ECE 302: Lecture 5.3 Covariance

Prof Stanley Chan

School of Electrical and Computer Engineering
Purdue University

Joint Expectation

Recall:

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

How about the expectation for two variables?

Definition

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

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

if X and Y are discrete, or

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

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

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

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,

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

Proof. Just apply the definition of covariance:

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

□

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:

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

Properties

Theorem

For any X and Y,

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

Proof.

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

Correlation Coefficient

Definition

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

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

- ρ is always between 0 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

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]. \tag{6}$$

Proof.

$$\begin{aligned}
\mathbb{E}[XY] &= \sum_y \sum_x xy p_{X,Y}(x,y) \\
&= \sum_y \sum_x xy p_X(x) p_Y(y) \\
&= \left(\sum_x x p_X(x)\right)\left(\sum_y y p_Y(y)\right) = \mathbb{E}[X]\mathbb{E}[Y]. \qquad \square
\end{aligned}$$

Properties

Theorem

Consider the following two statements:

a. *X and Y are independent;*

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

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 $\not\Leftarrow$ uncorrelated.

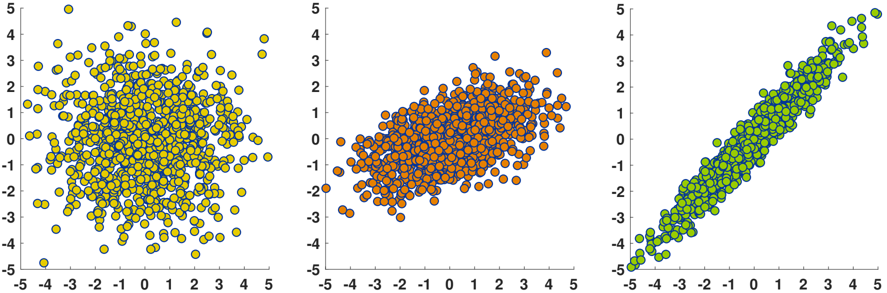
Computation

Theory:

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

Practice:

$$\widehat{\rho} = \frac{\frac{1}{N}\sum_{n=1}^N x_n y_n - \overline{x}\,\overline{y}}{\sqrt{\frac{1}{N}\sum_{n=1}^N (x_n - \overline{x})^2}\sqrt{\frac{1}{N}\sum_{n=1}^N (y_n - \overline{y})^2}}, \tag{7}$$



(a) $\widehat{\rho} = -0.0038$ (b) $\widehat{\rho} = 0.5321$ (c) $\widehat{\rho} = 0.9656$

Summary

Covariance is:

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

Correlation coefficient is

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

What is the relationship between independence and uncorrelated?

- Independence $\Rightarrow$ uncorrelated.
- Independence $\not\Leftarrow$ uncorrelated.

Questions?