§ 4.4协方差和相关系数

问题 对于二维随机变量(X,Y):

$\boxed{\text{已知联合分布}}$ $\overset{\longrightarrow}{\not\longleftarrow}$ $\boxed{\text{边缘分布}}$

这说明对于二维随机变量，除了每个随机变量各自的概率特性以外，相互之间可能还有某种联系.
问题是用一个什么样的数去反映这种联系.

数 $E\big((X-E(X))(Y-E(Y))\big)$

反映了随机变量X,Y之间的某种关

协方差和相关系数的定义

定义 称 $E\big((X-E(X))(Y-E(Y))\big)$

为 X,Y 的协方差． 记为

$$\operatorname{cov}(X,Y)=E\big((X-E(X))(Y-E(Y))\big)$$

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

称 $\begin{pmatrix} D(X) & \operatorname{cov}(X,Y) \\ \operatorname{cov}(X,Y) & D(Y) \end{pmatrix}$

为（X, Y）的协方差矩

高维协方差矩阵

若$D(X)>0, D(Y)>0$,称

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

为X,Y的 相关系数，记

为

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

事实上， $\rho_{XY}=\operatorname{cov}(X^*,Y^*)$

若 $\rho_{XY}=0$, 称 X,Y 不相

协方差和相关系数的计算

——利用函数的期望计算协方差

- 若 (X,Y) 为离散型，

$$\mathrm{cov}(X,Y)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(x_i-E(X))(y_j-E(Y))p_{ij}$$

- 若 (X,Y) 为连续型，

$$\mathrm{cov}(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-E(X))(y-E(Y))f(x,y)dxdy$$

例1 已知 X, Y 的联合分布为

Y \ X (p_{ij})	1	0
1	p	0
0	0	q

$0 < p < 1$
$p + q = 1$

求 $\text{cov}(X, Y)$, □XY

解

X	1	0
P	p	q

Y	1	0
P	p	q

XY	1	0
P	p	q

$$E(X) = p,\ E(Y) = p,$$

$$D(X) = pq,\ D(Y) = pq,$$

$$E(XY) = p,\ D(XY) = pq,$$

$$\mathrm{cov}(X,Y) = pq,\ \rho_{XY} = 1$$

例2 (自学，记住结论)

设 $(X,Y)\sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\rho)$，求 ρ_{XY}

解 $\operatorname{cov}(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)f(x,y)\,dxdy$

令 $\dfrac{x-\mu_1}{\sigma_1}=s$

$\dfrac{y-\mu_2}{\sigma_2}=t$

$$=\frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} st\, e^{-\frac{1}{2(1-\rho^2)}(s-\rho t)^2-\frac{1}{2}t^2}\,dsdt$$

令 $s-\rho t=u$

$$=\frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} t(\rho t+u)\, e^{-\frac{u^2}{2(1-\rho^2)}-\frac{1}{2}t^2}\,dudt$$

$$=\frac{\sigma_1\sigma_2\rho}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}e^{-\frac{u^2}{2(1-\rho^2)}}du\int_{-\infty}^{+\infty}t^2e^{-\frac{1}{2}t^2}dt$$

$$=\sigma_1\sigma_2\rho$$

$$\rho_{XY}=\rho$$

若 $(X,Y)\sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\rho)$，则 X,Y 相互独立 $\Longleftrightarrow$ X,Y 不相关

例3 设 $\Theta \sim U(0,2\pi)$，$X=\cos\Theta$，$Y=\cos(\Theta+\alpha)$，

α 是给定的常数，求 ρ_{XY}

解

$$f_\Theta(t)=\begin{cases}\dfrac{1}{2\pi}, & 0<t<2\pi,\\ & \text{其他}\end{cases}$$

$$E(X)=\int_0^{2\pi}\cos t\cdot\frac{1}{2\pi}dt=0,$$

$$E(Y)=\int_0^{2\pi}\cos(t+\alpha)\cdot\frac{1}{2\pi}dt=0,$$

$$E(XY)=\int_0^{2\pi}\cos(t)\cos(t+\alpha)\cdot\frac{1}{2\pi}dt=\frac{1}{2}\cos\alpha$$

$$\operatorname{cov}(X,Y)=\frac{1}{2}\cos\alpha$$

$$E(X^2)=\int_0^{2\pi}\cos^2 t\cdot\frac{1}{2\pi}dt=\frac{1}{2},\qquad D(X)=\frac{1}{2},$$

$$E(Y^2)=\int_0^{2\pi}\cos^2(t+\alpha)\cdot\frac{1}{2\pi}dt=\frac{1}{2},\qquad D(Y)=\frac{1}{2},$$

$\longrightarrow \quad \rho_{XY}=\cos\alpha$

若 $\alpha=0,\ \rho_{XY}=1 \rightarrow Y=X$

若 $\alpha=\pi,\ \rho_{XY}=-1 \rightarrow Y=-X$

$\rightarrow$ X,Y 有线性关系 $\rightarrow$ $|\rho_{XY}|=1$

若 $\alpha=\dfrac{\pi}{2},\dfrac{3\pi}{2},\ \rho_{XY}=0$ X,Y 不相关, 但 X,Y 不独立,

此时 X,Y 没有线性关系，但有函数关系 $X^2+Y^2=1$

协方差和相关系数的性质

协方差的性质

- $\operatorname{cov}(X,Y)=\operatorname{cov}(Y,X)=E(XY)-E(X)E(Y)$
- $\operatorname{cov}(aX,bY)=ab\operatorname{cov}(X,Y)$
- $\operatorname{cov}(X+Y,Z)=\operatorname{cov}(X,Z)+\operatorname{cov}(Y,Z)$
- $\operatorname{cov}(X,X)=D(X)$

相关系数的性质

- $|\rho_{XY}| \le 1$ 简证
- $|\rho_{XY}| = 1$

Y 与 X 有线性关系的概率等于1，这种线性关系为

$$P\left(\frac{Y-E(Y)}{\sqrt{D(Y)}} = \pm\frac{X-E(X)}{\sqrt{D(X)}}\right) = 1$$

简证

- $\rho_{XY}=0 \Longleftrightarrow X, Y$ 不相关

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

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

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

X, Y 相互独立 $\Longrightarrow$ X, Y 不相关

$\not\Longleftarrow$

若 X, Y 服从二维正态分布，

X, Y 相互独立 $\Longleftrightarrow$ X, Y 不相关

例5 设 $(X,Y)\sim N(1,4;\,1,4;\,0.5)$, $Z=X+Y$,

求 □XZ

解

$$E(X)=E(Y)=1, D(X)=D(Y)=4,$$

$$\rho_{XY}=\frac{1}{2},\quad \operatorname{cov}(X,Y)=2$$

$$\operatorname{cov}(X,Z)=\operatorname{cov}(X,X)+\operatorname{cov}(X,Y)=6$$

$$D(Z)=D(X+Y)$$

$$=D(X)+D(Y)+2\operatorname{cov}(X,Y)=12$$

$$\rho_{XZ}=\frac{6}{2\sqrt{12}}=\frac{\sqrt{3}}{2}$$

作业习题四

A组：20，21