Probability

Prof. Wells

STA 209, 4/7/23

Outline

In this lecture, we will. . .

Outline

In this lecture, we will. . .

- Introduce key definitions for probability theory
- Define conditional probability
- Investigate Bayes Rule for conditional probabilities

Section 1

Probability Theory

Random Processes and Events

Probability Theory is the study and quantification of uncertainty and randomness in outcomes of repeated experiments.

Random Processes and Events

Probability Theory is the study and quantification of uncertainty and randomness in outcomes of repeated experiments.

- A **random process** is one which we know what results *could* happen, but don't know which particular result *will* happen.

Random Processes and Events

Probability Theory is the study and quantification of uncertainty and randomness in outcomes of repeated experiments.

- A **random process** is one which we know what results *could* happen, but don't know which particular result *will* happen.
- An **event** is a statement that can be ascertained based on the result of a particular random process.

Random Processes and Events

Probability Theory is the study and quantification of uncertainty and randomness in outcomes of repeated experiments.

- A **random process** is one which we know what results *could* happen, but don't know which particular result *will* happen.
- An **event** is a statement that can be ascertained based on the result of a particular random process.
    - Example: Suppose we roll a 6-sided die. One event is *"the die rolls a 6"*. Another is *"the die rolls an odd number"*.

Random Processes and Events

Probability Theory is the study and quantification of uncertainty and randomness in outcomes of repeated experiments.

- A **random process** is one which we know what results *could* happen, but don't know which particular result *will* happen.
- An **event** is a statement that can be ascertained based on the result of a particular random process.
  - Example: Suppose we roll a 6-sided die. One event is *"the die rolls a 6"*. Another is *"the die rolls an odd number"*.
  - The statement *"it will rain at 4pm in Grinnell on 4/7"* is not an event for this particular process; we cannot determine whether it occurs based on the results of the die roll

Random Processes and Events

Probability Theory is the study and quantification of uncertainty and randomness in outcomes of repeated experiments.

- A **random process** is one which we know what results *could* happen, but don't know which particular result *will* happen.
- An **event** is a statement that can be ascertained based on the result of a particular random process.
  - Example: Suppose we roll a 6-sided die. One event is *"the die rolls a 6"*. Another is *"the die rolls an odd number"*.
  - The statement *"it will rain at 4pm in Grinnell on 4/7"* is not an event for this particular process; we cannot determine whether it occurs based on the results of the die roll
  - For brevity, we'll often denote events using capital letters like A, B, C.

Random Processes and Events

Probability Theory is the study and quantification of uncertainty and randomness in outcomes of repeated experiments.

- A **random process** is one which we know what results *could* happen, but don't know which particular result *will* happen.
- An **event** is a statement that can be ascertained based on the result of a particular random process.
    - Example: Suppose we roll a 6-sided die. One event is *"the die rolls a 6"*. Another is *"the die rolls an odd number"*.
    - The statement *"it will rain at 4pm in Grinnell on 4/7"* is not an event for this particular process; we cannot determine whether it occurs based on the results of the die roll
    - For brevity, we'll often denote events using capital letters like A, B, C.
    - For example, we might say: Let A be the event "the sum of values on two dice equals 7"

Probability

- (The Law of Large Numbers) The **probability** of a particular event is the proportion of times the event would occur, if we observed the random process an *arbitrarily large* number of times.

Probability

- (The Law of Large Numbers) The **probability** of a particular event is the proportion of times the event would occur, if we observed the random process an *arbitrarily large* number of times.
  - To say that *a coin has 50% probability of landing heads*, means that...

Probability

- (The Law of Large Numbers) The **probability** of a particular event is the proportion of times the event would occur, if we observed the random process an *arbitrarily large* number of times.
  - To say that *a coin has 50% probability of landing heads*, means that. . .
  - In a large number of coin flips, we expect the proportion of heads to be close to 0.5, and that this proportion should get closer to 0.5 with additional flips.

Probability

- (The Law of Large Numbers) The **probability** of a particular event is the proportion of times the event would occur, if we observed the random process an *arbitrarily large* number of times.
  - To say that *a coin has 50% probability of landing heads*, means that...
  - In a large number of coin flips, we expect the proportion of heads to be close to 0.5, and that this proportion should get closer to 0.5 with additional flips.
- Since probabilities are defined in terms of proportions, they will always be values between 0 and 1.

Probability

- (The Law of Large Numbers) The **probability** of a particular event is the proportion of times the event would occur, if we observed the random process an *arbitrarily large* number of times.
  - To say that *a coin has 50% probability of landing heads*, means that...
  - In a large number of coin flips, we expect the proportion of heads to be close to 0.5, and that this proportion should get closer to 0.5 with additional flips.
- Since probabilities are defined in terms of proportions, they will always be values between 0 and 1.
- For brevity, we'll represent statements like *the probability of the event "the coin lands heads" is 50%* using the notation:

$$P(\text{the coin lands heads}) = 0.5 \quad \text{or} \quad P(\text{Heads}) = 0.5 \quad \text{or} \quad P(H) = 0.5$$

Law of Large Numbers

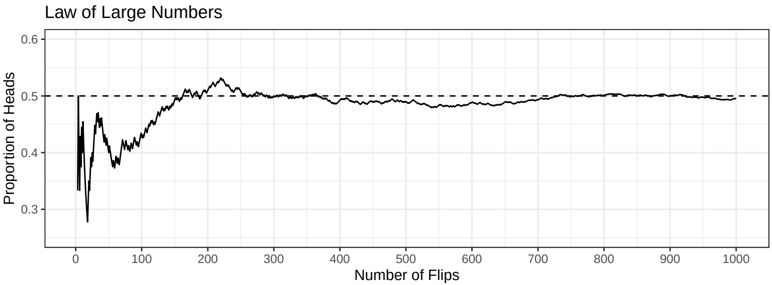
- The probability of an event refers to the long-run tendency of the proportion.

Law of Large Numbers

- The probability of an event refers to the long-run tendency of the proportion.
  - It is possible for the observed proportion to deviate from the expected probability in a finite number of trials.
  - But the magnitude of these deviations decreases as the number of trials increases.

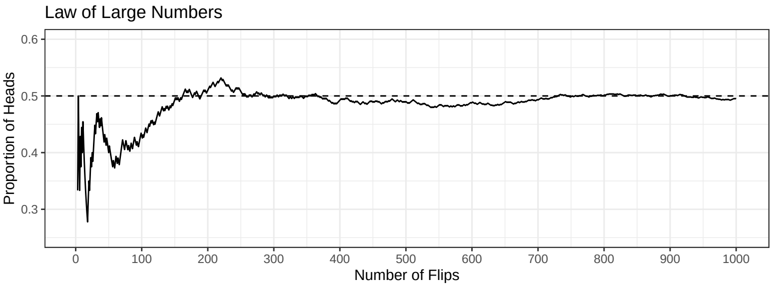
Law of Large Numbers

- The probability of an event refers to the long-run tendency of the proportion.
  - It is possible for the observed proportion to deviate from the expected probability in a finite number of trials.
  - But the magnitude of these deviations decreases as the number of trials increases.



Law of Large Numbers

- The probability of an event refers to the long-run tendency of the proportion.
  - It is possible for the observed proportion to deviate from the expected probability in a finite number of trials.
  - But the magnitude of these deviations decreases as the number of trials increases.



- Note that the proportion of heads deviates (significantly) from 0.5 during the first 50 flips, but stabilizes around 0.5 by 1000 flips.

Probability Models

- In order to know the probability of an event, we would need to witness the random process an infinite number of times.
  - Therefore, we will *never* know the precise probability of a real-world event.
  - Instead, we construct *models* which we hope are good representations of random phenomenon.

Probability Models

- In order to know the probability of an event, we would need to witness the random process an infinite number of times.
  - Therefore, we will *never* know the precise probability of a real-world event.
  - Instead, we construct *models* which we hope are good representations of random phenomenon.
- A **probability model** has two components:
  1. A list of the possible results of a random process (called **events**)
  2. A rule (called the **probability function**) that assigns to each event a probability between 0 and 1, in a consistent manner.

Probability Models

- In order to know the probability of an event, we would need to witness the random process an infinite number of times.
  - Therefore, we will *never* know the precise probability of a real-world event.
  - Instead, we construct *models* which we hope are good representations of random phenomenon.
- A **probability model** has two components:
  1. A list of the possible results of a random process (called **events**)
  2. A rule (called the **probability function**) that assigns to each event a probability between 0 and 1, in a consistent manner.
- Whenever we discuss probability, we are always (either explicitly or implicitly) defining a probability model.

Probability Models

- In order to know the probability of an event, we would need to witness the random process an infinite number of times.
  - Therefore, we will *never* know the precise probability of a real-world event.
  - Instead, we construct *models* which we hope are good representations of random phenomenon.
- A **probability model** has two components:
  1. A list of the possible results of a random process (called **events**)
  2. A rule (called the **probability function**) that assigns to each event a probability between 0 and 1, in a consistent manner.
- Whenever we discuss probability, we are always (either explicitly or implicitly) defining a probability model.
- A large component of statistical inference is comparing observed data to the results we would expect under a certain probability model.

The Addition Rule

- Two events A and B are said to be **mutually exclusive** (or disjoint) if it is not possible for both to occur at the same time.

The Addition Rule

- Two events A and B are said to be **mutually exclusive** (or disjoint) if it is not possible for both to occur at the same time.
  - Let A denote the event "a 1 is rolled on 6-sided die" and B denote the event "a 2 is rolled on 6-sided die". Then A and B are disjoint events describing a single die roll.

The Addition Rule

- Two events A and B are said to be **mutually exclusive** (or disjoint) if it is not possible for both to occur at the same time.
  - Let A denote the event "a 1 is rolled on 6-sided die" and B denote the event "a 2 is rolled on 6-sided die". Then A and B are disjoint events describing a single die roll.

Theorem (Addition Rule)

The probability that at least one event occurs in a pair of disjoint events is the sum of their individual probabilities:

$$P(A \text{ } or \text{ } B) = P(A) + P(B)$$

The Addition Rule

- Two events A and B are said to be **mutually exclusive** (or disjoint) if it is not possible for both to occur at the same time.
  - Let A denote the event "a 1 is rolled on 6-sided die" and B denote the event "a 2 is rolled on 6-sided die". Then A and B are disjoint events describing a single die roll.

Theorem (Addition Rule)

The probability that at least one event occurs in a pair of disjoint events is the sum of their individual probabilities:

$$P(A \text{ or } B) = P(A) + P(B)$$

- When a fair 6-sided die is rolled, each number 1 - 6 has equal chance of appearing. What is the probability that a number whose name starts with f is rolled?

The Addition Rule

- Two events A and B are said to be **mutually exclusive** (or disjoint) if it is not possible for both to occur at the same time.
  - Let A denote the event "a 1 is rolled on 6-sided die" and B denote the event "a 2 is rolled on 6-sided die". Then A and B are disjoint events describing a single die roll.

Theorem (Addition Rule)

The probability that at least one event occurs in a pair of disjoint events is the sum of their individual probabilities:

$$P(A \textit{ or } B) = P(A) + P(B)$$

- When a fair 6-sided die is rolled, each number 1 - 6 has equal chance of appearing. What is the probability that a number whose name starts with f is rolled?

$$P(\text{ starts with f }) = P(\text{ roll a 4 or roll a 5 }) = P(\text{ roll a 4 }) + P(\text{ roll a 5 }) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

Complementary Events

- The complement to an event A (denoted A^c) is the event that occurs exactly when the original does not.

Complementary Events

- The complement to an event A (denoted A^c) is the event that occurs exactly when the original does not.
  - If A is the event that a 1 is rolled on a 6-sided die, then A^c is the event that a 2,3,4,5, or 6 is rolled.

Complementary Events

- The complement to an event A (denoted A^c) is the event that occurs exactly when the original does not.
  - If A is the event that a 1 is rolled on a 6-sided die, then A^c is the event that a 2,3,4,5, or 6 is rolled.

Theorem (Complement Rule)

The probability that the complement of an event occurs is 1 minus the probability of the event:

$$P(A^c) = 1 - P(A)$$

Complementary Events

- The complement to an event A (denoted A^c) is the event that occurs exactly when the original does not.
  - If A is the event that a 1 is rolled on a 6-sided die, then A^c is the event that a 2,3,4,5, or 6 is rolled.

Theorem (Complement Rule)

The probability that the complement of an event occurs is 1 minus the probability of the event:

$$P(A^c) = 1 - P(A)$$

- What is the probability that any number other than a 1 is rolled on a fair 6-sided die?

Complementary Events

- The complement to an event A (denoted A^c) is the event that occurs exactly when the original does not.
  - If A is the event that a 1 is rolled on a 6-sided die, then A^c is the event that a 2,3,4,5, or 6 is rolled.

Theorem (Complement Rule)

The probability that the complement of an event occurs is 1 minus the probability of the event:

$$P(A^c) = 1 - P(A)$$

- What is the probability that any number other than a 1 is rolled on a fair 6-sided die?

$$P(\text{ roll something other than a 1 }) = 1 - P(\text{ roll a 1}) = 1 - \frac{1}{6} = \frac{5}{6}$$

Section 2

Conditional Probability

Coffee or Tea?

A survey was given to 100 Intro Stats students in 2022. Some results are summarized below:

Coffee or Tea?

A survey was given to 100 Intro Stats students in 2022. Some results are summarized below:

	Coffee	Tea	total
First-year	7	10	17
Sophomore	25	20	45
Junior	13	12	25
Senior	8	5	13
total	53	47	100

Coffee or Tea?

A survey was given to 100 Intro Stats students in 2022. Some results are summarized below:

	Coffee	Tea	total
First-year	7	10	17
Sophomore	25	20	45
Junior	13	12	25
Senior	8	5	13
total	53	47	100

1. What is the probability that a random student prefers coffee?

Coffee or Tea?

A survey was given to 100 Intro Stats students in 2022. Some results are summarized below:

	Coffee	Tea	total
First-year	7	10	17
Sophomore	25	20	45
Junior	13	12	25
Senior	8	5	13
total	53	47	100

1. What is the probability that a random student prefers coffee?
2. What is the probability that a random student was a sophomore?

Coffee or Tea?

A survey was given to 100 Intro Stats students in 2022. Some results are summarized below:

	Coffee	Tea	total
First-year	7	10	17
Sophomore	25	20	45
Junior	13	12	25
Senior	8	5	13
total	53	47	100

1. What is the probability that a random student prefers coffee?
2. What is the probability that a random student was a sophomore?
3. What is the probability that a random student was a sophomore who preferred coffee?

Coffee or Tea?

A survey was given to 100 Intro Stats students in 2022. Some results are summarized below:

	Coffee	Tea	total
First-year	7	10	17
Sophomore	25	20	45
Junior	13	12	25
Senior	8	5	13
total	53	47	100

1. What is the probability that a random student prefers coffee?
2. What is the probability that a random student was a sophomore?
3. What is the probability that a random student was a sophomore who preferred coffee?
4. What is the probability that a random sophomore preferred coffee?

Conditional Probability

- The **conditional probability** of an event A given another event B is

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

Conditional Probability

- The **conditional probability** of an event A given another event B is

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

  - In the previous example, what is the conditional probability that a student prefers coffee, given that the student is a sophomore?

Conditional Probability

- The **conditional probability** of an event A given another event B is

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

  - In the previous example, what is the conditional probability that a student prefers coffee, given that the student is a sophomore?

- How can we use conditional probability to find the probability that two events both occur?

Theorem (General Multiplication Rule)

For any events A and B,

$$P(A \textit{ and } B) = P(A|B)P(B) = P(B|A)P(A)$$

Conditional Probability

- The **conditional probability** of an event A given another event B is

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

  - In the previous example, what is the conditional probability that a student prefers coffee, given that the student is a sophomore?

- How can we use conditional probability to find the probability that two events both occur?

Theorem (General Multiplication Rule)

For any events A and B,

$$P(A \text{ and } B) = P(A|B)P(B) = P(B|A)P(A)$$

- Find the probability that a randomly chosen student is a sophomore who likes coffee.

Conditioning and Independence

- We say that two events are independent if knowing that one occurs doesn't change the probability that the other occurs

Conditioning and Independence

- We say that two events are independent if knowing that one occurs doesn't change the probability that the other occurs

Theorem (Criteria for Independence)

Two events A and B are independent exactly when

$$P(A|B) = P(A) \quad and \quad P(B|A) = P(B)$$

Conditioning and Independence

- We say that two events are independent if knowing that one occurs doesn't change the probability that the other occurs

Theorem (Criteria for Independence)

Two events A and B are independent exactly when

$$P(A|B) = P(A) \quad \textit{and} \quad P(B|A) = P(B)$$

- In the previous example, are the events that "a student is a sophomore" and "a student prefers coffee" independent?

Conditioning and Independence

- We say that two events are independent if knowing that one occurs doesn't change the probability that the other occurs

Theorem (Criteria for Independence)

Two events A and B are independent exactly when

$$P(A|B) = P(A) \quad and \quad P(B|A) = P(B)$$

- In the previous example, are the events that "a student is a sophomore" and "a student prefers coffee" independent?
- **Note!** Saying that two events are *disjoint* is not the same as saying they are independent.

Conditioning and Independence

- We say that two events are independent if knowing that one occurs doesn't change the probability that the other occurs

Theorem (Criteria for Independence)

Two events A and B are independent exactly when

$$P(A|B) = P(A) \quad \textit{and} \quad P(B|A) = P(B)$$

- In the previous example, are the events that "a student is a sophomore" and "a student prefers coffee" independent?
- **Note!** Saying that two events are *disjoint* is not the same as saying they are independent.
  - Disjoint events cannot simultaneously occur; while for independent events, knowing that one occurs gives no information about whether the other occurs.

Conditioning and Independence

- We say that two events are independent if knowing that one occurs doesn't change the probability that the other occurs

Theorem (Criteria for Independence)

Two events A and B are independent exactly when

$$P(A|B) = P(A) \quad \textit{and} \quad P(B|A) = P(B)$$

- In the previous example, are the events that "a student is a sophomore" and "a student prefers coffee" independent?
- **Note!** Saying that two events are *disjoint* is not the same as saying they are independent.
  - Disjoint events cannot simultaneously occur; while for independent events, knowing that one occurs gives no information about whether the other occurs.
  - If two events are disjoint, and if you know one has occurred, then you automatically know the other cannot occur. Disjoint events are as far away from independence as possible!

Multiplication Rule for Independent Events

- The general multiplication rule is simpler in the case when two events are independent:

Theorem (Independent Multiplication Rule)

If events A and B are independent, the

$$P(A \text{ and } B) = P(A)P(B)$$

Multiplication Rule for Independent Events

- The general multiplication rule is simpler in the case when two events are independent:

Theorem (Independent Multiplication Rule)

If events A and B are independent, the

$$P(A \text{ and } B) = P(A)P(B)$$

- Suppose we flip a fair coin twice. What is the probability that both flips are heads?

Multiplication Rule for Independent Events

- The general multiplication rule is simpler in the case when two events are independent:

Theorem (Independent Multiplication Rule)

If events A and B are independent, the

$$P(A \text{ } and \text{ } B) = P(A)P(B)$$

- Suppose we flip a fair coin twice. What is the probability that both flips are heads?
  - Let A be the event that the first flip is heads and B be the event that the second is heads.

Multiplication Rule for Independent Events

- The general multiplication rule is simpler in the case when two events are independent:

Theorem (Independent Multiplication Rule)

If events A and B are independent, the

$$P(A \textit{ and } B) = P(A)P(B)$$

- Suppose we flip a fair coin twice. What is the probability that both flips are heads?
  - Let A be the event that the first flip is heads and B be the event that the second is heads.
  - Since the result of the first flip has no bearing on the second flip, then A and B are independent.

Multiplication Rule for Independent Events

- The general multiplication rule is simpler in the case when two events are independent:

Theorem (Independent Multiplication Rule)

If events A and B are independent, the

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

- Suppose we flip a fair coin twice. What is the probability that both flips are heads?
  - Let A be the event that the first flip is heads and B be the event that the second is heads.
  - Since the result of the first flip has no bearing on the second flip, then A and B are independent.
  - Moreover, since we have a fair coin, then $P(A) = P(B) = \frac{1}{2}$ and so

$$P(A \text{ and } B) = P(A)P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Practice with Probability

- Suppose we obtain a random sample of 4 Grinnell students: $\{A, B, C, D\}$ and wish to form new *bootstrap sample*. What is the probability that student A is *not* a member of the bootstrap sample?

Practice with Probability

- Suppose we obtain a random sample of 4 Grinnell students: $\{A, B, C, D\}$ and wish to form new *bootstrap sample*. What is the probability that student A is *not* a member of the bootstrap sample?
  - To create the bootstrap, we sample 4 times with replacement from the original sample.

Practice with Probability

- Suppose we obtain a random sample of 4 Grinnell students: $\{A, B, C, D\}$ and wish to form new *bootstrap sample*. What is the probability that student A is *not* a member of the bootstrap sample?
  - To create the bootstrap, we sample 4 times with replacement from the original sample.
  - In order for A not to be in the bootstrap sample, A must not be selected first, nor second, nor third, nor fourth.

Practice with Probability

- Suppose we obtain a random sample of 4 Grinnell students: $\{A, B, C, D\}$ and wish to form new *bootstrap sample*. What is the probability that student A is *not* a member of the bootstrap sample?
  - To create the bootstrap, we sample 4 times with replacement from the original sample.
  - In order for A not to be in the bootstrap sample, A must not be selected first, nor second, nor third, nor fourth.
  - To calculate the probability that A is not selected first, we use the complement rule:

$$P(A \text{ not 1st}) = 1 - P(A \text{ 1st}) = 1 - \frac{1}{4} = \frac{3}{4}$$

Practice with Probability

- Suppose we obtain a random sample of 4 Grinnell students: $\{A, B, C, D\}$ and wish to form new *bootstrap sample*. What is the probability that student A is *not* a member of the bootstrap sample?
  - To create the bootstrap, we sample 4 times with replacement from the original sample.
  - In order for A not to be in the bootstrap sample, A must not be selected first, nor second, nor third, nor fourth.
  - To calculate the probability that A is not selected first, we use the complement rule:

$$P(A \text{ not 1st}) = 1 - P(A \text{ 1st}) = 1 - \frac{1}{4} = \frac{3}{4}$$

Practice with Probability

- Suppose we obtain a random sample of 4 Grinnell students: $\{A, B, C, D\}$ and wish to form new *bootstrap sample*. What is the probability that student A is *not* a member of the bootstrap sample?
  - To create the bootstrap, we sample 4 times with replacement from the original sample.
  - In order for A not to be in the bootstrap sample, A must not be selected first, nor second, nor third, nor fourth.
  - To calculate the probability that A is not selected first, we use the complement rule:
  $$P(A \text{ not 1st}) = 1 - P(A \text{ 1st}) = 1 - \frac{1}{4} = \frac{3}{4}$$
  - Similarly, the probabilities that A is not 2nd, not 3rd, and not 4th are also $\frac{3}{4}$.

Practice with Probability

- Suppose we obtain a random sample of 4 Grinnell students: $\{A, B, C, D\}$ and wish to form new *bootstrap sample*. What is the probability that student A is *not* a member of the bootstrap sample?
  - To create the bootstrap, we sample 4 times with replacement from the original sample.
  - In order for A not to be in the bootstrap sample, A must not be selected first, nor second, nor third, nor fourth.
  - To calculate the probability that A is not selected first, we use the complement rule:

$$P(A \text{ not 1st}) = 1 - P(A \text{ 1st}) = 1 - \frac{1}{4} = \frac{3}{4}$$

  - Similarly, the probabilities that A is not 2nd, not 3rd, and not 4th are also $\frac{3}{4}$.
  - Since we are sampling with replacement, these 4 events are **independent**, so...

Practice with Probability

- Suppose we obtain a random sample of 4 Grinnell students: $\{A, B, C, D\}$ and wish to form new *bootstrap sample*. What is the probability that student A is *not* a member of the bootstrap sample?
  - To create the bootstrap, we sample 4 times with replacement from the original sample.
  - In order for A not to be in the bootstrap sample, A must not be selected first, nor second, nor third, nor fourth.
  - To calculate the probability that A is not selected first, we use the complement rule:

$$P(A \text{ not 1st}) = 1 - P(A \text{ 1st}) = 1 - \frac{1}{4} = \frac{3}{4}$$

  - Similarly, the probabilities that A is not 2nd, not 3rd, and not 4th are also $\frac{3}{4}$.
  - Since we are sampling with replacement, these 4 events are **independent**, so...

$$\begin{aligned} P(A \text{ not in bootstrap}) =& P(A \text{ not 1st}) \cdot P(A \text{ not 2nd}) \cdot P(A \text{ not 3rd}) \cdot P(A \text{ not 4th}) \\ =& \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} = \frac{81}{256} \approx 0.32 \end{aligned}$$

Is conditional probability symmetric?

Consider two events A and B. Is it always true that $P(A|B) = P(B|A)$?

Is conditional probability symmetric?

Consider two events A and B. Is it always true that $P(A|B) = P(B|A)$?

- Suppose we flip two coins. Let A be the event "the first flip is heads" and let B be the event "at least one flip is heads".

Is conditional probability symmetric?

Consider two events A and B. Is it always true that $P(A|B) = P(B|A)$?

- Suppose we flip two coins. Let A be the event "the first flip is heads" and let B be the event "at least one flip is heads".
  - The event B occurs if we get one of HH, HT, TH. So $P(B) = \frac{3}{4}$

Is conditional probability symmetric?

Consider two events A and B. Is it always true that $P(A|B) = P(B|A)$?

- Suppose we flip two coins. Let A be the event "the first flip is heads" and let B be the event "at least one flip is heads".
  - The event B occurs if we get one of HH, HT, TH. So $P(B) = \frac{3}{4}$
  - The event A occurs if we get one of HT or HH, so $P(A) = \frac{1}{2}$.

Is conditional probability symmetric?

Consider two events A and B. Is it always true that $P(A|B) = P(B|A)$?

- Suppose we flip two coins. Let A be the event "the first flip is heads" and let B be the event "at least one flip is heads".
  - The event B occurs if we get one of HH, HT, TH. So $P(B) = \frac{3}{4}$
  - The event A occurs if we get one of HT or HH, so $P(A) = \frac{1}{2}$.
  - The events A and B both occur if we get one of HT or HH, so $P(A \text{ and } B) = \frac{1}{2}$.

Is conditional probability symmetric?

Consider two events A and B. Is it always true that $P(A|B) = P(B|A)$?

- Suppose we flip two coins. Let A be the event "the first flip is heads" and let B be the event "at least one flip is heads".
  - The event B occurs if we get one of HH, HT, TH. So $P(B) = \frac{3}{4}$
  - The event A occurs if we get one of HT or HH, so $P(A) = \frac{1}{2}$.
  - The events A and B both occur if we get one of HT or HH, so $P(A \text{ and } B) = \frac{1}{2}$.

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{\frac{1}{2}}{\frac{3}{4}} = \frac{2}{3} \qquad P(B|A) = \frac{P(A \text{ and } B)}{P(A)} = \frac{\frac{1}{2}}{\frac{1}{2}} = 1$$

Is conditional probability symmetric?

Consider two events A and B. Is it always true that $P(A|B) = P(B|A)$?

- Suppose we flip two coins. Let A be the event "the first flip is heads" and let B be the event "at least one flip is heads".
  - The event B occurs if we get one of HH, HT, TH. So $P(B) = \frac{3}{4}$
  - The event A occurs if we get one of HT or HH, so $P(A) = \frac{1}{2}$.
  - The events A and B both occur if we get one of HT or HH, so $P(A \text{ and } B) = \frac{1}{2}$.

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{\frac{1}{2}}{\frac{3}{4}} = \frac{2}{3} \qquad P(B|A) = \frac{P(A \text{ and } B)}{P(A)} = \frac{\frac{1}{2}}{\frac{1}{2}} = 1$$

- Suppose we choose a random car in Iowa. Consider the events "The car is red" and "The car is a Ferrari".

Is conditional probability symmetric?

Consider two events A and B. Is it always true that $P(A|B) = P(B|A)$?

- Suppose we flip two coins. Let A be the event "the first flip is heads" and let B be the event "at least one flip is heads".
  - The event B occurs if we get one of HH, HT, TH. So $P(B) = \frac{3}{4}$
  - The event A occurs if we get one of HT or HH, so $P(A) = \frac{1}{2}$.
  - The events A and B both occur if we get one of HT or HH, so $P(A \text{ and } B) = \frac{1}{2}$.

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{\frac{1}{2}}{\frac{3}{4}} = \frac{2}{3} \qquad P(B|A) = \frac{P(A \text{ and } B)}{P(A)} = \frac{\frac{1}{2}}{\frac{1}{2}} = 1$$

- Suppose we choose a random car in Iowa. Consider the events "The car is red" and "The car is a Ferrari".
  - What is $P(\text{ car is red}|\text{ car is Ferrari})$?

Is conditional probability symmetric?

Consider two events A and B. Is it always true that $P(A|B) = P(B|A)$?

- Suppose we flip two coins. Let A be the event "the first flip is heads" and let B be the event "at least one flip is heads".
  - The event B occurs if we get one of HH, HT, TH. So $P(B) = \frac{3}{4}$
  - The event A occurs if we get one of HT or HH, so $P(A) = \frac{1}{2}$.
  - The events A and B both occur if we get one of HT or HH, so $P(A \text{ and } B) = \frac{1}{2}$.

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{\frac{1}{2}}{\frac{3}{4}} = \frac{2}{3} \qquad P(B|A) = \frac{P(A \text{ and } B)}{P(A)} = \frac{\frac{1}{2}}{\frac{1}{2}} = 1$$

- Suppose we choose a random car in Iowa. Consider the events "The car is red" and "The car is a Ferrari".
  - What is $P(\text{ car is red}|\text{ car is Ferrari})$?
  - But what is $P(\text{ car is Ferrari}|\text{ car is red})$?

Bayes' Rule

To relate $P(A|B)$ and $P(B|A)$, we use the following theorem:

Theorem (Bayes' Rule)

Let A and B be events.Then

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

Bayes' Rule

To relate $P(A|B)$ and $P(B|A)$, we use the following theorem:

Theorem (Bayes' Rule)

Let A and B be events. Then

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

- Why is this rule true?

Bayes' Rule

To relate $P(A|B)$ and $P(B|A)$, we use the following theorem:

Theorem (Bayes' Rule)

Let A and B be events. Then

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

- Why is this rule true?

$$P(A|B)P(B) = P(A \text{ and } B) = P(B|A)P(A)$$

Bayes' Rule

To relate $P(A|B)$ and $P(B|A)$, we use the following theorem:

Theorem (Bayes' Rule)

Let A and B be events. Then

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

- Why is this rule true?

$$P(A|B)P(B) = P(A \text{ and } B) = P(B|A)P(A)$$

- Under what circumstances will $P(A|B) = P(B|A)$?

$$\text{if } P(A) = P(B), \text{ then } P(A|B) = P(B|A)\frac{P(A)}{P(B)} = P(B|A)$$

Bayes' Rule

To relate $P(A|B)$ and $P(B|A)$, we use the following theorem:

Theorem (Bayes' Rule)

Let A and B be events. Then

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

- Why is this rule true?

$$P(A|B)P(B) = P(A \text{ and } B) = P(B|A)P(A)$$

- Under what circumstances will $P(A|B) = P(B|A)$?

$$\text{if } P(A) = P(B), \text{ then } P(A|B) = P(B|A)\frac{P(A)}{P(B)} = P(B|A)$$

- Under what circumstances will $P(A|B)$ be smaller than $P(B|A)$?

$$\text{if } P(A) < P(B), \text{ then } P(A|B) = P(B|A)\frac{P(A)}{P(B)} < P(B|A)$$

Bayes Rule and P-Values

- Recall: The p-value for a hypothesis test is the probability of obtaining a test statistic more extreme than the one observed, if the null hypothesis were true.

Bayes Rule and P-Values

- Recall: The p-value for a hypothesis test is the probability of obtaining a test statistic more extreme than the one observed, if the null hypothesis were true.
- In the language of conditional probability:

$$p\text{-value} = P(\text{ extreme statistic } | H_0 \text{ is true})$$

Bayes Rule and P-Values

- Recall: The p-value for a hypothesis test is the probability of obtaining a test statistic more extreme than the one observed, if the null hypothesis were true.
- In the language of conditional probability:

$$p\text{-value} = P(\text{ extreme statistic } | H_0 \text{ is true})$$

Bayes Rule and P-Values

- Recall: The *p*-value for a hypothesis test is the probability of obtaining a test statistic more extreme than the one observed, if the null hypothesis were true.
- In the language of conditional probability:

$$p\text{-value} = P(\text{ extreme statistic } | H_0 \text{ is true})$$

- From the previous slide

$$P(\text{ extreme statistic } | H_0 \text{ is true}) \neq P(H_0 \text{ is true} | \text{ extreme statistic })$$

Bayes Rule and P-Values

- Recall: The p-value for a hypothesis test is the probability of obtaining a test statistic more extreme than the one observed, if the null hypothesis were true.
- In the language of conditional probability:

$$p\text{-value} = P(\text{ extreme statistic } | H_0 \text{ is true})$$

- From the previous slide

$$P(\text{ extreme statistic } | H_0 \text{ is true}) \neq P(H_0 \text{ is true} | \text{ extreme statistic })$$

- We cannot say that the p-value is the probability that the null hypothesis is true.

Bayes Rule and P-Values

- Recall: The p-value for a hypothesis test is the probability of obtaining a test statistic more extreme than the one observed, if the null hypothesis were true.
- In the language of conditional probability:

$$p\text{-value} = P(\text{ extreme statistic } | H_0 \text{ is true})$$

- From the previous slide

$$P(\text{ extreme statistic } | H_0 \text{ is true}) \neq P(H_0 \text{ is true} | \text{ extreme statistic })$$

- We cannot say that the p-value is the probability that the null hypothesis is true.
- Instead, using Bayes Rule

$$P(H_0 \text{ is true} | \text{ extreme statistic }) = \frac{P(\text{ extreme statistic } | H_0 \text{ is true}) P(H_0 \text{ is true})}{P(\text{ extreme statistic })}$$
$$= p\text{-value} \cdot \frac{P(H_0 \text{ is true})}{P(\text{ extreme statistic })}$$

Bayes Rule and P-Values

- Recall: The p-value for a hypothesis test is the probability of obtaining a test statistic more extreme than the one observed, if the null hypothesis were true.
- In the language of conditional probability:

$$p\text{-value} = P(\text{ extreme statistic } | H_0 \text{ is true})$$

- From the previous slide

$$P(\text{ extreme statistic } | H_0 \text{ is true}) \neq P(H_0 \text{ is true} | \text{ extreme statistic })$$

- We cannot say that the p-value is the probability that the null hypothesis is true.
- Instead, using Bayes Rule

$$\begin{aligned} P(H_0 \text{ is true} | \text{ extreme statistic }) &= \frac{P(\text{ extreme statistic } | H_0 \text{ is true}) P(H_0 \text{ is true})}{P(\text{ extreme statistic })} \\ &= p\text{-value} \cdot \frac{P(H_0 \text{ is true})}{P(\text{ extreme statistic })} \end{aligned}$$