Chapter 4: Expected Values (期望值)

- The Expected Value of a Random Variable(随机变量的期望)
- Variance and Standard Deviation(方差和标准差)
- Covariance and Correlation Coefficient (协方差和相关系数)
- Conditional Expectation (条件期望)

Chapter 4: Expected Values (期望值)

- The Expected Value of a Random Variable(随机变量的期望)
- Variance and Standard Deviation(方差和标准差)
- **Covariance and Correlation Coefficient(协方差和相关系数)**
- Conditional Expectation (条件期望)

Relationship between r.v. X and r.v. Y

For r.v.s X and Y

$$D(X+Y)=D(X)+D(Y)+2E\left[(X-E(X))(Y-E(Y))\right]$$

When X and Y are independent,

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

Analysis: If X and Y are independent, then

$$E\left[(X-E(X))(Y-E(Y))\right]=0.$$

Therefore, if

$$E\left[(X-E(X))(Y-E(Y))\right]\neq 0,$$

then X and Y cannot be independent. They must be related.

If X and Y are not independent, how to describe their relationship ?

Definition: If the variances of X and Y both exists, denote

$$\boxed{Cov(X,Y)\triangleq E\left[(X-E(X))(Y-E(Y))\right]}$$

Then $Cov(X,Y)$ is called the Covariance (协方差) of X and Y.

Basic properties of Covariance（协方差的基本特性）

$$\boxed{Cov(X,Y) \triangleq E\left[(X-E(X))(Y-E(Y))\right]}$$

1. If X and Y are independent ⇒ $Cov(X,Y)=0$
2. $Cov(X,Y)=Cov(Y,X)$
3. $D(X)=E(X-E(X))^2=Cov(X,X)$
4. ✓ $\boxed{Cov(X,Y)=E(XY)-E(X)E(Y)}$

It is an important formula that is used often for calculation.

Proof:

$$\begin{aligned} Cov(X,Y) &= E\left[(X-E(X))(Y-E(Y))\right] \\ &= E[XY-XE(Y)-YE(X)+E(X)E(Y)] \\ &= E(XY)-E[XE(Y)]-E[YE(X)]+E(X)E(Y) \\ &= E(XY)-E(X)E(Y) \end{aligned}$$

5. ✓ $D(X+Y)=D(X)+D(Y)+2E\left[(X-E(X))(Y-E(Y))\right]$

$$=D(X)+D(Y)+2Cov(X,Y)$$

Basic properties of Covariance（协方差的基本特性）

$$\boxed{Cov(X,Y) \triangleq E\big[\big(X-E(X)\big)\big(Y-E(Y)\big)\big]}$$

6. For any constant a and b

$$\begin{aligned} Cov(aX,bY) &= E\big[\big(aX-E(aX)\big)\big(bY-E(bY)\big)\big] \\ &= abE\big[\big(X-E(X)\big)\big(Y-E(Y)\big)\big] \\ &= abCov(X,Y) \end{aligned}$$

✓ 7. $Cov(X_1+X_2,Y) = Cov(X_1,Y) + Cov(X_2,Y)$

Theorem: Bilinear (双线性) property

Suppose that $U = a + \sum_{i=1}^{n} b_i X_i$ and $V = c + \sum_{j=1}^{m} d_j Y_j$, then

$$\boxed{Cov(U,V) = \sum_{i=1}^{n}\sum_{j=1}^{m} b_i d_j Cov(X_i,Y_j)}$$

The meaning of Covariance

$\because X, Y$ are independent $\Rightarrow$ $Cov(X, Y) = 0$

$\therefore Cov(X, Y) \neq 0$ $\Rightarrow$ X, Y are not independent

$\Rightarrow$ X, Y must be related in some way

$Cov(X, Y) = 0$ $\overset{?}{\Rightarrow}$ X, Y are independent

No!

Example: Assume that r.v. (X,Y) follows the Uniform distribution in the circular region $G: x^2+y^2\le 1$. Compute $Cov(X,Y)$, are X, Y independent?

Independence: X and Y are independent $\Leftrightarrow$ $f(x,y) =_{a.e.} f_X(x)f_Y(y)$

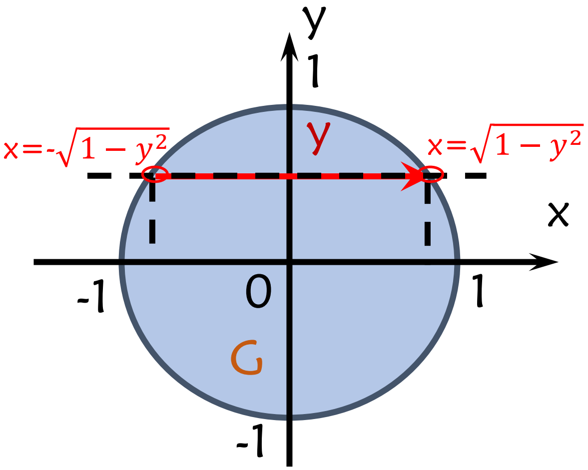
Answer: The density function of (X,Y) is

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

The marginal density functions of X and Y are:

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

$$f_X(x)=\begin{cases}\frac{2}{\pi}\sqrt{1-x^2}, & |x|<1\\ 0, & |x|\ge 1\end{cases}$$



$\because f(x,y)\ne f_X(x)f_Y(y)$

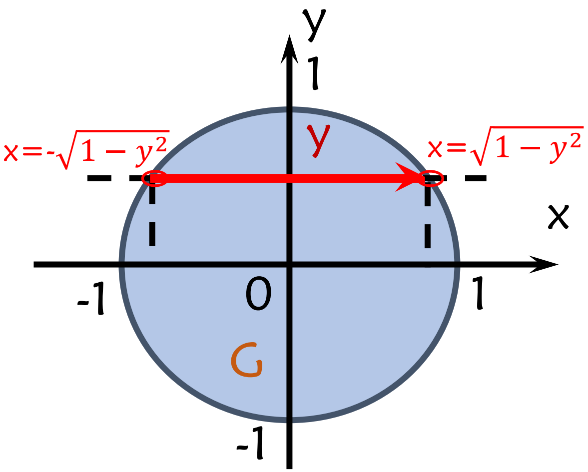
$\therefore X,Y$ are not independent.

Example: Assume that r.v. (X,Y) follows the Uniform distribution in the circular region $G: x^2+y^2\le 1$. Compute $Cov(X,Y)$, are X, Y independent?

Answer: Since

$$f_X(x)=\begin{cases}\frac{2}{\pi}\sqrt{1-x^2}, & |x|<1\\ 0, & |x|\ge 1\end{cases}$$

$$f_Y(y)=\begin{cases}\frac{2}{\pi}\sqrt{1-y^2}, & |y|<1\\ 0, & |y|\ge 1\end{cases}$$



$$\therefore E(X)=\int_{-\infty}^{\infty} x f_X(x)dx=0$$

$$E(Y)=\int_{-\infty}^{\infty} y f_Y(y)dy=0$$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy=\frac{1}{\pi}\int_{x^2+y^2\le 1} xydxdy=0$$

$$\therefore Cov(X,Y)=E(XY)-E(X)E(Y)=0$$

Therefore, X and Y are not independent, but $Cov(X,Y)=0$.

Correlation Coefficient

Consider r.v.s after "standardization(单位化/标准化)":

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

Then

$$Cov(X^*,Y^*) = \frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}}$$

Definition:

$$\boxed{\rho_{XY} \triangleq Cov(X^*,Y^*) = \frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}}}$$

ρ_{XY} is called the **correlation coefficient (相关系数)** of X and Y, which is a dimensionless quantity (无量纲数值).

What the relationship between X and Y does ρ_{XY} represent ?

Answer: Consider the linear relation between X and Y. Use r.v.

$$\hat{Y} = a + bX \ (a, b \text{ are constants})$$

to represent r.v. Y approximately, and consider the Mean Squared Error (均方误差)

$$e = E\left[\left(Y - \hat{Y}\right)^2\right] = E\left[\left(Y - (a + bX)\right)^2\right]$$

$$= E\left(Y^2\right) + b^2 E\left(X^2\right) + a^2 - 2bE(XY) + 2abE(X) - 2aE(Y)$$

Let
$$\begin{cases} \dfrac{\partial e}{\partial a} = 2a + 2bE(X) - 2E(Y) = 0 \\ \dfrac{\partial e}{\partial b} = 2bE\left(X^2\right) - 2E(XY) + 2aE(X) = 0 \end{cases}$$

Then get
$$\begin{cases} b_0 = \dfrac{Cov(X,Y)}{D(X)} \\ a_0 = E(Y) - b_0 E(X) = E(Y) - E(X)\dfrac{Cov(X,Y)}{D(X)} \end{cases}$$

$$\boxed{\begin{cases} b_0 = \dfrac{Cov(X,Y)}{D(X)} \\ a_0 = E(Y) - b_0E(X) = E(Y) - E(X)\dfrac{Cov(X,Y)}{D(X)} \end{cases}}$$

Therefore:

$$\begin{aligned}
\min_{a,b} e &= \min_{a,b} E\left[\left(Y-(a+bY)\right)^2\right] = E\left[\left(Y-(a_0+b_0X)\right)^2\right] \\
&= E\left[\left(Y-(E(Y)-b_0E(X)+b_0X)\right)^2\right] \\
&= E\left[\left(Y-E(Y)\right)^2\right] + b_0^2E\left[\left(X-E(X)\right)^2\right] - 2b_0Cov(X,Y) \\
&= D(Y) + b_0^2D(X) - 2b_0Cov(X,Y) \\
&= D(Y) + b_0Cov(X,Y) - 2b_0Cov(X,Y) = D(Y) - b_0Cov(X,Y) \\
&= D(Y)\left[1 - \frac{[Cov(X,Y)]^2}{D(X)D(Y)}\right] \\
&= D(Y)(1-\rho_{XY}^2)
\end{aligned}$$

Important relations (重要关系式)

In conclusion,

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

$$= D(Y)(1-\rho_{XY}^2)$$

Theorem ① $|\rho_{XY}| \le 1$

② $|\rho_{XY}| = 1 \Leftrightarrow Y =_{a.e.} a + bX$ (a, b are constants)

The real meaning of ρ_{XY}

$$\because \min_{a,b} e = \min_{a,b} E\left[\left(Y-(a+bX)\right)^2\right] = D(Y)(1-\rho_{XY}^2)$$

$\therefore$ When $|\rho_{XY}|$ is large, the mean square error e is small.

The relation between X and Y is closer to linear

Specifically, when $|\rho_{XY}|=1$, the relation between X and Y is almost perfectly linear.

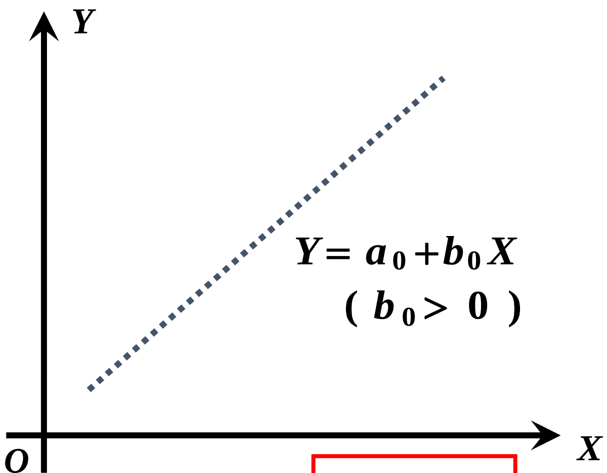
Otherwise, when $|\rho_{XY}|$ is small, the linear relation between X and Y is weak.

Definition:

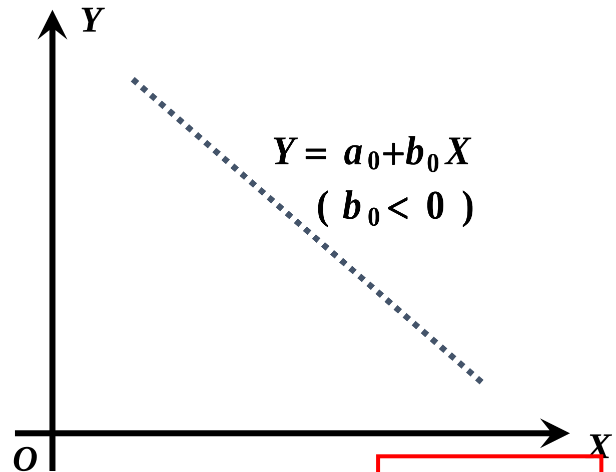
$$\left.\begin{array}{l}\text{When } \rho_{XY} = 1,\ X \text{ and } Y \text{ are positively correlated}\\ \text{When } \rho_{XY} = -1,\ X \text{ and } Y \text{ are negatively correlated}\end{array}\right\}\ \text{Correlated}$$

When $\rho_{XY} = 0$, X and Y are not correlated

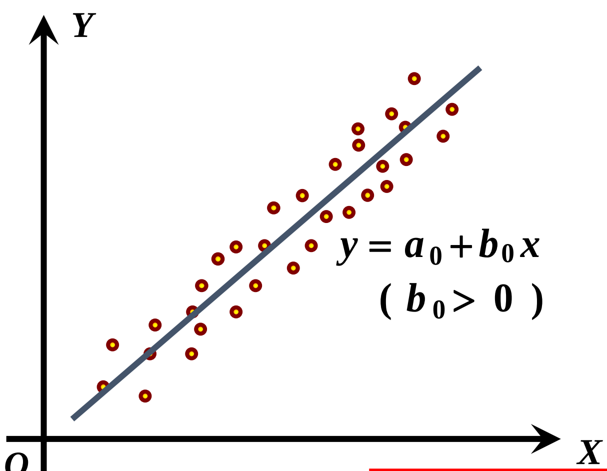
Figures showing the linear relation between X and Y



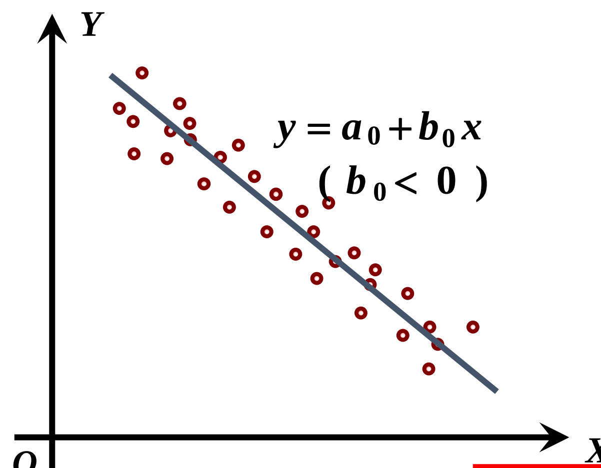
Positively correlated $\boxed{\rho_{XY} = 1}$



Negatively correlated $\boxed{\rho_{XY} = -1}$

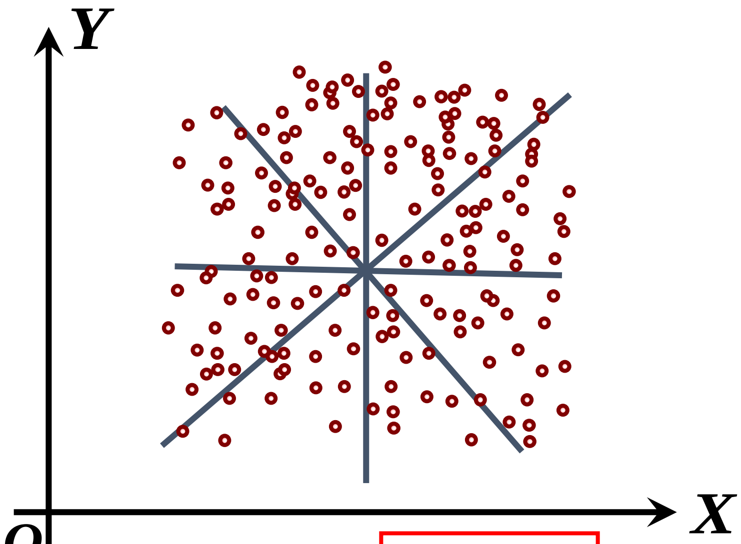


Weak positively correlated $\boxed{0 < \rho_{XY} < 1}$



Weak negatively correlated $\boxed{-1 < \rho_{XY} < 0}$

Figures showing the linear relation between X and Y



Not correlated $\boxed{\rho_{XY} = 0}$

Example: Assume that $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, what is the correlation coefficient of X and Y?

Review

1. $f(x,y)$

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

2. $X\sim N(\mu_1,\sigma_1^2), Y\sim N(\mu_2,\sigma_2^2)$

3. $E(X)=\mu_1, E(Y)=\mu_2$

   $D(X)=\sigma_1^2, D(Y)=\sigma_2^2$

Guess $\rho_{XY}=\rho$? We can prove it

Example: Assume that $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, what is the correlation coefficient of X and Y?

Proof:

$$
\begin{aligned}
Cov(X,Y) &= E[(X-\mu_1)(Y-\mu_2)] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)f(x,y)dxdy \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\sigma_1\sigma_2\sqrt{1-\rho^2}tu+\rho\sigma_1\sigma_2u^2\right)e^{-(t^2+u^2)/2}dtdu
\end{aligned}
$$

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

Let $t=\dfrac{1}{\sqrt{1-\rho^2}}\left(\dfrac{y-\mu_2}{\sigma_2}-\rho\dfrac{x-\mu_1}{\sigma_1}\right)$ and $u=\dfrac{x-\mu_1}{\sigma_1}$

Example: Assume that $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, what is the correlation coefficient of X and Y?

Proof:

$$Cov(X,Y) = E[(X-\mu_1)(Y-\mu_2)]$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\sigma_1\sigma_2\sqrt{1-\rho^2}tu + \rho\sigma_1\sigma_2u^2\right)e^{-(t^2+u^2)/2}dtdu$$

$$= \frac{\rho\sigma_1\sigma_2}{2\pi}\int_{-\infty}^{\infty}e^{-t^2/2}dt\cdot\int_{-\infty}^{\infty}u^2e^{-u^2/2}du$$

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

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

$$\therefore \rho_{XY} = \frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}} = \frac{\rho\sigma_1\sigma_2}{\sigma_1\sigma_2} = \rho$$

Relationship between independence and correlation

X and Y are independent $\overset{?}{\rightleftarrows}$ X and Y have no correlation

X and Y have no relationship at all, including linear relation.

X and Y don't have linear relation but may have other relation.

Special case Assume that $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, then

X and Y are independent $\Leftrightarrow$ $\rho = 0$ $\Leftrightarrow$ $\rho_{XY} = 0$

$\Leftrightarrow$ X and Y have no correlation

Homework

P171: 54、60 and the extra question

Extra： Assume that the density function for X is

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

(1) Compute $E(X)$ and $D(X)$.

(2)Are X and $|X|$ independent or not? State your reason.

(3)Are X and $|X|$ correlated or not? State your reason

Supplementary Questions

1. Suppose that the joint density function of (X,Y) is

$$f(x,y)=\begin{cases}\dfrac{x+y}{8}, & 0\le x\le 2, 0\le y\le 2\\ 0, & \text{otherwise}\end{cases}$$

Compute $E(X)$, $E(Y)$, $Cov(X,Y)$, ρ_{XY}, $D(X+Y)$.

2. X and Y are independent random variables which both follow the normal distribution $N(\mu,\sigma^2)$. If $Z=\alpha X+\beta Y, W=\alpha X-\beta Y$, compute $Cov(Z,W)$ and ρ_{ZW}.