леонид Владимирович

ЛЕКЦИИ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ.

Учебное пособие

Редактор Елизарова С.А.

Подписано в печать _____________ Формат 60х90 $^{1/16}$ Бумага газетная. Печать офсетная. Уч.изд.л _____. Усл.печ.л ______. Тираж ______ экз.
Заказ №______
Нижегородский государственный архитектурно-строительный университет. 603600, Н. Новгород, ул. Ильинская, 65
Полиграфический центр ННГАСУ. 603600, Н. Новгород, ул. Ильинская, 65