Lecture Note: Week 12

MATH 461: Probability Theory, Spring 2021
Daesung Kim

Lecture 28. Covariance (Sec 7.4)

Covariance
The covariance between X and Y is defined by

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

Proposition 1. (i) $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$, $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$, *and* $\mathrm{Cov}(aX+b,Y) = a\,\mathrm{Cov}(X,Y)$.

(ii) *We have*

$$\mathrm{Cov}(\sum_{i=1}^n X_i, \sum_{j=1}^n Y_j) = \sum_{i=1}^n \mathrm{Cov}(X_i, Y_j).$$

In particular,

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

Correlation coefficient
The correlation between X and Y is defined by

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

We say X and Y are uncorrelated if $\rho(X,Y) = 0$.

Proposition 2. (i) $-1 \le \rho(X,Y) \le 1$.

(ii) *For* $a > 0$, $\rho(aX+b,Y) = \rho(X,Y)$.

(iii) *If* $\rho(X,Y) = \pm 1$, *then* $Y = aX + b$.

Example 3. Toss a fair coin 3 times. Let X be the number of heads, and Y be the number of tails. Find $\mathrm{Cov}(X,Y)$.

Example 4. Let $X_1, X_2, \cdots$ be independent random variables with common mean μ and common variance σ_2. Set $Y_n = X_n + 2X_{n+1}$ for $n \ge 1$. For $j \ge 0$, find $\mathrm{Cov}(Y_n, Y_{n+j})$ and $\rho(Y_n, Y_{n+j})$.

Lecture 29. Conditional Expectation (Sec 7.5)

Suppose X and Y are jointly discrete random variables with joint pmf $p(x, y)$. For y with $p_Y(y) = \mathbb{P}(Y = y) > 0$, the conditional expectation of X given $Y = y$ is defined by

$$\mathbb{E}[X|Y = y] = \sum_x x\mathbb{P}(X = x|Y = y) = \sum_x x p_{X|Y}(x|y).$$

Example 5. If X and Y are independent binomial random variables with identical parameters n and p, calculate the conditional expected value of X given that $X + Y = m$.

Suppose X and Y are jointly continuous random variables with joint pdf $f(x, y)$. For y with $f_Y(y) => 0$, the conditional expectation of X given $Y = y$ is defined by

$$\mathbb{E}[X|Y = y] = \int x f_{X|Y}(x|y)\, dx.$$

Example 6. Suppose that the joint density of X and Y is given by

$$f(x, y) = \begin{cases} \frac{e^{-\frac{x}{y}} e^{-y}}{y}, & 0 < x, y < \infty \\ 0, & \text{otherwise.} \end{cases}$$

Compute $\mathbb{E}[X|Y = y]$.

Remark 7. Note that the conditional expectation $\mathbb{E}[X|Y = y]$ is a function of y, that is, $g(y) = \mathbb{E}[X|Y = y]$. Using this, we define a random variable $g(Y) = \mathbb{E}[X|Y]$.

Computing expectations by conditioning
For random variables X and Y,

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[\mathbb{E}[X|Y]] \\ &= \sum_y \mathbb{E}[X|Y = y]\mathbb{P}(Y = y) \qquad \text{(discrete case)} \\ &= \int \mathbb{E}[X|Y = y] f_Y(y)\, dy \qquad \text{(jointly continuous case).} \end{aligned}$$

Example 8. A miner is trapped in a mine containing 3 doors. The first door leads to a tunnel that will take him to safety after 3 hours of travel. The second door leads to a tunnel that will return him to the mine after 5 hours of travel. The third door leads to a tunnel that will return him to the mine after 7 hours. If we assume that the miner is at all times equally likely to choose any one of the doors, what is the expected length of time until he reaches safety?

Computing probabilities by conditioning
For an event A and a random variable Y,

$$\begin{aligned} \mathbb{P}(A) = \mathbb{E}[I_A] &= \mathbb{E}[\mathbb{E}[I_A|Y]] \\ &= \sum_y \mathbb{P}(X|Y = y)\mathbb{P}(Y = y) \qquad \text{(discrete case)} \\ &= \int \mathbb{P}(X|Y = y) f_Y(y)\, dy \qquad \text{(jointly continuous case).} \end{aligned}$$

Example 9. Suppose that we are to be presented with n distinct prizes, in sequence. After being presented with a prize, we must immediately decide whether to accept it or to reject it and consider the next prize. The only information we are given when deciding whether to accept a prize is the relative rank of that prize compared to ones already seen. That is, for instance, when the fifth prize is presented, we learn how it compares with the four prizes we've already seen. Suppose that once a prize is rejected, it is lost, and that our objective is to maximize the probability of obtaining the best prize. Assuming that all $n!$ orderings of the prizes are equally likely, how well can we do?

Example 10. Suppose that the number of people entering a department store on a given day is a random variable with mean 50. Suppose further that the amounts of money spent by these customers are independent random variables having a common mean of $8. Finally, suppose also that the amount of money spent by a customer is also independent of the total number of customers who enter the store. What is the expected amount of money spent in the store on a given day?

References

[SR] Sheldon Ross, *A First Course in Probability*, 9th Edition, Pearson

Department of Mathematics, University of Illinois at Urbana-Champaign
E-mail address: daesungk@illinois.edu