



哈爾濱工業大學

# 协方差和相关系数、矩



# 协方差



方差的性质3.

$$D(X+Y) = D(X) + D(Y) + 2E\{[X-E(X)][Y-E(Y)]\};$$

若 $X$ 与 $Y$ 独立, 则  $D(X+Y) = D(X) + D(Y)$ ;

即  $E\{[X-E(X)][Y-E(Y)]\} = 0$

等价于, 若  $E\{[X-E(X)][Y-E(Y)]\} \neq 0$ ,

则 $X$ 与 $Y$ 不独立.

$X$ 与 $Y$ 的协方差

# 协方差



■ **定义1** 若 $E\{[X-E(X)][Y-E(Y)]\}$ 存在, 称它为随机变量 $X$ 和 $Y$ 的协方差, 记为 $\text{Cov}(X,Y)$ , 即

$$\text{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}$$

此时

$$D(X \pm Y) = D(X) + D(Y) \pm 2\text{Cov}(X,Y).$$

- ◆ 当 $\text{Cov}(X,Y)>0$ 时, 称 $X$ 与 $Y$ 正相关;
- ◆ 当 $\text{Cov}(X,Y)<0$ 时, 称 $X$ 与 $Y$ 负相关;
- ◆ 当 $\text{Cov}(X,Y)=0$ 时, 称 $X$ 与 $Y$ 不相关.

# 协方差的性质



$$\text{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}$$

1.  $\text{Cov}(X,Y)=\text{Cov}(Y,X)$ ,  $\text{Cov}(X,a)=0$ ,
2.  $D(X)=\text{Cov}(X,X)$ ;
3.  $\text{Cov}(aX,bY)=ab\text{Cov}(X,Y)$   $a,b$ 是常数;
4.  $\text{Cov}(X_1+X_2,Y)=\text{Cov}(X_1,Y)+\text{Cov}(X_2,Y)$ .

## 协方差的计算公式

$$\text{Cov}(X,Y)=E(XY)-E(X)E(Y)$$

5. 若 $X$ 与 $Y$ 独立, 则 $\text{Cov}(X,Y)=0$ . 反之不成立.

例1 设随机变量 $X$ 的概率密度为

$$f(x) = \frac{1}{2}e^{-|x|}, \quad -\infty < x < +\infty,$$

试证： $X$ 与 $|X|$ 不相关，但也不独立.

证  $\text{Cov}(X, |X|) = EX|X| - EXE|X|$

$$E(X|X|) = \int_{-\infty}^{+\infty} x|x| \cdot \frac{1}{2}e^{-|x|} dx = 0,$$

$$E(X) = \int_{-\infty}^{+\infty} x \cdot \frac{1}{2}e^{-|x|} dx = 0, \quad \text{Cov}(X, |X|) = 0,$$

可得 $X$ 与 $|X|$ 不相关.



例1 设随机变量 $X$ 的概率密度为

$$f(x) = \frac{1}{2}e^{-|x|}, \quad -\infty < x < +\infty,$$

试证： $X$ 与 $|X|$ 不相关，但也不独立.

对任意常数 $a$ 有：

$$\{|X| \leq a\} \subset \{X \leq a\} \text{ 且 } P(X \leq a) < 1,$$

$$\begin{aligned} P(X \leq a, |X| \leq a) &= P(|X| \leq a) \\ &> P(|X| \leq a)P(X \leq a). \end{aligned}$$

说明 $X$ 与 $|X|$ 不独立.

# 协方差



## 练习

$$\text{Cov}(X+2Y, 3X-4)=?$$

$$\begin{aligned} &= \text{Cov}(X, 3X) + \text{Cov}(X, -4) + \text{Cov}(2Y, 3X) \\ &\quad + \text{Cov}(2Y, -4) \end{aligned}$$

$$= 3D(X) + 0 + 2 \times 3\text{Cov}(Y, X) + 0$$

$$= 3D(X) + 6\text{Cov}(Y, X).$$

# 相关系数



■ 定义2 若 $\text{Cov}(X,Y)$ 存在, 且 $D(X)>0, D(Y)>0$ ,  
称

$$\rho_{XY} = \frac{\text{Cov}(X,Y)}{\sqrt{D(X)D(Y)}}.$$

为随机变量 $X$ 和 $Y$ 的相关系数.

在不致引起混淆时, 记 $\rho_{XY}$ 为 $\rho$ .



# 相关系数的性质



1.  $|\rho| \leq 1$ ;

2.  $|\rho| = 1 \Leftrightarrow$  存在常数  $a, b$ , 使  $P(Y = a + bX) = 1$ .

$\rho = 1$  时,  $b > 0$ ;  $\rho = -1$  时,  $b < 0$ .

✚ 相关系数是表示两个随机变量之间线性相关程度的一个数字特征. (无量纲)

✚ 协方差也是表示两个随机变量之间线性相关程度的一个数字特征. (有量纲)

# 相关系数的含义



$|\rho|$ 的值越接近于1,  $Y$ 与 $X$ 的线性相关程度越高;

$|\rho|$ 的值越接近于0,  $Y$ 与 $X$ 的线性相关程度越弱;

当 $\rho=0$ ,称 $X$ 与 $Y$ 不相关, 说明 $X$ 与 $Y$ 之间无线性关系.

$X$ 与 $Y$ 独立  $\longleftrightarrow$   $X$ 与 $Y$ 不相关

当相关系数存在时, 有

$$\rho=0 \Leftrightarrow \text{Cov}(X,Y)=0 \Leftrightarrow E(XY)=E(X)E(Y)$$

$$\Leftrightarrow D(X \pm Y) = D(X) + D(Y).$$



**例2** 某箱装有100件产品，其中一、二和三等品分别为80、10和10件，现在从中随机抽取一件，记  $X_i = \begin{cases} 1, & \text{抽到}i\text{等品,} \\ 0, & \text{其他.} \end{cases} (i = 1, 2, 3).$

求(1)  $(X_1, X_2)$ 的分布列及边缘分布列;  
(2)  $X_1$ 与 $X_2$ 的相关系数  $\rho$  .



(2) 在第19讲中，我们已经求得了 $(X_1, X_2)$ 的分布列及边缘分布列为

$$E(X_1) = 0.8, E(X_2) = 0.1$$

$$D(X_1) = 0.8 \times 0.2 = 0.16,$$

$$D(X_2) = 0.1 \times 0.9 = 0.09.$$

$$E(X_1 X_2) = 0 \times 0 \times 0.1 + 0 \times 1 \times 0.1$$

$$+ 1 \times 0.8 + 1 \times 1 \times 0 = 0,$$

$$\text{Cov}(X_1, X_2) = E(X_1 X_2) - E(X_1)E(X_2) = -0.08.$$

$$\rho = \frac{\text{Cov}(X_1, X_2)}{\sqrt{D(X_1)}\sqrt{D(X_2)}} = \frac{-0.08}{\sqrt{0.16}\sqrt{0.09}} = -\frac{2}{3}.$$

<del><math>X_1</math></del> $X_2$	0	1	$p_{\cdot j}$
0	0.1	0.8	0.9
1	0.1	0	0.1
$p_{i \cdot}$	0.2	0.8	1



设  $(X, Y) \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho)$ , 可以求得  $X$  与  $Y$  的协方差为

$$\text{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\} = \sigma_1 \sigma_2 \rho.$$

则  $\rho_{XY} = \rho$ .

✚ 当  $(X, Y)$  服从二维正态分布时, 有

$X$  与  $Y$  独立  $\iff X$  与  $Y$  不相关. (证明在21讲)

🌺 一般情况下, 独立与不相关并不是等价的.

## 原点矩、中心矩



■ **定义3** 若 $E(X^k)$  ( $k=1,2,\dots$ )存在, 则称 $E(X^k)$  为 $X$ 的 **$k$ 阶原点矩**, 记为 **$\alpha_k=E(X^k)$** ;

若 $E[X-E(X)]^k$  ( $k=1,2,\dots$ )存在, 则称  $E[X-E(X)]^k$  为 $X$ 的 **$k$ 阶中心矩**, 记为  **$\beta_k=E[X-E(X)]^k$**  ;

$E(X)$ 为1阶原点矩;  $D(X)$ 为2阶中心矩.

## 原点矩、中心矩



■ **定义4** 若 $E(X^k Y^l)$  ( $k, l=1, 2, \dots$ )存在, 则称 $E(X^k Y^l)$ 为 $X$ 和 $Y$ 的 **$k+l$ 阶混合原点矩**, 记为  
 $\alpha_{k,l}=E(X^k Y^l)$ ;

若 $E\{[X-E(X)]^k [Y-E(Y)]^l\}$  ( $k, l=1, 2, \dots$ )存在, 则称 $E\{[X-E(X)]^k [Y-E(Y)]^l\}$ 为 $X$ 与 $Y$ 的 **$k+l$ 阶混合中心矩**, 记为 **$\beta_{k,l}=E\{[X-E(X)]^k [Y-E(Y)]^l\}$** .

■ 协方差为1+1阶混合中心矩.



谢 谢！