Задача. Известны законы распределения случайных величин X и Y – числа очков, выбиваемых 1-ми 2-м стрелками. Необходимо выяснить, какой из двух стрелков стреляет лучше. Найти числовые характеристики ДСВ.

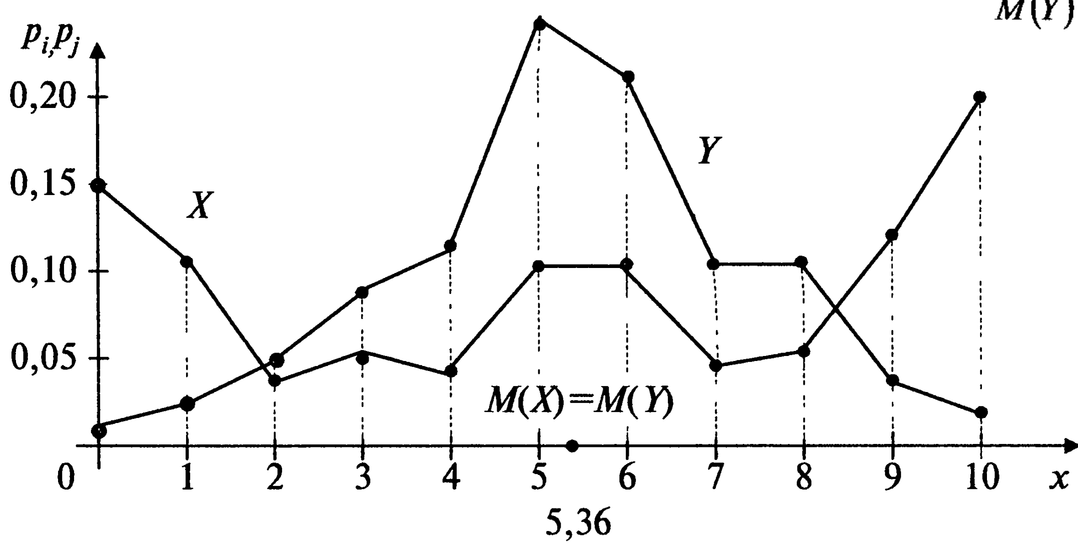
X:

x_i	0	1	2	3	4	5	6	7	8	9	10
p_i	0,15	0,11	0,04	0,05	0,04	0,10	0,10	0,04	0,05	0,12	0,20

Y:

y_j	0	1	2	3	4	5	6	7	8	9	10
p_j	0,01	0,03	0,05	0,09	0,11	0,24	0,21	0,10	0,10	0,04	0,02

Очевидно, что из двух стрелков лучше стреляет тот, кто в среднем выбивает большее количество очков.



Расчет числовых характеристик ДСВ:

$$M(X) = 0\cdot 0{,}15 + 1\cdot 0{,}10 + 2\cdot 0{,}04 + \ldots + 9\cdot 0{,}12 + 10\cdot 0{,}20 = 5{,}36,$$

$$M(Y) = 0\cdot 0{,}01 + 1\cdot 0{,}03 + 2\cdot 0{,}05 + \ldots + 9\cdot 0{,}04 + 10\cdot 0{,}02 = 5{,}36,$$

$$D(X) = (0 - 5{,}36)^2\cdot 0{,}15 + (1 - 5{,}36)^2\cdot 0{,}11 + \ldots + (10 - 5{,}36)^2\cdot 0{,}20 = 13{,}61,$$

$$\sigma_x = \sqrt{D(X)} = 3{,}69;$$

$$D(Y) = (0 - 5{,}36)^2\cdot 0{,}01 + (1 - 5{,}36)^2\cdot 0{,}03 + \ldots + (10 - 5{,}36)^2\cdot 0{,}02 = 4{,}17,$$

$$\sigma_y = \sqrt{D(Y)} = 2{,}04.$$

Лекция 5

Законы распределения

- *Математическое ожидание, дисперсия, среднее квадратическое отклонение и другие числа*, призванные в сжатой форме выразить наиболее существенные черты распределения, называются ***числовыми характеристиками случайной величины***.

- Сама величина X – ***случайная***, а ее числовые характеристики являются величинами ***неслучайными***, постоянными.

1. Функция распределения случайной величины

- Описание случайной величины X с помощью закона распределения не является единственным и не универсально.
- Оно неприменимо для непрерывной случайной величины: 1) нельзя перечислить все бесконечное несчетное множество ее значений; 2) вероятности каждого отдельно взятого значения непрерывной случайной величины равны нулю.
- Для описания закона распределения случайной величины X возможен другой подход:

рассматривать не вероятности событий $X=x$ для разных x (как это имеет место в ряде распределения), а вероятности события $X<x$, где x – текущая переменная. Вероятность $P(X<x)$, очевидно, зависит от x, т.е. является некоторой функцией от x.

1. Функция распределения случайной величины

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

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

- Функцию $F(x)$ иногда называют интегральной функцией распределения или интегральным законом распределения.
- Геометрически функция распределения интерпретируется как вероятность того, что случайная точка X попадет левее заданной точки x

Пример. Дан ряд распределения случайной величины

X:

x_i	1	4	5	7
p_i	0,4	0,1	0,3	0,2

Найти и изобразить графически её функцию распределения.

Пример. Дан ряд распределения случайной величины

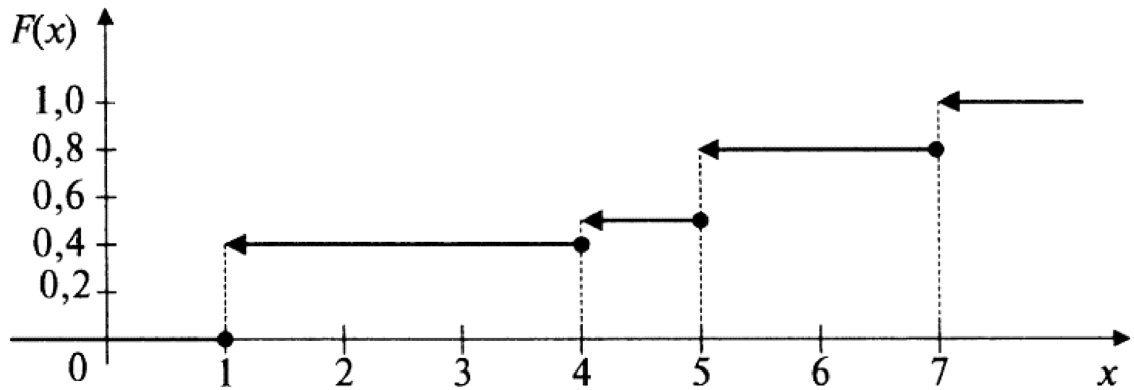
X:

x_i	1	4	5	7
p_i	0,4	0,1	0,3	0,2

Найти и изобразить графически её функцию распределения.

Решение. Будем задавать различные значения x и находить для них $F(x) = P(X < x)$.

$$F(x)=\begin{cases} 0 & \text{при} \quad x \le 1, \\ 0{,}4 & \text{при} \quad 1 < x \le 4, \\ 0{,}5 & \text{при} \quad 4 < x \le 5, \\ 0{,}8 & \text{при} \quad 5 < x \le 7, \\ 1{,}0 & \text{при} \quad x > 7. \end{cases}$$



Функция распределения любой дискретной случайной величины есть разрывная ступенчатая функция, скачки которой происходят в точках, соответствующих возможным значениям случайной величины и равны вероятностям этих значений.

Сумма всех скачков функции $F(x)$ равна 1.

1. Функция распределения случайной величины

Общие свойства функции распределения

- **1.** Функция распределения случайной величины есть неотрицательная функция, заключенная между нулем и единицей:

$$0 \le F(x) \le 1.$$

- **2.** Функция распределения случайной величины есть неубывающая функция на всей числовой оси.
- **3.** На минус бесконечности функция распределения равна нулю, на плюс бесконечности равна единице, т.е.

$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0, \qquad F(+\infty) = \lim_{x \to +\infty} F(x) = 1.$$

- 4. Вероятность попадания случайной величины в интервал $[x_1, x_2)$ (включая x_1) равна приращению ее функции распределения на этом интервале, т.е.

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

2. Законы распределения ДСВ

- Особенно важными являются ДСВ, которые принимают значения из множества целых неотрицательных чисел 0, 1, 2, ...

Эти величины описывают реальные задачи.

- К наиболее распространенным законам распределения ДСВ относят:
- ***биномиальное распределение***,
- ***распределение Пуассона,***
- ***геометрическое распределение***.

2. Законы распределения ДСВ

- *Определение.* Дискретная случайная величина X имеет биномиальный закон распределения с параметрами n и p, если она принимает значения $0, 1, 2,..., m,..., n$ с вероятностями $P(X = m) = C_n^m p^m q^{n-m}$

где $0<p<1,\ q=1-p.$

X	0	1	2	...	$n-1$	n
P	q^n	$C_m^1 pq^{n-1}$	$C_m^2 p^2 q^{n-2}$	...	$C_m^{m-1} p^{n-1} q$	p^n

- *Теорема*. Математическое ожидание и дисперсия случайной величины X, распределенной по биномиальному закону:

$$M(X) = nq\ ; \qquad D(X) = npq; \qquad \sigma = \sqrt{npq}$$

Числовые характеристики ДСВ X при биноминальном распределении

2. Законы распределения ДСВ

- Очевидно, что определение биномиального закона корректно, так как основное свойство ряда распределения $\sum_{i=0}^{n} p_i = 1$ выполнено, так как сумма в левой части является сумой всех членов разложения бинома Ньютона:

$$q^n + C_n^1 pq^{n-1} + C_n^2 p^2 q^{n-2} + \ldots + C_n^m p^m q^{n-m} + \ldots + p^n = (q+p)^n = 1^n = 1$$

Следствие. Математическое ожидание частости $\frac{m}{n}$ события в *n* независимых испытаниях, в каждом из которых оно может наступить с одной и той же вероятностью *p*, равно *p*, т.е. $M\left(\frac{m}{n}\right) = p$, а дисперсия $D\left(\frac{m}{n}\right) = \frac{pq}{n}$.

Доказательство. Частость события $\frac{m}{n}$ есть $\frac{X}{n}$, т.е. $\frac{m}{n} = \frac{X}{n}$, где *X* – случайная величина, распределенная по биномиальному закону. Поэтому

$$M\left(\frac{m}{n}\right) = M\left(\frac{X}{n}\right) = \frac{1}{n} M(X) = \frac{1}{n} \cdot np = p, \quad D\left(\frac{m}{n}\right) = D\left(\frac{X}{n}\right) = \frac{1}{n^2} D(X) = \frac{1}{n^2} \cdot npq = \frac{pq}{n}.$$

2. Законы распределения ДСВ

- *Смысл аргументов* в функциях *f(x)* и *Φ(x)*, содержащихся в локальной и интегральной теоремах Муавра-Лапласа:
- 1) аргумент *x* функции *f(x)* есть отклонение числа $X=m$ появления события A в n независимых испытаниях, распределенного по биномиальному закону, от его среднего значения $M(X)=np$, выраженное в стандартных отклонениях

$$\sigma_x = \sqrt{D(X)} = \sqrt{npq}.$$

- 2) аргумент *x* функции *Φ*(*x*), рассматриваемой в следствии интегральной теоремы Муавра-Лапласа, есть отклонение A частости m/n события A в n независимых испытаниях от его вероятности p в отдельном испытании, выраженное в стандартных отклонениях

$$\sigma\left(\frac{m}{n}\right) = \sqrt{D\left(\frac{m}{n}\right)} = \sqrt{\frac{pq}{n}}.$$

2. Законы распределения ДСВ

- Если в схеме повторных независимых испытаний $n \to \infty$, а число p близко к 0, и кроме того $np \to \lambda$ при $n \to \infty$, тогда ДСВ X, которая определяет количество появлений определённого события с схеме Бернулли имеет **распределение Пуассона**, которое задаётся таблицей:

X	0	1	2	...	$n-1$	n
p	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2}{2!}e^{-\lambda}$	...	$\frac{\lambda^{n-1}}{(n-1)!}e^{-\lambda}$	$\frac{\lambda^n}{n!}e^{-\lambda}$

Определение. Дискретная случайная величина X имеет закон распределения Пуассона с параметром $\lambda > 0$, если она принимает значения 0, 1, 2,..., m,... (бесконечное, но счетное множество значений) с вероятностями

$$P(X = m) = \frac{\lambda^m e^{-\lambda}}{m!} = P_m(\lambda)$$

2. Законы распределения ДСВ

- **Определение.** Дискретная случайная величина $X=m$ имеет геометрическое распределение с параметром p, если она принимает значения 1,2,..., m, ... (бесконечное, но счетное множество значений) с вероятностями

$$P(X=m)=pq^{m-1}$$

где $0<p<1$, $q=1-p$.

- *Случайная величина X=m, имеющая геометрическое распределение, представляет собой число m испытаний, проведенных по схеме Бернулли, с вероятностью p наступления события в каждом испытании до первого положительного исхода.*

- Числовые характеристики ДСВ X, которые имеют геометрический закон распределения:

$$M(X)=\frac{1}{p} \qquad D(X)=\frac{q}{p^2}$$

2. Законы распределения ДСВ

Ряд геометрического распределения:

X	1	2	3	…	m	…
p	p	pq	pq^2	…	pq^{m-1}	…

Вероятности pq образуют *геометрическую* прогрессию с первым членом p и знаменателем q (отсюда название *«геометрическое распределение»*).

Определение геометрического распределения корректно, так как сумма ряда

$$\sum_{i=1}^{\infty} p_i = p + pq + \ldots + pq^{m-1} + \ldots = p\left(1 + q + \ldots + q^{m-1} + \ldots\right) = p\frac{1}{1-q} = 1,$$

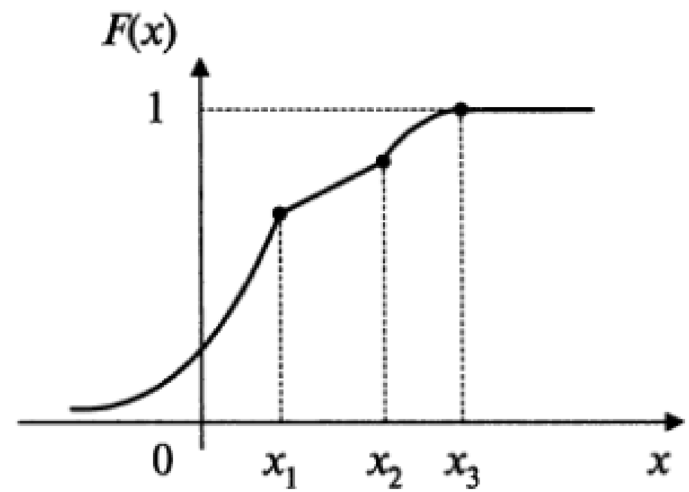
(так как $\frac{1}{1-q} = \frac{1}{p}$ сумма ряда $\sum_{m=1}^{\infty} q^{m-1}$ при $|q| < 1$).

3. Непрерывные случайные величины. Плотность вероятности

- ***Определение.*** Случайная величина X называется непрерывной, если ее функция распределения непрерывна в любой точке и дифференцируема всюду, кроме, быть может, отдельных точек.

Будем рассматривать пространство элементарных событий как совокупность всех точек числовой оси. В этом случае введенная ранее функция распределения имеет вид:

$$F(x) = P(X > x)$$



Пусть функция распределения является непрерывной. Вероятность того, что в результате испытаний случайная величина X примет значение a, где a - произвольное действительное число

Теорема. Вероятность любого отдельно взятого значения непрерывной случайной величины равна нулю.

3. Непрерывные случайные величины. Плотность вероятности

- Из приведенной выше теоремы следует, что нулевой вероятностью могут обладать и возможные события, так как событие, состоящее в том, что случайная величина X приняла конкретное значение a, является возможным.
- ***Следствие.*** Если X - непрерывная случайная величина, то вероятность попадания случайной величины в интервал (x_1, x_2) не зависит от того, является этот интервал открытым или закрытым, т.е.

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

Определение. Плотностью вероятности (плотностью распределения или просто плотностью) $f(x)$ непрерывной случайной величины X называется производная ее функции распределения $f(x) = \lim\limits_{\Delta x \to \infty} \frac{(x \le X \le x + \Delta x)}{\Delta x} = \lim\limits_{\Delta x \to \infty} \frac{F(x + \Delta x) - F(x)}{\Delta x} = F'(x)$

3. Непрерывные случайные величины. Плотность вероятности

- ***Свойства плотности вероятности***

1. Плотность вероятности является неотрицательной функцией.

2. $F(x) = P(X < x) = \int_{-\infty}^{x} f(U)dU$.

$$F(x) = F(x) - F(-\infty)$$

Геометрически полученная вероятность равна площади фигуры, ограниченной сверху кривой распределения и опирающейся на отрезок [a, b]

3. $P(a \le x \le b) = \int_{a}^{b} f(x)dx$

$$P(a \le x \le b) = F(b) - F(a) = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx = \int_{a}^{b} f(x)dx$$

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

$$\int_{-\infty}^{\infty} f(x)dx = P(X < \infty) = F(\infty) - F(-\infty) = 1 - 0 = 1$$

Следствие: Если пространством элементарных событий является отрезок числовой оси, то пространство элементарных событий формально можно распространить на всю числовую ось, положив вне отрезка значение плотности вероятности равное 0.

4. Вероятностные характеристики непрерывных случайных величин

- В случае НСВ математическое ожидание, дисперсия и среднее квадратичное отклонение имеют тот самый вид и те же свойства, но рассчитываются по другим формулам.
- Если $f(x)$ плотность распределения вероятностей X, то $M(X)$ находят по формуле:

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

- *Дисперсия*, как и в случае ДСВ, вычисляется по формуле

$$D(X)=M((X-M(X)^{2})$$

что в случае НСВ имеет вид

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

- Для расчёта удобно использовать формулу

$$D(X)=\int_{-\infty}^{+\infty} x^{2} f(x)dx-\left(\int_{-\infty}^{+\infty} xf(x)dx\right)^{2}$$

- *Среднее квадратичное отклонение* НСВ определяют следующим образом: $\sigma=\sqrt{D(X)}$

4. Вероятностные характеристики непрерывных случайных величин

- ***Модой*** *(Мо)* называется значение случайной величины, которое встречается чаще всего, т.е. имеет максимальную вероятность (для дискретной случайной величины) или максимум функции плотности вероятности в данной точке (при непрерывной случайной величине).
- Одна и та же величина может иметь несколько мод. Однако возможно, что случайная величина и не имеет моды (если все её значения имеют одинаковую вероятность (равномерное распределение).
- ***Медиана***. Определим сначала понятие квантиля непрерывной случайной величины. Корень уравнения

$F(x) = p$, где $F(x)$ - функция распределения и $0<p<1$, называется p-квантилем x_p .

- По определению функции распределения $F(x)$ получаем

$$P(X < Me) = \frac{1}{2} \quad \text{и отсюда} \quad P(X > Me) = \frac{1}{2}$$

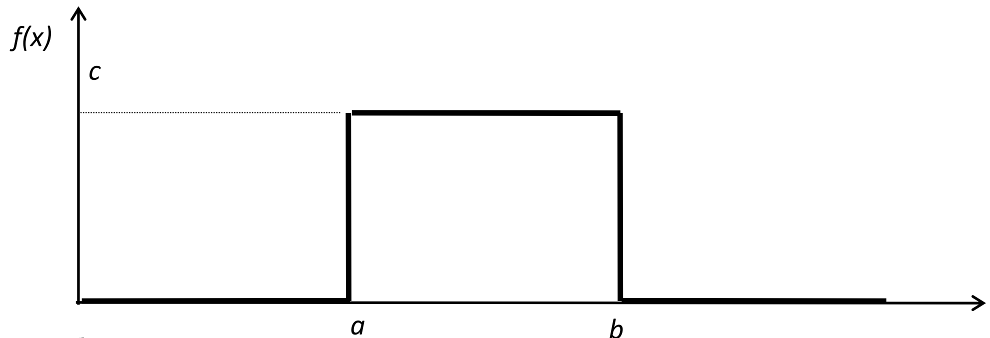
- Таким образом, медиана делит область значений случайной величины на две равные по вероятности части.

5. Законы распределения НСВ

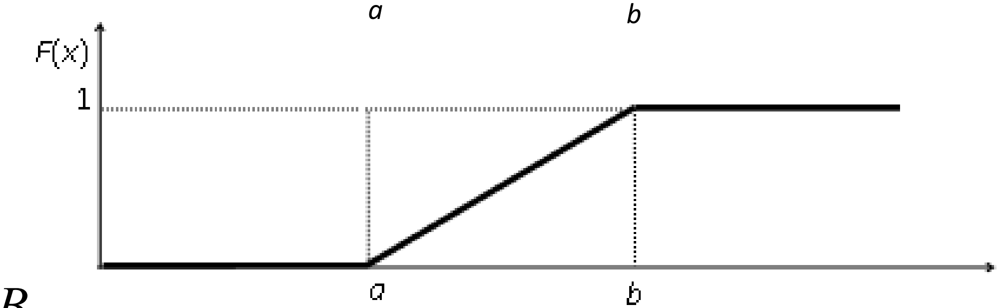
- ***Равномерное распределение***

Определение. Непрерывная случайная величина X имеет равномерный закон распределения на отрезке [a, b], если ее плотность вероятности $f(x)$ постоянна на этом отрезке и равна нулю вне его, т.е.

$$f(x)=\begin{cases} c=\dfrac{1}{b-a},\_\text{если}\_x\in[a;b]; \\ 0,\_\text{если}\_x\notin[a;b]. \end{cases}$$



$$F(x)=\begin{cases} 0,\_x\le a; \\ \dfrac{x-a}{x-b},\_a<x\le b; \\ 1,\_x>b. \end{cases}$$



Числовые характеристики равномерно распределенной *НСВ*

$$M(X)=\frac{a+b}{2}; \qquad D(X)=\frac{(b-a)^2}{12}.$$

Вероятность того, что равномерно распределённая НСВ попадёт в промежуток [x_1; x_2] при условии $a\le x_1<x_2\le b$ высчитывается по формуле

$$P(x_1<X<x_2)=\frac{x_1-x_2}{b-a}$$

5. Законы распределения НСВ

- Равномерный закон распределения используется при анализе ошибок округления при проведении числовых расчетов (например, ошибка округления числа до целого распределена равномерно на отрезке [-0,5; +0,5]), в ряде задач массового обслуживания, при статистическом моделировании наблюдений, подчиненных заданному распределению.
- Так, случайная величина X, распределенная по равномерному закону на отрезке [0;1], называемая случайным числом от 0 до 1, служит исходным материалом для получения случайных величин с любым законом распределения.

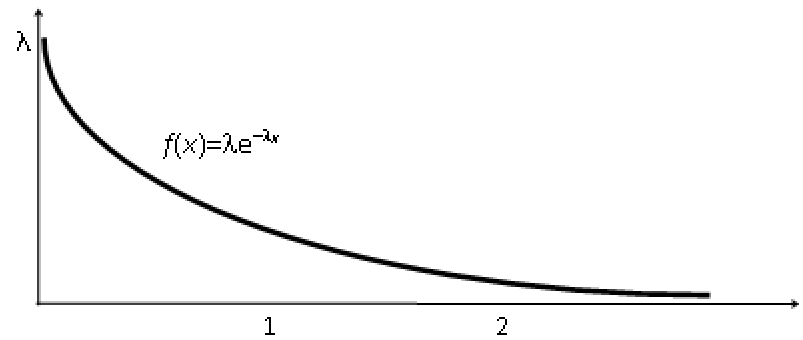
5. Законы распределения НСВ

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

$$f(x)=\begin{cases}\lambda\cdot e^{-\lambda x},\_при\_x\ge o;\\ 0,\_при\_x<0.\end{cases}$$

- *Интегральная функция распределения для НСВ, имеющей показательное распределение* задаётся формулой

$$F(x)=\begin{cases}1-e^{-\lambda x},\_при\_x\ge 0;\\ 0,\_при\_x<0.\end{cases}$$



- *Числовые характеристики:*

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

- *Вероятность того, что распределённая по показательному закону НСВ попадёт в интервал* $(a;b)$ при условии $0<a<b$ вычисляется по формуле

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

5. Законы распределения НСВ

- Показательный закон распределения играет большую роль в *теории массового обслуживания* и *теории надежности*.
- Так, например, интервал времени T между двумя соседними событиями в простейшем потоке имеет показательное распределение с параметром λ - интенсивностью потока.

5. Законы распределения НСВ

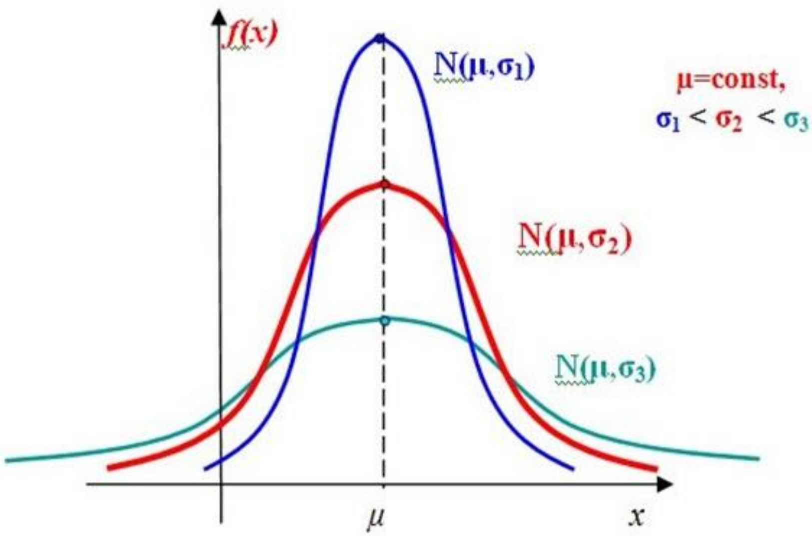
- ***Нормальный закон распределения***

Нормальный закон распределения наиболее часто встречается на практике.

Главная особенность, выделяющая его среди других законов, состоит в том, что он является *предельным* законом, к которому приближаются другие законы распределения при весьма часто встречающихся типичных условиях

Определение. Непрерывная случайная величина X имеет нормальный закон распределения (закон Гаусса) с параметрами μ и σ если ее плотность вероятности имеет вид:

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



Кривую нормального закона распределения называют нормальной или гауссовой кривой.

5. Законы распределения НСВ

- *Интегральная функция нормального распределения* имеет вид

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

- *При $\mu = 0$, $\sigma = 1$ нормальная кривая называется нормированной и НСВ X имеет стандартное или нормированное распределение.*

- Числовые характеристики НСВ X, распределенной по нормальному закону

$$M(X)=\mu \qquad\qquad D(X)=\sigma^2$$

- Вероятность того, что нормально распределенная величина попадёт в промежуток (c; d)

$$P(c<X<d)=\Phi\left(\frac{d-\mu}{\sigma}\right)-\Phi\left(\frac{c-\mu}{\sigma}\right)$$

где *$\Phi(x)$ функция Лапласа*, которая задаётся формулой

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

5. Законы распределения НСВ

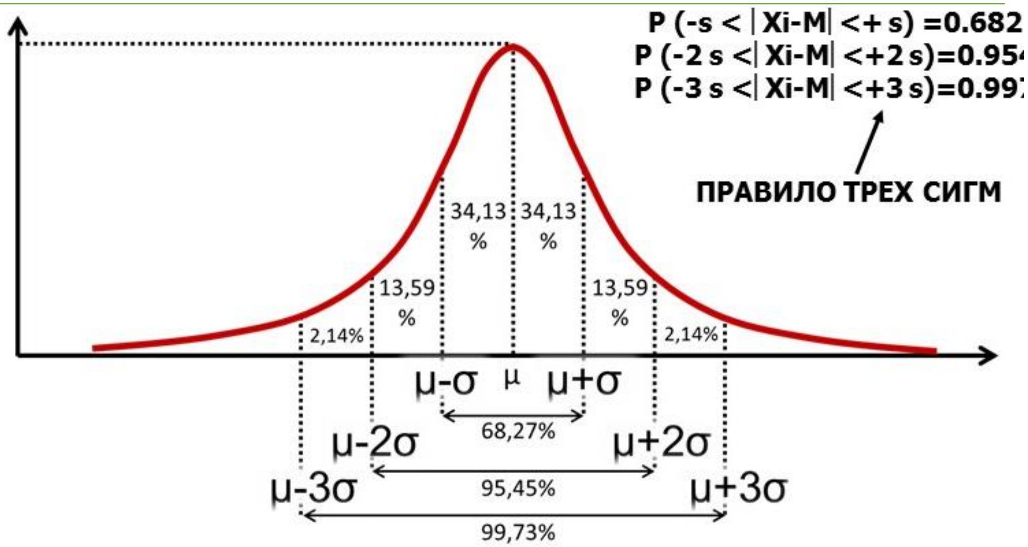
- Для вычисления вероятности отклонения нормально распределенной СВ от своего математического ожидания μ на наперед заданную величину δ используют формулу

$$P(|X-\mu|<\delta)=2\Phi\left(\frac{\delta}{\sigma}\right)$$

Правило трёх сигм: Если СВ X распределена нормально, то вероятность того, что абсолютная величина отклонения *X* от математического ожидания стремится к нулю, то есть событие

$$|X-\mu|<3\sigma$$

практически достоверно.



Если случайная величина распределена нормально, то абсолютная величина ее отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения.

На практике: если распределение случайной величины неизвестно, но условие, указанное в данном правиле выполняется, то есть основание предполагать, что случайная величина распределена нормально.