

17. Коэффициент корреляции и его свойства.



Status

Completed

Ковариация

Ковариацией $cov(X, Y)$ случайных величин X, Y называется число

$$cov(X, Y) = E((X - EX)(Y - EY))$$

Свойства ковариации

1. $cov(X, Y) = E(XY) - EXEY$
2. $cov(X, X) = DX$
3. $cov(X, Y) = cov(Y, X)$
4. $cov(cX, Y) = c \cdot cov(X, Y)$

Коэффициент корреляции

Коэффициентом корреляции $p(X, Y)$ случайных величин X, Y дисперсии которых существуют и отличны от нуля, называется число

$$p(X, Y) = \frac{cov(X, Y)}{\sqrt{DX} \sqrt{DY}}$$

Свойства коэффициента корреляции

1. Если X, Y - независимы, то $p(X, Y) = 0$
2. $|p(X, Y)| \leq 1$

Докажем свойство (2). Рассмотрим преобразование $\widehat{\xi} = \frac{\xi - E\xi}{\sqrt{D\xi}}$ случайной величины, называемое *стандартизацией*. Случайная величина $\widehat{\xi}$ имеет нулевое математическое ожидание и единичную дисперсию:

$$E\widehat{\xi} = E \frac{\xi - E\xi}{\sqrt{D\xi}} = \frac{E\xi - E\xi}{\sqrt{D\xi}} = 0; \quad E\widehat{\xi}^2 = D\widehat{\xi} = D \frac{\xi - E\xi}{\sqrt{D\xi}} = \frac{D(\xi - E\xi)}{D\xi} = 1.$$

Коэффициент корреляции теперь запишется проще: $\rho(\xi, \eta) = E(\widehat{\xi} \cdot \widehat{\eta})$.

Далее, неравенство $(x \pm y)^2 \geq 0$ равносильно неравенству

$$-\frac{1}{2}(x^2 + y^2) \leq xy \leq \frac{1}{2}(x^2 + y^2).$$

Подставив в него $\widehat{\xi}$ вместо x , $\widehat{\eta}$ вместо y и взяв математические ожидания всех частей неравенства, получим свойство (2):

$$-1 = -\frac{1}{2}E(\widehat{\xi}^2 + \widehat{\eta}^2) \leq \rho(\xi, \eta) = E(\widehat{\xi} \cdot \widehat{\eta}) \leq \frac{1}{2}E(\widehat{\xi}^2 + \widehat{\eta}^2) = 1. \quad (20)$$

3. $|p(X, Y)| = 1 \Leftrightarrow X, Y$ - линейно связаны ($P(X = aY + b) = 1$)

Докажем свойство (3). В одну сторону утверждение проверяется непосредственно: если $\eta = a\xi + b$, то

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

Докажем вторую часть свойства (3): если $|\rho(\xi, \eta)| = 1$, то существуют числа $a \neq 0$ и b такие, что $P(\eta = a\xi + b) = 1$.

Рассмотрим сначала случай $\rho(\xi, \eta) = \rho(\widehat{\xi}, \widehat{\eta}) = 1$. Тогда второе неравенство в формуле (20) превращается в равенство:

$$E(\widehat{\xi} \cdot \widehat{\eta}) = \frac{1}{2}E(\widehat{\xi}^2 + \widehat{\eta}^2), \quad \text{т.е.} \quad E(\widehat{\xi} - \widehat{\eta})^2 = 0.$$

Если математическое ожидание неотрицательной случайной величины $(\widehat{\xi} - \widehat{\eta})^2$ равно нулю, то $(\widehat{\xi} - \widehat{\eta})^2 = 0$, п.н. Поэтому с единичной вероятностью

$$\frac{\eta - E\eta}{\sqrt{D\eta}} = \frac{\xi - E\xi}{\sqrt{D\xi}}, \quad \eta = \frac{\sqrt{D\eta}}{\sqrt{D\xi}} \xi + E\eta - \frac{\sqrt{D\eta}}{\sqrt{D\xi}} E\xi = a\xi + b.$$

В случае $\rho(\xi, \eta) = -1$ нужно рассмотреть первое неравенство в формуле (20) и повторить рассуждения. Тем самым теорема 33 доказана. \square