

ECE 302: Lecture 5.3 Covariance

Prof Stanley Chan

School of Electrical and Computer Engineering
Purdue University



Joint Expectation

Recall:

$$\mathbb{E}[X] = \int_{\Omega} xf_X(x)dx.$$

How about the expectation for two variables?

Definition

Let X and Y be two random variables. The **joint expectation** is

$$\mathbb{E}[XY] = \sum_{y \in \Omega_Y} \sum_{x \in \Omega_X} xy \cdot p_{X,Y}(x,y) \quad (1)$$

if X and Y are discrete, or

$$\mathbb{E}[XY] = \int_{y \in \Omega_Y} \int_{x \in \Omega_X} xy \cdot f_{X,Y}(x,y)dxdy \quad (2)$$

if X and Y are continuous. Joint expectation is also called **correlation**.

Outline

- Joint PDF and CDF
- Joint Expectation
- Conditional Distribution
- Conditional Expectation
- Sum of Two Random Variables
- Random Vectors
- High-dimensional Gaussians and Transformation
- Principal Component Analysis

Today's lecture

- What is joint expectation?
- Why define joint expectation as $\mathbb{E}[XY]$?
- Independence
- Covariance
- Correlation

Covariance

Definition

Let X and Y be two random variables. Then the covariance of X and Y is

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)], \quad (3)$$

where $\mu_X = \mathbb{E}[X]$ and $\mu_Y = \mathbb{E}[Y]$.

Remark:

$$\text{Cov}(X, X) = \mathbb{E}[(X - \mu_X)(X - \mu_X)] = \text{Var}[X]$$

Covariance

Theorem

Let X and Y be two random variables. Then,

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \quad (4)$$

Proof. Just apply the definition of covariance:

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}[(X - \mu_X)(Y - \mu_Y)] \\ &= \mathbb{E}[XY - X\mu_Y - Y\mu_X + \mu_X\mu_Y] = \mathbb{E}[XY] - \mu_X\mu_Y. \quad \square \end{aligned}$$

Properties

Theorem

For any X and Y ,

a. $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.

Proof. Recall the definition of joint expectation:

$$\begin{aligned}\mathbb{E}[X + Y] &= \sum_y \sum_x (x + y)p_{X,Y}(x, y) \\ &= \sum_y \sum_x xp_{X,Y}(x, y) + \sum_y \sum_x yp_{X,Y}(x, y) \\ &= \sum_x x \left(\sum_y p_{X,Y}(x, y) \right) + \sum_y y \left(\sum_x p_{X,Y}(x, y) \right) \\ &= \sum_x xp_X(x) + \sum_y yp_Y(y) = \mathbb{E}[X] + \mathbb{E}[Y].\end{aligned}$$

Properties

Theorem

For any X and Y ,

b. $\text{Var}[X + Y] = \text{Var}[X] + 2\text{Cov}(X, Y) + \text{Var}[Y]$.

Proof.

$$\begin{aligned}\text{Var}[X + Y] &= \mathbb{E}[(X + Y)^2] - \mathbb{E}[X + Y]^2 = \mathbb{E}[(X + Y)^2] - (\mu_X + \mu_Y)^2 \\ &= \mathbb{E}[X^2 + 2XY + Y^2] - (\mu_X^2 + 2\mu_X\mu_Y + \mu_Y^2) \\ &= \mathbb{E}[X^2] - \mu_X^2 + \mathbb{E}[Y^2] - \mu_Y^2 + 2(\mathbb{E}[XY] - \mu_X\mu_Y) \\ &= \text{Var}[X] + 2\text{Cov}(X, Y) + \text{Var}[Y]. \quad \square\end{aligned}$$

Correlation Coefficient

Definition

Let X and Y be two random variables. The **correlation coefficient** is

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}[X]\text{Var}[Y]}}. \quad (5)$$

- ρ is always between -1 and 1, i.e., $-1 \leq \rho \leq 1$. This is due to the cosine angle definition.
- When $X = Y$ (fully correlated), $\rho = +1$.
- When $X = -Y$ (negatively correlated), $\rho = -1$.
- When X and Y uncorrelated then $\rho = 0$.

Independence

Theorem

If X and Y are independent, then

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]. \quad (6)$$

Proof.

$$\begin{aligned} \mathbb{E}[XY] &= \sum_y \sum_x xyp_{X,Y}(x,y) \\ &= \sum_y \sum_x xyp_X(x)p_Y(y) \\ &= \left(\sum_x xp_X(x) \right) \left(\sum_y yp_Y(y) \right) = \mathbb{E}[X]\mathbb{E}[Y]. \quad \square \end{aligned}$$

Properties

Theorem

Consider the following two statements:

- X and Y are independent;
- $\text{Cov}(X, Y) = 0$.

It holds that (a) implies (b), but (b) does not imply (a). Thus, independence is a stronger condition than correlation.

Proof: See ebook.

What is the relationship between independence and uncorrelated?

- Independence \Rightarrow uncorrelated.
- Independence \nLeftarrow uncorrelated.

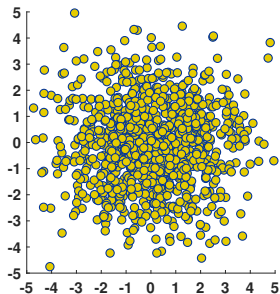
Computation

Theory:

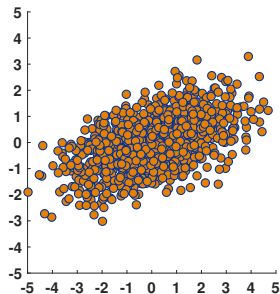
$$\rho = \frac{\mathbb{E}[XY] - \mu_X \mu_Y}{\sigma_X \sigma_Y}.$$

Practice:

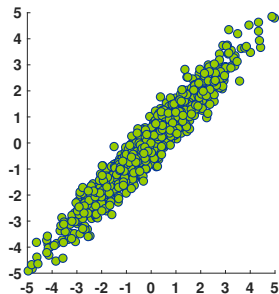
$$\hat{\rho} = \frac{\frac{1}{N} \sum_{n=1}^N x_n y_n - \bar{x} \bar{y}}{\sqrt{\frac{1}{N} \sum_{n=1}^N (x_n - \bar{x})^2} \sqrt{\frac{1}{N} \sum_{n=1}^N (y_n - \bar{y})^2}}, \quad (7)$$



(a) $\hat{\rho} = -0.0038$



(b) $\hat{\rho} = 0.5321$



(c) $\hat{\rho} = 0.9656$

Summary

Covariance is:

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)], \quad (8)$$

Correlation coefficient is

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}[X]\text{Var}[Y]}}. \quad (9)$$

What is the relationship between independence and uncorrelated?

- Independence \Rightarrow uncorrelated.
- Independence \nLeftarrow uncorrelated.

Questions?