MTH 224, Spring 2024

Instructor: Bahman Angoshtari

Lecture 12

Section 4.2: covariance and correlation coefficient

12.1. Expected value of a function of two or more random variables (continued)

Example 12.1. (a) Show that

$$(\mathbb{E}[UV])^2 \le \mathbb{E}\left[U^2\right]\mathbb{E}\left[V^2\right], \qquad \text{(the Cauchy–Schwarz inequality)}$$

for any two random variable U and V.
(b) Show that $(\mathbb{E}[UV])^2 = \mathbb{E}\left[U^2\right]\mathbb{E}\left[V^2\right]$ if and only if $tU + V = 0$ for some constant $t \neq 0$.

Solution. (a) Define $f(t) = \mathbb{E}\left[(tU+V)^2\right]$. Clearly, $f(t) \ge 0$ for all $t \in \mathbb{R}$. On the other hand, if we expand the expectation, we get:

$$0 \le f(t) = \mathbb{E}\left[t^2U^2 + 2tUV + V^2\right] = \mathbb{E}\left[U^2\right]\cdot t^2 + 2\mathbb{E}[UV]\cdot t + \mathbb{E}\left[V^2\right].$$

The right hand side is a quadratic function of t that is always non-negative. Therefore, its discriminant is non-positive:

$$(2\mathbb{E}[UV])^2 - 4\mathbb{E}\left[U^2\right]\mathbb{E}\left[V^2\right] \le 0$$

or, equivalently,

$$(\mathbb{E}[UV])^2 \le \mathbb{E}\left[U^2\right]\mathbb{E}\left[V^2\right],$$

as claimed.

(b) Note that if $U = -tV$ for some $t \neq 0$, then

$$(\mathbb{E}[UV])^2 = \mathbb{E}[UV]\mathbb{E}[UV] = \mathbb{E}[-\frac{U^2}{t}]\mathbb{E}[-tV^2] = \mathbb{E}\left[U^2\right]\mathbb{E}\left[V^2\right].$$

Conversely, if $(\mathbb{E}[UV])^2 = \mathbb{E}\left[U^2\right]\mathbb{E}\left[V^2\right]$, then the discrimiant of the quadratic equation $f(t) = 0$ is zero. Thus, we must have $\mathbb{E}\left[(tU+V)^2\right] = 0$ for some constant t. However, this can only be true if $tU + V = 0$, which is what we wanted to show.

12.2. Covariance and correlation

Recall that $\sigma_X^2 = \mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$. For two random variables, we define:

Definition 12.2. The covariance of X and Y is

$$\sigma_{XY} = \mathrm{Cov}(X,Y) = \mathbb{E}\Big[(X - \mathbb{E}[X])\cdot(Y - \mathbb{E}[Y])\Big].$$

The correlation coefficient of X, Y is defined as:

$$\rho_{X,Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}.$$

We say that X and Y are positively correlated if $\rho_{X,Y} > 0$, negatively correlated if $\rho_{X,Y} < 0$, and uncorrelated if $\rho_{X,Y} = 0$.

Properties of Covariance.

(1) $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$.

(2) $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$, $\mathrm{Cov}(aX,bY) = ab\mathrm{Cov}(X,Y)$, and $\mathrm{Cov}(X+a,Y+b) = \mathrm{Cov}(X,Y)$.

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

(4) $-1 \leq \rho_{X,Y} \leq 1$ for any X and Y. Furthermore, $\rho_{XY} = \pm 1$ if and only if $Y = aX + b$ for some constants a and b.

(5) For any $X_1, \ldots X_n$ and $Y_1, \ldots, Y_m$, we have: $\mathrm{Cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j\right) = \sum_i^n \sum_j^m \mathrm{Cov}(X_i, Y_j)$.

(6) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$. More generally,

$$\mathrm{Var}\left(\sum_{i=1}^n X_i\right) = \sum_i^n \mathrm{Var}(X_i) + 2\sum\sum_{i<j} \mathrm{Cov}(X_i, X_j).$$

- Properties (1) and (2) are easy to prove using the definition of covariance. You will prove properties (3) and (4) in Homework 6. Below, we show properties (5) and (6).

Proof of (5).

$$\mathrm{Cov}\left(\sum_i X_i, \sum_j Y_j\right) = \mathbb{E}\left[\left(\sum_i X_i - \mathbb{E}\left[\sum_i X_i\right]\right)\left(\sum_j Y_j - \mathbb{E}\left[\sum_j X_j\right]\right)\right]$$

$$= \mathbb{E}\left[\left(\sum_i (X_i - \mathbb{E}[X_i])\right)\left(\sum_j (Y_j - \mathbb{E}[Y_j])\right)\right] = \mathbb{E}\left[\sum_i \sum_j (X_i - \mathbb{E}[X_i])(Y_j - \mathbb{E}[Y_j])\right]$$

$$= \sum_i \sum_j \mathbb{E}[(X_i - \mathbb{E}[X_i])(Y_j - \mathbb{E}[Y_j])] = \sum_i \sum_j \mathrm{Cov}(X_i, Y_j) \quad \square$$

Proof of (6).

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

$$= \sum_i^n \mathrm{Var}(X_i) + 2\sum\sum_{i<j} \mathrm{Cov}(X_i, X_j).$$

Example 12.3. Let X be the # of bedrooms and Y be the number of laptops in an apartment. The joint pmf of X and Y are as follows:

$X\backslash Y$	0	1	2	p_X
0	0.05	0.12	0.03	0.2
1	0.07	0.1	0.08	0.25
2	0.02	0.26	0.27	0.55
p_Y	0.14	0.48	0.38	1

Find $\operatorname{Cov}(X, Y)$.

SOLUTION. $\mathbb{E}[X] = 0 \cdot 0.2 + 1 \cdot 0.25 + 2 \cdot 0.55 = 1.35$, and $\mathbb{E}[Y] = 1.24$. Also,

$$\mathbb{E}[XY] = \sum_{x=0}^{2} \sum_{y=0}^{2} xy \cdot p(x, y) = 1 \cdot 1 \cdot 0.1 + 1 \cdot 2 \cdot 0.08 + 2 \cdot 1 \cdot 0.26 + 2 \cdot 2 \cdot 0.27 = 1.86.$$

Therefore,

$$\operatorname{Cov}(X, Y) = 1.86 - 1.35 \cdot 1.24 = 0.186.$$

This means a positive correlation (as one can expect).