Lecture 22 – Independence and Classification

DSC 40A, Fall 2022 @ UC San Diego

Dr. Truong Son Hy, with help from **many others**

Agenda

- Independence.
- Conditional independence.
- Classification

Review: Partition

- A set of events $E_1, E_2, \ldots, E_k$ is a **partition** of S if each outcome in S is in exactly one E_i.
- The law of total probability states that if A is an event and $E_1, E_2, \ldots, E_k$ is a **partition** of S, then

$$
\begin{aligned}
P(A) &= P(E_1) \cdot P(A|E_1) + P(E_2) \cdot P(A|E_2) + \ldots + P(E_k) \cdot P(A|E_k) \\
&= \sum_{i=1}^{k} P(E_i) \cdot P(A|E_i)
\end{aligned}
$$

Review: Bayes' theorem

- Bayes' theorem states that

$$P(B|A) = \frac{P(B) \cdot P(A|B)}{P(A)}$$

- We often re-write the denominator $P(A)$ in Bayes' theorem using the law of total probability:

$$P(B|A) = \frac{P(B) \cdot P(A|B)}{\sum_{i=1}^{k} P(E_i) \cdot P(A|E_i)}$$

Independence

Updating probabilities

- Bayes' theorem describes how to update the probability of one event, given that another event has occurred.

$$P(B|A) = \frac{P(B) \cdot P(A|B)}{P(A)}$$

  - $P(B)$ can be thought of as the "**prior**" probability of B occurring, before knowing anything about A.
  - $P(B|A)$ is sometimes called the "**posterior**" probability of B occurring, given that A occurred.
- What if knowing that A occurred doesn't change the probability that B occurs? In other words, what if

$$P(B|A) = P(B)$$

Independent events

- A and B are **independent events** if one event occurring does not affect the chance of the other event occurring.

$$P(B|A) = P(B) \qquad P(A|B) = P(A)$$

- Otherwise, A and B are **dependent events**.

- Using Bayes' theorem, we can show that if one of the above statements is true, then so is the other.

Proof

- Suppose $P(B|A) = P(B)$, given Bayes' theorem, we have:

$$P(B|A) = \frac{P(B) \cdot P(A|B)}{P(A)} = P(B),$$

that leads to:

$$\frac{P(A|B)}{P(A)} = 1 \Leftrightarrow P(A|B) = P(A).$$

Proof

- Suppose $P(B|A) = P(B)$, given Bayes' theorem, we have:

$$P(B|A) = \frac{P(B) \cdot P(A|B)}{P(A)} = P(B),$$

  that leads to:

$$\frac{P(A|B)}{P(A)} = 1 \Leftrightarrow P(A|B) = P(A).$$

- Suppose $P(A|B) = P(A)$, given Bayes' theorem, we have:

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B)} = P(A),$$

  that leads to:

$$\frac{P(B|A)}{P(B)} = 1 \Leftrightarrow P(B|A) = P(B).$$

Independent events

- **Equivalent definition:** A and B are independent events if

$$P(A \cap B) = P(A) \cdot P(B)$$

- To check if A and B are independent, use whichever is easiest:
  - $P(B|A) = P(B)$.
  - $P(A|B) = P(A)$.
  - $P(A \cap B) = P(A) \cdot P(B)$.

Mutual exclusivity and independence

Discussion Question

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

A) Yes
B) No
C) It depends on A and B

Mutual exclusivity and independence

Discussion Question

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

A) Yes
B) No
C) It depends on A and B

Answer: B) No. Why?

Mutual exclusivity and independence

When two events (call them A and B) are **mutually exclusive**, it is impossible for them to happen together:

$$P(A \cap B) = 0$$

When two events are **independent**:

$$P(A \cap B) = P(A) \cdot P(B)$$

Thus, if they are both **mutually exclusive** and **independent** then at least one of them must have zero probability.

Example: Venn diagrams

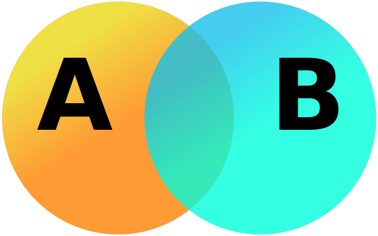
For three events A, B, and C, we know that

- A and C are independent,
- B and C are independent,
- A and B are mutually exclusive,
- $P(A \cup C) = \frac{2}{3}$, $P(B \cup C) = \frac{3}{4}$, $P(A \cup B \cup C) = \frac{11}{12}$.

Find $P(A)$, $P(B)$, and $P(C)$.

Example: Venn diagrams

Venn diagram:



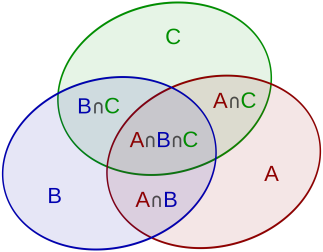
Sets A and B:

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Events A and B:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Example: Venn diagrams



Sets A, B, and C:

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

Events A, B and C:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

Example: Venn diagrams

A and C are independent:

$$P(A \cap C) = P(A) \cdot P(C)$$

B and C are independent:

$$P(B \cap C) = P(B) \cdot P(C)$$

A and B are mutually exclusive:

$$P(A \cap B) = 0 \Rightarrow P(A \cap B \cap C) = 0$$

$P(A \cup C) = 2/3$:

$$P(A) + P(C) - P(A \cap C) = P(A) + P(C) - P(A) \cdot P(C) = \frac{2}{3}$$

Example: Venn diagrams

$P(B \cup C) = 3/4$:

$$P(B) + P(C) - P(B \cap C) = P(B) + P(C) - P(B) \cdot P(C) = \frac{3}{4}$$

$P(A \cup B \cup C) = 11/12$:

$$P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) = \frac{11}{12}$$

Because $P(A \cap B) = P(A \cap B \cap C) = 0$, we have:

$$P(A) + P(B) + P(C) - P(A \cap C) - P(B \cap C) = \frac{11}{12}$$

We re-arrange the terms a bit:

$$\Big(P(A) + P(C) - P(A \cap C)\Big) + \Big(P(B) + P(C) - P(B \cap C)\Big) - P(C) = \frac{11}{12}$$

Example: Venn diagrams

We get:

$$P(A \cup C) + P(B \cup C) - P(C) = \frac{2}{3} + \frac{3}{4} - P(C) = \frac{11}{12}$$

Thus:

$$P(C) = \frac{1}{2}$$

Furthermore:

$$P(A) + P(C) - P(A) \cdot P(C) = \frac{1}{2}P(A) + \frac{1}{2} = \frac{2}{3} \Rightarrow P(A) = \frac{1}{3}$$

$$P(B) + P(C) - P(B) \cdot P(C) = \frac{1}{2}P(B) + \frac{1}{2} = \frac{3}{4} \Rightarrow P(B) = \frac{1}{2}$$

Example: cards

♥: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♦: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♣: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♠: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A

- Suppose you draw two cards, one at a time.
  - A is the event that the first card is a heart.
  - B is the event that the second card is a club.

- If you draw the cards **with** replacement, are A and B independent?

Example: cards

♥: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♦: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♣: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♠: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A

- Suppose you draw two cards, one at a time.
  - A is the event that the first card is a heart.
  - B is the event that the second card is a club.
- If you draw the cards **with** replacement, are A and B independent? Yes. Because $P(B|A) = P(B) = \frac{1}{4}$.
- If you draw the cards **without** replacement, are A and B independent?

Example: cards

♥: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♦: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♣: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♠: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A

- Suppose you draw two cards, one at a time.
  - A is the event that the first card is a heart.
  - B is the event that the second card is a club.

- If you draw the cards **with** replacement, are A and B independent? Yes. Because $P(B|A) = P(B) = \frac{1}{4}$.

- If you draw the cards **without** replacement, are A and B independent? No. Because $P(B|A) = \frac{13}{51} \neq P(B) = \frac{1}{4}$.

Example: cards

♥: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♦: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♣: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♠: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A

- Suppose you draw one card from a deck of 52.
  - A is the event that the card is a heart.
  - B is the event that the card is a face card (J, Q, K).
- Are A and B independent?

Example: cards

♥: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♦: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♣: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♠: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A

- Suppose you draw one card from a deck of 52.
  - A is the event that the card is a heart.
  - B is the event that the card is a face card (J, Q, K).

- Are A and B independent? Yes. Because:

$$P(A) = \frac{13}{52} = \frac{1}{4}$$

$$P(B) = \frac{12}{52} = \frac{3}{13}$$

$$P(A \cap B) = \frac{3}{52} = \frac{1}{4} \cdot \frac{3}{13} = P(A) \cdot P(B)$$

Assuming independence

- Sometimes we assume that events are independent to make calculations easier.
- Real-world events are almost never exactly independent, but they may be close.

Example: breakfast

1% of UCSD students are DSC majors. 25% of UCSD students eat avocado toast for breakfast. Assuming that being a DSC major and eating avocado toast for breakfast are independent:

1. What percentage of DSC majors eat avocado toast for breakfast?

Example: breakfast

1% of UCSD students are DSC majors. 25% of UCSD students eat avocado toast for breakfast. Assuming that being a DSC major and eating avocado toast for breakfast are independent:

1. What percentage of DSC majors eat avocado toast for breakfast? 25%

2. What percentage of UCSD students are DSC majors who eat avocado toast for breakfast?

Example: breakfast

1% of UCSD students are DSC majors. 25% of UCSD students eat avocado toast for breakfast. Assuming that being a DSC major and eating avocado toast for breakfast are independent:

1. What percentage of DSC majors eat avocado toast for breakfast? 25%

2. What percentage of UCSD students are DSC majors who eat avocado toast for breakfast? 0.25%

Conditional independence

Conditional independence

- Sometimes, events that are dependent *become* independent, upon learning some new information.

- Or, events that are independent can become dependent, given additional information.

Example: cards

♥: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♦: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♣: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, A
♠: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A

- Your dog ate the King of Clubs. Suppose you draw one card from a deck of 51.
  - A is the event that the card is a heart.
  - B is the event that the card is a face card (J, Q, K).

- Are A and B independent?

Example: cards

♥: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♦: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♣: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, A
♠: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A

- Your dog ate the King of Clubs. Suppose you draw one card from a deck of 51.
  - A is the event that the card is a heart.
  - B is the event that the card is a face card (J, Q, K).

- Are A and B independent? No. Because:

$$P(A) = \frac{13}{51}$$

$$P(B) = \frac{11}{51}$$

$$P(A \cap B) = \frac{3}{51} \neq P(A) \cdot P(B) = \frac{13 \cdot 11}{51^2}$$

Example: cards

♥: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♦: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♣: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, A
♠: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A

- Your dog ate the King of Clubs. Suppose you draw one card from a deck of 51.
  - A is the event that the card is a heart.
  - B is the event that the card is a face card (J, Q, K).
- Suppose you learn that the card is red. Are A and B independent given this new information?

Example: cards

♥: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♦: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
♣: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, A
♠: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A

- Your dog ate the King of Clubs. Suppose you draw one card from a deck of 51.
  - A is the event that the card is a heart.
  - B is the event that the card is a face card (J, Q, K).
- Suppose you learn that the card is red. Are A and B independent given this new information? Yes. Why?

Example: cards

We have:

$$P(A|\text{red}) = \frac{13}{26} = \frac{1}{2}$$

$$P(B|\text{red}) = \frac{6}{26} = \frac{3}{13}$$

$$P(A \cap B|\text{red}) = \frac{3}{26} = \frac{1}{2} \cdot \frac{3}{13} = P(A|\text{red}) \cdot P(B|\text{red})$$

Therefore, A and B are **independent conditioned that the card is red**.

Conditional independence

- Recall that A and B are independent if

$$P(A \cap B) = P(A) \cdot P(B)$$

- A and B are **conditionally independent** given C if

$$P((A \cap B)|C) = P(A|C) \cdot P(B|C)$$

- Given that C occurs, this says that A and B are independent of one another.

Assuming conditional independence

- Sometimes we assume that events are conditionally independent to make calculations easier.
- Real-world events are almost never exactly conditionally independent, but they may be close.

Example: Harry Potter and TikTok

Suppose that 50% of UCSD students like Harry Potter and 80% of UCSD students use TikTok. What is the probability that a random UCSD student likes Harry Potter and uses TikTok, assuming that these events are conditionally independent given that a person is a UCSD student?

Example: Harry Potter and TikTok

Suppose that 50% of UCSD students like Harry Potter and 80% of UCSD students use TikTok. What is the probability that a random UCSD student likes Harry Potter and uses TikTok, assuming that these events are conditionally independent given that a person is a UCSD student?

40%

Independence vs. conditional independence

- Is it reasonable to assume conditional independence of
  - liking Harry Potter
  - using TikTok

  given that a person is a UCSD student?
- Is it reasonable to assume independence of these events in general, among all people?

Discussion Question

Which assumptions do you think are reasonable?

A) Both
B) Conditional independence only
C) Independence (in general) only
D) Neither

Independence vs. conditional independence

- Is it reasonable to assume conditional independence of
  - liking Harry Potter
  - using TikTok

  given that a person is a UCSD student?
- Is it reasonable to assume independence of these events in general, among all people?

Discussion Question

Which assumptions do you think are reasonable?

A) Both
B) Conditional independence only
C) Independence (in general) only
D) Neither

Answer: B) Conditional independence only.

Independence vs. conditional independence

In general, there is **no relationship** between independence and conditional independence. All of these are possibilities, given three events A, B, and C.

- A and B are independent, and are conditionally independent given C.
- A and B are independent, and are conditionally dependent given C.
- A and B are dependent, and are conditionally independent given C.
- A and B are dependent, and are conditionally dependent given C.

Example: constructing events

- Consider a sample space $S = \{1, 2, 3, 4, 5, 6\}$ where all outcomes are equally likely.
- For each scenario, specify events A, B, and C that satisfy the given conditions. (e.g. $A = \{2, 5, 6\}$)
- Choose events that are neither impossible nor certain, i.e. $0 < P(A), P(B), P(C) < 1$.

Scenario 1: A and B **are not** independent. A and B **are** conditionally independent given C.

Scenario 1: A and B **are not** independent. A and B **are** conditionally independent given C.

$$P(A \cap B) \neq P(A) \cdot P(B)$$

$$P(A \cap B | C) = P(A | C) \cdot P(B | C)$$

Let's aim to get $P(A|C) = 1/2$, $P(B|C) = 1/2$ and $P(A \cap B|C) = 1/4$. For example:

$$A|C = \{1, 2\}, \quad B|C = \{2, 3\}, \quad A \cap B|C = \{2\}$$

and C is condition so that the sample is less than or equal to 4. We can set $A = \{1, 2, 5\}$ and $B = \{2, 3, 6\}$. Obviously:

$$P(A \cap B) = P(2) = \frac{1}{6}$$

that is not equal to

$$P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

Example: constructing events

- Consider a sample space $S = \{1, 2, 3, 4, 5, 6\}$ where all outcomes are equally likely.
- For each scenario, specify events A, B, and C that satisfy the given conditions. (e.g. $A = \{2, 5, 6\}$)
- Choose events that are neither impossible nor certain, i.e. $0 < P(A), P(B), P(C) < 1$.

Scenario 2: A and B **are not** independent. A and B **are not** conditionally independent given C.

Scenario 2: A and B **are not** independent. A and B **are not** conditionally independent given C.

We can set $A = \{1, 2, 5\}$ and $B = \{2, 3, 6\}$. For the condition, let change C to the sample is less than or equal to 3. We get:

$$P(A|C) = \frac{2}{3},$$

$$P(B|C) = \frac{2}{3},$$

and:

$$P(A \cap B|C) = \frac{1}{3} \neq P(A|C) \cdot P(B|C) = \frac{4}{9}.$$

Example: constructing events

- Consider a sample space $S = \{1, 2, 3, 4, 5, 6\}$ where all outcomes are equally likely.
- For each scenario, specify events A, B, and C that satisfy the given conditions. (e.g. $A = \{2, 5, 6\}$)
- Choose events that are neither impossible nor certain, i.e. $0 < P(A), P(B), P(C) < 1$.

Scenario 3: A and B **are** independent. A and B **are** conditionally independent given C.

Scenario 3: A and B **are** independent. A and B **are** conditionally independent given C.

Let's aim to construct so that:

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{1}{3}, \quad P(A \cap B) = \frac{1}{6},$$

$$P(A|C) = \frac{1}{2}, \quad P(B|C) = \frac{1}{2}, \quad P(A \cap B|C) = \frac{1}{4}.$$

For example:

$$A = \{1, 2, 5\},$$

$$B = \{2, 4\},$$

and C is the condition so that the sample is less than or equal to 4.

Example: constructing events

- Consider a sample space $S = \{1, 2, 3, 4, 5, 6\}$ where all outcomes are equally likely.
- For each scenario, specify events A, B, and C that satisfy the given conditions. (e.g. $A = \{2, 5, 6\}$)
- Choose events that are neither impossible nor certain, i.e. $0 < P(A), P(B), P(C) < 1$.

Scenario 4: A and B **are** independent. A and B **are not** conditionally independent given C.

Scenario 4: A and B **are** independent. A and B **are not** conditionally independent given C.

We can keep $A = \{1, 2, 5\}$ and C = the sample is less than or equal to 4 as in Scenario 3. But we change $B = \{2, 6\}$. We have:

$$P(A \cap B) = \frac{1}{6} = P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{3},$$

but

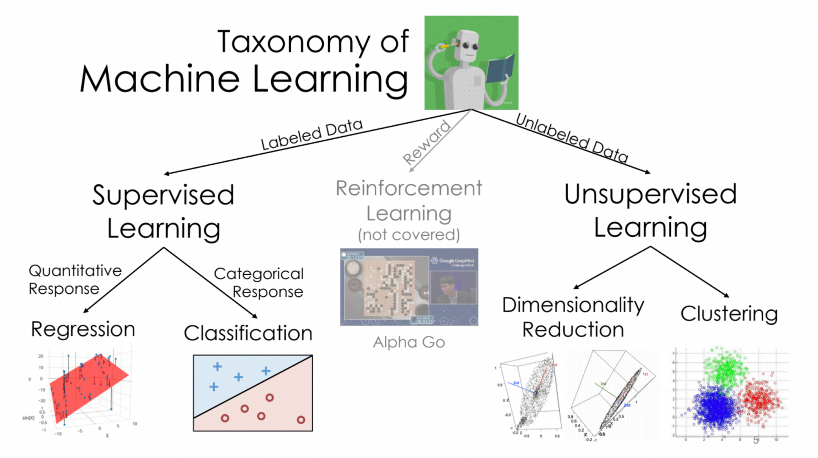
$$P(A \cap B|C) = \frac{1}{4} \neq P(A|C) \cdot P(B|C) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}.$$

Review

- Two events A and B are independent when knowledge of one event does not change the probability of the other event.
  - Equivalent conditions: $P(B|A) = P(B)$, $P(A|B) = P(A)$, $P(A \cap B) = P(A) \cdot P(B)$.
- Two events A and B are conditionally independent if they are independent given knowledge of a third event, C.
  - Condition: $P((A \cap B)|C) = P(A|C) \cdot P(B|C)$.
- In general, there is no relationship between independence and conditional independence.
- **Next:** Using Bayes' theorem and conditional independence to solve the **classification problem** in machine learning.

Classification

Taxonomy of machine learning



Classification problems

- Like with regression, we're interested in mkaing predictions based on data we've already collected (called **training data**).
- The difference is that the response variable is **categorical**.
- Categories are called **classes**.
- Example classification problems:
  - Deciding whether a patient has kidney disease.
  - Identifying handwritten digits.
  - Determining whether an avocado is ripe.
  - Predicting whether credit card activity is fraudulent.

Example: avocados

You have a green-black avocado, and want to know if it is ripe.

color	ripeness
bright green	unripe
green-black	ripe
purple-black	ripe
green-black	unripe
purple-black	ripe
bright green	unripe
green-black	ripe
purple-black	ripe
green-black	ripe
green-black	unripe
purple-black	ripe

Question: Based on this data, would you predict that your avocado is ripe or unripe?

Example: avocados

You have a green-black avocado, and want to know if it is ripe. Based on this data, would you predict that your avocado is ripe or unripe?

color	**ripeness**
bright green	unripe
green-black	ripe
purple-black	ripe
green-black	unripe
purple-black	ripe
bright green	unripe
green-black	ripe
purple-black	ripe
green-black	ripe
green-black	unripe
purple-black	ripe

Strategy: Calculate two probabilities:

$$P(\text{ripe}|\text{green-black})$$

$$P(\text{unripe}|\text{green-black})$$

Then, predict the class with a **larger** probability.