DSC 40A: Theoretical Foundations of Data Science

Lecture 23: Independence and Conditional Independence
(Sawyer's version)

November 19, 2025

Announcements

- Gal is out sick today, so Sawyer will be lecturing for her.
- Note: Gal's Wednesday office hours are cancelled. There will be an extra OH during Week 10 to make up for the cancellation.
- **Homework 6** is due this **Friday, November 21st**.
- **Homework 7** will be released this Friday and is due **Wednesday, December 3rd**. No slip days!
- Next Wednesday's lecture will take place as usual.

Bayes' Theorem

On Monday you learned about **Bayes' theorem**, which tells us how to update probabilities in light of new information.

Theorem Bayes' Theorem

Let S be a sample space and $A, B \subseteq S$ any two events such that $\mathbb{P}(A), \mathbb{P}(B) > 0$. Then

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\,\mathbb{P}(A)}{\mathbb{P}(B)}.$$

Furthermore, we can write

$$\frac{\mathbb{P}(B \mid A)\,\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \mid A)\,\mathbb{P}(A)}{\mathbb{P}(B \mid A)\,\mathbb{P}(A) + \mathbb{P}\left(B \mid \overline{A}\right)\mathbb{P}\left(\overline{A}\right)}.$$

Bayes' Theorem

Sometimes, the knowledge of an event A occurring makes another event B more or less likely to occur. Other times, the occurrence of A has no effect on the likelihood of B occurring.

Example Coin Flips

Suppose we flip two fair coins. Let

$$A = \{\text{first coin is heads}\},$$
$$B = \{\text{second coin is heads}\}.$$

The sample space consists of four equally likely outcomes:

$$S = \{(H,H),(H,T),(T,H),(T,T)\}.$$

Then by the definition of conditional probability,

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(\{(H,H)\})}{\mathbb{P}(\{(H,H),(H,T)\})} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} = \mathbb{P}(B).$$

Thus, knowing that the first coin is heads does not change the probability that the second coin is heads; in this case, A and B are *independent*.

Independence

In many experiments, we encounter events that can occur without influencing each other in any meaningful way.

We say that two events A and B are **independent** if

$$\mathbb{P}(A \mid B) = \mathbb{P}(A), \text{ and } \mathbb{P}(B \mid A) = \mathbb{P}(B).$$

Intuitively, learning that B occurred does not change the probability of A. Otherwise, the events are **dependent**.

Independence

Note that by Bayes' theorem, if $\mathbb{P}(A \mid B) = \mathbb{P}(A)$, then it must also hold that

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B)\,\mathbb{P}(B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A)\,\mathbb{P}(B)}{\mathbb{P}(A)} = \mathbb{P}(B).$$

Moreover, in this case,

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(A \mid B)\,\mathbb{P}(B) = \mathbb{P}(A)\,\mathbb{P}(B).$$

It can also be shown that the converse holds: if $\mathbb{P}(A \text{ and } B) = \mathbb{P}(A)\,\mathbb{P}(B)$, then $\mathbb{P}(A \mid B) = \mathbb{P}(A)$ and $\mathbb{P}(B \mid A) = \mathbb{P}(B)$. Thus, A and B are independent if and only if

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(A)\,\mathbb{P}(B).$$

Independence

In other words, two events A, B are said to be independent if one and hence *both* of the following conditions holds:

$$\mathbb{P}(A \mid B) = \mathbb{P}(A) \text{ and } \mathbb{P}(B \mid A) = \mathbb{P}(B),$$

or equivalently,

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(A)\,\mathbb{P}(B).$$

The second definition is a bit more friendly since it doesn't assume that $\mathbb{P}(A)$ and $\mathbb{P}(B)$ are nonzero.

A collection of multiple events $A_1, A_2, \ldots, A_n$ is **mutually independent** if every finite intersection factorizes into the product of its probabilities.

Concept check

Suppose A and B are two events with nonzero probability. Is it possible for A and B to be *both* mutually exclusive *and* independent?

(a) Yes, A and B *CAN* be both mutually exclusive *and* independent.

(b) No, A and B *CANNOT* be both mutually exclusive *and* independent.

Answer: (b). If A and B are mutually exclusive, then $\mathbb{P}(A \text{ and } B) = 0$. But if they are independent, then $\mathbb{P}(A \text{ and } B) = \mathbb{P}(A)\,\mathbb{P}(B) > 0$ since both probabilities are nonzero. This is a contradiction.

Concept check

Suppose $\mathbb{P}(A) = 0$ and B is any event. Which statement is true?

(a) A and B can never be independent.

(b) A and B are always independent.

(c) A and B are independent only if $\mathbb{P}(B) = 0$.

(d) We cannot say anything about independence in this case.

Answer: (b). Since $\mathbb{P}(A) = 0$, we have $\mathbb{P}(A \text{ and } B) = 0 = \mathbb{P}(A)\,\mathbb{P}(B)$ for *every* B, so A is (degenerately) independent of all events.

Independence

Example Coin Flips Revisited

Let's revisit the coin flip example. Once again the sample space is given by

$$S = \{(H, H), (H, T), (T, H), (T, T)\},$$

with all events equally likely. Let

$$A = \{\text{first coin is heads}\} = \{(H, H), (H, T)\},$$
$$B = \{\text{second coin is heads}\} = \{(H, H), (T, H)\}.$$

Then

$$\mathbb{P}(A) = \tfrac{1}{2}, \quad \mathbb{P}(B) = \tfrac{1}{2}, \quad \mathbb{P}(A \text{ and } B) = \mathbb{P}(\{(H, H)\}) = \tfrac{1}{4},$$

so $\mathbb{P}(A \text{ and } B) = \mathbb{P}(A)\,\mathbb{P}(B)$ and A, B are independent.

Example Coin Flips Revisited

Now let

$$C = \{\text{at least one head}\} = \{(H, H), (H, T), (T, H)\}.$$

Then

$$\mathbb{P}(C) = \tfrac{3}{4},$$
$$\mathbb{P}(A \text{ and } C) = \mathbb{P}(A) = \tfrac{1}{2},$$
$$\mathbb{P}(A)\,\mathbb{P}(C) = \tfrac{1}{2} \times \tfrac{3}{4} = \tfrac{3}{8} \neq \tfrac{1}{2}.$$

Thus A and C are not independent.

Example Yahtzee!

Suppose we roll two fair six-sided dice. The sample space is given by

$$S = \{(1,1),(1,2),\ \dots\ (6,5),(6,6)\},$$
$$\mathbb{P}(\{\,(i,j)\,\}) = \tfrac{1}{36} \text{ for each } (i,j) \in S.$$

Define the events

$$A = \{\text{first die is even}\},$$
$$B = \{\text{second die is even}\},$$
$$C = \{\text{sum of the two dice is even}\}.$$

Since there are three even faces on each die, one checks

$$\mathbb{P}(A) = \mathbb{P}(B) = \tfrac{18}{36} = \tfrac{1}{2},$$
$$\mathbb{P}(C) = \tfrac{18}{36} = \tfrac{1}{2}.$$

Example Yahtzee!

Moreover,

$$\begin{aligned}\mathbb{P}(A \text{ and } B) &= \mathbb{P}(\{\text{both dice even}\}) \\ &= \tfrac{3\times 3}{36} = \tfrac{9}{36} \\ &= \tfrac{1}{4} = \mathbb{P}(A)\,\mathbb{P}(B)\,,\end{aligned}$$

and similarly

$$\begin{aligned}\mathbb{P}(A \text{ and } C) &= \tfrac{1}{4} = \mathbb{P}(A)\,\mathbb{P}(C) \\ \mathbb{P}(B \text{ and } C) &= \tfrac{1}{4} = \mathbb{P}(B)\,\mathbb{P}(C)\end{aligned}$$

However,

$$\begin{aligned}A \text{ and } B \text{ and } C &= \{\text{both dice even}\}, \\ \mathbb{P}(A \text{ and } B \text{ and } C) &= \tfrac{9}{36} = \tfrac{1}{4} \neq \tfrac{1}{2} \times \tfrac{1}{2} \times \tfrac{1}{2} = \tfrac{1}{8}.\end{aligned}$$

Thus A, B, C are **pairwise independent but not mutually independent**.

Independence in practice

Sometimes we assume that events are independent to make calculations easier.

Real-world events are almost never exactly independent, but the independence assumption can be a useful approximation in many cases.

Example: Event A corresponds to a randomly selected student having blue eyes, and event B corresponds to a randomly selected student being left-handed. These events are not truly independent, but they are close enough that we can often treat them as independent for practical purposes.

Putting independence to the test

This quick poll will allow us to test the claim on the last slide.

Conditional Independence

We often need to consider events that are independent only after conditioning on some background information. Or, independent events can become dependent once we condition on some other event.

Events A and B are **conditionally independent** given C when

$$\mathbb{P}(A \text{ and } B \mid C) = \mathbb{P}(A \mid C)\,\mathbb{P}(B \mid C),$$

provided $\mathbb{P}(C) > 0$.

This notion will be useful when we study Naïve Bayes classifiers, which will be covered in the next portion of the course.

Poker

Example Two of a kind

Consider a standard deck of 52 playing cards. We draw three cards at random without replacement. Consider the following events:

$$A = \{\text{hand contains the Ace of spade}\},$$
$$B = \{\text{hand contains the Ace of heart}\},$$
$$C = \{\text{hand contains exactly one spade and one heart}\}.$$

Are A and B independent? We can check:

$$\mathbb{P}(A) = \frac{\mathrm{C}(51,2)}{\mathrm{C}(52,3)} = \frac{51 \times 50 \times 6}{2 \times 52 \times 51 \times 50} = \frac{3}{52},$$
$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(B)} = \frac{\frac{\mathrm{C}(50,1)}{\mathrm{C}(52,3)}}{\frac{\mathrm{C}(51,2)}{\mathrm{C}(52,3)}} = \frac{\mathrm{C}(50,1)}{\mathrm{C}(51,2)} = \frac{50 \times 2}{51 \times 50} = \frac{2}{51}.$$

Thus, A and B are not independent since $\mathbb{P}(A) = \frac{50}{221} \neq \frac{3}{51} = \mathbb{P}(A \mid B)$. This makes sense because knowing that the hand contains the Ace of hearts makes it less likely that it also contains the Ace of spades.

Example Two of a kind

However, conditioning on C, the events A and B become independent. First observe that

$$\mathbb{P}(C) = \frac{\binom{13}{1}\binom{13}{1}\binom{26}{1}}{\binom{52}{3}} = \frac{13 \times 13 \times 26 \times 6}{52 \times 51 \times 50} \approx 0.20$$

Now given C, the spade in the hand is equally likely to be any of the 13 spades, and the heart is equally likely to be any of the 13 hearts. Therefore

$$\mathbb{P}(A \mid C) = \frac{1}{13}, \qquad \mathbb{P}(B \mid C) = \frac{1}{13}.$$

For both aces to appear under C, we must choose the Ace of spades as the spade and the Ace of hearts as the heart, so

$$\mathbb{P}(A \text{ and } B \mid C) = \frac{1}{13} \cdot \frac{1}{13} = \frac{1}{169} = \mathbb{P}(A \mid C)\,\mathbb{P}(B \mid C).$$

Thus A and B are conditionally independent given C.

Concept check

We flip two fair coins. Let

$$A = \{\text{first coin is heads}\}, \qquad B = \{\text{second coin is heads}\}.$$

We know A and B are independent.
Now condition on the event $C = \{\text{exactly one head in total}\}$.
Are A and B independent *given* C?

Answer: No. Under C, the only outcomes are (H, T) and (T, H), each with probability $1/2$. Then

$$\mathbb{P}(A \mid C) = \mathbb{P}(B \mid C) = \tfrac{1}{2}, \quad \mathbb{P}(A \text{ and } B \mid C) = 0 \neq \tfrac{1}{2} \cdot \tfrac{1}{2}.$$

So A and B become dependent once we condition on C.

Chess Tournament

Sometimes rather than proving conditional independence, we are told to assume it and use this fact to solve problems.

Example

A chess player is entering a weekend tournament and is about to play two games on her first day.

She is either having a good day (denoted "G"), or a bad day (denoted "B"); and she is more likely to win if she's having a good day.

The sample space of outcomes corresponding to her two games is given by

$$S = \{(G, W, W), (B, W, W), (G, L, W), (B, L, W), \\ (G, W, L), (B, W, L), (G, L, L), (B, L, L)\},$$

where "W" denotes a win and "L" denotes a loss.

Chess Tournament

Example

For example, (B, W, L) denotes the outcome where she has a bad day, wins the first game, and loses the second.

On a typical day, the probability she has a good or a bad day is given by

$$P(\{\text{Good day}\}) = \mathbb{P}(\{(G, W, W), (G, L, W), \dots\}) = 0.6,$$
$$P(\{\text{Bad day}\}) = \mathbb{P}(\{(B, W, W), (B, L, W), \dots\}) = 0.4.$$

Let $A, B \subseteq S$ be the events given by

$$A = \{\text{she wins in round 1}\}, \qquad B = \{\text{she wins in round 2}\}.$$

Suppose that *conditional on* the player's form, the outcomes of the two rounds are independent, with

$$P(A \mid \{\text{Good day}\}) = P(B \mid \{\text{Good day}\}) = 0.7,$$
$$P(A \mid \{\text{Bad day}\}) = P(B \mid \{\text{Bad day}\}) = 0.4.$$

Example

Now suppose we wish to compute the following probabilities:

$$\mathbb{P}(A), \quad \mathbb{P}(B), \quad \mathbb{P}(A \text{ and } B).$$

and determine whether A and B are independent unconditionally. By the *conditional* independence assumption, we have

$$\mathbb{P}(A \text{ and } B \mid \{\text{Good day}\}) = 0.7 \times 0.7 = 0.49,$$
$$\mathbb{P}(A \text{ and } B \mid \{\text{Bad day}\}) = 0.4 \times 0.4 = 0.16.$$

Hence by the law of total probability,

$$\mathbb{P}(A) = 0.6 \times 0.7 + 0.4 \times 0.4 = 0.42 + 0.16 = 0.58,$$
$$\mathbb{P}(B) = 0.6 \times 0.7 + 0.4 \times 0.4 = 0.58,$$
$$\mathbb{P}(A \text{ and } B) = 0.6 \times 0.49 + 0.4 \times 0.16 = 0.294 + 0.064 = 0.358.$$

Example

Since

$$\mathbb{P}(A)\,\mathbb{P}(B) = 0.58^2 = 0.3364 \neq 0.358 = \mathbb{P}(A \text{ and } B),$$

the events A and B are not independent (even though they are conditionally independent given the form of the day).

Concept check

In the chess example, we found

$$\mathbb{P}(A) = \mathbb{P}(B) = 0.58, \qquad \mathbb{P}(A \text{ and } B) = 0.358.$$

Is winning round 1 associated with a *higher* or *lower* probability of winning round 2 compared to the marginal chance of winning round 2?

(a) Higher: winning round 1 makes winning round 2 more likely.

(b) Lower: winning round 1 makes winning round 2 less likely.

(c) Neither: there is no association.

Answer: (a). Since

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(A)} = \frac{0.358}{0.58} \approx 0.62 > 0.58 = \mathbb{P}(B),$$

a win in round 1 makes a win in round 2 more likely, even though the two rounds were conditionally independent given the day type.