

Оглавление

1	Основные понятия	3
1.1	Событие. Вероятность события	3
1.2	Основные теоремы теории вероятностей	9
1.3	Повторение опытов	14
2	Случайные величины	15
3	Различные законы распределения величин	25
3.1	Равномерные распределения	25
3.2	Биномиальные распределения	27
3.3	Распределение Пуассона	28
3.4	Показательное распределение	30
3.5	Нормальное распределение	31
4	Системы случайных величин	35
4.1	Функция распределения системы и плотность вероятности	35
4.2	Числовые характеристики системы из двух случайных величин	39
4.3	Нормальный закон распределения систем случайных величин	42
5	Функции случайных величин	43
5.1	Числовые характеристики функций случайных величин	43
5.2	Законы распределения функций случайной величины	46
6	Предельные теоремы теории вероятностей	49
7	Математическая статистика	52
7.1	Гистограмма. Статистическая функция распределения.	52
7.2	Точечные оценки параметров распределения	54
7.3	Интервальные оценки параметров	57
7.4	Оценки числовых характеристик случайных величин	60
7.5	Проверка гипотез	61
7.6	Метод моментов	63
7.7	Принцип максимального правдоподобия	64
7.8	Метод наименьших квадратов	65
7.9	Линеаризация функций	66
8	Случайные процессы	67
8.1	Основные определения	68

8.2	Стационарный случайный процесс	69
8.3	Преобразование случайной функции линейной системой	72

1 Основные понятия

1.1 Событие. Вероятность события

При рассмотрении опытов в которых могут быть различные исходы, результаты опытов будем называть событиями. Отличаем составные (разложимые) и элементарные (неразложимые) события.

Пример

Выпало 6 очков при броске двух игральных костей – составное разложение на (1, 5) или (2,4) или (3,3) или (4,2) или (5,1).

О том, что мы понимаем под элементарным событием надо предварительно условиться. Это неопределяемое понятие (как точка в геометрии). Они определяют идеализированный опыт. По определению, каждый неразложимый исход идеализированного опыта представляется одним и только одним элементарным событием. Совокупность всех элементарных событий называется пространством элементарных событий (σ). Элементарное событие – точки этого пространства. Событие – множество точек.

Совокупность точек представляет все те исходы, при которых происходит событие A , полностью описывают это событие. Любое множество точек A нашего пространства можно назвать событием; оно происходит или нет в зависимости от того, принадлежит или нет множеству A точка, представляющая исходный опыт.

Пример

Число курящих среди 100 человек. Пространство элементарных событий – множество чисел 0, 1, 2, . . . 100.

Определение

Невозможное событие – это событие, которое в результате данного опыта не может произойти. Обозначим его \emptyset . То есть, запись $A = \emptyset$ означает, что A не содержит элементарных событий.

Пример

Рассмотрим систему, состоящую из 6 атомов Н. Выбираем один атом. A – то есть что это атом He – невозможное событие. Но это понятие относительно (надо заранее договориться о том, какую систему рассматриваем.) Если в системе допускаются ядерные реакции, то оно уже возможно (но это другая система)

Определение

Достоверное событие – то, которое в результате опыта обязательно произойдет. То есть $A = (\sigma)$.

Пример

Достоверное событие – выпадение ≤ 6 очков при броске одной игральной кости.

Определение

Событие, состоящее из всех точек, не содержащих событие A , называется событием противоположным A и обозначается \bar{A} .

$$\bar{\sigma} = \emptyset.$$

A – выпадение орла, \bar{A} – решки.

Определение

Суммой (объединением) $A + B$ ($A \cup B$) событий A и B назовем событие, которое состоит в том, что имеет место или A , или B , или (A и B). То есть это объединение множеств точек A и B .

Определение

Произведением (пересечением) $A \cdot B$ ($A \cap B$) событий A и B называется событие, которое состоит в том, что имеет место и A и B (одновременно) – пересечение множества точек A и B .



Определение

События A и B несовместны если $(A \cup B) = \emptyset$.

(То есть не могут произойти одновременно)

Пример

Бросок кости A – не менее 3 очков, \bar{A} – менее 3 очков (1 или 2), $A + B = \sigma$, $A \cdot B = 3$ или 4 очка.

Пример

Выстрел по мишени. A – попадание, B – промах, $A \cdot B = \emptyset$.

$$A + B = B + A$$

$$(A + B) + C = A + (B + C)$$

$$A \cdot B = B \cdot A$$

Определение

Пространство элементарных событий называется дискретным, если оно состоит из конечного числа точек, или из бесконечного числа точек, которые могут быть занумерованы последовательно (Счетное число точек).

Пример

Предыдущий пример – конечное число точек.

Теперь попробуем ввести вероятность, то есть число, которое характеризует степень объективной возможности события.

Определение

Пусть дано дискретное пространство элементарных событий σ с точками $E_1, E_2, E_3 \dots$. Предполагаем, что с каждой точкой E_i (событием) связано число, называемое вероятностью E_i и обозначаемое $P(E_i)$, такое что:

$$1) P(E_i) \geq 0$$

$$2) P(E_1) + P(E_2) + \dots = 1$$

Вероятность любого события A есть сумма вероятностей элементарных событий из которых оно состоит.

$$1) P(\sigma) = 1$$

$$2) P(\emptyset) = 0$$

$$3) 0 \leq P(A) \leq 1$$

Как определить вероятность события в общей теории не постулируется. Об этом надо специально договариваться. Чаще всего встречается схема случаев.

Пусть пространство элементарных событий состоит из n точек, причем все они равновозможны, то есть по условиям симметрии есть основание считать, что ни одно из них не является объективно более возможным, чем другие. Напомним, кроме того, что элементарные события несовместны. Такие элементарные события обычно называют случаями.

Пример

Орел и решка при броске монеты. Появляется для любой из карт тщательно перетасованной колоды.

Пусть событие состоит из m точек (эти m случаев называются благоприятными событию A). Тогда вероятность $P(A) = \frac{m}{n}$

Пример

Бросок игральной кости. A – выпадение четного числа очков.

$n = 6, m = 3$ (2, 4, 6), следовательно, $P(A) = \frac{3}{6} = \frac{1}{2}$

В других ситуациях, не сводящихся к схеме случаев, вероятность определяется по другому (например плотник, землемер, штурман измеряют расстояния – одно и то же, но делают это по разному). При этом все способы с корнями уходят в опыт.

Пусть производится n опытов, в каждом из которых может появиться событие. Частотой события называется отношение числа опытов, в которых появилось к общему числу опытов. Частоту часто называют статистической вероятностью.

$$0 \leq P^*(A) \leq 1, P^*(A) = \frac{m}{n}$$

Так определенная статистическая вероятность носит случайный характер. Но при росте n она стабильно около некоторого значения. При $n \rightarrow \infty$ с практической достоверностью (то есть, вероятность ошибки сколь угодно мала) можно утверждать, что частота события будет сколько угодно мало отличаться от вероятности его в отдельном опыте. Более подробно это рассмотрим потом.

Факты из комбинаторики

Число размещений с повторениями. $\overline{A}_n^k = n^k$, где n – количество типов элементов – группы по k элементов с учетом порядка (число трёхзначных чисел в десятичной системе счисления равно $10^3 - 10^2$).

Число размещений без повторений $A_n^k = \frac{n!}{(n-k)!}$

Пример

12 человек участвует в соревновании. Сколько вариантов распределения медалей. $\overline{A}_{12}^3 = 12 \cdot 11 \cdot 10 = 1320$

Число сочетаний без повторений $C_n^k = \frac{n!}{(n-k)!(k)!}$

Пример

12 команд. Сколько способов сформировать финальную группу из 3 команд без учета мест?

$$C_{12}^3 = \frac{12 \cdot 11 \cdot 10}{2 \cdot 3} = 220$$

Число перестановок из n различных элементов $P_n = n! = A_n^n$

Число перестановок из $n = n_1 + n_2 + n_3 + \dots + n_k$ элементов, $n_1 - 1$ типа, $n_2 - 2$ типа, \dots $n_k - k$ -го типа,

$$P(n_1, n_2, \dots, n_k) = \frac{n!}{n_1! n_2! \dots n_k!}$$

Пример

8 ладей расставляются на доске. Какова вероятность, что никакие две не бьют друг – друга?

Сколько способов расположить ладей на шахматной доске что бы они не били друг друга? На каждой горизонтальной по одной. Пусть на первой горизонтали она стоит на позиции a_1 , на второй на позиции a_2 и так далее.

(a_1, a_2, \dots, a_8) – перестановка чисел $1, 2, \dots, 8$

То есть благоприятных случаев $P_8 = 8!$

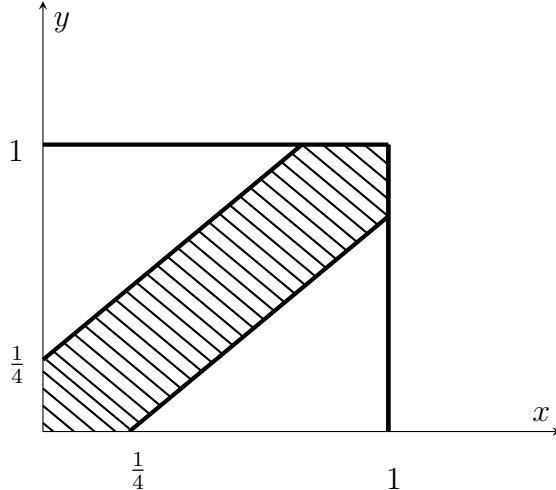
$$P = \frac{8!}{8^2 \cdot (8^2 - 1) \cdot (8^2 - 2) \cdot \dots \cdot (8^2 - 7)} \approx 9 \cdot 10^{-6}$$

Как определить вероятность если пространство элементарных событий не является конечным? Часто здесь имеет смысл метод геометрической вероятности. Если пространство σ может быть изображено геометрической фигуры и по условию опыта вероятность попадания точки (элементарного события) в любую часть области σ пропорционально мере этой части (длине,

площади, объему . . .) и не зависит от ее расположения и формы, то вероятность события A определяется как $P(A) = \frac{S_A}{S}$, где S_A – мера части области, попадание в которую благоприятствует событию, S – мера всей области.

Пример

Двое договорились встретиться в определенном месте между 17 и 18 часами. Пришедший первым ждет второго 15 минут, после чего уходит. Определить вероятность встречи, если время прихода каждого независимо и равномерно в течении этого часа.



Благоприятные исходы : $|x - y| \leq \frac{1}{4}$

$$\frac{1}{4} \leq x - y \leq \frac{1}{4}$$

$$S = 1 - (1 - \frac{1}{4})^2$$

$$P = \frac{1 - \frac{9}{16}}{1} = \frac{7}{16}$$

Парадокс де-Мере-Паскаля

Что вероятнее: при 3 бросках игральной кости получить в сумме: 11 или 12 очков?

Рассуждение де-Мере: Суммы 11 и 12 образуются при выпадении на костях следующих цифр: $12 = 6 + 5 + 1 = 6 + 3 + 3 = 5 + 4 + 3 = 5 + 5 + 2 =$

$4 + 4 + 4$ (то есть 6 вариантов);

$11 = 6 + 4 + 1 = 6 + 3 + 2 = 5 + 5 + 1 = 5 + 4 + 2 = 5 + 3 + 3 =$

$4 + 4 + 2$ (то есть 6 вариантов);

То есть 11 и 12 должны быть равновероятны, но на опыте 11 появляется чаще.

На ошибку указал Паскаль: необходимо учитывать все возможные комбинации цифр, дающие в сумме 11 или 12.

Например $6 + 5 + 1 = 6 + 1 + 5 = 5 + 1 + 6 = 5 + 6 + 1 = 1 + 5 + 6$

$= 1 + 6 + 5$ (то есть 6 способов $= 3!$). Аналогично, $6 + 4 + 2$ ($6 = 3!$ способов), $6 + 3 + 3$ (3 способа), $5 + 4 + 3$ (6 способов), $4 + 4 + 4$ (1 способ).
 $P_{11} = \frac{6+6+6+3+3+3}{6^3} > P_{12}$.

Сравнение статистик Больцмана, Бозе – Энштейна, Ферми – Дирака

Дано k частиц и l ячеек ($l > k$)

Найти вероятность того что:

- 1) в определённых k ячейках окажется по 1 частице
- 2) в каких то ячейках окажется по одной частице

Статистика Больцмана

Ей подчиняется обычный газ.

Условия:

- а) частицы различны
- б) в любой ячейке может находиться сколько угодно частиц

1) Общее число исходов l^k (Так как любую частицу можно положить в любую ячейку). Благоприятных исходов $k!$, так как частицы различны и их можно переставлять.

Значит $P_1 = \frac{k!}{l^k}$

2) Теперь можно k ячеек выбирать из l . (То есть число сочетаний k из общего числа l). То есть число благоприятных исходов равно $C_l^k \cdot k!$

$P_2 = \frac{C_l^k \cdot k!}{l^k}$

Статистика Бозе-Энштейна

Ей подчиняется обычный газ.

Условия:

- а) частицы различны
- б) в любой ячейке может находиться сколько угодно частиц

Общее число исходов.

Переставив ячейки в ряд, границы определим перегородками, которых $l + 1$.

Если поменять местами две частицы, то нового распределения не получится, так как частицы неразличимы. Если поменять местами две перегородки, то тоже ничего нового не получится, так как все перегородки одинаковы. Если же поменять местами перегородку и частицу, то получится новое распределение. Две крайние перегородки закреплены, поэтому в перестановке участвует $l - 1$ перегородок и к частиц, то есть $k + 1 - l$ элементов.

Число перестановок равно : $\frac{(k+l-1)!}{k! \cdot (l-1)!}$

Число благоприятных исходов равно:

1) Так как перестановка частиц не дает нового распределения, то благоприятный исход один (при фиксированных ячейках).

То есть $P_1 = \frac{1}{\frac{(k+l-1)!}{k! \cdot (l-1)!}} = \frac{k! \cdot (l-1)!}{(k+l-1)!}$

2) Число благоприятных исходов равно числу способов выбрать k ячеек из l , где будут частицы: $C_l^k = \frac{l!}{k! \cdot (l-k)!}$

$P_2 = \frac{l! \cdot (k+l-1)!}{k! \cdot (l-k)! \cdot k! \cdot (l-1)!}$

Статистика Ферми–Дирака

Ей подчинен, например, электронный газ.

а) частицы неразличимы

б) в ячейке может находиться не более одной частицы (принцип Паули).

Общее число исходов – это число способов выбрать из l ячеек k , где будут частицы, то есть $C_l^k = \frac{l!}{k! \cdot (l-k)!}$

Число благоприятных исходов

1) Так как k ячеек определены, а частицы неразличимы, то благоприятный исход один.

$P_1 = \frac{1}{C_l^k} = \frac{k! \cdot (l-k)!}{l!}$

2) Число благоприятных исходов равно числу способов выбрать k заполненных ячеек из l равно C_l^k , следовательно $P_2 = \frac{C_l^k}{C_l^k} = 1$

1.2 Основные теоремы теории вероятностей

Теорема сложения вероятностей

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

Доказательство (для дискретного пространства элементарных событий): Что – бы получить $P(A + B)$ надо сложить вероятности точек входящих в A , и точек,

входящих в B , на каждую по одному разу. Точки из AB сосчитали дважды, поэтому их надо вычесть.

Если A и B несовместны то $P(A + B) = P(A) + P(B)$.

Если A_1, A_2, \dots, A_n – попарно несовместны, то $P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$

■

Определение:

Говорят, что события A_1, A_2, \dots, A_n образуют полную группу, если $A_1 + A_2 + \dots + A_n = \sigma$.

Следствие 1

Если события A_1, A_2, \dots, A_n образуют полную группу попарно несовместных событий, то $\sum_{i=1}^n P(A_i) = 1$

Следствие 2

Сумма вероятностей противоположных событий равна 1: $P(A) + P(\bar{A}) = 1$

Если событий три:

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

$$P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{i,j} P(A_i A_j) + \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n)$$

Пример

Есть 100 карточек с числами 1, 2, 3, ..., 100. Случайно выбирается одно из них. Событие A – число делится на 2, B – делится на 3. Найти вероятность $P(A + B)$.

$$\begin{aligned} P(A) &= \frac{50}{100}, P(B) = \frac{33}{100}. \\ AB - \text{делится на 6: } P(AB) &= \frac{16}{100}. \\ P(A + B) &= \frac{50}{100} + \frac{33}{100} - \frac{16}{100} = \frac{67}{100}. \end{aligned}$$

Теорема умножения вероятностей

Определение

Событие A называется независимым от события B , если вероятность события A не зависит от того, произошло событие B или нет, и зависимым, если вероятности A меняются в зависимости от того, произошло событие B или нет.

Пример 1

Бросание 2 монет. A – орел на первой монете, B – орел на второй монете. Эти события независимы.

Пример 2

Охотник, имеющий один патрон попадает в цель (событие A). Событие B – лев ловит добычу – зависимые, если добыча – охотник.

Определение

Вероятность события A , вычисленная при условии, что имело место событие B – называется условной вероятностью события A , и обозначается $P(A|B)$, или $P_B(A)$ (считается $P(B) \neq 0$). Условие независимости события A от B это $P(A|B) = P(A)$.

Теорема

$$P(A \cdot B) = P(A) \cdot P(B|A).$$

Доказательство:



Для схемы случаев. Пусть в результате опыта возможно n исходов, событие A благоприятных m исходов, событие B – k исходов, и событие A , и событие B – l исходов. $P(A \cdot B) = \frac{l}{n}$, $P(A) = \frac{m}{n}$. Если известно, что A происходит, то из n остальных возможных исходов, из них l блокируют событие B , то есть $P(A|B) = \frac{l}{m}$. То есть $P(A \cdot B) = P(A) \cdot P(B|A)$.

Ясно, что $P(A \cdot B) = P(B) \cdot P(A|B)$

■

Следствие 1

Если событие A не зависит от события B , то и событие B не зависит от события A .

Известно, что $P(A) = P(A|B)$. Считаем, что $P(A) \neq 0$.

$$\begin{cases} P(AB)=P(A)P(B|A) \\ P(AB)=P(B)P(A|B) \end{cases}$$

Следовательно, $P(A)P(B|A) = P(B)P(A)$, следовательно
(при условии $P(A) \neq 0$) $P(B|A) = P(B)$

То есть, свойство зависимости или независимости взаимно.

Следствие 2

Для независимых событий $P(A \cdot B) = P(A) \cdot P(B)$

Замечание

Эта формула может быть взята за определение независимых событий.

Теорема

Умножение для n событий.

$$P(A_1, A_2, \dots, A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 A_2) \dots P(A_n|A_1, A_2, \dots, A_{n-1})$$

Определение

События A_1, A_2, \dots, A_n называются независимыми в совокупности, если :

$$P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_n)$$

Замечание

Это определение эквивалентно следующему: события независимы в совокупности, если любое из них не зависит от совокупности остальных.

Замечание

Независимость в совокупности не эквивалентна попарной независимости событий

Пример 1 (Пример Бернштейна):

Имеется 4 шара. Красного, жёлтого, зелёного и трехцветный (содержащий все три цвета). (Модификация - 4 студента, имеющие задолженности по разным предметам). Событие К – вынули шар, на котором есть красный цвет, событие Ж – есть жёлтый, событие З – зелёный. $P(K) = \frac{1}{2} = P(Ж) = P(З)$

Вероятность того, что вынули шар одновременно с двумя цветами равна $P(K \cdot Ж) = \frac{1}{4} = P(K) \cdot P(Ж)$, $P(K \cdot З) = P(K) \cdot P(З) = \frac{1}{4}$, $P(Ж \cdot З) = P(Ж) \cdot P(З) = \frac{1}{4}$.

Вероятность того, что выпадет три цвета равна $P(K \cdot Ж \cdot З) = \frac{1}{4} \neq P(K) \cdot P(Ж) \cdot P(З) = \frac{1}{8}$ То есть они попарно независимы, но зависимы в совокупности.

Пример 2

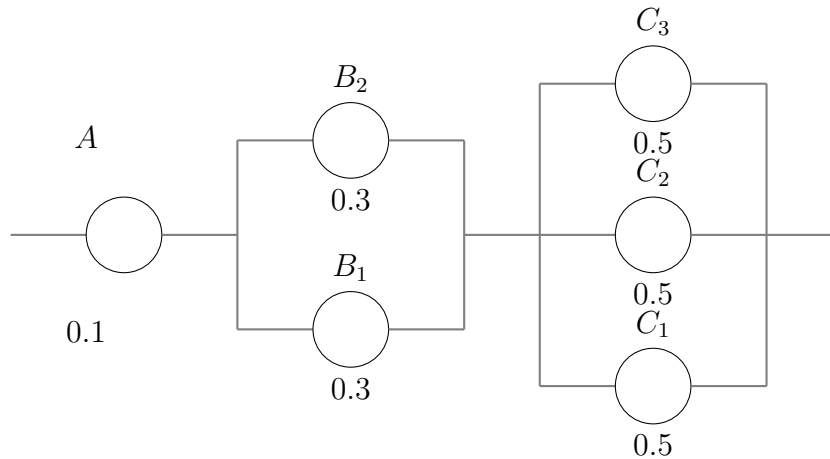
Техническое устройство отказывает с вероятностью $p = 0.5$. Сколько раз его надо продублировать, что бы вероятность отказа установки была $q < 0.1$?

$$P = p^n < 0.1 = q, n > \frac{\ln(q)}{\ln(p)}$$

Суммы и произведение вероятностей часто работают вместе.

Пример 3

Пусть все элементы отказывают независимо от комбинаций других. Чему равна вероятность отказа цепи?



$$P(A + B_1B_2 + C_1C_2C_3) = P(A) + P(B_1) \cdot P(B_2) + P(C_1C_2C_3) - P(AB_1B_2) - P(AC_1C_2C_3) - P(B_1B_2C_1C_2C_3) + P(AB_1B_2C_1C_2C_3)$$

Формула полной вероятности

Определение

Говорят, что события H_1, H_2, \dots, H_n образуют полную группу, если $H_1 + H_2 + \dots + H_n$ – достоверное событие.

Пусть H_1, H_2, \dots, H_n – полная группа попарно несовместимых событий. Эти события будем называть гипотезами. Пусть надо найти вероятность события A , которое может произойти вместе с одной из гипотез.

Тогда $P(A) = \sum_{i=1}^n P(H_i) \cdot P(A|H_i)$. (эта формула носит имя Томаса Байеса)

Так как H_1, \dots, H_n – полная группа. $A = H_1A + H_2A + \dots + H_nA$.

H_1, \dots, H_n – попарно несовместные, следовательно H_1A, \dots, H_nA – несовместны, следовательно $P(A) = P(H_1A) + \dots + P(H_nA) = \sum_{i=1}^n P(H_iA)$, следовательно, по

теореме умножения $P(A) = \sum_{i=1}^n P(H_i) \cdot P(A|H_i)$

Пример

По самолету производится 3 выстрела. Вероятность попадания при первом 0.4, при втором – 0.5, при третьем – 0.7. Для вывода самолета из строя заведомо достаточно 3 попаданий. При первом попадании самолет выходит из строя с вероятностью 0.2, при двух – 0.6. Найти вероятность того, что в результате

трех выстрелов самолет будет выведен из строя.

H_i – в самолет попал i -й снаряд, $P(H_0) = 0.6 \cdot 0.5 \cdot 0.3 = 0.09$.

$P(H_1) = 0.4 \cdot 0.5 \cdot 0.3 + 0.6 \cdot 0.5 \cdot 0.3 + 0.6 \cdot 0.5 \cdot 0.7 = 0.36$

$P(H_2) = 0.6 \cdot 0.5 \cdot 0.7 + 0.4 \cdot 0.5 \cdot 0.7 + 0.4 \cdot 0.5 \cdot 0.5 = 0.41$

$P(H_3) = 0.4 \cdot 0.5 \cdot 0.7 = 0.14$

$P(A) = 0.36 \cdot 0.2 + 0.41 \cdot 0.6 + 0.14 \cdot 1 = 0.458$

Формула Байеса

Пусть событие A может произойти с одним из n попарно несовместимых событий H_1, H_2, \dots, H_n , образующих полную группу. Вероятности гипотез до опыта известны – $P(H_1), \dots, P(H_n)$ (априорные вероятности). Произведен опыт, в результате которого произошло событие A . Как следует изменить вероятности гипотез в связи с появлением A (то есть найти апостериорные вероятности $P(H_i|A)$)? $P(AH_i) = P(A) \cdot P(H_i|A) = P(H_i) \cdot P(A|H_i)$, следовательно $P(H_i|A) = \frac{P(H_i) \cdot P(A|H_i)}{P(A)}$.

Итого: $P(H_i|A) = \frac{P(H_i) \cdot P(A|H_i)}{\sum_{j=1}^n P(H_j) \cdot P(A|H_j)}$

Пример

В первой урне 5 белых и 10 черных шаров. Во второй – 3 белых и 7 черных. Из второй в первую переложили 1 шар, а затем из второй вынули один шар. Оказалось что он белый. Найти вероятность того, что был переложён белый шар

H_1 – переложили белый, H_2 – черный, $P(H_1) = \frac{3}{10}$, $P(H_2) = \frac{7}{10}$. $P(A|H_1) = \frac{6}{11}$, $P(A|H_2) = \frac{5}{11}$

$P(A) = \frac{3}{10} \cdot \frac{6}{11} + \frac{7}{10} \cdot \frac{5}{11}$, $P(H_1|A) = \frac{\frac{3}{10} \cdot \frac{6}{11}}{\frac{3}{10} \cdot \frac{6}{11} + \frac{7}{10} \cdot \frac{5}{11}} = \frac{18}{18+35}$

Пример

Известно, что 5% мужчин и 0.25% женщин – дальтоники. Наугад выбранное лицо страдает дальтонизмом. Какова вероятность того, что это мужчина?

$P(H_1) = P(H_2) = \frac{1}{2}$

$P(A|H_1) = 0.05$, $P(A|H_2) = 0.0025$, $P(H_1|A) = \frac{\frac{1}{2} \cdot 0.05}{\frac{1}{2} \cdot 0.05 + \frac{1}{2} \cdot 0.0025} = \frac{20}{21}$

1.3 Повторение опытов

В самом начале развития теории вероятностей выяснилась фундаментальная роль одной математической схемы, изученной швейцарцем Якобом Бернулли.

Схема такая: проведем последовательность испытаний, в каждом из которых вероятность события A одна и та же (p). Испытания независимы, то есть вероятность выявления события A в каждом из них не зависит от того, появилось оно или нет в других испытаниях.

Пример

2 игрока играют в шахматы 3 партии. Вероятность выигрыша первого $p = \frac{2}{3}$. Найти вероятность того, что он выиграет 2 партии.

Это можно осуществить :

$$P(A_1 A_2 \bar{A}_3 + A_1 \bar{A}_2 A_3 + \bar{A}_1 A_2 A_3) = \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} + \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}.$$

В общем виде: вероятность $P_n(m)$ того, что в n опытах событие произойдёт m раз равна: $\bar{A}_1 A_2 \dots A_m \bar{A}_{m+1} \dots \bar{A}_n + A_1 A_2 \dots \bar{A}_m A_{m+1} \bar{A}_{m+2} \dots \bar{A}_n + \dots + \bar{A}_1 A_2 \dots \bar{A}_{n-m} A_{n-m+1} \dots A_n$. В каждую комбинацию A входит m раз, \bar{A} входит $n - m$ раз. Число комбинаций C_n^m , все комбинации несовместны, следовательно $P_n(m) = p^m q^{n-m} + p^m q^{n-m} + \dots = C_n^m p^m q^{n-m}$, где $q = 1 - p$.

Так как $P_n(m)$ представляет из себя член разложения бинома $(q + p)^n$ распределение вероятностей такого вида называется биномиальным распределением.

Пример

Два шахматных игрока, 10 результативных партий (ничьи не учитываются), Вероятность выигрыша первого – $\frac{2}{3}$, второго – $\frac{1}{3}$. Найти вероятность выигрыша всей игры первым?

$$P_1 = P_{10}(6) + P_{10}(7) + P_{10}(8) + P_{10}(9) + P_{10}(10) = \frac{2^6}{3^{10}} \cdot (210 + 240 + 180 + 80 + 16) = \frac{2^6 \cdot 241}{3^9}$$

$$P_2 = P_{10}(6) + P_{10}(0) + P_{10}(1) + P_{10}(2) + P_{10}(3) + P_{10}(4) = \frac{1507}{3^9} \cdot (210 + 240 + 180 + 80 + 16)$$

То есть вероятность выигрыша первой партии у первого в два раза больше чем у второго, вероятность выигрыша матча у первого в 10 раз больше, чем у второго.

Замечание

$$\sum_{m=0}^n P_n(m) = 1, (p + q)^n = 1, p + q = 1, \text{ следовательно, } (p + q)^n = \sum_{m=0}^n P_n(m) = \sum_{m=0}^n C_n^m p^m q^{(n-m)} - \text{Бином Ньютона.}$$

2 Случайные величины

Определение

Функция, определённая на пространстве элементарных событий называется случайной величиной.

Определение

Случайная величина называется дискретной, если она определена на дискретном пространстве элементарных событий.

Замечание

Лучше бы было называть случайные величины функциями случая

Пример

Дискретная случайная величина. Число тузов у одного игрока при игре в бридж, число совпадающих дней рождения в группе из n человек.

Обозначать случайные величины будем буквами X, Y, \dots и их значения x, y, \dots . Пусть X – случайная величина. x_1, x_2, \dots – ее значения. Совокупность всех элементарных событий, на которых X принимает значение x_i образует событие $X = x_i$. Его вероятность обозначается $P(X = x_j) = p_j$. Соотношение, устанавливающее связь между значениями случайных величин и их вероятностями называется законом распределения случайной величины. Самой простой формой закона распределения для случайной величины является ряд распределения то есть таблица распределения, в которой сведены значения случайной величины и их вероятности.

x_i	x_1	x_2	\dots
p_i	p_1	p_2	\dots

Для наглядности это часто изображают на графике и точки соединяют отрезками прямых. Получившаяся фигура называется многоугольником распределения.



Так как события $X = x_j$ – несовместимы и образуют полную группу, то $\sum_{i=1}^n p_i = 1$.

Пример

Баскетболист бросает мяч в кольцо до первого попадания, либо пока не сделано 3 броска. Вероятность попадания при одном броске равна 0.7.

X – число бросков.

x_i	1	2	3
p_i	0.7	$0.3 \cdot 0.7 = 0.21$	$0.3 \cdot 0.3 = 0.09$

Функция распределения

Вероятность $P(X = x)$ часто использовать неудобно. Используют вероятность $P(X < x)$.

Определение

Функция распределения (интегральная функция распределения или интегральный закон распределения или распределение накопленной вероятности) – это функция на вещественной оси, определяемая следующим образом: $F(x) = P(X < x)$

$F(x)$ полностью характеризует случайную величину с вероятностной точки зрения, то есть это форма закона распределения.

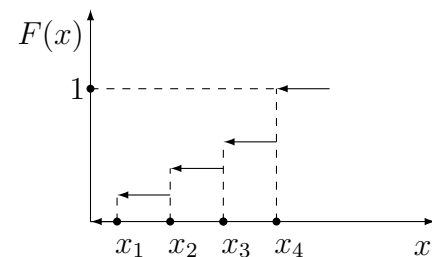
Свойства:

- 1) $F(x)$ – неубывающая функция (то есть, при $x_1 < x_2$ $F(x_1) \leq F(x_2)$)
- 2) $\lim_{x \rightarrow -\infty} F(x) = 0$
- 3) $\lim_{x \rightarrow +\infty} F(x) = 1$
- 4) $F(x)$ непрерывна слева. $F(x_i) = \lim_{x \rightarrow x_i - 0} F(x)$

Определение

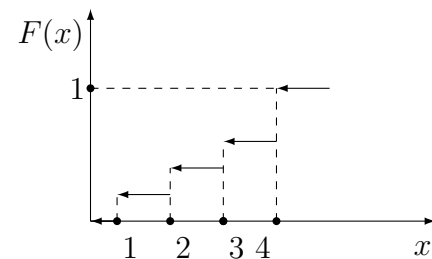
Случайная величина называется непрерывной, если ее функция распределения непрерывна.

Для дискретной случайной величины: $F(x) = \sum_{x_j < x} P(X = x_j)$



Пример

Со стр 11 про баскетболиста



Пример

Случайная величина – площадь разрушений, наносимых цели бомбой. Значение

этой случайной величины непрерывно заполняет промежуток от 0 до πR^2 (R – радиус разрушительного действия). Но в точках 0 и πR^2 у функции распределения скачки, так как этим значениям соответствуют конечные вероятности (вероятности положений 1 и 3 соответственно) круга разрыва.

Вероятность попадания точки на отрезок.

$$P(\alpha \leq x < \beta) = P(x < \beta) - P(x < \alpha) = F(\beta) - F(\alpha)$$

Замечание

$$P(\alpha \leq x \leq \beta) = F(\beta) - F(\alpha) + P(x = \beta)$$

$$P(x = \alpha) = \lim_{\beta \rightarrow \alpha} P(\alpha \leq x < \beta) = \lim_{\beta \rightarrow \alpha} (F(\beta) - F(\alpha))$$

Если слева величина непрерывна, то $P(x = \alpha) = 0$.

Плотность распределения

Пусть непрерывная случайная величина имеет дифференцируемую функцию распределения. Тогда $P(x < X < x + \Delta x) = F(x + \Delta x) - F(x)$

Определение

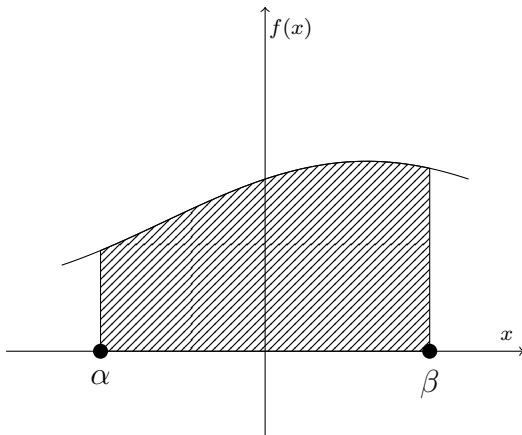
$f(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x < X < x + \Delta x)}{\Delta x} = F'(x)$ – плотность вероятности (плотность распределения).

$F(x)$ – первообразная для $f(x)$. Так как $F(-\infty) = 0$, то константа в первообразной выбирается однозначно.

$$F(x) = \int_{-\infty}^x f(x) dx$$

Найдем $P(\alpha < X < \beta)$ через $f(x)$. Так как вероятность одного значения непрерывной случайной величины равна нулю: $P(\alpha < X < \beta) =$

$$= P(\alpha \leq X < \beta) = F(\beta) - F(\alpha) = \int_{-\infty}^{\beta} f(x) dx - \int_{-\infty}^{\alpha} f(x) dx = \int_{\alpha}^{\beta} f(x) dx.$$



Геометрически $P(\alpha < x < \beta)$ – это площадь.

Свойства плотности вероятности:

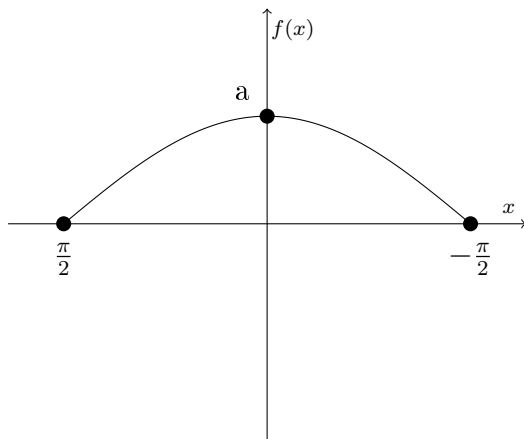
1) $f(x) \geq 0$ (следует из того, что $F(x)$ неубывает)

2) $\int_{-\infty}^{\infty} f(x)dx = 1$ (так как $F(\infty) = 1$)

То есть график плотности вероятности выше оси абсцисс, а полная площадь равна 1.

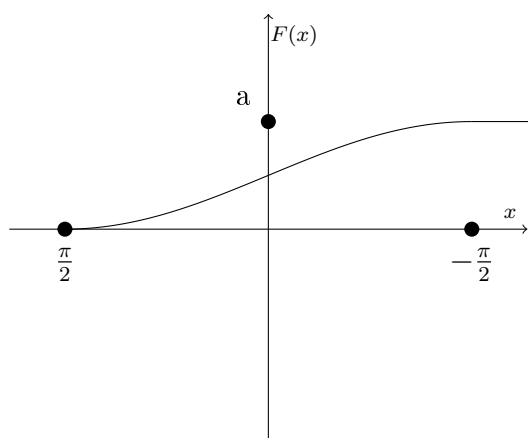
Пример

$$f(x) = \begin{cases} a \cdot \cos(x), & -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ 0, & x < -\frac{\pi}{2} \text{ или } x > \frac{\pi}{2} \end{cases}$$



$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} a \cdot \cos x dx = 2a = 1, \text{ следовательно, } a = \frac{1}{2}$$

$$F(x) = \begin{cases} 0, & x < -\frac{\pi}{2} \\ \frac{1}{2} \cdot (\sin(x) + 1), & -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ 1, & x > \frac{\pi}{2} \end{cases}$$



Числовые характеристики случайных величин.

Функция распределения или плотность полностью описывает случайную величину с вероятностной точки зрения.

Характеристика	Дискретная случайная величина X	Непрерывная случайная величина X
Математическое ожидание	$\sum_k x_k p_k$	$\int_{-\infty}^{\infty} x f(x) dx$
Дисперсия	$\sum_k (x_k - M(x))^2 p_k$	$\int_{-\infty}^{\infty} (x - M(x))^2 f(x) dx$
Начальный момент порядка S	$\sum_k x_k^s p_k$	$\int_{-\infty}^{\infty} x^s f(x) dx$
Центральный момент порядка S	$\sum_k (x_k - M(x))^s p_k$	$\int_{-\infty}^{\infty} (x - M(x))^s f(x) dx$
Мода	наиболее вероятное значение	значение, в котором f(x) максимально
Медиана	не определяем – редко используется (такое значение x_k , что $p_1 + \dots + p_{k-1} \leq \frac{1}{2}$ и $p_1 + \dots + p_k > \frac{1}{2}$)	такое $X = Me$, что $P(X < Me) =$ $= P(X > Me)$
Среднеквадратичное отклонение $\sigma(x)$	$\sigma(x) = \sqrt{D(x)}$	$\sigma(x) = \sqrt{D(x)}$
Коэффициент симметрии S_k	$S_k = \frac{\mu_3}{\sigma^3}$	$S_k = \frac{\mu_3}{\sigma^3}$
Эксцесс E_x	$E_x = \frac{\mu_4}{\sigma^4}$	$E_x = \frac{\mu_4}{\sigma^4} - 3$

Определение

Математическое ожидание дискретной случайной величины (среднее значение) – это число $M(X)$ ($E(X)$, \bar{X} , $< X >$), определяется по формуле $M(X) = \sum_k x_k p_k$, при условии, что ряд абсолютно сходится. Если ряд абсолютно расходится, то говорят что X не имеет конечного математического ожидания. Пусть случайная величина может принимать одно из N равновероятных

значений x_1, x_2, \dots, x_n . Тогда $p_i = \frac{1}{N}$ и $M(X) = \frac{\sum_{i=1}^N x_i}{N}$ – среднее арифметическое. Если значение не является равновероятным, то надо брать взвешенное среднее, что и сделано в определении. Аналогично с точечными массами и центром масс.

Замечание

$M(X)$ может не быть значением случайной величины.

Пример

Испытываются однотипные приборы. Вероятность каждого пройти испытание равно p и независимы. Испытания заканчиваются после выхода из строя первого же прибора. X – число произведенных испытаний. Чему равно $M(X)$?

$$P(X = k) = q \cdot p^{k-1}, \quad q = 1 - p.$$

x_i	1	2	3	...
p_i	q	qp	qp^2	...

$$M(X) = 1 \cdot q + 2q \cdot p + 3q \cdot p^2 + \dots + k \cdot q \cdot p^{k-1} = q(1 + 2p + 3p^2 + \dots)' = q \cdot \left(\frac{p}{1-p}\right)' = \frac{q}{(1-p)^2} = \frac{1}{q}$$

Пример

Пример отсутствия конечного математического ожидания – "Петербургская игра". Бросается монета до тех пор, пока не выпадет орёл. Если это случается при бросании с номером r , то игрок получит 2^r рублей. $x_r = 2^r$, $p_r = 2^{-r}$.

$$\sum_{r=1}^{\infty} x_r p_r = \sum_{r=1}^{\infty} 1 = \infty$$

Определение

Мода – координата максимума $f(x)$. Если кривая распределения имеет более одного максимума, то распределение называется полимодальным.

Определение

Медиана – $P(X < Me) = P(X > Me)$, для непрерывной случайной величины.

Определение

Квантиль порядка p – это значение x_p , соответствующее значению функции распределения равному p ($F(x_p) = p$). Медиана – квантиль порядка $\frac{1}{2}$.

Определение

Центрированной случайной величиной X , соответствующей величине X называется отклонение случайной величины X ее математического ожидания (A). Как оценить разброс X относительно среднего?

Можно рассмотреть $M(A)$, но он оказывается равен 0.

$$M(X - m_x) = \sum_{i=1}^n (x_i - m_x) p_i = \sum_{i=1}^n x_i p_i - m_x \sum_{i=1}^n p_i = m_x - m_x = 0$$

Поэтому используют $M((X - m_x)^2) = D(X) = Var(X)$

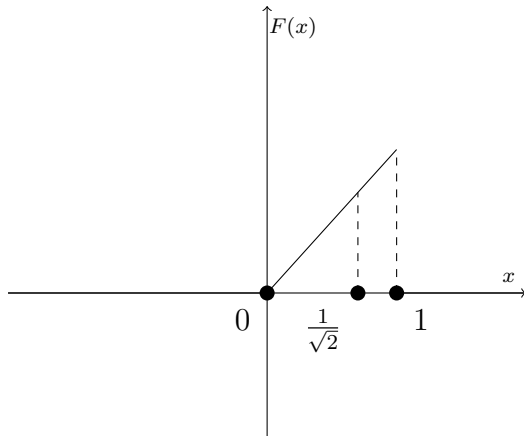
$$D(X) = \sum_{i=1}^n (x_i - m_x)^2 p_i = \sum_{i=1}^n x_i^2 p_i - 2m_x \sum_{i=1}^n x_i p_i + m_x^2 \sum_{i=1}^n p_i = \sum_{i=1}^n x_i^2 p_i - m_x^2$$

$$\text{Для непрерывных величин: } D(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - m_x^2$$

Разброс характеризуется средним квадратичным отклонением: $\sigma(x) = \sqrt{D(x)}$

Пример

$$f(x) = \begin{cases} 2x, & x \in [0, 1] \\ 0, & x \notin [0, 1] \end{cases} \quad (1)$$



$$M(X) = \int_0^1 x \cdot 2x dx = \frac{2}{3}$$

$$D(x) = \int_0^1 x^2 \cdot 2x dx - \frac{4}{9} = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

$$\sigma(x) = \frac{1}{3\sqrt{2}} \text{ Чему равно } x_p?$$

$$F(x) = \int_0^x 2t dt = x^2, \quad x^2 = p, \text{ следовательно, } x_p = \sqrt{p};$$

$$M_e = x_{\frac{1}{2}} = \frac{1}{\sqrt{2}}$$

$$\text{Начальный момент порядка } S: \alpha_S = \sum_k x_k^s p_k = \int_{-\infty}^{\infty} x^s f(x) dx$$

$$\text{Центральный момент порядка } S: \mu_2 = \sum_k (x_k - M(x))^s p_k = \int_{-\infty}^{\infty} (x - M(x))^s f(x) dx$$

$$\text{Центральный и начальный моменты связаны: } \mu_2 = \alpha_2 - m_x^2$$

$$\mu_3 = \sum_i^n (x_i - m_x)^3 p_i = \sum_i x_i^3 p_i - 3m_x \sum_i x_i^2 p_i + 3m_x^2 \sum_i x_i p_i - m_x^3 \sum_i p_i =$$

$$= \alpha_3 - 3m_x \alpha_2 + 2m_x^3$$

Можно определить моменты не относительно m_x , а относительно точки a :

$$M((x - a)^s).$$

Однако центрированные имеют преимущество.

$$\text{В частности: } D(X) = \min_a M((x - a)^2)$$

Доказательство:

$$M(X - a)^2 = \int_{-\infty}^{\infty} (x - m_x + m_x - a)^2 f(x) dx = \int_{-\infty}^{\infty} (x - m_x)^2 f(x) dx + 2(m_x - a) \cdot \int_{-\infty}^{\infty} (x - m_x) f(x) dx + (m_x - a)^2 \int_{-\infty}^{\infty} f(x) dx = D(X) + (m_x - a)^2,$$

следовательно минимум достигается при $m_x = a$

■

Нечетные центральные моменты характеризуют симметрию распределения (кроме первого, который всегда ноль). Поэтому для характеристики симметрии выбирают третий центральный момент. Коэффициент симметрии: $S_k(x) = \frac{\mu_3}{\sigma^3}$.

Если симметрия относительно m_x , то $S_k = 0$.

Четвертый момент служит для характеристики крутости, то есть островершинности или плосковершинности распределения. За стандартное распределение, с которым проводится сравнение принято нормальное распределение: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-m_x)^2}{2\sigma^2}}$.

Его эксцесс = $E_x = \frac{\mu_4}{\sigma^4} - 3$. Для нормального $E_x = 0$.

Пример

$$f(x) = \begin{cases} 2x, x \in [0, 1] \\ 0, x \notin [0, 1] \end{cases}$$

$$\alpha_3 = 2 \int_0^1 x^4 dx = \frac{2}{5}, \mu_3 = \alpha_3 - 3m_x\alpha_2 + 2m_x^3 = \frac{2}{5} - 3 \cdot \frac{2}{3} \cdot \frac{1}{2} + 2 \cdot \left(\frac{2}{3}\right)^3 = \frac{1}{135}$$

$$S_k = \frac{\mu_3}{\sigma^3} = -\frac{2}{5}\sqrt{2}$$

$$\mu_4 = 2 \int_0^1 x^4 dx = 2 \int_0^1 \left(x - \frac{2}{3}\right)^5 dx + \frac{4}{3} \int_0^1 \left(x - \frac{2}{3}\right)^4 dx = \frac{1}{3^7} - \frac{2^6}{3^7} + \frac{4}{15} \cdot \frac{4^0}{3^5} + \frac{2^7}{5 \cdot 3^6} = \frac{2^2}{3^3} - \frac{2^8}{3^3} + \frac{16}{5 \cdot 3^2} + \frac{2^9}{5 \cdot 3^2} > 0.$$

Помимо важных отдельных моментов существенных значений имеет полный набор моментов (μ_n или α_n). Часто проще найти полный набор, чем само распределение. А при весьма общих условиях набор моментов полностью определяет распределение вероятности.

Теорема

Если две плотности вероятности $f_1(x)$ и $f_2(x)$ – непрерывные случайные величины имеют одинаковые моменты α и функция $f_1(x) - f_2(x)$ представляется сходящимся на $(-\infty; \infty)$ рядом по степеням x , то $f_1(x) = f_2(x)$.

Доказательство:

Пусть $f_1(x) - f_2(x) = c_0 + c_1x + c_2x^2 + \dots =$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} (f_1(x) - f_2(x))^2 dx = \int_{-\infty}^{\infty} (f_1(x) - f_2(x))(c_0 + c_1x + c_2x^2 + \dots) dx = \\
&= c_0 \int_{-\infty}^{\infty} (f_1(x) - f_2(x)) dx + c_1 \int_{-\infty}^{\infty} (xf_1(x) - xf_2(x)) dx + \\
&+ c_2 \int_{-\infty}^{\infty} (x^2f_1(x) - x^2f_2(x)) dx + \dots = \\
&= c_0(1 - 1) + c_1(\alpha_1 - \alpha_1) + c_2(\alpha_2 - \alpha_2) + \dots = 0, \text{ следовательно, } f_1(x) = f_2(x).
\end{aligned}$$

■

Замечание

Для "смешанной" случайной величины формулы дискретной и непрерывной объединяются для значений $0 \leq x < 1$ – плотность $f(x) = \frac{1}{2}$.

$$M(X) = 0 \cdot \frac{3}{8} + 1 \cdot \frac{1}{8} + \int_0^1 x \cdot \frac{1}{2} dx = \frac{3}{8}$$

3 Различные законы распределения величин

3.1 Равномерные распределения

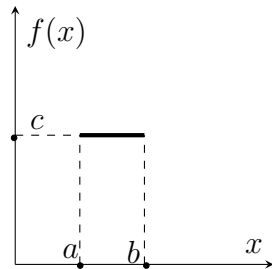
Равномерно распределённая непрерывная случайная величина – случайная величина, у которой плотность вероятности постоянна на некотором интервале, а вне него равна нулю.

Пример

Колесо рулетки. Зафиксируем некоторое направление. Тогда угол под которым она останавливается – случайная величина, равномерно распределённая в интервале $(0, 2\pi)$.

Пример

Производится взвешивание на весах с ценой деления 1 грамм. Если оказывается что вес заключен от k до $k + 1$ грамм, тогда его можно принять за $(k + \frac{1}{2})$ и считать, что при этом допущена ошибка X – случайная величина, равномерно распределённая на интервале $(-\frac{1}{2}, \frac{1}{2})$ грамм.



Пусть распределение X равномерно на (a, b) .

$$f(x) = \begin{cases} c, x \in [a, b] \\ 0, x \notin [a, b] \end{cases}$$

$$\int_a^b c dx = c(b-a) = 1, \text{ значит } c = \frac{1}{b-a}$$

Итого,

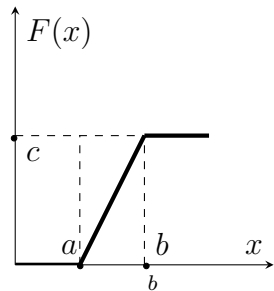
$$f(x) = \begin{cases} \frac{1}{b-a}, x \in [a, b] \\ 0, x \notin [a, b] \end{cases}$$

Значит,

$$f(x) = \begin{cases} c, x \in [a, b] \\ 0, x \notin [a, b] \end{cases}$$

Итого,

$$F(x) = \begin{cases} 0, x \leq a \\ \frac{x-a}{b-a}, a < x < b \\ 1, x \geq b \end{cases}$$



$$M(X) = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$$

В силу симметричности $M_e = \frac{a+b}{2} = M(x)$

Моды это распределение не имеет.

$$D(X) = \frac{1}{b-a} \int_a^b (x - \frac{a+b}{2})^2 dx = \frac{(b-a)^2}{12}$$

$$\sigma(X) = \sqrt{D(X)} = \frac{b-a}{2\sqrt{3}}$$

Распределение симметрично, следовательно асимметрия равна 0 ($S_k = \frac{\mu_3}{\sigma^3} = 0$)

$$\mu_4 = \frac{1}{b-a} \int_a^b (x - \frac{a+b}{2})^4 dx = \frac{(b-a)^4}{80}, \text{ значит } E_x = \frac{\mu_4}{\sigma^4} - 3 = -1.2$$

Вероятность попадания на интервал $(\alpha, \beta) \subset (a, b)$:

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} \frac{1}{(b-a)} dx = \frac{\beta-\alpha}{b-a} - \text{соответствует тому, что используется для вычисления вероятности в методе геометрической вероятности.}$$

3.2 Биномиальные рапределения

Произведение n однотипных опытов, в каждом вероятность события равна p . Случайная величина – число реализаций события.

Случайная величина X принимает целые значения от 0 до n .

$$P(X = m) = C_n^m p^m q^n, \quad q = (1 - p).$$

$$1) \sum_{m=0}^n C_n^m p^m q^{n-m} = 1 - \text{условие нормирования.}$$

$$M(X) = \sum_{m=0}^n \frac{mn!}{(n-m)!m!} p^m q^{n-m} = np \sum_{m=0}^n \frac{(n-1)!}{(n-m)!(m-1)!} p^{m-1} q^{n-m} =$$

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

$$\alpha_2(X) = \sum_{m=0}^n m^2 P(X = m) = \sum_{m=2}^n m(m-1) P(X = m) + \sum_{m=1}^n m P(X = m) =$$

$$= n(n-1)p^2 \sum_{m=2}^n \frac{(n-1)!}{(n-m)!(m-2)!} p^{m-2} q^{n-m} + np = n(n-1)p^2 \sum_{s=0}^{n-2} \frac{(n-2)!}{(n-s-2)!s!} p^s q^{n-s-2} + np =$$

$$= n(n-1)p^2 \sum_{s=0}^k \frac{k!}{(k-s)!s!} p^{m-2} q^{k-s} + np = n(n-1)p^2 + np = n(n-1)p^2 + np.$$

$$D(X) = \alpha_2(x) - m_x^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p) = npq.$$

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

$$\mu_3(X) = npq(q-p), \text{ следовательно асимметрия } S_k = \frac{npq(q-p)}{(npq)^{\frac{3}{2}}} = \frac{q-p}{\sqrt{npq}} = \frac{1-2p}{\sqrt{np(1-p)}}.$$

$$S_k < 0 \text{ если } p > \frac{1}{2}$$

$$S_k = 0 \text{ если } p = \frac{1}{2}$$

$$S_k > 0 \text{ если } p < \frac{1}{2}$$

Если p фиксировано, то $\lim_{n \rightarrow \infty} S_k = 0$ для любого p .

Мода M – целое число, определяемое из двойного неравенства $np - q \leq M \leq np + p$.

Если целое, то две моды: $np + p$ и $np - q$.

3.3 Распределение Пуассона

Пусть нас интересует вероятность того, что за данный промежуток времени произойдёт m событий. При этом выполнены следующие условия:

1) Произойдёт событие или нет в момент времени t не зависит от истории событий, предшествующих моменту t

2) Вероятность отдельного события за малый интервал времени Δt возрастает пропорционально длительности интервала, то есть вероятность отдельного события за интервал $(t, t + \Delta t)$ равна $\lambda \Delta t + O(\Delta t)$, $O(\Delta t)$ – бесконечно малое, более высокого порядка малости, чем Δt . λ – среднее число событий на единицу времени (длины)

3) Вероятность двух или большего числа событий за $(t, t + \Delta t)$ есть $o(\Delta t)$.

Найдем вероятность того, что в интервале $(0, t)$ не произойдёт ни одного события – $P_0(t)$

За промежуток $(0, t + \Delta t)$ не произойдёт ни одного события, если не будет событий в интервалах $(0, t)$ и $(t, t + \Delta t)$, то есть, $P_0(t + \Delta t) = P_0(t)(1 - \alpha \Delta t + o(\Delta t))$ следовательно, $\frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = -\alpha P_0(t) + \frac{O(\Delta t)}{\Delta t}$, следовательно, при $\Delta t \rightarrow 0$ $\frac{dP_0(t)}{dt} = -\alpha P_0(t)$, следовательно $\ln(P_0) = -\alpha t + c$, следовательно $P_0(t) = Ae^{-\alpha t}$ При $t = 0$ $P_0(0) = 1$, следовательно, $A = 1$, то есть $P_0(t) = e^{-\alpha t}$

Вероятность того, что за $(0, t)$ произойдёт событие $P_1(t)$. Тут две возможности: либо произойдёт в $(0, t)$, либо произойдёт в $(t, t + \Delta t)$, следовательно $P_1(t + \Delta t) = P_1(t)(1 - \alpha \Delta t + o(\Delta t)) + P_0(t)(\alpha \Delta t + o(\Delta t))$, следовательно $\frac{P_1(t + \Delta t) - P_1(t)}{\Delta t} = -\lambda P_1(t) + \lambda P_0(t) + \frac{o(\Delta t)}{\Delta t}$, следовательно $\frac{dP_1(t)}{dt} = \lambda P_1(t) + \lambda e^{-\lambda t}$.

Его решение: $P_1(t) = \lambda t e^{-\lambda t}$.

Для m событий на $(0, t)$: $\frac{dP_m(t)}{dt} = \lambda P_m(t) + \lambda P_{m-1}(t)$.

Его решение: $P_m(t) = \frac{(\alpha t)^m}{m!} e^{-\alpha t}$.

Это распределение называется распределение Пуассона.

То есть, случайная величина X , распределенная по Пуассону, принимает целые значения от 0 до ∞ с вероятностью $P_m = P(X = m) = \frac{a^m}{m!} e^{-a}$, a – параметр

распределения. $\sum_{m=0}^{\infty} P_m = e^{-a} \sum_{m=0}^{\infty} \frac{a^m}{m!} = e^{-a} e^a = 1$ (условие нормирования

выполнено). $M(X) = \sum_{m=0}^{\infty} m \frac{a^m}{m!} e^{-a} = \left(\sum_{m=1}^{\infty} \frac{a^{m-1}}{(m-1)!} \right) a e^{-a} = a e^{-a} \sum_{k=0}^{\infty} \frac{a^k}{k!} = a e^{-a} e^a = a$.

$\alpha_2(X) = \sum_{m=0}^{\infty} m^2 \frac{a^m}{m!} e^{-a} = a \sum_{m=1}^{\infty} m \frac{a^{m-1}}{(m-1)!} e^{-a} = a \sum_{m=1}^{\infty} ((m-1) + 1) \frac{a^{m-1}}{(m-1)!} e^{-a} =$

$$= a \left(\underbrace{\sum_{m=1}^{\infty} (m-1) \frac{a^{m-1}}{(m-1)!} e^{-a}}_a + \underbrace{\sum_{m=1}^{\infty} \frac{a^{m-1}}{(m-1)!} e^{-a}}_1 \right) = a(a+1)$$

$$D(X) = \alpha_2 - m_x^2 = a^2 + a - a^2 = a. \text{ То есть } D(X) = M(X).$$

Теорема

$$\lim_{n \rightarrow \infty} C_n^m p^m (1-p)^{n-m} = \frac{a^m}{m!} e^{-a}$$

Доказательство.

$$C_n^m \frac{a^m}{n} (1 - \frac{a}{n})^{n-m} = \frac{(n)(n-1) \dots (n-m+1)}{m!} \frac{a^m}{n^m} \frac{(1 - \frac{a}{n})^n}{(1 - \frac{a}{n})^m} = \frac{(n)(n-1) \dots (n-m+1)}{n^m} \frac{a^m}{m!} \frac{(1 - \frac{a}{n})^n}{(1 - \frac{a}{n})^m}$$

стремится к $\frac{a^m}{m!} e^{-a}$

$(1 - \frac{a}{n})^n = ((1 - \frac{a}{n})^{\frac{n}{a}})^a$ стремится к e^{-a}

■

Бросание точек на прямую

Пусть:

- 1) Точки распределены статистически равномерно со средней плотностью λ (на единицу длины, площади, объема).
- 2) точки попадают в неперекрывающиеся области независимым образом.
- 3) точки появляются по одиночке, а не парами.

Тогда число точек, попавших в область D распределено по Пуассону:

$$P(X = m) = \frac{a^m}{m!} e^{-a}, \text{ где } a = \lambda l (\lambda S_D, \lambda V_D)$$

В этом случае говорят, что точки образуют пуассоновское поле.

Вероятность того, что на отрезок l попадет хотя бы одна точка :

$$P(X \geq 1) = 1 - e^{-a}$$

Пример

Число осколков, попадающих в малоразмерную цель при заданном положении точки разрыва, распределение по Пуассону.

Средняя плотность осколков поля равна 3 осколка на квадратный метр.

Площадь цели 0.5 метров квадратных. Для поражения цели достаточно попасть в нее хотя-бы одним осколком. Найти вероятность поражения.

$$a = \lambda s = 1.5$$

$$P(X \geq 1) = 1 - e^{-1.5} \approx 1 - 0.223 = 0.777$$

Мода

Легко найти при фиксированном m максимум по a :

$$(a^m e^{-a})' = m a^{m-1} e^{-a} - a^m e^{-a} = 0, \text{ следовательно } m = a.$$

3.4 Показательное распределение

Если число точек попавших на интервал длинны t распределено по Пуассону с параметром λt , то расстояние между соседними событиями есть непрерывная случайная величина, распределенная по показательному закону.

$$f(t) = \begin{cases} 0, & t < 0 \\ \lambda e^{-\lambda t}, & t > 0 \end{cases}$$

$$M(X) = \int_0^{\infty} t \lambda e^{-\lambda t} dt = -\lambda t \frac{1}{\lambda} e^{-\lambda t} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda t} dt = \frac{1}{\lambda}$$

$$D(X) = \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt - \frac{1}{\lambda^2} = -t^2 e^{-\lambda t} \Big|_0^{\infty} - \frac{2}{\lambda} t e^{-\lambda t} \Big|_0^{\infty} - \frac{2}{\lambda^2} e^{-\lambda t} \Big|_0^{\infty} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Пример

Длина свободного пробега (Феллер 2, с. 23)

Пример

Устойчивость неудач (Феллер 2, с. 29).

Показательное распределение связано со временем ожидания (точки – по Пуассону), следовательно, расстояния между ними (время ожидания) – показательное распределение. Рассмотрим очередь. Пусть X_0 – моё время ожидания. Пусть мои друзья подвергли себя опыту того же типа (независимо один от другого). Обозначим их результаты как X_1, X_2, \dots . В каждом случае очереди считаем одинаковыми. (То есть X_0, X_1, \dots – взаимно независимые случайные величины с одним и тем же распределением (например, показательным, хотя это несущественно). Пусть последовательность X_0, X_1, \dots – неограниченная. Чтобы оценить размер моей неудачи, я спрашиваю, как много времени должно пройти, прежде чем один из моих друзей испытает большую неудачу (событие $X_k = X_0$ имеет нулевую вероятность, им пренебрегаем)? То есть вводим время ожидания N как значение первого индекса n , такого, что $X_n > X_0$. Событие $\{N > n - 1\}$ происходит, если и только если максимальный член строки X_0, X_1, \dots, X_{n-1} является начальным – по соображениям симметрии вероятность этого события $= \frac{1}{n}$. Событие $P(N = n) = \frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)}$

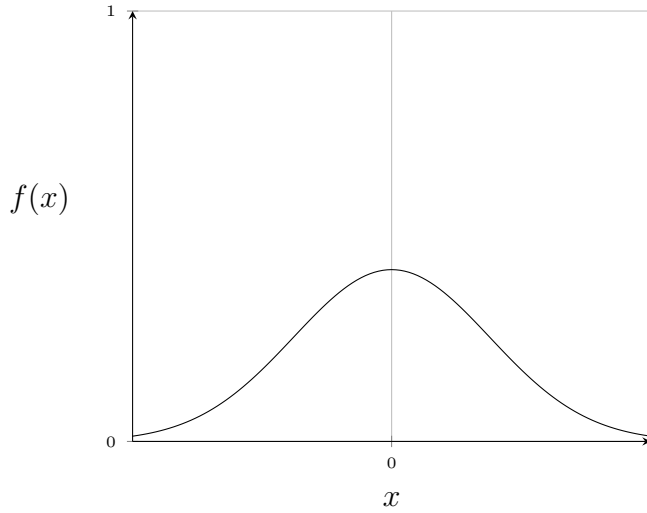
Найдем сколько в среднем надо провести испытаний, что бы побить мой рекорд неудач, то есть найдем математическое ожидание N . $M(N) = \sum_{n=1}^{\infty} n \cdot \frac{1}{n(n+1)} = \infty$.

(Объяснение того факта, что всегда кажется, что именно тебе не везет больше всех).

Замечание

Величина X_k не обязательно показательно распределена. Достаточно, что бы они распределены одинаково и были независимы. Тот факт, что это не зависит от закона распределения X_k , используется в статистике для проверки независимости X_k .

3.5 Нормальное распределение



По нормальному закону распределения случайная ошибка.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

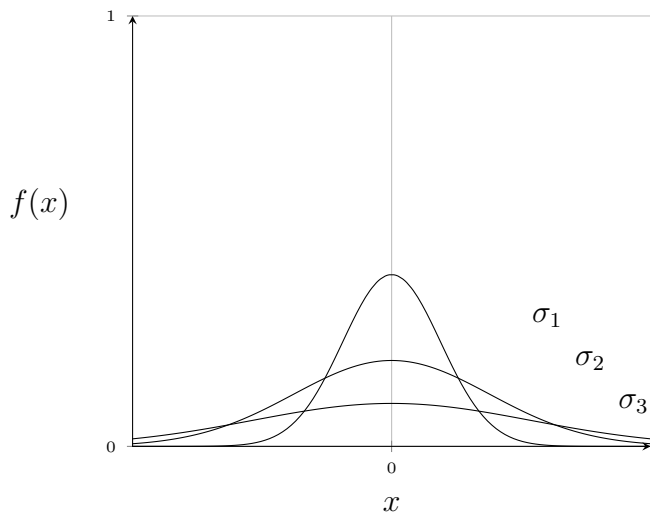
Выясним смысл параметров σ и m .

$$M(X) = \int_{-\infty}^{\infty} x f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} dx. \text{ Пусть } t = \frac{x-m}{\sigma\sqrt{2}}.$$

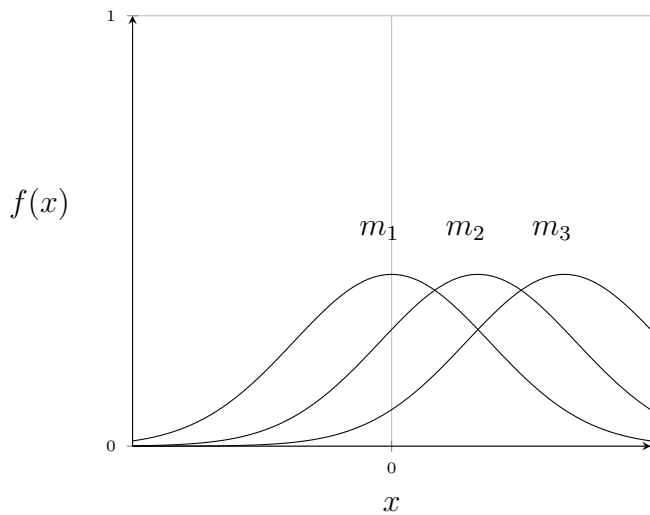
$$\text{Тогда } M(X) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\sigma\sqrt{2}t + m) e^{-t^2} dt = \underbrace{\frac{\sqrt{2}\sigma}{\sqrt{\pi}} \int_{-\infty}^{\infty} t e^{-t^2} dt}_0 + \underbrace{\frac{m}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2} dt}_{\sqrt{\pi}} = m$$

$$D(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-m)^2 e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2} dt = \frac{\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t \cdot 2t e^{-t^2} dt =$$

$$= \frac{\sigma^2}{\sqrt{\pi}} \left(\underbrace{-te^{-t^2}}_0 \Big|_{-\infty}^{\infty} + \underbrace{\int_{-\infty}^{\infty} e^{-t^2} dt}_{\sqrt{\pi}} \right) = \sigma^2$$



$$\sigma_1 < \sigma_2 < \sigma_3$$



$$m_1 < m_2 < m_3$$

$M = Me = m$ – мода и медиана совпадают с $M(X)$.

$$\mu_s(x) = \int_{-\infty}^{\infty} (x-m)^s f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-m)^s e^{-\frac{(x-m)^2}{2\sigma^2}} dx . \text{ Пусть } t = \frac{x-m}{\sigma\sqrt{2}}. \text{ Тогда}$$

$$\mu_s(x) = \frac{(\sigma\sqrt{2})^s}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^s e^{-t^2} dt = \frac{(\sigma\sqrt{2})^s}{\sqrt{\pi}} \left(\underbrace{-\frac{1}{2} e^{-t^2} t^{s-1} \Big|_{-\infty}^{\infty}}_0 + \frac{s-1}{2} \int_{-\infty}^{\infty} t^{s-2} e^{-t^2} dt \right) =$$

$$= \frac{(s-1)(\sigma\sqrt{2})^s}{2\sqrt{\pi}} \int_{-\infty}^{\infty} t^{s-2} e^{-t^2} dt = (s-1)\sigma^2 \mu_{s-2},$$

$$\text{так как } \mu_{s-2} = \frac{(\sigma\sqrt{2})^{s-2}}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^{s-2} e^{-t^2} dt$$

То есть имеет место рекуррентное соотношение: $\mu_s = (s-1)\sigma^2 \mu_{s-2}$
 $\mu_0 = 1$ (по условию нормировки). $\mu_1 = 0$. Пользуясь рекуррентным соотношением, получаем что $\mu_{2k+1} = 0$.

Четные: $\mu_2 = \sigma^2, \mu_4 = 3\sigma^4, \mu_6 = 15\sigma^6 \dots$

Итого: $\mu_{2s} = (2s-1)!!\sigma^{2s}$, то есть все чётные центральные моменты однозначно выражаются через σ . Итак, асимптота $S_k = \frac{\mu_3}{\sigma^3} = 0$

$\mu_4 = 3\sigma^4$, следовательно $Ex = \frac{\mu_4}{\sigma^4} - 3 = 0$

Функция нормального распределения

$$F(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx.$$

$$\text{Пусть } t = \frac{x-m}{\sigma}. \text{ Тогда } F(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-m}{\sigma}} e^{-\frac{t^2}{2}} dt = \Phi^*\left(\frac{x-m}{\sigma}\right)$$

Эта функция затабулирована. Иногда вместо нее рассматривают функцию

$$\phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \text{ и } \phi_1(x) = \frac{1}{2\sqrt{\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$$

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha) = \Phi^*\left(\frac{\beta-m}{\sigma}\right) - \Phi^*\left(\frac{\alpha-m}{\sigma}\right).$$

Свойства: 1) $\Phi(-\infty) = 0$;

2) $\Phi(\infty) = 1$;

3) $\Phi^*(x)$ – неубывает

4) $\Phi^*(-x) = 1 - \Phi^*(x)$; $\Phi(0) = \frac{1}{2}$

$$\Phi^*(x) = \Phi_1(x) + \frac{1}{2}$$

$$\Phi(x) = 2\Phi_1(\sqrt{2}x)$$



Теорема

$$\text{При } \lim_{x \rightarrow -\infty} 1 - \Phi^*(x) \sim \frac{1}{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

точнее для всех $x > 0$ справедливо: $(x^{-1} - x^{-3}) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} < 1 - \Phi^*(x) < x^{-1} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

Доказательство:

Очевидно, $(1 - 3x^{-4}) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} < \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} < (1 + x^{-2}) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Теперь проинтегрируем это неравенство и получим то что требуется в условии, так как член этого неравенства – произведение членов неравенства из условия теоремы.

"Правило трех сигма:" Для нормального распределения величина

$P(|(X - m)| \leq 3\sigma) = 0,9975 \approx 1$, то есть, с практической достоверностью значение нормального распределения случайной величины X попадают в интервал $m - 3\sigma \leq X \leq m + 3\sigma$.

■

Пример

X – распределена по нормальному распределению с $m = 10$, $\sigma = 10$. Найти $P(5 < X < 20)$. Какую точность отклонения ε случайной величины X от m можно гарантировать с вероятностью 0.9.

Решение

$$P(5 < X < 20) = \Phi^*\left(\frac{20-10}{10}\right) - \Phi^*\left(\frac{5-10}{10}\right) = \Phi^*(1) + \Phi^*\left(\frac{1}{2}\right) - 1 = 0.84 - 0.31 = 0.53$$

$$P(|x - 10| < \varepsilon) \geq 0.9 \text{ следовательно } P(|(x - 10)| < \varepsilon) = 2\Phi^*\left(\frac{\varepsilon}{10}\right) - 1 \geq 0.9$$

Из равенства $\Phi^*\left(\frac{\varepsilon}{10}\right) = 0.95$, следовательно $\frac{\varepsilon}{10} \approx 1.645$, так как функция $\Phi^*(x)$ возрастает, в качестве ε можно взять 16.5.

4 Системы случайных величин

Определение

Вектор-функция на пространстве элементарных событий называется случайным вектором.

Замечание

Вектор-функция может иметь значения в пространствах различной размерности.

Координаты случайного вектора в фиксированном базисе – случайные величины. Так что случайный вектор – система случайных величин.

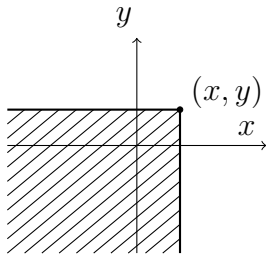
Пример

- 1) Место падения снаряда на плоскости – случайный двумерный вектор.
- 2) Осколок от деления ядра урана характеризуется несколькими параметрами: массой, импульсом, энергией. (это случайный трехмерный вектор).

4.1 Функция распределения системы и плотность вероятности

Для двух случайных величин X, Y .

$F(x, y) = P(X < x, Y < y)$ – это вероятность попадания в угол.



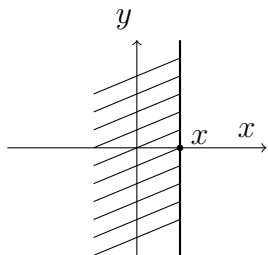
Свойства:

- 1) Формула распределений компонент:

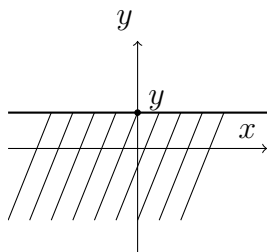
$$F_1(x) = \lim_{y \rightarrow \infty} F(x, y)$$

$$F_2(y) = \lim_{x \rightarrow \infty} F(x, y)$$

$F_1(x) :$



$F_2(x) :$



2) $F(x, y)$ неубывающая функция по каждому аргументу.

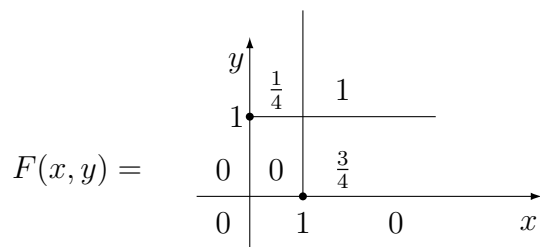
3) $F(x, -\infty) = F(-\infty, y) = F(-\infty, -\infty) = 0$

4) $F(\infty, \infty) = 1$

Пример

По мишени производится один выстрел, вероятность попадания $= \frac{3}{4}$, X – число попаданий, Y – число промахов, $F(x, y) = ?$.

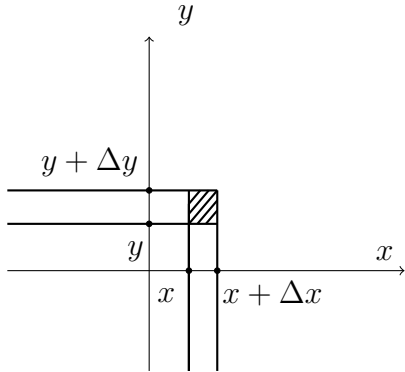
$y \backslash x$	0	1
0	0	$\frac{3}{4}$
1	$\frac{1}{4}$	0



Дискретная система случайных величин X, Y характеризуется таблицей распределения : $p_{ij} = P(X = x_i, Y = y_i)$

x	x_i
y	y_i
	p_{ij}

Совместная плотность вероятности системы случайных величин



Для непрерывной случайной величины : $f(x, y) = \lim_{\substack{\Delta y \rightarrow \infty \\ \Delta x \rightarrow \infty}} \frac{P((X; Y) \in D)}{\Delta x \Delta y}$ – совместная плотность вероятности (вероятностное распределение).

Пусть $F(x, y)$ дважды непрерывно дифференцируема.

$$P((X, Y) \in D) = F(x + \Delta x, y + \Delta y) - F(x + \Delta x, y) - F(x, y + \Delta y) + F(x, y)$$

$$\begin{aligned} f(x, y) &= \lim_{\substack{\Delta y \rightarrow \infty \\ \Delta x \rightarrow \infty}} \frac{F(x + \Delta x, y + \Delta y) - F(x + \Delta x, y) - F(x, y + \Delta y) + F(x, y)}{\Delta x \Delta y} = \\ &= \frac{\delta^2 F(x, y)}{\delta x \delta y} \end{aligned}$$

Из определения следует, что $P((X, Y) \in \Omega) = \iint_{\Omega} f(x, y) dx dy$

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy$$

Свойства:

$$1) f(x, y) \geq 0$$

$$2) \iint_{-\infty}^{\infty} f(x, y) dx dy = 1$$

Пример

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

$$F(x, y) = \frac{1}{\pi^2} \int_{-\infty}^x \int_{-\infty}^y \frac{dx dy}{(1+x^2)(1+y^2)} = \left(\frac{1}{\pi} \operatorname{arctg} x + \frac{1}{2}\right) \left(\frac{1}{\pi} \operatorname{arctg} y + \frac{1}{2}\right)$$

$$P(0 \leq x \leq 1; -1 \leq y \leq 1) = \frac{1}{\pi^2} \int_0^1 dx \int_{-1}^1 dy \frac{dxdy}{(1+x^2)(1+y^2)} = \frac{1}{8}$$

Плотность вероятности компонент
(маргинальные (частные) плотности)

$$F_1(x) = F(x, \infty) = \int_{-\infty}^x \int_{-\infty}^{\infty} f(x, y) dxdy$$

$$f_1(x) = F'_1(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

$$\text{Аналогично: } f_2(y) = F'_2(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Определим условные плотности вероятности: $f_2(y|x) = \frac{f(x,y)}{f_1(x)}$; $f_1(x|y) = \frac{f(x,y)}{f_2(y)}$
Соотношение $f(x, y) = f_1(x)f_2(y|x)$ соответствует теореме умножения вероятностей. Смысл: $f_2(y|x)$ – это плотность распределения случайной величины Y при условии, что $X = x$.

Для дискретной случайной величины: $p_i = \sum_j p_{ij}$

$$P(X = x_i) = \sum_j P(X = x_i, Y = y_j)$$

Определение

Случайные величины X, Y называются независимыми, если $f(x, y) = f_1(x)f_2(y)$ (для непрерывных случайных величин)

$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$ (для дискретных случайных величин)

Это означает, что $f_2(y|x) = f_2(y)$ и $f_1(x|y) = f_1(x)$, то есть закон распределения случайной величины Y не зависит от того, какое значение приняла величина X .

Пример

$f(x, y) = \frac{1}{\pi^2(1+x^2)(1+y^2)}$, $f_1(x) = \frac{1}{\pi(1+x^2)}$, $f_2(y) = \frac{1}{\pi(1+y^2)}$ – независимые случайные величины.

Функции распределения и плотности систем из n случайных величин

$$F(x_1, x_2, \dots, x_n) = P(X_1 < x_1, X_2 < x_2, \dots, X_n < x_n)$$

Свойства те же, что и при $n = 2$.

$$f(x_1, x_2, \dots, x_n) = \frac{\delta^n F(x_1, x_2, \dots, x_n)}{\delta x_1 \delta x_2 \dots \delta x_n}$$

$$F_1(x_1) = F(x_1, \infty, \dots, \infty)$$

$$f_1(x_1) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_2 \dots dx_n$$

Случайные величины независимы, если $f(x_1, x_2, \dots, x_n) = f_1(x_1)f_2(x_2) \dots f_n(x_n)$

4.2 Числовые характеристики системы из двух случайных величин

Основные характеристики системы (X, Y) :
 $M(X), M(Y), D(X), D(Y), K_{XY}$ – они уже определены, но их можно находить, не вычисляя $f_1(x), f_2(y)$.

Дискретная случайная величина	Непрерывная случайная величина
$M(X) = \sum_{i,j} x_i p_{i,j}$	$\iint_{-\infty}^{\infty} x f(x, y) dx dy$
$M(Y) = \sum_{i,j} y_j p_{i,j}$	$\iint_{-\infty}^{\infty} y f(x, y) dx dy$
$D(X) = \sum_{i,j} (x_i - M(X))^2 p_{i,j}$	$\iint_{-\infty}^{\infty} (x - M(X))^2 f(x, y) dx dy$ $= \iint_{-\infty}^{\infty} x^2 f(x, y) dx dy$
$K_{X,Y} = \sum_{i,j} (x_i - M(X))(y_j - M(Y)) p_{i,j}$ $= cov(X, Y)$	$\iint_{-\infty}^{\infty} (x - M(X))(y - M(Y)) f(x, y) dx dy$

Теорема

Корреляционный момент независимых случайных величин равен нулю

Доказательство:

Для непрерывных случайных величин:

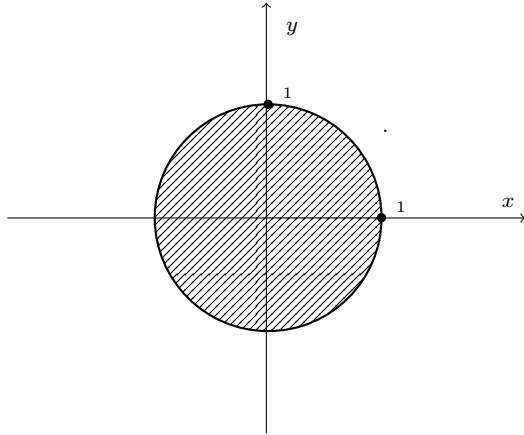
Они независимы, значит, $f(x, y) = f_1(x)f_2(y)$

$$K_{XY} = \int_{-\infty}^{\infty} (x - M(X)) f_1(x) dx \int_{-\infty}^{\infty} (y - M(Y)) f_2(y) dy = 0$$

$$r_{XY} = \frac{K_{XY}}{\sigma_x \sigma_y}, \sigma_x = \sqrt{D(X)}, \sigma_y = \sqrt{D(Y)}.$$

■

Пример



$$f(x, y) = \begin{cases} \frac{1}{\pi}, & x^2 + y^2 < 1 \\ 0, & x^2 + y^2 > 1 \end{cases}$$

Величины являются зависимыми (если $X = 0$, то Y может с равными вероятностями принимать значения на $[-1, 1]$, если $x = 1$, то Y может принять одно единственное значение $Y = 0$).

$M(X) = M(Y) = 0$ по соображениям симметрии.

$$\begin{aligned} K_{XY} &= \iint_C x f(x, y) dx dy = \frac{1}{\pi} \iint_C xy dx dy = \frac{1}{\pi} \iint_{C_1} xy dx dy + \frac{1}{\pi} \iint_{C_2} xy dx dy + \\ &+ \frac{1}{\pi} \iint_{C_3} xy dx dy + \frac{1}{\pi} \iint_{C_4} xy dx dy + \frac{1}{\pi} \iint_{C_5} xy dx dy = 0. \end{aligned}$$

(Так как интегралы одинаковы по величине и отличаются знаком.)

То есть независимые следовательно некоррелирующие, обратное неверно.

$$-1 \leq r_{xy} \leq 1$$

Если $Y = aX + b$, то $r_{XY} = \pm 1$

Если $a > 0$, то $+1$, если $a < 0$ то -1

Докажем это позже.

Пример

а) Вес и рост человека связаны положительной корреляцией

б) Возраст взрослого человека и количество волос у него на голове связаны отрицательной корреляцией

$$\alpha_{k,s}(X, Y) = \sum_{i,j} x_i^k y_j^s p_{ij} = \iint_{-\infty}^{\infty} x^k y^s f(x, y) dx dy - \text{начальный момент корреляции.}$$

$$\mu_{k,s}(X, Y) = \sum_{i,j} (x_i - M(X))^k (y_j - M(Y))^s p_{ij} = \iint_{-\infty}^{\infty} (x - M(X))^k (y - M(Y))^s f(x, y) dx dy$$

Числовые характеристики систем n случайных величин

X_1, X_2, \dots, X_n — их характеристики: n математических ожиданий (m_1, m_2, \dots, m_n) , n дисперсий (D_1, \dots, D_n) , $n(n-1)$ корреляционных моментов: $K_{ij} = M((X_i - m_i)(X_j - m_j))$ ($D_i = K_{ii}$)

Корреляционные матрицы K

$$K = \begin{pmatrix} D_1 & K_{12} & K_{13} & \dots & K_{1n} \\ K_{21} & D_2 & K_{23} & \dots & K_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ K_{n1} & K_{n2} & K_{n3} & \dots & D_n \end{pmatrix}$$

$K_{i,j} = K_{j,i}$, поэтому часто пишут только половину матрицы:

$$K = \begin{pmatrix} D_1 & K_{12} & K_{13} & \dots & K_{1n} \\ & D_2 & K_{23} & \dots & K_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & & D_n \end{pmatrix}$$

Если X_1, X_2, \dots, X_n — не коррелированы, то корреляционная матрица:

$$K = \begin{pmatrix} D_1 & 0 & 0 & \dots & 0 \\ 0 & D_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & D_n \end{pmatrix}$$

Коэффициент корреляции $r_{ij} = \frac{k_{ij}}{\sigma_i \sigma_j}$

Нормированная корреляционная матрица

$$K = \begin{pmatrix} 1 & r_{12} & r_{13} & \dots & r_{1n} \\ r_{21} & 1 & r_{23} & \dots & r_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & r_{n3} & \dots & 1 \end{pmatrix}$$

– характеризует именно коррелированность, безотносительно к их рассеиванию.

4.3 Нормальный закон распределения систем случайных величин

(X, Y) – нормированные случайные величины.

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)} \cdot \left(\frac{(x-m_x)^2}{\sigma_x^2} + \frac{2r(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2} \right)}$$

Можно показать, что m_x, m_y – математические ожидания,

σ_x, σ_y – среднеквадратичные отклонения, r – коэффициент корреляции X и Y .

Теорема

Для нормального распределения случайной величины X и Y независимость эквивалентна некоррелированности.

Доказательство:

Из независимости следует некоррелируемость для любой случайной величины.

Пусть X, Y – нормированны, и $r = 0$, следовательно:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}} e^{-\frac{(y-m_y)^2}{2\sigma_y^2}} = f_1(x)f_2(y), \text{ то есть } X \text{ и } Y \text{ – независимы.}$$

■

Регрессия

Рассмотрим условную плотность вероятности $f_2(y|x)$. Введем условное

$$\text{математическое ожидание: } M(Y|X=x) = \int_{-\infty}^{\infty} y f_2(y|x) dy = \frac{1}{f_1(x)} \int_{-\infty}^{\infty} y f(x, y) dy,$$

в предположении, что интеграл сходится абсолютно.

$M(Y|X)$ – это случайная величина (она является функцией только от X , а не от Y , которая называется регрессией Y на X).

Это соотношение не определено, когда $f_1(x) = 0$. Пусть в этом случае

$M(Y|X = x) = 0$, тогда $M(Y|X = x)$ определено, когда $f_1(x) = 0$. Пусть в этом случае $M(Y|X = x) = 0$, тогда $M(Y|X)$ определено для любой непрерывной плотности распределения. Можно ввести аналогичным образом условную дисперсию.
 $Var(Y|X) = D(Y|X)$.

Для нормального закона распределения:

$$f_2(y|x) = \frac{1}{\sigma_y \sqrt{1-r^2} \sqrt{2\pi}} e^{-\frac{1}{2(1-r^2)} \cdot \left(\frac{(x-m_x)^2}{\sigma_x^2} - r \frac{(y-m_y)^2}{\sigma_y^2} \right)} = \frac{e^{-\frac{1}{2(1-r^2)} \cdot \left(\frac{(x-m_x)^2}{\sigma_x^2} - r \frac{(y-m_y)^2}{\sigma_y^2} \right)}}{\sigma_y \sqrt{1-r^2} \sqrt{2\pi}}$$

То есть это плотность нормального закона с $m_{y|x} = m_y + r \frac{\sigma_y}{\sigma_x} (x - m_x)$ и среднеквадратичное отклонения $\sigma_{y|x} = \sigma_y \sqrt{1-r^2}$

То есть от x в условном законе зависит лишь $m_{y|x}$, но не $\sigma_{y|x}$. Откладывая $m_{y|x}$ на оси ординат при различных x получим линию регрессии Y на X : $y = m_y + r \frac{\sigma_y}{\sigma_x} (x - m_x)$. Для нормального распределения регрессия линейна.

Пример

Гальтона – характеризует эмпирическое значение формул линейной регрессии. Пусть X_1 и X_2 – характеризуют рост (измеренный по отклонению от среднего в сантиметрах) соответственно отцов и сыновей в некоторой популяции. Рост случайно выбранного сына – нормальная случайная величина с $M(X_2) = 0$ и $D(X_2) = \sigma_2^2$, отца – с $M(X_1) = 0, D(X_1) = \sigma_1^2$. Но в подпопуляции сыновей, отцы которых имеют фиксированный рост x , рост сыновей представляется нормальной величиной с (по формулам регрессии) математическим ожиданием $r \frac{\sigma_2}{\sigma_1} x$ и дисперсией $\sigma_2^2(1-r^2) < \sigma_2^2$.

Таким образом, регрессия X_2 на X_1 показывает как много статистической информации относительно X_2 содержится в наблюдении над X_1 .

5 Функции случайных величин

5.1 Числовые характеристики функций случайных величин

$$Y = \varphi(X), M(Y) = M(\varphi(X))$$

x_i	x_1	x_2	...
p_i	p_1	p_2	...

$\varphi(x_i)$	$\varphi(x_1)$	$\varphi(x_2)$...
p_i	p_1	p_2	...

Для непрерывной случайной величины: $M(\varphi(X, Y)) = \int_{-\infty}^{\infty} \varphi(x) f(x) dx$

Для дискретной случайной величины: $M(\varphi(X, Y)) = \sum_{i,j} \varphi(x_i, y_j) p_{i,j} dx$

Пример

На плоскости задан отрезок длины l , вращающаяся случайным образом так, что все направления одинаково вероятны. Найти среднее значение длины проекции отрезка на неподвижную ось.

$Y = l|\cos \alpha|$, α распределенна равномерно от 0 до 2π

$$M(l|\cos \alpha|) = \int_0^{2\pi} l|\cos \alpha| \frac{d\alpha}{2\pi} = \frac{2l}{\pi} \int_0^{2\pi} \cos \alpha d\alpha = \frac{2l}{\pi}$$

$$D(l|\cos \alpha|) = \frac{4}{2\pi} \int_0^{\frac{\pi}{2}} l^2 \cos^2 \alpha d\alpha - \frac{4l^2}{\pi^2} = \frac{2l^2}{\pi} \int_0^{\frac{\pi}{2}} (\frac{1}{2} + \frac{1}{2} \cos 2\alpha) d\alpha - \frac{4l^2}{\pi^2} = \frac{l^2}{2} - \frac{4l^2}{\pi^2}$$

Пример

В процессе слежения радиолокатором за определенным объектом пятно изображающее объект, все время удерживается в пределах экрана (круг радиуса R) и занимает на нем случайное положение с постоянной плотностью вероятности. Найти среднее расстояние до центра экрана.

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

$$M(r) = \iint_K \sqrt{x^2 + y^2} \frac{dx dy}{\pi R^2} = \frac{1}{\pi R^2} \int_0^{2\pi} d\varphi \int_0^R \rho^2 d\rho = \frac{l}{3} R$$

Теоремы о числовых характеристиках

$$1) M(C) = C$$

$$2) D(C) = 0$$

$$3) M(cX) = cM(X)$$

Доказательство:

$$M(cX) = \int_{-\infty}^{\infty} cxf(x)dx = c \int_{-\infty}^{\infty} xf(x)dx = xM(X)$$

$$4) D(cX) = c^2 D(X)$$

Доказательство:

$$\overline{D(cX)} = M((cX - M(cX))) = M((cX - cM(X))^2) = M(c^2(X - M(X))^2) = c^2 M((X - M(X))^2) = c^2 D(X)$$

$$5) M(X + Y) = M(X) + M(Y), M(\sum x_i) = \sum M(x_i)$$

Доказательство:

$$\begin{aligned} M(X + Y) &= \iint_{-\infty}^{\infty} (x + y) f(x, y) dx dy = \iint_{-\infty}^{\infty} x f(x, y) dx dy + \iint_{-\infty}^{\infty} y f(x, y) dx dy = \\ &= \int_{-\infty}^{\infty} dx x \int_{-\infty}^{\infty} f(x, y) dy + \int_{-\infty}^{\infty} dy y \int_{-\infty}^{\infty} f(x, y) dx = \int_{-\infty}^{\infty} dx x f_1(x) + \int_{-\infty}^{\infty} dy y f_1(y) = \\ &= M(X) + M(Y) \end{aligned}$$

$$6) M(\sum a_i X_i + b) = \sum a_i M(x_i) + b$$

$$7) D(X + Y) = D(X) + D(Y) + 2K_{XY}$$

Доказательство:

$Z = X + Y$. Перейдем к центрированным величинам.

$$8) D\left(\sum_{i=1}^n a_i X_i + b\right) = \sum_{i=1}^n a_i^2 D(X_i) + 2 \sum_{i < j}^n a_i a_j K_{i,j}$$

Доказательство:

$$Y_i = a_i X_i,$$

$$\begin{aligned} \sum_{i=1}^n a_i X_i + b &= \sum_{i=1}^n Y_i + b, \text{ следовательно } D\left(\sum_{i=1}^n a_i X_i + b\right) = \sum_{i=1}^n D(Y_i) + \sum_{i < j} K_{i,j}^y = \\ &= \sum_{i=1}^n a_i^2 D(X_i) + 2 \sum_{i < j} K_{i,j}^y \end{aligned}$$

$$9) M(XY) = M(X)M(Y) + K_{XY}$$

Доказательство:

$$K_{XY} = M(XY) - M((X - m_x)(Y - m_y)) = M(XY) - m_x M(Y) - m_y M(X) + m_x m_y = M(XY) - M(X)M(Y)$$

■

Для двух некоррелирующих величин: $M(XY) = M(X)M(Y)$. Если произведение более 2х случайных величин, то некоррелируемости недостаточно (требуется зануление некоторых высших смешных моментов), что заведомо выполняется при независимости случайных величин.

Независимы, следовательно $M(\prod_i x_i) = \prod_i M(x_i)$

10) Для независимых случайных величин X, Y, Z .

$$D(XY) = D(X)D(Y) + m_x D(Y) + m_y^2 D(X)$$

Доказательство:

$XY = Z, D(XY) = M(Z^2) = ((Z - m_Z)^2) = M((XY - m_x m_y)^2)$ Величины независимы, следовательно $m_z = m_x m_y$

$$M((XY - m_x m_y)^2) = M(X^2 Y^2) - 2m_x m_y M(XY) + m_x^2 m_y^2$$

X, Y независимы, следовательно X^2 и Y^2 независимы, следовательно $M(X^2 Y^2) = M(X^2)M(Y^2)$ и $M(XY) = m_x m_y$, следовательно

$D(XY) = M(X^2)M(Y^2) - m_x^2 m_y^2, M(X^2) = D(X) + m_x^2, M(Y^2) = D(Y) + m_y^2$, следовательно утверждение теоремы верно.

■

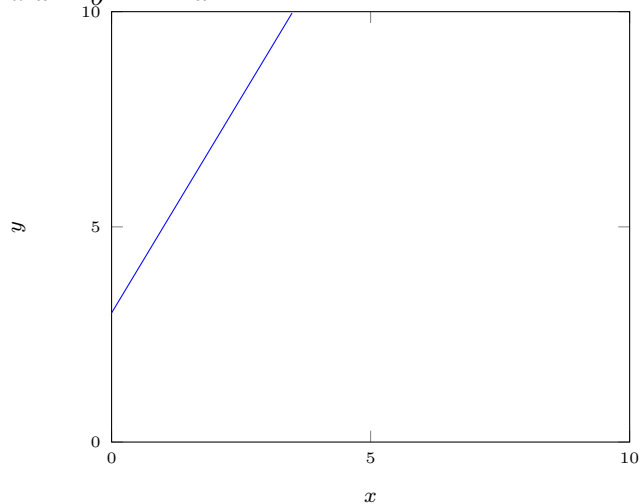
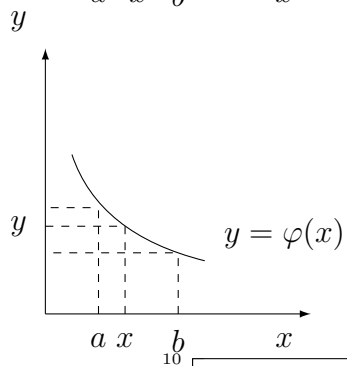
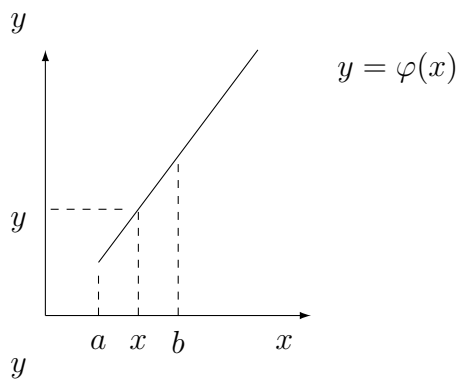
5.2 Законы распределения функций случайной величины

$Y = \varphi(x)$. Пусть φ непрерывна и дифференцируема. X – непрерывная случайная величина. Пусть значения случайной величины X лежат на (a, b) , и φ на (a, b) монотонна.

а) φ монотонно возрастает на (a, b) . $f(x)$ – плотность вероятности X , $g(y)$ – плотность вероятности Y , $G(y)$ – функция распределения Y .

$G(y) = P(Y < y) = P(a < X < x)$ (так как φ – возрастающая функция.)

$$P(a < X < x) = \int_a^x f(x) dx.$$



Пусть $\psi(y)$ – обратная функция к $\varphi(x)$. Тогда $G(Y) = \int_a^{\psi(y)} f(x)dx$. $g(y) = G'(y) =$

$$= f(\psi(y))\psi'(y)$$

б) φ – монотонно убывает на (a, b)

$$G(y) = P(Y < y) = P(x < X < b) = \int_{\psi(y)}^b f(x)dx$$

$$g(y) = G'(y) = -f(\psi(y)) \cdot \psi'(y)$$

Общая формула: $g(y) = f(\psi(y)) |\psi'(y)|$

Пример

Пушка, расположенная на расстоянии l от бесконечной прямолинейной стены, равномерно вращается. В произвольный момент времени при каждом обороте производится выстрел. Найти плотности распределения точек попаданий на стене.

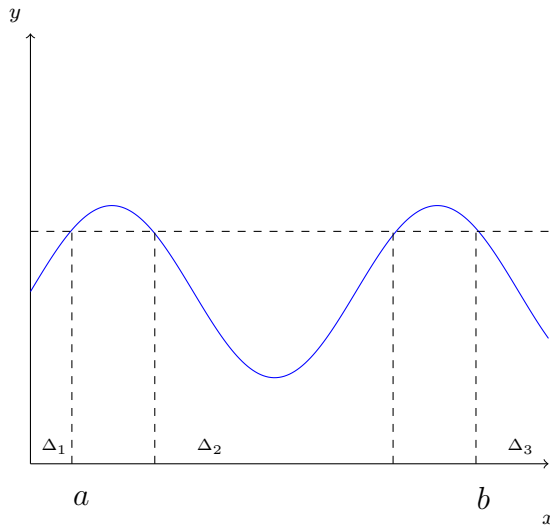
$x = l \cdot \tan \varphi$. Функция распределения равномерна на $(-\frac{\pi}{2}, \frac{\pi}{2})$:

$$f(\varphi) = \begin{cases} \frac{1}{\pi}, & \varphi \in [-\frac{\pi}{2}; \frac{\pi}{2}] \\ 0, & \varphi \notin [-\frac{\pi}{2}; \frac{\pi}{2}] \end{cases}$$

$g(x) = \frac{1}{\pi} |(arctg \frac{x}{l})'| = \frac{l}{\pi(l^2 + x^2)}$ – распределение Коши.

Пусть φ немонотонно дифференцируемая функция на (a, b) .

$$G(y) = P(Y < y) = P(X \in \Delta_1(y)) + P(X \in \Delta_2(y)) + \dots = \sum_i \int_{\Delta_i(y)} f(x) dx$$



Пример

Для получения формулы для $g(y)$ оказывается удобнее действовать по-другому. Разобьем (a, b) на интервалы монотонности φ .

$$G(y) = P(Y < y) = P((Y < y) \cap (x \in (a, a_1))) + P((Y < y) \cap (x \in (a_1, a_2))) + \dots + P((Y < y) \cap (x \in (a_{n-1}, b)))$$

$g(y) = G'(y)$. Для каждого слагаемого производная находится по формуле для монотонной функции. Пусть на i_m интервале монотонности обратная функция

к $\varphi(x)$ – это $\psi_i(y)$. Тогда $g(y) = \sum_{i=1}^n f(t_i(y)) |\psi(y)|$

Пример

$f(x) = \frac{1}{\pi(1+x^2)}$, $Y = X^2$, $X = \pm\sqrt{Y}$ (два интервала монотонности)

$$\begin{cases} g(y) = \frac{1}{\pi(1+y)} \left| \frac{1}{2\sqrt{y}} \right| + \frac{1}{\pi(1+y)} \left| -\frac{1}{2\sqrt{y}} \right| = \frac{1}{\pi(1+y)} \left| \frac{1}{\sqrt{y}} \right|, y > 0 \\ g(y) = 0, y \leq 0 \end{cases}$$

Для дискретной случайной величины $Y = \varphi(x)$ надо найти все $y_i = \varphi(x_i)$ и если среди y_i есть совпадающие, то надо сложить соответствующие вероятности.

x_1	x_2	x_3	...
p_1	p_2	p_3	...
$\varphi(x_1)$	$\varphi(x_2)$	$\varphi(x_3)$...

переходит в

y_1	y_2	y_3	...
$\sum_{y_1=\varphi(x_i)} p_i$	$\sum_{y_2=\varphi(x_i)} p_i$	$\sum_{y_3=\varphi(x_i)} p_i$...

Пример

x_i	$-\frac{\pi}{4}$	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	$\frac{5\pi}{4}$
p_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$ \cos x_i $	$\frac{\sqrt{2}}{2}$	1	$\frac{\sqrt{2}}{2}$	0	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$

$Y = \cos(x)$

y_i	0	$\frac{\sqrt{2}}{2}$	1
p_i	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$

6 Предельные теоремы теории вероятностей

Предельные теоремы связаны с большим числом однородных опытов или с большим числом складывающихся случайных воздействий. Суть "закона больших чисел" в том, что при большом числе случайных явлений средний их результат стабилизируется и практически перестает быть случайным.

Неравенство Чебышева

Пусть X – случайная величина с математическим ожиданием m_x , дисперсией D_x .

Тогда $P(|X - m_x| \geq \alpha) \leq \frac{D_x}{\alpha^2}$

Доказательство:

Для непрерывных случайных величин.

$$D_x = \int_{-\infty}^{\infty} (x - m_x)^2 f(x) dx \geq \int_{|x-m_x|>\alpha} |x - m_x|^2 f(x) dx \geq \\ \geq \alpha^2 \cdot \int_{|x-m_x|>\alpha} f(x) dx = \alpha^2 P(|X - m_x| \geq \alpha)$$

Для дискретных случайных величин – аналогично.

■

Пример

$$P(|X - m_x| \geq 3\sigma_x) \leq \frac{D_x}{9\sigma_x^2} = \frac{1}{9}$$

Видно, что оценка по приближению Чебышева завышена. Например для нормального распределения это $0.003 \ll \frac{1}{9}$

Пусть X – случайная величина с m_x и D_x

X_1 – значение X в первом опыте. X_2 – во втором и так далее. $Y = \frac{\sum_{i=1}^n x_i}{n}$

$$M(Y) = \frac{1}{n} \sum_{i=1}^n M(X_i) = \frac{1}{n} n m_x = m_x$$

$$D(Y) = \frac{1}{n^2} \sum_{i=1}^n D(X_i) = \frac{D_x}{n}$$

Теорема(Чебышева)

При неограниченном увеличении числа независимых опытов среднее арифметическое наблюдаемых значений случайной величины, имеющих ограниченную дисперсию D_x , сходится по вероятности к её математическому ожиданию, то есть $\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n x_i - m_x\right| < \varepsilon\right) = 1$, где ε – сколь угодно маленькое положительное число.

Доказательство:

Y имеет $m_y = m_x$, $D_y = \frac{1}{n} D_x$ Неравенство Чебышева для Y :

$$P(|Y - m_y| \geq \varepsilon) \leq \frac{D_y}{\varepsilon^2} = \frac{D_x}{n\varepsilon^2}. \text{ Пусть } \varepsilon \text{ и } \Delta \text{ сколь угодно малые. Выберем } n \text{ так, что } \frac{D_x}{n\varepsilon^2} < \Delta. \text{ Тогда } P\left(\left|\frac{1}{n} \sum_{i=1}^n x_i - m_x\right| \geq \varepsilon\right) < \Delta, \text{ переходим к}$$

противоположному событию, $P\left(\left|\frac{1}{n}\sum_{i=1}^n x_i - m_x\right| < \varepsilon\right) > 1 - \Delta$.

■

Теорема (Бернулли)

При неограниченном увеличении числа опытов n частота события A сходится по вероятности к его вероятности p .

Доказательство:

Частота $P^* = \frac{m}{n}$. X_1 – число появлений события A в одном опыте, X_2 – в двух и так далее. Закон распределения для всех X_i – одинаков:

x_i	0	1
p_i	q	p

$M(X_i) = P$, $D(X_i) = pq$.

$P^* = \frac{1}{n}\sum_i^n x_i$, следовательно, по теореме Чебышева $P(|P^* - p| < \varepsilon) > 1 - \delta$

■

Замечание

Математическое ожидание и дисперсия биномиального распределения

находятся проще с помощью системы случайных величин $X = \sum_{i=1}^n X_i$, $M(X_i) =$

$= p$, $D(X_i) = pq$. $M(\sum_{i=1}^n X_i) = np$, $D(\sum_{i=1}^n X_i) = npq$ – для независимых случайных величин

Центральная предельная теорема

Если X_1, X_2, \dots, X_n – независимые случайные величины, имеющие один и тот же закон распределения с математическим ожиданием m и дисперсией σ^2 , то при

неограниченном увеличении n закон распределения $Y_n = \sum_{i=1}^n X_k$ неограниченно

приближается к нормальному.

Частный случай этой теоремы – теорема Муавра-Лапласа.

Если произведение n независимых опытов, в каждом из которых событие A появляется с вероятностью p , то справедливо соотношение:

$$P(\alpha < Y < \beta) \approx \Phi^*\left(\frac{\beta - np}{\sqrt{npq}}\right) - \Phi^*\left(\frac{\alpha - np}{\sqrt{npq}}\right)$$

Y – число появлений события A в опытах $q = 1 - p$.

Пример

60% изготовленных изделий – первого сорта. Приёмщик берет наугад 200 изделий. Чему равна вероятность того, что среди них от 120 до 150 изделий первого сорта?

$$P(120 < Y < 150) \approx \Phi\left(\frac{150-120}{\sqrt{48}}\right) - \Phi\left(\frac{120-120}{\sqrt{48}}\right) = \Phi(4.33) - \Phi(0) \approx 0.5$$

7 Математическая статистика

7.1 Гистограмма. Статистическая функция распределения.

Математическая статистика имеет дело с анализом экспериментальных данных, получаемых в результате наблюдения массовых случайных явлений.

Основные задачи:

- 1) Задача определения закона распределения случайной величины по статистическим данным. В связи с ограниченностью экспериментального материала необходимо научиться отбирать в нем существенные и отбрасывать то, что не присуще закону распределения.
- 2) Задача проверки правдоподобия гипотез. Например, согласуются ли экспериментальные данные с тем, что случайная величина распределена по данному закону, или существует или нет зависимость между двумя случайными величинами.
- 3) Задача отыскания неизвестных параметров распределения. Часто закон распределения находить не требуется, или он известен заранее. Требуется находить лишь числовые характеристики.

Пусть надо определить закон распределения некоторой случайной величины X . Производится ряд измерений величины X . Совокупность полученных результатов называется простой статистической совокупностью (выборкой). Происхождение названия: часто в статистике рассматривается распределение какого-то признака среди большой группы объектов (например вес коровы).

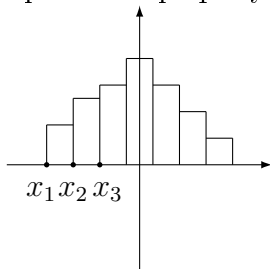
Множество всех объектов при этом называется генеральной совокупностью. Для оценки признака выбирают часть этой совокупности (её называют выборкой). Например, Остап Бендер изучал распределение бриллиантов в стульях и проводил выборку из генеральной совокупности.

Определение

Статистическая функция распределения случайной величины X – это частота события $X < x$ в данной выборке: $F^*(x) = P^*(X < x)$. Это ступенчатая функция. По теореме Бернулли, при n стремящемся к бесконечности, частота событий стремится к вероятности, $F^*(x)$ сходится по вероятности к $F(x)$. При большом объеме (n) выборки строить так $F^*(x)$ трудоёмко. Там действуют так: весь интервал полученных значений X разбивают на разряды (интервалы). Подсчитывают число наблюдений, приходящихся на данных интервал m_i и находят частоту $p_i^* = \frac{m_i}{n}$.

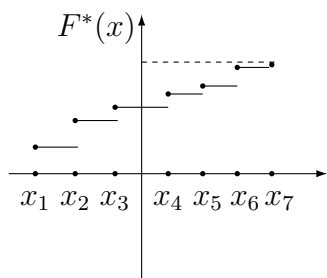
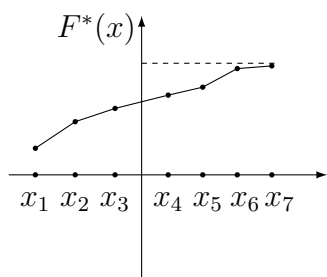
I_i	(x_1, x_2)	(x_2, x_3)	\dots	(x_k, x_{k+1})
p_i^*	p_1^*	p_2^*	\dots	p_k^*

Эта таблица называется статистическим рядом. Длины разрядов можно брать и одинаковыми, и различными. Графически, статистический ряд оформляется в виде гистограммы. На каждом интервале откладывается вверх прямоугольник высоты $\frac{p_i^*}{l_i}$, где l_i – длина интервала. (на длинну интервала можно не делить , если они все одинаковые). Площадь гистограммы равна единице. Если число опытов и число разрядов увеличивать, то гистограмма буде стремиться к плавной кривой – графику плотности вероятности случайной величины X .



Для построения статистической функции распределения обычно приписывают p_i^* середине i -го интервала $\frac{x_i + x_{i+1}}{2} = x_i^*$. И дальше строят ступенчатую функцию для данного статистического ряда. Иногда строят не ступенчатую, а например функцию соединяя значения, соответствующие концам интервала

Можно ввести по аналогии с вероятностными – статистические числовые характеристики. Однако, в большей степени они интересны не сами по себе, а той информацией, которую они содержат об измеряемой величине.



7.2 Точечные оценки параметров распределения

Часто вид распределения известен заранее, а определить надо его параметры (например для нормального m и σ), или закон распределения нас не интересует, а нужны лишь его числовые характеристики. Если при этом после статистической обработки мы предлагаем для данного параметра в качестве приближения одно число, то это число называется точечной оценкой параметра. Пусть у случайной величины X закон распределения содержит параметр a . Мы производим n независимых измерений величины X и получаем значения X_1, X_2, \dots, X_n . Их можно рассматривать как n экземпляров случайной величины X , то есть n независимых случайных величин с тем же законом распределения. На основе X_1, X_2, \dots, X_n мы находим оценку \hat{a} параметра $a = \hat{a} = \hat{a}(X_1, X_2, \dots, X_n)$, то есть оценка это случайная величина. Её закон распределения зависит от закона распределения X и от числа опытов. Какие оценки хороши?

Определение

Оценка \hat{a} называется состоятельной, если при $n \rightarrow \infty$ она сходится по вероятности к a

Определение

Оценка \hat{a} называется несмещённой, если её математическое ожидание равно a . $M(\hat{a}) = a$. Это говорит об отсутствии систематических ошибок при оценивании.

Определение

Несмещённая оценка называется эффективной, если её дисперсия минимальна

по сравнению с другими оценками.

Оценки математического ожидания и дисперсии

Пусть случайная величина X имеет $M(X) = m$ и $D(X) = D$ (M и D неизвестны). X_1, X_2, \dots, X_n – наблюдаемые величины X .

$$\hat{m} = \frac{\sum_{i=1}^n x_i}{n}$$

Это состоятельная оценка так как по закону больших чисел, при $n \rightarrow \infty$ \hat{m} сходится по вероятности к m . Оценка \hat{m} несмещённая, так как

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

$$\text{Дисперсия оценки: } D(\hat{m}) = \frac{1}{n^2} \sum_{i=1}^n D(X_i) = \frac{D}{n}$$

Отсюда видно, зачем надо производить несколько измерений, а потом усреднять – уменьшать дисперсию (точность увеличивается в \sqrt{n} раз).

Эффект оценки зависит от вида распределения. Если X распределена нормально, то оценка \hat{m} эффективна. Для других законов это может не выполняться. Если n велико (практически $n > 30$) и x_i некоррелируют, то по центральной предельной теореме \hat{m} подчиняется нормальному распределению со средним значением m и среднеквадратичным отклонением $\sqrt{\frac{D}{n}}$

Оценка дисперсии.

$$\text{Можно предположить: } S^2 = \frac{\sum_{i=1}^n (X_i - \hat{m})^2}{n}, \quad \hat{m} = \frac{\sum_{i=1}^n x_i}{n}$$

1) Оценка состоятельна

$$S^2 = \frac{\sum_{i=1}^n (X_i)^2}{n} - \hat{m}^2$$

Это среднее арифметическое n наблюдений значений величины X^2 . Он сходится по вероятности (по закону больших чисел) к $M(X^2) = \alpha_2(X)$. Второй член сходится по вероятности к m^2 . То есть $S^2 \xrightarrow{n \rightarrow \infty} \alpha_2(x) - m^2 = D$ (по вероятности)

2) Оценка S^2 – смещена.

$$S^2 = \frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2 = \frac{\sum_{i=1}^n x_i^2}{n} - \frac{\sum_{i=1}^n x_i^2}{n^2} - 2 \frac{\sum_{i < j} x_i x_j}{n^2} = \frac{n-1}{n} \sum_{i=1}^n x_i^2 - 2 \frac{\sum_{i < j} x_i x_j}{n^2}$$

$$M(S^2) = \frac{n-1}{n^2} \sum_{i=1}^n M(X_i^2) - \frac{2}{n^2} \sum_{i < j} M(X_i X_j)$$

S^2 не зависит от выбора начала координат. Пусть начало координат в точке m .

Тогда $M(X_i^2) = M(\hat{X}_i^2) = D$, $\sum_{i=1}^n M(X_i^2) = nD$, $M(X_i X_j) = M(\hat{X}_i \hat{X}_j) = K_{ij} = 0$.

(так как опыты независимы, следовательно некоррелируют).

Значит, $M(S^2) = \frac{n-1}{n} D \neq D$

Что бы оценка стала несмещённой, ее надо умножить на $\frac{n}{n-1}$.

$$\hat{D} = \frac{\sum_{i=1}^n ((x_i - \hat{m})^2)}{n-1} - \text{несмещённая оценка.} \quad \left(\hat{D} = \frac{\sum_{i=1}^n x_i^2}{n-1} - \hat{m}^2 \frac{n}{n-1} \right)$$

При больших значениях $n - S^2$ и \hat{D} различаются мало.

При известном математическом ожидании (m) несмещённая оценка дисперсии: $\hat{D} = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2$.

(То есть $\frac{1}{n-1}$ происходит из-за того, что была оценкой \hat{m} , а не m).

$$M \left(\frac{1}{n} \sum_{i=1}^n (X_i - m)^2 \right) = \frac{1}{n} \sum_{i=1}^n M((X_i - m)^2) = \frac{1}{n} \sum_{i=1}^n D = D, \quad \text{то есть она несмещённая.}$$

Пример

Для определения точности измерительного прибора (лазерного дальномера) систематическая ошибка которого практически равна нулю, было произведено 5 независимых измерений.

$x_i = 2781$ метра, 2836 метра, 2807 метра, 2763 метра, 2858 метра. Определить несмещённую оценку дисперсий ошибок измерительного прибора, если значение измеряемой величины:

а) известно и равно 2800 м

б) неизвестно

Решение

$$\text{a) } \hat{D}(X) = \frac{1}{n} \sum_{j=1}^n (x_j - m)^2 = \frac{6439}{5} = 1287.8 \text{ квадратных метров}$$

$$\text{b) } \hat{m} = \frac{1}{n} \sum_{j=1}^n (x_j) = 2809 \text{ метров}$$

$$\hat{D}(X) = \frac{1}{n} \sum_{j=1}^n (x_j - \hat{m})^2 = \frac{6034}{4} = 1508.5 \text{ квадратных метров}$$

7.3 Интервальные оценки параметров

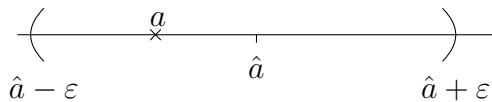
Если для оценки параметра предлагается интервал, в который его значение попадает с заданной вероятностью, то такая оценка называется интервальной.

Определение

Доверительным интервалом называется интервал, который с заданной доверительной вероятностью β накрывает оцениваемый параметр a .

Оценка \hat{a} случайна, и поэтому замена a на \hat{a} может привести к ошибкам. Интервалы оценки вводятся для того, что бы можно было судить о надёжности оценки. Для симметричного доверительного интервала его ширина 2ε определяется условием: $P(|\hat{a} - a| \leq \varepsilon) = \beta$

Доверительный интервал: $(\hat{a} - \varepsilon, \hat{a} + \varepsilon) = I_\beta$



Замечание

Здесь a не является случайным, случаен интервал, положение которого задаётся случайным положением \hat{a} . $\hat{a} - \varepsilon$ и $\hat{a} + \varepsilon$ – доверительные границы.

Пусть проведено n измерений величины X (с $M(X) = m$, $D(X) = D$)

$$\text{Оценки: } \hat{m} = \frac{1}{n} \sum_{i=1}^n X_i, \hat{D} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{m})^2$$

Построим доверительный интервал I_β для математического ожидания \hat{m} – сумма n независимых одинаково распределённых случайных величин X_i .

По центральной предельной теореме при большом n закон распределения \hat{m} близок к нормальному с математическим ожиданием m и дисперсией $\frac{D}{n}$. Пусть дисперсия D известна. Найдём доверительные границы $P(|\hat{m} - m| < \varepsilon_\beta) = \beta$.

Так как \hat{m} распределен нормально с $\sigma_{\hat{m}} = \sqrt{\frac{D}{n}}$

$$\beta = P(|\hat{m} - m| < \varepsilon_\beta) = 2\Phi^*\left(\frac{\varepsilon_\beta}{\sigma_{\hat{m}}}\right) - 1.$$

$$\Phi^*(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

Находим t_β по таблице функции $\Phi^*(x)$: $2\Phi^*(t_\beta) - 1 = \beta$

$$\Phi^*(t_\beta) = \frac{\beta+1}{2}$$

$$\text{Тогда, } \varepsilon_\beta = t_\beta \sqrt{\frac{D}{n}}$$

$$\text{Доверительный интервал: } (\hat{m} - t_\beta \sqrt{\frac{D}{n}}, \hat{m} + t_\beta \sqrt{\frac{D}{n}})$$

Если D неизвестна, то в качестве её ориентировочного значения можно взять \hat{D} и положить приближенно $\sigma_{\hat{m}} = \sqrt{\frac{\hat{D}}{n}}$

Но при этой замене возникает дополнительная ошибка. Для точного определения доверительного интервала необходимо знать вид закона распределения X . Пусть, например, X распределено по нормальному закону.

$$\text{Произведено } n \text{ измерений. } \hat{m} = \frac{1}{n} \sum_{i=1}^n X_i, \hat{D} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{m})^2.$$

Оказывается, что величина $T = \sqrt{n} \frac{\hat{m} - m}{\sqrt{\hat{D}}}$ подчиняется закону распределения Стьюдента с $(n - 1)$ степенями свободы, не зависящих от неизвестных параметров m и D , плотность которого $S_{n-1} = \frac{\Gamma(\frac{n}{2})}{\sqrt{(n-1)\pi}} (1 + \frac{t^2}{n-1})^{-\frac{n}{2}}$,

$$\text{где } \Gamma(x) - \text{гамма-функция Эйлера. (} \Gamma(x) = \int_0^\infty u^{x-1} e^{-u} du \text{.)}$$

Найдем доверительный интервал соответствующий вероятности β :

$$P(|\hat{m} - m| < \varepsilon) = \beta \Leftrightarrow P\left(\frac{\sqrt{n}|\hat{m} - m|}{\sqrt{\hat{D}}} < \frac{\varepsilon_\beta}{\sqrt{\frac{\hat{D}}{n}}}\right) = \beta \Leftrightarrow P\left(|T| < \frac{\varepsilon_\beta}{\sqrt{\frac{\hat{D}}{n}}}\right) = \beta.$$

Найдем t_β из : $P(|T| < t_\beta) = \beta$

$$P(|T| < t_\beta) = \int_{-t_\beta}^{t_\beta} S_{n-1}(t) dt = \beta.$$

$$(\text{Замечание: } S_{n-1} - \text{четная функция, поэтому } \int_{-t_\beta}^{t_\beta} S_{n-1}(t) dt = 2 \int_0^{t_\beta} S_{n-1}(t) dt)$$

$t_\beta(n)$ называется коэффициент Стьюдента. Для них существуют таблицы.)

Определив по таблице t_β находим полуширину доверительного интервала $\varepsilon_\beta = t_\beta \sqrt{\frac{\hat{D}}{n}}$. Доверительный интервал при неизвестной дисперсии: $(\hat{m} - t_\beta \sqrt{\frac{\hat{D}}{n}}, \hat{m} + t_\beta \sqrt{\frac{\hat{D}}{n}})$, t_β – коэффициент Стьюдента при $(n - 1)$ степени свободы.

Пример

Оцениваем средний рост мужчины. Выборка 10 человек. Результат: 160, 160, 167, 170, 173, 176, 178, 178, 181, 181.

Найдем $\hat{m} = 172.4$. $\hat{D} = \frac{1}{9} \sum_{i=1}^{10} (x_i - 172.4)^2 = 62.9$

$\sqrt{\frac{\hat{D}}{n}} = 2.51$. Если $\beta = 0.95$, то $t_\beta(9) = 1.83$. $\varepsilon_\beta = 1.83 \cdot 2.51 \approx 4.6$, то есть $m = 172.4 \pm 4.6$ с вероятностью 0.95.

Замечание

Физики обычно предпочитают приводить результат \pm одно среднее квадратичное отклонение. Из таблиц распределения Стьюдента следует, что доверительный интервал полушириной в одно среднее квадратичное отклонение содержит истинный результат с вероятностью всего 0.66.

Доверительный интервал для дисперсии

\hat{D} – сумма n случайных величин вида: $\frac{(x_i - \hat{m})^2}{n-1}$. Они зависимы, так как \hat{m} зависит от всех X_i . Но при $n \rightarrow \infty$ распределение их суммы приближается к нормальному. При $n \in (20, 30)$ его уже приближенно можно считать нормальным. Опишем приближённый способ построения доверительного интервала. Найдем характеристики этого приближенного нормального распределения. $M(\hat{D}) = D$. (оценка \hat{D} несмещенна.)

$$D(\hat{D}) = \frac{1}{n} \mu_4 - \frac{n-3}{n(n-1)} D^2$$

Вместо D можно взять \tilde{D} . Брать вместо μ_4 его оценку $\frac{1}{n} \sum_{i=1}^n (x_i - \hat{m})^4$ плохо, так

как при небольшом n ошибка будет велика. Если известен закон распределения X , то μ_4 можно выразить через D . Для нормального распределения X : $\mu_4 = 3D^2$, следовательно, $D(\hat{D}) = \frac{3}{n} D^2 - \frac{n-3}{n(n-1)} D^2 = \frac{2}{n-1} D^2 \approx \frac{2}{n-1} \hat{D}^2$. $\sigma_{\hat{D}} = \sqrt{\frac{2}{n-1}} \hat{D}$.

Пусть доверительная вероятность β . Найдем t_β : $\Phi^*(t_\beta) = \frac{\beta+1}{2}$, по таблице доверительного интервала для дисперсии: $(\hat{D} - t_\beta \sigma_{\hat{D}}, \hat{D} + t_\beta \sigma_{\hat{D}})$

Точное построение доверительного интервала для дисперсии

Известно, что $V = \frac{(n-1)\hat{D}}{D}$ имеет распределение χ^2 с $(n-1)$ степеней свободы, плотность которой:

$$K_{m-1}(V) = \begin{cases} 0, & V < 0 \\ \frac{1}{2^{\frac{n-1}{2}} \Gamma(\frac{n-1}{2})} V^{\frac{n-1}{2}-1} e^{-\frac{V}{2}}, & V > 0 \end{cases}$$

Будем выбирать интервал l_β так, чтобы вероятности выхода V за пределы l_β влево и вправо были одинаковы и равны.

$\frac{\alpha}{2} = \frac{1-\beta}{2}$. Распределение χ^2 затабулировано. Здесь приведем числа χ^2 такие, что $P(V > \chi^2) = p$, r – число степеней свободы. ($r = n-1$).

Зафиксируем $r = n-1$ и найдем по таблице два значения χ^2 , одно соответствует $p_1 = \frac{\alpha}{2}$, другое $p_2 = 1 - \frac{\alpha}{2}$. Обозначим их χ_1^2 и χ_2^2 . Интервал $l_\beta = (\chi_2^2, \chi_1^2)$.

Доверительный интервал для дисперсии $P(D_1 < D < D_2) = \beta$ – он будет тогда, когда $V = \frac{(n-1)\hat{D}}{D}$ попадет в l_β .

Если $V < \chi_1^2$, то $\frac{(n-1)\hat{D}}{D} M < \chi_1^2$ следовательно $\frac{(n-1)\hat{D}}{\chi_1^2} < D$.

Если $V > \chi_2^2$, то $\frac{(n-1)\hat{D}}{D} M > \chi_2^2$ следовательно $\frac{(n-1)\hat{D}}{\chi_2^2} > D$.

То есть доверительный интервал для дисперсии: $\left(\frac{(n-1)\hat{D}}{\chi_1^2}, \frac{(n-1)\hat{D}}{\chi_2^2} \right)$

7.4 Оценки числовых характеристик случайных величин

Система (X, Y) . Результат n опытов $(x_1, y_1), \dots, (x_n, y_n)$.

Несмещенные оценки:

$$\hat{m}_x = \frac{1}{n} \sum_{i=1}^n x_i, \quad \hat{m}_y = \frac{1}{n} \sum_{i=1}^n y_i,$$

$$\hat{D}_x = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{m}_x)^2, \quad \hat{D}_y = \frac{1}{n-1} \sum_{i=1}^n (y_i - \hat{m}_y)^2.$$

$$\hat{K}_{XY} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{m}_x)(y_i - \hat{m}_y)$$

7.5 Проверка гипотез

Рассмотрим задачу проверки опытных данных заданому закону распределения.

Критерий согласия χ^2 Пирсона по измерению X

Пусть в результате опытов (независимых) получены результаты, которые сведены в статистический ряд. (p_i^* – частоты)

x_1, x_2	...	x_k, x_{k+1}
p_1^*	...	p_k^*
p_1	...	p_k

Мы предполагаем, что X распределена по некоторому "теоретическому" закону. На основании его вычисляем теоретические вероятности попадания случайной величины в $(x_1, x_2), \dots, (x_k, x_{k+1})$. Мера расхождения

$U = \sum_{i=1}^k c_i (p_i^* - p_i)^2$. "Весовые" коэффициенты надо выбирать обратно пропорционально вероятностям, так как одно и то же значение $p^* - p$ более значимо если само p_i мало. Возьмем $c_i = \frac{n}{p_i}$. Тогда U будет распределено по "закону χ^2 ".

Замечание

Пусть X_i – независимые случайные величины, распределенные, по нормальному закону с $M(X_i) = 0, \sigma_{x_i} = 1$. Тогда величина $U = \sum_{i=1}^r x_i^2$ распределена по закону

с χ^2 с r степенями свободы. У нас: $U = \chi^2 = \sum_{i=1}^k \frac{n(p_i^* - p_i)^2}{p_i} = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i}$, так как $p_i^* = \frac{m_i}{n}$.

Число степеней свободы r равно числу разрядов k минус число независимых условий, наложенных на частоты p_i^* . Одно из таких условий, которое всегда

накладывается : $\sum_{i=1}^k p_i^* = 1$

Если все параметры "теоретического" распределения заранее известны, то больше условий нет. Если эти параметры мы оцениваем из результатов опытов,

то добавляется еще условие, например $\sum_{i=1}^k x_i^* p_i^* = m_x, \sum_{i=1}^k (x_i^* - \hat{m}_x)^2 p_i^* = D_x$, и соответственно уменьшается число степеней свободы.

Рассмотрим на примере действие критерия χ^2 .

Пример

20 лет собирали сведения о количестве офицеров русской армии, погибших в результате гибели под ними коня. (на основании ежегодных докладов 10 армейских корпусов, то есть 200 донесений).

Пусть X – число погибших в одном корпусе за год.

Количество погибших, x_i	0	1	2	3	4	5	всего
Количество донесений m_i с указанием x_i погибших	109	65	22	3	1	0	200
Ожидаемая (по пуассону) частота p_i	108.7	66.3	20.2	4.1	0.6	0.07	200

Полное число погибших: $0 \cdot 109 + 1 \cdot 65 + 2 \cdot 22 + 3 \cdot 3 + 4 \cdot 1 = 122$.

Среднее число погибших: $\frac{122}{200} = 0.61$. Проверим гипотезу о том, что эти данные описывают распределение Пуассона: $P(x = x_i) = p_i = \frac{\mu^i e^{-\mu}}{i!}$. Для определения p_i надо оценить одним параметром μ . В качестве μ возьмем $\hat{\mu} = 0.61$. np_i приведен в третьей строке. Так как ожидаемые частоты для $x_i > 2$ малы, объединим их $U = \chi^2 = \sum \frac{(m_i - np_i)^2}{np_i} = \frac{(0.3)^2}{108.7} + \frac{(1.3)^2}{66.3} + \frac{(1.8)^2}{20.2} + \frac{(0.8)^2}{4.8} = 0.32$.

x_i	0	1	2	≥ 3
m_i	109	65	22	4
np_i	108.7	66.3	20.2	4.8

Есть 2 линейки соотношений, связывающие частоты: $\sum m_i = 200$ и $\sum x_i m_i = 122 = \hat{\mu}$.

То есть число степеней свободы: $r = 4 - 1 - 1 = 2$. То есть U распределено примерно по закону χ^2 с двумя степенями свободы. U распределено примерно по закону χ^2 с двумя степенями свободы.

Из таблицы распределения χ^2 с $r = 2$ находим, что с вероятностью 0.95 χ^2 должно лежать в интервале от 0 до 6.0 так как наше значение 0.32 попадает в этот незначимый интервал, то проверяемая гипотеза не отвергается. В таблице указана вероятность p того, что $\chi^2 \geq \chi_p^2$ при числе степеней свободы r . Если бы мы получили для χ^2 значение 6.1 или больше, то должны были бы считать его значимым, так как если проверяемая гипотеза верна, то такие значения χ^2 может принимать с вероятностью не более 0.05. 0.05 называется уровнем значимости. Обычно, схема действий такая: По экспериментальным данным составляется функция U и на основе "нулевой гипотезы" H вычисляется её распределение. Затем проверяют, как изменится U , если гипотезы H неверна (верна \bar{H}). Обычно U в этом случае намного больше. На рисунке показано два распределения величины U , соответствующие двум конкурирующим гипотезам

$H(a)$ и $\hat{H}(\hat{a})$. Если из опыта получилось большое значение U , то можно сказать, то либо справедлива гипотеза H , но произошло чрезвычайно маловероятное событие, либо более правдоподобна гипотеза \bar{H} . Если оставаться в рамках одной гипотезы, то результат следует признать значимым в том смысле, что он выходит за рамки обычных статистических ошибок.

Определение

Отвергая нулевую гипотезу H , когда она справедлива, мы совершаем ошибку первого рода. Принимая нулевую гипотезу H , когда справедлива конкурирующая гипотеза \bar{H} , мы совершаем ошибку второго рода. Вероятность ошибки первого рода называется уровнем значимости (α). Если вероятность ошибки второго рода равна β , то $1 - \beta$ называют мощностью критерия.

Часто возможен компромисс между уровнем значимости и мощностью критерия. Если $U > U_\alpha$, мы отвергаем H и принимаем \bar{H} , если $U < U_\alpha$ – наоборот.

7.6 Метод моментов

Пусть функция распределения случайной величины X зависит от k параметров a_1, a_2, \dots, a_k . Надо выбрать оценки этих параметров по результатам n опытов. Теоретические моменты $(M(X), D(X), \dots, \mu_k(x))$ являются функциями от этих параметров. По результатам опытов находим оценки моментов $\hat{M}(X), \hat{D}(X), \dots, \hat{\mu}_k(x)$. Оценки параметров a_1, a_2, \dots, a_k находим решая систему уравнений:

$$\begin{cases} m_x = \hat{m}_x \\ D(X)(a_1, \dots, a_k) = \hat{D}(X) \\ \dots \\ \mu_k(x)(a_1, \dots, a_k) = \hat{\mu}_k(X) \end{cases}$$

Пример

Найти точечные оценки параметров a, b равномерно распределенных в интервале (a, b) по результатам n опытов. По результатам n опытов находим $\hat{m}_x = \bar{x}$, и $\hat{D}(X) = \hat{\sigma}_x^2$.

Теоретические моменты:

$$M(X) = \frac{a+b}{2} = \bar{X}$$

$$D(X) = \frac{(b-a)^2}{12} = \hat{D}_x = \hat{\sigma}_x^2.$$

Решая эту систему, находим оценки для a и b : $\hat{a} = \bar{X} - \hat{\sigma}_x \sqrt{3}$, $\hat{b} = \bar{X} + \hat{\sigma}_x \sqrt{3}$.

Замечание

Об ошибке прибора при измерении. Если в паспорте прибора задан гарантированный его производителем интервал $\pm\Delta$, из которого не выходят предельные отклонения (ошибки). Тогда дисперсия ошибки $D = \frac{\Delta^2}{3}$. Если задан класс точности прибора (а), то $\Delta = \frac{a \cdot b}{100}$, где b – максимальное значение шкалы прибора.

7.7 Принцип максимального правдоподобия

Пусть есть урна, в которой неизвестная смесь черных и белых шаров.

Вероятность вытянуть белый шар – p , может принимать любое значение от 0 до 1.

Хотим изменить p , вынимая шар из урны. Пусть мы вытащили шар n раз (с возвращением) и белый шар оказался в r случаях. Вероятность этого события $l(p) = C_n^r p^r (1-p)^{n-r} = l(p|r)$ эта функция при неизвестном r (и n) и неизвестном p называется функцией правдоподобия. Выбираем то значение p , которое с наибольшей вероятностью приводит к наблюдаемому значению r . Когда параметр p дискретен, надо вычислить для разных p функцию правдоподобия и выбрать то p , при котором функция правдоподобия максимальна. Если p непрерывен, то можно искать максимум с помощью дифференцирования: $\frac{dl}{dp} = C_n^r (r p^{r-1} (1-p)^{n-r} - (n-r) p^r (1-p)^{n-r-1}) = 0$
 $r - rp - np - rp = 0, p = \frac{r}{n}$.

Определение

Логарифмическая функция правдоподобия:

$L(p|r) = \ln l(p|r) = r \ln p + (n-r) \ln(1-p)$. (здесь отброшено постоянное слагаемое, а в $l(a|r)$ – постоянный множитель, который не влияет на максимум.)
 $L(p|r)$ имеет максимум при том же значении p , что и $l(p|r)$.

Более общий случай:

Пусть X – дискретная случайная величина со значениями $x_1 \dots x_l$ и

$P(X = x_j; \theta) = p_j(\theta)$. $\theta = (\theta_1, \dots, \theta_k)$ – неизвестные параметры. Пусть m_1, \dots, m_l – число появлений значений x_1, \dots, x_l в выборке X_1, X_2, \dots, X_n . Тогда функция правдоподобия значений $l(\theta|x_1, x_2, \dots, x_n) = p_1^{m_1}(\theta) \cdot p_2^{m_2}(\theta) \cdot \dots \cdot p_l^{m_l}(\theta)$.
 $L(\theta|x_1, x_2, \dots, x_n) = \ln l(\theta|x_1, x_2, \dots, x_n)$

Пусть случайная величина X имеет непрерывное распределение с плотностью $f(x|\theta)$, θ – единственный параметр. Плотность вероятности выборки из n измерений (n независимых случайных величин):

$$\prod_{i=1}^n f(x_i|\theta) = l(\theta|x_1, x_2, \dots, x_n)$$

Логарифмическая функция правдоподобия: $L(\theta|x_1, \dots, x_n) = \sum_{i=1}^n \ln f(x_i|\theta)$

Если параметров несколько, то $\theta = (\theta_1, \theta_2, \dots, \theta_k)$. Параметры выбираются из условия максимума функции l (или L)

Пример

$$f(x|m, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

$$l(\mu, \sigma^2|x_1, \dots, x_n) = (\sigma^2)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (x_i - m)^2}{2\sigma^2}}$$

$$L(\mu, \sigma^2|x_1, \dots, x_n) = -\frac{n}{2} \ln \sigma^2 - \sum_{i=1}^n (x_i - m)^2 \frac{(x_i - m)^2}{2\sigma^2}$$

$$\frac{\partial L}{\partial m} = 0, \frac{\partial L}{\partial \sigma^2} = 0$$

Отсюда находим $\hat{m}, \hat{\sigma}^2$

$$\frac{\partial L}{\partial m} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \hat{m}) = 0, \text{ следовательно } \hat{m} = \bar{x}$$

$$\frac{\partial L}{\partial \sigma^2} = -\frac{n}{2} \frac{1}{\hat{\sigma}} + \frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n (x_i - m)^2 = 0, \text{ следовательно } \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

То есть тут оказывается смещённая оценка дисперсии.

7.8 Метод наименьших квадратов

Пусть в эксперименте определяется зависимость величины Y от величины X . Надо найти функцию $Y = \varphi(x)$ по экспериментальным данным. Часто вид $\varphi(x)$ известен (например прямая или парабола), а надо найти параметры этой кривой. Пусть истинная зависимость $Y = \varphi(x)$. Результат опыта Y_i – случайная величина, распределённая по нормальному закону с $M(Y_i) = \varphi(x_i)$ и $\sigma_{y_i} = \sigma_i$. Пусть все $\sigma_i = \sigma$. (Точность измерения везде одинакова)

$$\text{Плотность } y_i : f_i(y_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y_i - \varphi(x_i))^2}{2\sigma^2}}.$$

$$-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \varphi(x_i))^2$$

Составим функцию правдоподобия $Ce^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \varphi(x_i))^2}$, где C – константа зависящая от σ и n .

Максимум функции правдоподобия достигается тогда, когда $\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \varphi(x_i))^2$ – минимально, то есть, что бы данная совокупность наблюдений y_1, y_2, \dots, y_n

была наивероятнейшей, надо выбрать функцию $\varphi(x)$ так, что бы сумма квадратов отклонений наблюдаемых значений y_i от $\varphi(x_i)$ была минимальной.

$$\left(\sum_{i=1}^n (y_i - \varphi(x_i))^2 = \min\right)$$

Пусть функция φ зависит от параметров $a, b, \dots : y = \varphi(x, a, b, c, \dots)$

Параметры выбираются из условия $\sum_{i=1}^n (y_i - \varphi(x, a, b, c, \dots))^2$ — минимально.

Условие минимума: производные по $a, b, c \dots = 0$

$$\begin{cases} \sum_{i=1}^n (y_i - \varphi(x_i, a, b, c, \dots)) \frac{\partial \varphi}{\partial a}(x_i, a, b, c, \dots) \\ \sum_{i=1}^n (y_i - \varphi(x_i, a, b, c, \dots)) \frac{\partial \varphi}{\partial b}(x_i, a, b, c, \dots) \\ \dots \end{cases}$$

Уравнений получается столько, сколько параметров. Решаем систему относительно $a, b, c \dots$

Пример

$$Y = aX + b$$

$$\frac{\delta \varphi}{\delta a} = x, \quad \frac{\delta \varphi}{\delta b} = 1$$

$$\begin{cases} \sum_{i=1}^n (y_i - (ax_i + b))x_i = 0 \\ \sum_{i=1}^n (y_i - (ax_i + b)) = 0 \end{cases}$$

Систему можно решать относительно a, b в общем виде:

$$a = \frac{\hat{K}_{XY}}{\hat{D}_x}, \quad b = \hat{m}_y - a\hat{m}_x$$

7.9 Линеаризация функций

Функция одной переменной.

Пусть известны m_X и D_X и практически возможные значения X сосредоточены в малой окрестности математического ожидания. Надо найти m_X, D_X , если $Y = \varphi(X)$

В малой окрестности m_X $\Phi(X)$ мало отличается от линейной. По формуле

Тейлора : $Y \approx \varphi(m_X) + \varphi'(m_X)(X - m_X)$

Для линейной функции характеристики находить умеем: $m_X = \varphi(m_X)$;

$$D_Y = |\varphi'(m_X)|^2 D_X.$$

Функция нескольких переменных:

Есть система X_1, X_2, \dots, X_n с характеристиками $m_{X_1}, m_{X_2}, \dots, m_{X_n}, D_{X_1}, \dots, D_{X_n}, K_{ij}$.

$$Y = \varphi(X_1, X_2, \dots, X_n).$$

По формуле Тейлора до первого порядка:

$$y \approx \varphi(m_{X_1}, m_{X_2}, \dots, m_{X_n}) + \sum_{i=1}^n \varphi_{X_i}(m_{X_1}, m_{X_2}, \dots, m_{X_n})(X_i - m_{X_i})$$

Для линейной функции находим:

$$m_y = \varphi(m_{X_1}, m_{X_2}, \dots, m_{X_n}). \quad D_Y = \sum_{i=1}^n \left(\left(\frac{\partial \varphi}{\partial X_i} \right) (m_{X_1}, m_{X_2}, \dots, m_{X_n}) \right)^2 D_{X_i} +$$

$$+ 2 \sum_{i < j} \left(\frac{\partial \varphi}{\partial X_i} \right) (m_{X_1}, m_{X_2}, \dots, m_{X_n}) \cdot \left(\frac{\partial \varphi}{\partial X_j} \right) (m_{X_1}, m_{X_2}, \dots, m_{X_n}) K_{ij}$$

Если X_1, \dots, X_n не коррелируют, то $D_Y = \sum_{i=1}^n \left(\frac{\partial \varphi}{\partial X_i} (m_{X_1}, m_{X_2}, \dots, m_{X_n}) \right)^2$, то

$$\text{есть ошибка } \sigma_y = \sqrt{\sum_{i=1}^n \left(\frac{\partial \varphi}{\partial X_i} (m_{X_1}, m_{X_2}, \dots, m_{X_n}) \right)^2 \sigma_{X_i}^2}$$

Пример

Оценить количество тепла, выделяемого в проводнике сопротивления R за время T при токе I : $Q = I^2 R T$ и дисперсию оценки $\hat{\sigma}_Q$, если $I = 10$ ампер, $\hat{R} = 30$ см, $\hat{T} = 10$ минут, $\hat{\sigma}_I = 0.1$ ампер, $\hat{\sigma}_R = 0.2$ сантиметра, $\hat{\sigma}_T = 0.5$ секунд. $\hat{Q} = \hat{I}^2 \hat{R} \hat{T} = 100 \cdot 30 \cdot 600$ джоулей = 432 ккал.

$$D(\hat{Q}) = \left(\frac{\partial \hat{Q}}{\partial I} \right) \sigma_I^2 + \left(\frac{\partial \hat{Q}}{\partial R} \right) \sigma_R^2 + \left(\frac{\partial \hat{Q}}{\partial T} \right) \sigma_T^2 = 2 \hat{I} \hat{R} \hat{T} \sigma_I^2 + \hat{I}^2 \hat{T} \sigma_R^2 + \hat{I}^2 \hat{R} \sigma_T^2$$

$$\sigma(\hat{Q}) = \sqrt{D(\hat{Q})} = \sqrt{(2 \cdot 12 \cdot 10^6)^2 \cdot 0.01 + (6 \cdot 10^3)^2 \cdot 0.04 + (3 \cdot 10^3)^2 \cdot 0.25} =$$

$$= \sqrt{8.7} \cdot 10^3 \text{ ккал.}$$

8 Случайные процессы

При изучении различных явлений мы сталкиваемся с процессами, течение которых заранее не можем предсказать: колебание высоты полета, движение отдельной молекулы в газе, различных бактерий в среде. В каждом случае речь идет о движении точки в каком то пространстве (реальном или придуманном),

а это есть функция времени со значением в этом пространстве. То есть модель случайного процесса – функция от t , значения которой – случайные величины.

8.1 Основные определения

Определение

Случайной функцией называется семейство случайных величин, зависящих от параметра t , пробегающего произвольное множество Ω . В случае когда Ω – подмножество вещественной прямой случайная функция называется случайным процессом (t интерпретируется как время), когда Ω состоит из целых чисел, говорят о случайной последовательности. Зафиксировав t мы получаем случайную величину, зафиксировав элементарное событие, получаем неслучайную функцию от t , она называется реализацией случайной функции (траекторией). Проводя несколько опытов, получаем семейство реализаций (например, при нескольких полетах определяем отклонение высоты от заданой). Зафиксируем в несколько моментов t , получим систему случайных величин, то есть случайная функция – обобщение системы случайных величин. Не будем рассматривать законы распределения случайной функции, остановимся только на характеристиках. Если числовые характеристики случайной величины – числа, то соответствующие характеристики случайных функций – функции (неслучайные).

Определение

Математическое ожидание $M(X(t)) = m_X(t)$ случайной функции $X(t)$ называется неслучайная функция $m_X(t)$, которая при каждом значении t равна математическому ожиданию соответствующего сечения случайной функции, дисперсия $D_X(t) = D(X(t))$. Этих характеристик явно не достаточно. Например, процессы показанные на рисунках, будут иметь примерно одинаковые математические ожидания и дисперсии. Чтобы уловить их разницу, надо указать связь между значениями случайной функции $X(t)$ в различные моменты времени. Ее характеризует корреляционная функция.

Определение

Корреляционной (автокорреляционной) функцией случайной функции $X(t)$ называется неслучайная функция двух аргументов, $K_X(t, t')$, которая при каждой паре t, t' равна корреляционному моменту соответствующих сечений случайной функции: $K_X(t, t') = M(\overset{\circ}{X}(t) \cdot \overset{\circ}{X}(t')) = M((X(t) - m_X(t))(X(t') - m_X(t')))$.

Если $t = t'$, то $k_X(t, t) = D_X(t)$

Свойство: $K_X(t, t') = K_X(t', t)$ (из свойства корреляции моментов)

Определение

Нормированная корреляционная функция : $r_x(t, t') = \frac{K_x(t, t')}{\sigma_X(t)\sigma_X(t')}$, $r_X(t, t) = 1$.

Простейшая характеристика закона распределения случайного процесса (функции) – одномерный закон распределения случайной функции $X(t)$: для любого t задается закон распределения $X(T)$, то есть функция двух параметров: $f(x, t)$. Зависимость случайных величин $X(t)$ при различных t показывается двумерный закон распределения: $f(x_1, x_2, t_1, t_2)$ – закон распределения системы $X_1(t), X_2(t)$. Можно дальше увеличивать число аргументов и получить различные конечномерные законы распределения.

8.2 Стационарный случайный процесс

Определение

Случайный процесс $X(t)$, $t \in \Omega \subseteq \mathbb{R}$ называется стационарным, если для любого вещественного h его конечномерное распределение не меняются при сдвиге на h :

$f(x_1, \dots, x_n, t_1, \dots, t_n) = f(x_1, \dots, x_n, t_1 + h, \dots, t_n + h)$, если $t_1, \dots, t_n, t_1 + h, \dots, t_n + h \in \Omega$.

В частности, у стационарного случайного процесса не меняются случайные числовые характеристики.

Пример

Стационарный случайный процесс – отклонение высоты полеты от установившейся.

$m_X(t) = m_X = \text{const}$, $D_X(t) = D_X = \text{const}$.

Корреляционный момент $K_X(t, t + \tau)$ у стационарного процесса не должен зависеть от t :

$K_X(t, t + \tau) = k_X(\tau)$, то есть это функция одного аргумента.

Свойство:

1) $k_x(\tau) = k_x(-\tau)$

$K_X(t, t') = K_X(t', t)$, следовательно, если $\tau = t' - t$, $k_x(\tau) = k_x(-\tau)$.

2) $K_X(0) = D_X = \text{const}$

Спектральное разложение стационарного случайного процесса

а) Пусть стационарный случайный процесс наблюдается как на конечном интервале $(0, T)$. Корреляционная функция $K_X(\tau)$ (АКФ) чётна. $k_x(\tau) = k_x(-\tau)$, следовательно она представима в виде ряда Фурье.

$$k_X(\tau) = \frac{D_0}{2} + \sum_{i=1}^{\infty} D_n \cos \omega_n \tau.$$

$$D_n = \frac{2}{T} \int_0^T k_X(\tau) \cos(\omega_n \tau) d\tau$$

$$D_X = \sum_{i=1}^{\infty} D_n.$$

Спектральный стационарный случайный процесс

б) бесконечный интервал

Стационарный случайный процесс на $(0, \infty)$. Так как $k_X(\tau)$ чётна, её можно представить в виде интеграла Фурье:

$$k_X(\tau) = \int_0^{\infty} \hat{S}_X(\omega) \cos \omega \tau d\omega, \text{ где } \hat{S}_X(\omega) = \frac{2}{\pi} \int_0^{\infty} k_X(\tau) \cos \omega \tau d\tau - \text{спектральная}$$

плотность (спектральная плотность дисперсии). $D_X = \int_0^{\infty} S_X(\omega) d\omega$ - то есть

площадь под кривой.

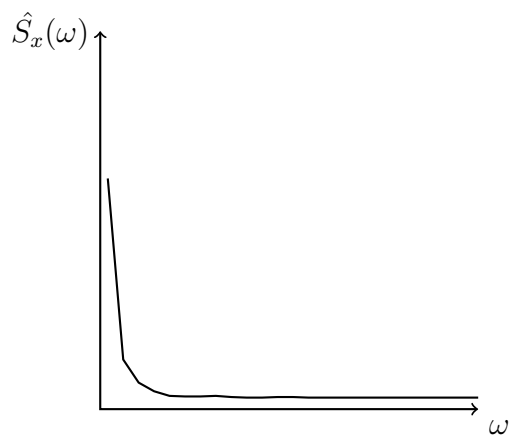
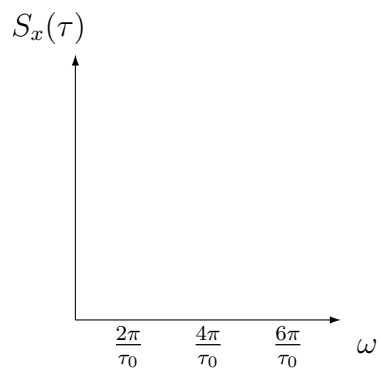
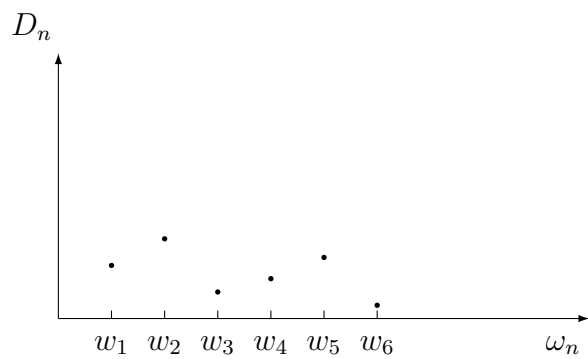
Пример

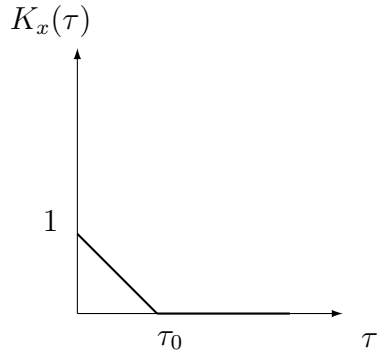
$$k_X(\tau) = \begin{cases} 1 - \frac{\tau}{\tau_0}, & 0 < \tau < \tau_0 \\ 0, & \tau > \tau_0 \end{cases}$$

$$\hat{S}_X(\omega) = \frac{2}{\pi} \int_0^{\infty} k_X(\tau) \cos \omega \tau d\tau = \frac{2}{\pi} \int_0^{\tau_0} (1 - \frac{\tau}{\tau_0}) \cos \omega \tau d\tau = \frac{2}{\pi \tau_0 \omega^2} (1 - \cos \omega \tau)$$

в) Спектральная плотность в канонической форме:

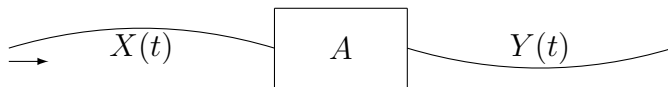
$$K_X(\tau) = \int_{-\infty}^{\infty} S_X(\omega) e^{i\omega \tau} d\omega, \text{ где } S_X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} k_X(\tau) e^{-i\omega \tau} d\tau, D_X = \int_{-\infty}^{\infty} S_X(\omega) d\omega.$$





8.3 Преобразование случайной функции линейной системой

Пусть есть прибор (динамическая система), на вход которой поступает непрерывно сигнал $x(t)$ (воздействие), а на выходе получаем непрерывно сигнал $y(t)$ (реакция). Система A осуществляется преобразование: $y(t) = A(x(t))$. (например, умножает на константу – усилитель, интегрирует или дифференцирует сигнал – интегратор или дифференциатор цепочки, и тому подобное). Реально сигнал $x(t)$ присутствует всегда вместе со случайной ошибкой, то есть надо рассматривать преобразование вместе со случайной ошибкой, то есть надо рассмотреть преобразование случайной функции $X(t)$ в случайную функцию $Y(t)$. $Y(t) = A(X(t))$, A – оператор динамической системы. Мы рассматриваем линейные системы (в которых оператор A линеен): $A(x_1(t) + x_2(t)) = A(x_1(t)) + A(x_2(t))$, $A(cx_1(t)) = c \cdot A(x_1(t))$



Пример

$$1) y(t) = Ax(t) = \frac{dx(t)}{dt}$$

$$2) y(t) = \int_0^t x(t) dt$$

Чаще всего в технических задачах оказывается связь $y(t)$ с $x(t)$ через решение линейного дифференциального уравнения.

$$a_n \frac{d^n y(t)}{dt^n} + a_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \dots + a_1 \frac{dy(t)}{dt} + a_0 y(t) = \\ = b_m \frac{d^m x(t)}{dt^m} + b_{m-1} \frac{d^{m-1} y(t)}{dt^{m-1}} + \dots + b_1 \frac{dy(t)}{dt} + b_0 y(t)$$

Если параметры линейной системы не зависят от t , то она называется стационарной. Рассмотрим преобразование стационарной линейной системы, описываемой линейным дифференциальным уравнением с постоянными коэффициентами.

Решение $y(t)$ это общее решение однородного (свободные колебания) + частичное решение неоднородного (вынужденные колебания). На практике чаще всего встречаются устойчивые системы, в которых свободные колебания затухают при t стремящимся к ∞ . Поэтому ограничимся рассмотрением только вынужденных колебаний.

Пусть на входе $x(t) = Ue^{i\omega t}$. Ищем $y(t)$ в виде $y(t) = U\Phi(i\omega)e^{i\omega t}$. Подставляем в уравнение, получаем: $\Phi(\omega)(a_n(i\omega)^n + a_{n-1}(i\omega)^{n-1} + \dots + a_1 i\omega + a_0)e^{i\omega t} = (b_m(i\omega)^m + b_{m-1}(i\omega)^{m-1} + \dots + b_1 i\omega + b_0)e^{i\omega t}$.

Сокращаем на $e^{i\omega t}$ и вводим обозначение $A_n(i\omega)$, $B_m(i\omega)$.

$$\Phi(i\omega) = \frac{B_m(i\omega)}{A_n(i\omega)}$$

Математическое ожидание случайной величины $X(t)$ – это гармонические колебания нулевой частоты, следовательно $\Phi(0) = \frac{B_m(0)}{A_n(0)} = \frac{b_0}{a_0}$, следовательно $m_Y = \frac{b_0}{a_0} m_X$ – математическое ожидание на выходе.

Рассмотрим $\dot{X}(t) = X(t) - m_X$. Пусть $\dot{X}(t)$ задана на $(0, T)$. Представим её в виде спектрального разложения: $\dot{X}(t) = \sum_{k=-\infty}^{\infty} U_k e^{i\omega_k t}$. U_k – некоррелируемая

случайная величина, дисперсии которой образуют спектр $X(t)$. $X_k(t) = U_k e^{i\omega_k t}$, следовательно $Y_k(t) = U_k \Phi(i\omega_k) e^{i\omega_k t}$. По принципу суперпозиции (система

линейна): $\dot{Y}(t) = \sum_{k=-\infty}^{\infty} U_k \Phi(i\omega_k) e^{i\omega_k t} = \sum_{k=-\infty}^{\infty} W_k e^{i\omega_k t}$

Найдем спектр $Y(t)$, то есть дисперсии W_k (W_k некоррелируют и имеют нулевое математическое ожидание).

$$D(W_k) = M(|U_k \Phi(i\omega_k)|^2) = M(|U_k|^2 |\Phi(i\omega_k)|^2) = |\Phi(i\omega_k)|^2 M(|U_k|^2), \text{ так как } M(|U_k|^2) = D_k, \text{ то } D(W_k) = |\Phi(i\omega_k)|^2 D_k.$$

То есть при преобразовании каждая из ординат спектра умножается на квадрат модуля частотной характеристики. Если сигнал задан на бесконечном промежутке, то аналогично получается закон преобразования спектральной плотности: $S_Y(\omega) = |\Phi(i\omega)|^2 S_X(\omega)$

Итак, описываем прохождение стационарного случайного сигнала $X(t)$ с m_x и АКФ $K_x(\tau)$ так:

$$1) m_Y = \frac{b_0}{a_0} m_X$$

$$2) \text{ по } k_X(\tau) \text{ находим } S_X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} k_X(\tau) e^{i\omega\tau} d\tau$$

$$3) \text{ находим частотную характеристику системы } |\Phi(i\omega)|^2 = \frac{|B_m(i\omega)|^2}{|A_n(i\omega)|^2}$$

$$4) S_Y(\omega) = |\Phi(i\omega)|^2 S_X(\omega)$$

$$5) k_Y(\tau) = \int_{-\infty}^{\infty} S_Y(\omega) e^{i\omega\tau} d\omega$$

Замечание

$$\text{Если нужна дисперсия, } D_Y = k_Y(0) \int_{-\infty}^{\infty} S_Y(\omega) d\omega.$$

Пример

Пусть работа линейной системы описывается $\frac{dy}{dt} + y = x(t)$
 $X(t)$ имеет математическое ожидание m_X и $k_X(\tau) = D_X e^{-\alpha|\tau|}$ ($1 > \alpha > 0$). Найти m_Y и $k_Y(\tau)$

$$S_X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} k_X(\tau) e^{-i\omega\tau} d\tau = \frac{D_X \alpha}{\pi(\alpha^2 + \omega^2)}$$

$$\Phi(i\omega) = \frac{1}{i\omega + 1}, |\Phi(i\omega)|^2 = \frac{1}{\omega^2 + 1}$$

$$S_Y(\omega) = |\Phi(i\omega S)|^2 S_X(\omega) = \frac{D_X}{\pi} \frac{\alpha}{(\alpha^2 + \omega^2)(\omega^2 + 1)}$$

$$k_Y(\tau) = \frac{D_X}{\pi} \int_{-\infty}^{\infty} \frac{\alpha e^{i\omega\tau}}{(\alpha^2 + \omega^2)(\omega^2 + 1)} d\omega =$$

$$= \begin{cases} \frac{D_X}{\pi} \cdot 2\pi i \sum_{\substack{\text{(вычисляется в верхней полуплоскости)}}} \text{Res} \frac{\alpha e^{i\omega\tau}}{(\alpha^2 + \omega^2)(\omega^2 + 1)}, & \tau > 0 \\ -\frac{D_X}{\pi} \cdot 2\pi i \sum_{\substack{\text{(вычисляется в нижней полуплоскости)}}} \text{Res} \frac{\alpha e^{i\omega\tau}}{(\alpha^2 + \omega^2)(\omega^2 + 1)}, & \tau < 0 \end{cases}$$

$$= \begin{cases} 2iD_X \alpha \left(\frac{e^{i \cdot i\alpha\tau}}{(i\alpha + i\alpha)((i\alpha)^2 + 1)} + \frac{e^{ii\tau}}{(\alpha^2 + i^2)(i+i)} \right), & \tau > 0 \\ -2iD_X \alpha \left(\frac{e^{i \cdot (-i\alpha)\tau}}{(-i\alpha - i\alpha)((-i\alpha)^2 + 1)} + \frac{e^{i(-i)\tau}}{(\alpha^2 + (-i)^2)(-i-i)} \right), & \tau < 0 \end{cases}$$

$$\begin{aligned}
&= \begin{cases} D_X \alpha \left(\frac{e^{-\alpha\tau}}{\alpha(1-\alpha^2)} + \frac{e^{-\tau}}{(\alpha^2-1)} \right), & \tau > 0 \\ D_X \alpha \left(\frac{e^{\alpha\tau}}{(\alpha)(1-\alpha^2)} + \frac{e^{\tau}}{\alpha^2-1} \right), & \tau < 0 \end{cases} \\
&= D_X \frac{e^{-\alpha|\tau|} - \alpha e^{-|\tau|}}{(1-\alpha^2)}
\end{aligned}$$

При $\tau = 0$ интеграл вычисляется точно так же (только на основании не леммы Жордана, а убывании рациональных подынтегральных функций быстрее чем $\frac{1}{r}$ или разложении на простейшие дроби). При $\tau = 0$, $D_Y = \frac{D_X}{1+\alpha}$

Приложение

Значение нормальной функции распределения.

$$\Phi^*(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2} dt$$

x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ
-0.00	0.50000	40	-0.30	0.3821	38	-0.60	0.2743	34
-0.01	0.4960	40	-0.31	0.3783	38	-0.61	0.2709	33
-0.02	0.4920	40	-0.32	0.3745	38	-0.62	0.2667	33
-0.03	0.4880	40	-0.33	0.3707	38	-0.63	0.2643	32
-0.04	0.4840	39	-0.34	0.3669	37	-0.64	0.2611	33
-0.05	0.4801	40	-0.35	0.3632	38	-0.65	0.2578	32
-0.06	0.4762	40	-0.36	0.3594	37	-0.66	0.2546	32
-0.07	0.4721	40	-0.37	0.3557	37	-0.67	0.2514	31
-0.08	0.4681	40	-0.38	0.3520	37	-0.68	0.2483	32
-0.09	0.4641	39	-0.39	0.3483	37	-0.69	0.2451	31
-0.10	0.4602	40	-0.40	0.3446	37	-0.70	0.2420	31
-0.11	0.4562	40	-0.41	0.3409	37	-0.71	0.2389	31
-0.12	0.4522	39	-0.42	0.3372	36	-0.72	0.2358	31
-0.13	0.4483	40	-0.43	0.3336	36	-0.73	0.2327	30
-0.14	0.4443	39	-0.44	0.3300	36	-0.74	0.2297	31
-0.15	0.4404	40	-0.45	0.3264	36	-0.75	0.2266	30
-0.16	0.4364	39	-0.46	0.3228	36	-0.76	0.2236	30
-0.17	0.4325	39	-0.47	0.3192	36	-0.77	0.2206	29
-0.18	0.4286	39	-0.48	0.3156	35	-0.78	0.2177	29
-0.19	0.4247	40	-0.49	0.3121	36	-0.79	0.2148	29
-0.20	0.4207	39	-0.50	0.3085	35	-0.80	0.2119	29
-0.21	0.4168	39	-0.51	0.3050	35	-0.81	0.2090	29
-0.22	0.4129	39	-0.52	0.3015	34	-0.82	0.2061	28
-0.23	0.4090	38	-0.53	0.2981	35	-0.83	0.2033	28
-0.24	0.4052	39	-0.54	0.2946	34	-0.84	0.2005	28
-0.25	0.4013	39	-0.55	0.2912	35	-0.85	0.1977	28
-0.26	0.3974	38	-0.56	0.2877	34	-0.86	0.1949	27
-0.27	0.3936	39	-0.57	0.2843	33	-0.87	0.1922	28
-0.28	0.3897	38	-0.58	0.2810	34	-0.88	0.1894	27
-0.29	0.3859	39	-0.59	0.2776	33	-0.89	0.1867	26

x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ
-0.90	0.1841	27	-1.30	0.0968	17	-1.70	0.0446	10
-0.91	0.1814	26	-1.31	0.051	17	-1.71	0.0436	9
-0.92	0.1788	26	-1.32	0.0934	16	-1.72	0.0427	9
-0.93	0.1762	26	-1.33	0.0918	17	-1.73	0.0418	9
-0.94	0.1736	25	-1.34	0.0901	16	-1.74	0.0409	8
-0.95	0.1711	25	-1.35	0.0885	16	-1.75	0.0401	9
-0.96	0.1685	25	-1.36	0.0869	16	-1.76	0.0392	8
-0.97	0.1660	25	-1.37	0.0853	15	-1.77	0.0384	9
-0.98	0.1635	24	-1.38	0.0838	15	-1.78	0.0375	8
-0.99	0.1611	24	-1.39	0.0823	15	-1.79	0.0367	8
-1.00	0.1587	24	-1.40	0.0808	15	-1.80	0.0359	8
-1.01	0.1563	24	-1.41	0.0793	15	-1.81	0.0351	7
-1.02	0.1539	24	-1.42	0.0078	14	-1.82	0.0344	8
-1.03	0.1515	23	-1.43	0.0764	15	-1.83	0.0336	7
-1.04	0.1492	23	-1.44	0.0749	14	-1.84	0.0329	7
-1.05	0.1469	23	-1.45	0.0735	14	-1.85	0.0322	8
-1.06	0.1446	23	-1.46	0.0721	13	-1.86	0.0314	7
-1.07	0.1423	22	-1.47	0.0708	14	-1.87	0.0307	6
-1.08	0.1401	22	-1.48	0.0694	1	-1.88	0.0301	7
-1.09	0.1379	22	-1.49	0.0681	13	-1.89	0.0294	6
-1.10	0.1357	22	-1.50	0.0668	13	-1.90	0.0288	7
-1.11	0.1335	21	-1.51	0.0655	12	-1.91	0.0281	7
-1.12	0.1314	22	-1.52	0.0643	13	-1.92	0.0274	6
-1.13	0.1292	21	-1.53	0.0630	12	-1.93	0.0268	6
-1.14	0.1271	20	-1.54	0.0618	12	-1.94	0.0262	6
-1.15	0.1251	21	-1.55	0.0606	12	-1.95	0.0256	6
-1.16	0.1230	20	-1.56	0.0594	12	-1.96	0.0250	6
-1.17	0.1210	20	-1.57	0.0582	11	-1.97	0.0244	5
-1.18	0.1190	20	-1.58	0.0571	12	-1.98	0.0239	6
-1.19	0.1170	19	-1.59	0.0559	11	-1.99	0.0233	5
-1.20	0.1151	20	-1.60	0.0548	11	-2.00	0.0228	49
-1.21	0.1131	19	-1.61	0.0537	11	-2.10	0.0179	40
-1.22	0.1112	19	-1.62	0.0526	10	-2.20	0.0139	32
-1.23	0.1093	18	-1.63	0.0516	11	-2.30	0.0107	25
-1.24	0.1075	19	-1.64	0.0505	10	-2.40	0.0082	20

x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ
-1.25	0.1056	18	-1.65	0.0495	10	-2.50	0.0062	15
-1.26	0.1038	18	-1.66	0.0485	10	-2.60	0.0047	12
-1.27	0.1020	17	-1.67	0.0475	10	-2.70	0.0035	9
-1.28	0.1003	18	-1.68	0.0465	10	-2.80	0.0026	7
-1.29	0.0985	17	-1.69	0.0455	9	-2.90	0.0019	5

x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ
-3.00	0.0014	4	0.30	0.6179	38	0.70	0.7580	31
-3.10	0.0010	3	0.31	0.6217	38	0.71	0.7611	31
-3.20	0.0007	2	0.32	0.6255	38	0.72	0.7642	31
-3.30	0.0005	2	0.33	0.6293	38	0.73	0.7673	30
-3.40	0.0003	1	0.34	0.6331	37	0.74	0.7703	31
-3.50	0.0002	0	0.35	0.6368	38	0.75	0.7734	30
-3.60	0.0002	1	0.36	0.6406	37	0.76	0.7764	30
-3.70	0.0001	0	0.37	0.6443	37	0.77	0.7794	29
-3.80	0.0001	1	0.38	0.6480	37	0.78	0.7823	29
-3.90	0.0000		0.39	0.6517	37	0.79	0.7852	29
0.00	0.5000	40	0.40	0.6554	37	0.80	0.7881	29
0.01	0.5040	40	0.40	0.6591	37	0.81	0.7910	29
0.02	0.5080	40	0.40	0.6628	36	0.82	0.7939	28
0.03	0.5120	40	0.40	0.6664	36	0.83	0.7967	28
0.04	0.5160	39	0.40	0.6700	36	0.84	0.7995	28
0.05	0.5199	40	0.40	0.6736	36	0.85	0.8023	28
0.06	0.5239	40	0.40	0.6772	36	0.86	0.8051	27
0.07	0.5279	40	0.40	0.6808	36	0.87	0.8078	28
0.08	0.5319	40	0.40	0.6844	35	0.88	0.8106	27
0.09	0.5359	40	0.39	0.6879	36	0.89	0.8133	26
0.10	0.5398	40	0.50	0.6915	35	0.90	0.8159	27
0.11	0.5438	40	0.51	0.6950	35	0.91	0.8186	26
0.12	0.5478	39	0.52	0.6985	34	0.92	0.8212	26
0.13	0.5517	40	0.53	0.7019	35	0.93	0.8238	26
0.14	0.5557	39	0.54	0.7054	34	0.94	0.8264	25
0.15	0.5596	40	0.55	0.7088	35	0.95	0.8289	26
0.16	0.5636	39	0.56	0.7123	34	0.96	0.8315	25
0.17	0.5675	39	0.57	0.7157	33	0.97	0.8340	25
0.18	0.5714	39	0.58	0.7190	34	0.98	0.8365	24
0.19	0.5753	40	0.59	0.7224	33	0.99	0.8389	24
0.20	0.5793	39	0.60	0.7257	34	1.00	0.8413	24
0.21	0.5832	39	0.61	0.7291	33	1.01	0.8437	24
0.22	0.5871	39	0.62	0.7324	33	1.02	0.8461	24
0.23	0.5910	39	0.63	0.7357	32	1.03	0.8485	23
0.24	0.5948	39	0.64	0.7389	33	1.04	0.8508	23

x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ
0.25	0.5987	39	0.65	0.7422	32	1.05	0.8531	23
0.26	0.6026	38	0.66	0.7454	32	1.06	0.8554	23
0.27	0.6064	39	0.67	0.7486	31	1.07	0.8577	22
0.28	0.6103	38	0.68	0.7517	32	1.08	0.8599	22
0.29	0.6141	38	0.69	0.7549	31	1.09	0.8621	22
1.10	0.8643	22	1.50	0.9332	13	1.90	0.9713	6
1.11	0.8665	21	1.51	0.9345	12	1.91	0.9719	7
1.12	0.8686	22	1.52	0.9357	13	1.92	0.9726	6
1.13	0.8708	21	1.53	0.9370	12	1.93	0.9732	6
1.14	0.8729	20	1.54	0.9382	12	1.94	0.9738	6
1.15	0.8749	21	1.54	0.9394	12	1.95	0.9744	6
1.16	0.8770	20	1.56	0.9406	12	1.96	0.9750	6
1.17	0.8790	20	1.57	0.9418	11	1.97	0.9756	5
1.18	0.8810	20	1.58	0.9429	12	1.98	0.9761	6
1.19	0.8830	19	1.59	0.9441	11	1.99	0.9767	5

x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ
1.20	0.8849	20	1.60	0.9452	11	2.00	0.9772	49
1.21	0.8869	19	1.61	0.9463	11	2.10	0.9821	40
1.22	0.8888	19	1.62	0.9474	10	2.20	0.9861	32
1.23	0.8907	18	1.63	0.9484	11	2.30	0.9893	25
1.24	0.8925	19	1.64	0.9495	10	2.40	0.9918	20
1.25	0.8944	18	1.65	0.9505	10	2.50	0.9938	15
1.26	0.8962	18	1.66	0.9515	10	2.60	0.9953	12
1.27	0.8980	17	1.67	0.9525	10	2.70	0.9965	9
1.28	0.8997	18	1.68	0.9535	10	2.80	0.9974	7
1.29	0.9015	17	1.69	0.9545	9	2.90	0.9981	5
1.30	0.9032	17	1.70	0.9554	10	3.00	0.9986	4
1.31	0.9049	17	1.71	0.9564	9	3.10	0.9990	3
1.32	0.9066	16	1.72	0.9573	9	3.20	0.9993	2
1.33	0.9082	17	1.73	0.9582	9	3.30	0.9995	2
1.34	0.9099	16	1.74	0.9591	8	3.40	0.9997	1
1.35	0.9115	16	1.75	0.9599	9	3.50	0.9998	0
1.36	0.9131	16	1.76	0.9608	8	3.60	0.9998	1
1.37	0.9147	15	1.77	0.9616	9	3.70	0.9999	0
1.38	0.9162	15	1.78	0.9625	8	3.80	0.9999	1
1.39	0.9177	15	1.79	0.9633	8	3.90	1.0000	
1.40	0.9192	15	1.80	0.9641	8			
1.41	0.9207	15	1.81	0.9649	7			
1.42	0.9222	14	1.82	0.9656	8			
1.43	0.9236	15	1.83	0.9664	7			
1.44	0.9251	14	1.84	0.9671	7			
1.45	0.9265	14	1.85	0.9678	8			
1.46	0.9279	13	1.86	0.9686	7			
1.47	0.9292	14	1.87	0.9693	6			
1.48	0.9306	13	1.88	0.9699	7			
1.49	0.9319	13	1.89	0.9706	7			

Значение χ^2 в зависимости от r и p

$r \backslash p$	0.99	0.98	0.95	0.90	0.80	0.70	0.50
1	0.000	0.001	0.004	0.016	0.064	0.148	0.455
2	0.020	0.040	0.103	0.211	0.466	0.713	1.386
3	0.115	0.185	0.352	0.584	1.005	1.424	2.370
4	0.297	0.429	0.711	1.064	1.649	2.200	3.360
5	0.554	0.752	1.145	1.610	2.34	3.000	4.350
6	0.872	1.134	1.635	2.200	3.070	3.830	5.350
7	1.239	1.564	2.170	2.830	3.820	4.670	6.350
8	1.646	2.030	2.730	3.490	4.590	5.530	7.340
9	2.090	2.530	3.320	4.170	5.380	6.390	8.340
10	2.560	3.060	3.940	4.860	6.180	7.270	9.340
11	3.500	3.610	4.580	5.580	6.990	8.150	10.34
12	3.570	4.180	5.230	6.300	7.810	9.030	11.34
13	4.110	5.890	7.040	8.630	9.930	12.34	15.12
14	4.660	5.370	6.570	7.790	9.470	10.82	13.34
15	5.230	5.980	7.260	8.550	10.31	11.72	14.34
16	5.810	6.610	7.960	9.310	11.15	12.62	15.34
17	6.410	7.260	8.670	10.08	12.00	13.53	16.34
18	7.020	7.910	9.390	10.86	12.86	14.44	17.34
19	7.630	8.570	10.11	11.65	13.72	15.35	18.34
20	8.260	9.240	10.85	12.44	14.58	16.27	19.34
21	8.900	9.920	11.59	13.24	15.44	17.18	20.3
22	9.540	10.60	12.34	14.04	16.31	18.10	21.3
23	10.20	11.29	13.09	14.85	17.19	19.02	22.3
24	10.86	11.99	13.85	15.66	18.06	19.94	23.3
25	11.52	12.70	14.61	16.47	18.94	20.90	24.3
26	12.20	13.41	15.38	17.29	19.82	21.80	25.3
27	12.88	14.12	16.15	18.11	20.70	22.70	26.3
28	13.56	14.84	16.93	18.94	21.60	23.60	27.3
29	14.26	15.57	17.71	19.77	22.50	24.60	28.3
30	14.95	16.31	18.49	20.60	23.49	25.50	29.3

Значение χ^2 в зависимости от r и p

$r \backslash p$	0.30	0.20	0.10	0.05	0.02	0.01	0.001
1	1.074	1.642	2.71	3.84	5.41	6.64	10.83
2	2.410	3.22	4.60	5.99	7.82	9.21	13.82
3	3.660	4.62	6.25	7.82	9.84	11.34	16.27
4	4.880	5.99	7.78	9.49	11.67	13.28	18.46
5	6.060	7.29	9.24	11.07	13.39	15.09	20.5
6	7.230	8.56	11.03	12.59	15.03	16.81	22.5
7	8.380	9.80	12.24	14.07	16.62	18.48	24.3
8	9.520	11.03	13.44	15.51	18.17	20.1	26.1
9	10.66	12.24	14.63	16.92	19.68	21.7	27.9
10	11.78	13.44	15.99	18.31	21.2	23.2	29.6
11	12.90	14.63	17.28	19.68	22.6	24.7	31.3
12	14.01	15.81	18.55	21.0	24.1	26.2	32.9
13	15.12	16.98	19.81	22.4	25.5	27.7	34.6
14	16.22	18.15	21.1	23.7	26.9	29.1	36.1
15	17.32	19.31	22.3	25.0	28.3	30.6	37.7
16	18.42	20.5	23.5	26.3	29.6	32.0	39.3
17	19.51	21.6	24.8	27.6	31.0	33.4	40.8
18	20.6	22.8	26.0	28.9	32.3	34.8	42.3
19	21.7	23.9	27.2	30.1	33.7	36.2	43.8
20	22.9	25.0	28.4	31.4	35.0	37.6	45.3
21	23.9	26.2	29.6	32.7	36.3	38.9	46.8
22	24.9	27.3	30.8	33.9	37.7	40.3	48.3
23	26.0	28.4	32.0	35.2	39.0	41.6	49.7
24	27.1	29.6	33.2	36.4	40.3	43.0	51.2
25	28.2	30.7	34.4	37.7	41.7	44.3	52.6
26	29.2	31.8	35.6	38.9	42.9	45.6	54.1
27	30.3	32.9	36.7	40.1	44.1	47.0	55.5
28	31.4	34.0	37.9	41.3	45.4	48.3	56.9
29	32.5	35.1	39.1	42.6	46.7	49.6	58.3
30	33.5	36.2	40.3	43.8	48.0	50.9	59.7

$$\Phi^*(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2} dt$$

$$\varphi(x) = \Phi(x)' = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

x	$\Phi(x)$	$\varphi(x)$	x	$\Phi(x)$	$\varphi(x)$
0.0	0.50000	0.39894	2.1	0.98214	0.04398
0.1	0.53983	0.39695	2.2	0.98610	0.03547
0.2	0.57926	0.39104	2.3	0.98928	0.02833
0.3	0.61791	0.38139	2.4	0.99180	0.02239
0.4	0.65542	0.36827	2.5	0.99379	0.01753
0.5	0.69145	0.35207	2.6	0.99534	0.01358
0.6	0.72575	0.33322	2.7	0.99635	0.01042
0.7	0.75804	0.31225	2.8	0.99744	0.00792
0.8	0.78814	0.28969	2.9	0.99813	0.00595
0.9	0.81594	0.26609	3.0	0.99865	0.00443
1.0	0.84134	0.24197	3.1	0.99903	0.00327
1.1	0.86433	0.21785	3.2	0.99931	0.00238
1.2	0.88493	0.19419	3.3	0.99952	0.00172
1.3	0.90320	0.17137	3.4	0.99966	0.00123
1.4	0.91924	0.14937	3.5	0.99977	0.00087
1.5	0.93319	0.12925	3.6	0.99984	0.00061
1.6	0.94520	0.11092	3.7	0.99989	0.00042
1.7	0.95534	0.09405	3.8	0.99993	0.00029
1.8	0.96107	0.07895	3.9	0.99995	0.00020
1.9	0.97128	0.06569	4.0	0.99997	0.00013
2.0	0.97725	0.05399			