Случайные величины

Основные законы распределения

Случайные величины

Случайной называется величина, которая в результате испытания может принять то или иное значение, заранее не известное, и зависящее от случайных причин, которые заранее не могут быть учтены.

Формально:

случайная величина $\boldsymbol{X}$ – это числовая функция $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{\omega})$, определенная на множестве элементарных событий Ω, $\omega \in \Omega$.

Функция $\boldsymbol{X}(\boldsymbol{\omega})$ должна быть такова, чтобы для любого $\boldsymbol{x}$ событие $\boldsymbol{X}(\boldsymbol{\omega}) < \boldsymbol{x}$ принадлежало полю событий

была определена вероятность этого события.

Случайные величины:

- дискретные;
- непрерывные.

Дискретные случайные величины

Дискретной называется случайная величина, которая принимает отдельные, изолированные возможные значения с определенными вероятностями.

Более строго:
случайная величина $\boldsymbol{X}$ называется *дискретной*, если существует конечное или счетное множество чисел $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ \dots\ ,$ таких что

$$\boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x}_i) = \boldsymbol{p}_i \geq 0, \quad \boldsymbol{i} = 1, 2, 3, \dots ,$$

$$\boldsymbol{p}_1 + \boldsymbol{p}_2 + \boldsymbol{p}_3 + \ \dots\ = 1.$$

Законом распределения дискретной случайной величины называется соответствие между возможными значениями этой случайной величины и их вероятностями.

Способы задания закона распределения:

- табличный (ряд распределения);
- графический (многоугольник распределения);
- аналитический;

Только для дискретных СВ

- функция распределения (интегральная функция).

Для дискретных и непрерывных СВ

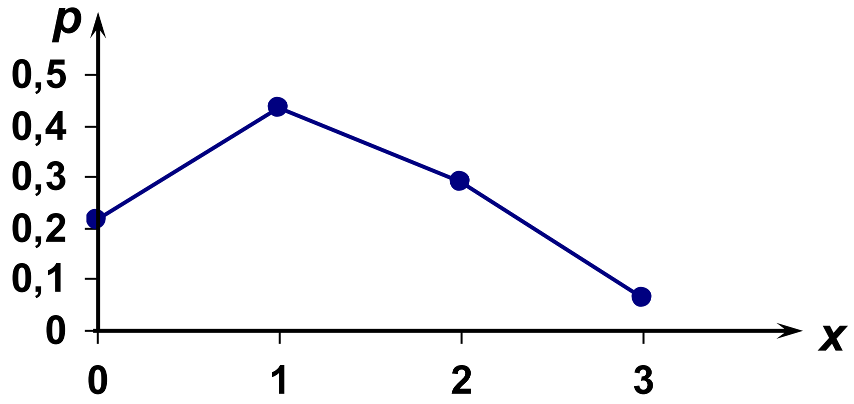
Пример.

Случайная величина ***X*** – число попаданий в мишень при условии, что сделано 3 выстрела, вероятность попадания в каждом выстреле равна 0,4.

- Ряд распределения.

X	0	1	2	3
p	0,216	0,432	0,288	0,064

- Многоугольник распределения.



- Аналитическое задание.

$$P(X=i) = C_3^i (0{,}4)^i (1-0{,}4)^{3-i}, \quad i = 0, 1, 2, 3.$$

Функцией распределения (*интегральной функцией*) случайной величины X называется функция $F_X(x)$:

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

Основные свойства функции распределения:

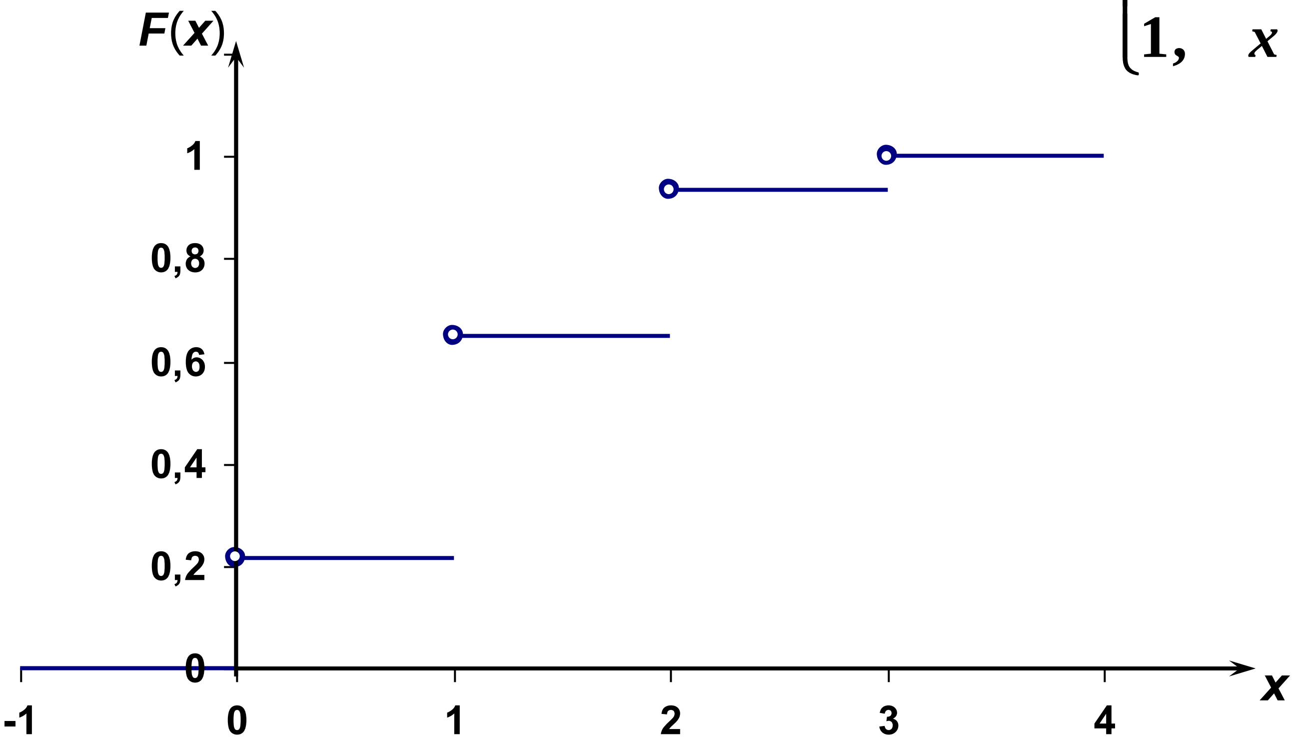
1) для любого x $0 \le F(x) \le 1$;
2) из $x_1 < x_2$ следует $F(x_1) \le F(x_2)$;
3) $\lim\limits_{x \to -\infty} F(x) = 0, \quad \lim\limits_{x \to +\infty} F(x) = 1;$
4) для любого x $\lim\limits_{x \to -0} F(x + \Delta x) = F(x).$ **Непрерывность слева в каждой точке**

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

Пример.

Функция распределения случайной величины X из предыдущего примера.

$$F(x) = \begin{cases} 0, & x \le 0; \\ 0{,}216; & 0 < x \le 1; \\ 0{,}648; & 1 < x \le 2; \\ 0{,}936; & 2 < x \le 3; \\ 1, & x > 3. \end{cases}$$



Непрерывные случайные величины

Непрерывной называется случайная величина, функция распределения которой всюду непрерывна.

Непрерывная СВ может принимать все значения из некоторого конечного или бесконечного промежутка.

Для непрерывной случайной величины $\boldsymbol{X}$:

- для любого $\boldsymbol{x_0}$ $\boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x_0}) = 0$;
- $\boldsymbol{P}(\boldsymbol{x_1} \le \boldsymbol{X} < \boldsymbol{x_2}) = \boldsymbol{P}(\boldsymbol{x_1} < \boldsymbol{X} < \boldsymbol{x_2}) = \boldsymbol{P}(\boldsymbol{x_1} < \boldsymbol{X} \le \boldsymbol{x_2}) =$
  $= \boldsymbol{P}(\boldsymbol{x_1} \le \boldsymbol{X} \le \boldsymbol{x_2})$.

Формы задания закона распределения непрерывной случайной величины:

- функция распределения (интегральная функция);

Для дискретных и непрерывных СВ

- плотность распределения вероятностей (дифференциальная функция).

Только для непрерывных СВ

Пусть функция $\boldsymbol{F(x)}$ дифференцируема.

Функция

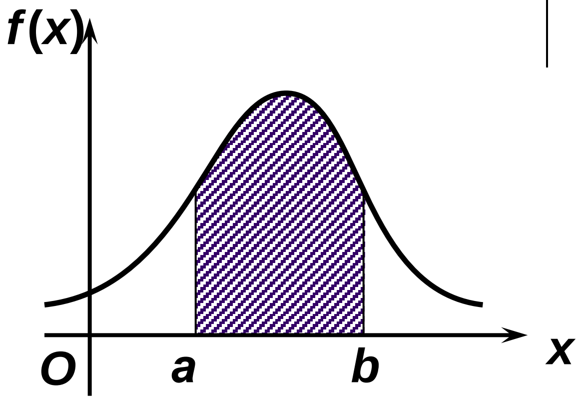
$$\boldsymbol{f(x) = F'(x)}$$

называется *плотностью распределения вероятностей* (*плотностью вероятности*, *дифференциальной функцией*).

График функции $\boldsymbol{f(x)}$ называется кривой распределения.

Для непрерывной СВ ***X***:

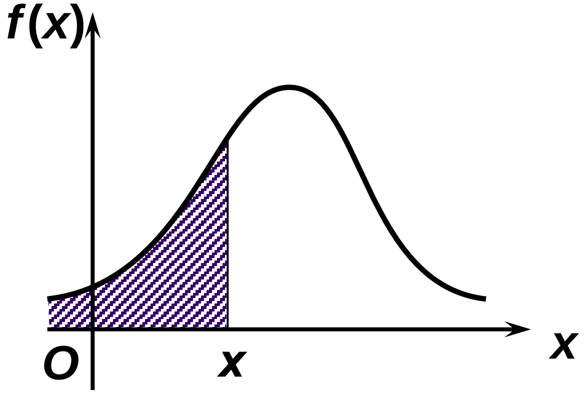
$$P\left(x \in \langle a,\ b\rangle\right) = \int_a^b f(x)dx.$$



Если известна функция ***f***(***x***), то

$$F(x) = P(X < x) = P\left(X \in (-\infty,\ x)\right) =$$

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



Свойства плотности распределения:

1) ***f***(***x***) ≥ 0 для любого ***x***;

2) $$\int\limits_{-\infty}^{+\infty} f(x)dx = 1.$$

В частности, если все возможные значения СВ ***X*** принадлежат промежутку < ***a***, ***b*** >, то

$$\int\limits_{a}^{b} f(x)dx = 1.$$

Значение x_p, для которого выполняется равенство

$$F(x_p) = p,$$

называется *квантилью*, *отвечающей заданному уровню вероятности* p (или $p\cdot$*100%-ной квантилью*).

Для непрерывной СВ квантили x_p существуют для любых p, $0 < p < 1$.

Обычно используются квантили x_p для значений p

0,01; 0,05; 0,1; 0,9; 0,95; 0,99.

Важнейшие числовые характеристики случайных величин

Закон распределения рассматриваемой случайной величины не всегда известен.

Во многих задачах достаточно знать отдельные числовые параметры, характеризующие существенные свойства случайной величины – ее числовые характеристики.

- **Математическое ожидание (среднее значение) СВ.**

  Является «представителем» СВ и заменяет ее в грубо ориентировочных расчетах.

  Обозначение: ***M*(*X*)**.

Для дискретной СВ

$M(X)$ существует, если ряд сходится абсолютно

$$M(X)=\sum_{i=1}^{n}x_i\cdot p_i \quad \text{или} \quad M(X)=\sum_{i=1}^{\infty}x_i\cdot p_i.$$

Для непрерывной СВ

$M(X)$ существует, если интеграл сходится абсолютно

$$M(X)=\int_{-\infty}^{+\infty}x\cdot f(x)dx$$

или $$M(X)=\int_{a}^{b}x\cdot f(x)dx$$

Если все возможные значения *X* принадлежат промежутку < *a*, *b* >

Размерность $M(X)$ совпадает с размерностью случайной величины X.

Основные свойства математического ожидания:

C – постоянная (неслучайная) величина

- $M(C) = C$;
- $M(X+C) = M(X) + C$;
- $M(C \cdot X) = C \cdot M(X)$;
- если $M(X)$ и $M(Y)$ существуют, то
$$M(X+Y) = M(X) + M(Y);$$
- если случайные величины X и Y независимы и $M(X)$ и $M(Y)$ существуют, то
$$M(X \cdot Y) = M(X) \cdot M(Y).$$

Закон распределения одной из них не зависит от того, какие возможные значения приняла другая

- **Дисперсия СВ**.

Характеризует степень рассеянности значений СВ вокруг математического ожидания.

Обозначение: ***D***(***X***).

$$D(X) = M(X - M(X))^2.$$

Размерность ***D***(***X***) равна квадрату размерности СВ.

Теорема.

$$D(X) = M(X^2) - (M(X))^2.$$

Для дискретной СВ

$$D(X) = \sum_{i=1}^{n} x_i^2 \cdot p_i - (M(X))^2 \quad \text{или} \quad D(X) = \sum_{i=1}^{\infty} x_i^2 \cdot p_i - (M(X))^2.$$

Для непрерывной СВ

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

или

$$D(X) = \int_{a}^{b} x^2 \cdot f(x)dx - (M(X))^2$$

Если все возможные значения *X* принадлежат промежутку < *a*, *b* >

Основные свойства дисперсии:

C – постоянная (неслучайная) величина

- $D(C) = 0$;
- $D(X+C) = D(X)$;
- $D(C \cdot X) = C^2 \cdot D(X)$;
- если случайные величины X и Y независимы, то

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

- если случайные величины X и Y независимы, то

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

- **Среднее квадратическое отклонение**.

  Характеризует степень рассеянности значений СВ вокруг математического ожидания.

  Обозначение: $\boldsymbol{\sigma}(\boldsymbol{X})$.

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

  Размерность $\boldsymbol{\sigma}(\boldsymbol{X})$ совпадает с размерностью СВ $\boldsymbol{X}$.

Основные законы распределения

Дискретные случайные величины

- **Биномиальный закон распределения.**

  СВ $\boldsymbol{X}$ имеет биномиальное распределение, если ее возможные значения

  $$x_1 = 0,\ x_2 = 1,\ x_3 = 2,\ \ldots,\ x_n = n-1,\ x_{n+1} = n,$$

  а соответствующие им вероятности определяются по формуле Бернулли:

  $$P(X = i) = C_n^i p^i q^{n-i}, \quad i = 0, 1, 2, \ldots, n,$$

  где $0 < p < 1$, $q = 1 - p$.

  $\boldsymbol{n}$ и $\boldsymbol{p}$ – параметры биномиального распределения.

Условия возникновения.

Пусть производится ***n*** независимых испытаний, в каждом из которых событие ***A*** появляется с одной и той же вероятностью ***p***.

СВ ***X*** – число появлений события ***A*** в ***n*** испытаниях – имеет биномиальное распределение с параметрами ***n*** и ***p***.

Числовые характеристики.

$$M(X) = np, \quad D(X) = npq,$$

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

- **Распределение Пуассона**.

  СВ $\boldsymbol{X}$ имеет распределение Пуассона, если ее возможные значения

$$x_1 = 0,\ \ x_2 = 1,\ \ x_3 = 2,\ \ \dots,$$

(бесконечное, но счетное множество значений),

а соответствующие им вероятности определяются формулой

$$P(X = k) = \frac{a^k}{k!}e^{-a}, \quad k = 0, 1, 2, \ \dots,$$

где $\boldsymbol{a} > 0$ – некоторое число, называемое параметром закона Пуассона.

Числовые характеристики.

$$M(X) = D(X) = a,$$

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

Характеристическое свойство распределения Пуассона

Условия возникновения.

1. Распределение Пуассона – предельное для биномиального, когда

$$\begin{cases} \boldsymbol{n} \to \infty, \\ \boldsymbol{p} \to 0, \end{cases} \quad \text{причем} \quad \lim_{\substack{n\to\infty \\ p\to 0}} np = a.$$

Другое название закона – *закон редких явлений*.

На практике распределение Пуассона с параметром $\boldsymbol{a} = \boldsymbol{np}$ может приближенно применяться вместо биномиального, когда число опытов $\boldsymbol{n}$ очень велико, а вероятность $\boldsymbol{p}$ наступления события $\boldsymbol{A}$ в каждом опыте очень мала.

2. Последовательность случайных моментов возникновения некоторых однородных событий называется *потоком событий*.

Поток событий может обладать следующими свойствами.

1) *Стационарность*.

Вероятность появления того или иного числа событий на интервале времени длины τ зависит только от *длины* этого интервала и не зависит от того, *где* именно на оси времени ***Ot*** расположен этот участок.

Среднее число событий, появляющихся в единицу времени, постоянно.
Обозначим его $\boldsymbol{\lambda}$ и будем называть *интенсивностью* (*плотностью*) потока.

2) *Ординарность*.

Вероятность появления на малом промежутке времени ***Δt*** двух и более событий пренебрежимо мала по сравнению с вероятностью появления на нем одного события.

Грубо говоря: события происходят поодиночке, а не парами, тройками, и т. д.

3) *Отсутствие последействия*.

Для любых неперекрывающихся промежутков времени число событий, попадающих на один из этих промежутков, не зависит от числа событий, попадающих на другие.

События наступают независимо друг от друга

Поток событий, обладающий свойствами стационарности, ординарности и отсутствия последействия, называется *простейшим* (*стационарным пуассоновским*) потоком.

Пусть имеется простейший поток событий с плотностью $\boldsymbol{\lambda}$.

СВ $\boldsymbol{X}$ – число событий, возникающих на промежутке времени длины $\boldsymbol{\tau}$, имеет распределение Пуассона с параметром $\boldsymbol{a} = \boldsymbol{\lambda\tau}$:

$$P(X = k) = \frac{(\lambda\tau)^k}{k!} \cdot e^{-\lambda\tau}, \quad k = 0, 1, 2, \ldots$$

- **Геометрическое распределение**.

  СВ $\boldsymbol{X}$ имеет геометрическое распределение, если ее возможные значения

  $$x_1 = 1,\ \ x_2 = 2,\ \ x_3 = 3,\ \ \ldots,$$

  (бесконечное, но счетное множество значений),

  а соответствующие им вероятности определяются формулой

  $$P(X = k) = p \cdot q^{k-1}, \quad k = 1, 2, \ \ldots,$$

  где $0 < p < 1$, $q = 1 - p$.

Условия возникновения.

Пусть производятся независимые испытания, в каждом из которых событие ***A*** появляется с одной и той же вероятностью ***p***.

СВ ***X*** – число испытаний, которые нужно провести до первого появления события ***A*** – имеет геометрическое распределение с параметром ***p***.

Числовые характеристики.

$$M(X) = \frac{1}{p}, \quad D(X) = \frac{1-p}{p^2}.$$

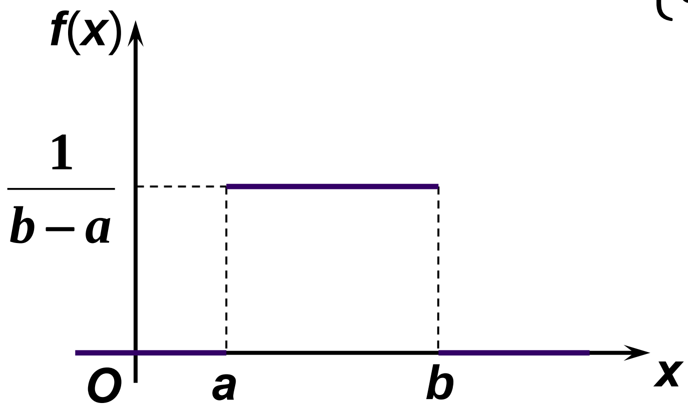
Основные законы распределения

Непрерывные случайные величины

- **Равномерный закон распределения**.

  СВ ***X*** имеет равномерное на (***a***, ***b***) распределение, если плотность ее распределения имеет вид

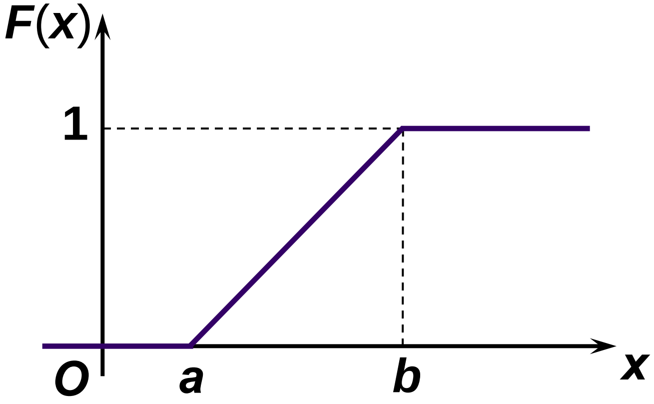
$$f(x)=\begin{cases}\dfrac{1}{b-a} & \text{при } x\in(a,\ b),\\ 0 & \text{при } x\notin(a,\ b).\end{cases}$$



Все возможные значения СВ ***X*** принадлежат интервалу (***a***, ***b***), причем в пределах этого интервала плотность вероятности постоянна

Функция распределения (интегральная функция).

$$F(x) = \begin{cases} 0 & \text{при } x \le a, \\ \dfrac{x-a}{b-a} & \text{при } a < x \le b, \\ 1 & \text{при } x > b. \end{cases}$$



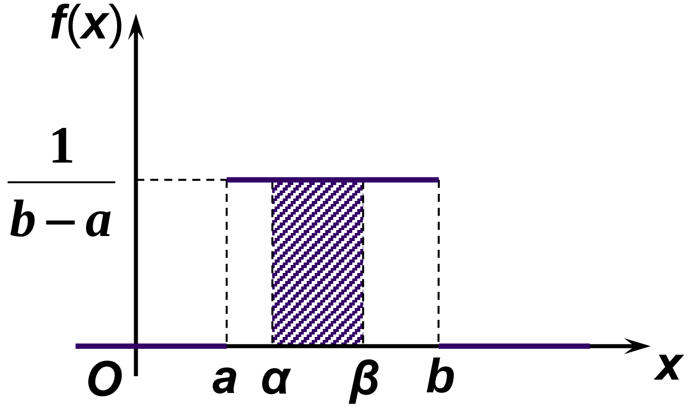
Числовые характеристики.

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

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

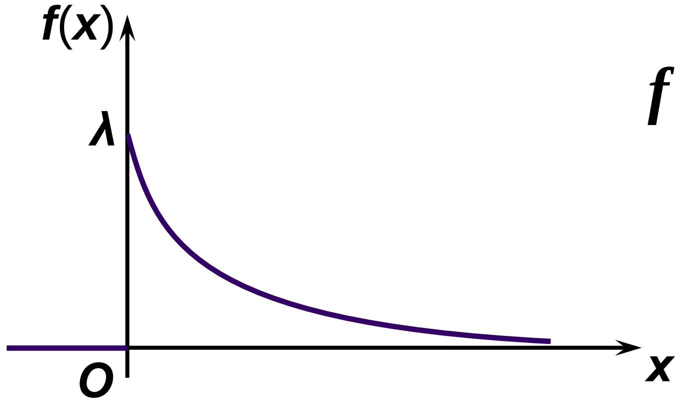
Вероятность попадания СВ в интервал $(\alpha,\ \beta) \subset (a,\ b).$

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha) = \frac{\beta - \alpha}{b - a}.$$



- **Показательное (экспоненциальное) распределение**.

  СВ ***X*** имеет показательное распределение, если плотность распределения имеет вид

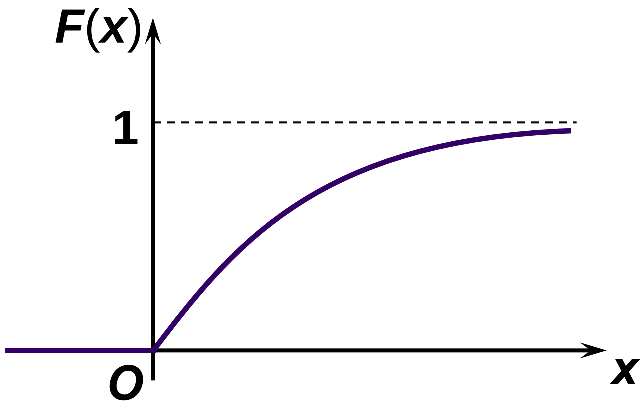


$$f(x)=\begin{cases}0 & \text{при } x<0,\\ \lambda\cdot e^{-\lambda x} & \text{при } x\ge 0,\end{cases}$$

  где $\boldsymbol{\lambda} > 0$ – параметр показательного распределения.

Функция распределения (интегральная функция).

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



Числовые характеристики.

$$M(X) = \frac{1}{\lambda}, \quad D(X) = \frac{1}{\lambda^2},$$

$$\sigma(X) = \frac{1}{\lambda}.$$

Вероятность попадания СВ в интервал (***a***, ***b***), ***a*** > 0, ***b*** > 0.

$$P(a < X < b) = F(b) - F(a) = e^{-\lambda a} - e^{-\lambda b}.$$

Условия возникновения.

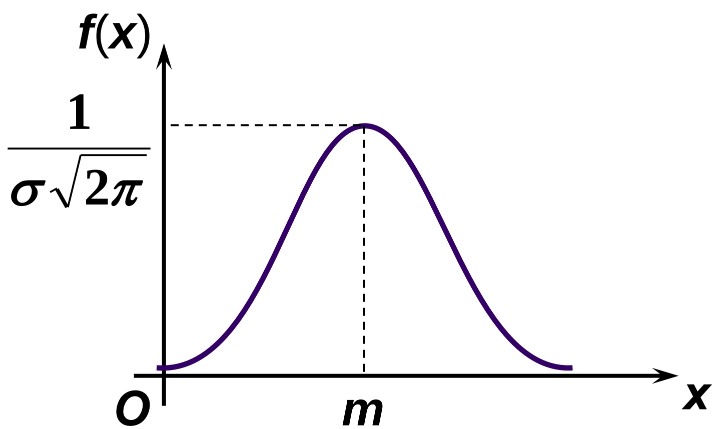
Пусть имеется простейший поток событий с плотностью λ.

СВ ***T*** – интервал времени между двумя соседними событиями в потоке – имеет показательное распределение с параметром λ.

- **Нормальный закон распределения (закон Гаусса)**.

  СВ ***X*** имеет нормальное распределение, если плотность распределения имеет вид

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



Кривая нормального распределения называется *нормальной кривой* или *кривой Гаусса*

m и ***σ*** – параметры нормального распределения.

Числовые характеристики.

$$M(X) = m, \quad D(X) = \sigma^2,$$
$$\sigma(X) = \sigma.$$

Функция распределения (интегральная функция).

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

Нормированным (*стандартным*) нормальным распределением называется нормальное распределение с параметрами ***m*** = 0 и ***σ*** = 1.

Плотность нормированного распределения

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

интегральная функция

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt .$$

Обе функции табулированы

Функция распределения нормального закона с параметрами ***m*** и ***σ*** может быть выражена через табулированные функции.

Примеры: $$F(x) = F_0\left(\frac{x-m}{\sigma}\right);$$

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x-m}{\sigma}\right),$$

где $$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt \quad -$$

интегральная функция Лапласа.

Вероятность попадания СВ в интервал (***a***, ***b***), ***a*** > 0, ***b*** > 0.

$$P(a < X < b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-m}{\sigma}}^{\frac{b-m}{\sigma}} e^{-\frac{t^2}{2}} dt.$$

Может быть выражена через табулированные функции. Например:

$$P(a < X < b) = F_0\left(\frac{b-m}{\sigma}\right) - F_0\left(\frac{a-m}{\sigma}\right) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right).$$

В частности,

$$P(|X-m| < \delta) = P(m-\delta < X < m+\delta) = 2\Phi\left(\frac{\delta}{\sigma}\right) = 2F_0\left(\frac{\delta}{\sigma}\right) - 1.$$

Условия возникновения.

Установление этих условий – в различных формах центральной предельной теоремы (ЦПТ).

Нормальный закон распределения возникает, когда рассматриваемая СВ может быть представлена в виде суммы достаточно большого числа независимых (или слабо зависимых) элементарных слагаемых, каждое из которых в отдельности сравнительно мало влияет на сумму.

Такая ситуация часто встречается на практике

широкая распространенность нормального закона.

Одна из самых простых форм ЦПТ:

ЦПТ для одинаково распределенных слагаемых.

Если X_1, X_2, … , X_n, … – независимые СВ, имеющие один и то же закон распределения с математическим ожиданием m и дисперсией σ^2, то при неограниченном увеличении n закон распределения суммы

$$Y_n = \sum_{k=1}^{n} X_k$$

неограниченно приближается к нормальному.