STA 103 Lecture 5: Covariance

Instructor: Wookyeong Song

Department of Statistics, University of California, Davis

Aug 12th, 2025

Computing $E\left(f(X,Y)\right)$

- Recall our shortcut for computing $E(f(X))$:

$$E(f(X)) = \sum_x f(x)p_X(x).$$

- A similar shortcut works with a joint PMF $p(x,y)$ for two RVs X and Y:

$$E(f(X,Y)) = \sum_{(x,y)} f(x,y)p(x,y),$$

$$E(f(X)) = \sum_{(x,y)} f(x)p_X(x),$$

$$E(f(Y)) = \sum_{(x,y)} f(y)p_Y(y).$$

- Let (X,Y) have the joint PMF:

$X \backslash Y$	0	1	
0	0.06	0.04	
1	0.54	0.36	
			1

- **Question**: Find $E(X+Y)$, $E(XY)$, and $\mathrm{Var}(XY)$.

Computing $E\left(f(X,Y)\right)$

Joint PMF

$X \backslash Y$	0	1	
0	0.06	0.04	
1	0.54	0.36	
			1

- Using our shortcut:

$$\begin{aligned} E(X+Y) &= \sum_{(x,y)} (x+y)p(x,y) \\ &= (0+0)\times 0.06 + (1+0)\times 0.54 \\ &+ (0+1)\times 0.04 + (1+1)\times 0.36. \end{aligned}$$

- Similarly,

$$\begin{aligned} E(XY) &= \sum_{(x,y)} (xy)p(x,y) \\ &= (0\cdot 0)\times 0.06 + (1\cdot 0)\times 0.54 \\ &+ (0\cdot 1)\times 0.04 + (1\cdot 1)\times 0.36. \end{aligned}$$

Computing $E\left(f(X, Y)\right)$

- We know that $\mathrm{Var}(XY) = E(X^2Y^2) - (E(XY))^2$. Then,

$$\begin{aligned} E(X^2Y^2) &= \sum_{(x,y)} (x^2y^2)p(x, y) \\ &= (0^2 \cdot 0^2) \times 0.06 + (1^2 \cdot 0^2) \times 0.54 \\ &+ (0^2 \cdot 1^2) \times 0.04 + (1^2 \cdot 1^2) \times 0.36. \end{aligned}$$

  Then,

$$\mathrm{Var}(XY) = E(X^2Y^2) - (E(XY))^2.$$

Conditional Joint PMF

- Revisit the joint PMF for (X, Y):

$X \backslash Y$	0	1	$p_X(x)$
0	$1/8$	0	1/8
1	$2/8$	$1/8$	3/8
2	0	$3/8$	3/8
3	$1/8$	0	1/8
$p_Y(y)$	4/8	4/8	1

- For a given draw of (X, Y), suppose you told that $X \geq 1$ and $Y = 1$ but not the exact values of X and Y.

- **Question 2**: Your best prediction for X given $X \geq 1$ and $Y = 1$, i.e., $E(X \mid X \geq 1 \text{ and } Y = 1)$?

- **Question 3**: Your typical prediction error for X given $X \geq 1$ and $Y = 1$, i.e., $\mathrm{sd}(X \mid X \geq 1 \text{ and } Y = 1)$?

Conditional Joint PMF

The Joint PMF for $P\left(X = x, Y = y \mid X \geq 1 \text{ and } Y = 1\right)$

$X \backslash Y$	0	1	$p_X\left(x \mid X \geq 1 \text{ and } Y = 1\right)$
0	0	0	$0 + 0 =$0
1	0	1/4	$0 + 1/4 =$1/4
2	0	3/4	$0 + 3/4 =$ 3/4
3	0	0	$0 + 0 =$ 0
			1

- **Answer for Question 2**: Using the shortcut

$$
\begin{aligned}
E(X \mid X \geq 1 \text{ and } Y = 1) &= \sum_x x p_X(x \mid X \geq 1 \text{ and } Y = 1). \\
&= (0 \cdot 0) + (1 \cdot 1/4) + (2 \cdot 3/4) + (3 \cdot 0) \\
&= 7/4.
\end{aligned}
$$

Conditional Joint PMF

The Joint PMF for $P\left(X=x, Y=y \mid X \geq 1 \text{ and } Y=1\right)$

$X \backslash Y$	0	1	$p_X\left(x \mid X \geq 1 \text{ and } Y=1\right)$
0	0	0	$0+0=0$
1	0	1/4	$0+1/4=1/4$
2	0	3/4	$0+3/4=3/4$
3	0	0	$0+0=0$
			1

- **Answer for Question 3**: Using the shortcut

$$
\begin{aligned}
& \operatorname{Var}(X \mid X \geq 1 \text{ and } Y=1) \\
& = E(X^2 \mid X \geq 1 \text{ and } Y=1) - (E(X \mid X \geq 1 \text{ and } Y=1))^2 \\
& = \sum_x x^2 p_X(x \mid X \geq 1 \text{ and } Y=1) - \left(\sum_x x p_X(x \mid X \geq 1 \text{ and } Y=1)\right)^2. \\
& = (0^2 \cdot 0) + (1^2 \cdot 1/4) + (2^2 \cdot 3/4) + (3^2 \cdot 0) - (7/4)^2 \\
& = 13/4 - (7/4)^2 = 3/16.
\end{aligned}
$$

Conditional Joint PMF

The Joint PMF for $P\left(X=x, Y=y \mid X \geq 1 \text{ and } Y=1\right)$

$X \backslash Y$	0	1	$p_X\left(x \mid X \geq 1 \text{ and } Y=1\right)$
0	0	0	$0+0=0$
1	0	1/4	$0+1/4=1/4$
2	0	3/4	$0+3/4=3/4$
3	0	0	$0+0=0$
			1

- **Answer for Question 3 (Continue)**: Then we have

$$\begin{aligned}\mathrm{sd}(X \mid X \geq 1 \text{ and } Y=1) &= \sqrt{\mathrm{Var}(X \mid X \geq 1 \text{ and } Y=1)} \\ &= \sqrt{\frac{3}{16}} = \frac{\sqrt{3}}{4}.\end{aligned}$$

Covariance

- **Motivation**: If two RVs X and Y are not independent (i.e., they are dependent), there are many ways to summarize how dependent they are. The most common is with covariance.

- **Definition**: If X and Y are two RVs then the covariance between X and Y is defined as

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) \quad \text{(easy to compute)},$$
$$\mathrm{Cov}(X, Y) = E\left((X - E(X))(Y - E(Y))\right) \quad \text{(easy to understand)}.$$

- **Interpretation**: It measures how much X and Y "co-vary together."

Properties of Covariance

- We will see that Cov is very closely related to Var. Indeed notice the similarity of the definitions:

$$\begin{aligned}\mathrm{Cov}(X,Y) &= E(XY) - E(X)E(Y)\\ \mathrm{Var}(X) &= E(X\cdot X) - E(X)E(X) = E(X^2) - (E(X))^2.\end{aligned}$$

- Notice therefore

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

- If X and Y are independent, then

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

  However, $\mathrm{Cov}(X,Y) = 0$ does not imply the independence between X and Y.

Properties of Covariance

- **Formula**:

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

- If X and Y are independent, then

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y),$$
$$\mathrm{Var}(X-Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Revisit Properties of Expectation and Variance

Recall the formula:

- $E(X+Y) = E(X) + E(Y)$.
- $E(aX+b) = aE(X) + b$.
- $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$.
- $\mathrm{Var}(aX+b) = \mathrm{Var}(aX) = a^2\mathrm{Var}(X)$.
- $\mathrm{sd}(aX+b) = \sqrt{\mathrm{Var}(aX+b)} = \sqrt{a^2\mathrm{Var}(X)} = |a|\mathrm{sd}(X)$.

Revisit Properties of Expectation and Variance

Suppose we have arbitrary RVs X_1 and X_2. Then consequence of the formula,

- $E(X_1 + X_2) = E(X_1) + E(X_2)$.
- If X_1 and X_2 are independent then

$$\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2).$$

- **Example (Investment Portfolio Risk Hedging)**: We have two investments:

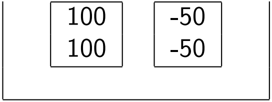
$$X_1 = \text{Profit from first investment,}$$
$$X_2 = \text{Profit from second investment.}$$

Revisit Properties of Expectation and Variance

- **Example (Investment Portfolio Risk Hedging)**:

$$\text{Total Profit } = X_1 + X_2.$$

- Case 1:



Suppose we select $\boxed{\begin{array}{c}X_1\\X_2\end{array}}$.

- What is expectation of total profit, $E(X_1 + X_2)$?

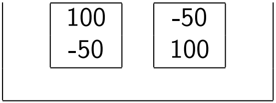
- What is volatility (variance) of total profit, $\mathrm{Var}(X_1 + X_2)$?

Revisit Properties of Expectation and Variance

- **Example (Investment Portfolio Risk Hedging)**:

$$\text{Total Profit } = X_1 + X_2.$$

- Case 2:



Suppose we select $\boxed{\begin{matrix} X_1 \\ X_2 \end{matrix}}$.

- What is expectation of total profit, $E(X_1 + X_2)$?

- What is risk (variance) of total profit, $\mathrm{Var}(X_1 + X_2)$?

Revisit Properties of Expectation and Variance

- **Example (Investment Portfolio Risk Hedging)**:

$$\text{Total Profit } = X_1 + X_2.$$

- **Results**: Expected profits $E(X_1 + X_2)$ are the same in both Case 1 and 2. But, risk (variance) of profit $\mathrm{Var}(X_1 + X_2)$ in Case 2 is lower than that in Case 1.

- How to lower the risk (variance) of total profit $X_1 + X_2$? We know that

$$\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + 2\mathrm{Cov}(X_1, X_2).$$

  The first and second term $\mathrm{Var}(X_1)$ and $\mathrm{Var}(X_2)$ are always ≥ 0. However, the third covariance term $\mathrm{Cov}(X_1, X_2)$ can be negative.

- If $\mathrm{Cov}(X_1, X_2)$ is negative, it reduces the risk of total profit $X_1 + X_2$. If $\mathrm{Cov}(X_1, X_2)$ is positive, it adds the risk of total profit $X_1 + X_2$.

Revisit Properties of Expectation and Variance

- **Example (Investment Portfolio Risk Hedging)**:

$$\text{Total Profit } = X_1 + X_2.$$

- **Interpretation**: Negative $\mathrm{Cov}(X_1, X_2)$ means that the direction of movement between X_1 and X_2 should be opposite (e.g., stock and bond). The negative dependence between X_1 and X_2 in Case 2 act like a hedge and reduce risk.

- **Interpretation**: Positive $\mathrm{Cov}(X_1, X_2)$ means that the direction of movement between X_1 and X_2 should be the same (e.g., Coin Base stock in NASDAQ and Bitcoin). The positive dependence between X_1 and X_2 in Case 2 is more aggressive portfolio with higher risk.

- Case 1 (HW): check $\mathrm{Cov}(X_1, X_2) > 0$.

- Case 2 (HW): check $\mathrm{Cov}(X_1, X_2) < 0$.