§3. 协方差及相关系数

1、定义

称 COV(X,Y)= E(X-EX)(Y-EY) 为随机变量 X,Y 的**协方差**.

特别 COV（X，X）=DX

注意 1: COV（X，Y）=E(XY)- E(X)E(Y)

注意 2: D(aX+bY)= $a^2DX+b^2DY+2abCOV(X,Y)$

特别 $D(X\pm Y)=DX+DY\pm 2COV(X,Y)$

$$\rho_{XY}=\frac{COV(X,Y)}{\sqrt{DX}.\sqrt{DY}}$$

称为随机变量 X,Y 的**相关系数**。

注意 3： X,Y 不相关 $\Leftrightarrow \rho_{XY}=0 \Leftrightarrow COV(X,Y)=0$

$\Leftrightarrow E(XY)=EXEY \Leftrightarrow D(X\pm Y)=DX+DY$

定理：若 X ， Y 独立，则 X,Y 不相关

证明：由数学期望的性质有

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

又 E(X-EX)=0, E(Y-EY)=0

所以 E(X-EX)(Y-EY)=0 ， 即 $\rho_{XY}=0$

注意 4 ：若 E(X-EX)(Y-EY¥ 0, 即 EXY-EXE¥ 0 ， 则 X ， Y 一定相关，且 X ， Y 一定不独立。

返回主目录

但是，X ，Y 不相关，不一定有 X ，Y 相互独立

例 1 ：设（ X ，Y ）的概率密度为

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

则　X ，Y 不相关，但 X ，Y 不相互独立。

证明： $E(X)=\iint\limits_{x^2+y^2\le 1} xf(x,y)dxdy=\iint\limits_{x^2+y^2\le 1} x\frac{1}{\pi}dxdy=0$

同理 $E(Y)=0$　　$E(XY)=\iint\limits_{x^2+y^2\le 1} xyf(x,y)dxdy=0$

$$\therefore\quad E(XY)=E(X)E(Y)$$

返回主目录

即　X，Y 不相关。

但 $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \begin{cases} \frac{2}{\pi}\sqrt{1-x^2} & ,-1 \le x \le 1 \\ 0 & ，其他 \end{cases}$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx = \begin{cases} \frac{2}{\pi}\sqrt{1-y^2} & ,-1 \le y \le 1 \\ 0 & ，其他 \end{cases}$$

$\therefore \quad f_X(x) f_Y(y) \ne f(x,y)$ 即　X，Y 不相互独立。

2、协方差的性质

1) COV(X,Y)=COV(Y,X);

2) COV(aX，bY)=abCOV(X,Y);

3) COV(X+Y,Z)=COV(X,Z)+COV(Y,Z);

返回主目录

例 2

设 X，Y **是二个随机变量，已知** $DX=1$，$DY=4$，$\text{cov}(X, Y)=1$，**记**

$$\xi = X-2Y, \quad \eta = 2X-Y$$

试求：$\rho_{\xi,\eta}$.

解： $D\xi = D(X-2Y) = DX+4DY-4\text{cov}(X, Y)$

$=1+4\times4-4\times1 \quad =13$

$D\eta = D(2X-Y) = 4DX+DY-4\text{cov}(X, Y)$

$=4\times1+4-4\times1$

$=4$

返回主目录

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

$$= 2\operatorname{cov}(X, X) - 4\operatorname{cov}(Y, X) - \operatorname{cov}(X, Y) + 2\operatorname{cov}(Y, Y)$$

$$= 2DX - 5\operatorname{cov}(X, Y) + 2DY$$

$$= 2\times 1 - 5\times 1 + 2\times 4$$

$$= 5$$

所以，

$$\rho_{\xi, \eta} = \frac{\operatorname{cov}(\xi, \eta)}{\sqrt{D\xi}\sqrt{D\eta}}$$

$$= \frac{5}{\sqrt{13}\sqrt{4}} = \frac{5\sqrt{13}}{26}$$

3、相关系数的性质

1) $|\rho_{XY}| \le 1$

2) $|\rho_{XY}| = 1 \Leftrightarrow$ 存在常数 a,b 使 P{Y=a+bX}=1.

证明：令：$e = E[Y-(a+bX)]^2$

$$= EY^2 + b^2EX^2 + a^2 - 2aEY - 2bEXY + 2abEX$$

求 a,b 使 e 达到最小

令 $$\begin{cases} \dfrac{\partial e}{\partial a} = 2a + 2bEX - 2EY = 0 \\ \dfrac{\partial e}{\partial b} = 2bEX^2 - 2EXY + 2aEX = 0 \end{cases}$$

将 $a = EY - bEX$，代入第二个方程得

$$bEX^2 - EXY + (EY - bEX)EX = 0, \text{ 故 } b = \frac{EXY - EXEY}{EX^2 - (EX)^2}$$

解得

$$b_0=\frac{COV(X,Y)}{DX};$$

$$a_0=EY-b_0EX=EY-EX\cdot\frac{COV(X,Y)}{DX}$$

$$\min_{a,b}E[Y-(a+bX)]^2=E[Y-(a_0+b_0X)]^2$$

$$=E\left(Y-EY+EX\frac{COV(X,Y)}{DX}-X\cdot\frac{COV(X,Y)}{DX}\right)^2$$

$$=E\left((Y-EY)-(X-EX)\cdot\frac{COV(X,Y)}{DX}\right)^2$$

$$=DY+DX\cdot\frac{COV^2(X,Y)}{(DX)^2}-2COV(X,Y)\cdot\frac{COV(X,Y)}{DX}$$

$$=DY+\frac{COV^2(X,Y)}{DX}-2\frac{COV^2(X,Y)}{DX}$$

返回主目录

$$= DY - \frac{COV^2(X,Y)}{DX} = DY - \frac{\rho_{XY}^2 \cdot DXDY}{DX} = (1-\rho_{XY}^2)DY$$

即 $$\min_{a,b} E[Y-(a+bX)]^2 = (1-\rho_{XY}^2)DY$$

由上式得： 1) $1-\rho_{XY}^2 \quad 0,\quad \left|\rho_{XY}\right| \leq 1$

2) 若 $\left|\rho_{XY}\right| = 1$， 则 $E[Y-(a_0+b_0X)]^2 = 0$

从而 $$D[Y-(a_0+b_0X)] + (E[Y-(a_0+b_0X)])^2$$

$$= E[Y-(a_0+b_0X)]^2 = 0$$

所以 $D[Y-(a_0+b_0X)] = 0,\qquad E[Y-(a_0+b_0X)] = 0$

故 $P\{Y-(a_0+b_0X)=0\}=1$

即 $P\{Y=a_0+b_0X\}=1$

返回主目录

[illegible]

$$P\{Y-(a^*+b^*X)=0\}=1,$$

故 $E[Y-(a^*+b^*X)]^2=0$ 而

$$0=E[Y-(a^*+b^*X)]^2=\min_{a,b}E[Y-(a+bX)]^2=(1-\rho_{XY}^2)D(Y)$$

则 $1-\rho_{XY}^2=0, |\rho_{XY}|=1$。

总之：相关系数的性质

1) $|\rho_{XY}| \le 1$;

2) $|\rho_{XY}| = 1 \Leftrightarrow p\{Y = bX + a\} = 1$

当 $b > 0$ 时，$\rho_{XY} = 1$；　当 $b < 0$ 时，$\rho_{XY} = -1$.

说　明

相关系数是表征随机变量 X 与 Y 之间线性关系紧密程度的量.

当 $|\rho_{X,Y}| = 1$ 时，X 与 Y 之间以概率1存在着线性关系；

当 $|\rho_{X,Y}|$ 越接近于0时，X 与 Y 之间的线性关系越弱；

当 $|\rho_{X,Y}| = 0$ 时，X 与 Y 之间不存在线性关系(不相关).

X 与 Y 之间没有线性关系并不表示它们之间没有关系。

例 3

将一枚硬币重复抛掷 n 次，以 X 和 Y 分别表示正面向上和反面向上的次数，则 X 和 Y 的相关系数为

(A) –1, (B) 0, (C) $\frac{1}{2}$, (D) 1

解： $\because\ X+Y=n$，即 $Y=n-X$

$\therefore b=-1<0$ $\quad\therefore\ \rho_{XY}=-1.$

故 (A) 正确。

返回主目录

例 4　设 (X,Y) 服从二维正态分布，求 ρ_{XY} :

$$f(x, y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

由上述知：$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$，$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$

$$EX=\mu_1, DX=\sigma_1^2, EY=\mu_2, DY=\sigma_2^2,$$

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

$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right]^2}dydx$$

返回主目录

令 $t=\frac{1}{\sqrt{1-\rho^2}}\left[\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right],\quad u=\frac{x-\mu_1}{\sigma_1}$

则　$x-\mu_1=\sigma_1 u$，　$y-\mu_2=(t\sqrt{1-\rho^2}+\rho u)\sigma_2$

$$J=\begin{vmatrix}\frac{\partial t}{\partial x} & \frac{\partial t}{\partial y}\\ \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\end{vmatrix}^{-1}=\begin{vmatrix}-\frac{1}{\sqrt{1-\rho^2}}\frac{\rho}{\sigma_1} & \frac{1}{\sqrt{1-\rho^2}}\frac{1}{\sigma_2}\\ \frac{1}{\sigma_1} & 0\end{vmatrix}^{-1}$$

$$=\left(-\frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}}\right)^{-1}=-\sigma_1\sigma_2\sqrt{1-\rho^2}$$

返回主目录

$$x-\mu_1=\sigma_1 u\,,\qquad J=-\sigma_1\sigma_2\sqrt{1-\rho^2}$$
$$y-\mu_2=(t\sqrt{1-\rho^2}+\rho u)\sigma_2$$

$$\boldsymbol{COV(X,Y)=}$$

$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right]^2}dydx$$

$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sigma_1\sigma_2 u(t\sqrt{1-\rho^2}+\rho u)e^{-\frac{u^2}{2}-\frac{t^2}{2}}\left|-\sigma_1\sigma_2\sqrt{1-\rho^2}\right|dtdu$$

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

$$\boldsymbol{=\frac{\rho\sigma_1\sigma_2}{2\pi}\sqrt{2\pi}\cdot\sqrt{2\pi}+0=\rho\sigma_1\sigma_2}$$

返回主目录

故$\rho_{XY}=\rho$。

4、二维正态分布的性质

设(X,Y)服从二维正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

则 X，Y独立$\Leftrightarrow \rho_{XY}=\rho$ =0$\Leftrightarrow$X，Y不相关。

返回主目录

§4 矩　协方差矩阵

1、定义

若 EX^k 存在，称之为 X 的 k 阶原点矩。

[illegible]

若$E(X-EX)^k(Y-EY)^l$ 存在，称之为 X 和 Y 的 k+l 阶混合中心矩。

所以 EX 是一阶原点矩，DX 是二阶中心矩，协方差 Cov(X,Y)是二阶混合中心矩。

协方差矩阵

定义　设二维随机变量(X_1, X_2)的四个二阶矩为

$$c_{11}=E[X_1-E(X_1)]^2=D(X_1)$$

$$c_{12}=E[X_1-E(X_1)][X_2-E(X_2)]=COV(X_1,X_2)$$

$$c_{21}=E[X_2-E(X_2)][X_1-E(X_1)]=COV(X_2,X_1)$$

$$c_{22}=E[X_2-E(X_2)]^2=D(X_2)$$

则矩阵

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}=\begin{bmatrix} D(X_1) & COV(X_1,X_2) \\ COV(X_2,X_1) & D(X_2) \end{bmatrix}$$

称为二维随机变量(X_1, X_2)的 协方差矩阵

返回主目录

设$(X_1,\ X_2) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

$$c_{11} = D(X_1) = \sigma_1^2$$

$$c_{12} = COV(X_1, X_2) = \rho\sigma_1\sigma_2$$

$$c_{21} = COV(X_2, X_1) = \rho\sigma_2\sigma_1$$

$$c_{22} = D(X_2) = \sigma_2^2$$

则**协方差矩阵**

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 \end{bmatrix}$$

返回主目录

5、n 维正态分布的性质

1) n 维随机变量 $(X_1,\cdots,X_n)$ 服从 n 维正态分布 $\Leftrightarrow X_1,\cdots,X_n$ 的任意线性组合 $l_1X_1+\cdots+l_nX_n$ 服从一维正态分布。

2) 若 $X_1,\cdots,X_n$ 服从 n 维正态分布，$Y_1,\cdots,Y_k$ 是 $X_j\ (j=1,\cdots,n)$ 的线性函数，则 $(Y_1,\cdots,Y_k)$ 也服从多维正态分布

3) 若 $X_1,\cdots,X_n$ 服从 n 维正态分布，则 $X_1,\cdots,X_n$ 相互独立 $\Leftrightarrow X_1,\cdots,X_n$ 两两不相关

4) 相互独立的一维 正态随机变量的线性组合服从正态分布

返回主目录

例 4 (1) 设 X,Y 独立，$X \sim N(1,4), Y \sim N(2,9)$,

求 :$2X-Y$ 的分布；

(2) 若 $(X,Y) \sim N(1,2,4,9,\frac{1}{2})$, 即 $X \sim N(1,4), Y \sim N(2,9)$

且 $\rho_{XY}=\frac{1}{2}$,　　求 :$2X-Y$ 的分布；

解：(1)　$E(2X-Y)=2EX-EY=0$

$$D(2X-Y)=4DX+DY=4\times4+9=25$$

则 :$2X-Y \sim N(0,25)$

(2)　$D(2X-Y)=4DX+DY-2\times2COV(X,Y)$

$$=25-4\rho_{XY}\sqrt{DX}\sqrt{DY}=25-4\times\frac{1}{2}\times2\times3=13$$

则 :$2X-Y \sim N(0,13)$

返回主目录

例 1

设随机变量 $X \sim N(0, \sigma^2)$，试求 $E(X^n)$．

解：

令：$Y = \dfrac{X - EX}{\sqrt{DX}} = \dfrac{X}{\sigma}$　　则　$Y \sim N(0, 1)$．

所以，

$$E(X^n) = \sigma^n E(Y^n) = \sigma^n \int_{-\infty}^{+\infty} y^n f_Y(y)dy = \frac{\sigma^n}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^n e^{-\frac{y^2}{2}} dy$$

(1)．当n为奇数时，由于被积函数是奇函数，所以

$E(X^n) = 0$．

返回主目录

(2)．当n为偶数时，由于被积函数是偶函数，所以

$$EX^n=\frac{2\sigma^n}{\sqrt{2\pi}}\int_0^{+\infty}y^n e^{-\frac{y^2}{2}}dy$$

令：$\frac{y^2}{2}=t$，则　$y=\sqrt{2}\sqrt{t}$，

$$dy=\frac{\sqrt{2}}{2}t^{-\frac{1}{2}}dt=2^{-\frac{1}{2}}t^{-\frac{1}{2}}dt$$

其中　$\Gamma(t)=\int_0^{\infty}x^{t-1}e^{-x}dx.$

$$EX^n=\frac{2\sigma^n}{\sqrt{2\pi}}2^{\frac{n}{2}-1}\int_0^{+\infty}t^{\frac{n-1}{2}}e^{-t}dt$$

$$=2^{\frac{n}{2}}\frac{\sigma^n}{\sqrt{\pi}}\int_0^{+\infty}t^{\frac{n+1}{2}-1}e^{-t}dt=2^{\frac{n}{2}}\frac{\sigma^n}{\sqrt{\pi}}\Gamma(\frac{n+1}{2})$$

返回主目录

$$=\frac{2^{\frac{n}{2}}\sigma^n}{\sqrt{\pi}}\Gamma\left(\frac{n+1}{2}\right)$$

利用$\Gamma-$函数的性质：$\Gamma(r+1)=r\Gamma(r)$，得

$$E(X^n)=\frac{2^{\frac{n}{2}}\sigma^n}{\sqrt{\pi}}\cdot\frac{n-1}{2}\Gamma\left(\frac{n-1}{2}\right)=\frac{2^{\frac{n}{2}}\sigma^n}{\sqrt{\pi}}\cdot\frac{n-1}{2}\cdot\frac{n-3}{2}\Gamma\left(\frac{n-3}{2}\right)$$

$$=\frac{2^{\frac{n}{2}}\sigma^n}{\sqrt{\pi}}\cdot\frac{n-1}{2}\cdot\frac{n-3}{2}\cdot\cdots\cdot\frac{1}{2}\Gamma\left(\frac{1}{2}\right)$$

$$=\frac{2^{\frac{n}{2}}\sigma^n}{\sqrt{\pi}}\cdot\frac{(n-1)!!}{2^{\frac{n}{2}}}\sqrt{\pi}=\sigma^n(n-1)!!$$

因而，

$$E\left(X^{n}\right)=\begin{cases}\sigma^{n}(n-1)!! & n\text{为偶数}\\ 0 & n\text{为奇数}\end{cases}$$

其中，

$$n!!=\begin{cases}1\cdot 3\cdot 5\cdot\cdots\cdot n & n\text{为奇数}\\ 2\cdot 4\cdot 6\cdot\cdots\cdot n & n\text{为偶数}\end{cases}$$

特别，若　$X\sim N(0,\ 1)$，则

$$E\left(X^{n}\right)=\begin{cases}(n-1)!! & n\text{为偶数}\\ 0 & n\text{为奇数}\end{cases},\quad n=4\text{时},\quad EX^{4}=3.$$

返回主目录

第四章　小　结

1 阐述了数学期望、方差的概念及背景，要掌握它们的性质与计算，会求随机变量函数的数学期望和方差。

2 要熟记两点分布、二项分布、泊松分布、均匀分布、指数分布和正态分布的数学期望与方差。

3 给出了契比雪夫不等式，要会用契比雪夫不等式作简单的概率估计。

4 引进了协方差、相关系数的概念，要掌握它们的性质与计算。

5 要掌握二维正态随机变量的不相关与独立的等价性。

6 给出了矩与协方差矩阵。

作业：　$P_{116-117}$ 25,26,28,30,31,33,34.

返回主目录

第四章 习题课

例 1 设随机变量 $X \sim \begin{bmatrix} 0 & 1 \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}$, $Y \sim \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

$COV(X,Y) = \frac{1}{8}$, 则 X 与 Y 的联合分布为________ 。

解： $E(X) = \frac{3}{4}, E(Y) = \frac{1}{2}$,

$$E(XY) = COV(X,Y) + E(X)E(Y)$$

$$= \frac{1}{8} + \frac{3}{8} = \frac{1}{2},$$

$$E(XY) = P\{XY = 1\} = P\{X = 1, Y = 1\}$$

$$\therefore P\{X = 1, Y = 1\} = \frac{1}{2}$$

$X \backslash Y$	0	1	$p_{i\cdot}$
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$
$p_{\cdot j}$	$\frac{1}{2}$	$\frac{1}{2}$	

返回主目录

例 2　若 $X \sim N(\mu,\sigma^2), Y \sim N(\mu,\sigma^2)$，且它们独立，求：$E|X-Y|, D|X-Y|$

分析：$E|X-Y| = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x-y| f(x,y)dxdy$

解：令 $Z = X - Y$，则 $Z \sim N(0, 2\sigma^2)$，

$$E(Z) = 0, \quad D(Z) = 2\sigma^2, E(Z^2) = 2\sigma^2.$$

$$E|X-Y| = E(|Z|) = \int_{-\infty}^{+\infty} |z| \frac{1}{\sqrt{2\pi}\sqrt{2}\sigma} e^{-\frac{z^2}{2\cdot 2\sigma^2}} dz$$

$$= 2\int_0^{+\infty} z \frac{1}{\sqrt{2\pi}\sqrt{2}\sigma} e^{-\frac{z^2}{2\cdot 2\sigma^2}} dz \quad = \frac{2\sigma}{\sqrt{\pi}}$$

返回主目录

例 2（续）

$$D|X-Y| = E|X-Y|^2 - (E|X-Y|)^2$$

$$= E|Z|^2 - (E|Z|)^2$$

$$= 2\sigma^2(1-\frac{2}{\pi})$$