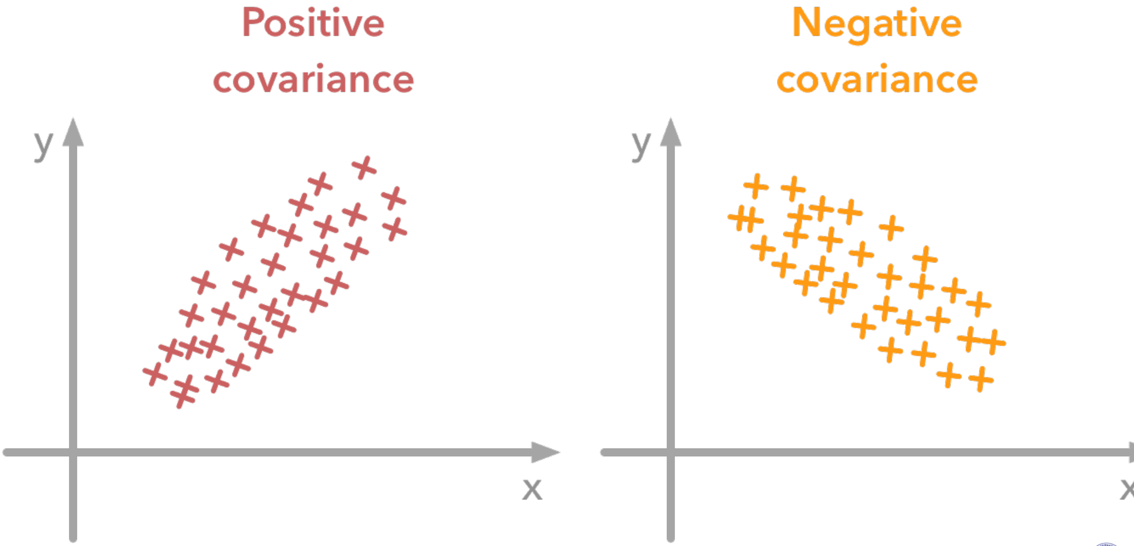
Variance - Covariance

November 10, 2020



How two RV vary together or relationship between the RV?

$$E(g(X,Y)) = \begin{cases} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dxdy, \\ \qquad\qquad\qquad\qquad X, Y \text{ continuous} \\ \sum_{x,y} g(x,y) p_{X,Y}(x,y), \\ \qquad\qquad\qquad\qquad X, Y \text{ discrete} \end{cases}$$

X and Y are two RVs. The **covariance** of X and Y is

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

Alternative formula

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

X and Y are two RVs. The **covariance** of X and Y is

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

Alternative formula

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

2 indicator RVs

$$X = \begin{cases} 1, & \text{component 1 fails} \\ 0, & otherwise \end{cases}$$

$$Y = \begin{cases} 1, & \text{component 2 fails} \\ 0, & otherwise \end{cases}$$

- $E(X) = P(X = 1)$
- $E(Y) = P(Y = 1)$
-
$$XY = \begin{cases} 1, & X = Y = 1 \\ 0, & otherwise \end{cases}$$

  $E(XY) = P(X = 1, Y = 1)$
-

- $$\begin{aligned}\operatorname{Cov}(X,Y) &= E(XY) - E(X)E(Y)\\ &= P(X=1,Y=1) - P(X=1)P(Y=1)\end{aligned}$$

- $$\begin{aligned}&\operatorname{Cov}(X,Y) > 0\\ &\Leftrightarrow P(X=1,Y=1) - P(X=1)P(Y=1) > 0\\ &\Leftrightarrow \frac{P(X=1,Y=1)}{P(X=1)} > P(Y=1)\\ &\Leftrightarrow P(Y=1|X=1) > P(Y=1)\end{aligned}$$

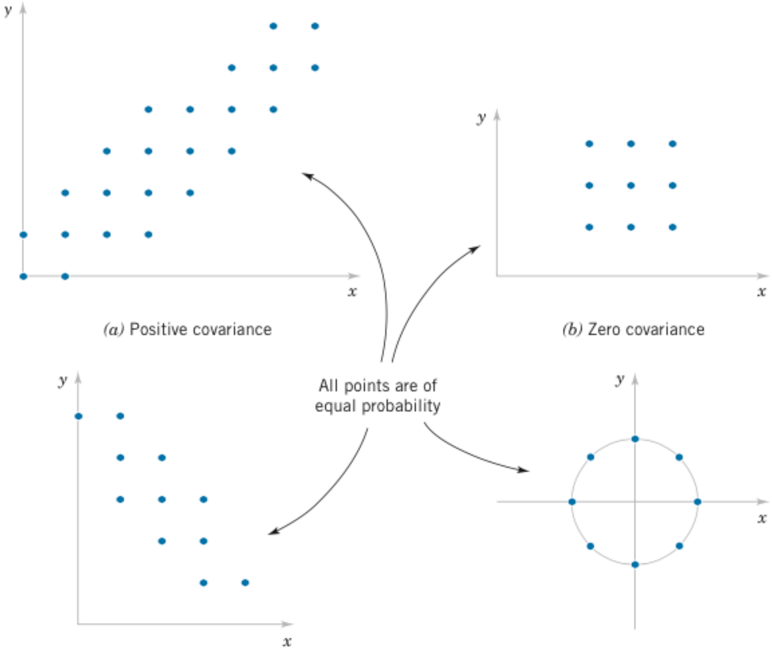
- $$\begin{aligned}&\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)\\&= P(X=1, Y=1) - P(X=1)P(Y=1)\end{aligned}$$

- $$\begin{aligned}&\mathrm{Cov}(X,Y) > 0\\&\Leftrightarrow P(X=1, Y=1) - P(X=1)P(Y=1) > 0\\&\Leftrightarrow \frac{P(X=1, Y=1)}{P(X=1)} > P(Y=1)\\&\Leftrightarrow P(Y=1|X=1) > P(Y=1)\end{aligned}$$

- The covariance shows the relation between X and Y
- If $\mathrm{Cov}(X, Y) > 0$ then Y tends to increase when X increases
- If $\mathrm{Cov}(X < Y) < 0$ then Y tends to decrease when X increases

- The covariance shows the relation between X and Y
- If $\text{Cov}(X, Y) > 0$ then Y tends to increase when X increases
- If $\text{Cov}(X < Y) < 0$ then Y tends to decrease when X increases

- The covariance shows the relation between X and Y
- If $\text{Cov}(X, Y) > 0$ then Y tends to increase when X increases
- If $\text{Cov}(X < Y) < 0$ then Y tends to decrease when X increases



(a) Positive covariance

(b) Zero covariance

- $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$
- $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$
- $\mathrm{Cov}(aX,Y) = a\,\mathrm{Cov}(X,Y)$
- $\mathrm{Cov}(X+Y,Z) = \mathrm{Cov}(X,Z) + \mathrm{Cov}(Y,Z)$

- $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
- $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
- $\mathrm{Cov}(aX, Y) = a\,\mathrm{Cov}(X, Y)$
- $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$

- $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$
- $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$
- $\mathrm{Cov}(aX,Y) = a\,\mathrm{Cov}(X,Y)$
- $\mathrm{Cov}(X+Y,Z) = \mathrm{Cov}(X,Z) + \mathrm{Cov}(Y,Z)$

- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- $\text{Cov}(X, X) = \text{Var}(X)$
- $\text{Cov}(aX, Y) = a\,\text{Cov}(X, Y)$
- $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$

1.

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$$

2. $X_1, \ldots, X_n$: RVs

$$\mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j)$$

1.

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$$

2. $X_1, \ldots, X_n$: RVs

$$\mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j)$$

Recall: X and Y are independence if for all x, y

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

If so then

$$\mathrm{Cov}(X, Y) = 0$$

$$
\begin{aligned}
E[XY] &= \sum_j \sum_i x_i y_j P\{X = x_i, Y = y_j\} \\
&= \sum_j \sum_i x_i y_j P\{X = x_i\} P\{Y = y_j\} \quad \text{by independence} \\
&= \sum_y y_j P\{Y = y_j\} \sum_i x_i P\{X = x_i\} \\
&= E[Y]E[X]
\end{aligned}
$$

$\text{Cov}(X, Y) = 0$ does not imply that X and Y are independent.

- If $X_i, \ldots, X_n$ are independent then

$$\mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

- If $X_i, \ldots, X_n$ are i.i.d then

$$\mathrm{Var}(\sum_{i=1}^{n} X_i) = n\,\mathrm{Var}(X_1)$$

$$\rho_{XY} = \frac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var}(X)}\sqrt{\operatorname{Var}(Y)}}$$

use to compare linear relationship between pair of RVs

- $-1 \leq \rho_{XY} \leq 1$
- If $\rho > 0$ (or $\rho < 0$) then the values of $x - E(X)$ and $Y - E(Y)$ "tend" to have the same (or opposite) sign
- $\rho = \pm 1$ if and only if Y is a linear function of X

$$X - E(Y) = c(X - E(X))$$

for some c

Example

- n independent tosses of a biased coin
- X: number of heads
- Y: number of tails
- $X + Y = n$
- $E(X) + E(Y) = n$
- $X - E(X) = Y - E(Y)$

- n independent tosses of a biased coin
- X: number of heads
- Y: number of tails
- $X + Y = n$
- $E(X) + E(Y) = n$
- $X - E(X) = Y - E(Y)$

- $\operatorname{Var}(X) = E(X - E(X))^2 = E(Y - E(Y))^2 = \operatorname{Var}(Y)$
- $\operatorname{Cov}(X, Y) = E((X - E(X)))(Y - E(Y)) = -E(X - E(X))^2 = -\operatorname{Var}(X)$
- $$\begin{aligned}\rho_{X,Y} &= \frac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var}(X)}\sqrt{\operatorname{Var}(Y)}} \\ &= \frac{-\operatorname{Var}(X)}{\sqrt{\operatorname{Var}(X)}\sqrt{\operatorname{Var}(X)}} = -1\end{aligned}$$

Let $X_1 \hookrightarrow \mathcal{N}(\mu_1, \sigma_1^2)$, $X_2 \hookrightarrow \mathcal{N}(\mu_2, \sigma_2^2)$ be 2 normal random variables with covariance σ_{12} then the random vector $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ is a bivariate normal distribution if the joint pdf is given by

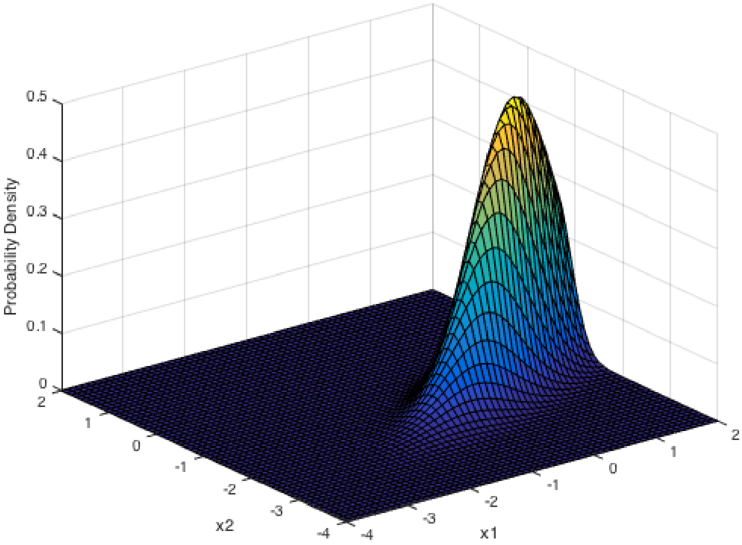
$$f(x_1, x_2) = \frac{1}{2\pi \det(\Sigma)^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$$

where

- $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$
- $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$ is the variance - covariance matrix of X

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{1-\rho^2}\left(\frac{(x_1-\mu_1)^2}{2\sigma_1^2} + \frac{(x_2-\mu_2)^2}{2\sigma_2^2} - 2\rho\frac{(x_1-\sigma_1)(x_2-\sigma_2)}{\sigma_1\sigma_2}\right)}$$

where $\rho = \frac{\sigma_{12}}{\sigma_1\sigma_2}$ is the correlation coefficient of X_1 and X_2


Probability Density
0.5
0.4
0.3
0.2
0.1
0
x2
2
1
0
-1
-2
-3
-4
x1
-4
-3
-2
-1
0
1
2

Let X and Y be jointly normal random variables with parameters $\mu_X = 1$, $\sigma_X^2 = 1$, $\mu_Y = 0$, $\sigma_Y^2 = 4$, and $\rho = 1/2$.

1. Find $P(2X + Y \leq 3)$.
2. Find $Cov(X + Y, 2X - Y)$
3. Find $P(Y > 1|X = 2)$.

(X, Y) is a bivariate normal distribution if and only if $aX + bY$ is normally distributed for all a, b

Z_1, Z_2 are independent $\mathcal{N}(0,1)$. Let

$$\begin{cases} X = Z_1 \\ Y = \rho Z_1 + \sqrt{1-\rho^2} Z_2 \end{cases}$$

1. Show that (X, Y) are bivariate normal.
2. Find $\rho(X, Y)$.
3. Find the joint PDF of X and Y.

1. For any a and b, we have

$$W = aX + bY = (a + \rho b)Z_1 + b\sqrt{1 - \rho^2}Z_2 \sim \mathcal{N}(0, (a + \rho b)^2 + b^2(1 - \rho^2))$$

2. $Var(X) = 1$, $Var(Y) = \rho^2 + (1 - \rho^2) = 1$ and by linear property of covariance, we have

$$Cov(X, Y) = \rho \underbrace{Cov(Z_1, Z_1)}_{Var(Z_1)} + \sqrt{1 - \rho^2}Cov(\underbrace{Z_1, Z_2}_{independent}) = \rho + 0 = \rho$$

3.

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{1 - \rho^2}} e^{-\frac{1}{1-\rho^2}\left(\frac{(x_1-\mu_1)^2}{2} + \frac{(x_2-\mu_2)^2}{2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{2}\right)}$$

$$\begin{cases} X = \sigma_1 Z_1 + \mu_1 \\ Y = \sigma_2(\rho Z_1 + \sqrt{1-\rho^2} Z_2) + \mu_2 \end{cases}$$

where Z_1, Z_2 are independent $\mathcal{N}(0, 1)$
Then $(X, Y) \sim \mathcal{N}(\mu, \Sigma)$ with

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

and

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2 \end{pmatrix}$$

with $\sigma_{12} = \rho\sigma_1\sigma_2$