COVARIÂNCIA

Sejam X e Y v.a. com distribuição de probabilidade conjunta f(x,y). A *covariância* de X e Y é

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

e pode ser escrita como

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

Exemplo

X: número de moradores
Y: número de televisores

Qual a covariância de (X,Y)?

Função de probabilidade conjunta e marginais

f(x,y)		x		
y	1	2	3	h(y)
0	0,02	0	0	0,02
1	0,20	0,08	0,05	0,33
2	0,10	0,20	0,15	0,45
3	0	0,10	0,10	0,20
g(x)	0,32	0,38	0,30	1

$$COV[X,Y] = E[XY] - E[X]E[Y]$$

$$\hookrightarrow E[XY] = \sum_x \sum_y xy\, f(x,y)$$
$$= 0\times 1\times 0{,}02 + 0\times 2\times 0 + 0\times 3\times 0 + \cdots$$
$$= 3{,}91$$

$$\hookrightarrow E[X] = \sum_x x\, g(x) = 1\times 0{,}32 + 2\times 0{,}38 + 3\times 0{,}30$$
$$= 1{,}98$$

$$\hookrightarrow E[Y] = \sum_y y\, h(y) = 0\times 0{,}02 + 1\times 0{,}33 + 2\times 0{,}45 + 3\times 0{,}20 = 1{,}83$$

$$\Rightarrow Cov[X,Y] = 3{,}91 - 1{,}98\times 1{,}83$$
$$= 3{,}91 - 3{,}6234 = 0{,}2866$$

Exemplo

Considere as intensidades elétricas de duas componentes de um sistema eletrônico, X e Y, que são v.a. A função densidade de probabilidade conjunta de X e Y é dada por

$$f(x,y) = \begin{cases} x + \frac{y}{2} + xy, & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{c.c.} \end{cases}$$

Qual a covariância de (X,Y)?

$$\mathrm{COV}[X,Y] = E[XY] - E[X]E[Y]$$

$$\hookrightarrow E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\, dx\, dy$$

$$= \int_0^1\int_0^1 xy\left(x + \frac{y}{2} + xy\right) dx\, dy = \cdots\cdots = \frac{13}{36}$$

$\hookrightarrow E[X] = \int_{-\infty}^{\infty} x\, g(x)\, dx$

$= \int_{0}^{1} x\left(\frac{3x}{2} + \frac{1}{4}\right) dx = \cdots = \frac{5}{8}$

$\hookrightarrow E[Y] = \int_{-\infty}^{\infty} y\, h(y)\, dy$

$= \int_{0}^{1} y\left(y + \frac{1}{2}\right) dy = \cdots = \frac{7}{12}$

$\Rightarrow \mathrm{Cov}[XY] = \frac{13}{36} - \frac{5}{8} \times \frac{7}{12} = 0{,}3611 - 0{,}3646 = -0{,}0035$

TEOREMA

Se X e Y são v.a. independentes então

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

COROLÁRIO

Se X e Y são v.a. independentes então

$$Cov(X,Y) = 0$$

Demonstração

$$\mathrm{Cov}[x,y] = E[xy] - E[x]E[y]$$

como x e y independentes $\Rightarrow$ $E[xy] = E[x]E[y]$

Assim,

$$\mathrm{Cov}[x,y] = E[x]E[y] - E[x]E[y]$$

$$\Rightarrow \mathrm{cov}[x,y] = 0 \,//$$

PERGUNTA: Se Cov(X,Y) = 0 então X e Y são independentes?

Vamos ver um exemplo para responder.

f(x,y)		x		
y	0	1	2	h(y)
1	3/20	3/20	2/20	8/20
2	1/20	1/20	2/20	4/20
3	4/20	1/20	3/20	8/20
g(x)	8/20	5/20	7/20	1

$\hookrightarrow E[xy] = \sum_x \sum_y xy\, f(x,y)$

$= 1 \times 0 \times \frac{3}{20} + 1 \times 1 \times \frac{3}{20} + \dots = \frac{38}{20}$

$\hookrightarrow E[x] = \sum_x x\, g(x)$

$= 0 \times \frac{8}{20} + 1 \times \frac{5}{20} + 2 \times \frac{7}{20} = \frac{19}{20}$

$\hookrightarrow E[y] = \sum_y y\, h(y)$

$= 1 \times \frac{8}{20} + 2 \times \frac{4}{20} + 3 \times \frac{8}{20} = 2$

$$\Rightarrow \operatorname{Cov}[x,y] = \frac{38}{20} - \frac{19}{20} \times 2 = 0 \,//$$

Mas,

$$f(0,1) = \frac{3}{20} \neq g(0) \times h(1) = \frac{4}{25} \,//$$

Exemplo

X: número de moradores
Y: número de televisores

X e Y são independentes?

Função de probabilidade conjunta e marginais

f(x,y)		x		
y	1	2	3	h(y)
0	0,02	0	0	0,02
1	0,20	0,08	0,05	0,33
2	0,10	0,20	0,15	0,45
3	0	0,10	0,10	0,20
g(x)	0,32	0,38	0,30	1

$f(1,0) = 0{,}02 \neq g(1)\,h(0) = 0.0064$

X e Y não são independentes.

MATRIZ DE VARIÂNCIA E COVARIÂNCIA

A *matriz de variância e covariância* de duas v.a. X e Y é dada por

$$V(X,Y) = \begin{pmatrix} V(X) & Cov(X,Y) \\ Cov(Y,X) & V(Y) \end{pmatrix}$$

Como Cov(X,Y) = Cov(Y,X) podemos escrever

$$V(X,Y) = \begin{pmatrix} V(X) & Cov(X,Y) \\ Cov(X,Y) & V(Y) \end{pmatrix}$$

Exemplo

X: número de moradores
Y: número de televisores

A matriz de variância e covariância de X e Y é dada por

Função de probabilidade conjunta e marginais

f(x,y)		x		
y	1	2	3	h(y)
0	0,02	0	0	0,02
1	0,20	0,08	0,05	0,33
2	0,10	0,20	0,15	0,45
3	0	0,10	0,10	0,20
g(x)	0,32	0,38	0,30	1

$\hookrightarrow V[X] = E[X^2] - (E[X])^2$

$\cdot\ E[X^2] = \sum_x x^2 g(x)$

$= \cdots = 4{,}54$

$\cdot\ E[X] = 1{,}98$

$\Rightarrow V[X] = 4{,}54 - (1{,}98)^2$

$= 0{,}6196$

$\hookrightarrow V[Y] = E[Y^2] - (E[Y])^2$

$\cdot\ E[Y^2] = \sum_y y^2 h(y)$

$= \cdots = 3{,}93$

$\cdot\ E[Y] = 1{,}83$

$\Rightarrow V[Y] = 3{,}93 - (1{,}83)^2$

$= 0{,}5811$ //

$V[X,Y] = \begin{pmatrix} 0{,}6196 & 0{,}2866 \\ 0{,}2866 & 0{,}5811 \end{pmatrix}$

Exemplo

Considere as intensidades elétricas de duas componentes de um sistema eletrônico, X e Y, que são v.a. A função densidade de probabilidade conjunta de X e Y é dada por

$$f(x,y) = \begin{cases} x + \frac{y}{2} + xy, & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{c.c.} \end{cases}$$

A matriz de variância-covariância de X e Y é

$$V[X,Y] = \begin{pmatrix} V[X] & cov[X,Y] \\ cov[Y,X] & V[Y] \end{pmatrix}$$

$\hookrightarrow V[X] = E[X^2] - (E[X])^2$

$\cdot\ E[X^2] = \int_{-\infty}^{\infty} x^2 g(x)\, dx$

$= \int_0^1 x^2 \left(\frac{3x}{2} + \frac{1}{4}\right) dx$

$= \dots = 0{,}4583$

$\cdot\ E[X] = 0{,}6250$

$\Rightarrow V[X] = 0{,}4583 - (0{,}6250)^2$

$= 0{,}0677$

$\hookrightarrow V[Y] = E[Y^2] - (E[Y])^2$

$\cdot\ E[Y^2] = \int_{-\infty}^{\infty} y^2 h(y)\, dy$

$= \int_0^1 y^2 \left(y + \frac{1}{2}\right) dy$

$= \dots = 0{,}4167$

$\cdot\ E[Y] = 0{,}5833$

$\Rightarrow V[Y] = 0{,}4167 - (0{,}5833)^2$

$= 0{,}0764$

Logo,

$$V[X,Y] = \begin{pmatrix} 0{,}0677 & -0{,}0035 \\ -0{,}0035 & 0{,}0764 \end{pmatrix}$$

CORRELAÇÃO

Sejam X e Y v.a. então a *correlação populacional* entre X e Y é dada por

$$Cor(X,Y) = \rho_{XY} = \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}}$$

CARACTERÍSTICAS DA CORRELAÇÃO

- É uma medida admensional.
- Pertence ao intervalo [-1,1].
- Se igual a -1 indica uma relação linear perfeita e decrescente.
- Se igual a 1 indica uma relação linear perfeita e crescente.
- Se as v.a. são independentes sua correlação é zero.
- Se é zero, nada podemos afirmar quanto a dependência das v.a.
- CUIDADO: Não confundir com a correlação amostral conhecida como coeficiente de correlação de Pearson.

Exemplo

Sejam X uma v.a. e $Y = 2X + 1$. Obtenha a correlação de (X,Y).

COMBINAÇÃO LINEAR DE V.A. NORMAIS INDEPENDENTES

Sejam X e Y duas v.a. Independentes tais que
$X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ e $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$. Então

$$aX + bY \sim \mathcal{N}(a\mu_x + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2)$$

para a e b constantes.

CASO GERAL

Sejam $X_1, X_2, \ldots, X_n$ tais que

$$X_i \sim \mathcal{N}(\mu_i, \sigma_i^2), i = 1, \ldots, n.$$

Então

$$Y = a_1 X_1 + a_2 X_2 + \ldots + a_n X_n \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$$

em que

$$\mu_Y = \sum_{i=1}^{n} a_i \mu_i \text{ e } \sigma_Y^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2$$

CASO PARTICULAR

Conisere que $\mu_i = \mu \text{ e } \sigma_i^2 = \sigma^2, i = 1, \ldots, n$, e $a_i = \frac{1}{n}, i = 1, \ldots, n.$ Assim

$$Y = \sum_{i=1}^{n} \frac{1}{n} X_i = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad \therefore Y = \overline{X}$$

Pelo resultado anterior $\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$