CSE 312
Foundations of Computing II

Lecture 6: Conditional Probability and Bayes Theorem

Slide Credit: Based on Stefano Tessaro’s slides for 312 19au incorporating ideas from Alex Tsun’s and Anna Karlin’s slides for 312 20su and 20au
Announcement

• PSet 1 due tonight
  – Submit both coding and written portion on Gradescope.
  – If working in a pair, remember to add your partner to your submissions!

• PSet 2 posted on website, due next Thursday

• No class or OH on Monday 1/18 (MLK Day)
Review Probability

**Definition.** A **sample space** $\Omega$ is the set of all possible outcomes of an experiment.

**Examples:**
- Single coin flip: $\Omega = \{H, T\}$
- Two coin flips: $\Omega = \{HH, HT, TH, TT\}$
- Roll of a die: $\Omega = \{1, 2, 3, 4, 5, 6\}$

**Definition.** An **event** $E \subseteq \Omega$ is a subset of possible outcomes.

**Examples:**
- Getting at least one head in two coin flips: $E = \{HH, HT, TH\}$
- Rolling an even number on a die: $E = \{2, 4, 6\}$
Review Axioms of Probability

Let $\Omega$ denote the sample space and $E, F \subseteq \Omega$ be events. Note this is more general to any probability space (not just uniform).

Axiom 1 (Non-negativity): $P(E) \geq 0$
Axiom 2 (Normalization): $P(\Omega) = 1$
Axiom 3 (Countable Additivity): If $E$ and $F$ are mutually exclusive, then $P(E \cup F) = P(E) + P(F)$

Corollary 1 (Complementation): $P(E^c) = 1 - P(E)$
Corollary 2 (Monotonicity): If $E \subseteq F$, $P(E) \leq P(F)$
Corollary 3 (Inclusion-Exclusion): $P(E \cup F) = P(E) + P(F) - P(E \cap F)$
Agenda

• Conditional Probability
• Bayes Theorem
• Law of Total Probability
• Bayes Theorem + Law of Total Probability
• More Examples
What’s the probability that someone likes ice cream **given** they like donuts?
Conditional Probability

**Definition.** The **conditional probability** of event $A$ given an event $B$ happened (assuming $P(B) \neq 0$) is

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

An equivalent and useful formula is

$$P(A \cap B) = P(A|B)P(B)$$
Conditional Probability Examples

In the popular, social video game Among Us, you are either a crewmate or an imposter. This game, you are an imposter. What is the probability you will win the game given that you are imposter?

\[ W = \text{You win a game} \]
\[ I = \text{You are the imposter in a game} \]

\[ P(W|I) = \frac{P(W \cap I)}{P(I)} \]
Reversing Conditional Probability

Question: Does $P(A|B) = P(B|A)$?

No!

• Let $A$ be the event you are wet
• Let $B$ be the event you are swimming

\[ P(A|B) = 1 \]
\[ P(B|A) \neq 1 \]
Example with Conditional Probability

Toss a red die and a blue die (both 6 sided and all outcomes equally likely). What is $P(B)$? What is $P(B|A)$?

|   | P(b) | P(B|A) |
|---|------|--------|
| a) | 1/6  | 1/6    |
| b) | 1/6  | 1/3    |
| c) | 1/6  | 3/36   |
| d) | 1/9  | 1/3    |
Gambler’s fallacy

Assume we toss 51 fair coins.
Assume we have seen 50 coins, and they are all “tails”.
What are the odds the 51st coin is “heads”?

\[ \mathcal{A} = \text{first 50 coins are “tails”} \]
\[ B = \text{first 50 coins are “tails”, 51st coin is ”heads”} \]

51st coin is independent of outcomes of first 50 tosses!

\[ P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1/2^{51}}{2/2^{51}} = \frac{1}{2} \]

Gambler’s fallacy = Feels like it’s time for “heads”!?
Agenda

• Conditional Probability
• Bayes Theorem
• Law of Total Probability
• Bayes Theorem + Law of Total Probability
• More Examples
Bayes Theorem

A formula to let us “reverse” the conditional.

**Theorem. (Bayes Rule)** For events $A$ and $B$, where $P(A), P(B) > 0$,

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

$P(A)$ is called the **prior** (our belief without knowing anything)

$P(A|B)$ is called the **posterior** (our belief after learning $B$)
Bayes Theorem Proof
Bayes Theorem Proof

By definition of conditional probability

\[ P(A \cap B) = P(A|B)P(B) \]

Swapping A, B gives

\[ P(B \cap A) = P(B|A)P(A) \]

But \( P(A \cap B) = P(B \cap A) \), so

\[ P(A|B)P(B) = P(B|A)P(A) \]

Dividing both sides by \( P(B) \) gives

\[ P(A|B) = \frac{P(B|A)P(A)}{P(B)} \]
Our First Machine Learning Task: Spam Filtering

Subject: “FREE $$$ CLICK HERE”

What is the probability this email is spam, given the subject contains “FREE”? Some useful stats:

- 10% of ham (i.e., not spam) emails contain the word “FREE” in the subject.
- 70% of spam emails contain the word “FREE” in the subject.
- 80% of emails you receive are spam.
Brain Break
Agenda

- Conditional Probability
- Bayes Theorem
- Law of Total Probability
- Bayes Theorem + Law of Total Probability
- More Examples
Partitions (Idea)

These events **partition** the sample space

1. They “cover” the whole space
2. They don’t overlap
**Definition.** Non-empty events $E_1, E_2, ..., E_n$ partition the sample space $\Omega$ if

(Exhaustive)

$$E_1 \cup E_2 \cup \cdots \cup E_n = \bigcup_{i=1}^{n} E_i = \Omega$$

(Pairwise Mutually Exclusive)

$$\forall i \forall i \neq j E_i \cap E_j = \emptyset$$
Law of Total Probability (Idea)

If we know $E_1, E_2, \ldots, E_n$ partition $\Omega$, what can we say about $P(F)$?
Law of Total Probability (LTP)

**Definition.** If events $E_1, E_2, ..., E_n$ partition the sample space $\Omega$, then for any event $F$

$$P(F) = P(F \cap E_1) + ... + P(F \cap E_n) = \sum_{i=1}^{n} P(F \cap E_i)$$

Using the definition of conditional probability $P(F \cap E) = P(F|E)P(E)$

We can get the alternate form of this that show

$$P(F) = \sum_{i=1}^{n} P(F|E_i)P(E_i)$$
Another Contrived Example

Alice has two pockets:

• **Left pocket:** Two red balls, two green balls
• **Right pocket:** One red ball, two green balls.

Alice picks a random ball from a random pocket.

[Both pockets equally likely, each ball equally likely.]
Sequential Process – Non-Uniform Case

• Left pocket: Two red, two green
• Right pocket: One red, two green.

\[ \mathbb{P}(R) = \mathbb{P}(R \cap \text{Left}) + \mathbb{P}(R \cap \text{Right}) \]  
\[ = \mathbb{P}(\text{Left}) \times \mathbb{P}(R | \text{Left}) + \mathbb{P}(\text{Right}) \times \mathbb{P}(R | \text{Right}) \]  
\[ = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{3} = \frac{1}{4} + \frac{1}{6} = \frac{5}{12} \]  

Law of total probability
Agenda

- Conditional Probability
- Bayes Theorem
- Law of Total Probability
- Bayes Theorem + Law of Total Probability
- More Examples
Our First Machine Learning Task: Spam Filtering

Subject: “FREE $$CLICK HERE”

What is the probability this email is spam, given the subject contains “FREE”? Some useful stats:

- 10% of ham (i.e., not spam) emails contain the word “FREE” in the subject.
- 70% of spam emails contain the word “FREE” in the subject.
- 80% of emails you receive are spam.
Bayes Theorem with Law of Total Probability

Bayes Theorem with LTP: Let $E_1, E_2, \ldots, E_n$ be a partition of the sample space, and $F$ an event. Then,

$$P(E_1|F) = \frac{P(F|E_1)P(E_1)}{P(F)} = \frac{P(F|E_1)P(E_1)}{\sum_{i=1}^{n} P(F|E_i)P(E_i)}$$

Simple Partition: In particular, if $E$ is an event with non-zero probability, then

$$P(E|F) = \frac{P(F|E)P(E)}{P(F|E)P(E) + P(F|E^C)P(E^C)}$$
Agenda

• Conditional Probability
• Bayes Theorem
• Law of Total Probability
• Bayes Theorem + Law of Total Probability
• More Examples
Example – Zika Testing

Usually no or mild symptoms (rash); sometimes severe symptoms (paralysis).

During pregnancy: may cause birth defects.

Suppose you took a Zika test, and it returns “positive”, what is the likelihood that you actually have the disease?

• Tests for diseases are rarely 100% accurate.
Example – Zika Testing

Suppose we know the following Zika stats

- A test is 98% effective at detecting Zika ("true positive")
- However, the test may yield a "false positive" 1% of the time
- 0.5% of the US population has Zika.

What is the probability you have Zika (event $Z$) if you test positive (event $T$).
Example – Zika Testing

Suppose we know the following Zika stats

- A test is 98% effective at detecting Zika (“true positive”)
- However, the test may yield a “false positive” 1% of the time
- 0.5% of the US population has Zika.

What is the probability you have Zika (event $Z$) if you test positive (event $T$).

Suppose we had 1000 people:

- 5 have Zika and test positive
- 985 do not have Zika and test negative
- 10 do not have Zika and test positive

$$\frac{5}{5 + 10} = \frac{1}{3} \approx 0.33$$
Philosophy – Updating Beliefs

While it’s not 98% that you have the disease, your beliefs changed **drastically**

Z = you have Zika
T = you test positive for Zika

Prior: $P(Z)$  

I have a 0.5% chance of having Zika

Receive positive test result

Posterior: $P(Z|T)$  

I now have a 33% chance of having Zika after the test!!!
Example – Zika Testing

Suppose we know the following Zika stats

– A test is 98% effective at detecting Zika (“true positive”)
– However, the test may yield a “false positive” 1% of the time
– 0.5% of the US population has Zika.

What is the probability you test negative (event $\overline{T}$) if you have Zika (event $Z$)?
Conditional Probability Define a Probability Space

The probability conditioned on $A$ follows the same properties as (unconditional) probability.

**Example.** $\mathbb{P}(B^c|A) = 1 - \mathbb{P}(B|A)$

**Formally.** $(\Omega, \mathbb{P})$ is a probability space + $\mathbb{P}(A) > 0$

$(A, \mathbb{P}(\cdot|A))$ is a probability space