# Math 141

## Lecture 2: More Probability!

Albyn Jones[1]

[1] Library 304
jones@reed.edu
www.people.reed.edu/∼jones/courses/141

# Outline

- Law of total probability
- Bayes' Theorem

# the Multiplication Rule, again

Recall that for any two events $A$ and $B$:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B) \cdot \mathbb{P}(B)$$

Suppose there are several events $B_1, B_2, \ldots B_k$, the multiplication rule applies to each one:

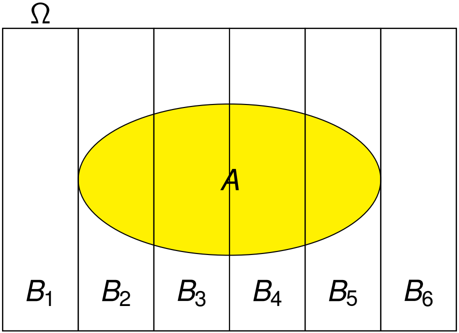
$$\mathbb{P}(A \cap B_1) = \mathbb{P}(A \mid B_1) \cdot \mathbb{P}(B_1)$$

$$\mathbb{P}(A \cap B_2) = \mathbb{P}(A \mid B_2) \cdot \mathbb{P}(B_2)$$

etc.

# A Partition

Definition: A partition of the sample space $\Omega$ is a collection of disjoint events $B_1, B_2, \ldots B_k$ whose union is $\Omega$. Such a partition divides any set $A$ into disjoint pieces:



# Application: The law of Total Probability

Suppose we have a partition $B_1, B_2, \ldots B_k$, then any other event $A$ is a union of its pieces:

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_k)$$

Those pieces are disjoint, so

$$\mathbb{P}(A) = \mathbb{P}(A \cap B_1) + \mathbb{P}(A \cap B_2) + \ldots + \mathbb{P}(A \cap B_k)$$

Applying the multiplcation rule gives:

$$\mathbb{P}(A) = \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) + \ldots + \mathbb{P}(A|B_k)\mathbb{P}(B_k)$$

# Example

Suppose we draw two cards from a well shuffled deck. One partition of the sample space is:

$$A_1 = \{\text{The first card is an Ace}\}$$

$$A_1^c = \{\text{The first card is not an Ace}\}$$

It is often useful to look for partitions of the sample space that simplify a computation!

# Example: The law of Total Probability

Consider a well shuffled card deck. What is the probability the second card in the deck is an ace?

Condition on whether the first card is an ace or not:

$$\mathbb{P}(A_2) = \mathbb{P}(A_2|A_1)\mathbb{P}(A_1) + \mathbb{P}(A_2|A_1^c)\mathbb{P}(A_1^c)$$

We can compute all of these terms easily:

$$\mathbb{P}(A_2) = \frac{3}{51}\frac{4}{52} + \frac{4}{51}\frac{48}{52}$$

$$= \frac{3 \cdot 4 + 48 \cdot 4}{51 \cdot 52} = \frac{(3 + 48) \cdot 4}{51 \cdot 52} = \frac{51 \cdot 4}{51 \cdot 52} = \frac{4}{52}$$

# Another Example: Mixtures

Suppose we have two hats: one has 4 red balls and 6 green balls, the other has 6 red and 4 green. We toss a fair coin, if heads, pick a random ball from the first hat, if tails from the second. What is the probability of getting a red ball?

$$\mathbb{P}(Red) = \mathbb{P}(Red|H) \cdot \mathbb{P}(H) + \mathbb{P}(Red|T) \cdot \mathbb{P}(T)$$

$$= \frac{4}{10}\frac{1}{2} + \frac{6}{10}\frac{1}{2} = \frac{1}{2}$$

The total probability of drawing a red ball is a weighted average of the two conditional probabilities, where the weights are the probabilities of each condition occuring.

# Go Timbers!

A soccer team wins 60% of its games when it scores the first goal, and 10% of its games when the opposing team scores first. If the team scores the first goal about 30% of the time, what fraction of the games does it win? Let $W$ be the event that the team wins, and $SF$ be the event that it scores first. Then the LTP says:

$$\mathbb{P}(W) = \mathbb{P}(W|SF) \cdot \mathbb{P}(SF) + \mathbb{P}(W|SF^c) \cdot \mathbb{P}(SF^c)$$

or

$$\mathbb{P}(W) = \frac{6}{10}\frac{3}{10} + \frac{1}{10}\frac{7}{10} = \frac{6 \times 3 + 1 \times 7}{100} = \frac{25}{100} = \frac{1}{4}$$

Bayes' Theorem is really just the definition of conditional probability dressed up with the Law of Total Probability.

Suppose we have a partition $B_1, B_2, \ldots B_k$, for which we know the probabilities $\mathbb{P}(A|B_i)$, and we wish to compute $\mathbb{P}(B_j|A)$.

$$\mathbb{P}(B_j|A) = \frac{\mathbb{P}(B_j \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_j) \cdot \mathbb{P}(B_j)}{\mathbb{P}(A)}$$

Now use the LTP to compute the denominator:

$$\mathbb{P}(B_j|A) = \frac{\mathbb{P}(A|B_j)\mathbb{P}(B_j)}{\mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) + \ldots + \mathbb{P}(A|B_k)\mathbb{P}(B_k)}$$

# Example: Screening for a rare disease

Suppose we have a blood test for a disease which has the following characteristics, where $P =$ 'tests positive', and $D =$ 'has the disease':

# Example: Screening for a rare disease

Suppose we have a blood test for a disease which has the following characteristics, where $P =$ 'tests positive', and $D =$ 'has the disease':

Sensitivity:

$$\mathbb{P}(P|D) = .99$$

# Example: Screening for a rare disease

Suppose we have a blood test for a disease which has the following characteristics, where $P =$ 'tests positive', and $D =$ 'has the disease':

Sensitivity:

$$\mathbb{P}(P|D) = .99$$

Specificity:

$$\mathbb{P}(P^c|D^c) = .99$$

Thus $\mathbb{P}(P|D^c) = .01$

# Example: Screening for a rare disease

Suppose we have a blood test for a disease which has the following characteristics, where $P =$ 'tests positive', and $D =$ 'has the disease':

Sensitivity:

$$\mathbb{P}(P|D) = .99$$

Specificity:

$$\mathbb{P}(P^c|D^c) = .99$$

Thus $\mathbb{P}(P|D^c) = .01$

Prevalence:

$$\mathbb{P}(D) = .01$$

# Example: Screening for a rare disease

Suppose we have a blood test for a disease which has the following characteristics, where $P =$ 'tests positive', and $D =$ 'has the disease':

Sensitivity:

$$\mathbb{P}(P|D) = .99$$

Specificity:

$$\mathbb{P}(P^c|D^c) = .99$$

Thus $\mathbb{P}(P|D^c) = .01$

Prevalence:

$$\mathbb{P}(D) = .01$$

Question: What is the probability that someone who tests positive actually has the disease?

# Use Bayes!

Clue: we want to get from $\mathbb{P}(P|D)$ to $\mathbb{P}(D|P)$.

$$\mathbb{P}(D|P) = \frac{\mathbb{P}(P|D)\mathbb{P}(D)}{\mathbb{P}(P|D)\mathbb{P}(D) + \mathbb{P}(P|D^c)\mathbb{P}(D^c)}$$

Substituting, we have

$$\mathbb{P}(D|P) = \frac{.99 \times .01}{.99 \times .01 + .01 \times .99} = \frac{1}{2}$$

Roughly half of those who test positive do not have the disease!

# What Happened?

- Imagine we have 1000 people: about 990 don’t have the disease, and 10 do.

# What Happened?

- Imagine we have 1000 people: about 990 don't have the disease, and 10 do.
- Of the 10 who have the disease, with high probability all 10 test positive.

# What Happened?

- Imagine we have 1000 people: about 990 don’t have the disease, and 10 do.
- Of the 10 who have the disease, with high probability all 10 test positive.
- Of the 990 who don’t,1%, or about 10, test positive.

# Example: Screening for a COMMON disease

Using the same notation as before:

# Example: Screening for a COMMON disease

Using the same notation as before:

Sensitivity:

$$\mathbb{P}(P|D) = .99$$

# Example: Screening for a COMMON disease

Using the same notation as before:

Sensitivity:

$$\mathbb{P}(P|D) = .99$$

Specificity:

$$\mathbb{P}(P|D^c) = .01$$

# Example: Screening for a COMMON disease

Using the same notation as before:

Sensitivity:

$$\mathbb{P}(P|D) = .99$$

Specificity:

$$\mathbb{P}(P|D^c) = .01$$

Prevalence:

$$\mathbb{P}(D) = .10$$

# Example: Screening for a COMMON disease

Using the same notation as before:

Sensitivity:

$$\mathbb{P}(P|D) = .99$$

Specificity:

$$\mathbb{P}(P|D^c) = .01$$

Prevalence:

$$\mathbb{P}(D) = .10$$

Question: What is the probability that someone who tests positive actually has the disease?

# Use Bayes again!

Same Clue: we want to get from $\mathbb{P}(P|D)$ to $\mathbb{P}(D|P)$.

$$\mathbb{P}(D|P) = \frac{\mathbb{P}(P|D)\mathbb{P}(D)}{\mathbb{P}(P|D)\mathbb{P}(D) + \mathbb{P}(P|D^c)\mathbb{P}(D^c)}$$

Substituting, we have

$$\mathbb{P}(D|P) = \frac{.99 \times .10}{.99 \times .10 + .01 \times .90} = \frac{.099}{.099 + .009} \approx .917$$

The difference is due to the baseline rate $\mathbb{P}(D)$.

# What Happened This Time?

- Imagine we have 1000 people: about 900 don't have the disease, and 100 do.

# What Happened This Time?

- Imagine we have 1000 people: about 900 don't have the disease, and 100 do.
- Of the 100 who have the disease, almost all, or about 99 test positive.

# What Happened This Time?

- Imagine we have 1000 people: about 900 don't have the disease, and 100 do.
- Of the 100 who have the disease, almost all, or about 99 test positive.
- Of the 900 who don't,1%, or about 9, test positive.

# What Happened This Time?

- Imagine we have 1000 people: about 900 don't have the disease, and 100 do.
- Of the 100 who have the disease, almost all, or about 99 test positive.
- Of the 900 who don't,1%, or about 9, test positive.
- Hence the advice from Public Health officials: normally we don't bother screening the general population for rare diseases, but we might consider screening 'high risk' groups.

# A probability Puzzle: Let's Make a Deal!

You're given the choice of three doors: Behind one door is a Valuable Prize; behind the others, Something Useless.

# A probability Puzzle: Let's Make a Deal!

You're given the choice of three doors: Behind one door is a Valuable Prize; behind the others, Something Useless.

- You pick a door, say No. 1.

# A probability Puzzle: Let's Make a Deal!

You're given the choice of three doors: Behind one door is a Valuable Prize; behind the others, Something Useless.

- You pick a door, say No. 1.
- The host, who knows what's behind the doors, opens another door, say No. 3, revealing Something Useless.

# A probability Puzzle: Let's Make a Deal!

You're given the choice of three doors: Behind one door is a Valuable Prize; behind the others, Something Useless.

- You pick a door, say No. 1.
- The host, who knows what's behind the doors, opens another door, say No. 3, revealing Something Useless.
- The host then asks you if you would like to switch to door No. 2.

# A probability Puzzle: Let's Make a Deal!

You're given the choice of three doors: Behind one door is a Valuable Prize; behind the others, Something Useless.

- You pick a door, say No. 1.
- The host, who knows what's behind the doors, opens another door, say No. 3, revealing Something Useless.
- The host then asks you if you would like to switch to door No. 2.
- What is the probability that the Valuable Prize is behind door No. 2?

# Summary

The Law of Total Probability

For a partition $B_1, B_2, \ldots B_k$ of $\Omega$, the probability of any other event $A$ is a weighted average of its conditional probabilities:

$$\mathbb{P}(A) = \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) + \ldots + \mathbb{P}(A|B_k)\mathbb{P}(B_k)$$

Bayes' Theorem

For a partition $B_1, B_2, \ldots B_k$ of $\Omega$, we can reverse the order of conditioning on an event $A$ to update our probabilities for events in the partition:

$$\mathbb{P}(B_j|A) = \frac{\mathbb{P}(A|B_j)\mathbb{P}(B_j)}{\mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) + \ldots + \mathbb{P}(A|B_k)\mathbb{P}(B_k)}$$