Probability

Nate Wells

Math 141, 3/30/22

Outline

In this lecture, we will. . .

Outline

In this lecture, we will...

- Discuss the Law of Total Probability and Bayes' Rule
- Define and investigate Random Variables

Section 1

Conditional Probability

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

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

Theorem (General Multiplication Rule)

For any events A and B,

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

Conditional Probability

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

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

Theorem (General Multiplication Rule)

For any events A and B,

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

- Suppose we draw two cards (without replacement) from a standard 52 card deck. What is the probability that both cards are aces?

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

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

- In the previous example, are the events that "1st card is an Ace" and "2nd card is an Ace" independent?

Multiplication Rule for Independent Events

- The general multiplication rule simplifies considerably in the case when two events are independent:

Theorem (Independent Multiplication Rule)

If events A and B are independent, the

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

Multiplication Rule for Independent Events

- The general multiplication rule simplifies considerably 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 coin and roll a 6-sided die at the same time. Let A be the event that the coin flips heads and B be the event that the die rolls a 1.
    - Since the coin presumably gives us no information about the die, we say A and B are independent.

Multiplication Rule for Independent Events

- The general multiplication rule simplifies considerably 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 coin and roll a 6-sided die at the same time. Let A be the event that the coin flips heads and B be the event that the die rolls a 1.
  - Since the coin presumably gives us no information about the die, we say A and B are independent.
  - What is the probability that A and B both occur?

The Law of Total Probability

One useful trick for computing probabilities is the following:

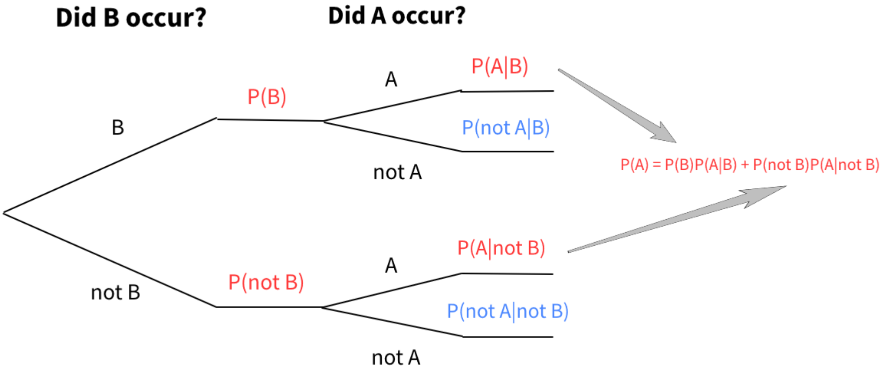
Theorem (The Law of Total Probability)

Let A and B be events. Then

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

- We can often represent the Law of Total Probability using a Tree Diagram:

Tree Diagrams



Lost Marbles

Two boxes contain a different number of red and blue marbles. The first box contains 20% red marbles while the second contains 80% red marbles. Suppose we select a marble from box 1 25% of the time and a marble from box 2 75% of the time. **What is the probability that a red marble is selected?**

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 $P(A \text{ and } B) = \frac{1}{2}$.
  - Then

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

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 $P(A \text{ and } B) = \frac{1}{2}$.
  - Then

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

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

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

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

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?
- Under what circumstances will $P(A|B) = P(B|A)$?
- Under what circumstances will $P(A|B)$ be much larger than $P(B|A)$? Much smaller?
- Suppose $P(B|A) = 1$.
  - What does this suggest about A and 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?
- Under what circumstances will $P(A|B) = P(B|A)$?
- Under what circumstances will $P(A|B)$ be much larger than $P(B|A)$? Much smaller?
- Suppose $P(B|A) = 1$.
  - What does this suggest about A and B?
  - What is $P(A|B)$ in this case?

Testing for a Rare Disease

Suppose that we have a rapid COVID-19 test that correctly diagnoses a person who **does not** have COVID 99% of the time, and correctly diagnoses a patient who **does** have COVID 80% of the time. Suppose a person takes this test and receives a positive diagnosis. Assume that the overall prevalence of COVID at the time of the test was 0.5%. **What is the probability that the person has COVID**?

Section 2

Random Variables

Definitions

A **random variable** is a numeric quantity whose value depends on the result of a random process.

Definitions

A **random variable** is a numeric quantity whose value depends on the result of a random process.

- We use capital letters at the end of the alphabet (W, X, Y, Z) to denote random variables.
    - We use lowercase letters (w, x, y, z) to denote the particular values of a random variable

Definitions

A **random variable** is a numeric quantity whose value depends on the result of a random process.

- We use capital letters at the end of the alphabet (W, X, Y, Z) to denote random variables.
  - We use lowercase letters (w, x, y, z) to denote the particular values of a random variable
- We use equations to express events associated to random variables.
  - I.e "$X = 5$" represents the event "The random variable X takes the value 5".

Definitions

A **random variable** is a numeric quantity whose value depends on the result of a random process.

- We use capital letters at the end of the alphabet (W, X, Y, Z) to denote random variables.
  - We use lowercase letters (w, x, y, z) to denote the particular values of a random variable
- We use equations to express events associated to random variables.
  - I.e "$X = 5$" represents the event "The random variable X takes the value 5".
- Events associated to variables have probabilities of occurring.
  - $P(X = 5) = .5$ means X has 50% probability of taking the value 5.

Types of Random Variables

There are two main types of random variables:

1. **Discrete** variables can take only finitely many different values.
2. **Continuous** variables can take values equal to any real number in an interval.

Types of Random Variables

There are two main types of random variables:

1. **Discrete** variables can take only finitely many different values.
2. **Continuous** variables can take values equal to any real number in an interval.

- Examples of discrete variables:
  - The number of credits a randomly chosen Reed student is taking.
  - The number of vegetarians in a random sample of 10 people.
  - The results of a coin flip, where 0 indicates Tails and 1 indicates Heads.

Types of Random Variables

There are two main types of random variables:

1. **Discrete** variables can take only finitely many different values.
2. **Continuous** variables can take values equal to any real number in an interval.

- Examples of discrete variables:
  - The number of credits a randomly chosen Reed student is taking.
  - The number of vegetarians in a random sample of 10 people.
  - The results of a coin flip, where 0 indicates Tails and 1 indicates Heads.
- Examples of continuous variables:
  - The temperature of my office at a particular time of the day.
  - The amount of time it takes a radioactive particle to decay.

Types of Random Variables

There are two main types of random variables:

1. **Discrete** variables can take only finitely many different values.
2. **Continuous** variables can take values equal to any real number in an interval.

- Examples of discrete variables:
  - The number of credits a randomly chosen Reed student is taking.
  - The number of vegetarians in a random sample of 10 people.
  - The results of a coin flip, where 0 indicates Tails and 1 indicates Heads.
- Examples of continuous variables:
  - The temperature of my office at a particular time of the day.
  - The amount of time it takes a radioactive particle to decay.
- Some discrete variables can be well-described by continuous variables:
  - The height of a random person selected from a large population.
  - The proportion of heads in a long sequence of coin flips.

The Distribution of a Random Variable

- Recall that *data* variables have distributions, which tell us...
    - the values the variable takes, and the *frequency* the variable takes those values.

The Distribution of a Random Variable

- Recall that *data* variables have distributions, which tell us. . .
  - the values the variable takes, and the *frequency* the variable takes those values.
- But *random* variables also have distributions, which tell us. . .
  - the values the variable can take, and the *probability* the variable takes those values.

The Distribution of a Random Variable

- Recall that *data* variables have distributions, which tell us...
  - the values the variable takes, and the *frequency* the variable takes those values.
- But *random* variables also have distributions, which tell us...
  - the values the variable can take, and the *probability* the variable takes those values.
- Suppose I play a casino game, where that the amount of money I win (in cents) has the following distribution:

value	1	5	10	25
probability	.3	.4	.2	.1

The Distribution of a Random Variable

- Recall that *data* variables have distributions, which tell us. . .
  - the values the variable takes, and the *frequency* the variable takes those values.
- But *random* variables also have distributions, which tell us. . .
  - the values the variable can take, and the *probability* the variable takes those values.
- Suppose I play a casino game, where that the amount of money I win (in cents) has the following distribution:

value	1	5	10	25
probability	.3	.4	.2	.1

- Suppose instead that I have a purse filled with the following 100 coins:

value	1	5	10	25
frequency	30	40	20	10

The Distribution of a Random Variable

- Recall that *data* variables have distributions, which tell us...
  - the values the variable takes, and the *frequency* the variable takes those values.
- But *random* variables also have distributions, which tell us...
  - the values the variable can take, and the *probability* the variable takes those values.
- Suppose I play a casino game, where that the amount of money I win (in cents) has the following distribution:

value	1	5	10	25
probability	.3	.4	.2	.1

- Suppose instead that I have a purse filled with the following 100 coins:

value	1	5	10	25
frequency	30	40	20	10

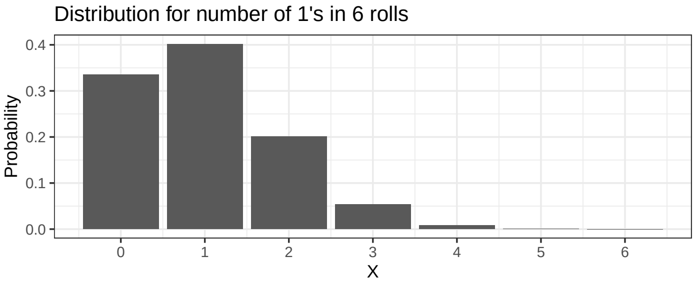
- Playing the casino game is very similar to drawing a random coin from the purse.

Visualizing Discrete Distributions

- We often use bar charts to visualize the distribution of discrete random variables.

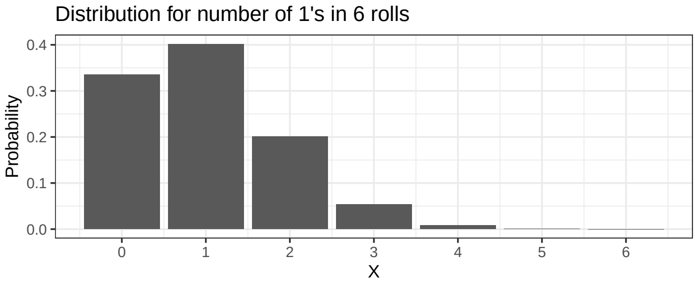
Visualizing Discrete Distributions

- We often use bar charts to visualize the distribution of discrete random variables.
  - Suppose a fair 6-sided die is rolled 6 times. Let X be the number of 1s rolled. The distribution of X is given by:



Visualizing Discrete Distributions

- We often use bar charts to visualize the distribution of discrete random variables.
  - Suppose a fair 6-sided die is rolled 6 times. Let X be the number of 1s rolled. The distribution of X is given by:



- Heights of bars are probabilities
  - This is analogous to rescaling a histogram to have heights equal to proportions, rather than counts

Expected Value

The **expected value** (or mean) of a discrete random variable X is

$$E[X] = x_1 P(X = x_1) + x_2 P(X = x_2) + \dots x_n P(X = x_n) = \sum_{i=1}^{n} x_i P(X = x_i)$$

Expected Value

The **expected value** (or mean) of a discrete random variable X is

$$E[X] = x_1 P(X = x_1) + x_2 P(X = x_2) + \ldots x_n P(X = x_n) = \sum_{i=1}^{n} x_i P(X = x_i)$$

- The expected value of X is the sum of the value X can take, weighted by the probability it takes those values.

Expected Value

The **expected value** (or mean) of a discrete random variable X is

$$E[X] = x_1 P(X = x_1) + x_2 P(X = x_2) + \dots x_n P(X = x_n) = \sum_{i=1}^{n} x_i P(X = x_i)$$

- The expected value of X is the sum of the value X can take, weighted by the probability it takes those values.
- Suppose we have a data set consisting of values $\{1, 1, 2, 2, 2, 2, 3, 4, 5, 5\}$. Let X be a value chosen from this data set randomly. What is the expected value of X?

Expected Value

The **expected value** (or mean) of a discrete random variable X is

$$E[X] = x_1 P(X = x_1) + x_2 P(X = x_2) + \dots x_n P(X = x_n) = \sum_{i=1}^{n} x_i P(X = x_i)$$

- The expected value of X is the sum of the value X can take, weighted by the probability it takes those values.
- Suppose we have a data set consisting of values $\{1, 1, 2, 2, 2, 2, 3, 4, 5, 5\}$. Let X be a value chosen from this data set randomly. What is the expected value of X?

$$\begin{aligned} E[X] =& 1P(X = 1) + 2P(X = 2) + 3P(X = 3) + 4P(X = 4) + 5P(X = 5) \\ =& 1\frac{2}{10} + 2\frac{4}{10} + 3\frac{1}{10} + 4\frac{1}{10} + 5\frac{2}{10} = \frac{27}{10} = 2.7 \end{aligned}$$

Expected Value

The **expected value** (or mean) of a discrete random variable X is

$$E[X] = x_1 P(X = x_1) + x_2 P(X = x_2) + \dots x_n P(X = x_n) = \sum_{i=1}^{n} x_i P(X = x_i)$$

- The expected value of X is the sum of the value X can take, weighted by the probability it takes those values.
- Suppose we have a data set consisting of values $\{1, 1, 2, 2, 2, 2, 3, 4, 5, 5\}$. Let X be a value chosen from this data set randomly. What is the expected value of X?

$$\begin{aligned} E[X] =& 1P(X = 1) + 2P(X = 2) + 3P(X = 3) + 4P(X = 4) + 5P(X = 5) \\ =& 1\frac{2}{10} + 2\frac{4}{10} + 3\frac{1}{10} + 4\frac{1}{10} + 5\frac{2}{10} = \frac{27}{10} = 2.7 \end{aligned}$$

- But also notice that

$$\begin{aligned} E[X] =& \frac{1}{10}\left(1 \cdot 2 + 2 \cdot 4 + 3 \cdot 1 + 4 \cdot 1 + 5 \cdot 2\right) \\ =& \frac{1}{10}\left(1 + 1 + 2 + 2 + 2 + 2 + 3 + 4 + 5 + 5\right) \end{aligned}$$

The Law of Large Numbers, again

Previously, we said that by the Law of Large numbers, the proportion of times an outcome occurs in a long sequence of trials is close to the probability for that outcome.

The Law of Large Numbers, again

Previously, we said that by the Law of Large numbers, the proportion of times an outcome occurs in a long sequence of trials is close to the probability for that outcome.

This is a generalization:

Theorem (The Law of Large Numbers)

Let X be a random variable whose value depends on a random experiment. Suppose the experiment is repeated n times and let $\bar{x}_n$ denote the arithmetic mean of the values of X in each trial. As n gets larger, the arithmetic mean $\bar{x}_n$ approaches the expected value $E[X]$ of that variable.

A Roll of the Die

Suppose we roll a fair 6-sided die. What is the expected value of the result?

A Roll of the Die

Suppose we roll a fair 6-sided die. What is the expected value of the result?

- Suppose we roll the same die 1000 times and keep track of the running arithmetic mean of the results. . .

A Roll of the Die

Suppose we roll a fair 6-sided die. What is the expected value of the result?

- Suppose we roll the same die 1000 times and keep track of the running arithmetic mean of the results. . .