

**СПЕЦГЛАВЫ
МАТЕМАТИКИ: ТЕОРИЯ
ВЕРОЯТНОСТЕЙ;
МАТЕМАТИЧЕСКАЯ
СТАТИСТИКА;
ИНТЕГРАЛЬНЫЕ
ПРЕОБРАЗОВАНИЯ**

Лекции: Кориков

Анатолий Михайлович

Пр. занятия: Ефремов

Александр Александрович

Томск, 2015

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Литература:

- Феллер В. Введение в теорию вероятностей и её приложения. В 2-х томах. Т. 1: М.: Мир, 1984. – 528 с.
- Вентцель Е.С. Теория вероятностей: Учебник для вузов/ 10-е изд., стереотип. – М.: Высшая школа, 2005. – 576 с.
- Письменный Д.Т. Конспект лекций по теории вероятностей, математической статистике и случайным процессам. – М.: Айрис-Пресс, 2006. – 287 с.
- Гмурман В.Е. Теория вероятностей и математическая статистика/ Учебное пособие для вузов – М.: Высшая школа, 2003. – 480с. (34 экз. в библиотеке ТУСУР)
- Пугачев В.С. Теория вероятностей и математическая статистика. – М.: Физматлит, 2002. – 496с.
- Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие для вузов/ 7-е изд., доп. – М.: Высшая школа, 2003 – 406 с.
- Вентцель Е.С. Овчаров Л.А. Задачи и упражнения по теории вероятностей: Учебное пособие для вузов/ - 3-е изд., стереотип. – М.: Академия. –2005. – 439 с.
- Сборник задач по теории вероятностей, математической статистике и теории случайных функций (под ред. А.А. Свешникова), – М.: Наука, 1970. – 656с.

1. СОБЫТИЯ, АЛГЕБРА СОБЫТИЙ

Случайные события

Под событием в теории вероятностей понимается всякий факт, который в результате опыта может произойти или не произойти.

Примеры событий:

A – появился герб при бросании монеты;

B – появление трех гербов при трехкратном бросании монеты;

C – попадание в цель при выстреле;

D – появление туза при извлечении карты из колоды; и т.д.

Рассматривая вышеперечисленные события, мы видим, что каждое из них обладает какой-то степенью возможности: одни – большей, другие – меньшей. Причем, для некоторых событий мы сразу же можем решить, какое из них более, а какое менее возможно. Чтобы количественно сравнить между собой события по степени их возможности, очевидно нужно с каждым событием связать определенное число, которое тем больше, чем более возможно событие. Такое число мы называем **вероятностью события**.

1. СОБЫТИЯ, АЛГЕБРА СОБЫТИЙ

Случайные события

Рассмотрим множество событий E , которые можно наблюдать в некотором эксперименте.

Выделим, прежде всего, два специальных события:

- достоверное событие (U), которое обязательно происходит в эксперименте,
- невозможное событие (V), которое не может произойти в эксперименте никогда.

Для каждого события A из E введем **противоположное событие** \bar{A} , которое состоит в том, что событие A не произошло.

Операции алгебры событий совпадают с операциями над множествами и, так же как и в теории множеств, хорошо иллюстрируются кругами Эйлера.

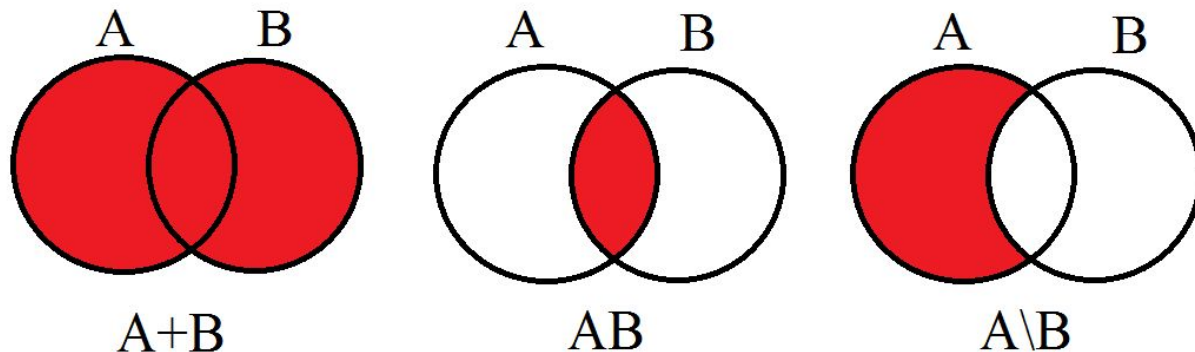
1. СОБЫТИЯ, АЛГЕБРА СОБЫТИЙ

Случайные события

Событие $A+B$ ($A \cup B$), заключающееся в том, что из двух событий A и B происходит по крайней мере одно (либо A , либо B , либо A и B вместе), называется суммой (или объединением) событий A и B .

Событие $A \cdot B$ ($A \cap B$), заключающееся в том, что события A и B происходят одновременно, называется произведением (или пересечением) событий A и B .

Событие $A-B$ ($A \setminus B$) называется разностью событий A и B ; оно заключается в том, что происходит A и не происходит B .



1. СОБЫТИЯ, АЛГЕБРА СОБЫТИЙ

Пространство элементарных событий

Предположим, что среди всех возможных событий A , которые в данном опыте по воле случая происходят или не происходят, можно выделить совокупность так называемых элементарных событий или элементарных исходов, обладающих следующими свойствами:

- во-первых, все они взаимоисключают друг друга, т.е. являются непересекающимися;
- во-вторых, в результате данного опыта обязательно происходит одно из этих элементарных событий;
- в-третьих, каково бы ни было событие A , по наступившему элементарному исходу всегда можно судить о том, происходит или не происходит это событие.

Элементарные исходы будем обозначать буквой e , а их совокупность **E** называется пространством элементарных событий.

1. СОБЫТИЯ, АЛГЕБРА СОБЫТИЙ

Пространство элементарных событий

Достоверное событие U , наступающее в результате любого из элементарных исходов e , при отождествлении событий множеством совпадает с пространством, т.е. $U = E$.

Невозможное событие V , не наступающее ни при каком элементарном исходе e , совпадает с пустым множеством и обозначается $V = \emptyset$.

Два события A и B несовместимы (или несовместны), если $A \cap B = \emptyset$ (т.е. событие невозможно).

События A_1, A_2, \dots, A_n образуют полную группу событий, если они

• попарно несовместны

$$A_i \cap A_j = \emptyset, \quad i \neq j;$$

• и

$$A_1 \cap A_2 \cap \dots \cap A_n = \bigcap_{i=1}^n A_i = E$$

1. СОБЫТИЯ, АЛГЕБРА СОБЫТИЙ

Статистическое определение вероятности

Испытанием называется эксперимент, который можно (хотя бы принципиально) провести в одинаковых условиях любое число раз.

Простейший результат испытания называется элементарным событием или исходом. При испытании неизбежно наступает какой-то исход и только один.

Если событие может привести к n различным равновозможным исходам и если в m случаях появится признак A , то относительная частота (частость) события A обозначается $r(A)$ и равна отношению m к n :

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

статистическое (комбинаторное)
определение вероятности

1. СОБЫТИЯ, АЛГЕБРА СОБЫТИЙ

Статистическое определение вероятности

Событие A , для которого относительная частота $r(A)$ при достаточно больших n мало отличается от некоторого фиксированного числа, не зависящего от серии проводимых испытаний, называется статически устойчивым.

Вероятностью статически устойчивого случайного события A называется число $P(A)$, около которого группируются относительные частоты этого события в длинных сериях независимых испытаний:

$$Pr(A) = r_n(A) \rightarrow \infty.$$

- 1) $0 \leq P(A) \leq 1$
- 2) $P(\emptyset) = 0$
- 3) $P(E) = 1$

1. СОБЫТИЯ, АЛГЕБРА СОБЫТИЙ

Геометрическое определение вероятности

Геометрическое определение вероятности применяется в тех случаях, когда множество всех исходов (возможных и благоприятных) бесконечно и эти исходы определяются одним или несколькими числовыми параметрами.

Геометрической вероятностью события A называется отношение меры области, благоприятствующей появлению события A , к мере всей области:

$$P(A) = \frac{mes(g)}{mes(G)}$$

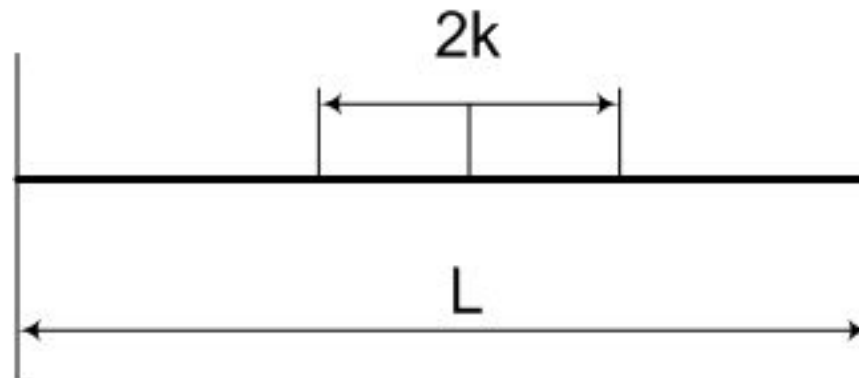
1. СОБЫТИЯ, АЛГЕБРА СОБЫТИЙ

Геометрическое определение вероятности

Пример 1

Предположим, что на отрезок длиной L действительной прямой наугад бросается точка, которую обозначим x . Какова вероятность того, что она отклонится не дальше, чем на расстояние k , от середины указанного отрезка?

Решение. Здесь имеется бесконечное множество возможных исходов: ведь точка x может попасть в любую точку рассматриваемого отрезка длиной L . Кроме того, условия опыта таковы, что x с одинаковой вероятностью может оказаться в любой точке этого отрезка, расположенного на оси абсцисс.



1. СОБЫТИЯ, АЛГЕБРА СОБЫТИЙ

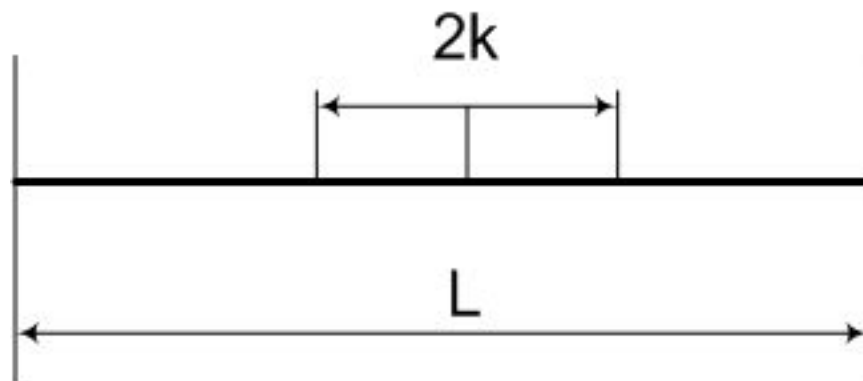
Геометрическое определение вероятности

Пример 1 (продолжение)

Событие A – точка x находится от середины отрезка на расстоянии не больше k , наступает в результате попадания в любую точку, отстоящую от середины не далее, чем на величину k .

«Доля» таких точек на всем отрезке может быть определена как отношение $L(A) / L$, где L – длина всего рассматриваемого отрезка. $L(A) = 2k$ – длина отрезка, попадание в который влечет за собой наступление события A . Таким образом, искомая вероятность $P(A)$ равна:

$$P(A) = \frac{L(A)}{L} = \begin{cases} \frac{2k}{L} & \text{если } k < \frac{L}{2} \\ 1 & \text{если } k \geq \frac{L}{2} \end{cases}$$

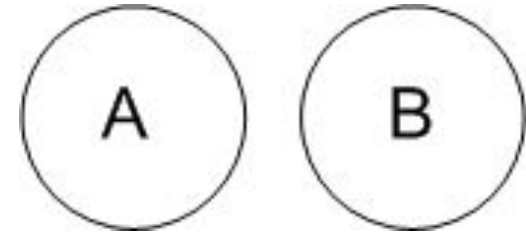


2. ВЕРОЯТНОСТЬ

Теорема сложения вероятностей несовместных событий

Вероятность суммы конечного числа несовместных событий A_1, A_2, \dots, A_n равна сумме вероятностей этих событий:

$$P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i).$$



Если события A_1, A_2, \dots, A_n образуют полную группу попарно несовместных событий, то сумма их вероятностей равна единице:

$$\sum_{i=1}^n P(A_i) = 1.$$

Противоположными событиями называются два несовместных события, составляющие (образующие) полную группу A и \bar{A} . Сумма вероятностей противоположных событий равна единице:

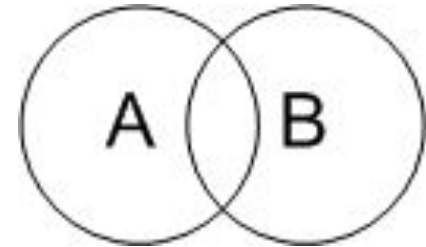
$$P(A) + P(\bar{A}) = 1.$$

2. ВЕРОЯТНОСТЬ

Теорема сложения вероятностей совместных событий

Теорема. Вероятность появления хотя бы одного из двух совместных событий равна сумме вероятностей этих событий без вероятности их совместного появления:

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



Доказательство. Событие $A+B$ наступит, если наступит одно из трех несовместных событий: $\bar{A}B, A\bar{B}, AB$. По теореме сложения вероятностей несовместных событий имеем:

$$P(A + B) = P(\bar{A}B) + P(A\bar{B}) + P(AB).$$

Событие A произойдет, если наступит одно из двух несовместных событий: $A\bar{B}, AB$. Вновь применяя теорему, получим

$$P(A) = P(A\bar{B}) + P(AB);$$

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

2. ВЕРОЯТНОСТЬ

Теорема сложения вероятностей совместных событий

Доказательство (продолжение).

Аналогично для события B :

$$P(B) = P(\bar{A}B) + P(AB);$$

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

Отсюда:

$$\begin{aligned} P(A+B) &= P(A\bar{B}) + P(\bar{A}B) + P(AB) = \\ &= P(A) - P(AB) + P(B) - P(AB) + P(AB) = \\ &= P(A) + P(B) - P(AB). \end{aligned}$$

Следствие: Вероятность произведения двух совместных событий равна

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

2. ВЕРОЯТНОСТЬ

Независимость событий

Событие A называется *независимым* от события B , если вероятность события A не зависит от того, произошло событие B или нет.

Событие A называется *зависимым* от события B , если вероятность события A меняется в зависимости от того, произошло событие B или нет.

Пример 2

В урне два белых шара и один черный. Два человека последовательно вынимают по одному шару, не возвращая их в урну.

Рассмотрим события:

A – появление белого шара у первого человека,

B – появление белого шара у второго человека.

Решение. Вероятность события A равна $2/3$. Если стало известно, что событие A произошло, то в урне осталось два шара, из которых только один белый. Тогда вероятность события B становится равной $1/2$. Из этого заключаем, что событие B зависит от события A .

2. ВЕРОЯТНОСТЬ

Условная вероятность

Вероятность события В , вычисленная при условии, что имело место другое событие А, называется условной вероятностью события В и обозначается: $P(B / A)$.

Для Примера 2: $P(A) = 2/3$, $P(B / A) = 1/2$.

С учетом определения условной вероятности можно записать условия зависимости и независимости событий.

Если $P(A \setminus B) = P(A)$, то события А и В являются независимыми.

Если $P(A \setminus B) \neq P(A)$, то события А и В являются зависимыми.

Возвращаясь к Примеру 2, можно добавить, что если после извлечения шар возвращается в урну, то события А и В – независимые; если шар не возвращается в урну – зависимые.

2. ВЕРОЯТНОСТЬ

Теорема умножения вероятностей

Теорема. Вероятность произведения двух событий равна произведению вероятностей одного из них на условную вероятность другого, вычисленную при условии, что первое имело место:

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

Следствие 1. Если событие A не зависит от события B , то и событие B не зависит от события A .

Следствие 2. Вероятность произведения независимых событий равна произведению вероятностей этих событий:

$$P(A_1 \cdot A_2 \cdot \dots \cdot A_n) = \prod_{i=1}^n P(A_i).$$

2. ВЕРОЯТНОСТЬ

Формула полной вероятности

Следствием обеих основных теорем – теоремы сложения вероятностей и теоремы умножения вероятностей – является так называемая **формула полной вероятности**.

Пусть требуется определить вероятность некоторого события A , которое может произойти или не произойти вместе с одним из событий H_1, H_2, \dots, H_n , образующих полную группу несовместных событий. События $H_i, i = 1, 2, \dots, n$ будем называть **гипотезами**.

Теорема. Вероятность события A равна сумме произведений вероятности гипотезы на соответствующую условную вероятность этого события:

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A / H_i).$$

2. ВЕРОЯТНОСТЬ

Формула Байеса

Формула Байеса позволяет «переставить причину и следствие»: по известному факту события вычислить вероятность того, что оно было вызвано данной причиной.

События, отражающие действие «причин», в данном случае называют гипотезами, так как они — предполагаемые события, повлекшие данное.

Безусловную вероятность справедливости гипотезы называют **априорной** (насколько вероятна причина вообще), а условную — с учетом учетом факта произошедшего события — **апостериорной** (насколько вероятной оказалась причина с учетом данных о событии).



Томас Байес
1701-1761

2. ВЕРОЯТНОСТЬ

Формула Байеса

Сформулируем задачу. Имеется полная группа несовместных событий (гипотез) H_1, H_2, \dots, H_n .

Вероятности этих гипотез известны и равны $P(H_1), P(H_2), \dots, P(H_n)$.

Произведен опыт, в результате которого наблюдалось событие A .

Спрашивается, как следует изменить вероятности гипотез в связи с появлением этого события?

Фактически нам необходимо найти условную вероятность $P(A) > 0$ для каждой гипотезы. Из теоремы умножения вероятностей имеем:

$$P(A \cdot H_i) = P(A) \cdot P(H_i / A) = P(H_i) \cdot P(A / H_i),$$
$$i = 1, 2, \dots, n.$$

2. ВЕРОЯТНОСТЬ

Формула Байеса

Отсюда

$$P(A) \cdot P(H_i / A) = P(H_i) \cdot P(A / H_i), \quad i = 1, 2, \dots, n.$$

Разделим на $P(A) > 0$ левую и правую часть уравнения, тогда окончательно получим

$$P(H_i / A) = \frac{P(H_i) \cdot P(A / H_i)}{P(A)}, \quad i = 1, 2, \dots, n.$$

Выражая $P(A)$ с помощью формулы полной вероятности получим **формулу Байеса**:

$$P(H_i / A) = \frac{P(H_i) \cdot P(A / H_i)}{\sum_{i=1}^n P(H_i) \cdot P(A / H_i)}, \quad i = 1, 2, \dots, n.$$

2. ВЕРОЯТНОСТЬ

Формула полной вероятности

Пример 3.

Имеются три одинаковые урны. В первой урне находятся два белых и один черный шар. Во второй урне – три белых и один черный, а в третьей урне – два белых и два черных. Какова вероятность того, что некто подойдет и из произвольной урны извлечет белый шар?

Решение. Рассмотрим 3 гипотезы:

- H_1 - выбор первой урны;
- H_2 - выбор второй урны;
- H_3 - выбор третьей урны.

Событие A - вынут белый шар.

Из условия задачи следует, что гипотезы равновозможны:

$$P(H_1) = P(H_2) = P(H_3) = 1/3.$$



2. ВЕРОЯТНОСТЬ

Формула полной вероятности

Пример 3 (продолжение).

Если случайно подойти к первой урне, то вероятность извлечь из нее белый шар равна $2/3$.

Рассуждая аналогичным образом, вычислим условные вероятности события A при этих гипотезах соответственно:

$$P(A|H_1) = 2/3, P(A|H_2) = 3/4, P(A|H_3) = 1/2.$$

По формуле полной вероятности окончательно получим:

$$P(A) = \frac{1}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{3}{4} + \frac{1}{3} \cdot \frac{1}{2} = \frac{23}{36} \approx 0.639.$$

2. ВЕРОЯТНОСТЬ

Формула Байеса

Пример 4.

Имеются три одинаковые урны. В первой урне находятся два белых и один черный шар. Во второй урне – три белых и один черный, а в третьей урне – два белых и два черных. Из некоторой урны был извлечен белый шар. Какова вероятность того, что шар был извлечен из второй урны?

Решение. Аналогично предыдущей задаче, мы рассматриваем 3 равновероятные гипотезы:

H_1 - выбор первой урны;

H_2 - выбор второй урны;

H_3 - выбор третьей урны.

$$P(H_1) = P(H_2) = P(H_3) = 1/3.$$

Событие А - вынут белый шар.



2. ВЕРОЯТНОСТЬ

Формула Байеса

Пример 4 (продолжение).

Необходимо определить вероятность того, что белый шар был извлечен именно из второй урны, т.е. $P(H_2/A)$. По формуле Байеса:

$$P(H_2 / A) = \frac{P(H_2) \cdot P(A / H_2)}{\sum_{i=1}^3 P(H_i) \cdot P(A / H_i)}.$$

Т.к. во второй урне находились 3 белых шара и один черный, то

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

Знаменатель формулы Байеса представляет собой полную вероятность события A , найденную в прошлой задаче:

$$\sum_{i=1}^3 P(H_i) \cdot P(A / H_i) = \frac{23}{36}.$$

Следовательно:

$$P(H_2 / A) = \frac{1}{4} \cdot \frac{36}{23} = \frac{9}{23} \approx 0.391.$$

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ

Схема Бернулли

Если производится несколько испытаний (опытов), причем вероятность события A в каждом испытании не зависит от исходов других испытаний, то такие испытания называются независимыми относительно события A .

В схеме Бернулли рассматривается серия, состоящая из n независимых испытаний, каждое из которых имеет лишь два исхода:

наступление какого-то события A (успех)
или его не наступление \bar{A} (неудача).

Причем вероятность успеха при одном испытании $P(A) = p$ – постоянна и не зависит от номера испытания.

Следовательно, вероятность неуспеха $P(\bar{A}) = 1 - p = q$ тоже постоянна.



Якоб Бернулли
1654-1705

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ

Схема Бернулли

Сформулируем задачу – вычислить вероятность того, что при n испытаниях событие A осуществится ровно k раз и, следовательно, не осуществится $(n - k)$ раз.

По теореме умножения вероятностей независимых событий искомая вероятность будет равна

$$P = p \cdot p \cdot \dots \cdot p \cdot q \cdot q \cdot \dots \cdot q = p^k q^{n-k}.$$

Однако интересующее нас событие (k успехов при n опытах) может произойти не только одним способом. Число возможных вариантов (комбинаций) выборки k элементов из n вычисляется по формуле сочетаний:

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

Окончательно получим:

$$P_n(k) = C_n^k p^k q^{n-k} = \frac{n! p^k q^{n-k}}{k!(n-k)!} \quad - \text{ формула Бернулли.}$$

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ

Схема Бернулли

Пусть

$$P_n(m_1 \leq k \leq m_2)$$

означает вероятность того, что в n испытаниях схемы Бернулли успех наступит не менее чем m_1 раз, и не более чем m_2 раз. При этом

$$0 \leq m_1 \leq m_2 \leq n.$$

Так как события, соответствующие различному числу успехов попарно несовместны, то имеет место формула:

$$P_n(m_1 \leq k \leq m_2) = \sum_{i=m_1}^{m_2} P_n(i).$$

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ

Схема Бернулли (обобщение)

В классической схеме Бернулли подразумевается, что в результате каждого испытания некоторое событие может либо произойти, либо не произойти, т.е. возможны только 2 исхода.

Пусть каждое из n испытаний имеет $m > 2$ попарно несовместных возможных исходов $(A_j, j=1,2,\dots,m)$, составляющих полную группу событий.

Вероятности наступления каждого события p_j в общем случае различны и удовлетворяют условию

$$\sum_{i=1}^m p_i = 1.$$

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ

Схема Бернулли (обобщение)

В этих условиях для произвольно заданных целых неотрицательных чисел k_i таких, что

$$\sum_{i=1}^m k_i = n,$$

определим вероятность того, что при n испытаниях исход A_1 наступит ровно k_1 раз, исход A_2 – k_2 раз, и т.д., исход A_m произойдет k_m раз:

$$P_n(k_1, k_2, \dots, k_m) = \frac{n!}{k_1! k_2! \dots k_m!} p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}.$$

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ

Схема Бернулли

Вероятность того, что при n испытаниях событие A осуществится **ровно** **k** раз:

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

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

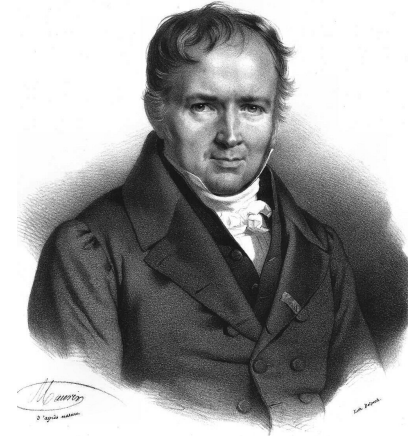
3. ПОВТОРЕНИЕ ИСПЫТАНИЙ (Пуассоновское приближение для биномиального распределения) Теорема Пуассона

Формула Бернулли удобна для вычисления лишь при сравнительно небольшом числе испытаний n .

При больших значениях n пользоваться этой формулой затруднительно.

Еще большая проблема возникает, если в схеме Бернулли число испытаний велико, а вероятность успеха мала.

Рассмотрим классическую схему Бернулли с большим числом испытаний и малой вероятностью появления события:



Симеон Дени
Пуассон
1781-1840

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ (Пуассоновское приближение ...)

Теорема Пуассона

Пусть $n \rightarrow \infty$, $p_n \rightarrow 0$ так, что $np_n \rightarrow \lambda > 0$.

Тогда для любого $k \geq 0$ вероятность получить k успехов в n испытаниях схемы Бернулли с вероятностью успеха p_n стремится к величине

$$\frac{\lambda^k}{k!} e^{-\lambda}.$$

Действительно,

$$\begin{aligned} P_n(k) &= C_n^k p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \\ &= \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!} \left[\frac{n!}{n^k (n-k)!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \right] \cong \\ &\cong \frac{\lambda^k}{k!} e^{-\lambda}. \end{aligned}$$

Пуассоновское приближение для биномиального распределения

$$k = 0 \quad P_n(0) = \left(1 - \frac{\lambda}{n}\right)^n$$

$$\log P_n(0) = n \log\left(1 - \frac{\lambda}{n}\right) = -\lambda - \frac{\lambda^2}{2n} - \dots$$

$$P_n(0) \approx e^{-\lambda} \quad P_n(k) / P_n(k-1) = [\lambda - (k-1)p] / kq \approx \lambda / k$$

$$P_n(1) \approx \lambda P_n(0) \approx \lambda e^{-\lambda}$$

$$P_n(2) \approx (1/2)\lambda P_n(1) \approx (1/2)\lambda^2 e^{-\lambda}$$

... ..

$$P_n(k) \approx \left(\frac{\lambda^k}{k!}\right) e^{-\lambda}$$

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ (Пуассоновское приближение ...)

Теорема Пуассона

Распределение Пуассона:

$$P_n(k) \cong \frac{\lambda^k}{k!} e^{-\lambda}$$

При решении конкретных задач понятия число испытаний – «велико» и вероятность успеха – «мала» – субъективные. При более строгом подходе воспользуемся оценкой погрешности формулы Пуассона:

$$\Psi = \left| C_n^k p^k (1-p)^{n-k} - \frac{\lambda^k}{k!} e^{-\lambda} \right| \leq np^2 = \frac{\lambda^2}{n}.$$

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ (Пуассоновское приближение ...)

Теорема Пуассона (

Пример 5.

На предприятии изготовлено и отправлено заказчику 100 000 бутылок пива. Вероятность того, что бутылка может оказаться битой, равна 0.0001. Какова вероятность того, что в отправленной партии будет ровно три битых бутылки?

Решение. Воспользуемся формулой Пуассона, учитывая, что $n=100000$, $p=0,0001$ и $\lambda=np=10$:

$$P_{100000}(3) = \frac{\lambda^k}{k!} e^{-\lambda} = \frac{10^3}{3!} e^{-10} \approx 0.007567.$$

Вычислим погрешность:

$$\Psi = \frac{\lambda^2}{n} = \frac{10^2}{100000} = 0.001.$$

Т.о., искомая вероятность равна:

$$P_{100000}(3) \pm \Psi = 0.007567 \pm 0.001.$$

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ (Нормальное приближение для биномиального распределения)

Локальная теорема Муавра-Лапласа

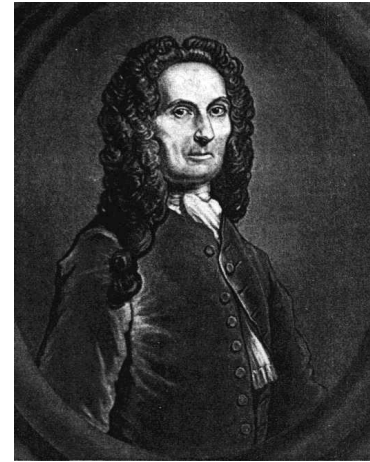
Несмотря на элементарность формулы Бернулли при большом числе испытаний непосредственное вычисление по ней связано с большой вычислительной работой (погрешностью).

Разрешить эту проблему поможет локальная теорема Муавра-Лапласа:

Если вероятность p наступления события A в каждом испытании постоянна и отлична от 0 и 1, то вероятность $P_n(k)$ того, что событие A произойдет k раз в n независимых испытаниях при достаточно большом числе n , приближенно равна

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \cdot \varphi\left(\frac{k - np}{\sqrt{npq}}\right),$$

где $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$ - функция Гаусса.



Абрахам де Муавр
1667-1754



Пьер-Симон Лаплас
1749-1827

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ (Нормальное приближение для биномиального распределения)

Локальная теорема Муавра-Лапласа

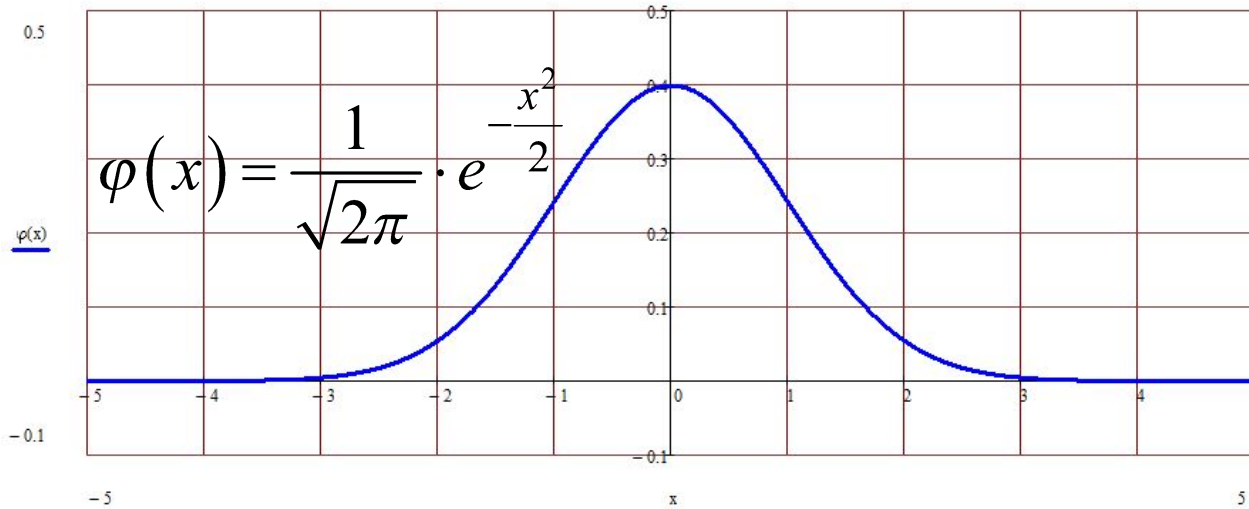
$$P_n(k) \approx \frac{1}{\sqrt{npq}} \cdot \varphi\left(\frac{k - np}{\sqrt{npq}}\right)$$

Формула, приведенная в теореме Муавра-Лапласа, тем точнее, чем больше n .

Вычисление по этой формуле дает незначительную погрешность уже при выполнении условия $npq \geq 20$.

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ (Нормальное приближение для биномиального распределения)

Функция Гаусса



Карл Фридрих Гаусс
1777-1855

Свойства:

$\varphi(x) = \varphi(-x)$ - четная функция;

$\varphi(x)$ - монотонно убывающая при $x \in [0; \infty)$;

$\varphi(x) \rightarrow 0$ при $x \rightarrow \infty$;

$\varphi(x) \leq 0.0001$ при $x \geq 4$.

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ (Нормальное приближение для биномиального распределения)

Интегральная теорема Муавра-Лапласа

Если вероятность p наступления события A в каждом испытании постоянна и отлична от 0 и 1, то вероятность того, что число k наступления события A в n независимых испытаниях заключено в пределах от m_1 до m_2 (включительно) при достаточно большом числе n приближенно равна

$$P_n(m_1 \leq k \leq m_2) = \sum_{i=m_1}^{m_2} C_n^i p^i q^{n-i} \approx \frac{1}{\sqrt{2\pi}} \cdot \int_{\alpha}^{\beta} e^{-\frac{x^2}{2}} dx,$$

где

$$\alpha = \frac{m_1 - np}{\sqrt{npq}}, \quad \beta = \frac{m_2 - np}{\sqrt{npq}}.$$

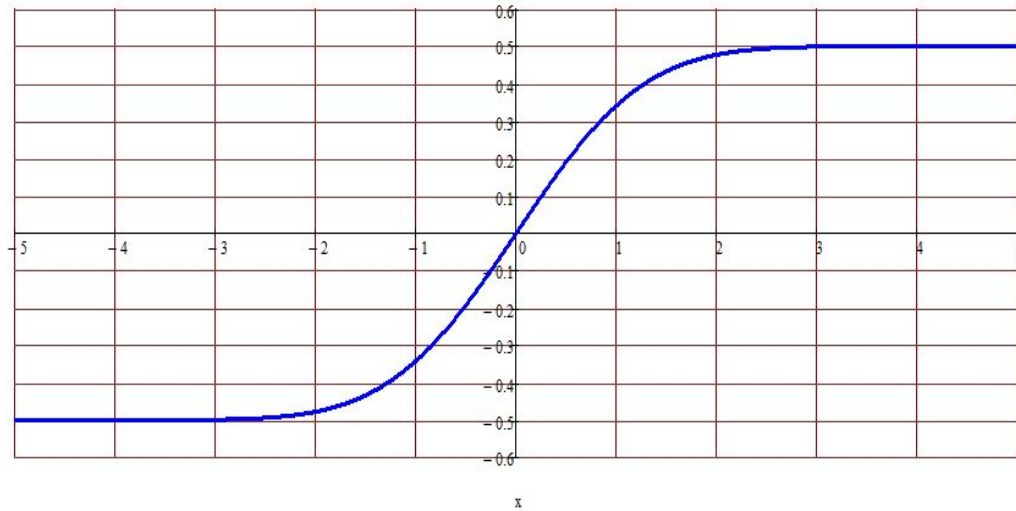
Введем новую функцию (см. далее)



3. ПОВТОРЕНИЕ ИСПЫТАНИЙ (Нормальное приближение для биномиального распределения)

Функция Лапласа

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^x e^{-\frac{t^2}{2}} dt$$



Свойства:

$\Phi_0(x) = -\Phi_0(-x)$ - нечетная функция;

$\Phi_0(x)$ - монотонно возрастающая функция;

$\Phi_0(x) \rightarrow \pm 0.5$ при $x \rightarrow \pm\infty$;

$\Phi_0(x) \approx 0.5$ при $x > 4$.

3. ПОВТОРЕНИЕ ИСПЫТАНИЙ (Нормальное приближение для биномиального распределения)

Интегральная теорема Муавра-Лапласа

С учетом свойств функции Лапласа получим выражение для интегральной теоремы Муавра-Лапласа:

$$P_n(m_1 \leq k \leq m_2) \approx \frac{1}{\sqrt{2\pi}} \cdot \int_{\alpha}^{\beta} e^{-\frac{x^2}{2}} dx = \Phi_0(\beta) - \Phi_0(\alpha),$$

где

$$\alpha = \frac{m_1 - np}{\sqrt{npq}}, \quad \beta = \frac{m_2 - np}{\sqrt{npq}}.$$

Интегральная формула, как и локальная тем точнее, чем больше n . При условии $npq \geq 20$ интегральная формула дает незначительную погрешность вычисления вероятностей.

Значения функции Лапласа в учебниках обычно приводятся в табличном виде в приложениях.

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Классификация случайных величин

Числовая величина ζ , значение которой меняется в зависимости от случая, называется случайной величиной (СВ).

Если мы рассматриваем вероятностное пространство E элементарных исходов e ($e \in E$), тогда случайной величиной будем называть

функцию от элементарных исходов:

$$\zeta = \zeta(e)$$

Функция, определенная на пространстве элементарных событий, называется случайной величиной.

Различают два основных типа случайных величин:

- дискретные;
- непрерывно распределенные.

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Дискретные случайные величины

Дискретная величина $\zeta = \zeta(e)$ в зависимости от элементарных исходов e принимает конечное или счетное число различных значений x с соответствующими вероятностями:

$$P_{\zeta}(x) = P(\zeta = x).$$

$$\zeta = \zeta(e)$$

$$\{\zeta = x\} = \{e : \zeta(e) = x\}, \quad e \in E$$

Вероятность события $x_1 \leq \zeta \leq x_2$, заключающееся в том, что СВ ζ примет одно из значений x , лежащее в указанных пределах, есть

$$P\{x_1 \leq \zeta \leq x_2\} = \sum_{x_1}^{x_2} P_{\zeta}(x).$$

В данной формуле суммирование производится по конечному или счетному числу значений x , которые может принимать дискретная случайная величина ζ .

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Дискретные случайные величины

Законом распределения СВ называется всякое соотношение, устанавливающее связь между возможными значениями СВ и соответствующими им вероятностями.

Простейшей формой задания этого закона является таблица, в которой перечислены возможные значения СВ и соответствующие им вероятности:

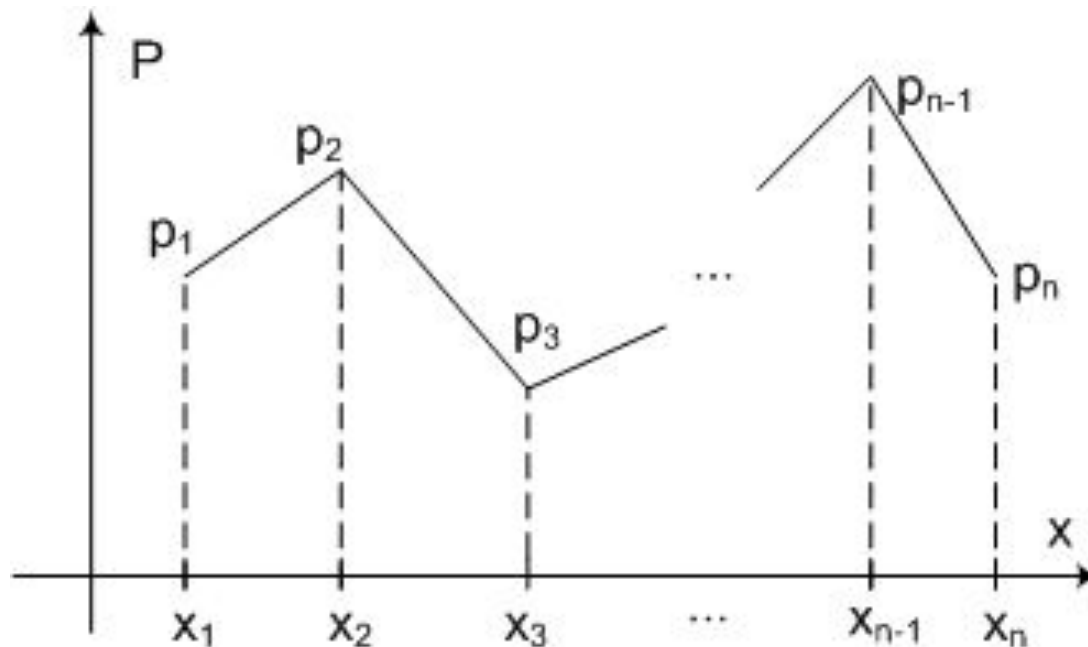
x_i	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_n

Такую таблицу будем называть **рядом распределения дискретной СВ.**

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Дискретные случайные величины

Чтобы придать ряду распределения более наглядный вид часто прибегают к его графическому отображению. Такое представление СВ называется многоугольником распределения.



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Дискретные случайные величины

Очевидно, что ряд распределения не универсальная характеристика. Нетрудно убедиться, что для непрерывной СВ такую характеристику построить нельзя (так как СВ имеет бесчисленное множество значений). Поэтому составить таблицу, в которой бы были перечислены все возможные значения СВ невозможно.

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Интегральная функция распределения

Для количественного описания распределения вероятностей удобно воспользоваться не вероятностью события $X = x$, а вероятностью события $X < x$, где x - некоторая текущая переменная.

Вероятность этого события, очевидно, зависит от x , и является некоторой функцией от x . Эта функция называется функцией распределения случайной величины X и обозначается $F(x)$:

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

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

Функция распределения – самая универсальная характеристика СВ. Она существует как для дискретных, так и непрерывных СВ. Функция распределения полностью характеризует СВ с вероятностной точки зрения и является одной из форм закона распределения.

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Общие свойства интегральной функции распределения

1) Функция распределения $F(x)$ есть неубывающая функция своего аргумента, то есть при $x_2 \geq x_1$

$$F(x_2) \geq F(x_1);$$

2) При $t \rightarrow -\infty$

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

3) При $t \rightarrow \infty$

$$F(\infty) = 1;$$

4)

$$P(x_1 \leq x \leq x_2) = F(x_2) - F(x_1).$$

График функции распределения в общем случае представляет собой график неубывающей функции, значение которой начинается от 0 и доходит до 1, причем в отдельных точках функция **может иметь разрыв**.

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Общие свойства интегральной функции распределения

Зная ряд распределения дискретной СВ, можно легко построить функцию распределения этой величины. Действительно:

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

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Непрерывные случайные величины

Случайная величина X называется **непрерывной**, если ее пространством элементарных событий является вся числовая ось (либо отрезок (отрезки) числовой оси), а **вероятность наступления любого элементарного события равна нулю**.

Для непрерывной случайной величины вероятность попасть на интервал равна

$$P(a \leq X \leq b) = F(b) - F(a).$$

Пусть имеется непрерывная СВ X с функцией распределения $F(x)$, которую мы предполагаем непрерывной и дифференцируемой. Вычислим вероятность попадания этой СВ на участок от x до $x + \Delta x$, то есть приращение функции распределения на этом участке:

$$P(x \leq X \leq x + \Delta x) = F(x + \Delta x) - F(x).$$

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Плотность распределения вероятностей

Найдем отношение этой вероятности к длине участка, то есть среднюю вероятность, приходящуюся на единицу длины на этом участке, и устремим Δx к 0. В пределе получим производную функции распределения:

$$\lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = F'(x) = f(x).$$

Функция $f(x)$ – производная функции распределения характеризует плотность, с которой распределяются значения СВ в данной точке.

Эта функция называется **плотностью распределения (плотностью вероятности)** непрерывной СВ X .

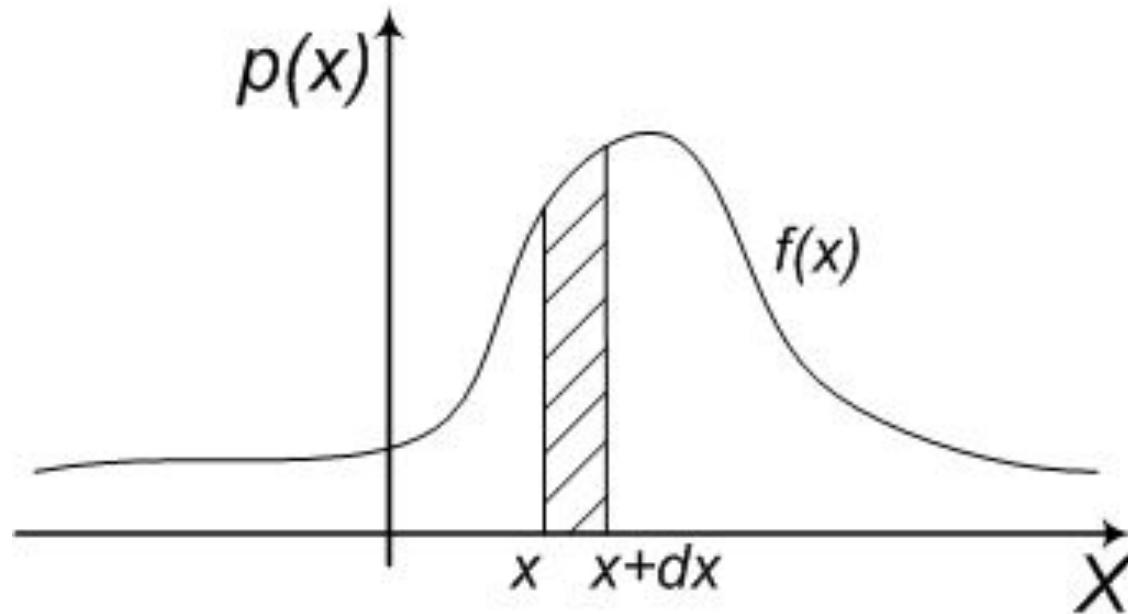
Плотность распределения является одной из форм закона распределения. Эта форма не является универсальной, так как $f(x)$ существует **только для непрерывных СВ**.

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Плотность распределения (геометрический смысл)

Рассмотрим непрерывную СВ X с плотностью распределения $f(x)$ и элементарный участок dx , примыкающий к точке x . Вероятность попадания СВ X на этот элементарный участок (с точностью до бесконечно малых высшего порядка) равна $f(x)dx$.

Геометрически – это площадь элементарного прямоугольника, опирающегося на отрезок dx .

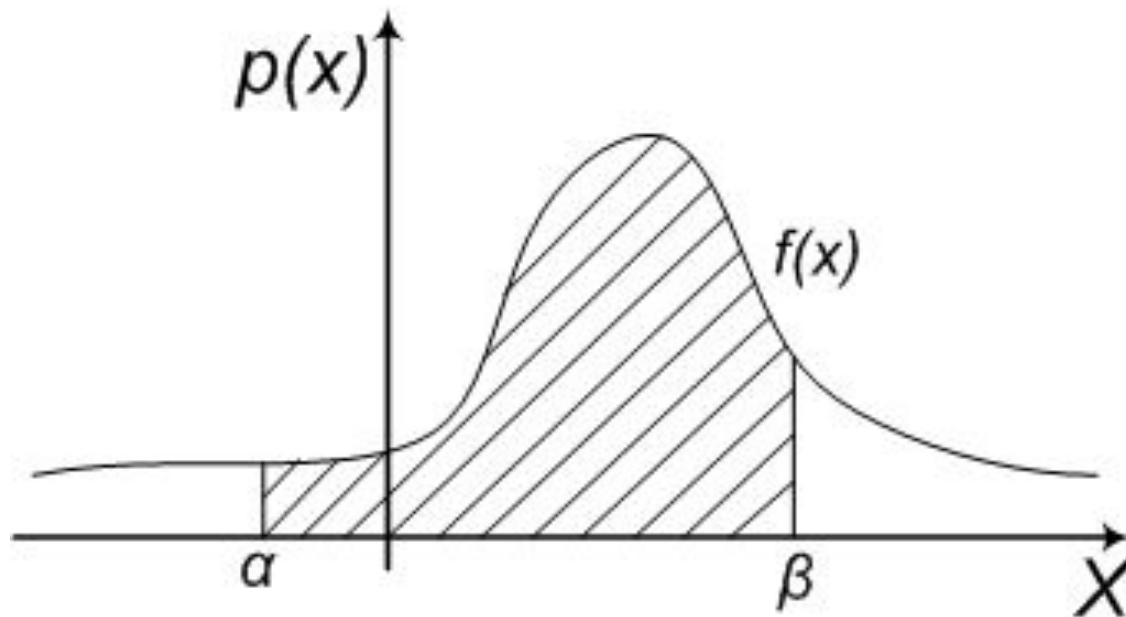


4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Плотность распределения (геометрический смысл)

Выразим вероятность попадания СВ X на отрезок от α до β через плотность распределения. Очевидно, она равна сумме элементов вероятности на всем участке, то есть интегралу:

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



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

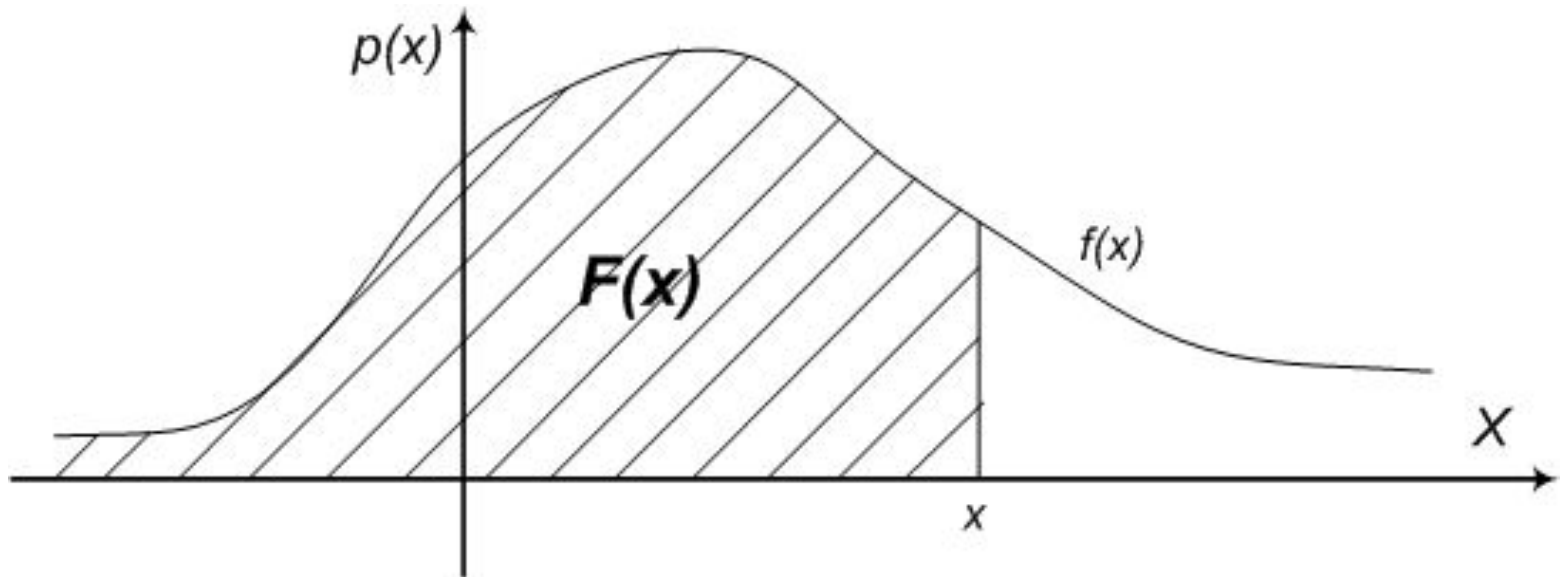
Функция распределения (геометрический смысл)

Выразим функцию распределения через плотность. Согласно определению

$$F(x) = P(X < x) = P(-\infty < X < x).$$

Учитывая, что $P(\alpha \leq X \leq \beta) = \int_{\alpha}^{\beta} f(x) dx$, получим

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



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Плотность распределения

1. Так как функция распределения – неубывающая, функция плотности распределения:

$$f(x) = F'(x) \geq 0.$$

2. Площадь под функцией плотности распределения всегда равна 1:

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

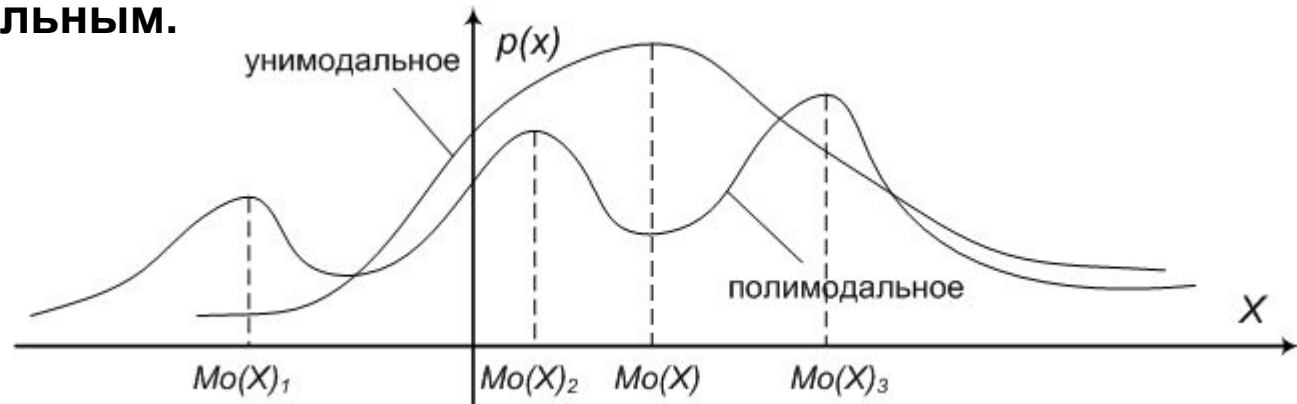
4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Характеристики положения случайной величины. Мода.

На практике в теории вероятностей применяют характеристики положения случайных величин, отражающие те или другие особенности распределения.

Модой $Mo(X)$ случайной величины X называется ее наиболее вероятное значение (для которого вероятность p_i или плотность вероятности $f(x)$ достигает максимума).

Если вероятность или плотность вероятности достигает максимума в одной точке, распределение называется **унимодальным**, если же максимум достигается в нескольких точках, распределение называется **полимодальным**.



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

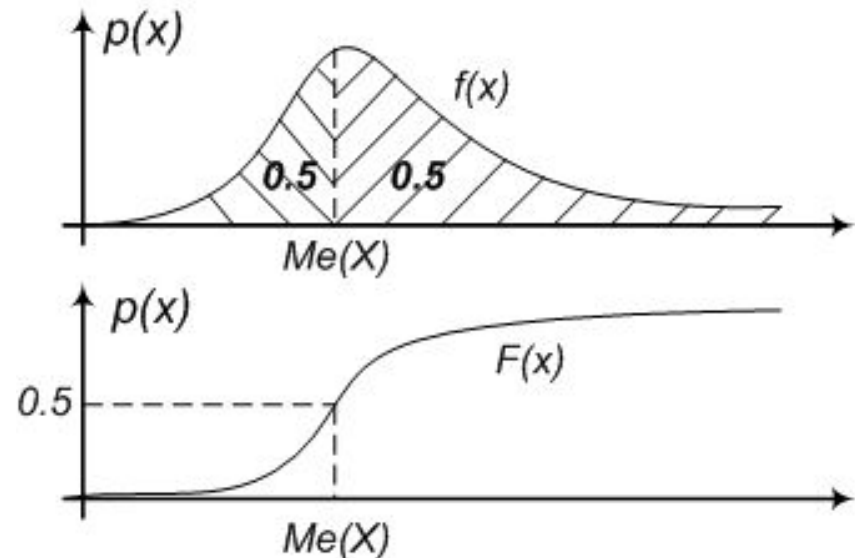
Характеристики положения случайной величины. Медиана.

Медианой $Me(X)$ случайной величины X называется такое ее значение, при котором вероятность того, что СВ $X < Me$ одинаково вероятна тому, что СВ $X > Me$, и будет равна 0.5.

$$P(-\infty \leq x \leq Me) = P(Me \leq x \leq \infty) = 0.5;$$

$$\int_{-\infty}^{Me} f(x) dx = \int_{Me}^{\infty} f(x) dx = 0.5.$$

Геометрически медиана – это абсцисса точки, в которой площадь, ограниченная кривой распределения, делится пополам (функция распределения равна 0,5).



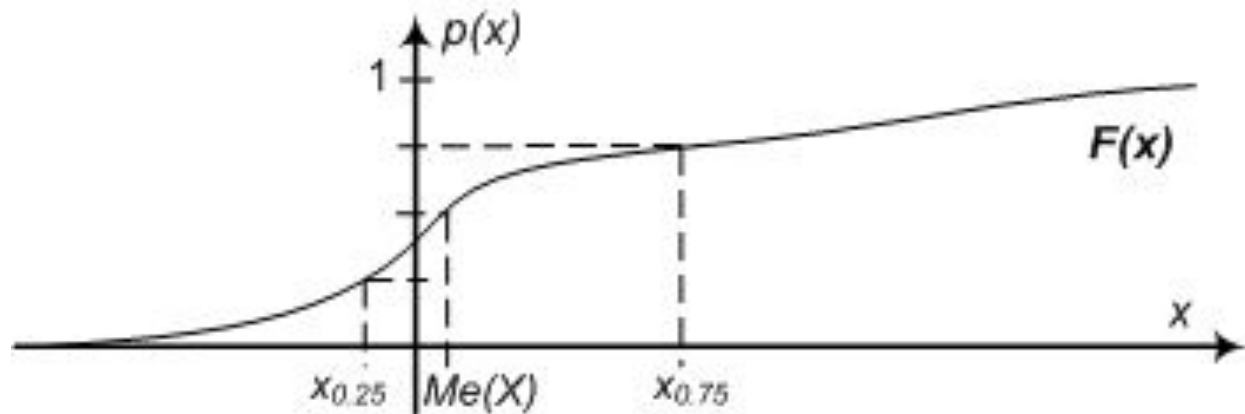
4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Характеристики положения случайной величины. Квантиль.

Квантилью уровня q (или q -квантилью) называется такое значение x_q случайной величины, при котором функция ее распределения принимает значение, равное q , то есть

$$F(x_q) = P(X < x_q) = q.$$

Некоторые квантили получили особое название. Очевидно, что определенная выше медиана случайной величины есть квантиль уровня 0.5, то есть $Me(X) = x_{0,5}$. Квантили $x_{0,25}$ и $x_{0,75}$ получили название нижней (первой) и верхней (третьей) квантилей соответственно.



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Характеристики положения случайной величины. Процентная точка.

С понятием квантили тесно связано понятие **процентной точки**.

Под **$q\%$ -ной точкой** подразумевается квантиль x_{1-q} , то есть такое значение случайной величины X , при котором

$$P(X \geq x_{1-q}) = q.$$

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Характеристики положения случайной величины.

Пример 6

Найти моду, медиану, квантиль $x_{0,3}$ и 30%-ную точку СВ X с плотностью вероятности

$$f(x) = 3x^2, \quad 0, x \in [1]$$

Решение.

Для нахождения моды распределения необходимо найти максимум функции плотности. Однако эта функция возрастает на интервале $[0,1]$, следовательно, максимум достигается при $x = Mo(X) = 1$.

Медиана – это такое значение $Me(X)$, которое делит площадь под кривой плотности распределения на две равные части. Следовательно

$$\int_{-\infty}^{Me(X)} f(x) dx = \int_0^{Me(X)} 3x^2 dx = \left(x^3 \right) \Big|_0^{Me(X)} = Me(X)^3 = 0.5.$$



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Характеристики положения случайной величины.

Пример 6 (продолжение)

$$\int_{-\infty}^{Me(X)} f(x) dx = Me(X)^3 = 0.5.$$

Отсюда

$$Me(X) = \sqrt[3]{0.5} \approx 0.79.$$

Функция распределения есть интеграл от функции плотности:

$$F(x) = \int_{-\infty}^x f(\tau) d\tau = \int_{-\infty}^0 0 d\tau + \int_0^x 3\tau^2 d\tau = x^3.$$

Квантиль уровня 0,3:

$$F(x_{0.3}) = 0.3 \Rightarrow x_{0.3}^3 = 0.3 \Rightarrow x_{0.3} \approx 0.67.$$

30%-ная точка (квантиль уровня 0,7):

$$F(x_{0.7}) = 0.7 \Rightarrow x_{0.7}^3 = 0.7 \Rightarrow x_{0.7} \approx 0.89.$$

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Математическое ожидание.

Математическим ожиданием (или средним значением) случайной величины X называется постоянное число, обозначаемое символами m_x , $M[X]$ или \bar{x} , и определяемое равенством:

$$\bar{x} = m_x = M[X] = \begin{cases} \sum_{-\infty}^{\infty} x_i p_i & \text{для дискретной СВ} \\ \int_{-\infty}^{\infty} x f(x) dx & \text{для непрерывной СВ} \end{cases}$$

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Математическое ожидание. Свойства.

1. Математическое ожидание константы равно самой константе:

$$M [C] = C .$$

2. Постоянный множитель можно выносить за знак математического ожидания:

$$M [k \cdot X] = k \cdot M [X] .$$

3. Математическое ожидание алгебраической суммы конечного числа случайных величин равно сумме их математических ожиданий:

$$M \left[\sum_{i=1}^n X_i \right] = \sum_{i=1}^n M [X_i] .$$

4. Математическое ожидание произведения конечного числа **независимых** случайных величин равно произведению их математических ожиданий:

$$M \left[\prod_{i=1}^n X_i \right] = \prod_{i=1}^n M [X_i] .$$



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Математическое ожидание. Свойства.

5. Математическое ожидание отклонения случайной величины от ее математического ожидания равно нулю. Пусть математическое ожидание СВ X равно a , тогда:

$$M[X - M[X]] = M[X - a] = M[X] - a = a - a = 0.$$

Математическое ожидание – одна из характеристик положения СВ.

С этой точки зрения математическое ожидание СВ – есть некоторое число, являющееся как бы ее «представителем» и заменяющее СВ при грубых (ориентировочных) расчетах.

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Моменты случайной величины.

Начальный момент порядка s случайной величины X определяется выражением:

$$\alpha_s(x) = \begin{cases} \sum_{-\infty}^{\infty} x_i^s p_i & \text{для дискретной СВ} \\ \int_{-\infty}^{\infty} x^s f(x) dx & \text{для непрерывной СВ} \end{cases} .$$

Нетрудно убедиться, что математическое ожидание представляет собой **первый начальный момент**.

Используя символ математического ожидания, начальный момент порядка s можно представить в следующем виде:

$$\alpha_s(x) = M [X^s]$$

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Моменты случайной величины.

Пусть имеется СВ X с математическим ожиданием m_x . Введем новое понятие.

Центрированной случайной величиной, соответствующей величине X , называется отклонение СВ X от ее математического ожидания:

$$\dot{X} = X - m_x.$$

Моменты центрированной СВ называются **центрльными моментами**.

Нетрудно показать, что математическое ожидание центрированной СВ равно нулю:

$$M[\dot{X}] = M[X - m_x] = m_x - m_x = 0.$$

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Моменты случайной величины.

Центральным моментом порядка s случайной величины X называется математическое ожидание соответствующей центрированной СВ, возведенной в степень s :

$$\mu_s(x) = M[X^s] = M[(X - m_x)^s] = \begin{cases} \sum_{-\infty}^{\infty} (x - m_x)^s p(x) & \text{для дискретной СВ} \\ \int_{-\infty}^{\infty} (x - m_x)^s f(x) dx & \text{для непрерывной СВ} \end{cases}$$

Очевидно, что для любой СВ
центральный момент первого порядка
равен нулю.

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Дисперсия.

Второй центральный момент СВ называется дисперсией и обозначается $D[X]$. Дисперсией $D[X]$ случайной величины X называется математическое ожидание квадрата ее отклонения от математического ожидания:

$$\begin{aligned} \mu_2(x) = D[X] &= M[X^2] = M[(X - m_x)^2] = \\ &= \begin{cases} \sum_{-\infty}^{\infty} (\text{для дискретной СВ})^2, \\ \int_{-\infty}^{\infty} (\text{для непрерывной СВ})^2. \end{cases} \end{aligned}$$

Дисперсия СВ характеризует рассеяние (вариацию, разброс) этой величины относительно ее математического ожидания.

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Дисперсия.

Дисперсия $D[X]$ имеет размерность квадрата случайной величины, что не всегда удобно. Поэтому в качестве показателя рассеяния используют также величину равную квадратному корню из $D[X]$.

Средним квадратическим отклонением (стандартным отклонением или стандартом) σ_x случайной величины X называется арифметическое значение корня квадратного из ее дисперсии:

$$\sigma_x = \sqrt{D[X]}.$$

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Свойства дисперсии.

1. Дисперсия константы равна нулю: $D[C] = 0$.

2. Постоянный множитель можно выносить за знак дисперсии, возведя его при этом в квадрат:

$$D[CX] = C^2 D[X].$$

3. Дисперсия алгебраической суммы конечного числа независимых СВ равна сумме их дисперсий:

$$D\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n D[X_i].$$

За. Дисперсия разности равна сумме дисперсий:

$$D[X - Y] = D[X] + D[Y].$$



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Свойства дисперсии.

4. Второй центральный момент случайной величины равен разности между вторым начальным моментом и квадратом первого начального момента этой случайной величины. Другими словами:

Дисперсия случайной величины равна разности между математическим ожиданием квадрата случайной величины и квадратом ее математического ожидания:

$$D[X] = M[(X - m_x)^2] = M[X^2] - m_x^2 = \alpha_2(x) - \alpha_1^2(x).$$

5. Дисперсия произведения независимых СВ X и Y равна произведению дисперсии X на дисперсию Y плюс произведение квадрата математического ожидания СВ X на дисперсию Y плюс произведение квадрата математического ожидания СВ Y на дисперсию X .

$$D[X \cdot Y] = D[X]D[Y] + m_x^2 D[Y] + m_y^2 D[X].$$

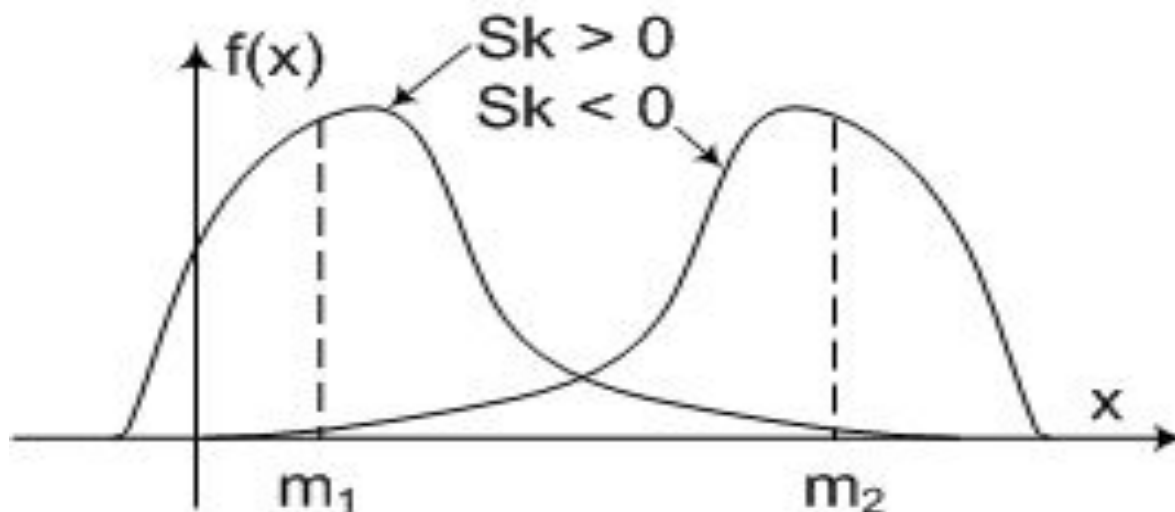
4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Асимметрия.

Третий центральный момент μ_3 служит для характеристики **асимметрии** (скошенности) **распределения**. Так как третий центральный момент имеет размерность куба случайной величины, то чтобы получить безразмерную характеристику, третий центральный момент делят на куб среднего квадратического отклонения СВ X :

$$Sk = \frac{\mu_3}{\sigma^3}.$$

Величина Sk (англ. *skewness*) называется **коэффициентом асимметрии** случайной величины



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

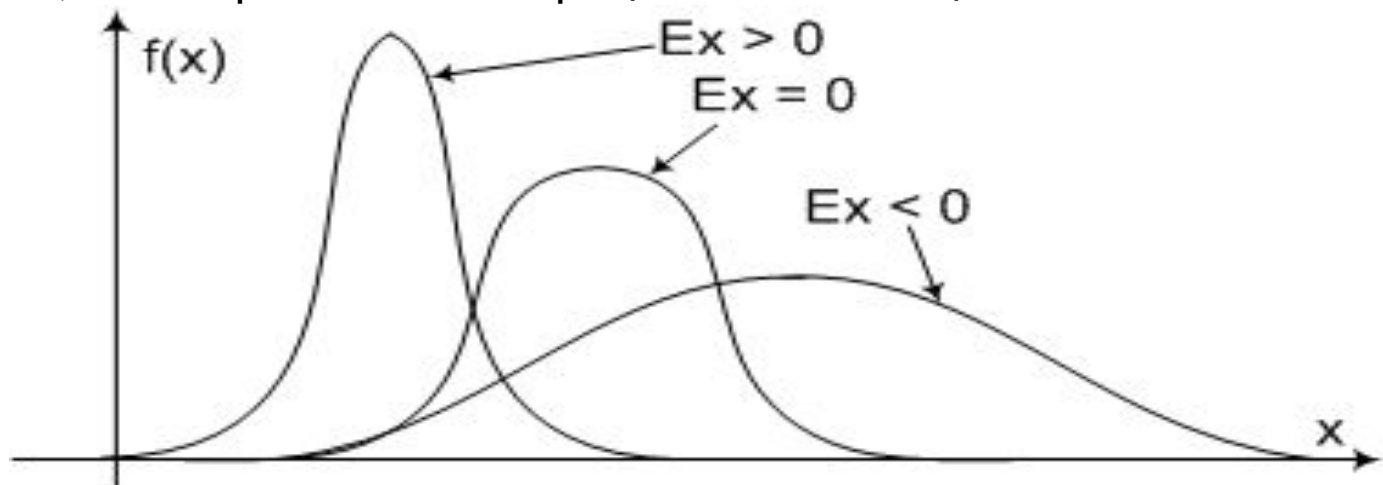
Эксцесс.

Четвертый центральный момент μ_4 служит для характеристики островершинности (крутости) распределения.

Эксцессом (коэффициентом эксцесса) СВ называется число

$$Ex = \frac{\mu_4}{\sigma^4} - 3.$$

Число 3 вычитается потому, что для нормального распределения это отношение равно 3. Т.о., распределения более островершинные, чем нормальное имеют положительный эксцесс, распределения с меньшей крутостью, чем нормальное – отрицательный эксцесс.



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Моменты случайной величины.

Пример 7

Непрерывная СВ X распределена равномерно на отрезке $[a, b]$. Определить математическое ожидание, дисперсию, коэффициенты асимметрии и эксцесса.

Решение:

Прежде всего, определим плотность распределения. Из условия задачи известно, что

$$f(x) = \begin{cases} C = \text{const} & \text{если } x \in [a, b] \\ 0 & \text{если } x \notin [a, b] \end{cases}$$

Известно, что площадь под кривой плотности должна быть равна 1:

$$\int_{-\infty}^{\infty} f(x) dx = C \cdot \int_a^b dx = Cx \Big|_a^b = C \cdot (b - a) = 1. \quad \rightarrow$$

4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Моменты случайной величины.

Пример 7 (продолжение)

Отсюда,

$$C = \frac{1}{b-a}.$$

Тогда математическое ожидание:

$$\begin{aligned} m_x &= \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_a^b \frac{xdx}{b-a} = \frac{1}{b-a} \cdot \left(\frac{x^2}{2} \right) \Big|_a^b = \\ &= \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} = \frac{a+b}{2}. \end{aligned}$$



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Моменты случайной величины.

Пример 7 (продолжение)

Известно, что

$$D[X] = M[X^2] - m_x^2.$$

Тогда:

$$\begin{aligned} D[X] &= \frac{1}{b-a} \int_a^b x^2 dx - \left(\frac{a+b}{2}\right)^2 = \frac{1}{b-a} \left(\frac{x^3}{3}\right) \Big|_a^b - \left(\frac{a+b}{2}\right)^2 = \\ &= \frac{b^3 - a^3}{3(b-a)} - \frac{a^2 + 2ab + b^2}{4} = \frac{a^2 + ab + b^2}{3} - \frac{a^2 + 2ab + b^2}{4} = \\ &= \frac{4a^2 + 4ab + 4b^2 - 3a^2 - 6ab - 3b^2}{12} = \\ &= \frac{a^2 - 2ab + b^2}{12} = \frac{(b-a)^2}{12}. \end{aligned}$$



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Моменты случайной величины.

Пример 7 (продолжение)

Среднее квадратическое отклонение:

$$\sigma = \sqrt{D[X]} = \sqrt{\frac{(b-a)^2}{12}} = \frac{b-a}{2\sqrt{3}}.$$

Третий центральный момент:

$$\begin{aligned}\mu_3 &= \int_a^b (x - m_x)^3 \cdot f(x) dx = \frac{1}{b-a} \left(\frac{(x - m_x)^4}{4} \right) \Bigg|_a^b = \\ &= \frac{\left(b - \frac{a+b}{2} \right)^4 - \left(a - \frac{a+b}{2} \right)^4}{4(b-a)} = \frac{(b-a)^4 - (a-b)^4}{2^4 \cdot 4(b-a)} = 0.\end{aligned}$$



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Моменты случайной величины.

Пример 7 (продолжение)

Т.к. третий центральный момент равен нулю, то коэффициент асимметрии

$$Sk = \frac{\mu_3}{\sigma^3}$$

также равен нулю.

Четвертый центральный момент:

$$\mu_4 = \int_a^b (x - m_x)^4 \cdot f(x) dx = \frac{1}{b-a} \left(\frac{(x - m_x)^5}{5} \right) \Big|_a^b =$$

$$= \frac{\left(b - \frac{a+b}{2} \right)^5 - \left(a - \frac{a+b}{2} \right)^5}{5(b-a)} = \frac{(b-a)^5 - (a-b)^5}{2^5 \cdot 5(b-a)}$$

$$= \frac{(b-a)(b-a)^4 - (a-b)(a-b)^4}{32 \cdot 5(b-a)} = \frac{2(b-a)(b-a)^4}{32 \cdot 5(b-a)} = \frac{(b-a)^4}{80}.$$



4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Моменты случайной величины.

Пример 7 (продолжение)

Коэффициент эксцесса:

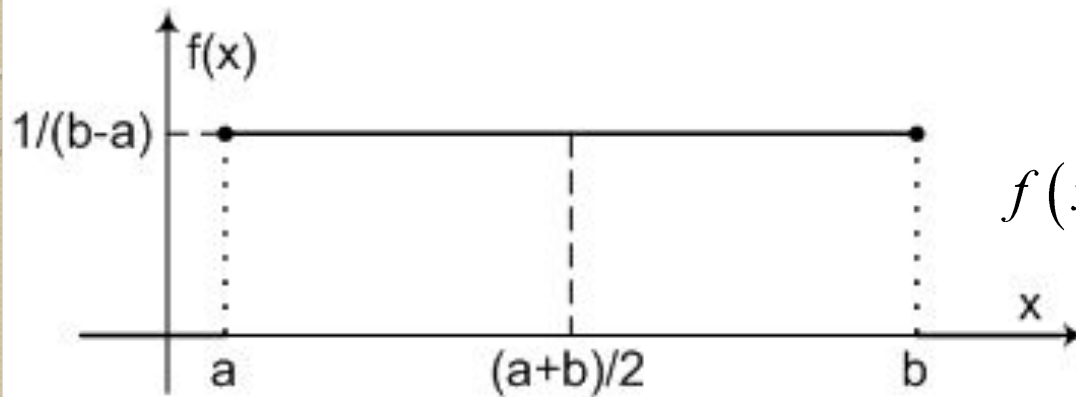
$$Ex = \frac{\mu_4}{\sigma^4} - 3 = \frac{(b-a)^4}{80} \cdot \frac{144}{(b-a)^4} - 3 = -1.2$$

В результате решения задачи мы определили основные характеристики непрерывного равномерного распределения для произвольных параметров a и b ($a < b$):

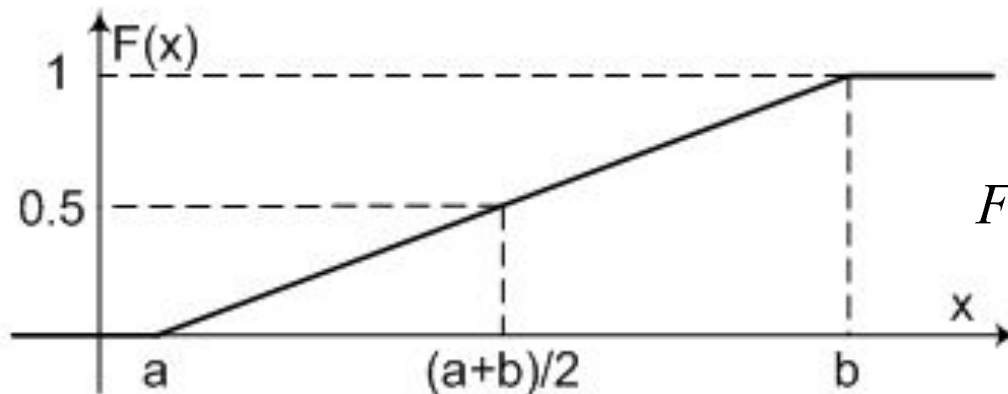
$$m_x = \frac{a+b}{2}; \quad D[X] = \frac{(b-a)^2}{12};$$
$$Sk = 0; \quad Ex = -1.2.$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Равномерное распределение.



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



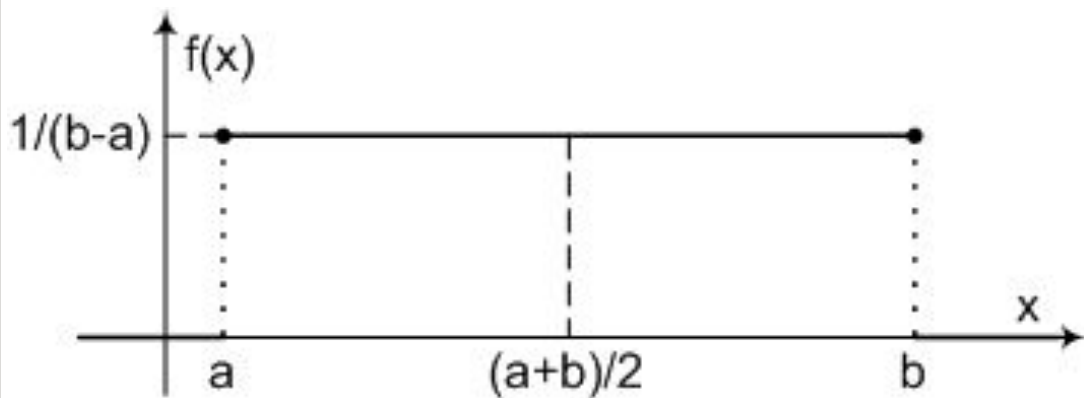
$$F(x) = \begin{cases} 0, & x < a; \\ \frac{x-a}{b-a}, & x \in [a, b]; \\ 1, & x > b. \end{cases}$$

Краткое обозначение: $U[a, b]$.

$U[0, 1]$ – стандартное равномерное распределение.

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Равномерное распределение.

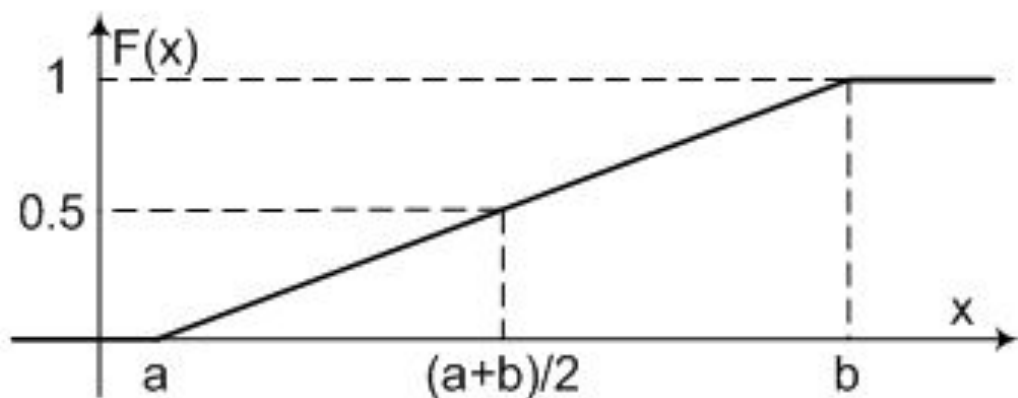


$$M[X] = Me(x) = \frac{a+b}{2};$$

$$D[X] = \frac{(b-a)^2}{12};$$

$$Sk = 0;$$

$$Ex = -1.2;$$



$$\sigma = \frac{b-a}{2\sqrt{3}}.$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Нормальный закон распределения (закон Гаусса) играет исключительно важную роль в теории вероятностей. Это наиболее часто встречающийся на практике закон распределения СВ.

Главная особенность, выделяющая закон Гаусса, состоит в том, что он является предельным законом, к которому приближаются другие законы при весьма часто встречающихся типичных условиях.

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Доказано, что сумма достаточно большого числа независимых (или слабо зависимых) случайных величин, подчиненных каким угодно законам распределения (при соблюдении некоторых весьма не жестких ограничениях) приближенно подчиняется нормальному закону. И это свойство выполняется тем точнее, чем большее количество СВ суммируется.

По нормальному закону распределены ошибки измерений, белый шум в электронике и т.п.

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Непрерывная случайная величина X имеет **нормальный закон распределения** (закон Гаусса) **с параметрами μ и σ** , если ее плотность вероятности определена на всей числовой оси x и имеет вид:

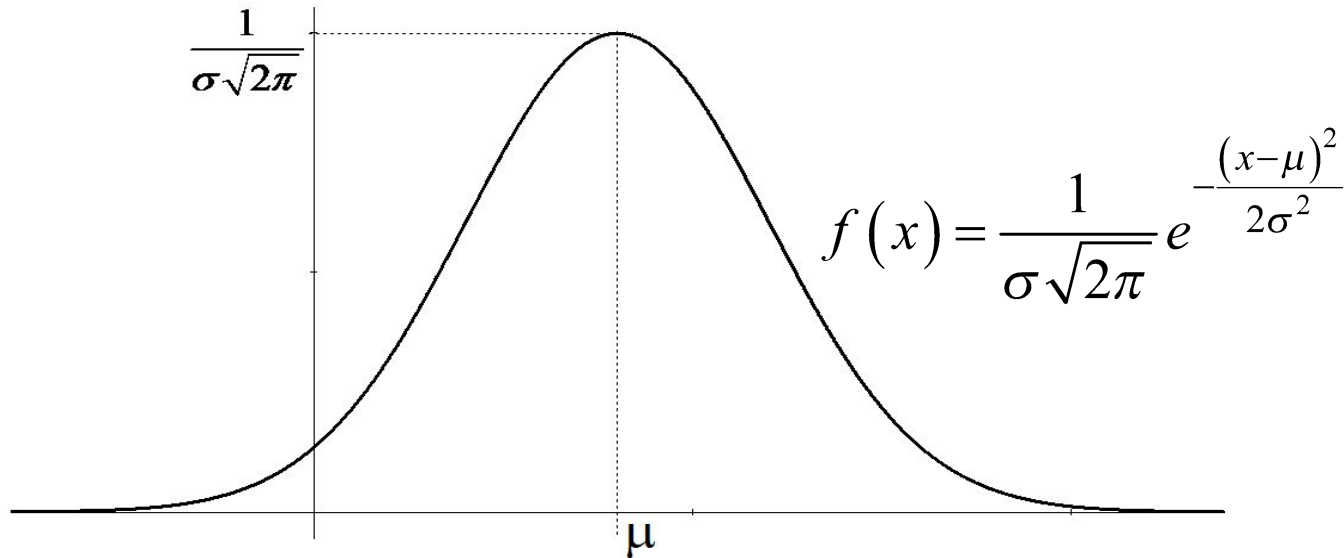
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Кривую нормального закона распределения называют нормальной или гауссовой кривой (гауссианой или гауссоидой). Гауссова кривая имеет симметричный колоколообразный вид с максимумом в точке $x=\mu$, причем сам максимум равен

$$\frac{1}{\sigma\sqrt{2\pi}}.$$



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Выясним смысл параметров нормального распределения. Для этого определим математическое ожидание и дисперсию СВ, распределенной нормально с параметрами μ и σ .

Вычислим математическое ожидание:

$$M[X] = \int_{-\infty}^{\infty} x \cdot f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Произведем замену переменных:

$$t = \frac{x - \mu}{\sigma\sqrt{2}}, \quad x = \sigma\sqrt{2}t + \mu, \quad dx = \sigma\sqrt{2}dt.$$



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Тогда

$$\begin{aligned} M[X] &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma\sqrt{2}t + \mu) \cdot e^{-t^2} \sigma\sqrt{2} dt = \\ &= \sigma\sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} t \cdot e^{-t^2} dt + \frac{\mu}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2} dt. \end{aligned}$$

Первый интеграл равен 0, поскольку подынтегральная функция является нечетной, а пределы интегрирования симметричны относительно начала координат.

Второй интеграл называется интегралом Пуассона-Эйлера; известно, что он равен $\sqrt{\pi}$.



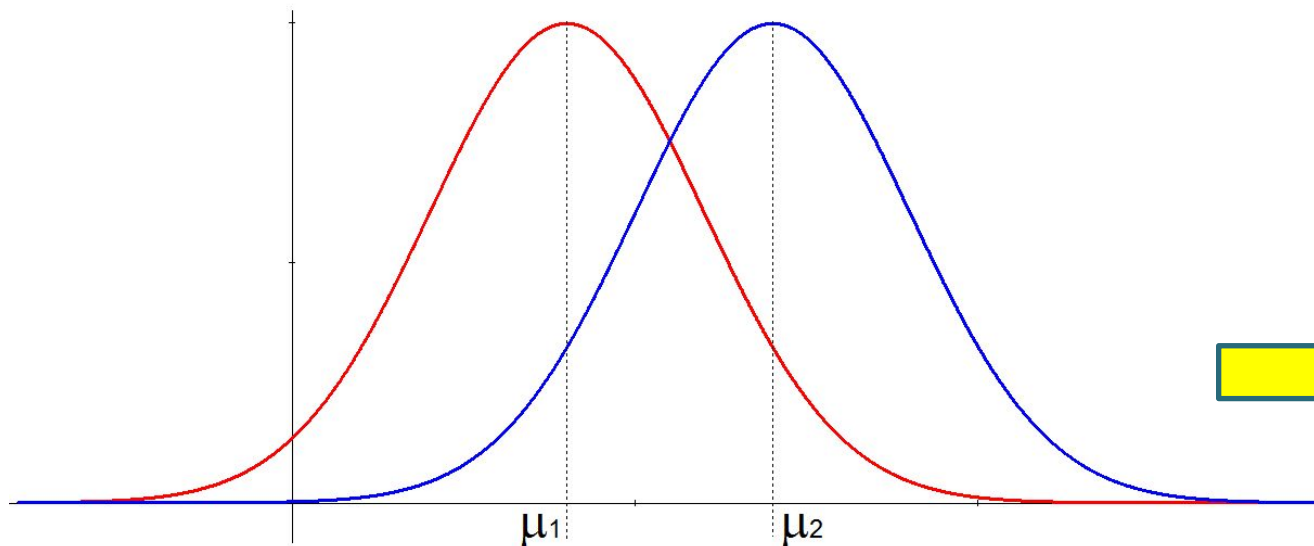
5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

С учетом этого

$$\begin{aligned} M[X] &= \sigma \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} t \cdot e^{-t^2} dt + \frac{\mu}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2} dt = \\ &= 0 + \frac{\mu}{\sqrt{\pi}} \sqrt{\pi} = \mu. \end{aligned}$$

Следовательно, параметр μ в функции плотности нормального распределения равен математическому ожиданию СВ.



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Вычислим теперь дисперсию:

$$\begin{aligned} D[X] &= \int_{-\infty}^{\infty} (x - M[X])^2 f(x) dx = \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x - \mu)^2 \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx. \end{aligned}$$

Произведя аналогичную замену переменных и взяв полученный интеграл по частям, в результате получим:

$$D[X] = \sigma^2.$$

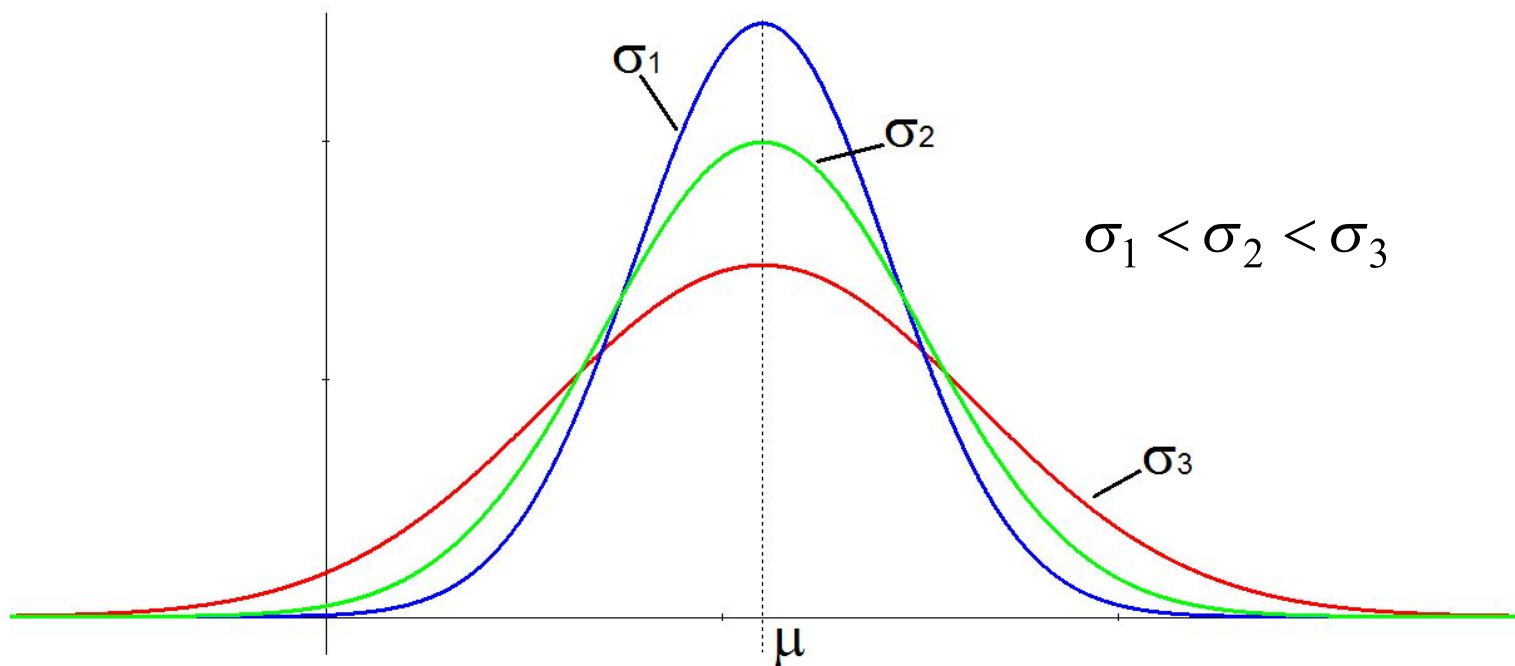


5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Следовательно, параметр σ нормального распределения есть ни что иное, как среднее квадратичное отклонение (квадратный корень из дисперсии).

Поскольку дисперсия есть мера рассеяния (разброса) СВ относительно мат. ожидания, параметр σ определяет, насколько сконцентрированы будут значения СВ около значения μ .



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Поскольку кривая плотности нормального распределения симметрична относительно мат. ожидания, коэффициент асимметрии нормального распределения равен 0:

$$Sk = 0.$$

Мы уже упоминали, что эксцесс (островершинность) распределений сравнивается с нормальным распределением, поэтому коэффициент эксцесса для нормального распределения равен 0:

$$Ex = 0.$$

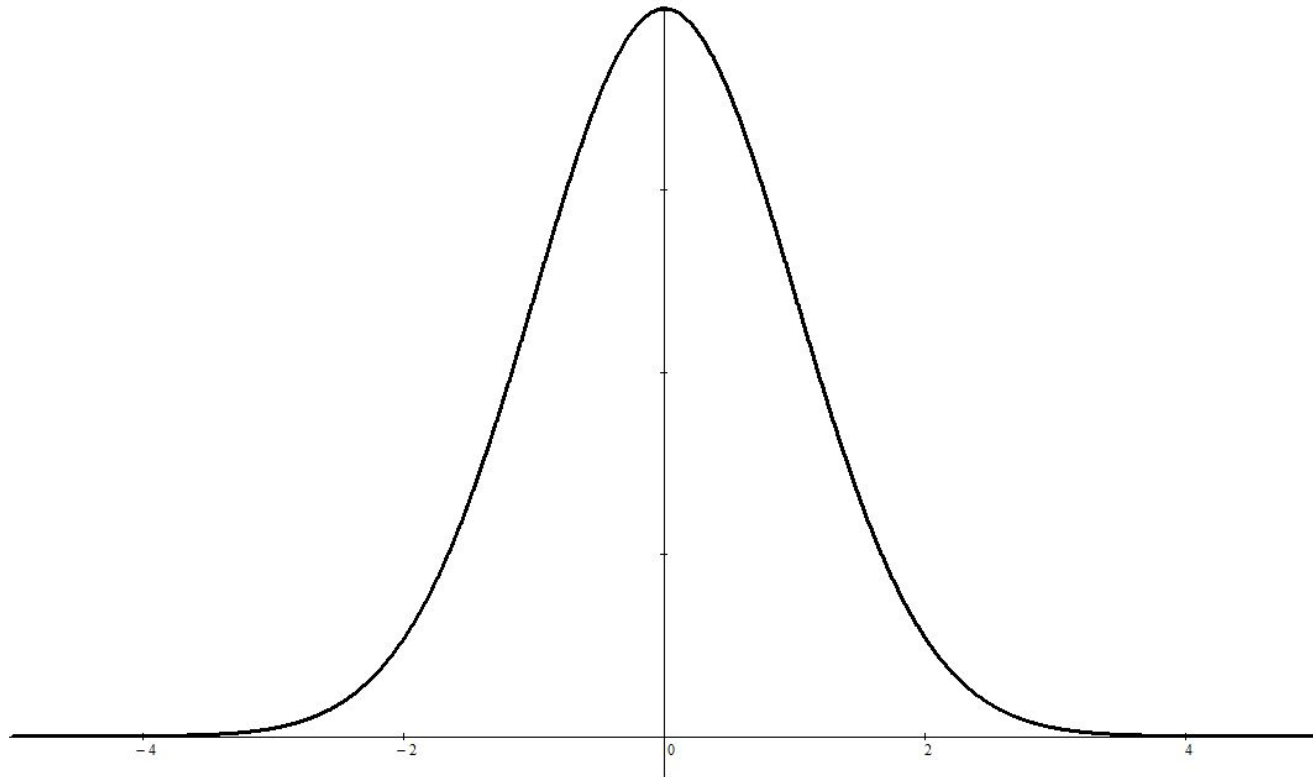
5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Краткое обозначение нормального распределения: $N(\mu, \sigma^2)$

Нормальное распределение $N(0, 1)$

называется стандартным нормальным распределением.



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Определим интегральную функцию нормального распределения:

$$F(x) = \int_{-\infty}^x f(t) dt = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Выполним замену переменной:

$$z = \frac{t - \mu}{\sigma}.$$

Тогда

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

Интеграл в этой формуле не выражается через элементарные функции



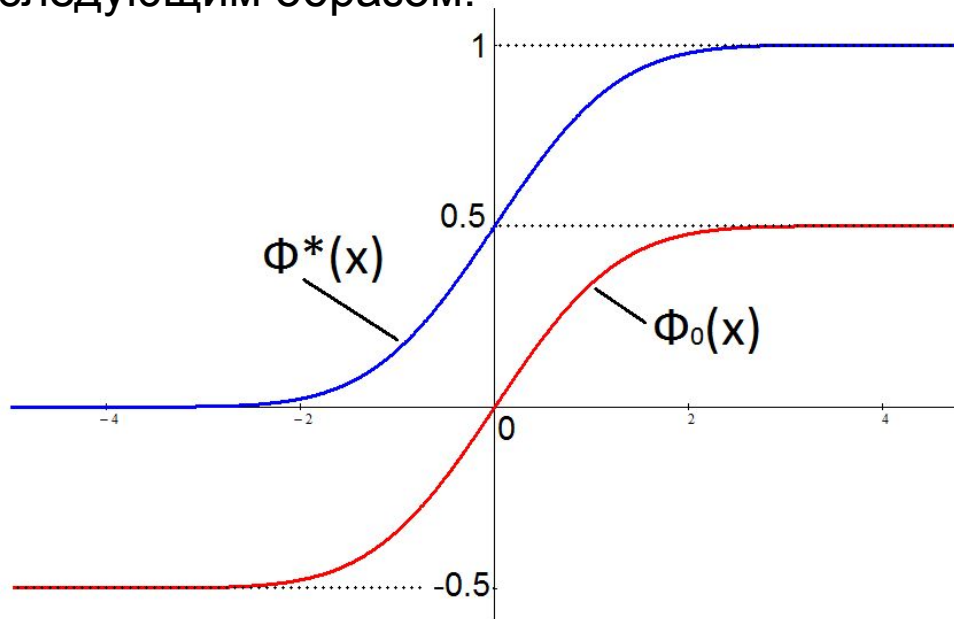
5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Для сравнения определим интегральную функцию стандартного нормального распределения:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz = \Phi^*(x).$$

Получившаяся функция называется (стандартной) нормальной функцией распределения. Она является табличной и связана с функцией Лапласа следующим образом:



$$\Phi^*(x) = 0.5 + \Phi_0(x).$$



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Стандартная нормальная функция распределения обладает следующими свойствами:

$$\Phi^* (-\infty) = 0;$$

$$\Phi^* (\infty) = 1;$$

$\Phi^* (x)$ — неубывающая функция;

$\Phi^* (-x) = 1 - \Phi^* (x)$, т.к. функция симметрична.

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Получается, что интегральную функцию нормального распределения с параметрами μ и σ можно выразить через стандартную функцию распределения:

$$F(x) = \Phi^* \left(\frac{x - \mu}{\sigma} \right).$$

Вероятность, что нормально распределенная случайная величина с параметрами μ и σ попадет на интервал $[\alpha, \beta]$:

$$\begin{aligned} P(\alpha \leq x \leq \beta) &= \frac{1}{\sigma \sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \\ &= \Phi^* \left(\frac{\beta - \mu}{\sigma} \right) - \Phi^* \left(\frac{\alpha - \mu}{\sigma} \right). \end{aligned}$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

На практике очень часто встречается задача вычисления вероятности попадания СВ X на участок симметричный относительно центра рассеивания μ . Рассмотрим такой участок длиной $2l$. Вычислим эту вероятность:

$$\begin{aligned} P(\mu - l \leq x \leq \mu + l) &= \Phi^* \left(\frac{\mu + l - \mu}{\sigma} \right) - \Phi^* \left(\frac{\mu - l - \mu}{\sigma} \right) = \\ &= \Phi^* \left(\frac{l}{\sigma} \right) - \Phi^* \left(\frac{-l}{\sigma} \right). \end{aligned}$$

С учетом свойств стандартной функции распределения:

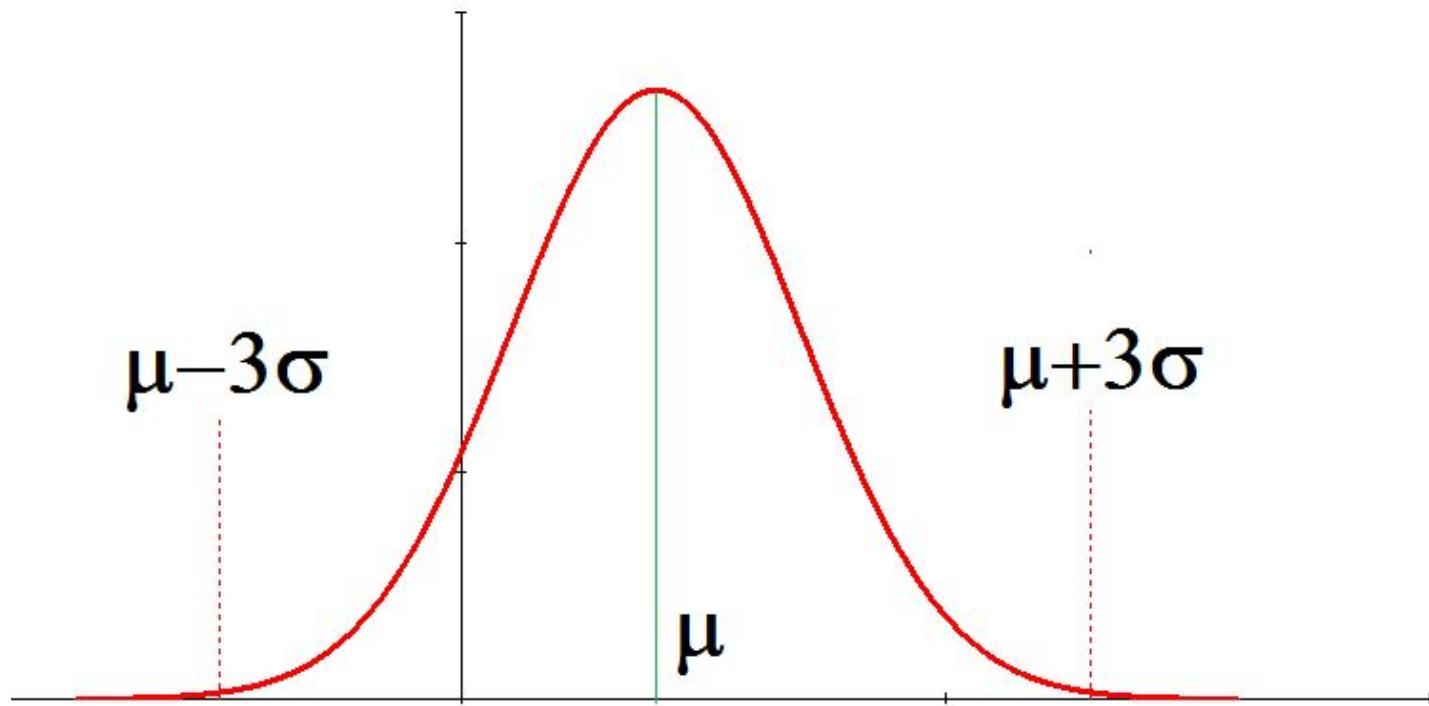
$$P(\mu - l \leq x \leq \mu + l) = 2\Phi^* \left(\frac{l}{\sigma} \right) - 1.$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Для нормального распределения характерно т.н. «правило трёх сигм»:

Значения нормально распределенной СВ X с параметрами μ и σ с вероятностью 0.9973 попадут в интервал $[\mu - 3\sigma; \mu + 3\sigma]$.



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Пример 8

Полагаем, что рост студентов – нормально распределенная случайная величина X с параметрами $\mu=173$ см и $\sigma^2=36$. Необходимо найти:

а) выражение плотности вероятности и функции распределения СВ;

б) доли костюмов 4-го роста (176–182 см) и 3-го роста(170–176 см), которые нужно предусмотреть в общем объеме производства;

в) квантиль $x_{0,7}$ и 10%-ную точку СВ.

Решение:

Функция плотности распределения:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{6\sqrt{2\pi}} e^{-\frac{(x-173)^2}{72}}.$$



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Пример 8 (продолжение)

Функция распределения:

$$\begin{aligned} F(x) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = \frac{1}{6\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-173)^2}{72}} dt = \\ &= \Phi^*\left(\frac{x-173}{6}\right). \end{aligned}$$

Долю костюмов 4-го роста (176–182 см) в общем объеме производства определим как

$$\begin{aligned} P(176 \leq x \leq 182) &= \Phi^*\left(\frac{182-173}{6}\right) - \Phi^*\left(\frac{176-173}{6}\right) = \\ &= \Phi^*(1.5) - \Phi^*(0.5) = 0.9332 - 0.6915 = 0.2417. \end{aligned}$$



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Пример 8 (продолжение)

Долю костюмов 3-го роста (170–176 см) можно определить аналогичным образом, но, если учесть, что данный интервал симметричен относительно мат. ожидания μ , то

$$P(\mu - 3 \leq x \leq \mu + 3) = 2\Phi^*\left(\frac{3}{6}\right) - 1 = 0.383.$$

Определим квантиль $x_{0,7}$:

$$F(x_{0,7}) = \Phi^*\left(\frac{x_{0,7} - 173}{6}\right) = 0.7 \Rightarrow \frac{x_{0,7} - 173}{6} = 0.525 \Rightarrow \\ \Rightarrow x_{0,7} \approx 176.$$

Это означает, что 70% студентов имеют рост до 176 см.



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Пример 8 (продолжение)

10%-ная точка есть квантиль $x_{0,9}$. Выполняя аналогичные вычисления, получим

$$F(x_{0,9}) = \Phi\left(\frac{x_{0,9} - 173}{6}\right) = 0.9 \Rightarrow \frac{x_{0,9} - 173}{6} = 1.282 \Rightarrow \\ \Rightarrow x_{0,9} \approx 180.7$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Пример 9

Средняя стоимость ценной бумаги составляет 2000 руб., а среднее квадратичное отклонение равно 100 руб. Предполагается, что цена имеет нормальное распределение. Определить вероятность того, что в день покупки цена будет заключена в пределах от 1800 руб. до 2300 руб. Найти с надежностью 0,9 интервал Δ изменения цены бумаги, симметричный относительно математического ожидания.



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Нормальное распределение (Распределение Гаусса).

Пример 9 (продолжение)

Решение:

$$\begin{aligned} P(1800 \leq x \leq 2300) &= \Phi^* \left(\frac{2300 - 2000}{100} \right) - \Phi^* \left(\frac{1800 - 2000}{100} \right) = \\ &= \Phi^*(3) - \Phi^*(-2) = 0.99865 - 0.02275 = 0.9759. \end{aligned}$$

$$\Phi^* \left(\frac{\Delta}{100} \right) - \Phi^* \left(\frac{-\Delta}{100} \right) = 0.9 \Rightarrow 2\Phi^* \left(\frac{\Delta}{100} \right) - 1 = 0.9 \Rightarrow$$

$$\Rightarrow \Phi^* \left(\frac{\Delta}{100} \right) = 0.95 \Rightarrow \frac{\Delta}{100} \approx 1.645 \Rightarrow \Delta = 164.5$$

Значит стоимость ценной бумаги заключена в интервале (1835,5; 2164,5) рублей.

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Распределение χ^2 («хи-квадрат»).

Пусть z_1, z_2, \dots, z_k – совместно независимые случайные величины, распределенные нормально с параметрами $\mu=0$ и $\sigma^2=1$ (стандартное нормальное распределение).

Тогда случайная величина

$$X = z_1^2 + z_2^2 + \dots + z_k^2 = \sum_{i=1}^k z_i^2, \quad k \in \mathbb{N}^+,$$

имеет распределение χ^2 («хи-квадрат») с k степенями свободы.

Краткая форма записи:

$$\chi^2(k) \text{ или } \chi_k^2$$

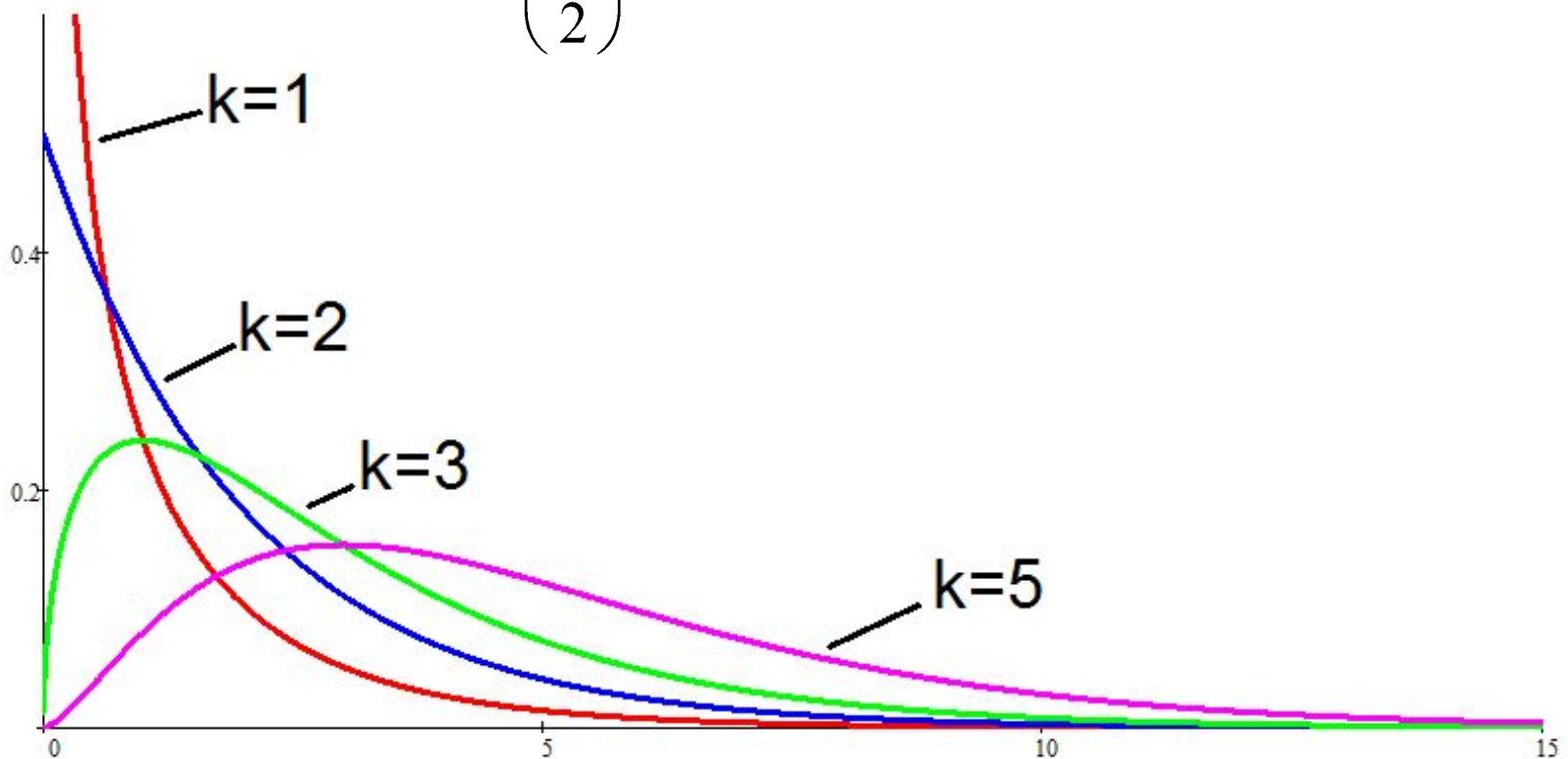
Распределение широко используется в статистике.

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Распределение χ^2 («хи-квадрат»).

Функция плотности распределения:

$$f(x) = \frac{1}{2^{0.5k} \Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}, \quad x \geq 0, k \in \mathbb{N}^+.$$



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Распределение χ^2 («хи-квадрат»).

В знаменателе функции плотности распределения стоит т.н. «гамма-функция» - не являющаяся элементарной, которая равна

$$\Gamma(n) = \int_0^{\infty} t^{n-1} e^{-t} dt, \quad n > 0.$$

Гамма-функция является табличной, однако полезно запомнить следующие ее свойства:

$$\Gamma(n) = (n-1)!, \quad n \in \mathbb{N}^+, n \neq 0;$$

$$\Gamma(n+1) = n\Gamma(n);$$

$$\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)!}{4^n n!} \sqrt{\pi}, \quad n \in \mathbb{N}^+.$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Распределение χ^2 («хи-квадрат»).

Вывод показателей распределения достаточно сложен, поэтому приведем значения этих показателей без доказательства.

Математическое ожидание:

$$M[X] = k$$

Дисперсия:

$$D[X] = 2k$$

Медиана:

$$Me(X) \approx k \left(1 - \frac{2}{9k} \right)^3$$

Мода:

$$Mo(X) = \max \{ 0, k - 2 \}$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Распределение χ^2 («хи-квадрат»).

Коэффициент асимметрии:

$$Sk = \sqrt{\frac{8}{k}}$$

Коэффициент эксцесса:

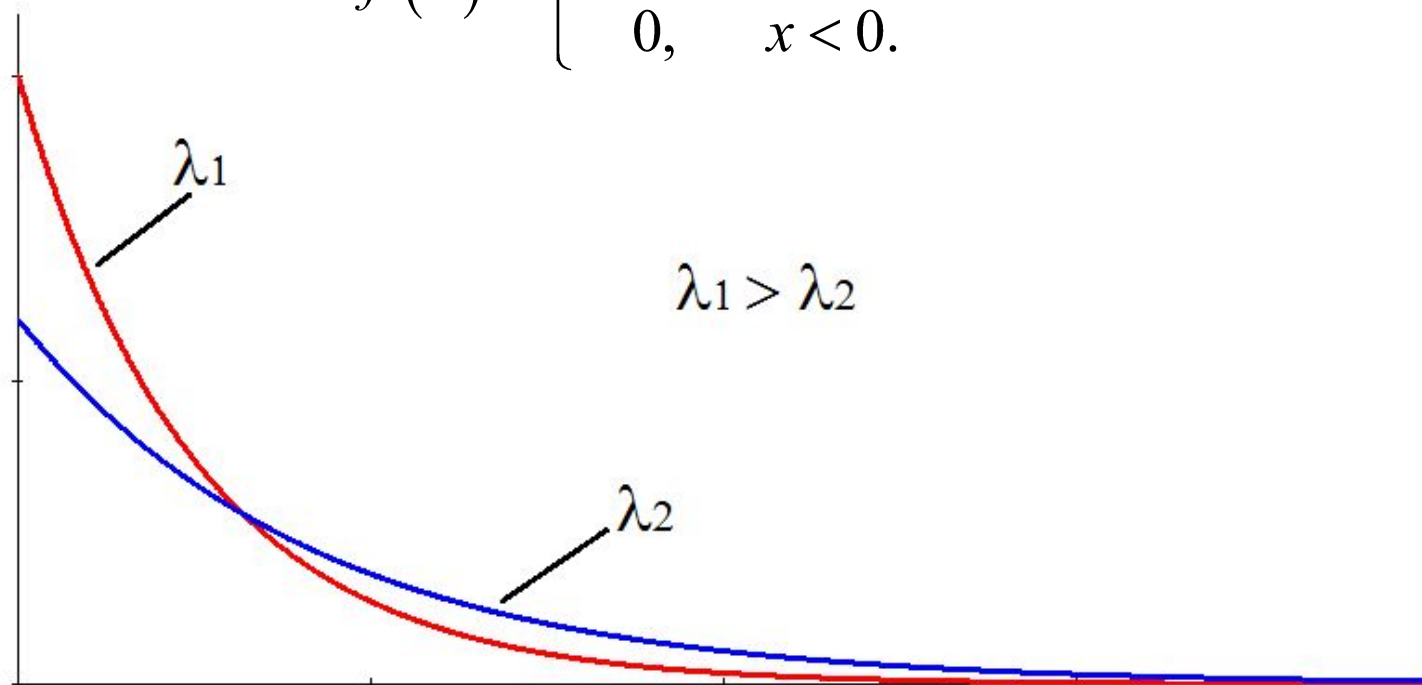
$$Ex = \frac{12}{k}$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Экспоненциальное (показательное) распределение

Непрерывная случайная величина X имеет показательный (экспоненциальный) закон распределения с параметром $\lambda > 0$, если ее плотность вероятности имеет вид:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

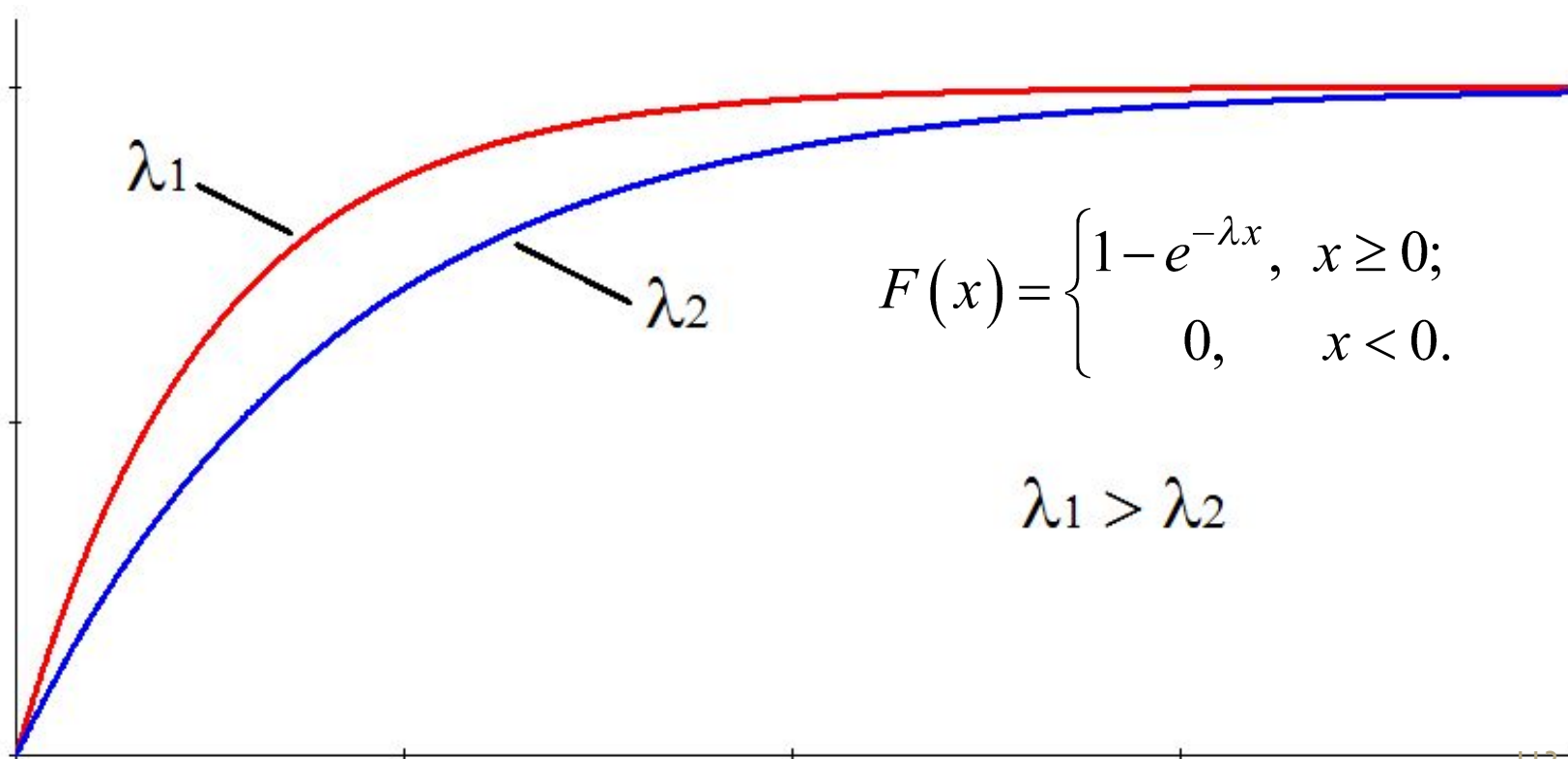


5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Экспоненциальное (показательное) распределение

Найдем интегральную функцию показательного распределения:

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^0 0 dt + \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}.$$



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Экспоненциальное (показательное) распределение

Основные показатели экспоненциального распределения:

$$M[X] = \frac{1}{\lambda} \qquad D[X] = \frac{1}{\lambda^2}$$

$$Me(X) = \frac{\ln 2}{\lambda} \qquad Mo(X) = 0$$

$$Sk = 2$$

$$Ex = 6$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Экспоненциальное (показательное) распределение

Пример 10:

Найти вероятность попадания экспоненциально распределенной СВ с параметром $\lambda=0.25$ в интервал $[a, 2]$.

Решение:

Известно, что

$$P(a \leq x \leq b) = F(b) - F(a).$$

Поэтому

$$\begin{aligned} P(a \leq x \leq 2) &= F(2) - F(a) = 1 - e^{-0.25 \cdot 2} - F(a) = \\ &= 0.3935 - F(a). \end{aligned}$$



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Экспоненциальное (показательное) распределение

Пример 10 (продолжение):

В свою очередь,

$$F(a) = \begin{cases} 0, & a \leq 0; \\ 1 - e^{-\lambda a}, & a > 0. \end{cases}$$

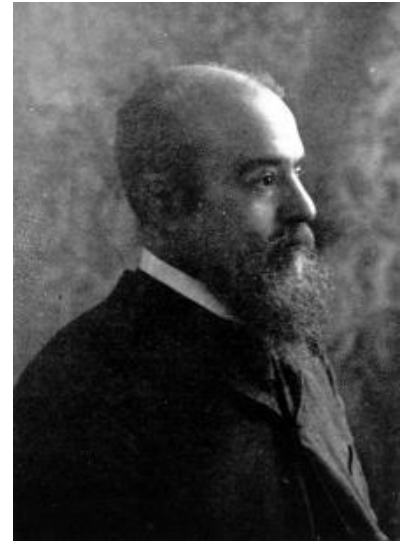
Поэтому

$$P(a \leq x \leq 2) = \begin{cases} 0.3935, & a \leq 0; \\ e^{-0.25a} - 0.6065, & a > 0. \end{cases}$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Распределение Парето

В практических задачах встречаются, так называемые, усеченные распределения, у которых из общего множества значений СВ устранены значения, большие или меньшие некоторого порогового уровня C_0 . В частности, такое распределение будет иметь заработная плата работника при условии, что ее значение не может быть меньше некоторой заданной величины.



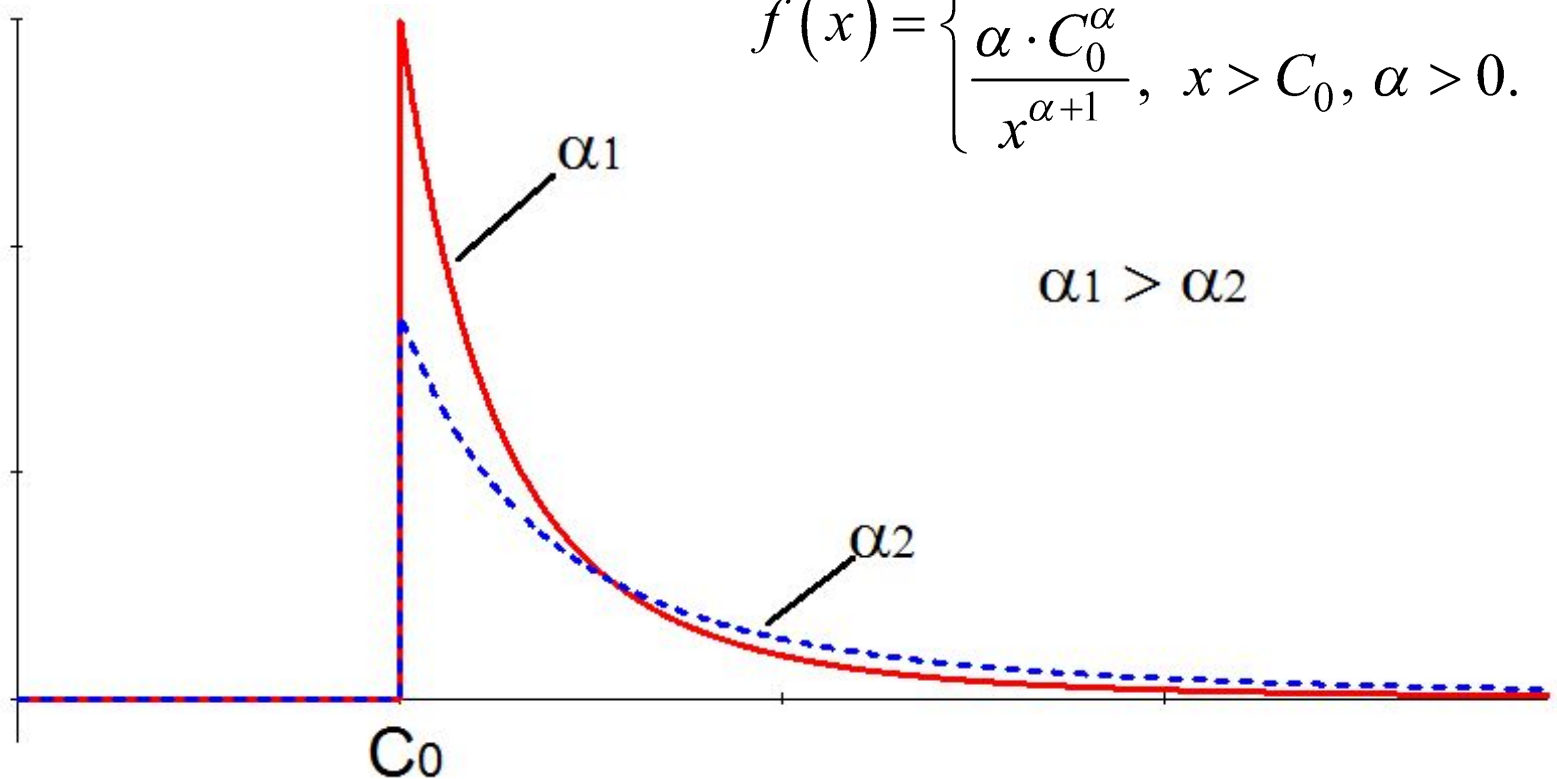
Вильфредо Парето
1848-1923

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Распределение Парето

Распределением Парето называется такое распределение, для которого плотность распределения вероятностей имеет вид:

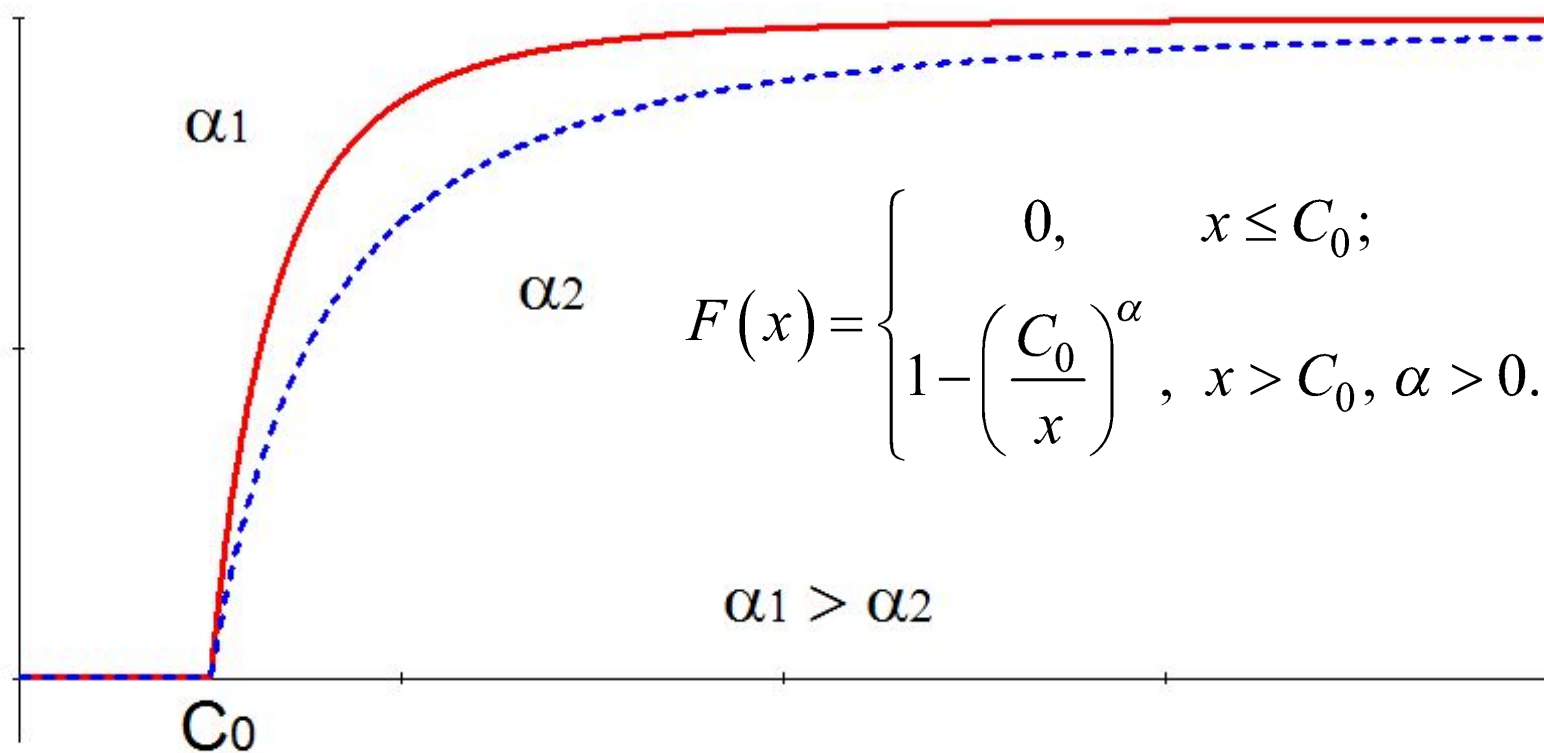
$$f(x) = \begin{cases} 0, & x \leq C_0; \\ \frac{\alpha \cdot C_0^\alpha}{x^{\alpha+1}}, & x > C_0, \alpha > 0. \end{cases}$$



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Распределение Парето

Интегральная функция распределения:



5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Распределение Парето

Основные показатели распределения:

$$M[X] = \frac{\alpha C_0}{\alpha - 1}, \quad \alpha > 1$$

$$Me(X) = C_0 \sqrt[\alpha]{2}$$

$$D[X] = \frac{\alpha C_0^2}{(\alpha - 1)^2 (\alpha - 2)}, \quad \alpha > 2$$

$$Mo(X) = C_0$$

$$Sk = \frac{2(1 + \alpha)}{\alpha - 3} \sqrt{\frac{\alpha - 2}{\alpha}}, \quad \alpha > 3$$

$$Ex = \frac{6(\alpha^3 + \alpha^2 - 6\alpha - 2)}{\alpha(\alpha - 3)(\alpha - 4)}, \quad \alpha > 4$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Распределение Парето

Пример 11:

Зарботная плата работника фирмы ограничена нижним пределом в размере 20 000 руб. и подчиняется закону Парето (x - зарботная плата в тысячах руб., $\alpha=3$). Необходимо записать плотность распределения СВ X , найти математическое ожидание уровня зарботной платы и ее среднее квадратическое отклонение.

Решение:

Плотность распределения:

$$f(x) = \frac{\alpha \cdot C_0^\alpha}{x^{\alpha+1}} = \frac{3 \cdot 20^3}{x^4} = \frac{24000}{x^4}.$$

5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Распределение Парето

Пример 11 (продолжение):

Математическое ожидание:

$$M[X] \text{ руб} = \frac{\alpha C_0}{\alpha - 1} = \frac{3 \cdot 20}{2} = 30 \text{ (руб.)}$$

Дисперсия:

$$D[X] = \frac{\alpha C_0^2}{(\alpha - 1)^2 (\alpha - 2)} = \frac{3 \cdot 20^2}{2^2 \cdot 1} = 300.$$

Среднее квадратическое отклонение:

$$\sigma = \sqrt{D[X]} \text{ руб} = \sqrt{300} \approx 17.3 \text{ (руб.)}$$

6. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Закон больших чисел

Свойство устойчивости массовых случайных явлений известно человечеству еще с глубоких времен. В какой бы области оно не проявлялось, суть его сводится к следующему: конкретные особенности каждого отдельного случайного явления почти не сказывается на среднем результате массы таких явлений. Именно эта устойчивость средних и представляет собой физическое содержание «закона больших чисел», понимаемого в **широком** смысле слова:

При очень большом числе случайных явлений средний их результат практически перестает быть случайным и может быть предсказан с большой степенью определенности.

6. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Закон больших чисел

Под законом больших чисел в узком смысле понимается ряд математических теорем, в каждой из которых при соблюдении определенных условий устанавливается факт приближения средних характеристик большого числа испытаний к некоторым определенным постоянным.

Различные формы закона больших чисел вместе с различными формами центральной предельной теоремы образуют совокупность так называемых предельных теорем теории вероятностей.

Предельные теоремы дают возможность не только осуществить научные прогнозы в области случайных явлений, но и оценивать точность этих прогнозов.

Неравенство Чебышева

Пусть имеется случайная величина X с мат. ожиданием m_X и дисперсией D_X :

$$X \Rightarrow (m_X, D_X)$$

Неравенство Чебышева утверждает, что каково бы ни было положительное число ε , вероятность того, что СВ X отклонится от своего математического ожидания не меньше, чем на ε , ограничена сверху величиной D_X / ε^2 :

$$P(|X - m_X| > \varepsilon) \leq \frac{D_X}{\varepsilon^2}.$$

Неравенство Чебышева дает **грубую оценку сверху** и утверждает, что для любой случайной величины X вероятность того, что она отклонится на ε от m_X меньше, чем дисперсия, деленная на ε^2 .



Пафнутий Львович
Чебышев
1821-1894

Неравенство Чебышева

Пример 12

Дана случайная величина X с мат. ожиданием m и дисперсией $D=\sigma^2$. Оценим сверху вероятность того, что X отклонится от мат. ожидания не менее, чем на 3σ .

Решение

Согласно неравенству Чебышева запишем:

$$P(|X - m_X| > 3\sigma) \leq \frac{\sigma^2}{9\sigma^2} = \frac{1}{9} \approx 0.111.$$

Известно, что для случайной величины, распределенной по нормальному закону, эта вероятность примерно равна 0.0027 (правило трёх сигма).

6. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Теорема Чебышева

Пусть имеется случайная величина X с мат. ожиданием m_X и дисперсией D_X :

$$X \Rightarrow (m_X, D_X)$$

Над этой величиной производится n независимых опытов и вычисляется среднее арифметическое всех наблюдаемых значений случайной величины X . Необходимо найти характеристики среднего арифметического – математическое ожидание и дисперсию. В результате первого опыта СВ X приняла значение x_1 , во втором опыте – x_2 , ..., в n -м опыте – x_n .



Теорема Чебышева

Рассмотрим среднее арифметическое этих значений:

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

СВ Y - линейная функция независимых СВ x_1, x_2, \dots, x_n .

$$M[Y] = \frac{1}{n} \sum_{i=1}^n M[x_i] = \frac{1}{n} n \cdot m_X = m_X.$$

$$D[Y] = \frac{1}{n^2} \sum_{i=1}^n D[x_i] = \frac{1}{n^2} n \cdot D_X = \frac{D_X}{n}.$$



6. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Теорема Чебышева

Таким образом $M[Y]$ не зависит от числа опытов (n), а дисперсия при больших n может стать сколь угодно малой, то есть СВ Y ведет себя почти не как случайная. Это свойство и устанавливает теорема Чебышева.

При достаточно большом числе независимых опытов среднее арифметическое наблюдаемых значений случайной величины сходится по вероятности к ее математическому ожиданию:

$$P\left(\left|\frac{1}{n}\sum_{i=1}^n M[x_i] - m_X\right| < \varepsilon\right) > 1 - \delta,$$

где ε и δ – произвольно малые положительные числа.

6. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Теорема Маркова

Если случайная величина X принимает только неотрицательные значения и имеет математическое ожидание m_X , то для любого положительного числа A верно неравенство:

$$P(x > A) \leq \frac{m_X}{A}.$$



Андрей Андреевич
Марков
1856-1922

Теорема Бернулли

Пусть производится n независимых опытов, в каждом из которых может произойти или не произойти событие A , вероятность которого в каждом опыте равна p . Теорема Бернулли утверждает, что:

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

Обозначим частоту события A в n опытах через p^* и запишем теорему Бернулли в виде формулы:

$$P(|p^* - p| < \varepsilon) > 1 - \delta,$$

где ε и δ – произвольно малые положительные числа.

6. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Теорема Бернулли

Общий смысл закона больших чисел — совместное действие большого числа одинаковых и независимых случайных факторов **приводит к результату, в пределе не зависящему от случая.**

На этом свойстве основаны методы оценки вероятности на основе анализа конечной выборки. Наглядным примером является прогноз результатов выборов на основе опроса выборки избирателей.

6. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Центральная предельная теорема

Закон больших чисел устанавливает факт приближения средних большого числа случайных величин к определенным постоянным. Но этим не ограничиваются закономерности, возникающие в результате суммарного действия случайных величин. Оказывается, что при некоторых условиях **совокупное действие случайных величин приводит** к определенному, а именно – к **нормальному закону распределения**.

Центральная предельная теорема представляет собой группу теорем, посвященных установлению условий, при которых возникает нормальный закон распределения. Среди этих теорем важнейшее место принадлежит **теореме Ляпунова**.

6. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Теорема Ляпунова

Если X_1, X_2, \dots, X_n - независимые случайные величины, у каждой из которых существует математическое ожидание $M[X_i] = a_i$, дисперсия $D[X_i] = \sigma_i^2$, абсолютный центральный момент третьего порядка m_i и при этом

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n m_i}{\left(\sum_{i=1}^n \sigma_i^2\right)^{1.5}} = 0,$$

то закон распределения суммы $Y_n = X_1 + X_2 + \dots + X_n$ при $n \rightarrow \infty$ неограниченно приближается к нормальному закону с математическим ожиданием

$$\sum_{i=1}^n a_i$$

и дисперсией

$$\sum_{i=1}^n \sigma_i^2.$$



Александр
Михайлович Ляпунов
1857-1918

Центральная предельная теорема

Если случайные величины X_1, X_2, \dots, X_n имеют мат. ожидания, то мат. ожидание их суммы существует и равно сумме их мат. ожиданий. Если X_1, X_2, \dots, X_n - случайные величины с конечными дисперсиями, то

$$\text{Var}(X_1 + \dots + X_n) = \sum_{i=1}^n \sigma_i^2 + 2 \sum_{j,k} \text{Cov}(X_j, X_k),$$

причем последняя сумма состоит из C_n^2 пар (X_j, X_k) , $j < k$.

Если X_1, X_2, \dots, X_n - независимые одинаково распределенные случайные величины, то

$$\sum_{i=1}^n a_i = na; \quad \sum_{i=1}^n \sigma_i^2 = n\sigma^2.$$

Центральная предельная теорема (т. Ляпунова [М.Э.- 1982-т. 3- с. 470])

Пусть независимые случайные величины

$$X_1, \dots, X_i, \dots$$

имеют конечные мат. ожидания MX_i

, дисперсии DX_i

и абсолютные моменты $M|X_i - MX_i|^{2+\delta}$, $\delta > 0$,

и пусть $B_n = \sum_{i=1}^n DX_i$ – дисперсия суммы X_1, \dots, X_n .

Тогда если при некотором

$\delta > 0$ выполняется условие

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n M |X_i - MX_i|^{2+\delta}}{B_n^{1+0,5\delta}} = 0, \quad (1)$$

то вероятность неравенства

$$x_1 < \frac{\sum_{i=1}^n (X_i - MX_i)}{\sqrt{B_n}} < x_2 \quad (2)$$

стремится при $n \rightarrow \infty$ к пределу

$$\frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-x^2/2} dx \quad (3)$$

равномерно относительно всех значений x_1 и x_2 .

Центральная предельная

теорема

Если неформально, то классическая центральная предельная теорема (т. Ляпунова) утверждает, что сумма n независимых одинаково распределённых случайных величин X_1, \dots, X_n имеет распределение близкое к $N(na, n\sigma^2)$.

Эквивалентно $Y = \frac{1}{n} \sum_{i=1}^n X_i$, имеет распределение

близкое к $N(a, \sigma^2 / n)$.

Так как функция распределения стандартного нормального распределения непрерывна, сходимость к этому распределению эквивалентна поточечной сходимости функций распределения к функции распределения стандартного нормального распределения.

Положив $Z_n = \frac{S_n - an}{\sigma\sqrt{n}}$, получаем,

$$F_{Z_n}(x) \rightarrow \Phi(x), \quad \forall x \in \mathbb{R}$$

где $\Phi(x)$

функция распределения стандартного нормального распределения.

Центральная предельная теорема (т. Линдеберга [В. Феллер, том I, с. 258])

Пусть $\{X_i\}$ - последовательность
независимых одинаково

распределенных случайных величин.

Предположим, что их математические
ожидания a и дисперсии σ^2 существуют,

и положим $S_n = X_1 + \dots + X_n$. Тогда

для

β

любого фиксированного

$$P\left\{\frac{S_n - na}{\sigma\sqrt{n}} < \beta\right\} \rightarrow N(\beta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta} e^{-t^2/2} dt.$$

Теорема Ляпунова

Пример 13

Определить вероятность того, что средняя продолжительность 100 производственных операций окажется в пределах от 46 до 49 с, если математическое ожидание одной операции равно 47,4 с, а среднее квадратичное отклонение – 4,9 с.

Решение

В этой задаче X - продолжительность наугад взятой производственной операции, $a = M[X] = 47,4$ с, $\sigma = 4,9$ с.



Теорема Ляпунова

Пример 13 (продолжение)

Средняя продолжительность 100 наугад взятых производственных операций

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i,$$

причем

$$M[\bar{X}] = 47.4 \text{ с}, \quad D[\bar{X}] = \frac{\sigma^2}{n} = \frac{4.9^2}{100} \text{ с}^2.$$

Вычислим теперь соответствующую вероятность:

$$\begin{aligned} P(46 \leq \bar{X} \leq 49) &= P\left(-\frac{1.4}{0.49} \leq \frac{\bar{X} - 47.4}{0.49} \leq \frac{1.6}{0.49}\right) = \\ &= P(-2.857 \leq Z \leq 3.265) = \\ &= \Phi^*(3.265) - (1 - \Phi^*(2.857)) = 0.9973. \end{aligned}$$



Теорема Ляпунова

Пример 13 (продолжение)

$$\begin{aligned} P(46 \leq \bar{X} \leq 49) &= P\left(-\frac{1.4}{0.49} \leq \frac{\bar{X} - 47.4}{0.49} \leq \frac{1.6}{0.49}\right) = \\ &= P(-2.857 \leq Z \leq 3.265) = \\ &= \Phi^*(3.265) - (1 - \Phi^*(2.857)) = 0.9973. \end{aligned}$$

При решении задачи на основании следствия из теоремы Ляпунова распределение централизованного и нормированного среднего приближенно заменено на распределение стандартной нормальной случайной величины $Z \sim N(0, 1)$.