第四章　随机变量的数字特征

1.	数学期望
2.	随机变量的方差
3.	协方差、相关系数和矩
4.	多维正态随机变量

电子科技大学数学科学学院 杜鸿飞　hongfeidu@qq.com

第4章3节　协方差、相关系数和矩

引 例 1

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

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

X、Y相互独立$\Rightarrow E\{[X-E(X)][Y-E(Y)]\}=0$

反过来，

$E\{[X-E(X)][Y-E(Y)]\}\neq 0 \Rightarrow X$、$Y$不相互独立

从而 $\boxed{E\{[X-E(X)][Y-E(Y)]\}}$

可在一定程度上反映X、Y之间的关系

第4章3节　协方差、相关系数和矩

一.协方差定义

定义 若$E\{[X-E(X)][Y-E(Y)]\}$存在，称

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

为随机变量X,Y的**协方差**.

$$D(X)=cov(X,X)$$

$$D(X\pm Y)=D(X)+D(Y)\pm 2cov(X,Y)$$

常用**计算公式**:

$$cov(X,Y)=E(XY)-E(X)E(Y)$$

一.协方差的性质

$$cov(X,Y) = cov(Y,X)$$

$$cov(aX + c, bY + d) = ab\, cov(X,Y)$$

$$cov(X_1 + X_2, Y) = cov(X_1, Y) + cov(X_2, Y)$$

练 习： 已知(X,Y)的联合概率密度为：

$$f(x,y) = \begin{cases} 1/\pi & x^2 + y^2 \le 1 \\ 0 & \text{其它} \end{cases}$$

则 $cov(X,Y) = \boxed{0}$

一.协方差矩阵

定义： 设n维随机变量$(X_1, X_2, \cdots, X_n)$的协方差

$$C_{ij} = cov(X_i, X_j)$$

均存在，则称矩阵

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

为$(X_1, X_2, \cdots, X_n)$的**协方差矩阵.**

第4章3节　协方差、相关系数和矩

一.协方差矩阵的性质

1) $c_{ii} = D(X_i) \qquad i = 1,2,\ldots,n$

2) $c_{ij} = c_{ji} \qquad i, j = 1,2,\ldots,n$

3) $c_{ij}^2 \le c_{ii} * c_{jj} \qquad i, j = 1,2,\ldots,n$

4）C非负定。

一.协方差的缺点

用协方差来衡量变量间关系存在的**缺点**：

变量纲影响很大

例如X表示身高，Y表示体重

X单位取米，Y单位取公斤计算出的协方差

与X单位取毫米，Y单位取克计算出的协方差

在数量级上相差了10^6！

思考：如何消除量纲的影响?

二.相关系数

随机变量经过**标准化**后可以消除量纲的影响：

$$X^* = \frac{X - E(X)}{\sqrt{D(X)}},\quad Y^* = \frac{Y - E(Y)}{\sqrt{D(Y)}}$$

$$\operatorname{cov}(X^*, Y^*) = \operatorname{cov}\left(\frac{X - E(X)}{\sqrt{D(X)}}, \frac{Y - E(Y)}{\sqrt{D(Y)}}\right)$$

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

二.相关系数

定义　设二维随机变量X,Y的$D(X)>0,D(Y)>0$

称
$$\rho_{XY}=\frac{cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$
为随机变量X与Y的**相关系数**.

注:　ρ_{XY}是一无量纲的量.

$$\rho_{XY}=E\left[X^*Y^*\right]=\mathrm{cov}\left(X^*,Y^*\right)=\rho_{X^*Y^*}$$

第4章3节　协方差、相关系数和矩

二.相关系数

$$\rho_{XY}=\frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)D(Y)}}=\frac{E\{[X-E(X)][Y-E(Y)]\}}{\sqrt{D(X)}\sqrt{D(Y)}}$$

$$=E\left[\frac{X-E(X)}{\sqrt{D(X)}}\frac{Y-E(Y)}{\sqrt{D(Y)}}\right]=E\left[X^{*}Y^{*}\right]$$

$$=\operatorname{cov}\left(X^{*},Y^{*}\right)=\rho_{X^{*}Y^{*}}$$

$$\because D\left(X^{*}\right)=D\left(Y^{*}\right)=1$$

二.相关系数性质

设随机变量X,Y的相关系数ρ存在，则

1） $|\rho|\le 1$

证　明

2） $|\rho| = 1 \Leftrightarrow X$与$Y$依概率为1线性相关，即

$$\exists \alpha, \beta\ (\alpha \neq 0)\quad s.t$$

$$P\{Y = \alpha X + \beta\} = 1$$

证　明

3）若$\xi = a_1X + b_1, \eta = a_2Y + b_2$ 则

$$\rho_{\xi\eta} = \frac{a_1a_2}{|a_1a_2|}\rho_{XY}$$

证　明

二.相关系数

相关系数是衡量两个随机变量之间线性相关程度的 数字特征.

定义: 设随机变量X,Y的相关系数存在,

1) $\rho_{XY}=1$ 称 X,Y 正相关;

2) $\rho_{XY}=-1$ 称 X,Y 负相关;

3) $\rho_{XY}=0$ 称 X,Y 不相关.

注:$\rho_{XY}=0$仅说明X,Y之间没有线性关系,但可以有其他非线性关系. 参见书上 例4.4.4

二.相关系数

定理: 若随机变量X与Y相互独立，则X与Y不相关，即　$\rho_{XY}=0$

注: 1）此定理的逆定理不一定成立，即由$\rho_{XY}=0$不一定能得到X与Y相互独立.

不相关但也不独立的例子

2）$(X,Y)\sim N(\mu_1,\sigma_1^2;\ \mu_2,\sigma_2^2;\ \rho)$则

X,Y相互独立 $\Leftrightarrow$ $\rho=0$ 见书 例4.4.6

例4.3.2　例4.3.3

二.相关系数

练习：将一枚硬币重复抛掷n次, X, Y分别表示正面朝上和反面朝上的次数,

则$\rho_{XY}=$ -1

注意到：$Y=n-X$, 则

$$\rho_{X,n-X}=-\rho_{XX}=-1$$

三.矩

定义： 设X为随机变量，若$E\left(|X|^k\right) < +\infty$,则称

$$\gamma_k = E\left(X^k\right) \qquad k = 1, 2, 3, \cdots$$

为X的**k阶原点矩.**

称$\alpha_k = E\left(|X|^k\right) \qquad k = 1, 2, 3, \cdots$

为X的**k阶绝对原点矩.**

第4章3节　协方差、相关系数和矩

三.矩

定义: 设X为随机变量，若$E\left[|X-E(X)|^k\right]<+\infty$，

则称　$\mu_k = E\left\{[X-E(X)]^k\right\}$　$k=1,2,3,\cdots$

为X的k阶中心矩.

称 $\beta_k = E\left[|X-E(X)|^k\right]$　$k=1,2,3,\cdots$

为X的k阶绝对中心矩.

γ_k与μ_k的关系:

$$\mu_1 = 0 \qquad \mu_2 = D(X)$$

$$\gamma_1 = E(X) \qquad \gamma_2 = \gamma_1^2 + \mu_2$$