Covariance and Correlation Coefficient

(4.2)

Covariance

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

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

The Covariance of X and Y is defined as follows:

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

Covariance

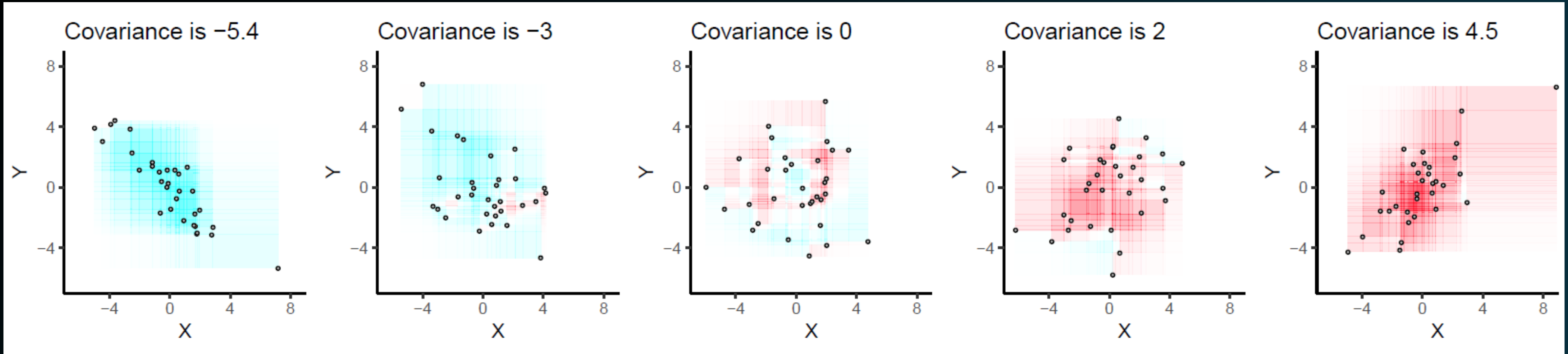
Cov[X,Y] =

$$
\begin{aligned}
E[(X-\mu_X)(Y-\mu_Y)] &= E(XY-\mu_X Y-\mu_Y X+\mu_X\mu_Y) \\
&= E(XY)-\mu_X E(Y)-\mu_Y E(X)+\mu_X\mu_Y
\end{aligned}
$$

$$
= E(XY)-\mu_X\mu_Y-\mu_Y\mu_X+\mu_X\mu_Y = E(XY)-\mu_X\mu_Y
$$

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

Covariance



Covariance is −5.4
Covariance is −3
Covariance is 0
Covariance is 2
Covariance is 4.5
8
4
0
−4
−4
0
4
8
X
Y

The correlation coefficient, ρ

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

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

- 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.

Calculating E[X]

E[X]

- Discrete
  - $$= \sum_x \sum_y x f(x, y)$$
  - $$= \sum_x x f_X(x)$$
- Continuous
  - $$= \int_{\mathcal{X}} \int_{\mathcal{Y}} x f(x, y)\, \mathrm{d}y\, \mathrm{d}x$$
  - $$= \int_{\mathcal{X}} x f(x)\, \mathrm{d}x$$

Calculating E[XY]

- E[XY]
  - Discrete $= \sum \sum xy \cdot f(x, y)$
  - Continuous $= \int_{\mathcal{X}} \int_{\mathcal{Y}} xy \cdot f(x, y)\, \mathrm{d}y\, \mathrm{d}x$

Covariance Example

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

Find Cor(X,Y).

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

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_0^x 3x dy = 3x^2, \qquad 0 \le x \le 1, \\
\Longrightarrow E_{f_X}[X] &= \int_{-\infty}^{\infty} x f_X(x)dx = \int_0^1 x \times 3x^2 dx = \left[\frac{3}{4}x^4\right]_0^1 = \frac{3}{4}, \\
E_{f_X}[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_0^1 x^2 \times 3x^2 dx = \left[\frac{3}{5}x^5\right]_0^1 = \frac{3}{5}, \\
\Longrightarrow Var_{f_X}[X] &= E_{f_X}[X^2] - \{E_{f_X}[X]\}^2 = \frac{3}{5} - \left\{\frac{3}{4}\right\}^2 = \frac{3}{80}.
\end{aligned}
$$

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

Covariance Example (continued)

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

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \int_y^1 3x dx = \left[\frac{3}{2}x^2\right]_y^1 = \frac{3}{2}(1-y^2), \qquad 0 \le y \le 1, \\
\Longrightarrow E_{f_Y}[Y] &= \int_{-\infty}^{\infty} y f_Y(y)dy = \int_0^1 y \times \frac{3}{2}(1-y^2)dy = \left[\frac{3}{2}\left(\frac{y^2}{2} - \frac{y^4}{4}\right)\right]_0^1 = \frac{3}{2}\left(\frac{1}{2} - \frac{1}{4}\right) = \frac{3}{8}, \\
E_{f_Y}[Y^2] &= \int_{-\infty}^{\infty} y^2 f_Y(y)dy = \int_0^1 y^2 \times \frac{3}{2}(1-y^2)dy = \left[\frac{3}{2}\left(\frac{y^3}{3} - \frac{y^5}{5}\right)\right]_0^1 = \frac{3}{2}\left(\frac{1}{3} - \frac{1}{5}\right) = \frac{1}{5}, \\
\Longrightarrow Var_{f_Y}[Y] &= E_{f_Y}[Y^2] - \{E_{f_Y}[Y]\}^2 = \frac{1}{5} - \left\{\frac{3}{8}\right\}^2 = \frac{19}{320},
\end{aligned}$$

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

Covariance Example (continued)

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

$$\begin{aligned}
E_{f_{X,Y}}[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x,y)\,dy\,dx = \int_0^1\int_0^x xy \times 3x\,dy\,dx \\
&= \int_0^1 \left\{\int_0^x y\,dy\right\} 3x^2\,dx = \int_0^1 \left[\frac{y^2}{2}\right]_0^x 3x\,dx = \int_0^1 \frac{x^2}{2} \times 3x^2\,dx \\
&= \frac{3}{2}\left[\frac{x^5}{5}\right]_0^1 = \frac{3}{10}, \\
\Longrightarrow Cov_{f_{X,Y}}[X,Y] &= E_{f_{X,Y}}[XY] - E_{f_X}[X]\,E_{f_Y}[Y] = \frac{3}{10} - \frac{3}{4} \times \frac{3}{8} = \frac{3}{160} \\
Corr_{f_{X,Y}}[X,Y] &= \frac{Cov_{f_{X,Y}}[X,Y]}{\sqrt{Var_{f_X}[X] \times Var_{f_Y}[Y]}} = \frac{\frac{3}{160}}{\sqrt{\frac{3}{80} \times \frac{19}{320}}} = 0.397.
\end{aligned}$$

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

notes

Textbook Example 4.2-1 (Covariance)

Let X and Y have the joint pmf

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

Find Cov(X,Y)

Read the textbook! (please)

Textbook Example 4.2-1 (Covariance)

Let X and Y have the joint pmf

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

Read the textbook! (please)

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

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