

§3.5 二维随机向量的数字特征

- 定义5.1. 假设 X, Y 的方差存在, 则称

$$E(X - EX)(Y - EY)$$

为 X 与 Y 的协方差, 记为 $\text{cov}(X, Y)$, σ_{XY} .

若 $\sigma_{XY} = 0$, 则称 X 与 Y 不相关.

- 注: 协方差存在, 因为

$$2(X - EX)(Y - EY) \leq (X - EX)^2 + (Y - EY)^2.$$

- 计算公式: $\text{cov}(X, Y) = E(XY) - (EX)(EY)$.

- 定理5.1. 假设 X, Y 的方差存在, 则称

$$(\text{cov}(X, Y))^2 \leq \text{var}(X) \cdot \text{var}(Y).$$

- 证: 若 $\text{var}(X) = 0$, 则 $X \equiv c$, 于是 $\text{cov}(X, Y) = 0$.

若 $\text{var}(X) > 0$, 则 $g(t)$ 的判别式 ≤ 0 , 其中

$$\begin{aligned} g(t) &:= E(t(X - EX) + (Y - EY))^2 \\ &= t^2 \text{var}(X) + 2t \text{cov}(X, Y) + \text{var}(Y) \geq 0. \end{aligned}$$

- 定义5.2. 设 $0 < \text{var}(X), \text{var}(Y) < \infty$, 则称

$$\frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)}\sqrt{\text{var}(Y)}}$$

为 X 与 Y 的相关系数, 记为 ρ_{XY} , 简记为 ρ .

- 定理5.2. (1) $|\rho| \leq 1$; (2) X 与 Y 独立, 则不相关, 从而 $\rho = 0$;
(3) $|\rho| = 1$ 当且仅当存在 a, b 使得 $Y = a + bX$.
- 证(3): $|\rho| = 1$ 当且仅当 $g(t)$ 的判别式为0, 即存在 t_0 使得

$$g(t_0) = E(t_0(X - EX) + (Y - EY))^2 = 0$$

$$\Leftrightarrow Y = -t_0 X + EY + t_0 EX$$

- 最优线性预测(定理5.3): 设 $0 < \text{var}(X), \text{var}(Y) < \infty$, 则

$$\min_{a, b \in \mathbb{R}} E(Y - (a + bX))^2 = \text{var}(Y)(1 - \rho_{XY}^2)$$

最小值点为:

$$b = \frac{\text{cov}(X, Y)}{\text{var}(X)}, \quad a = EY - bEX.$$

例5.2. 二维正态的密度:

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(u^2+v^2-2\rho uv)} \quad u = \frac{x - \mu_1}{\sigma_1}, v = \frac{y - \mu_2}{\sigma_2}.$$

