Lecture 21 – The law of total probability and Bayes' Theorem

DSC 40A, Fall 2022 @ UC San Diego

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

Agenda

- Partitions and the law of total probability.
- Bayes' theorem.

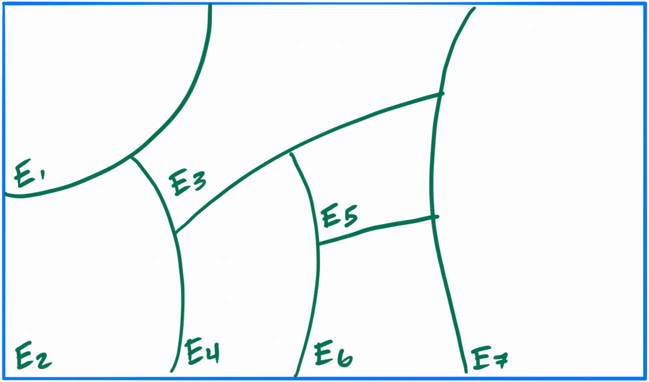
Partitions

- A set of events $E_1, E_2, ..., E_k$ is a **partition** of S if
  - $P(E_i \cap E_j) = 0$ for all unequal i, j.
  - $P(E_1 \cup E_2 \cup ... \cup E_k) = S$.
    - Equivalently, $P(E_1) + P(E_2) + ... + P(E_k) = 1$.
- In English, $E_1, E_2, ..., E_k$ is a partition of S if every outcome s in S is in **exactly** one event E_i.

Example partitions

- In getting to school, the events Walk, Bike, and Drive.
- In getting to school, the events Late and On-Time.
- In selecting an undergraduate student at random, the events Freshman, Sophomore, Junior, and Senior.
- In rolling a die, the events Even and Odd.
- In drawing a card from a standard deck of cards, the events Spades, Clubs, Hearts, and Diamonds.
- **Special case**: if A is an event and S is a sample space, A and $\bar{A}$ partition S.

Partitions, visualized

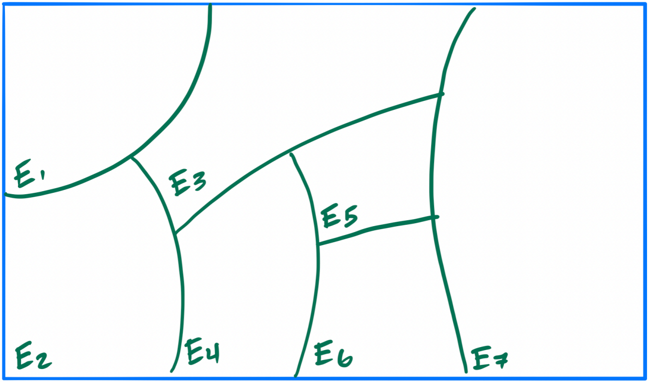


The law of total probability

- If A is an event and $E_1, E_2, \ldots, E_k$ is a **partition** of S, then

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

The law of total probability, visualized



The law of total probability

- If A is an event and $E_1, E_2, ..., E_k$ is a **partition** of S, then

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

- Since $P(A \cap E_i) = P(E_i) \cdot P(A|E_i)$ by the multiplication rule, an equivalent formulation is

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

	Late	Not Late
Walk	0.06	0.24
Bike	0.03	0.07
Drive	0.36	0.24

Discussion Question

Suppose someone is late to school. What is the probability that they walked? Choose the best answer.

A) Close to 0.05
B) Close to 0.15
C) Close to 0.3
D) Close to 0.4

	Late	Not Late
Walk	0.06	0.24
Bike	0.03	0.07
Drive	0.36	0.24

Discussion Question

Suppose someone is late to school. What is the probability that they walked? Choose the best answer.

A) Close to 0.05
B) Close to 0.15
C) Close to 0.3
D) Close to 0.4

Answer: B

$$P(\text{Walk}|\text{Late}) = \frac{P(\text{Walk, Late})}{P(\text{Late})}$$

That is equivalent to:

$$P(\text{Walk}|\text{Late}) = \frac{P(\text{Walk, Late})}{P(\text{Walk, Late}) + P(\text{Bike, Late}) + P(\text{Drive, Late})}$$

The result is:

$$P(\text{Walk}|\text{Late}) = \frac{0.06}{0.06 + 0.03 + 0.36} = \frac{0.06}{0.45} \approx 13\%$$

Bayes' theorem

Example: getting to school

- Now suppose you don't have that entire table. Instead, all you know is
  - $P(\text{Late}) = 0.45$.
  - $P(\text{Walk}) = 0.3$.
  - $P(\text{Late}|\text{Walk}) = 0.2$.
- Can you still find $P(\text{Walk}|\text{Late})$?

Example: getting to school

- Now suppose you don't have that entire table. Instead, all you know is
  - $P(\text{Late}) = 0.45$.
  - $P(\text{Walk}) = 0.3$.
  - $P(\text{Late}|\text{Walk}) = 0.2$.

- Can you still find $P(\text{Walk}|\text{Late})$?

Yes, because we know:

$$P(\text{Walk}, \text{Late}) = P(\text{Walk}|\text{Late}) \cdot P(\text{Late}) = P(\text{Late}|\text{Walk}) \cdot P(\text{Walk})$$

Example: getting to school

- Now suppose you don't have that entire table. Instead, all you know is
  - $P(\text{Late}) = 0.45$.
  - $P(\text{Walk}) = 0.3$.
  - $P(\text{Late}|\text{Walk}) = 0.2$.

- Can you still find $P(\text{Walk}|\text{Late})$?

Yes, because we know:

$$P(\text{Walk}, \text{Late}) = P(\text{Walk}|\text{Late}) \cdot P(\text{Late}) = P(\text{Late}|\text{Walk}) \cdot P(\text{Walk})$$

That leads to:

Example: getting to school

- Now suppose you don't have that entire table. Instead, all you know is
  - $P(\text{Late}) = 0.45$.
  - $P(\text{Walk}) = 0.3$.
  - $P(\text{Late}|\text{Walk}) = 0.2$.

- Can you still find $P(\text{Walk}|\text{Late})$?

Yes, because we know:

$$P(\text{Walk}, \text{Late}) = P(\text{Walk}|\text{Late}) \cdot P(\text{Late}) = P(\text{Late}|\text{Walk}) \cdot P(\text{Walk})$$

That leads to:

$$P(\text{Walk}|\text{Late}) = \frac{P(\text{Late}|\text{Walk}) \cdot P(\text{Walk})}{P(\text{Late})} = \frac{0.2 \cdot 0.3}{0.45} = \frac{2}{15}$$

Bayes' theorem

- Recall that the multiplication rule states that

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

- It also states that

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

- But since $A \cap B$ and $B \cap A$ are both "A and B", we have that

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

- Re-arranging yields **Bayes' theorem**:

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

Bayes' theorem and the law of total probability

- Bayes' theorem:

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

- Recall from earlier, for any sample space S, B and $\bar{B}$ partition S. Using the law of total probability, we can re-write $P(A)$ as

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) = P(B) \cdot P(A|B) + P(\bar{B}) \cdot P(A|\bar{B})$$

- This means that we can re-write Bayes' theorem as

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

Example: drug testing

A manufacturer claims that its drug test will **detect steroid use 95% of the time**. What the company does not tell you is that 15% of all steroid-free individuals also test positive (the false positive rate). 10% of the Tour de France bike racers use steroids. Your favorite cyclist just tested positive. What's the probability that they used steroids?

Summary:

$P(\text{positive}|\text{use}) = 0.95$

$P(\text{positive}|\text{nouse}) = 0.15$

$P(\text{use}) = 0.1$

$P(\text{nouse}) = 1 - P(\text{use}) = 0.9$

$P(\text{use}|\text{positive}) = ?$

Solution:

$$P(\text{use}|\text{positive}) = \frac{P(\text{use}, \text{positive})}{P(\text{positive})}$$

We have:

$P(\text{use}, \text{positive}) = P(\text{positive}|\text{use}) \cdot P(\text{use}) = 0.95 \cdot 0.1 = 0.095$

Summary:

$P(\text{positive}|\text{use}) = 0.95$

$P(\text{positive}|\text{nouse}) = 0.15$

$P(\text{use}) = 0.1$

$P(\text{nouse}) = 1 - P(\text{use}) = 0.9$

$P(\text{use}|\text{positive}) = ?$

Solution:

$$P(\text{use}|\text{positive}) = \frac{P(\text{use}, \text{positive})}{P(\text{positive})}$$

We have:

$P(\text{use}, \text{positive}) = P(\text{positive}|\text{use}) \cdot P(\text{use}) = 0.95 \cdot 0.1 = 0.095$

$P(\text{nouse}, \text{positive}) = P(\text{positive}|\text{nouse}) \cdot P(\text{nouse}) = 0.15 \cdot 0.9 = 0.135$

Summary:

$P(\text{positive}|\text{use}) = 0.95$
$P(\text{positive}|\text{nouse}) = 0.15$
$P(\text{use}) = 0.1$
$P(\text{nouse}) = 1 - P(\text{use}) = 0.9$
$P(\text{use}|\text{positive}) = ?$

Solution:

$$P(\text{use}|\text{positive}) = \frac{P(\text{use}, \text{positive})}{P(\text{positive})}$$

We have:
$P(\text{use}, \text{positive}) = P(\text{positive}|\text{use}) \cdot P(\text{use}) = 0.95 \cdot 0.1 = 0.095$
$P(\text{nouse}, \text{positive}) = P(\text{positive}|\text{nouse}) \cdot P(\text{nouse}) = 0.15 \cdot 0.9 = 0.135$
$P(\text{positive}) = P(\text{use}, \text{positive}) + P(\text{nouse}, \text{positive}) = 0.095 + 0.135 = 0.23$
$P(\text{use}|\text{positive}) = 0.095/0.23 \approx 41.3\%$

Example: blind burger taste test

- Your friend claims to be able to correctly guess a burger's restaurant after just one bite.
- The probability that she correctly identifies an In-n-Out Burger is 0.55, a Shake Shack burger is 0.75, and a Five Guys burger is 0.6.
- You buy 5 In-n-Out burgers, 4 Shake Shack burgers, and 1 Five Guys burger, choose one of the burgers randomly, and give it to her.
- **Question:** Given that she guessed it correctly, what's the probability she ate a Shake Shack burger?

Notations:
C = correct guess
I = In-n-Out
S = Shake Shack
F = Five Guys

Summary:
$P(C|I) = 0.55$
$P(C|S) = 0.75$
$P(C|F) = 0.6$
$P(I) = 0.5$
$P(S) = 0.4$
$P(F) = 0.1$
$P(S|C) =$?

We have:

$$P(C) = P(C, I) + P(C, S) + P(C, F)$$

That equals to:

$$P(C) = P(C|I)P(I) + P(C|S)P(S) + P(C|F)P(F) =$$

$$= 0.55 \times 0.5 + 0.75 \times 0.4 + 0.6 \times 0.1 = 0.635,$$

in which $P(C, S) = 0.75 \times 0.4 = 0.3$. Finally, we get:

$$P(S|C) = \frac{P(C, S)}{P(C)} = \frac{0.3}{0.635} \approx 47.24\%$$

Discussion Question

Consider any two events A and B. Which of the following is equal to

$$P(B|A) + P(\bar{B}|A)$$

A) $P(A)$
B) $1 - P(B)$
C) $P(B)$
D) $P(\bar{B})$
E) 1

Discussion Question

Consider any two events A and B. Which of the following is equal to

$$P(B|A) + P(\bar{B}|A)$$

A) $P(A)$
B) $1 - P(B)$
C) $P(B)$
D) $P(\bar{B})$
E) 1

Answer: E) 1

We have:

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

That equals to:

$$\frac{P(B, A) + P(\overline{B}, A)}{P(A)} = \frac{P(A)}{P(A)} = 1.$$

Summary

Summary

- 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

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

- 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.