LECTURE: COVARIANCE

1. Joint Expectation (continued)

Definition:

If Y_1 and Y_2 be two discrete random variables with joint pmf $p(y_1, y_2)$ then

$$E[g(Y_1, Y_2)] = \sum_{\text{all } y_1} \sum_{\text{all } y_2} g(y_1, y_2) p(y_1, y_2)$$

If Y_1 and Y_2 be two continuous random variables with joint density $f(y_1, y_2)$ then

$$E[g(Y_1, Y_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(y_1, y_2) f(y_1, y_2) dy_1 dy_2$$

A very useful special case is $g(Y_1, Y_2) = Y_1 Y_2$ which will lead to covariance (see below)

Example 1:

Let X and Y be random variables with joint density $f(x, y)$ where:

$$f(x, y) = \begin{cases} 8xy & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $E(XY)$

Here $g(XY) = XY$ and so

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy$$

As usual, looking at the picture, we get

$$\begin{aligned}
E(XY) &= \int_0^1 \int_0^x xy(8xy)dydx \\
&= 8\int_0^1 \int_0^x x^2y^2dydx \\
&= 8\int_0^1 x^2 \left[\frac{y^3}{3}\right]_{y=0}^{y=x} dx \\
&= 8\int_0^1 x^2 \left(\frac{x^3}{3}\right) dx \\
&= \frac{8}{3}\int_0^1 x^5dx \\
&= \frac{8}{3}\left[\frac{x^6}{6}\right]_0^1 \\
&= \frac{8}{18} = \frac{4}{9}
\end{aligned}$$

Fact:

Let Y_1 and Y_2 be two independent random variables. Then

$$E(Y_1Y_2) = E(Y_1)E(Y_2)$$

An analogous result holds for the product of any number of independent random variables.

2. Covariance and Correlation

Heuristically, two random variables are independent if their outcomes do not affect each other.

Suppose we have two random variables X and Y. There are two extreme cases to consider:

(1) X and Y are independent, so they don't affect each other at all, think for instance two coin flips.

(2) X and Y are completely dependent, i.e. the output of one random variable determines the output of the other random. Think for example if $Y = 2X$. In this case, knowledge of output of either random variable gives you knowledge of the output of the other random variable.

There is an entire spectrum between these two extremes. The **covariance** measures how much X and Y depend on each other:

Definition:

Let Y_1 and Y_2 be two random variables, then the **covariance** of Y_1 and Y_2 is

$$\operatorname{Cov}(Y_1, Y_2) = E[(Y_1 - E(Y_1))(Y_2 - E(Y_2))]$$

A larger covariance indicates a greater dependence between Y_1 and Y_2.

Usually $\operatorname{Cov}(Y_1, Y_2)$ depends on the units used for Y_1 and Y_2. To solve this problem, we standardize to get the **correlation coefficient**:

Definition:

Let Y_1 and Y_2 be two random variables with standard deviations σ_1 and σ_2. Then the **correlation coefficient** is:

$$\rho = \frac{\text{Cov}(Y_1, Y_2)}{\sigma_1 \sigma_2}$$

Notice that we always have $-1 \leq \rho \leq 1$ it measures the strength of the relationship between Y_1 and Y_2

Special Cases:

(1) If $\rho = 1$ then we have perfect correlation, like $Y_2 = 2Y_1$, all points of (Y_1, Y_2) fall on a straight line with positive slope.

(2) If $\rho = 0$ then there is no correlation between Y_1 and Y_2, they are uncorrelated

(3) If $\rho = -1$ then we have a perfect negative correlation, like $Y_2 = -2Y_1$, all points of (Y_1, Y_2) fall on a straight line with negative slope.

And any ρ-values in between indicate a correlation in between those extreme cases. For example, if $\rho = 0.5$, then Y_1 and Y_2 are somehow related, although not perfectly.

Warning: Correlation is **not** the same as independence! Independence implies uncorrelated but not the other way around (see below)

Facts:

(1) $\text{Cov}(Y_1, Y_2) = \text{Cov}(Y_2, Y_1)$

(2) $\text{Cov}(X, X) = \text{Var}(X)$

3. Magic Covariance Formula

Just like with variance, the covariance is not generally computed directly. Instead, we use the Magic Covariance Formula:

Magic Covariance Formula:

$$\text{Cov}(Y_1, Y_2) = E(Y_1Y_2) - E(Y_1)E(Y_2)$$

Why? Let $E(Y_1) = \mu_1$ and $E(Y_2) = \mu_2$, then

$$\begin{aligned}
\text{Cov}(Y_1, Y_2) &= E[(Y_1 - \mu_1)(Y_2 - \mu_2)] \\
&= E(Y_1Y_2 - \mu_1Y_2 - \mu_2Y_1 + \mu_1\mu_2) \\
&= E(Y_1Y_2) - \mu_1E(Y_2) - \mu_2E(Y_1) + \mu_1\mu_2 \\
&= E(Y_1Y_2) - \mu_1\mu_2 - \mu_2\mu_1 + \mu_1\mu_2 \\
&= E(Y_1Y_2) - \mu_1\mu_2
\end{aligned}$$

Example 2:

Let X and Y be random variables with joint density $f(x, y)$ where

$$f(x, y) = \begin{cases} 8xy & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $\text{Cov}(X, Y)$

Previously, we have found that $E(X) = 4/5$ and $E(Y) = 8/15$ and $E(XY) = 4/9$.

Using the Magic Covariance Formula,

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 4/9 - (4/5)(8/15) = 4/225$$

4. Independence

What happens if two random variables are independent?

Fact:

If Y_1 and Y_2 are independent, then $\text{Cov}(Y_1, Y_2) = 0$

Why?

$$\text{Cov}(Y_1, Y_2) = E(Y_1 Y_2) - E(Y_1)E(Y_2) = E(Y_1)E(Y_2) - E(Y_1)E(Y_2) = 0$$

Warning: The converse, is not generally true. In other words, if the covariance of two random variables is 0, we **cannot** conclude that they are independent.

The final result in this section concerns the variance of the sum of two random variables:

Fact:

Let Y_1 and Y_2 be two random variables. Then

$$\text{Var}(Y_1 + Y_2) = \text{Var}(Y_1) + \text{Var}(Y_2) + 2\,\text{Cov}(Y_1, Y_2)$$

If Y_1 and Y_2 are independent, then

$$\text{Var}(Y_1 + Y_2) = \text{Var}(Y_1) + \text{Var}(Y_2)$$

Why? To see this, we use the Magic Variance Formula:

$$\begin{aligned}
\text{Var}(Y_1 + Y_2) &= E[(Y_1 + Y_2)^2] - [E(Y_1 + Y_2)]^2 \\
&= E(Y_1^2 + 2Y_1Y_2 + Y_2^2) - [E(Y_1) + E(Y_2)]^2 \\
&= E(Y_1^2) + 2E(Y_1Y_2) + E(Y_2^2) - [E(Y_1)]^2 - 2E(Y_1)E(Y_2) - [E(Y_2)]^2 \\
&= (E(Y_1^2) - [E(Y_1)]^2) + (E(Y_2^2) - E(Y_2)]^2) + 2[E(Y_1Y_2) - E(Y_1)E(Y_2)] \\
&= \text{Var}(Y_1) + \text{Var}(Y_2) + 2\text{Cov}(Y_1, Y_2)
\end{aligned}$$

If Y_1 and Y_2 are independent, the covariance is 0, so we get

$$\text{Var}(Y_1 + Y_2) = \text{Var}(Y_1) + \text{Var}(Y_2) + 2(0) = \text{Var}(Y_1) + \text{Var}(Y_2)$$

We can extend the first result to the case of a sum of more than two random variables, but the result is cumbersome. In the second case, however, the result extends easily:

Fact:

If $Y_1, Y_2, \dots, Y_n$ are independent random variables, then

$$\text{Var}(Y_1 + Y_2 + \cdots + Y_n) = \text{Var}(Y_1) + \text{Var}(Y_2) + \cdots + \text{Var}(Y_n)$$

Congratulations, we are officially done with the probability part of this course! Onto the statistics part! ☺