Коэффициент корреляции и его свойства.

Def Ковариацией случайных величин X и Y наз. число

$$\operatorname{cov}(X,Y) = E[(X-EX)(Y-EY)]$$

Эту величину используют как индикатор зависимости между двумя случайными величинами.

Свойства: 1) $\operatorname{cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y$

$\triangle$ $\operatorname{cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)] = \mathbb{E}[XY - X\mathbb{E}Y - Y\mathbb{E}X + \mathbb{E}X\mathbb{E}Y] =$
$= \mathbb{E}XY - 2\mathbb{E}X\mathbb{E}Y + \mathbb{E}X\mathbb{E}Y = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y$ $\blacktriangle$

2) $\operatorname{cov}(X,X) = \mathbb{D}X$

3) $\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X)$

4) $\operatorname{cov}(cX,Y) = \operatorname{cov}(X,cY) = c\cdot\operatorname{cov}(X,Y)$

Def Коэффициентом корреляции $\rho(X,Y)$ случайных величин X и Y, дисперсии которых существуют и отличны от нуля, наз. число

$$\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{\mathbb{D}X\,\mathbb{D}Y}}$$

Свойства: 1) Если X,Y – независимы, то $\rho(X,Y) = 0$

2) $|\rho(X,Y)| \le 1$

3) $|\rho(X,Y)| = 1 \iff X,Y$ линейно связаны $(P(X = aY+b) = 1)$

Доказательство: 1) X, Y - незав $\Rightarrow cov(X,Y) = 0 \Rightarrow \rho(X,Y) = 0$

2) $\hat{\xi} = \dfrac{\xi - \mathbb{E}\xi}{\sqrt{\mathbb{D}\xi}}$ – стандартизация

- $\mathbb{E}\hat{\xi} = \dfrac{\mathbb{E}\xi - \mathbb{E}\xi}{\sqrt{\mathbb{D}\xi}} = 0$
- $\mathbb{D}\hat{\xi} = \mathbb{E}\hat{\xi}^2 = \mathbb{D}\left[\dfrac{\xi - \mathbb{E}\xi}{\sqrt{\mathbb{D}\xi}}\right] = \dfrac{\mathbb{D}(\xi - \mathbb{E}\xi)}{\mathbb{D}\xi} = \dfrac{\mathbb{D}\xi}{\mathbb{D}\xi} = 1$

Тогда $\rho(\xi,\eta) = \mathbb{E}(\hat{\xi}\cdot\hat{\eta})$

Далее, $(x \pm y)^2 \geq 0$ равносильно $-\frac{1}{2}(x^2+y^2) \leq xy \leq \frac{1}{2}(x^2+y^2)$

$-\frac{1}{2}(\hat{\xi}^2 + \hat{\eta}^2) \leq \hat{\xi}\cdot\hat{\eta} \leq \frac{1}{2}(\hat{\xi}^2+\hat{\eta}^2)$, возьмем мат ожидание от всех частей

$-1 = -\frac{1}{2}\mathbb{E}(\hat{\xi}^2 + \hat{\eta}^2) \leq \rho(\xi,\eta) \leq \frac{1}{2}\mathbb{E}(\hat{\xi}^2+\hat{\eta}^2) = 1$ (*) ▲

3) ($\Leftarrow$) Если $\eta = a\xi + b$ то

$$\rho(\xi, a\xi+b) = \frac{\mathbb{E}[\xi(a\xi+b)] - \mathbb{E}\xi\cdot\mathbb{E}(a\xi+b)}{\sqrt{\mathbb{D}\xi}\cdot\sqrt{\mathbb{D}(a\xi+b)}} = \frac{a\mathbb{E}\xi^2 + \mathbb{E}b - a(\mathbb{E}\xi)^2 - \mathbb{E}b}{\sqrt{\mathbb{D}\xi}\; a\cdot\sqrt{\mathbb{D}\xi}} = \frac{a\mathbb{D}\xi}{|a|\mathbb{D}\xi} = \begin{cases} 1 & a>0 \\ -1, & a<0 \end{cases}$$

($\Rightarrow$) $|\rho(\xi,\eta)| = 1 \Rightarrow \exists\, a \neq 0$ и b : $P(\eta = a\xi+b) = 1$

- $\rho(\xi,\eta) = \mathbb{E}(\hat{\xi},\hat{\eta}) = 1$

Тогда, в (*) $\mathbb{E}(\hat{\xi},\hat{\eta}) = \frac{1}{2}\mathbb{E}(\hat{\xi}^2 + \hat{\eta}^2) \Rightarrow \mathbb{E}(\hat{\xi}-\hat{\eta})^2 = 0$

По свойствам мат. ожидания из этого следует, что $(\hat{\xi}-\hat{\eta})^2 = 0$

Тогда $\hat{\xi} = \hat{\eta} \Rightarrow \dfrac{\xi - \mathbb{E}\xi}{\sqrt{\mathbb{D}\xi}} = \dfrac{\eta - \mathbb{E}\eta}{\sqrt{\mathbb{D}\eta}} \Rightarrow \eta = \xi\cdot\underbrace{\dfrac{\sqrt{\mathbb{D}\eta}}{\sqrt{\mathbb{D}\xi}}}_{a} \underbrace{- \dfrac{\sqrt{\mathbb{D}\eta}}{\sqrt{\mathbb{D}\xi}}\mathbb{E}\xi + \mathbb{E}\eta}_{b}$ ▲