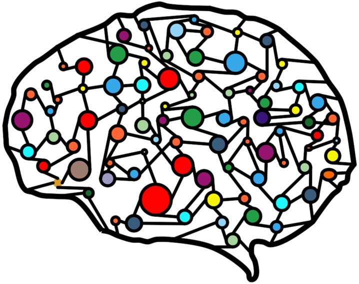
Module 22 – Independence and Conditional Independence



DSC 40A, Summer 2023

Agenda

- Independence.
- Conditional independence.

Independence

Independent events

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

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

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?

- 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 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?

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?

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

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).

- Suppose you learn that the card is red. Are A and B independent given this new information?

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 Discord

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

Independence vs. conditional independence

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

  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

In general, there is **no relationship** between independence and conditional independence. All four scenarios below are possible.

- **Scenario 1:** A and B **are** independent. A and B **are** conditionally independent given C.
- **Scenario 2:** A and B **are** independent. A and B **are not** conditionally independent given C.
- **Scenario 3:** A and B **are not** independent. A and B **are** conditionally independent given C.
- **Scenario 4:** A and B **are not** independent. A and B **are not** conditionally independent 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** independent. A and B **are** conditionally independent 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 2: A and B **are** independent. A and B **are not** conditionally independent 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 3: A and B **are not** independent. A and B **are** conditionally independent 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 4: A and B **are not** independent. A and B **are not** conditionally independent given C.

Summary

Summary

- 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 time:** Using Bayes' theorem and conditional independence to solve the **classification problem** in machine learning.