Лекция 6

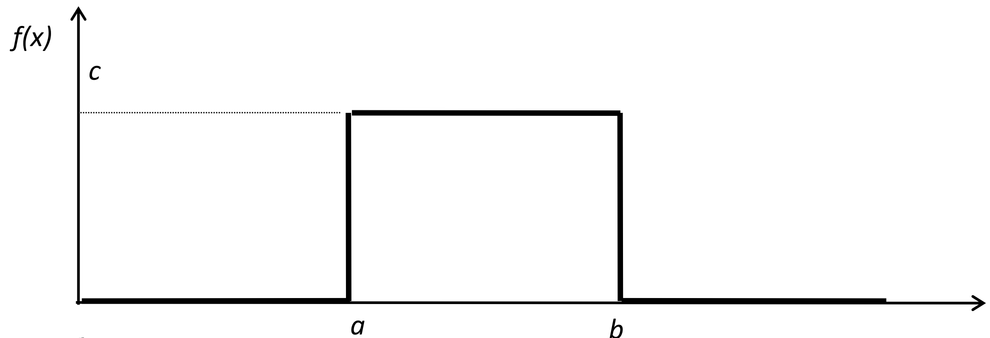
Законы распределения
Система двух случайных величин

Законы распределения НСВ

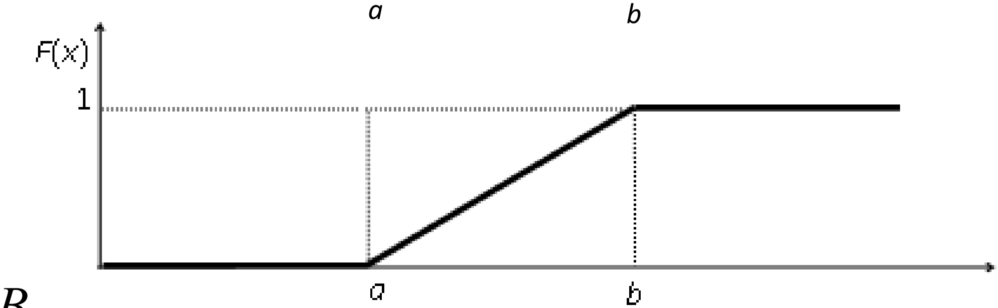
- ***Равномерное распределение***

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.

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



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



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

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.

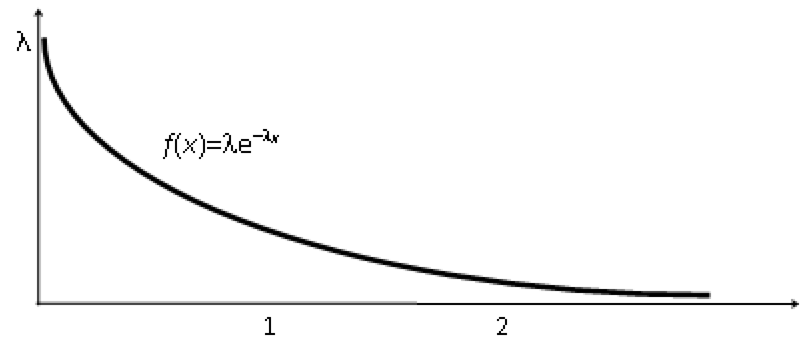
Законы распределения НСВ

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

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

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

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



- *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

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

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.

Законы распределения НСВ

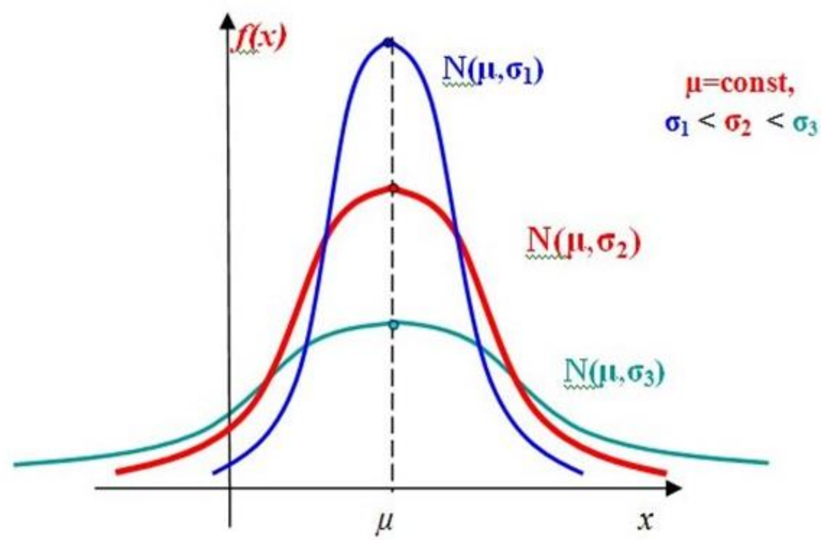
- ***Нормальный закон распределения***

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.

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 *Ф(х) функция Лапласа*, которая задаётся формулой

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

Законы распределения НСВ

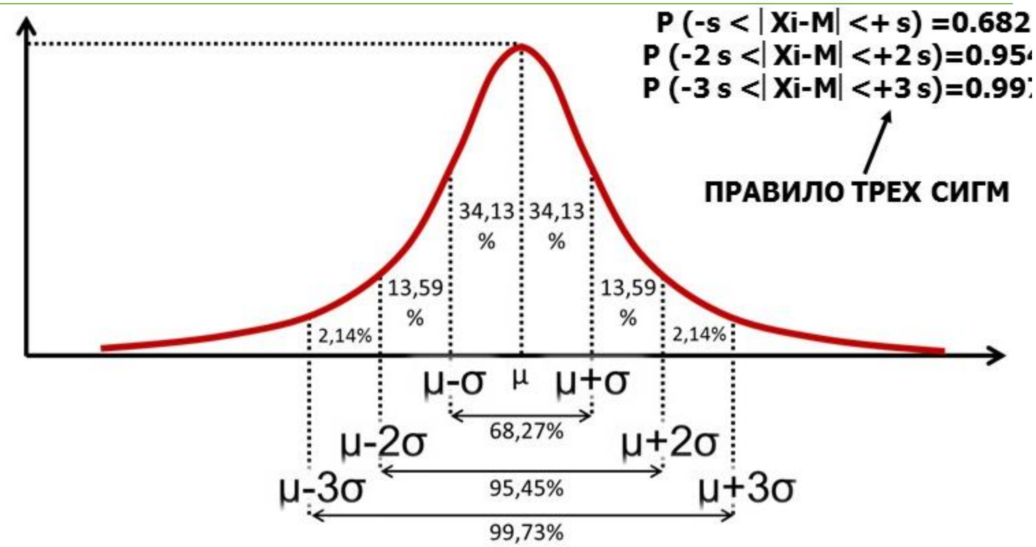
- Для вычисления вероятности отклонения нормально распределенной СВ от своего математического ожидания μ на наперед заданную величину δ используют формулу

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

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.

Числовые характеристики распределения НСВ

- ***Математическое ожидание*** *НСВ X* определяется по формуле:

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

- Если *НСВ X* определена на интервале *(a; b)*, то:

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

- ***Мода НСВ*** *X* будет определяться как максимум ее дифференциальной функции:

$$M_0(X) = \max_{(-\infty;+\infty)} f(x)$$

- ***Медиана*** определяется как значение случайной величины, которое делит площадь под дифференциальной функцией на две равные части

$$M_e(X)\text{: } P(x < M_e(X)) = P(x > M_e(X)) = \frac{1}{2}.$$

- ***Дисперсия НСВ***:

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

Все свойства дисперсии и математического ожидания, установленные для *ДСВ*, сохраняются для *НСВ*.

Числовые характеристики распределения НСВ

- ***Моменты случайных величин*.**
- *Начальным моментом* порядка s называется математическое ожидание степени s *СВ X:*

$$\alpha_s = M(X^s) \qquad (1)$$

Для *ДСВ*: $\alpha_s = \sum_{i=1}^{n} x_i^s p_i = x_1^s p_1 + x_1^s p_1 + \cdots + x_1^s p_n$.

Для *НСВ*: $\alpha_s = \int_{-\infty}^{+\infty} x^s f(x)\, dx$.

При s=1: $\alpha_1 = M(X) = m_x$, то есть, первый начальный момент - это математическое ожидание *СВ*.

Отклонение *СВ* от ее математического ожидания называется *центрированной СВ X*:

$$X = X - m_x.$$

Числовые характеристики распределения НСВ

- *Центральным моментом* порядка s *СВ* X называется математическое ожидание степени s, соответствующей центрированной *СВ*:

$$\mu_s = \mu(X^s) = M\left((x - m_x)^s\right). \quad (2)$$

- Для *ДСВ*: $\mu_s \sum_{i=1}^{n} (x_i - m_x)^s p_i = (x_1 - m_x)^s p_1 + \cdots + (x_n - m_x)^s p_n$

- Для *НСВ*: $\mu_s = \int_{-\infty}^{+\infty} (x - m_x)^s f(x)\, dx.$

При вычислении центральных моментов пользуются *формулами связи* между центральными и начальными моментами:

$$\begin{aligned} \mu_1 &= 0 \\ \mu_2 &= \alpha_2 - m_x^2, \\ \mu_3 &= \alpha_3 - 3m_x\alpha_2 + 2m_x^3, \\ \mu_4 &= \alpha_4 - 4m_x\alpha_3 + 6m_x^2\alpha_2 + 3m_x^4. \end{aligned} \quad (3)$$

Числовые характеристики распределения НСВ

Обычно рассматривают первые четыре центральных момента:

- 1) $\mu_1 = M(x - m_x) = 0$

– математическое ожидание центрированной СВ равно нулю;

- 2) $\mu_2 = M(x - m_x)^2 = D(x)$

второй центральный момент – это дисперсия;

- 3) $\mu_3 = M(x - m_x)^3$

– третий центральный момент может служить для характеристики асимметрии (скошенности распределения), обычно рассматривают безразмерный коэффициент асимметрии:

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

- 4) $\mu_4 = M(x - m_x)^4$

– четвертый центральный момент может служить для характеристики «крутости» или островершинности распределения, описывающейся с помощью эксцесса:

$$E_x = \frac{\mu_4}{\sigma^4} - 3. \quad (4)$$

Числовые характеристики распределения НСВ

- *Основным моментом* порядка s называется нормированный центральный момент порядка S:

$$r_s = \frac{\mu_s}{\sigma^s}, \qquad (5)$$

то есть $Sk = r_3$, $Ex = r_4 - 3$

1) Sk=0 - распределение симметрично $M_0(X)=M_e(X)=M(X)$,

- $Sk>0$ - распределение имеет положительную асимметрию $M_0(X)<M(X)$,
- $Sk<0$ - распределение имеет отрицательную асимметрию $M_0(X)>M(X)$.

2) Распределение имеет вершину:

- при $Ex=0$ - типа φ(x) (плотности распределения нормально распределенной СВ);
- при $Ex>0$ - более заостренную, чем φ(x);
- при $Ex<0$ - более плоскую, чем φ(x).

3) Фактически начальные и центральные моменты служат, для вычисления основных моментов, представляющих вполне определенные численные характеристики различных свойств случайных величин.

Система двух случайных величин

- В практических задачах приходится сталкиваться со случаями, когда результат описывается двумя и более случайными величинами, образующими систему случайных величин (случайный вектор,)

$(x_1, x_2, \ldots, x_n)$.

- Закон распределения *дискретной двумерной случайной* величины можно представить в виде таблицы, характеризующей собой совокупность всех значений случайных величин и соответствующих вероятностей:

	x_1	x_2	…	x_n	$\sum P(y_i)$
y_1	$P(x_1, y_1)$	$P(x_2, y_1)$	…	$P(x_n, y_1)$	$P(y_1)$
y_2	$P(x_1, y_2)$	$P(x_2, y_2)$	…	$P(x_n, y_2)$	$P(y_2)$
			…		
y_m	$P(x_1, y_m)$	$P(x_2, y_m)$	…	$P(x_n, y_m)$	$P(y_m)$
$\sum P(x_i)$	$P(x_1)$	$P(x_2)$	…	$P(x_n)$	1

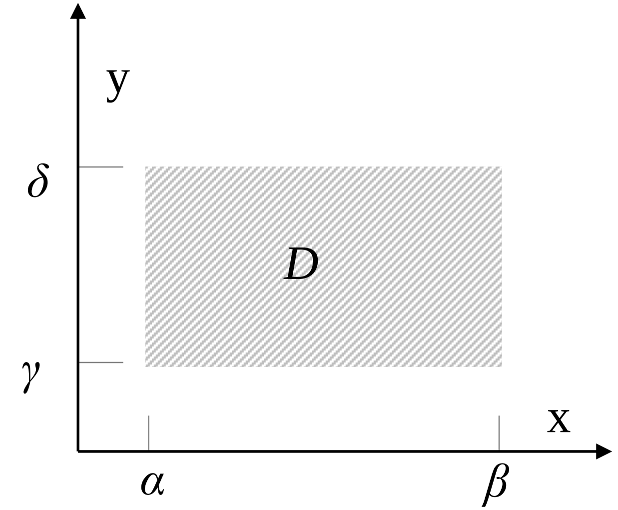
Система двух случайных величин

- В общем случае двумерная случайная величина задается в виде интегральной функции: $F(x, y) = P(X<x,\ Y<y)$, которая означает вероятность попадания двумерной случайной величины в квадрант левее и ниже точки с координатами (x, y).

Свойства интегральной функции:

1. F не убывает и непрерывна слева по каждому аргументу;
2. $F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty.) = 0$;
3. $F(+\infty, y) = F_2(y)$ - функция распределения случайной величины Y;
$F(x, +\infty) = F_1(x)$ - функция распределения случайной величины X;
4. $F(+\infty, +\infty) = 1$.

Вероятность попадания двумерной случайной величины в прямоугольник определяется, исходя из определения интегральной функции двумерной случайной величины:



$$P((x, y) \in D) = F(\beta, \delta) - F(a, \delta) - F(\beta, \gamma) + F(a, \gamma). \quad (6)$$

Система двух случайных величин

- Случайные величины X, Y *независимы,* если $F(x, y) = F_1(x) \cdot F_2(y)$. Дифференциальная функция системы двух непрерывных случайных величин определяется как вторая смешанная производная функции распределения:

$$f(x, y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = F''_{xy}(x, y). \quad (7)$$

***Свойства дифференциальной функции*:**

- 1) f(x,y)>0;
- 2) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1;$
- 3) $F(x, y) = \int_{-\infty}^{x} \int_{-\propto}^{y} f(x, y) dx dy.$

Геометрически свойство 2 означает, что объем тела, ограниченного поверхностью *f(x,y)* и плоскостью *XOY*, равен 1.

Если случайные величины x и y независимы, то

$$f(x,y)=f_1(x)f_2(y), \quad (8)$$

где $f_1(x)= F_1'(x)$, $f_2(y)= F_2'(y)$ - безусловные законы распределения.

Система двух случайных величин

- В противном случае:

$$f(x, y) = f_1(x)\, f(y/x), \quad \text{или} \quad f(x,y) = f_2(y) f(x/y),$$

(9) (10)

где $f(y/x) = \frac{f(x,y)}{f_{1(x)}}$ – условная дифференциальная функция *СВ У* при заданном значении $X=x$,

$f(x/y) = \frac{f(x,y)}{f_{2(x)}}$ – условная дифференциальная функция *СВ X* при заданном значении $Y=y$;

$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy$ и $f_2(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx$

– дифференциальные функции отдельных величин *X* и У, входящих в систему.

Числовые характеристики системы двух случайных величин

- *Начальным моментом порядка* s, h системы двух случайных величин X, $У$ называется математическое ожидание произведения степени s случайной величины X и степени h случайной величины $У$:

$$a_{s,h}, = M\{X^s Y^h\}. \qquad (11)$$

- *Центральным моментом порядка* s, h системы *СВ (X, Y)* называется математическое ожидание произведения степеней s, h соответствующих центрированных случайных величин:

$$\mu_{s,h} = M\left(X^s Y^h\right), \qquad (12)$$

где *X = X - M(X)*, *Y = Y - M(Y)* - центрированные случайные величины *X* и *Y*.

- *Основным моментом* порядка s, h системы *СВ (X, Y)* называется нормированный центральный момент порядкам s, h:

$$r_{s,h} = \frac{\mu_{s,h}}{\sigma_x^s \sigma_y^h} \qquad (13)$$

Числовые характеристики системы двух случайных величин

- Начальные моменты $a_{1.0}$, $a_{0.1}$:

$a_{1.0} = M(X^1\ Y^0) = M(X); \quad a_{0.1} = M(X^0 Y^1) = M(Y).$

- Вторые центральные моменты:

$$\mu_{2.0} = M(X^2 Y^0) = M(x - M\ (X))^2 = D(X),$$

- характеризует рассеяние случайных величин в направлении оси *OX*.

$$\mu_{0.2} = M(X^0 Y^2) = M(y - M\ (\text{У})\ )^2 = D(Y),$$

- характеризует рассеяние случайных величин в направлении оси *OY*.

Корреляционный момент. Коэффициент корреляции

- Особую роль в качестве характеристики совместной вариации случайных величин X и $У$ играет второй смешанный центральный момент, который называется ***корреляционным моментом*** (***ковариацией***):

$$\mu_{1.1} = M(X^1Y^1) = K(X,Y) = cov(X,Y) = M(XY) - M(X)M(Y). \quad (13)$$

- Корреляционный момент является мерой связи случайных величин. Если случайные величины X и У независимы, то математическое ожидание равно произведению их математических ожиданий:

$$M(XY)=M(X)\cdot M(Y), \text{ отсюда } cov\,(X, Y)=0.$$

- Если ковариация случайных величин не равна нулю, то говорят, что случайные величины коррелированы. Ковариация может принимать значения на всей числовой оси, поэтому в качестве меры связи используют основной момент порядка s=1, h=1, который называют *коэффициентом корреляции:*

$$r_{xy} = \frac{cov(x,y)}{\sigma(x)\sigma(y)} \quad (14)$$

где $\sigma(\mathrm{X}) = \sqrt{D(X)}\,,\ \sigma(Y) = \sqrt{D(Y)}$

Корреляционный момент. Коэффициент корреляции

Пример 1. Докажем, что если случайные величины *X* и У линейно зависимы, то коэффициент корреляции равен ±1.

Доказательство.

Пусть между случайными величинами *X* и *У* имеет место зависимость

$Y=AX+B$, где $M(X) = a$, $D(X) = \sigma^2$.

Тогда имеем: $M(У) = M(AX+B) = AM(X) + B = Aa+B$,

$$D(Y)= D(AX+B)= D(AX)+ D(B)= A^2 D(X)= A^2\sigma^2,$$

следовательно, $\sigma(Y) = \sqrt{A^2\sigma^2} = |A|\sigma,$

$$cov\,(X,У) = M\,((X - a)(Y - Aa - B)) = M\,((X - a)(AX+B - Aa - B))=$$
$$=AM((X-a)^2) = A\,D(X) = A\sigma^2.$$

Отсюда, $r_{xy} = \frac{A\sigma^2}{\sigma(|A|\sigma)} = \frac{A}{|A|} = \pm 1$, что и требовалось доказать.

Если между случайными величинами *X* и У существует линейная связь, то коэффициент корреляции равен ±1. Коэффициент корреляции служит мерой *линейной зависимости* между случайными величинами.

Корреляционный момент. Коэффициент корреляции

Свойства коэффициента корреляции:

- $-1 \leq r_{xy} \leq 1$.
- если $r_{xy} = \pm 1$, то случайные величины линейно зависимы;
- если $r_{xy} = 0$, то случайные величины не коррелированы, что не означает их независимости вообще.

Замечание. *Если случайные величины X и Y подчиняются нормальному закону распределения, то некоррелированность СВ X и Y означает их независимость.*

Первые моменты:

а) для дискретных *СВ*:

$$\mathrm{M}(\mathrm{X}) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i p_y ,$$

$$\mathrm{M}(Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} y_j p_y ,$$

$$D(\mathrm{X}) = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(x_i - M(X)\right)^2 p_{ij} ,$$

$$D(Y) = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(y_i - M(Y)\right)^2 p_{ij} ,$$

$$K(X,Y) = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(x_i - M(X)\right)\cdot\left(y_i - M(Y)\right)p_{ij}$$

б) для непрерывных *СВ*:

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

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

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

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

$$K(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(x - M(X)\right)\cdot\left(y - M(Y)\right)f(x,y)\,dx\,dy$$

Функции случайных величин

Закон распределения функции случайных величин

- Пусть имеется непрерывная случайная величина X с функцией плотности вероятности $f(x)$. Другая случайная величина Y связана со случайной величиной X функциональной зависимостью: $Y=\varphi(X)$, Случайная точка (X, Y) может находиться только на кривой $y=\varphi(x)$.
- Дифференциальная функция случайной величины Y определяется при условии, что $\varphi(x)$ - монотонна на интервале (a, b), тогда для функции $\varphi(x)$ существует обратная функция: $\varphi^{-1}=\psi$, $x = \psi(y)$.
- Обычно числовая прямая разбивается на n промежутков монотонности и обратная функция находится на каждом из них, поэтому:

$$g(y) = \sum_{i=1}^{n} f\left(\psi_i(y)\right) \cdot \left|\psi_i'(y)\right|, \quad (15)$$

$g(y)$ - дифференциальная функция *СВ Y*.

Функции случайных величин

Математическое ожидание и дисперсию *СВ Y* - функции случайной величины *X(Y=φ(X))*, имеющей дифференциальную функцию *f(x)*, можно определить по формулам:

- $M(Y) = M\big(\varphi(x)\big) = \int_{-\infty}^{+\infty} \varphi(x) f(x)\, dx,$ (16)
- $D(Y) = D\big(\varphi(x)\big) = \int_{-\infty}^{+\infty} \varphi^2(x) f(x)\, dx - M^2(Y).$ (17)

Пример 2. Случайная величина *X* подчиняется нормальному закону распределения с математическим ожиданием *a* и дисперсией σ^2, то есть дифференциальная функция имеет вид:

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

Найти дифференциальную функцию случайной величины $Y=X^2$.

Решение. На (0;∞), для $y=x^2$, обратная функция $x = \sqrt{y} = \psi_1$;
на (- ∞;0) - обратная функция $x = -\sqrt{y} = \psi_2$. По формуле (15):

$$g(y) = f\big(\psi_1(y)\big) \cdot |\psi_1'(y)| + f\big(\psi_2(y)\big) \cdot |\psi_2'(y)|$$
$$= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\sqrt{y}-a)^2}{2\sigma^2}} \cdot \frac{1}{2\sqrt{y}} + \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(-\sqrt{y}-a)^2}{2\sigma^2}} \cdot \frac{1}{2\sqrt{y}}.$$

При a=0 и σ=1: $g(y) = \frac{1}{\sqrt{2\pi v}} e^{-\frac{y}{2}}$.