第4.3节 ——协方差与相关系数

（1）协方差的定义及计算(*****)

（2）协方差的性质(*****)

（3）相关系数的定义及计算(*****)

（4）相关系数的性质(***)

（5）矩与协方差矩阵（*）

（6）N维正态分布的性质(***)

（7）不相关和相互独立的关系(*****)

◆ 协方差的定义及计算

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

计算： $\operatorname{cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}=E(XY)-E(X)E(Y)$

注：区别 $D(X)=E[X-E(X)]^2=E(X^2)-[E(X)]^2$

◆ 协方差的性质

定义计算：$\mathrm{cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}=E(XY)-E(X)E(Y)$

(1) $\mathrm{cov}(X,X)=D(X)$

(2) $\mathrm{cov}(X,Y)=\mathrm{cov}(Y,X)$

(3) $\mathrm{cov}(aX,bY)=ab\,\mathrm{cov}(X,Y)$

(4) $\mathrm{cov}(X,C)=0$

(5) $\mathrm{cov}(X_1\pm X_2,Y)=\mathrm{cov}(X_1,Y)\pm\mathrm{cov}(X_2,Y)$

例：$\mathrm{cov}(X_1+X_2,X_1-2X_2)=\quad D(X_1)-\mathrm{cov}(X_1,X_2)-2D(X_2)$

◆ 协方差的性质

定义计算： $\operatorname{cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}=E(XY)-E(X)E(Y)$

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

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

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

(4) $\operatorname{cov}(X,C)=0$

(5) $\operatorname{cov}(X_1\pm X_2,Y)=\operatorname{cov}(X_1,Y)\pm\operatorname{cov}(X_2,Y)$

(6) X,Y相互独立 $\Rightarrow \operatorname{cov}(X,Y)=0$ （注：反之不一定成立）

◆ 相关系数的定义及计算

协方差： $\operatorname{cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}=E(XY)-E(X)E(Y)$

定义及计算： $\rho_{XY}=\operatorname{cov}(X^*,Y^*)=\dfrac{\operatorname{cov}(X,Y)}{\sqrt{D(X)D(Y)}}$

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

定义： 当$\rho_{XY}=0$时，称X与Y不相关.

◆ 相关系数的性质

定义及计算： $\rho_{XY}=\operatorname{cov}\left(X^*,Y^*\right)=\dfrac{\operatorname{cov}\left(X,Y\right)}{\sqrt{D\left(X\right)D\left(Y\right)}}$

(1) $\left|\rho_{XY}\right|\le 1$

(2) X,Y相互独立$\Rightarrow\rho_{XY}=0$，即X,Y不相关.

注：不能由X,Y不相关$\Rightarrow X,Y$相互独立

例： 已知X服从$\left[-\pi,\pi\right]$上的均匀分布，$Y_1=\sin X, Y_2=\cos X$，判断Y_1, Y_2是否不相关.

提示： $$\rho_{XY}=\frac{\operatorname{cov}\left(X,Y\right)}{\sqrt{D\left(X\right)D\left(Y\right)}}$$

$$\operatorname{cov}\left(X,Y\right)=E\left(XY\right)-E\left(X\right)E\left(Y\right)$$

例： 已知X服从$\left[-\pi,\pi\right]$上的均匀分布，$Y_1=\sin X, Y_2=\cos X$，判断Y_1,Y_2是否不相关.

分析： $f_X\left(x\right)=\begin{cases}\dfrac{1}{2\pi}, -\pi\le x\le\pi\\ 0\ , \qquad \text{else}\end{cases}$

$$E\left(Y_1\right)=E\left(\sin X\right)=\int_{-\pi}^{\pi}\sin x\cdot\frac{1}{2\pi}dx=0$$

$$E\left(Y_2\right)=E\left(\cos X\right)=\int_{-\pi}^{\pi}\cos x\cdot\frac{1}{2\pi}dx=0$$

$$E\left(Y_1Y_2\right)=E\left(\sin X\cos X\right)=\int_{-\pi}^{\pi}\sin x\cos x\cdot\frac{1}{2\pi}dx=0$$

$$\operatorname{cov}\left(Y_1Y_2\right)=E\left(Y_1Y_2\right)-E\left(Y_1\right)E\left(Y_2\right)=0 \quad \Rightarrow \rho_{Y_1Y_2}=0\Rightarrow Y_1,Y_2\text{不相关.}$$

◆ N维正态分布的重要性质

(1) 设(X,Y)服从二维正态分布，则X,Y不相关$\Leftrightarrow X,Y$相互独立.

(2) 有限个相互独立的正态随机变量它们的任意线性函数仍然服从正态分布.

例：已知X,Y独立，且$X\sim N(1,2),Y\sim N(0,1)$

求$Z=2X-Y+3$的概率密度$f_Z(z)$.

◆矩的定义

(1) k阶原点矩： $E\left(X^k\right)$.

(2) k阶中心矩： $E\left\{\left[X-E\left(X\right)\right]^k\right\}$.

(3) $k+l$ 阶混合中心矩： $E\left\{\left[X-E\left(X\right)\right]^k\left[Y-E\left(Y\right)\right]^l\right\}$.

◆协方差矩阵

例：$\left(X_1,X_2\right)$的协方差矩阵定义为： $\begin{pmatrix} \mathrm{cov}\left(X_1,X_1\right) & \mathrm{cov}\left(X_1,X_2\right) \\ \mathrm{cov}\left(X_2,X_1\right) & \mathrm{cov}\left(X_2,X_2\right) \end{pmatrix}$

即：$\begin{pmatrix} D\left(X_1\right) & \mathrm{cov}\left(X_1,X_2\right) \\ \mathrm{cov}\left(X_1,X_2\right) & D\left(X_2\right) \end{pmatrix}$

◆ 不相关与相互独立的关系图

$$X,Y\text{不相关} \Leftrightarrow \rho_{XY} = 0$$
$$\Leftrightarrow \operatorname{cov}(X,Y) = 0$$
$$\Leftrightarrow D(X \pm Y) = D(X) + D(Y)$$
$$\Leftrightarrow E(XY) = E(X)E(Y)$$

◆ 不相关与相互独立的关系图

X,Y相互独立 $\Leftrightarrow$ 对$\forall x,y$有：$F(x,y)=F_X(x)F_Y(y)$

$\Leftrightarrow$ 对任意x,y, 有$P\{X\le x,Y\le y\}=P\{X\le x\}P\{Y\le y\}$

$\Leftrightarrow$ 对$\forall A,B$: $P\{X\in A,Y\in B\}=P\{X\in A\}P\{Y\in B\}$

$\Leftrightarrow$ 对$\forall g_1(x),g_2(y)$: $g_1(X),g_2(Y)$相互独立

$$\Leftrightarrow\begin{cases}\text{对}\forall x_i,y_j\text{有：}P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{Y=y_j\}\\ \text{对}\forall x,y\text{有：}f(x,y)=f_X(x)f_Y(y)\end{cases}$$

◆ 不相关与相互独立的关系图

X,Y不相关 $\Leftrightarrow \rho_{XY}=0$

$\Leftrightarrow \operatorname{cov}(X,Y)=0$

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

$\Leftrightarrow E(XY)=E(X)E(Y)$

X,Y相互独立 $\Leftrightarrow$ 对$\forall x,y$有：$F(x,y)=F_X(x)F_Y(y)$

$\Leftrightarrow$ 对任意x,y, 有$P\{X\le x,Y\le y\}=P\{X\le x\}P\{Y\le y\}$

$\Leftrightarrow$ 对$\forall A,B$: $P\{X\in A,Y\in B\}=P\{X\in A\}P\{Y\in B\}$

$\Leftrightarrow$ 对$\forall g_1(x),g_2(y)$: $g_1(X),g_2(Y)$相互独立

$$\Leftrightarrow \begin{cases} \text{对}\forall x_i,y_j\text{有：} P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{Y=y_j\} \\ \text{对}\forall x,y\text{有：} f(x,y)=f_X(x)f_Y(y) \end{cases}$$

特例： 当(X,Y)服从二维正态分布时，X,Y不相关 $\Leftrightarrow$ X,Y相互独立.

◆ 作业

习题4-3（Page115）：2,3,6

2 设X服从参数为2的泊松分布，$Y=3X-2$,求$E(Y),D(Y),\mathbf{cov}(X,Y)$及$\rho_{XY}$.

解：$\boldsymbol{E}(\boldsymbol{Y})=\boldsymbol{E}(3\boldsymbol{X}-2)=3\boldsymbol{E}(\boldsymbol{X})-2=3\times2-2=4$

$$\boldsymbol{D}(\boldsymbol{Y})=\boldsymbol{D}(3\boldsymbol{X}-2)=\boldsymbol{D}(3\boldsymbol{X})=9\boldsymbol{D}(\boldsymbol{X})=9\times2=18$$

$$\mathbf{cov}(X,Y)=\mathbf{cov}(X,3X-2)=\mathbf{cov}(X,3X)=3\,\mathbf{cov}(X,X)=3D(X)=3*2=6$$

$$\boldsymbol{E}(\boldsymbol{XY})=\boldsymbol{E}\left[\boldsymbol{X}(3\boldsymbol{X}-2)\right]=\boldsymbol{E}\left(3\boldsymbol{X}^2-2\boldsymbol{X}\right)=3\boldsymbol{E}\left(\boldsymbol{X}^2\right)-2\boldsymbol{E}(\boldsymbol{X})$$

$$=3\left(\boldsymbol{D}(\boldsymbol{X})+\boldsymbol{E}^2(\boldsymbol{X})\right)-2\boldsymbol{E}(\boldsymbol{X})=3\left(2+2^2\right)-2\times2=14$$

$$\mathbf{cov}(\boldsymbol{X},\boldsymbol{Y})=\boldsymbol{E}(\boldsymbol{XY})-\boldsymbol{E}(\boldsymbol{X})\boldsymbol{E}(\boldsymbol{Y})=14-2\times4=6$$

2 设X服从参数为2的泊松分布，$Y=3X-2$,求$E(Y)$,$D(Y)$,$\mathbf{cov}(X,Y)$及ρ_{XY}.

解：$\boldsymbol{E}(\boldsymbol{Y})=\boldsymbol{E}(3\boldsymbol{X}-2)=3\boldsymbol{E}(\boldsymbol{X})-2=3\times 2-2=4$

$$\boldsymbol{D}(\boldsymbol{Y})=\boldsymbol{D}(3\boldsymbol{X}-2)=\boldsymbol{D}(3\boldsymbol{X})=9\boldsymbol{D}(\boldsymbol{X})=9\times 2=18$$

$$\mathbf{cov}(X,Y)=\mathbf{cov}(X,3X-2)=\mathbf{cov}(X,3X)=3\,\mathbf{cov}(X,X)=3D(X)=3*2=6$$

$$\rho_{XY}=\frac{\mathbf{cov}(X,Y)}{\sqrt{D(X)D(Y)}}=\frac{6}{\sqrt{2\times 18}}=1$$

3 设$D(X)=16, D(Y)=25$，$\rho_{XY}=0.5$，求$D(X+Y)$及$D(X-Y)$.

$$\rho_{XY}=\frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)D(Y)}} \Rightarrow \operatorname{cov}(X,Y)=\rho_{XY}\cdot\sqrt{D(X)D(Y)}=0.5\cdot\sqrt{16\times 25}=10$$

$$D(X+Y)=D(X)+D(Y)+2\operatorname{cov}(X,Y)=16+25+2\times 10=61$$

$$D(X-Y)=D(X)+D(Y)-2\operatorname{cov}(X,Y)=16+25-2\times 10=21$$

6 设$X \sim N(\mu,\sigma^2), Y \sim N(\mu,\sigma^2)$，且$X,Y$相互独立，试求$Z_1$，$Z_2$的相关系数$\rho_{Z_1Z_2}$

$Z_1 = \alpha X + \beta Y$，$Z_2 = \alpha X - \beta Y, \alpha, \beta$不全为零.

提示：$\rho_{XY} = \dfrac{\operatorname{cov}(X,Y)}{\sqrt{D(X)D(Y)}}$

6 设$X\sim N\left(\mu,\sigma^2\right),Y\sim N\left(\mu,\sigma^2\right)$，且$X,Y$相互独立，试求$Z_1$，$Z_2$的相关系数$\rho_{Z_1Z_2}$

$Z_1=\alpha X+\beta Y$，$Z_2=\alpha X-\beta Y,\alpha,\beta$不全为零.

解：$\operatorname{cov}\left(Z_1,Z_2\right)=\operatorname{cov}\left(\alpha X+\beta Y,\alpha X-\beta Y\right)=\operatorname{cov}\left(\alpha X,\alpha X\right)-\operatorname{cov}\left(\beta Y,\beta Y\right)$

$$=D\left(\alpha X\right)-D\left(\beta Y\right)=\alpha^2D\left(X\right)-\beta^2D\left(X\right)=\left(\alpha^2-\beta^2\right)\sigma^2$$

$$\sqrt{D\left(Z_1\right)}\sqrt{D\left(Z_2\right)}=\sqrt{D\left(\alpha X+\beta Y\right)}\sqrt{D\left(\alpha X-\beta Y\right)}=\sqrt{\left(\alpha^2+\beta^2\right)\sigma^2}\sqrt{\left(\alpha^2+\beta^2\right)\sigma^2}$$

$$=\left(\alpha^2+\beta^2\right)\sigma^2$$

$$\rho_{Z_1Z_2}=\frac{\operatorname{cov}\left(Z_1,Z_2\right)}{\sqrt{D\left(Z_1\right)}\sqrt{D\left(Z_2\right)}}=\frac{\left(\alpha^2-\beta^2\right)\sigma^2}{\left(\alpha^2+\beta^2\right)\sigma^2}=\frac{\alpha^2-\beta^2}{\alpha^2+\beta^2}$$