$\cancel{E[x,y]}$ $E[xy]$

Covariance and Correlation Coefficient

(4.2)

Variance is a special case of covariance

σ_{xy} Covariance

Covariance

$$\mu_X = E(X); \ \mu_Y = E(Y)$$

$Var[X] = Cov[X, X]$

Var →

$$\sigma_X^2 = E[(X - \mu_X)^2]; \ \sigma_Y^2 = E[(Y - \mu_Y)^2]$$

$= E[(X-\mu_X)(X-\mu_X)]$

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

The Covariance of X and Y is defined as follows:

notation

$$Cov[X,Y] = E[(X - \mu_X)(Y - \mu_Y)] = \sigma_{XY}$$

$Var[2] = 0$

Is ~~Cov(X²) = Var(X)?~~ NO

$Cov(\underbrace{X, X}) = Var(X)$ ✓

$Cov(2,2) = E[(2-2)(2-2)] = 0$

$Cov[X,2] = E[(X-\mu_x)(2-2)] = 0$

Covariance

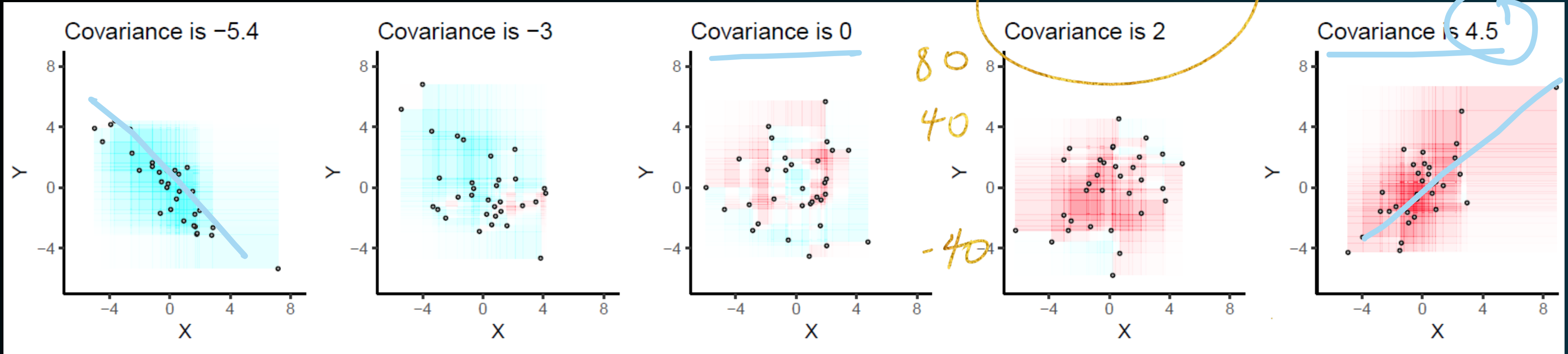
$$\underbrace{\mathrm{Cov}[X,Y]} = \underline{E[(X-\mu_X)(Y-\mu_Y)]} = E(XY - \mu_X Y - \mu_Y X + \mu_X\mu_Y)$$
$$= E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X\mu_Y$$
$$= E(XY) - \mu_X\mu_Y - \mu_Y\mu_X + \mu_X\mu_Y = E(XY) - \mu_X\mu_Y$$

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

$\mathrm{Var}[X] =$
$\mathrm{Cov}[X,X] = E[XX] - E[X]E[X]$

$$E[(x-\mu_x)(y-\mu_y)]$$

Covariance



-40 0 40 80

km

2

r^2

$a^2 = x^2 + y^2$

rho

The correlation coefficient, ρ

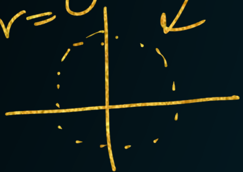
r

$$\rho = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

.9999

.3

$r=0$

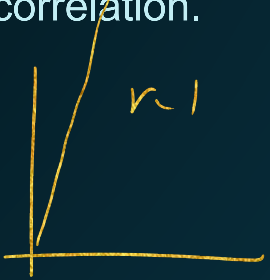


$$-1 \leq \rho \leq 1$$

$r=1$

- If $\rho_{XY} = 1$, X and Y are perfectly, positively, linearly correlated.
- If $\rho_{XY} = -1$, X and Y are perfectly, negatively, linearly correlated.
- If $\rho_{XY} = 0$, X and Y have no **linear** correlation.
- If $\rho_{XY} > 0$, X and Y have positive linear correlation.
- If $\rho_{XY} < 0$, X and Y have negative linear correlation.

$r=.9$



$r=1$

$r=1$

$y = x^2$

- A) If Cov(X,Y) = 0, does this mean rho = 0? yes
- B) If X and Y are independent, does this mean that Cov(X,Y) = 0? yes
- C) If Cov(X,Y) = 0, does this mean X and Y are independent? no

Calculating E[X]

$E[Y] = \sum_{x}\sum_{y} y\, f(x,y)$

E[X]

Discrete

$$= \sum_{x}\sum_{y} x f(x,y)$$

$$= \sum_{x} x f_X(x)$$

Continuous

$$= \int_{x}\int_{y} x f(x,y)\, dy\, dx$$

$$= \int_{x} x f(x)\, dx$$

Calculating E[XY]

E[XY]

- Discrete $= \sum \sum xy \cdot f(x,y)$
- Continuous $= \int_x \int_y xy \cdot f(x,y)\, dy\, dx$

Covariance Example

ρ_{xy}

$E[XY] - E[X]E[Y]$

Let $f_{XY}(x, y) = 3x, \quad 0 \le y \le x \le 1$

$\dfrac{Cov(X,Y)}{\sigma_x \sigma_y}$

→ Find Cor(X,Y).

The marginal pdfs, expectations and variances of X and Y are

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_0^x 3x dy = 3x^2, \qquad 0 \le x \le 1,$$

notation

$$\Longrightarrow E_{f_X}[X] = \int_{-\infty}^{\infty} x f_X(x)dx = \int_0^1 x \times 3x^2 dx = \left[\frac{3}{4}x^4\right]_0^1 = \frac{3}{4},$$

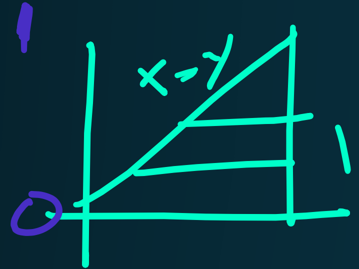
$$E_{f_X}[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_0^1 x^2 \times 3x^2 dx = \left[\frac{3}{5}x^5\right]_0^1 = \frac{3}{5},$$

$$\Longrightarrow Var_{f_X}[X] = E_{f_X}[X^2] - \{E_{f_X}[X]\}^2 = \frac{3}{5} - \left\{\frac{3}{4}\right\}^2 = \frac{3}{80}.$$

- Source: http://wwwf.imperial.ac.uk/~ayoung/m2s1/Covariancecalculations.PDF

Covariance Example (continued)

$$f_{XY}(x,y) = 3x, \qquad 0 \le y \le x \le 1$$



$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \int_y^1 3x dx = \left[\frac{3}{2}x^2\right]_y^1 = \frac{3}{2}(1-y^2), \qquad 0 \le y \le 1,$$

$$\implies E_{f_Y}[Y] = \int_{-\infty}^{\infty} y f_Y(y)dy = \int_0^1 y \times \frac{3}{2}(1-y^2)dy = \left[\frac{3}{2}\left(\frac{y^2}{2}-\frac{y^4}{4}\right)\right]_0^1 = \frac{3}{2}\left(\frac{1}{2}-\frac{1}{4}\right) = \boxed{\frac{3}{8}},$$

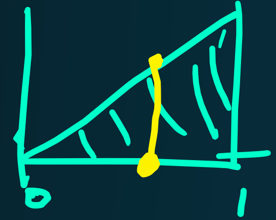
$$E_{f_Y}[Y^2] = \int_{-\infty}^{\infty} y^2 f_Y(y)dy = \int_0^1 y^2 \times \frac{3}{2}(1-y^2)dy = \left[\frac{3}{2}\left(\frac{y^3}{3}-\frac{y^5}{5}\right)\right]_0^1 = \frac{3}{2}\left(\frac{1}{3}-\frac{1}{5}\right) = \boxed{\frac{1}{5}},$$

$$\implies Var_{f_Y}[Y] = E_{f_Y}[Y^2] - \{E_{f_Y}[Y]\}^2 = \frac{1}{5} - \left\{\frac{3}{8}\right\}^2 = \boxed{\frac{19}{320}},$$

Source: http://wwwf.imperial.ac.uk/~ayoung/m2s1/Covariancecalculations.PDF

Covariance Example (continued)

$$f_{XY}(x,y) = 3x, \qquad 0 \le y \le x \le 1$$



$$E_{f_{X,Y}}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x,y)\,dy\,dx = \boxed{\int_0^1\int_0^x} xy \times 3x\,dy\,dx$$

$$= \int_0^1 \left\{\int_0^x y\,dy\right\} 3x^2\,dx = \int_0^1 \left[\frac{y^2}{2}\right]_0^x 3x\,dx = \int_0^1 \frac{x^2}{2} \times 3x^2\,dx$$

$$= \frac{3}{2}\left[\frac{x^5}{5}\right]_0^1 = \boxed{\frac{3}{10}},$$

σ_{xy}

$$\Longrightarrow Cov_{\cancel{f_{X,Y}}}[X,Y] = E_{\cancel{f_{X,Y}}}[XY] - E_{\cancel{f_X}}[X]\,E_{\cancel{f_Y}}[Y] = \boxed{\frac{3}{10} - \frac{3}{4} \times \frac{3}{8}} = \boxed{\frac{3}{160}}$$

ρ_{xy}

$$Corr_{f_{X,Y}}[X,Y] = \frac{Cov_{f_{X,Y}}[X,Y]}{\sqrt{Var_{f_X}[X] \times Var_{f_Y}[Y]}} = \frac{\frac{3}{160}}{\sqrt{\frac{3}{80} \times \frac{19}{320}}} = \boxed{0.397.}$$

Source: http://wwwf.imperial.ac.uk/~ayoung/m2s1/Covariancecalculations.PDF

notes

Textbook Example 4.2-1 (Covariance)

Discrete

Let X and Y have the joint pmf

$$f(x,y) = \frac{x+2y}{18}, \qquad x = 1,2, \qquad y = 1,2$$

Find Cov(X,Y)

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

① find $f(x)$ & $f(y)$

② use " " to find $E[X]$, $E[Y]$

③ $E[XY]$ $\displaystyle\sum_{x=1}^{2}\sum_{y=1}^{2} xy \cdot \left(\frac{x+2y}{18}\right)$

Read the textbook! (please)

Textbook Example 4.2-1 (Covariance)

Let X and Y have the joint pmf

$$f(x,y) = \frac{x+2y}{18}, \qquad x = 1,2, \qquad y = 1,2$$

(1,2) S

1,2 D

$$f(x) = \sum_{y=1}^{2} \frac{x+2y}{18} = \frac{x+2(1)}{18} + \frac{x+2(2)}{18}$$

$$= \frac{x+2}{18} + \frac{x+4}{18} = \frac{2x+6}{18} = \frac{x+3}{9}$$

$x = 1, 2$

Read the textbook! (please)

$$\mathrm{Cov}(X,Y) = \sum_{x=1}^{2}\sum_{y=1}^{2} xy\,\frac{x+2y}{18} - \left(\frac{14}{9}\right)\left(\frac{29}{18}\right)$$

$$= \frac{45}{18} - \frac{406}{162} = -\frac{1}{162}.$$