STA 131A: Introduction to Probability Theory

Lecture 20: Covariance and Conditional Variance

Dogyoon Song

Spring 2026, UC Davis

Announcements

Midterm 2 solutions and scores are posted online

- You may review your graded exam during TA office hours tomorrow or in discussion section on Thursday
- Remember that only your higher score on the two midterms will count toward your final course grade

To submit regrade requests

- If you believe a score should be adjusted, please email the TA **by noon on Friday, May 22** with:
  - The specific problem(s) for which you are requesting a regrade
  - A clear explanation of why your answer merits more credit

Where we are now

So far, we have studied:

- Probability fundamentals: events, conditioning, independence
- Discrete and continuous random variables
- Joint distributions, conditional distributions, and derived distributions

For the rest of the quarter, we study tools for **sums, dependence, and large-sample behavior**:

- Covariance and correlation
- Moment generating functions and random sums
- Deviation inequalities
- Limit theorems: WLLN and CLT

Agenda

Last time:

- Derived distributions
- Transformations of random variables
- Convolution for sums of independent random variables

Today: Dependence and variance decomposition

- Covariance
- Correlation coefficient
- Conditional variance and the law of total variance

Recall: Variance

The variance of a random variable X is

$$\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}[X^2] - \left(\mathbb{E}[X]\right)^2.$$

- $\mathbb{E}[X]$ measures the center of the distribution
- $\mathrm{Var}(X)$ measures the spread around the center
- $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$

Questions for today:

- How do we measure how two random variables vary together?
- How does dependence affect the variance of a sum?
- How can conditioning decompose uncertainty?

Covariance

Definition (Covariance)

For random variables X and Y, the **covariance** between X and Y is

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

Expanding the definition gives the useful computational formula

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

Interpretation:

- $\mathrm{Cov}(X, Y) > 0$: X and Y tend to move together
- $\mathrm{Cov}(X, Y) < 0$: one tends to be high when the other is low
- $\mathrm{Cov}(X, Y) = 0$: no linear co-movement, but not necessarily independence

Example: A tunable covariance

Example

Let $p \in [0,1]$. Let X and Y be independent random variables such that

$$X = \begin{cases} 1 & \text{with probability } 1/2, \\ -1 & \text{with probability } 1/2, \end{cases} \qquad \text{and} \qquad Y = \begin{cases} 1 & \text{with probability } p, \\ -1 & \text{with probability } 1-p, \end{cases}$$

Define $Z = XY$. Observe that

$$\mathbb{E}[X] = 0, \qquad \mathbb{E}[Z] = \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] = 0, \qquad XZ = X(XY) = X^2Y = Y.$$

Thus,

$$\mathrm{Cov}(X,Z) = \mathbb{E}[XZ] - \mathbb{E}[X]\mathbb{E}[Z] = \mathbb{E}[Y] = 2p-1.$$

and therefore

$$\mathrm{Cov}(X,Z) \begin{cases} > 0, & p > 1/2, \\ = 0, & p = 1/2, \\ < 0, & p < 1/2. \end{cases}$$

When $p > 1/2$, $Z = XY$ tends to have the same sign as X; when $p < 1/2$, the opposite.

Properties of covariance

For any random variables X, Y, Z and any scalars a, b,

- $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
- $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
- $\mathrm{Cov}(X, aY + b) = a\mathrm{Cov}(X, Y)$
- $\mathrm{Cov}(X, Y + Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$

The variance of a sum of random variables satisfies

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

- When $n = 2$,
$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$$
- Unlike expectations, variances add only when the covariance terms vanish

Example: Variance of a sum

Example

Suppose

$$\mathrm{Var}(X) = 4, \qquad \mathrm{Var}(Y) = 9, \qquad \mathrm{Cov}(X, Y) = -3.$$

Question: Compute $\mathrm{Var}(X + Y)$ and $\mathrm{Var}(X - Y)$.

Using $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$ and $\mathrm{Cov}(X, -Y) = -\mathrm{Cov}(X, Y)$, we get

$$\begin{aligned} \mathrm{Var}(X + Y) &= 4 + 9 + 2(-3) = 7, \\ \mathrm{Var}(X - Y) &= \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{Cov}(X, Y) \\ &= 4 + 9 - 2(-3) = 19. \end{aligned}$$

Interpretation: Negative covariance reduces the variability of $X + Y$, but increases the variability of $X - Y$.

Independence and zero covariance

If X and Y are independent, then

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y], \qquad \text{and therefore} \qquad \mathrm{Cov}(X, Y) = 0.$$

However, the converse is false:

$$\mathrm{Cov}(X, Y) = 0 \;\not\Rightarrow\; X, Y \text{ independent.}$$

Example

Let X be uniform on $\{-1, 0, 1\}$, and let

$$Y = X^2.$$

Then Y is determined by X, so X and Y are not independent.
Nevertheless,

$$\mathbb{E}[X] = 0, \qquad \mathbb{E}[XY] = \mathbb{E}[X^3] = 0, \qquad \text{and hence,} \qquad \mathrm{Cov}(X, Y) = 0.$$

Pop-up quiz: Covariance

Suppose

$$\mathrm{Var}(X) = 4, \qquad \mathrm{Var}(Y) = 9, \qquad \mathrm{Cov}(X, Y) = -3.$$

Question: What is $\mathrm{Var}(2X - Y)$?

A) 7

B) 19

C) 25

D) 37

Answer: D.

$$\begin{aligned} \mathrm{Var}(2X - Y) &= 4\mathrm{Var}(X) + \mathrm{Var}(Y) + 2(2)(-1)\mathrm{Cov}(X, Y) \\ &= 4(4) + 9 - 4(-3) = 37. \end{aligned}$$

Follow-up: What sign of $\mathrm{Cov}(X, Y)$ reduces $\mathrm{Var}(X + Y)$ and $\mathrm{Var}(X - Y)$ respectively?

Correlation coefficient

Definition

When $\mathrm{Var}(X) > 0$ and $\mathrm{Var}(Y) > 0$, the **correlation coefficient** between X and Y is

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

Covariance depends on the units of X and Y, but the correlation coefficient normalizes.

- When $\mathrm{Var}(X) > 0$ and $\mathrm{Var}(Y) > 0$,

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

Interpretation:

- $\rho > 0$: positive linear association
- $\rho < 0$: negative linear association
- $\rho = 0$: uncorrelated, but not necessarily independent
- $|\rho| = 1$: exact linear relationship

Example: Correlation under a linear relationship

Example

Let $a, b \in \mathbb{R}$ with $a \neq 0$, and suppose $\mathrm{Var}(X) > 0$ and

$$Y = aX + b.$$

Question: Compute the correlation coefficient between X and Y.

By definition,

$$\mathrm{Cov}(X, Y) = \mathrm{Cov}(X, aX + b) = a\mathrm{Var}(X),$$
$$\mathrm{Var}(Y) = a^2\mathrm{Var}(X).$$

Therefore,

$$\rho(X, Y) = \frac{a\mathrm{Var}(X)}{\sqrt{\mathrm{Var}(X)\, a^2\mathrm{Var}(X)}} = \frac{a}{|a|}. \qquad \text{and thus,} \qquad \rho(X, Y) = \begin{cases} 1, & a > 0, \\ -1, & a < 0. \end{cases}$$

Pop-up quiz: Correlation

Suppose $\rho(X, Y) = 0.6$, and define

$$U = 3X + 1, \qquad V = -2Y + 5.$$

Question: What is $\rho(U, V)$?

A) 0.6

B) -0.6

C) -1.2

D) Cannot be determined

Answer: B.
Adding constants does not change correlation. Multiplying one variable by a negative constant flips the sign:

$$\rho(3X + 1, -2Y + 5) = -\rho(X, Y) = -0.6.$$

Follow-up: What if $V = 2Y + 5$ instead?

Motivation: Conditional uncertainty

Variance measures uncertainty:

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2].$$

But if we observe another random variable Y, our uncertainty about X may change.

- $\mathbb{E}[X]$: unconditional center of X
- $\mathbb{E}[X \mid Y]$: center of X after observing Y
- $\mathrm{Var}(X)$: unconditional uncertainty
- $\mathrm{Var}(X \mid Y)$: remaining uncertainty after observing Y

Goal: decompose total uncertainty in X into what remains after conditioning and what is explained by Y.

Conditional variance

Recall that $\mathbb{E}[X \mid Y]$ is a random variable encoding the mean of X after Y is observed.

Similarly, conditional variance measures the remaining spread of X after Y is observed.

Definition

The **conditional variance** of X given $Y = y$ is

$$\mathrm{Var}(X \mid Y = y) = \mathbb{E}\left[(X - \mathbb{E}[X \mid Y = y])^2 \mid Y = y\right].$$

Equivalently,

$$\mathrm{Var}(X \mid Y = y) = \mathbb{E}[X^2 \mid Y = y] - (\mathbb{E}[X \mid Y = y])^2.$$

The conditional variance $\mathrm{Var}(X \mid Y)$ is a random variable

- When $Y = y$, its value is $\mathrm{Var}(X \mid Y = y)$.

Example: Conditional variance in a two-group model

Example

Let $Y \sim \mathrm{Bernoulli}(1/2)$. Suppose

$$X \mid Y = 0 \sim N(0,1), \qquad X \mid Y = 1 \sim N(2,4).$$

Question: Compute conditional mean and conditional variance for $Y = 0$ and $Y = 1$.

If $Y = 0$, then

$$\mathbb{E}[X \mid Y = 0] = 0, \qquad \mathrm{Var}(X \mid Y = 0) = 1.$$

If $Y = 1$, then

$$\mathbb{E}[X \mid Y = 1] = 2, \qquad \mathrm{Var}(X \mid Y = 1) = 4.$$

Observe that

$$\mathbb{E}[X \mid Y] = 2Y \qquad \text{and} \qquad \mathrm{Var}(X \mid Y) = 1 + 3Y.$$

Interpretation: observing Y reveals which group mean and within-group variance apply.

Note: $\mathbb{E}[X \mid Y]$ and $\mathrm{Var}(X \mid Y)$ are random variables as they depend on random label Y.

Wrap-up

Covariance measures joint linear variability: $\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.

- Correlation is normalized covariance:

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

- Independence implies zero covariance, but zero covariance does not imply independence.

Variance of sums: $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$

- Variances add when the relevant covariance terms are zero, e.g., under independence.

Conditional variance of X given $Y = y$ is

$$\mathrm{Var}(X \mid Y = y) = \mathbb{E}\left[(X - \mathbb{E}[X \mid Y = y])^2 \mid Y = y\right].$$

Suggested reading: [BT08, Ch. 4.2 & 4.3]

References

Dimitri Bertsekas and John N Tsitsiklis.
Introduction to probability, volume 1.
Athena Scientific, 2nd edition, 2008.