♥: 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

DSC 40A

Lecture 15

Independence

Last Time

- $P(A \mid B)$ = probability of A given that we know B has occurred.

- **Bayes' Theorem**:

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

The Bayesian View

- Bayesian view: probabilities quantify level of **belief**.
  - $P(A) = 1$; absolutely certain it will happen
  - $P(A) = 0$; absolute certain it will not
  - $P(A) = .75$; about 75% sure
- Bayes' Theorem allows us to "update" our beliefs when given new information.

Example

- In San Diego: 3,752 burglaries per year.
- Roughly 10 burglaries per night.
- 1.5 million people in San Diego.
- On any given night:

$$P(\text{Burglary}) = \frac{10}{1.5 \text{ million}} \approx 6 \times 10^{-7}$$

Example

- You hear your burglar alarm going off.
- How worried should you be?
- Assume:
  - If there is a burglary, there is a 95% of alarm.

$$P(\text{Alarm} \mid \text{Burglary}) = .95$$

  - If there isn't a burglary, there is a 1% chance of alarm.

$$P(\text{Alarm} \mid \text{No Burglary}) = .01$$

Example

$P(\text{Burglary}) = 6 \times 10^{-7}$ $P(\text{Alarm} \mid \text{Burglary}) = 0.95$ $P(\text{Alarm} \mid \text{No Burglary}) = 0.01$

$$P(\text{Burglary} \mid \text{Alarm}) = \frac{P(\text{Alarm} \mid \text{Burglary})\, P(\text{Burglary})}{P(\text{Alarm})}$$

$$= \frac{(.95)(6 \times 10^{-7})}{(.95)(6 \times 10^{-7}) + (.01)(1 - 6 \times 10^{-7})}$$

by Law of Total Prob

$$P(\text{Alarm}) = \underbrace{P(\text{Alarm} \mid \text{Burglary})}_{.95}\, \underbrace{P(\text{Burglary})}_{6 \times 10^{-7}} + \underbrace{P(\text{Alarm} \mid \text{No Burglary})}_{.01}\, P(\text{No Burglary})$$

$$= 1 - P(\text{Burglary})$$

$$= 1 - 6 \times 10^{-7}$$

Prior and Posterior Probabilities

- **Before** hearing the alarm, the probability of a burglary is

$$P(\text{Burglary}) = 6 \times 10^{-7}$$

- We call this the **prior** probability.
- **After** hearing the alarm, the probability increases:

$$P(\text{Burglary} \mid \text{Alarm}) = 5.6 \times 10^{-5}$$

- We call this the **posterior** probability.

Discussion Question

Now suppose $P(\text{Alarm} \mid \text{No Burglary}) = 10^{-5}$ instead of 0.01. What happens to the **posterior probability**, $P(\text{Burglary} \mid \text{Alarm})$?

A) It goes up.
B) It goes down.
C) Nothing; it stays the same.

Example

- Suppose $P(\text{Alarm} \mid \text{No Burglary}) = 10^{-5}$.
- Then $P(\text{Burglary} \mid \text{Alarm}) = 0.054 \approx 5\%$

“Updating” Probabilities

- $P(A)$ is our prior belief that A happens.

- $P(A \mid B)$ is our updated belief that A happens, now that we know B happens.

- Sometimes knowing that B happens doesn’t change anything.

Example

- We flip a fair coin twice.
- $P(\text{Second Flip = Heads}) = 1/2$
- $P(\text{Second Flip = Heads} \mid \text{First Flip = Heads}) = 1/2$

Independence

- We say that A and B are **independent** if knowing that B happens doesn't change our belief that A happens (and *vice versa*).

- Formally, A and B are independent if[1]:

$$P(A \mid B) = P(A)$$

- Equivalently: $P(B \mid A) = P(B)$.

- Equivalently, A and B are independent if:

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

[1] Assuming $P(B) > 0$

Example #1

Discussion Question

You throw two dice. A is the event that the first is a 6, B is the event that the sum is 10. Are these independent?

A) Yes.
B) No.

Example #2

Discussion Question

You draw two cards, one-at-a-time, **with** replacement. A is the event that the first card is a heart, B is the event that the second card is a club. Are these independent?

A) Yes.
B) No.

Example #3

Discussion Question

You draw two cards, one-at-a-time, **without** replacement. A is the event that the first card is a heart, B is the event that the second card is a club. Are these independent?

A) Yes.
B) No.

Example #4

Discussion Question

You draw one card from a deck of 52 cards. A is the event that the card is a heart, B is the event that the card is a face card (J,Q,K,A). Are these independent?

$A \cap B$

♥: 2, 3, 4, 5, 6, 7, 8, 9, 10, $\boxed{\text{J, Q, K, A}}$
♦: 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

Since $P(A \cap B) = P(A) \cdot P(B)$ A & B are indep. by definition

A) Yes.
B) No.

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

$P(B) = \frac{16}{52} = \frac{4}{13}$

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

Assuming Independence

- Sometimes we **assume** that events are independent to make calculation easier.

- Real-world events are almost never exactly independent, but may be "close".

- Example: A is event that a student is a data science major, B is event that they bike to campus.

Example

In general $P(\text{Bike}|\text{DSC}) = \frac{P(\text{Bike} \cap \text{DSC})}{P(\text{DSC})}$

2% of UCSD students are data science majors. 20% of UCSD students bike to campus. Assuming that biking to campus and being a DSC major are independent:

- What percentage of data science majors bike to campus?

$P(\text{Bike}|\text{DSC})$. Since independent $P(\text{Bike}|\text{DSC}) = P(\text{Bike})$
$= 20\%$

- What percentage of students are data science majors who bike to campus?

$P(\text{DSC} \cap \text{Bike})$ Since indep: $P(\text{DSC} \cap \text{Bike}) =$
$P(\text{DSC}) \cdot P(\text{Bike})$
$= (.02)(.20) = (.004)$

Conditional Independence

- Sometimes events A and B might not be independent.

- But they **become** independent upon learning some new information.

Example

Discussion Question

We've lost the King of Clubs! You draw one card from this deck of 51 cards. A is the event that the card is a heart, B is the event that the card is a face card (J,Q,K,A). Are these independent?

♥: 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

$P(A) = \frac{13}{51}$ $P(B) = \frac{15}{51}$ $P(A \cap B) = \frac{4}{51}$

$P(A)P(B) = \frac{13 \cdot 15}{51^2} \neq \frac{4}{51}$

A) Yes.
B) No.

Example

We've lost the King of Clubs! You draw one card from this deck of 51 cards. A is the event that the card is a heart, B is the event that the card is a face card (J,Q,K,A).

Now suppose you know that the card is red. Are A and B independent **given** this information?

♥: 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~~

Conditional Independence

- Let A, B, C be events. A and B are **conditionally independent** given C if

$$P(A \cap B \mid C) = P(A \mid C) \cdot P(B \mid C).$$

Example

Discussion Question

A box contains two coins: one is fair, and the other is not – both sides are Heads. A coin is selected at random and flipped ten times. Let A be the event that the first nine flips are Heads, and let B be the event that the tenth flip is Heads. Are A and B **independent**?

A) Yes.
B) No.

Example

Discussion Question

A box contains two coins: one is fair, and the other is not – both sides are Heads. A coin is selected at random and flipped ten times. Let A be the event that the first nine flips are Heads, and let B be the event that the tenth flip is Heads. Let C be the event that the coin is fair. Are A and B **conditionally independent** given C?

A) Yes.
B) No.

Relationship Between Independence and Conditional Independence

- There is none.

Assuming Conditional Independence

- Sometimes we **assume** that events are conditionally independent to make calculation easier.

- Real-world events are almost never exactly conditionally independent, but may be “close”.

- Example: A is the event that a person knows Python. B is the event that someone knows Bayes Theorem. C is the event that they are a data science major.

Example

Suppose 80% of data science majors know Python, and 70% know Bayes' Theorem. What is the probability that a randomly-selected major knows both, assuming that the events are conditionally independent given that they are a data science major?

$$P(\text{Python} \cap \text{Bayes} \mid \text{DSC}) = P(\text{Python} \mid \text{DSC})\, P(\text{Bayes} \mid \text{DSC})$$

$$= (.8)(.7)$$

$$= .56$$