THE UNIVERSITY OF TEXAS AT AUSTIN
Department of Statistics and Data Sciences
College of Natural Sciences

# SDS 321: Introduction to Probability and Statistics

## Lecture 2: Conditional probability

Purnamrita Sarkar
Department of Statistics and Data Science
The University of Texas at Austin

www.cs.cmu.edu/∼psarkar/teaching

# Example

Alice and Bob toss two fair coins separately. Denote the event that Alice gets a $H$ by $E_A$ and the event that Bob gets a $H$ by $E_B$. We know that $P(E_A) = P(E_B) = 1/2$.

# Example

Alice and Bob toss two fair coins separately. Denote the event that Alice gets a $H$ by $E_A$ and the event that Bob gets a $H$ by $E_B$. We know that $P(E_A) = P(E_B) = 1/2$.

- $P(E_A \cup E_B) = P(E_A) + P(E_B) = 1$. So its certain that at least one of them will get a head.

# Example

Alice and Bob toss two fair coins separately. Denote the event that Alice gets a $H$ by $E_A$ and the event that Bob gets a $H$ by $E_B$. We know that $P(E_A) = P(E_B) = 1/2$.

- $P(E_A \cup E_B) = P(E_A) + P(E_B) = 1$. So its certain that at least one of them will get a head.
- Whats wrong with this argument?

# Example

Alice and Bob toss two fair coins separately. Denote the event that Alice gets a $H$ by $E_A$ and the event that Bob gets a $H$ by $E_B$. We know that $P(E_A) = P(E_B) = 1/2$.

- $P(E_A \cup E_B) = P(E_A) + P(E_B) = 1$. So its certain that at least one of them will get a head.
- Whats wrong with this argument? But both could get tails!

# Example

Alice and Bob toss two fair coins separately. Denote the event that Alice gets a $H$ by $E_A$ and the event that Bob gets a $H$ by $E_B$. We know that $P(E_A) = P(E_B) = 1/2$.

- $P(E_A \cup E_B) = P(E_A) + P(E_B) = 1$. So its certain that at least one of them will get a head.
- Whats wrong with this argument? But both could get tails!
- Are $E_A$ and $E_B$ disjoint?

.

# Example

Alice and Bob toss two fair coins separately. Denote the event that Alice gets a $H$ by $E_A$ and the event that Bob gets a $H$ by $E_B$. We know that $P(E_A) = P(E_B) = 1/2$.

- $P(E_A \cup E_B) = P(E_A) + P(E_B) = 1$. So its certain that at least one of them will get a head.
- Whats wrong with this argument? But both could get tails!
- Are $E_A$ and $E_B$ disjoint? Not quite. The sample space you are considering is: $HT, TH, TT, HH$. So $E_A = \{HT, HH\}$ and $E_B = \{TH, HH\}$.

# Example

Alice and Bob toss two fair coins separately. Denote the event that Alice gets a $H$ by $E_A$ and the event that Bob gets a $H$ by $E_B$. We know that $P(E_A) = P(E_B) = 1/2$.

- $P(E_A \cup E_B) = P(E_A) + P(E_B) = 1$. So its certain that at least one of them will get a head.
- Whats wrong with this argument? But both could get tails!
- Are $E_A$ and $E_B$ disjoint? Not quite. The sample space you are considering is: $HT, TH, TT, HH$. So $E_A = \{HT, HH\}$ and $E_B = \{TH, HH\}$.
- So the union is really $P(E_A \cup E_B) = P(E_A) + P(E_B) - P(E_A \cap E_B) = 3/4$.

# Example

Alice and Bob toss two fair coins separately. Denote the event that Alice gets a $H$ by $E_A$ and the event that Bob gets a $H$ by $E_B$. We know that $P(E_A) = P(E_B) = 1/2$.

- $P(E_A \cup E_B) = P(E_A) + P(E_B) = 1$. So its certain that at least one of them will get a head.
- Whats wrong with this argument? But both could get tails!
- Are $E_A$ and $E_B$ disjoint? Not quite. The sample space you are considering is: $HT, TH, TT, HH$. So $E_A = \{HT, HH\}$ and $E_B = \{TH, HH\}$.
- So the union is really $P(E_A \cup E_B) = P(E_A) + P(E_B) - P(E_A \cap E_B) = 3/4$.
- May be what you are really thinking about is independent and not disjoint events.

# Partial information

- So far, we have assumed we know nothing about the outcome of our experiment, except for the information encoded in the probability law.
- Sometimes, however, we have **partial information** that may affect the likelihood of a given event.
  - The experiment involves rolling a die. You are told that the number is odd.
  - The experiment involves the weather tomorrow. You know that the weather today is rainy.
  - The experiment involves the presence or absence of a disease. A blood test comes back positive.
- In each case, knowing about some event $B$ (e.g. "it is raining today") changes our beliefs about event $A$ ("Will it rain tomorrow?").
- We want to **update** our probability law to incorporate this new knowledge.

# Conditional Probability

**Original problem**:

- What is the probability of some event $A$.
  - e.g. What is the probability that we roll a number less than 4?
- In other words, what is $P(A)$?
- This is given by our probability law.

# Conditional Probability

**Original problem**:

- What is the probability of some event $A$.
  - e.g. What is the probability that we roll a number less than 4?
- In other words, what is $P(A)$?
- This is given by our probability law.

**New problem**:

- Assuming event $B$ (equivalently given event $B$), what is the probability of event $A$?
  - e.g. Given that the number rolled is an odd number, what is the probability that it is less than 4?
- We call this the **conditional distribution** of $A$ given $B$.
- We write this as $P(A|B)$
- Read $|$ as "given" or "conditioned on the fact that".
- Our conditional probability is still describing "the probability of something", so we expect it to behave like a probability distribution.

# Conditional Probability

- Consider rolling a fair 6-sided die (uniform, discrete probability distribution).
- Let $A$ be the event "outcome is equal to 1'.
  - What is $P(A)$?
- Let's now assume that the number rolled is an odd number.
  - What is the set, $B$, that we are conditioning on?
- What do you think $P(A|B)$ should be?

# Conditional Probability

- You are conditioning on event $B$. So your new sample space is $\{1, 3, 5\}$ instead of $\{1, 2, 3, 4, 5, 6\}$.

# Conditional Probability

- You are conditioning on event $B$. So your new sample space is $\{1, 3, 5\}$ instead of $\{1, 2, 3, 4, 5, 6\}$.
- Each of these are equally likely. So $P(A|B) = 1/3$.

# Conditional Probability

- You are conditioning on event $B$. So your new sample space is $\{1, 3, 5\}$ instead of $\{1, 2, 3, 4, 5, 6\}$.
- Each of these are equally likely. So $P(A|B) = 1/3$.
    - Formally, if all outcomes are equally likely, we have

$$P(A|B) = \frac{\#\text{ elements in } A \cap B}{\#\text{ elements in } B}$$

# Conditional Probability

- You are conditioning on event $B$. So your new sample space is $\{1, 3, 5\}$ instead of $\{1, 2, 3, 4, 5, 6\}$.
- Each of these are equally likely. So $P(A|B) = 1/3$.
  - Formally, if all outcomes are equally likely, we have

$$P(A|B) = \frac{\#\text{ elements in } A \cap B}{\#\text{ elements in } B}$$

- More generally, $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$.

# Conditional Probability

- You are conditioning on event $B$. So your new sample space is $\{1, 3, 5\}$ instead of $\{1, 2, 3, 4, 5, 6\}$.
- Each of these are equally likely. So $P(A|B) = 1/3$.
  - Formally, if all outcomes are equally likely, we have

$$P(A|B) = \frac{\#\text{ elements in } A \cap B}{\#\text{ elements in } B}$$

- More generally, $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$.
- A conditional probability is only defined if $P(B) > 0$.

# Conditional Probability Axioms

- **Nonnegativity –** Check.
- **Normalization –**
- **Additivity –**

# Conditional Probability Axioms

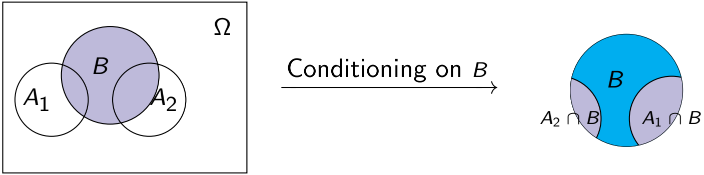
- **Nonnegativity** – Check.
- **Normalization** – Your new universe is now $B$, and we know that $P(B|B) = 1$.
- **Additivity** –

# Conditional Probability Axioms

- **Nonnegativity** – Check.
- **Normalization** – Your new universe is now $B$, and we know that $P(B|B) = 1$.
- **Additivity** – $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$ for two disjoint sets, $A_1$ and $A_2$.

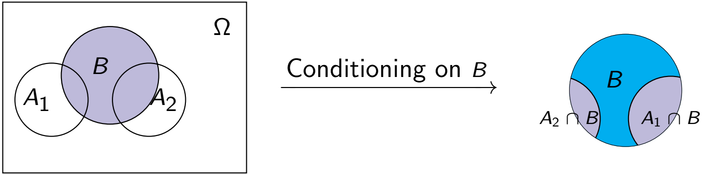
# Conditional Probability Axioms

- **Nonnegativity** – Check.
- **Normalization** – Your new universe is now $B$, and we know that $P(B|B) = 1$.
- **Additivity** – $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$ for two disjoint sets, $A_1$ and $A_2$.



# Conditional Probability Axioms

- **Nonnegativity** – Check.
- **Normalization** – Your new universe is now $B$, and we know that $P(B|B) = 1$.
- **Additivity** – $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$ for two disjoint sets, $A_1$ and $A_2$.




Using additivity, $P((A_1 \cup A_2) \cap B) = P(A_1 \cap B) + P(A_2 \cap B)$, so

$$P(A_1 \cup A_2|B) = \frac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)} = P(A_1|B) + P(A_2|B)$$

# Properties of conditional probability

If $P(B) > 0$,

- If $A_i$ for $i \in \{1, \ldots, n\}$ are all pairwise disjoint, then
  $P(\cup_{i=1}^{n} A_i | B) = \sum_{i=1}^{n} P(A_i | B)$.

# Properties of conditional probability

If $P(B) > 0$,

- If $A_i$ for $i \in \{1, \ldots, n\}$ are all pairwise disjoint, then $P(\cup_{i=1}^{n} A_i | B) = \sum_{i=1}^{n} P(A_i | B)$.
- If $A_1 \subseteq A_2$, then $P(A_1 | B) \leq P(A_2 | B)$.

# Properties of conditional probability

If $P(B) > 0$,

- If $A_i$ for $i \in \{1, \ldots, n\}$ are all pairwise disjoint, then $P(\cup_{i=1}^{n} A_i|B) = \sum_{i=1}^{n} P(A_i|B)$.
- If $A_1 \subseteq A_2$, then $P(A_1|B) \leq P(A_2|B)$.
- $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2|B)$

# Properties of conditional probability

If $P(B) > 0$,

- If $A_i$ for $i \in \{1, \ldots, n\}$ are all pairwise disjoint, then $P(\cup_{i=1}^n A_i | B) = \sum_{i=1}^n P(A_i|B)$.
- If $A_1 \subseteq A_2$, then $P(A_1|B) \leq P(A_2|B)$.
- $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2|B)$
- Union bound: $P(A_1 \cup A_2|B) \leq P(A_1|B) + P(A_2|B)$

# Properties of conditional probability

If $P(B) > 0$,

- If $A_i$ for $i \in \{1, \ldots, n\}$ are all pairwise disjoint, then $P(\cup_{i=1}^n A_i|B) = \sum_{i=1}^n P(A_i|B)$.
- If $A_1 \subseteq A_2$, then $P(A_1|B) \leq P(A_2|B)$.
- $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2|B)$
- Union bound: $P(A_1 \cup A_2|B) \leq P(A_1|B) + P(A_2|B)$
  - $P(\cup_{i=1}^n A_i|B) \leq \sum_{i=1}^n P(A_i|B)$.

# Example: Coin toss

Consider the experiment of tossing a fair coin three times. What is the probability of getting alternating heads and tails conditioned on the event that your first toss gives a head?

- Notation: Let $A := \{\text{Tosses yield alternating tails and heads}\}$ and $B := \{\text{The first toss is a head}\}$.
- We want $P(A|B)$.
- What if you wanted to follow the formula?

# Example: Coin toss

Consider the experiment of tossing a fair coin three times. What is the probability of getting alternating heads and tails conditioned on the event that your first toss gives a head?

- Notation: Let $A := \{\text{Tosses yield alternating tails and heads}\}$ and $B := \{\text{The first toss is a head}\}$.
- We want $P(A|B)$.
- What if you wanted to follow the formula?
  - Sample space for three coin tosses $\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.

# Example: Coin toss

Consider the experiment of tossing a fair coin three times. What is the probability of getting alternating heads and tails conditioned on the event that your first toss gives a head?

- Notation: Let $A := \{\text{Tosses yield alternating tails and heads}\}$ and $B := \{\text{The first toss is a head}\}$.
- We want $P(A|B)$.
- What if you wanted to follow the formula?
  - Sample space for three coin tosses $\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.
  - $A = \{HTH, THT\}$ and $A \cap B = \{HTH\}$.

# Example: Coin toss

Consider the experiment of tossing a fair coin three times. What is the probability of getting alternating heads and tails conditioned on the event that your first toss gives a head?

- Notation: Let $A := \{\text{Tosses yield alternating tails and heads}\}$ and $B := \{\text{The first toss is a head}\}$.
- We want $P(A|B)$.
- What if you wanted to follow the formula?
  - Sample space for three coin tosses $\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.
  - $A = \{HTH, THT\}$ and $A \cap B = \{HTH\}$.
  - We have $P(B) = 4/8$ and $P(A \cap B) = 1/8$ and so $P(A|B) = 1/4$.

# Example: Coin toss

Consider the experiment of tossing a fair coin three times. What is the probability of getting alternating heads and tails conditioned on the event that your first toss gives a head?

- Notation: Let $A := \{\text{Tosses yield alternating tails and heads}\}$ and $B := \{\text{The first toss is a head}\}$.
- We want $P(A|B)$.
- What if you wanted to follow the formula?
  - Sample space for three coin tosses $\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.
  - $A = \{HTH, THT\}$ and $A \cap B = \{HTH\}$.
  - We have $P(B) = 4/8$ and $P(A \cap B) = 1/8$ and so $P(A|B) = 1/4$.

  - Your new world/sample space when you condition is $B = \{HHT, HTH, HTT, HHH\}$. Each of these are equally likely. Out of these 1 event satisfies alternating heads and tails.

# Example: Coin toss

Consider the experiment of tossing a fair coin three times. What is the probability of getting alternating heads and tails conditioned on the event that your first toss gives a head?

- Notation: Let $A := \{\text{Tosses yield alternating tails and heads}\}$ and $B := \{\text{The first toss is a head}\}$.
- We want $P(A|B)$.
- What if you wanted to follow the formula?
  - Sample space for three coin tosses $\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.
  - $A = \{HTH, THT\}$ and $A \cap B = \{HTH\}$.
  - We have $P(B) = 4/8$ and $P(A \cap B) = 1/8$ and so $P(A|B) = 1/4$.

  - Your new world/sample space when you condition is $B = \{HHT, HTH, HTT, HHH\}$. Each of these are equally likely. Out of these 1 event satisfies alternating heads and tails.
  - So, $P(A|B) = 1/4$.

# Homework

- Read Sections 1.1, 1.2 and 1.3 of Bertsekas and Tsitsiklis.
- The first homework will be posted online today. It is due next Thursday by 5pm via Canvas.