CS 70 Discrete Mathematics and Probability Theory
Fall 2025 Sabin, Hug
DIS 11B

Covariance and Total Expectation Intro

Covariance: measure of the relationship between two RVs

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

The sign of $\operatorname{cov}(X,Y)$ illustrates how X and Y are related; a positive value means that X and Y tend to increase and decrease together, while a negative value means that X increases as Y decreases (and vice versa). A covariance of zero means that the two random variables are uncorrelated—there is no linear relationship between them.

Properties: for random variables X, Y, Z and constant a,

- $\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) + 2\operatorname{cov}(X,Y)$
- $\operatorname{cov}(X,X) = \operatorname{Var}(X)$
- $\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X)$
- Bilinearity: $\operatorname{cov}(X+Y,Z) = \operatorname{cov}(X,Z) + \operatorname{cov}(Y,Z)$ and $\operatorname{cov}(aX,Y) = a\operatorname{cov}(X,Y)$

Conditional Expectation: When we want to find the expectation of a random variable X conditioned on an event A, we use the following formula:

$$\mathbb{E}[X \mid A] = \sum_x x \cdot \mathbb{P}[(X = x) \mid A].$$

This is an application of the definition of expectation. We still consider all values of X but reweigh them based on their probability of occurring together with A.

Total Expectation: For any random variable X and events $A_1, A_2, \ldots, A_n$ that partition the sample space Ω,

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X \mid A_i]\,\mathbb{P}[A_i].$$

We can think of this as splitting the sample space into partitions (events) and looking at the expectation of X in each partition, weighted by the probability of that event occurring.

Often, we use another random variable to construct the partition. If Y is a random variable, then the events $Y = y_1, Y = y_2, \ldots$ partition the sample space, where $\{y_1, y_2, \ldots\}$ are all the possible values of Y. In this case, $\mathbb{E}[X \mid Y = y]$ is a function of Y: it takes inputs $y \in Y$ and outputs $f(y) = \mathbb{E}[X \mid Y = y]$. So $f(Y) = \mathbb{E}[X \mid Y]$ is itself a random variable.

1 Covariance

Note 21

(a) We have a bag of 5 red and 5 blue balls. We take two balls uniformly at random from the bag without replacement. Let X_1 and X_2 be indicator random variables for the events of the first and second ball being red, respectively. What is $\text{cov}(X_1, X_2)$? Recall that $\text{cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$.

(b) Now, we have two bags A and B, with 5 red and 5 blue balls each. Draw a ball uniformly at random from A, record its color, and then place it in B. Then draw a ball uniformly at random from B and record its color. Let X_1 and X_2 be indicator random variables for the events of the first and second draws being red, respectively. What is $\text{cov}(X_1, X_2)$?

Solution:

(a) We can use the formula $\text{cov}(X_1, X_2) = \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\,\mathbb{E}[X_2]$.

$$\begin{aligned}
\mathbb{E}[X_1] &= \frac{5}{10} \times 1 + \frac{5}{10} \times 0 = \frac{1}{2}, \\
\mathbb{E}[X_2] &= \frac{5}{10} \times 1 + \frac{5}{10} \times 0 = \frac{1}{2}, \\
\mathbb{E}[X_1 X_2] &= \frac{5}{10} \cdot \frac{4}{9} \times 1 + \left(1 - \frac{5}{10} \cdot \frac{4}{9}\right) \times 0 = \frac{2}{9}.
\end{aligned}$$

Therefore,

$$\text{cov}(X_1, X_2) = \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\,\mathbb{E}[X_2] = \frac{2}{9} - \frac{1}{2} \times \frac{1}{2} = -\frac{1}{36}.$$

(b) Again, we use the formula $\text{cov}(X_1, X_2) = \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]$.

$$\begin{aligned}
\mathbb{E}[X_1] &= \frac{5}{10} \times 1 + \frac{5}{10} \times 0 = \frac{1}{2} \\
\mathbb{E}[X_2] &= \left(\frac{5}{10} \times \frac{6}{11} + \frac{5}{10} \times \frac{5}{11}\right) \times 1 + \left(\frac{5}{10} \times \frac{5}{11} + \frac{5}{10} \times \frac{6}{11}\right) \times 0 = \frac{1}{2} \\
\mathbb{E}[X_1 X_2] &= \frac{5}{10} \times \frac{6}{11} \times 1 = \frac{30}{110}.
\end{aligned}$$

Therefore,

$$\mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2] = \frac{30}{110} - \frac{1}{4} = \frac{1}{44}.$$

Note that in part (a), if one event happened, the other would be less likely to happen, and thus the covariance was negative. Similarly, in part (b), if one event happened, the other would be more likely to happen, and thus the covariance was positive.

2 Correlation and Independence

Note 21

(a) What does it mean for two random variables to be uncorrelated?

(b) What does it mean for two random variables to be independent?

(c) Are all uncorrelated variables independent? Are all independent variables uncorrelated? If your answer is yes, justify your answer; if your answer is no, give a counterexample.

Solution:

(a) Recall that for two random variables X and Y,

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

Two random variables are uncorrelated iff their covariance is equal to zero. If X and Y are uncorrelated, then there is no linear relationship between them.

(b) Recall that two random variables X and Y are independent if and only if the following criteria are met (the three criteria are equivalent and connected by Bayes rule):

$$\begin{aligned}\mathbb{P}[X=x \mid Y=y] &= \mathbb{P}[X=x]\\ \mathbb{P}[Y=y \mid X=x] &= \mathbb{P}[Y=y]\\ \mathbb{P}[X=x, Y=y] &= \mathbb{P}[X=x]\,\mathbb{P}[Y=y]\end{aligned}$$

for all x, y such that $\mathbb{P}[X=x]$, $\mathbb{P}[Y=y] > 0$.

If X and Y are independent, any information about one variable offers no information whatsoever about the other variable.

(c) Note that if two random variables are independent, they must have no relationship whatsoever, including linear relationships; therefore they must be uncorrelated. The converse, however, is not true: two uncorrelated variables may not be independent. Consider two variables X and Y that follow a uniform joint distribution over the points $(1,0),(0,1),(-1,0),(0,-1)$. See Figure 1. Then

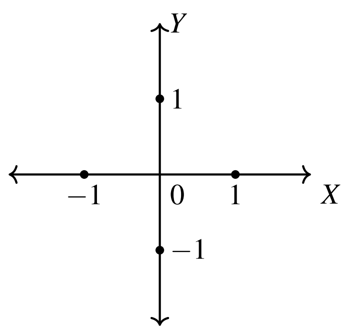

Figure 1: Choose one of the four points shown uniformly at random.

$$\operatorname{cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] = 0.$$

To see why, observe that $XY = 0$ always because at least one of X and Y is always 0, and furthermore $\mathbb{E}[X] = \mathbb{E}[Y] = 0$ because both X and Y are symmetric around 0. So, there is no linear relationship, but X and Y are not independent (for example, $\mathbb{P}[Y=0] = 1/2$ but $\mathbb{P}[Y=0 \mid X=1] = 1$).

3 Dice Games

Note 21 Suppose you roll a fair six-sided die. You read off the number showing on the die, then flip that many fair coins.

(a) If the result of your die roll is i, what is the expected number of heads you see?

(b) What is the expected number of heads you see?

Solution:

(a) The number of heads you get is binomially distributed with parameters i and $\frac{1}{2}$. Thus, the expected number of heads you see is $\frac{i}{2}$.

(b) Let D be the outcome of the die roll and H be the number of heads you get. We have that

$$\begin{aligned}\mathbb{E}[H] &= \sum_{i=1}^{6} \mathbb{E}[H|D=i] \cdot \mathbb{P}[D=i] \\ &= \sum_{i=1}^{6} \frac{i}{2} \cdot \frac{1}{6} \\ &= \frac{1}{12} \sum_{i=1}^{6} i\end{aligned}$$

We know that $\sum_{i=1}^{n} i$ comes out to $\frac{n(n+1)}{2}$, so $\mathbb{E}[H] = \frac{1}{12} \cdot \frac{6 \cdot 7}{2} = \frac{7}{4}$.

4 Number Game

Note 21 Sinho and Vrettos are playing a game where they each choose an integer uniformly at random from $[0, 100]$, then whoever has the larger number wins (in the event of a tie, they replay). However, Vrettos doesn't like losing, so he's rigged his random number generator such that it instead picks randomly from the integers between Sinho's number and 100. Let S be Sinho's number and V be Vrettos' number.

(a) What is $\mathbb{E}[S]$?

(b) What is $\mathbb{E}[V \mid S = s]$, where s is any constant such that $0 \le s \le 100$?

(c) What is $\mathbb{E}[V]$?

Solution:

(a) S is a (discrete) uniform random variable between 0 and 100, so its expectation is $\frac{0+100}{2} = 50$.

(b) If $S = s$, we know that V will be uniformly distributed between s and 100. Similar to the previous part, this gives us that $\mathbb{E}[V \mid S = s] = \frac{s+100}{2}$.

(c) With the law of total expectation, we have that

$$\begin{aligned}
\mathbb{E}[V] &= \sum_{s=0}^{100} \mathbb{E}[V \mid S = s] \cdot \mathbb{P}[S = s] \\
&= \sum_{s=0}^{100} \frac{s+100}{2} \cdot \frac{1}{101} \\
&= \frac{1}{202} \left(\sum_{s=0}^{100} s + \sum_{s=0}^{100} 100 \right)
\end{aligned}$$

The first summation comes out to $\frac{100(100+1)}{2} = 50 \cdot 101$; the second summation is just adding 100 to itself 101 times, so it comes out to $100 \cdot 101$. Plugging these values in, we get $\mathbb{E}[V] = 75$.

Alternate Solution:
Using the previous part and the Law of Total Expectation, we get

$$\begin{aligned}
\mathbb{E}[V] = \mathbb{E}[\mathbb{E}[V \mid S]] = \mathbb{E}\left[\frac{S+100}{2}\right] \\
&= \frac{\mathbb{E}[S]+100}{2} \\
&= \frac{150}{2} = 75.
\end{aligned}$$