Probability and Random Variables

Prof. Wells

STA 209, 4/10/23

Outline

In this lecture, we will. . .

Outline

In this lecture, we will...

- Investigate Bayes Rule for conditional probabilities
- Define and investigate random variables
- Compute the mean and standard deviation of 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,

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

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

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

- **Note!** Saying that two events are *independent* is not the same as saying they are *disjoint*.

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

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

- **Note!** Saying that two events are *independent* is not the same as saying they are *disjoint*.
  - 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

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

- **Note!** Saying that two events are *independent* is not the same as saying they are *disjoint*.
  - 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

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

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

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

- 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.

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?
  - 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:

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

  - 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:

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

- 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

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

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

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

- In general, the value of ratio $\frac{P(H_0 \text{ is true})}{P(\text{ extreme statistic })}$ is difficult to calculate, and requires significant prior knowledge about the data generation process.

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

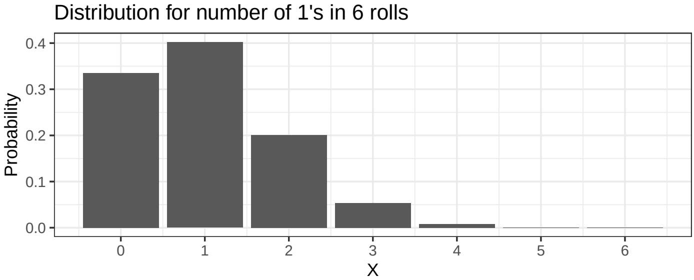
- Playing the casino game is very similar to drawing a random coin from the purse.

Visualizing Discrete Distributions

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

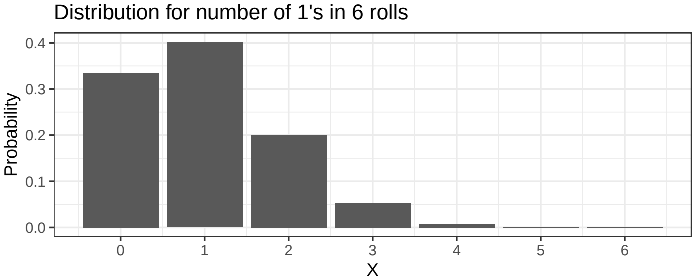
Visualizing Discrete Distributions

- 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

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

Expected Value

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

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

- 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

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

- 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

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

- 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

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

- 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

$$E[X] = \frac{1}{10}(2 \cdot 1 + 4 \cdot 2 + 1 \cdot 3 + 1 \cdot 4 + 2 \cdot 5) = \frac{1 + 1 + 2 + 2 + 2 + 2 + 3 + 4 + 5 + 5}{10}$$

Expected Value

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

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

- 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

$$E[X] = \frac{1}{10}(2 \cdot 1 + 4 \cdot 2 + 1 \cdot 3 + 1 \cdot 4 + 2 \cdot 5) = \frac{1 + 1 + 2 + 2 + 2 + 2 + 3 + 4 + 5 + 5}{10}$$

- The expected value of a random variable is the arithmetic mean of a data set, where each observation in the data occurs with frequency equal to its probability.

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

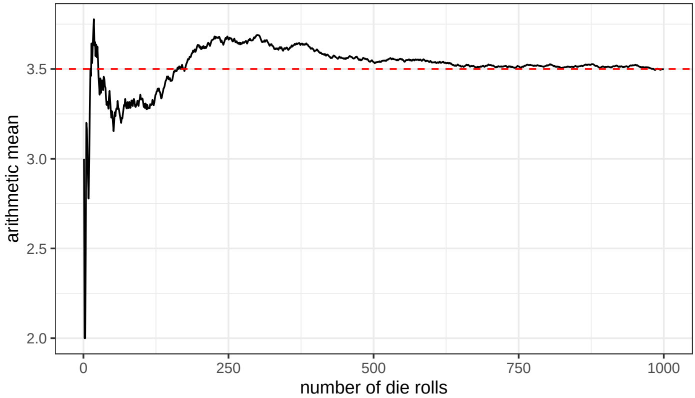
Suppose we roll a fair 6-sided die. What is the expected value of the result?

- 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...



Variance and Standard Deviation

The **variance** of a discrete random variable X with mean μ is

$$\begin{aligned}\mathrm{Var}(X) =&(x_1 - \mu)^2 P(X = x_1) + (x_2 - \mu)^2 P(X = x_2) + \cdots + (x_n - \mu)^2 P(X = x_n) \\ =&\sum_{i=1}^{n} (x_i - \mu)^2 P(X = x_i)\end{aligned}$$

Variance and Standard Deviation

The **variance** of a discrete random variable X with mean μ is

$$\begin{aligned}\mathrm{Var}(X) =&(x_1-\mu)^2 P(X=x_1) + (x_2-\mu)^2 P(X=x_2) + \cdots + (x_n-\mu)^2 P(X=x_n)\\ =&\sum_{i=1}^{n}(x_i-\mu)^2 P(X=x_i)\end{aligned}$$

- The variance of X is the sum the squared deviations of X from its mean μ, weighted by the corresponding probabilities.

Variance and Standard Deviation

The **variance** of a discrete random variable X with mean μ is

$$\begin{aligned}\mathrm{Var}(X) =& (x_1-\mu)^2 P(X=x_1) + (x_2-\mu)^2 P(X=x_2) + \cdots + (x_n-\mu)^2 P(X=x_n) \\ =& \sum_{i=1}^{n} (x_i-\mu)^2 P(X=x_i)\end{aligned}$$

- The variance of X is the sum the squared deviations of X from its mean μ, weighted by the corresponding probabilities.
    - Variables with low variance tend have values close to the mean, while those with high variance tend to have values farther from the mean.

Variance and Standard Deviation

The **variance** of a discrete random variable X with mean μ is

$$\begin{aligned}\mathrm{Var}(X) =&(x_1-\mu)^2 P(X=x_1) + (x_2-\mu)^2 P(X=x_2) + \cdots + (x_n-\mu)^2 P(X=x_n) \\ =&\sum_{i=1}^{n}(x_i-\mu)^2 P(X=x_i)\end{aligned}$$

- The variance of X is the sum the squared deviations of X from its mean μ, weighted by the corresponding probabilities.
  - Variables with low variance tend have values close to the mean, while those with high variance tend to have values farther from the mean.
- As with data variables, we define the **standard deviation** of a random variable X to be

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$$

Variance and Standard Deviation

The **variance** of a discrete random variable X with mean μ is

$$\begin{aligned}\mathrm{Var}(X) =&(x_1 - \mu)^2 P(X = x_1) + (x_2 - \mu)^2 P(X = x_2) + \cdots + (x_n - \mu)^2 P(X = x_n) \\ =&\sum_{i=1}^{n} (x_i - \mu)^2 P(X = x_i)\end{aligned}$$

- The variance of X is the sum the squared deviations of X from its mean μ, weighted by the corresponding probabilities.
  - Variables with low variance tend have values close to the mean, while those with high variance tend to have values farther from the mean.
- As with data variables, we define the **standard deviation** of a random variable X to be

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$$

- Often, σ^2 denotes the variance and σ to denotes the standard deviation of a variable.

Variance and Standard Deviation

The **variance** of a discrete random variable X with mean μ is

$$\begin{aligned}\mathrm{Var}(X) =& (x_1-\mu)^2 P(X=x_1) + (x_2-\mu)^2 P(X=x_2) + \cdots + (x_n-\mu)^2 P(X=x_n) \\ =& \sum_{i=1}^{n} (x_i-\mu)^2 P(X=x_i)\end{aligned}$$

- The variance of X is the sum the squared deviations of X from its mean μ, weighted by the corresponding probabilities.
  - Variables with low variance tend have values close to the mean, while those with high variance tend to have values farther from the mean.
- As with data variables, we define the **standard deviation** of a random variable X to be

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$$

- Often, σ^2 denotes the variance and σ to denotes the standard deviation of a variable.
- Show that the standard deviation of a variable X which takes value 1 with probability p and 0 with probability $1-p$ is

$$\mathrm{SD}(X) = \sqrt{p(1-p)}$$