

Math Camp

Justin Grimmer

Professor
Department of Political Science
Stanford University

September 17th, 2018

Where we're at

- Conditional Probability/Bayes' Rule

Where we're at

- Conditional Probability/Bayes' Rule
- Today: Random Variables

Where we're at

- Conditional Probability/Bayes' Rule
- Today: Random Variables
- Probability Mass Functions

Where we're at

- Conditional Probability/Bayes' Rule
- Today: Random Variables
- Probability Mass Functions
- Expectation, Variance

Where we're at

- Conditional Probability/Bayes' Rule
- Today: Random Variables
- Probability Mass Functions
- Expectation, Variance
- Famous Discrete Random Variables

Where we're at

- Conditional Probability/Bayes' Rule
- Today: Random Variables
- Probability Mass Functions
- Expectation, Variance
- Famous Discrete Random Variables
- A Brief Introduction to Markov Chains

Random Variable: Intuition

Recall the three parts of our probability model

Random Variable: Intuition

Recall the three parts of our probability model

- Sample Space

Random Variable: Intuition

Recall the three parts of our probability model

- Sample Space
- Events

Random Variable: Intuition

Recall the three parts of our probability model

- Sample Space
- Events
- Probability

Random Variable: Intuition

Recall the three parts of our probability model

- Sample Space
- Events
- Probability

Often, we are interested in some function of the sample space

Random Variable: Intuition

Recall the three parts of our probability model

- Sample Space
- Events
- Probability

Often, we are interested in some function of the sample space

- Number of incumbents who win

Random Variable: Intuition

Recall the three parts of our probability model

- Sample Space
- Events
- Probability

Often, we are interested in some function of the sample space

- Number of incumbents who win
- An indicator whether a country defaults on loans (1 if Default, 0 otherwise)

Random Variable: Intuition

Recall the three parts of our probability model

- Sample Space
- Events
- Probability

Often, we are interested in some function of the sample space

- Number of incumbents who win
- An indicator whether a country defaults on loans (1 if Default, 0 otherwise)
- Number of casualties in a war (rather than all outcomes of casualties)

Random Variable: Intuition

Recall the three parts of our probability model

- Sample Space
- Events
- Probability

Often, we are interested in some function of the sample space

- Number of incumbents who win
- An indicator whether a country defaults on loans (1 if Default, 0 otherwise)
- Number of casualties in a war (rather than all outcomes of casualties)

Random variables: functions defined on the **sample space**

Definition: Random Variable

Definition

Definition: Random Variable

Definition

*Random Variable: A Random variable X is a function from the sample space to **real numbers**. In notation,*

Definition: Random Variable

Definition

*Random Variable: A Random variable X is a function from the sample space to **real numbers**. In notation,*

$$X : \text{Sample Space} \rightarrow \mathcal{R}$$

Definition: Random Variable

Definition

*Random Variable: A Random variable X is a function from the sample space to **real numbers**. In notation,*

$$X : \text{Sample Space} \rightarrow \mathcal{R}$$

- X 's **domain** are all outcomes (Sample Space)

Definition: Random Variable

Definition

*Random Variable: A Random variable X is a function from the sample space to **real numbers**. In notation,*

$$X : \text{Sample Space} \rightarrow \mathcal{R}$$

- X 's **domain** are all outcomes (Sample Space)
- X 's **range** is the Real line (or some subset of it)

Definition: Random Variable

Definition

*Random Variable: A Random variable X is a function from the sample space to **real numbers**. In notation,*

$$X : \text{Sample Space} \rightarrow \mathcal{R}$$

- X 's **domain** are all outcomes (Sample Space)
- X 's **range** is the Real line (or some subset of it)
- Because X is defined on outcomes, makes sense to write $p(X)$ (we'll talk about this soon)

Example

Example

Treatment assignment:

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control
- X = Number of units received treatment

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control
- X = Number of units received treatment

Defining the function:

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control
- X = Number of units received treatment

Defining the function:

$$X = \left\{ \begin{array}{l} 0 \\ 1 \\ 2 \\ 3 \end{array} \right.$$

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control
- X = Number of units received treatment

Defining the function:

$$X = \begin{cases} 0 & \text{if } (C, C, C) \\ \vdots & \end{cases}$$

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control
- X = Number of units received treatment

Defining the function:

$$X = \begin{cases} 0 & \text{if } (C, C, C) \\ 1 & \text{if } (T, C, C) \text{ or } (C, T, C) \text{ or } (C, C, T) \end{cases} .$$

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control
- X = Number of units received treatment

Defining the function:

$$X = \begin{cases} 0 & \text{if } (C, C, C) \\ 1 & \text{if } (T, C, C) \text{ or } (C, T, C) \text{ or } (C, C, T) \\ 2 & \text{if } (T, T, C) \text{ or } (T, C, T) \text{ or } (C, T, T) \end{cases} .$$

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control
- X = Number of units received treatment

Defining the function:

$$X = \begin{cases} 0 & \text{if } (C, C, C) \\ 1 & \text{if } (T, C, C) \text{ or } (C, T, C) \text{ or } (C, C, T) \\ 2 & \text{if } (T, T, C) \text{ or } (T, C, T) \text{ or } (C, T, T) \\ 3 & \text{if } (T, T, T) \end{cases} .$$

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control
- X = Number of units received treatment

Defining the function:

$$X = \begin{cases} 0 & \text{if } (C, C, C) \\ 1 & \text{if } (T, C, C) \text{ or } (C, T, C) \text{ or } (C, C, T) \\ 2 & \text{if } (T, T, C) \text{ or } (T, C, T) \text{ or } (C, T, T) \\ 3 & \text{if } (T, T, T) \end{cases} .$$

In other words,

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control
- X = Number of units received treatment

Defining the function:

$$X = \begin{cases} 0 & \text{if } (C, C, C) \\ 1 & \text{if } (T, C, C) \text{ or } (C, T, C) \text{ or } (C, C, T) \\ 2 & \text{if } (T, T, C) \text{ or } (T, C, T) \text{ or } (C, T, T) \\ 3 & \text{if } (T, T, T) \end{cases} .$$

In other words,

$$X((C, C, C)) = 0$$

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control
- X = Number of units received treatment

Defining the function:

$$X = \begin{cases} 0 & \text{if } (C, C, C) \\ 1 & \text{if } (T, C, C) \text{ or } (C, T, C) \text{ or } (C, C, T) \\ 2 & \text{if } (T, T, C) \text{ or } (T, C, T) \text{ or } (C, T, T) \\ 3 & \text{if } (T, T, T) \end{cases} .$$

In other words,

$$X((C, C, C)) = 0$$

$$X((T, C, C)) = 1$$

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control
- X = Number of units received treatment

Defining the function:

$$X = \begin{cases} 0 & \text{if } (C, C, C) \\ 1 & \text{if } (T, C, C) \text{ or } (C, T, C) \text{ or } (C, C, T) \\ 2 & \text{if } (T, T, C) \text{ or } (T, C, T) \text{ or } (C, T, T) \\ 3 & \text{if } (T, T, T) \end{cases} .$$

In other words,

$$X((C, C, C)) = 0$$

$$X((T, C, C)) = 1$$

$$X((T, C, T)) = 2$$

Example

Treatment assignment:

- Suppose we have 3 units, flipping fair coin ($\frac{1}{2}$) to assign each unit
- Assign to T = Treatment or C = control
- X = Number of units received treatment

Defining the function:

$$X = \begin{cases} 0 & \text{if } (C, C, C) \\ 1 & \text{if } (T, C, C) \text{ or } (C, T, C) \text{ or } (C, C, T) \\ 2 & \text{if } (T, T, C) \text{ or } (T, C, T) \text{ or } (C, T, T) \\ 3 & \text{if } (T, T, T) \end{cases} .$$

In other words,

$$X((C, C, C)) = 0$$

$$X((T, C, C)) = 1$$

$$X((T, C, T)) = 2$$

$$X((T, T, T)) = 3$$

Another Example

Another Example

X = Number of Calls into congressional office in some period p

Another Example

X = Number of Calls into congressional office in some period p

- $X(c) = c$

Another Example

X = Number of Calls into congressional office in some period p

- $X(c) = c$

Outcome of Election

Another Example

X = Number of Calls into congressional office in some period p

- $X(c) = c$

Outcome of Election

- Define v as the proportion of vote the candidate receives

Another Example

X = Number of Calls into congressional office in some period p

- $X(c) = c$

Outcome of Election

- Define v as the proportion of vote the candidate receives
- Define $X = 1$ if $v > 0.50$

Another Example

X = Number of Calls into congressional office in some period p

- $X(c) = c$

Outcome of Election

- Define v as the proportion of vote the candidate receives
- Define $X = 1$ if $v > 0.50$
- Define $X = 0$ if $v < 0.50$

Another Example

X = Number of Calls into congressional office in some period p

- $X(c) = c$

Outcome of Election

- Define v as the proportion of vote the candidate receives
- Define $X = 1$ if $v > 0.50$
- Define $X = 0$ if $v < 0.50$

For example, if $v = 0.48$, then $X(v) = 0$

Another Example

X = Number of Calls into congressional office in some period p

- $X(c) = c$

Outcome of Election

- Define v as the proportion of vote the candidate receives
- Define $X = 1$ if $v > 0.50$
- Define $X = 0$ if $v < 0.50$

For example, if $v = 0.48$, then $X(v) = 0$

Big Question: How do we compute $P(X=1)$, $P(X=0)$, etc?

Probability Mass Function: Intuition

Go back to our experiment example—probability comes from probability of outcomes

Probability Mass Function: Intuition

Go back to our experiment example—probability comes from probability of outcomes

$$P(C, T, C) = P(C)P(T)P(C) = \frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{8}$$

Probability Mass Function: Intuition

Go back to our experiment example—probability comes from probability of outcomes

$$P(C, T, C) = P(C)P(T)P(C) = \frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{8}$$

That's true for all outcomes.

Probability Mass Function: Intuition

Go back to our experiment example—probability comes from probability of outcomes

$$P(C, T, C) = P(C)P(T)P(C) = \frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{8}$$

That's true for all outcomes.

$$p(X = 0) = P(C, C, C) = \frac{1}{8}$$

Probability Mass Function: Intuition

Go back to our experiment example—probability comes from probability of outcomes

$$P(C, T, C) = P(C)P(T)P(C) = \frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{8}$$

That's true for all outcomes.

$$p(X = 0) = P(C, C, C) = \frac{1}{8}$$

$$p(X = 1) = P(T, C, C) + P(C, T, C) + P(C, C, T) = \frac{3}{8}$$

Probability Mass Function: Intuition

Go back to our experiment example—probability comes from probability of outcomes

$$P(C, T, C) = P(C)P(T)P(C) = \frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{8}$$

That's true for all outcomes.

$$p(X = 0) = P(C, C, C) = \frac{1}{8}$$

$$p(X = 1) = P(T, C, C) + P(C, T, C) + P(C, C, T) = \frac{3}{8}$$

$$p(X = 2) = P(T, T, C) + P(T, C, T) + P(C, T, T) = \frac{3}{8}$$

Probability Mass Function: Intuition

Go back to our experiment example—probability comes from probability of outcomes

$$P(C, T, C) = P(C)P(T)P(C) = \frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{8}$$

That's true for all outcomes.

$$p(X = 0) = P(C, C, C) = \frac{1}{8}$$

$$p(X = 1) = P(T, C, C) + P(C, T, C) + P(C, C, T) = \frac{3}{8}$$

$$p(X = 2) = P(T, T, C) + P(T, C, T) + P(C, T, T) = \frac{3}{8}$$

$$p(X = 3) = P(T, T, T) = \frac{1}{8}$$

Probability Mass Function: Intuition

Go back to our experiment example—probability comes from probability of outcomes

$$P(C, T, C) = P(C)P(T)P(C) = \frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{8}$$

That's true for all outcomes.

$$p(X = 0) = P(C, C, C) = \frac{1}{8}$$

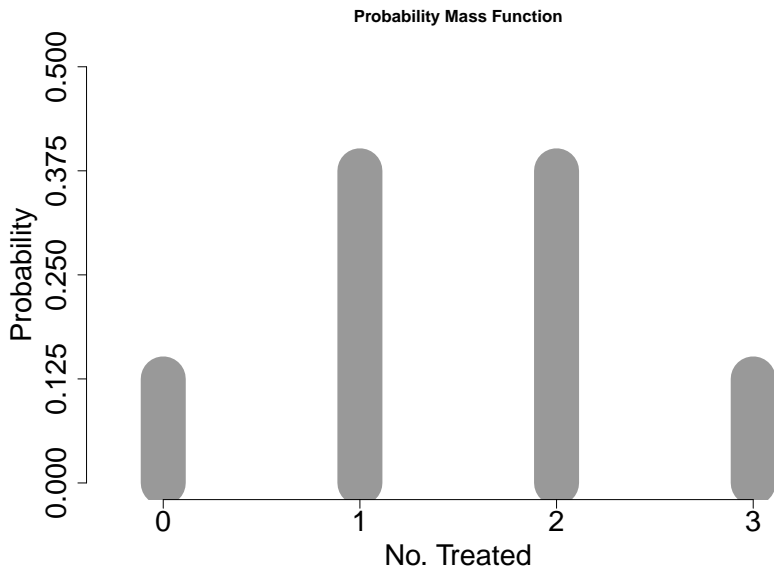
$$p(X = 1) = P(T, C, C) + P(C, T, C) + P(C, C, T) = \frac{3}{8}$$

$$p(X = 2) = P(T, T, C) + P(T, C, T) + P(C, T, T) = \frac{3}{8}$$

$$p(X = 3) = P(T, T, T) = \frac{1}{8}$$

$$p(X = a) = 0, \text{ for all } a \notin (0, 1, 2, 3)$$

Probability Mass Function: Intuition



Probability Mass Function: Intuition

Consider outcome of election:

- $X(v) = 1$ if $v > 0.5$ otherwise $X(v) = 0$
- $P(X = 1)$ then is equal to $P(v > 0.5)$

Probability Mass Function

If X is defined on an outcome space that is discrete (countable), we'll call it **discrete**.

Probability Mass Function

If X is defined on an outcome space that is discrete (countable), we'll call it **discrete**.

(Brief aside) Countable: A set is countable if there is a function that can map all its elements to the natural numbers $\{1, 2, 3, 4, \dots\}$ (one-to-one, injective). If it is onto (from S to all natural numbers, surjective), then we say the set is countably infinite

Probability Mass Function

If X is defined on an outcome space that is discrete (countable), we'll call it **discrete**.

Definition

*Probability Mass Function: For a **discrete** random variable X , define the probability mass function $p(x)$ as*

$$p(x) = P(X = x)$$

Probability Mass Function: Example 2

Probability Mass Function: Example 2

Topics:

Probability Mass Function: Example 2

Topics: distinct concepts (war in Afghanistan, national debt, fire department grants)

Probability Mass Function: Example 2

Topics: distinct concepts (war in Afghanistan, national debt, fire department grants)

Mathematically: Probability Mass Function on Words

Probability Mass Function: Example 2

Topics: distinct concepts (war in Afghanistan, national debt, **fire department grants**)

Mathematically: Probability Mass Function on Words Probability of using word, when discussing a topic

Probability Mass Function: Example 2

Topics: distinct concepts (war in Afghanistan, national debt, **fire department grants**)

Mathematically: Probability Mass Function on Words Probability of using word, when discussing a topic

Suppose we have a set of words:

Probability Mass Function: Example 2

Topics: distinct concepts (war in Afghanistan, national debt, **fire department grants**)

Mathematically: Probability Mass Function on Words Probability of using word, when discussing a topic

Suppose we have a set of words:

(afghanistan, fire, department, soldier, troop, war, grant)

Probability Mass Function: Example 2

Topics: distinct concepts (war in Afghanistan, national debt, **fire department grants**)

Mathematically: Probability Mass Function on Words Probability of using word, when discussing a topic

Suppose we have a set of words:

(afghanistan, fire, department, soldier, troop, war, grant)

Topic 1 (say, **war**):

Probability Mass Function: Example 2

Topics: distinct concepts (war in Afghanistan, national debt, **fire department grants**)

Mathematically: Probability Mass Function on Words Probability of using word, when discussing a topic

Suppose we have a set of words:

(afghanistan, fire, department, soldier, troop, war, grant)

Topic 1 (say, **war**):

$P(\text{afghanistan}) = 0.3$; $P(\text{fire}) = 0.0001$; $P(\text{department}) = 0.0001$;
 $P(\text{soldier}) = 0.2$; $P(\text{troop}) = 0.2$; $P(\text{war}) = 0.2997$; $P(\text{grant}) = 0.0001$

Probability Mass Function: Example 2

Topics: distinct concepts (war in Afghanistan, national debt, **fire department grants**)

Mathematically: Probability Mass Function on Words Probability of using word, when discussing a topic

Suppose we have a set of words:

(afghanistan, fire, department, soldier, troop, war, grant)

Topic 1 (say, **war**):

$P(\text{afghanistan}) = 0.3$; $P(\text{fire}) = 0.0001$; $P(\text{department}) = 0.0001$;

$P(\text{soldier}) = 0.2$; $P(\text{troop}) = 0.2$; $P(\text{war}) = 0.2997$; $P(\text{grant}) = 0.0001$

Topic 2 (say, **fire departments**):

Probability Mass Function: Example 2

Topics: distinct concepts (war in Afghanistan, national debt, **fire department grants**)

Mathematically: Probability Mass Function on Words Probability of using word, when discussing a topic

Suppose we have a set of words:

(afghanistan, fire, department, soldier, troop, war, grant)

Topic 1 (say, **war**):

$P(\text{afghanistan}) = 0.3$; $P(\text{fire}) = 0.0001$; $P(\text{department}) = 0.0001$;
 $P(\text{soldier}) = 0.2$; $P(\text{troop}) = 0.2$; $P(\text{war}) = 0.2997$; $P(\text{grant}) = 0.0001$

Topic 2 (say, **fire departments**):

$P(\text{afghanistan}) = 0.0001$; $P(\text{fire}) = 0.3$; $P(\text{department}) = 0.2$;
 $P(\text{soldier}) = 0.0001$; $P(\text{troop}) = 0.0001$; $P(\text{war}) = 0.0001$;
 $P(\text{grant}) = 0.2997$

Probability Mass Function: Example 2

Topics: distinct concepts (war in Afghanistan, national debt, **fire department grants**)

Mathematically: Probability Mass Function on Words Probability of using word, when discussing a topic

Suppose we have a set of words:

(afghanistan, fire, department, soldier, troop, war, grant)

Topic 1 (say, **war**):

$P(\text{afghanistan}) = 0.3$; $P(\text{fire}) = 0.0001$; $P(\text{department}) = 0.0001$;
 $P(\text{soldier}) = 0.2$; $P(\text{troop}) = 0.2$; $P(\text{war}) = 0.2997$; $P(\text{grant}) = 0.0001$

Topic 2 (say, **fire departments**):

$P(\text{afghanistan}) = 0.0001$; $P(\text{fire}) = 0.3$; $P(\text{department}) = 0.2$;
 $P(\text{soldier}) = 0.0001$; $P(\text{troop}) = 0.0001$; $P(\text{war}) = 0.0001$;
 $P(\text{grant}) = 0.2997$

Topic Models: take a set of documents and estimate topics.

Definition

Cumulative Mass (distribution) Function: For a random variable X , define the cumulative mass function $F(x)$ as,

$$F(x) = P(X \leq x)$$

- Characterizes how probability **cumulates** as X gets larger
- $F(x) \in [0, 1]$
- $F(x)$ is **non-decreasing**

Cumulative Mass Function: Example

Consider the three person experiment.

Cumulative Mass Function: Example

Consider the three person experiment. $P(T) = P(C) = 1/2$.

Cumulative Mass Function: Example

Consider the three person experiment. $P(T) = P(C) = 1/2$.
What is $F(2)$?

Cumulative Mass Function: Example

Consider the three person experiment. $P(T) = P(C) = 1/2$.
What is $F(2)$?

$$F(2) = P(X = 0) + P(X = 1) + P(X = 2)$$

Cumulative Mass Function: Example

Consider the three person experiment. $P(T) = P(C) = 1/2$.
What is $F(2)$?

$$\begin{aligned} F(2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} \end{aligned}$$

Cumulative Mass Function: Example

Consider the three person experiment. $P(T) = P(C) = 1/2$.
What is $F(2)$?

$$\begin{aligned} F(2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} \\ &= \frac{7}{8} \end{aligned}$$

Cumulative Mass Function: Example

Consider the three person experiment. $P(T) = P(C) = 1/2$.
What is $F(2)$?

$$\begin{aligned} F(2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} \\ &= \frac{7}{8} \end{aligned}$$

What is $F(2) - F(1)$?

Cumulative Mass Function: Example

Consider the three person experiment. $P(T) = P(C) = 1/2$.
What is $F(2)$?

$$\begin{aligned} F(2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} \\ &= \frac{7}{8} \end{aligned}$$

What is $F(2) - F(1)$?

$$\begin{aligned} F(2) - F(1) &= [P(X = 0) + P(X = 1) + P(X = 2)] \\ &\quad - [P(X = 0) + P(X = 1)] \end{aligned}$$

Cumulative Mass Function: Example

Consider the three person experiment. $P(T) = P(C) = 1/2$.
What is $F(2)$?

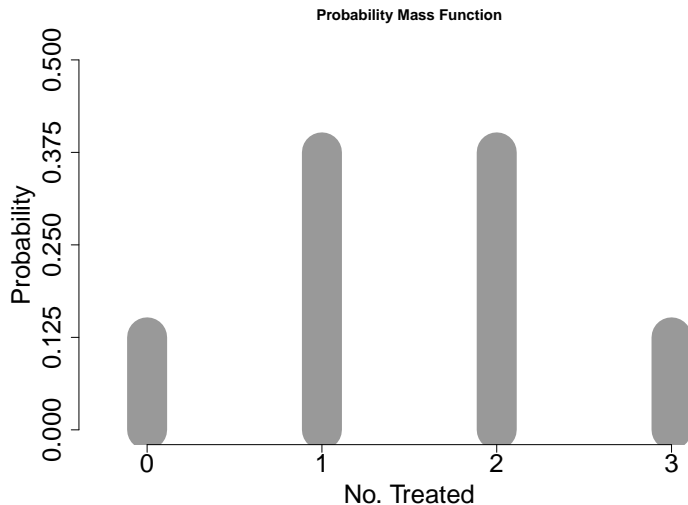
$$\begin{aligned} F(2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} \\ &= \frac{7}{8} \end{aligned}$$

What is $F(2) - F(1)$?

$$\begin{aligned} F(2) - F(1) &= [P(X = 0) + P(X = 1) + P(X = 2)] \\ &\quad - [P(X = 0) + P(X = 1)] \\ F(2) - F(1) &= P(X = 2) \end{aligned}$$

Cumulative Mass Function

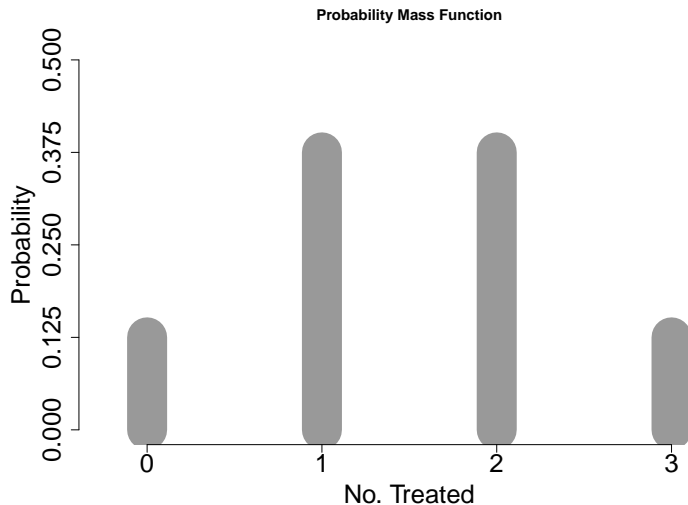
There is a close relationship between pmf's and cmf's.



Cumulative Mass Function

There is a close relationship between pmf's and cmf's.

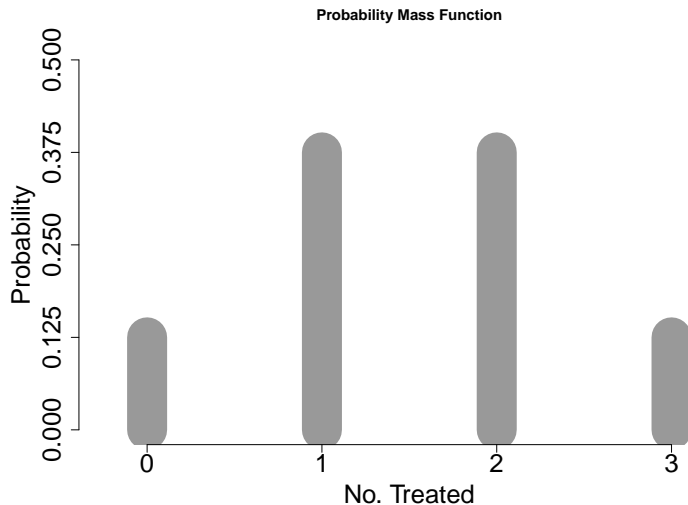
Consider Previous example:



Cumulative Mass Function

There is a close relationship between pmf's and cmf's.

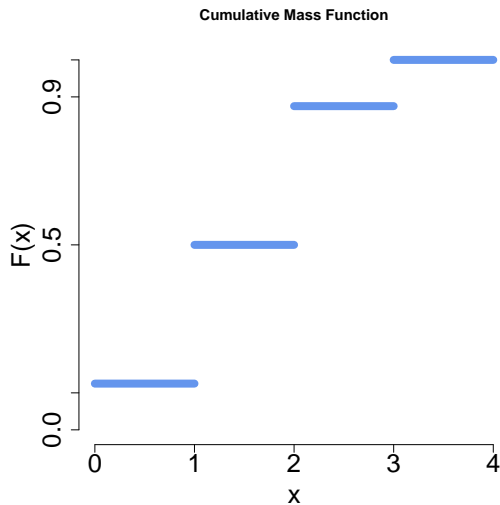
Consider Previous example:



Cumulative Mass Function

There is a close relationship between pmf's and cmf's.

Consider Previous example:

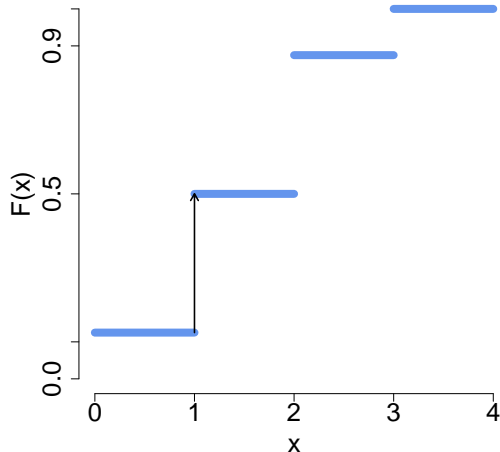


Cumulative Mass Function

There is a close relationship between pmf's and cmf's.

Consider Previous example:

Cumulative Mass Function

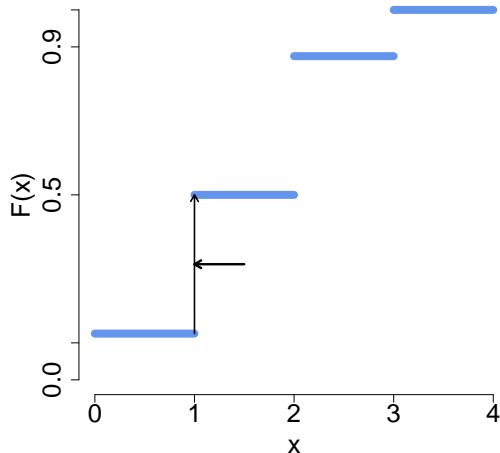


Cumulative Mass Function

There is a close relationship between pmf's and cmf's.

Consider Previous example:

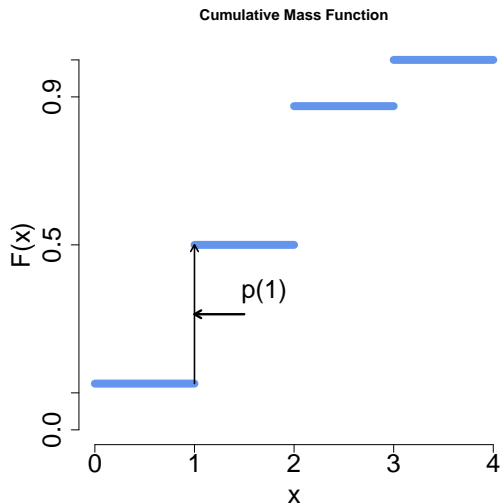
Cumulative Mass Function



Cumulative Mass Function

There is a close relationship between pmf's and cmf's.

Consider Previous example:



Expectation

What can we **expect** from a trial?

Expectation

What can we **expect** from a trial?

Value of random variable for any outcome

Expectation

What can we **expect** from a trial?

Value of random variable for any outcome

Weighted by the probability of observing that outcome

Expectation

What can we **expect** from a trial?

Value of random variable for any outcome

Weighted by the probability of observing that outcome

Definition

Expected Value: define the expected value of a function X as,

$$E[X] = \sum_{x:p(x)>0} xp(x)$$

Expectation

What can we **expect** from a trial?

Value of random variable for any outcome

Weighted by the probability of observing that outcome

Definition

Expected Value: define the expected value of a function X as,

$$E[X] = \sum_{x:p(x)>0} xp(x)$$

In words: for all values of x with $p(x)$ greater than zero, take the weighted average of the values

Expectation Example: Simple Experiment

Suppose again X is number of units assigned to treatment, in one of our previous example.

Expectation Example: Simple Experiment

Suppose again X is number of units assigned to treatment, in one of our previous example.

What is $E[X]$?

Expectation Example: Simple Experiment

Suppose again X is number of units assigned to treatment, in one of our previous example.

What is $E[X]$?

$$E[X]$$

Expectation Example: Simple Experiment

Suppose again X is number of units assigned to treatment, in one of our previous example.

What is $E[X]$?

$$E[X] = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8}$$

Expectation Example: Simple Experiment

Suppose again X is number of units assigned to treatment, in one of our previous example.

What is $E[X]$?

$$\begin{aligned} E[X] &= 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} \\ &= 1.5 \end{aligned}$$

Expectation Example: A Single Person Poll

Suppose that there is a group of N people.

Expectation Example: A Single Person Poll

Suppose that there is a group of N people.

- Suppose $M < N$ people approve of Barack Obama's performance as president

Expectation Example: A Single Person Poll

Suppose that there is a group of N people.

- Suppose $M < N$ people approve of Barack Obama's performance as president
- $N - M$ disapprove of his performance

Expectation Example: A Single Person Poll

Suppose that there is a group of N people.

- Suppose $M < N$ people approve of Barack Obama's performance as president
- $N - M$ disapprove of his performance

Define:

Expectation Example: A Single Person Poll

Suppose that there is a group of N people.

- Suppose $M < N$ people approve of Barack Obama's performance as president
- $N - M$ disapprove of his performance

Define:

Expectation Example: A Single Person Poll

Suppose that there is a group of N people.

- Suppose $M < N$ people approve of Barack Obama's performance as president
- $N - M$ disapprove of his performance

Define:

Draw one person i , with , $P(\text{Draw } i) = \frac{1}{N}$

Expectation Example: A Single Person Poll

Suppose that there is a group of N people.

- Suppose $M < N$ people approve of Barack Obama's performance as president
- $N - M$ disapprove of his performance

Define:

Draw one person i , with , $P(\text{Draw } i) = \frac{1}{N}$

$$X = \begin{cases} 1 & \text{if person Approves} \\ 0 & \text{if Disapproves} \end{cases} .$$

Expectation Example: A Single Person Poll

Suppose that there is a group of N people.

- Suppose $M < N$ people approve of Barack Obama's performance as president
- $N - M$ disapprove of his performance

Define:

Draw one person i , with , $P(\text{Draw } i) = \frac{1}{N}$

$$X = \begin{cases} 1 & \text{if person Approves} \\ 0 & \text{if Disapproves} \end{cases} .$$

$E[X]$?

Expectation Example: A Single Person Poll

Suppose that there is a group of N people.

- Suppose $M < N$ people approve of Barack Obama's performance as president
- $N - M$ disapprove of his performance

Define:

Draw one person i , with , $P(\text{Draw } i) = \frac{1}{N}$

$$X = \begin{cases} 1 & \text{if person Approves} \\ 0 & \text{if Disapproves} \end{cases} .$$

$E[X]$?

$$E[X] = 1 \times P(\text{Approve}) + 0 \times P(\text{Disapprove})$$

Expectation Example: A Single Person Poll

Suppose that there is a group of N people.

- Suppose $M < N$ people approve of Barack Obama's performance as president
- $N - M$ disapprove of his performance

Define:

Draw one person i , with , $P(\text{Draw } i) = \frac{1}{N}$

$$X = \begin{cases} 1 & \text{if person Approves} \\ 0 & \text{if Disapproves} \end{cases} .$$

$E[X]$?

$$\begin{aligned} E[X] &= 1 \times P(\text{Approve}) + 0 \times P(\text{Disapprove}) \\ &= 1 \times \frac{M}{N} \end{aligned}$$

Indicator Variables and Probabilities



Indicator Variables and Probabilities

Proposition



Indicator Variables and Probabilities

Proposition

Suppose A is an event. Define random variable I such that $I = 1$ if an outcome in A occurs and $I = 0$ if an outcome in A^c occurs. Then,



Indicator Variables and Probabilities

Proposition

Suppose A is an event. Define random variable I such that $I = 1$ if an outcome in A occurs and $I = 0$ if an outcome in A^c occurs. Then,

$$E[I] = P(A)$$



Indicator Variables and Probabilities

Proposition

Suppose A is an event. Define random variable I such that $I = 1$ if an outcome in A occurs and $I = 0$ if an outcome in A^c occurs. Then,

$$E[I] = P(A)$$

Proof.



Indicator Variables and Probabilities

Proposition

Suppose A is an event. Define random variable I such that $I = 1$ if an outcome in A occurs and $I = 0$ if an outcome in A^c occurs. Then,

$$E[I] = P(A)$$

Proof.

$$E[I] = 1 \times P(A) + 0 \times P(A^c)$$



Indicator Variables and Probabilities

Proposition

Suppose A is an event. Define random variable I such that $I = 1$ if an outcome in A occurs and $I = 0$ if an outcome in A^c occurs. Then,

$$E[I] = P(A)$$

Proof.

$$\begin{aligned} E[I] &= 1 \times P(A) + 0 \times P(A^c) \\ &= P(A) \end{aligned}$$



Functions of Random Variables

We might (or often) apply a function to a random variable $g(X)$.

Functions of Random Variables

We might (or often) apply a function to a random variable $g(X)$.
How do we compute $E[g(X)]$?

Functions of Random Variables

We might (or often) apply a function to a random variable $g(X)$.
How do we compute $E[g(X)]$?

Proposition

Functions of Random Variables

We might (or often) apply a function to a random variable $g(X)$.
How do we compute $E[g(X)]$?

Proposition

Expected value of a function of a random variable: Suppose X is a discrete random variable that takes on values x_i , $i = \{1, 2, \dots\}$, with probabilities $p(x_i)$.

Functions of Random Variables

We might (or often) apply a function to a random variable $g(X)$.
How do we compute $E[g(X)]$?

Proposition

Expected value of a function of a random variable: Suppose X is a discrete random variable that takes on values x_i , $i = \{1, 2, \dots\}$, with probabilities $p(x_i)$. If $g : X \rightarrow \mathcal{R}$, then its expected value $E[g(X)]$ is,

Functions of Random Variables

We might (or often) apply a function to a random variable $g(X)$.
How do we compute $E[g(X)]$?

Proposition

Expected value of a function of a random variable: Suppose X is a discrete random variable that takes on values x_i , $i = \{1, 2, \dots\}$, with probabilities $p(x_i)$. If $g : X \rightarrow \mathcal{R}$, then its expected value $E[g(X)]$ is,

$$E[g(X)] = \sum_i g(x_i)p(x_i)$$

Functions of Random Variables

Proof.



Functions of Random Variables

Proof.

Observation $g(X)$ is itself a random variable. Let's say it has unique values y_j ($j = 1, 2, \dots$)



Functions of Random Variables

Proof.

Observation $g(X)$ is itself a random variable. Let's say it has unique values y_j ($j = 1, 2, \dots$). So, we know that $E[g(X)] = \sum_j y_j P(g(X) = y_j)$.



Functions of Random Variables

Proof.

Observation $g(X)$ is itself a random variable. Let's say it has unique values y_j ($j = 1, 2, \dots$). So, we know that $E[g(X)] = \sum_j y_j P(g(X) = y_j)$. And we want to show that $\sum_i g(x_i) p(x_i)$ is equal to that.



Functions of Random Variables

Proof.

Observation $g(X)$ is itself a random variable. Let's say it has unique values y_j ($j = 1, 2, \dots$). So, we know that $E[g(X)] = \sum_j y_j P(g(X) = y_j)$. And we want to show that $\sum_i g(x_i) p(x_i)$ is equal to that.

$$\sum_i g(x_i) p(x_i) = \sum_j \sum_{i: g(x_i) = y_j} g(x_i) p(x_i)$$



Functions of Random Variables

Proof.

Observation $g(X)$ is itself a random variable. Let's say it has unique values y_j ($j = 1, 2, \dots$). So, we know that $E[g(X)] = \sum_j y_j P(g(X) = y_j)$. And we want to show that $\sum_i g(x_i) p(x_i)$ is equal to that.

$$\begin{aligned} \sum_i g(x_i) p(x_i) &= \sum_j \sum_{i: g(x_i) = y_j} g(x_i) p(x_i) \\ &= \sum_j \sum_{i: g(x_i) = y_j} y_j p(x_i) \end{aligned}$$



Functions of Random Variables

Proof.

Observation $g(X)$ is itself a random variable. Let's say it has unique values y_j ($j = 1, 2, \dots$). So, we know that $E[g(X)] = \sum_j y_j P(g(X) = y_j)$. And we want to show that $\sum_i g(x_i)p(x_i)$ is equal to that.

$$\begin{aligned}\sum_i g(x_i)p(x_i) &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i)p(x_i) \\ &= \sum_j \sum_{i:g(x_i)=y_j} y_j p(x_i) \\ &= \sum_j y_j \sum_{i:g(x_i)=y_j} p(x_i)\end{aligned}$$



Functions of Random Variables

Proof.

Observation $g(X)$ is itself a random variable. Let's say it has unique values y_j ($j = 1, 2, \dots$). So, we know that $E[g(X)] = \sum_j y_j P(g(X) = y_j)$. And we want to show that $\sum_i g(x_i) p(x_i)$ is equal to that.

$$\begin{aligned}\sum_i g(x_i) p(x_i) &= \sum_j \sum_{i: g(x_i) = y_j} g(x_i) p(x_i) \\ &= \sum_j \sum_{i: g(x_i) = y_j} y_j p(x_i) \\ &= \sum_j y_j \sum_{i: g(x_i) = y_j} p(x_i) \\ &= \sum_j y_j P(g(X) = y_j)\end{aligned}$$



Functions of Random Variables

Proof.

Observation $g(X)$ is itself a random variable. Let's say it has unique values y_j ($j = 1, 2, \dots$). So, we know that $E[g(X)] = \sum_j y_j P(g(X) = y_j)$. And we want to show that $\sum_i g(x_i) p(x_i)$ is equal to that.

$$\begin{aligned}\sum_i g(x_i) p(x_i) &= \sum_j \sum_{i: g(x_i) = y_j} g(x_i) p(x_i) \\ &= \sum_j \sum_{i: g(x_i) = y_j} y_j p(x_i) \\ &= \sum_j y_j \sum_{i: g(x_i) = y_j} p(x_i) \\ &= \sum_j y_j P(g(X) = y_j) \\ &= E[g(X)]\end{aligned}$$



Functions of Random Variables: Example

Let's suppose that X is the number of observations assigned to treatment (from our previous example).

Functions of Random Variables: Example

Let's suppose that X is the number of observations assigned to treatment (from our previous example).

Suppose that $g(X) = X^2$. What is $E[g(X)]$?

Functions of Random Variables: Example

Let's suppose that X is the number of observations assigned to treatment (from our previous example).

Suppose that $g(X) = X^2$. What is $E[g(X)]$?

$$E[g(X)] = E[X^2] = 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8}$$

Functions of Random Variables: Example

Let's suppose that X is the number of observations assigned to treatment (from our previous example).

Suppose that $g(X) = X^2$. What is $E[g(X)]$?

$$\begin{aligned} E[g(X)] = E[X^2] &= 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8} \\ &= 0 + \frac{3}{8} + \frac{12}{8} + \frac{9}{8} \end{aligned}$$

Functions of Random Variables: Example

Let's suppose that X is the number of observations assigned to treatment (from our previous example).

Suppose that $g(X) = X^2$. What is $E[g(X)]$?

$$\begin{aligned} E[g(X)] = E[X^2] &= 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8} \\ &= 0 + \frac{3}{8} + \frac{12}{8} + \frac{9}{8} \\ &= \frac{24}{8} = 3 \end{aligned}$$

Functions of Random Variables: Corollary

Corollary

*Suppose X is a random variable and a and b are **constants** (not random variables). Then,*

$$E[aX + b] = aE[X] + b$$

Proof.

Functions of Random Variables: Corollary

Corollary

*Suppose X is a random variable and a and b are **constants** (not random variables). Then,*

$$E[aX + b] = aE[X] + b$$

Proof.

$$E[aX + b] = \sum_{x:p(x)>0} (ax + b)p(x)$$

Functions of Random Variables: Corollary

Corollary

Suppose X is a random variable and a and b are *constants* (not random variables). Then,

$$E[aX + b] = aE[X] + b$$

Proof.

$$\begin{aligned} E[aX + b] &= \sum_{x:p(x)>0} (ax + b)p(x) \\ &= \sum_{x:p(x)>0} axp(x) + \sum_{x:p(x)>0} bp(x) \end{aligned}$$

Functions of Random Variables: Corollary

Corollary

Suppose X is a random variable and a and b are **constants** (not random variables). Then,

$$E[aX + b] = aE[X] + b$$

Proof.

$$\begin{aligned} E[aX + b] &= \sum_{x:p(x)>0} (ax + b)p(x) \\ &= \sum_{x:p(x)>0} axp(x) + \sum_{x:p(x)>0} bp(x) \\ &= a \sum_{x:p(x)>0} xp(x) + b \sum_{x:p(x)>0} p(x) \end{aligned}$$

Functions of Random Variables: Corollary

Corollary

Suppose X is a random variable and a and b are **constants** (not random variables). Then,

$$E[aX + b] = aE[X] + b$$

Proof.

$$\begin{aligned} E[aX + b] &= \sum_{x:p(x)>0} (ax + b)p(x) \\ &= \sum_{x:p(x)>0} axp(x) + \sum_{x:p(x)>0} bp(x) \\ &= a \sum_{x:p(x)>0} xp(x) + b \sum_{x:p(x)>0} p(x) \\ &= aE[X] + b(1) \end{aligned}$$

Variance

Expected value is a measure of **central tendency**.

Variance

Expected value is a measure of **central tendency**.

What about spread?

Variance

Expected value is a measure of **central tendency**.

What about spread? **Variance**

Variance

Expected value is a measure of **central tendency**.

What about spread? **Variance**

- For each value, we might measure distance from center

Variance

Expected value is a measure of **central tendency**.

What about spread? **Variance**

- For each value, we might measure distance from center
 - Euclidean distance, squared $d(x, E[x])^2 = (x - E[x])^2$

Variance

Expected value is a measure of **central tendency**.

What about spread? **Variance**

- For each value, we might measure distance from center
 - Euclidean distance, squared $d(x, E[x])^2 = (x - E[x])^2$
- Then, we might take weighted average of these distances,

Variance

Expected value is a measure of **central tendency**.

What about spread? **Variance**

- For each value, we might measure distance from center
 - Euclidean distance, squared $d(x, E[x])^2 = (x - E[x])^2$
- Then, we might take weighted average of these distances,

$$E[(X - E[X])^2] = \sum_{x:p(x)>0} (x - E[X])^2 p(x)$$

Variance

Expected value is a measure of **central tendency**.

What about spread? **Variance**

- For each value, we might measure distance from center
 - Euclidean distance, squared $d(x, E[x])^2 = (x - E[x])^2$
- Then, we might take weighted average of these distances,

$$\begin{aligned} E[(X - E[X])^2] &= \sum_{x:p(x)>0} (x - E[X])^2 p(x) \\ &= \sum_{x:p(x)>0} (x^2 p(x)) - \end{aligned}$$

Variance

Expected value is a measure of **central tendency**.

What about spread? **Variance**

- For each value, we might measure distance from center
 - Euclidean distance, squared $d(x, E[x])^2 = (x - E[x])^2$
- Then, we might take weighted average of these distances,

$$\begin{aligned} E[(X - E[X])^2] &= \sum_{x:p(x)>0} (x - E[X])^2 p(x) \\ &= \sum_{x:p(x)>0} (x^2 p(x)) - \\ &\quad 2E[X] \sum_{x:p(x)>0} (x p(x)) + E[X]^2 \sum_{x:p(x)>0} p(x) \end{aligned}$$

Variance

Expected value is a measure of **central tendency**.

What about spread? **Variance**

- For each value, we might measure distance from center
 - Euclidean distance, squared $d(x, E[X])^2 = (x - E[X])^2$
- Then, we might take weighted average of these distances,

$$\begin{aligned} E[(X - E[X])^2] &= \sum_{x:p(x)>0} (x - E[X])^2 p(x) \\ &= \sum_{x:p(x)>0} (x^2 p(x)) - \\ &\quad 2E[X] \sum_{x:p(x)>0} (xp(x)) + E[X]^2 \sum_{x:p(x)>0} p(x) \\ &= E[X^2] - 2E[X]^2 + E[X]^2 \end{aligned}$$

Variance

Expected value is a measure of **central tendency**.

What about spread? **Variance**

- For each value, we might measure distance from center
 - Euclidean distance, squared $d(x, E[X])^2 = (x - E[X])^2$
- Then, we might take weighted average of these distances,

$$\begin{aligned} E[(X - E[X])^2] &= \sum_{x:p(x)>0} (x - E[X])^2 p(x) \\ &= \sum_{x:p(x)>0} (x^2 p(x)) - \\ &\quad 2E[X] \sum_{x:p(x)>0} (xp(x)) + E[X]^2 \sum_{x:p(x)>0} p(x) \\ &= E[X^2] - 2E[X]^2 + E[X]^2 \\ &= E[X^2] - E[X]^2 \end{aligned}$$

Variance

Expected value is a measure of **central tendency**.

What about spread? **Variance**

- For each value, we might measure distance from center
 - Euclidean distance, squared $d(x, E[X])^2 = (x - E[X])^2$
- Then, we might take weighted average of these distances,

$$\begin{aligned} E[(X - E[X])^2] &= \sum_{x:p(x)>0} (x - E[X])^2 p(x) \\ &= \sum_{x:p(x)>0} (x^2 p(x)) - \\ &\quad 2E[X] \sum_{x:p(x)>0} (x p(x)) + E[X]^2 \sum_{x:p(x)>0} p(x) \\ &= E[X^2] - 2E[X]^2 + E[X]^2 \\ &= E[X^2] - E[X]^2 \\ &= \text{Var}(X) \end{aligned}$$

Variance

Definition

The variance of a random variable X , $\text{var}(X)$, is

$$\begin{aligned}\text{var}(X) &= E[(X - E[X])^2] \\ &= E[X^2] - E[X]^2\end{aligned}$$

- We will define the standard deviation of X , $\text{sd}(X) = \sqrt{\text{var}(X)}$
- $\text{var}(X) \geq 0$.

Variance Calculation

Continue the three person experiment, with $P(T) = P(C) = 1/2$.

Variance Calculation

Continue the three person experiment, with $P(T) = P(C) = 1/2$.
What is $\text{Var}(X)$?

Variance Calculation

Continue the three person experiment, with $P(T) = P(C) = 1/2$.

What is $\text{Var}(X)$?

We have two components to our variance calculation:

Variance Calculation

Continue the three person experiment, with $P(T) = P(C) = 1/2$.
What is $\text{Var}(X)$?

We have two components to our variance calculation:

$$E[X^2] = 3$$

Variance Calculation

Continue the three person experiment, with $P(T) = P(C) = 1/2$.
What is $\text{Var}(X)$?

We have two components to our variance calculation:

$$E[X^2] = 3$$

$$E[X]^2 = 1.5^2 = 2.25$$

Variance Calculation

Continue the three person experiment, with $P(T) = P(C) = 1/2$.
What is $\text{Var}(X)$?

We have two components to our variance calculation:

$$E[X^2] = 3$$

$$E[X]^2 = 1.5^2 = 2.25$$

$$\text{Var}(X) = E[X^2] - E[X]^2$$

Variance Calculation

Continue the three person experiment, with $P(T) = P(C) = 1/2$.
What is $\text{Var}(X)$?

We have two components to our variance calculation:

$$\begin{aligned} E[X^2] &= 3 \\ E[X]^2 &= 1.5^2 = 2.25 \\ \text{Var}(X) &= E[X^2] - E[X]^2 \\ &= 3 - 2.25 = 0.75 \end{aligned}$$

Variance Corollary

Corollary

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

Proof.



Variance Corollary

Corollary

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

Proof.

Define $Y = aX + b$. Now, we know that



Variance Corollary

Corollary

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

Proof.

Define $Y = aX + b$. Now, we know that

$\text{Var}(Y) = E[(Y - E[Y])^2]$. Let's substitute and use our other corollary



Variance Corollary

Corollary

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

Proof.

Define $Y = aX + b$. Now, we know that

$\text{Var}(Y) = E[(Y - E[Y])^2]$. Let's substitute and use our other corollary

$$\text{Var}(Y) = E[(aX + b - aE[X] - b)^2]$$



Variance Corollary

Corollary

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

Proof.

Define $Y = aX + b$. Now, we know that

$\text{Var}(Y) = E[(Y - E[Y])^2]$. Let's substitute and use our other corollary

$$\begin{aligned}\text{Var}(Y) &= E[(aX + b - aE[X] - b)^2] \\ &= E[(a^2X^2 - 2a^2XE[X] + a^2E[X]^2)]\end{aligned}$$



Variance Corollary

Corollary

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

Proof.

Define $Y = aX + b$. Now, we know that

$\text{Var}(Y) = E[(Y - E[Y])^2]$. Let's substitute and use our other corollary

$$\begin{aligned}\text{Var}(Y) &= E[(aX + b - aE[X] - b)^2] \\ &= E[(a^2X^2 - 2a^2XE[X] + a^2E[X]^2)] \\ &= a^2E[X^2] - 2a^2E[X]^2 + a^2E[X]^2\end{aligned}$$



Variance Corollary

Corollary

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

Proof.

Define $Y = aX + b$. Now, we know that

$\text{Var}(Y) = E[(Y - E[Y])^2]$. Let's substitute and use our other corollary

$$\begin{aligned}\text{Var}(Y) &= E[(aX + b - aE[X] - b)^2] \\ &= E[(a^2X^2 - 2a^2XE[X] + a^2E[X]^2)] \\ &= a^2E[X^2] - 2a^2E[X]^2 + a^2E[X]^2 \\ &= a^2(E[X^2] - E[X]^2)\end{aligned}$$



Variance Corollary

Corollary

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

Proof.

Define $Y = aX + b$. Now, we know that

$\text{Var}(Y) = E[(Y - E[Y])^2]$. Let's substitute and use our other corollary

$$\begin{aligned}\text{Var}(Y) &= E[(aX + b - aE[X] - b)^2] \\ &= E[(a^2X^2 - 2a^2XE[X] + a^2E[X]^2)] \\ &= a^2E[X^2] - 2a^2E[X]^2 + a^2E[X]^2 \\ &= a^2(E[X^2] - E[X]^2) \\ &= a^2\text{Var}(X)\end{aligned}$$



Famous Distributions

- Bernoulli
- Binomial
- Multinomial
- Poisson

Models of how world works.

Bernoulli Random Variable

Definition

Suppose X is a random variable, with $X \in \{0, 1\}$ and $P(X = 1) = \pi$. Then we will say that X is **Bernoulli** random variable,

$$p(k) = \pi^k(1 - \pi)^{1-k}$$

for $k \in \{0, 1\}$ and $p(k) = 0$ otherwise.

We will (equivalently) say that

$$Y \sim \text{Bernoulli}(\pi)$$

Bernoulli Random Variable

Suppose we flip a fair coin and $Y = 1$ if the outcome is Heads .

$$\begin{aligned} Y &\sim \text{Bernoulli}(1/2) \\ p(1) &= (1/2)^1(1 - 1/2)^{1-1} = 1/2 \\ p(0) &= (1/2)^0(1 - 1/2)^{1-0} = (1 - 1/2) \end{aligned}$$

Bernoulli Random Variable Moments

Suppose $Y \sim \text{Bernoulli}(\pi)$

Bernoulli Random Variable Moments

Suppose $Y \sim \text{Bernoulli}(\pi)$

$$\begin{aligned} E[Y] &= 1 \times P(Y = 1) + 0 \times P(Y = 0) \\ &= \pi + 0(1 - \pi) = \pi \end{aligned}$$

$$E[Y] = \pi$$

Bernoulli Random Variable Moments

Suppose $Y \sim \text{Bernoulli}(\pi)$

$$\begin{aligned} E[Y] &= 1 \times P(Y = 1) + 0 \times P(Y = 0) \\ &= \pi + 0(1 - \pi) = \pi \\ \text{var}(Y) &= E[Y^2] - E[Y]^2 \end{aligned}$$

$$E[Y] = \pi$$

Bernoulli Random Variable Moments

Suppose $Y \sim \text{Bernoulli}(\pi)$

$$\begin{aligned}E[Y] &= 1 \times P(Y = 1) + 0 \times P(Y = 0) \\&= \pi + 0(1 - \pi) = \pi \\ \text{var}(Y) &= E[Y^2] - E[Y]^2 \\ E[Y^2] &= 1^2 P(Y = 1) + 0^2 P(Y = 0)\end{aligned}$$

$$E[Y] = \pi$$

Bernoulli Random Variable Moments

Suppose $Y \sim \text{Bernoulli}(\pi)$

$$\begin{aligned}E[Y] &= 1 \times P(Y = 1) + 0 \times P(Y = 0) \\&= \pi + 0(1 - \pi) = \pi \\ \text{var}(Y) &= E[Y^2] - E[Y]^2 \\ E[Y^2] &= 1^2 P(Y = 1) + 0^2 P(Y = 0) \\&= \pi\end{aligned}$$

$$E[Y] = \pi$$

Bernoulli Random Variable Moments

Suppose $Y \sim \text{Bernoulli}(\pi)$

$$\begin{aligned}E[Y] &= 1 \times P(Y = 1) + 0 \times P(Y = 0) \\&= \pi + 0(1 - \pi) = \pi \\ \text{var}(Y) &= E[Y^2] - E[Y]^2 \\ E[Y^2] &= 1^2 P(Y = 1) + 0^2 P(Y = 0) \\&= \pi \\ \text{var}(Y) &= \pi - \pi^2\end{aligned}$$

$$E[Y] = \pi$$

Bernoulli Random Variable Moments

Suppose $Y \sim \text{Bernoulli}(\pi)$

$$\begin{aligned}E[Y] &= 1 \times P(Y = 1) + 0 \times P(Y = 0) \\&= \pi + 0(1 - \pi) = \pi \\ \text{var}(Y) &= E[Y^2] - E[Y]^2 \\ E[Y^2] &= 1^2 P(Y = 1) + 0^2 P(Y = 0) \\&= \pi \\ \text{var}(Y) &= \pi - \pi^2 \\&= \pi(1 - \pi)\end{aligned}$$

$$\begin{aligned}E[Y] &= \pi \\ \text{var}(Y) &= \pi(1 - \pi)\end{aligned}$$

Bernoulli Random Variable Moments

Suppose $Y \sim \text{Bernoulli}(\pi)$

$$\begin{aligned}E[Y] &= 1 \times P(Y = 1) + 0 \times P(Y = 0) \\&= \pi + 0(1 - \pi) = \pi \\ \text{var}(Y) &= E[Y^2] - E[Y]^2 \\ E[Y^2] &= 1^2 P(Y = 1) + 0^2 P(Y = 0) \\&= \pi \\ \text{var}(Y) &= \pi - \pi^2 \\&= \pi(1 - \pi)\end{aligned}$$

$$E[Y] = \pi$$

$\text{var}(Y) = \pi(1 - \pi)$ What is the maximum variance?

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose.

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose.
Define $Y = 1$ if the country wins and $Y = 0$ otherwise.

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose.
Define $Y = 1$ if the country wins and $Y = 0$ otherwise.

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose.

Define $Y = 1$ if the country wins and $Y = 0$ otherwise.

Then,

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose.

Define $Y = 1$ if the country wins and $Y = 0$ otherwise.

Then,

$$Y \sim \text{Bernoulli}(\pi)$$

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose. Define $Y = 1$ if the country wins and $Y = 0$ otherwise.

Then,

$$Y \sim \text{Bernoulli}(\pi)$$

Suppose country 1 is deciding whether to fight a war.

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose. Define $Y = 1$ if the country wins and $Y = 0$ otherwise.

Then,

$$Y \sim \text{Bernoulli}(\pi)$$

Suppose country 1 is deciding whether to fight a war. Engaging in the war will cost the country c .

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose. Define $Y = 1$ if the country wins and $Y = 0$ otherwise.

Then,

$$Y \sim \text{Bernoulli}(\pi)$$

Suppose country 1 is deciding whether to fight a war.

Engaging in the war will cost the country c .

If they win, country 1 receives B .

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose. Define $Y = 1$ if the country wins and $Y = 0$ otherwise.

Then,

$$Y \sim \text{Bernoulli}(\pi)$$

Suppose country 1 is deciding whether to fight a war.

Engaging in the war will cost the country c .

If they win, country 1 receives B .

What is 1's expected utility from fighting a war?

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose. Define $Y = 1$ if the country wins and $Y = 0$ otherwise.

Then,

$$Y \sim \text{Bernoulli}(\pi)$$

Suppose country 1 is deciding whether to fight a war.

Engaging in the war will cost the country c .

If they win, country 1 receives B .

What is 1's expected utility from fighting a war?

$$E[U(\text{war})] = (\text{Utility}|\text{win}) \times P(\text{win}) + (\text{Utility}|\text{lose}) \times P(\text{lose})$$

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose. Define $Y = 1$ if the country wins and $Y = 0$ otherwise.

Then,

$$Y \sim \text{Bernoulli}(\pi)$$

Suppose country 1 is deciding whether to fight a war.

Engaging in the war will cost the country c .

If they win, country 1 receives B .

What is 1's expected utility from fighting a war?

$$\begin{aligned} E[U(\text{war})] &= (\text{Utility}|\text{win}) \times P(\text{win}) + (\text{Utility}|\text{lose}) \times P(\text{lose}) \\ &= (B - c)P(Y = 1) + (-c)P(Y = 0) \end{aligned}$$

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose. Define $Y = 1$ if the country wins and $Y = 0$ otherwise.

Then,

$$Y \sim \text{Bernoulli}(\pi)$$

Suppose country 1 is deciding whether to fight a war.

Engaging in the war will cost the country c .

If they win, country 1 receives B .

What is 1's expected utility from fighting a war?

$$\begin{aligned} E[U(\text{war})] &= (\text{Utility}|\text{win}) \times P(\text{win}) + (\text{Utility}|\text{lose}) \times P(\text{lose}) \\ &= (B - c)P(Y = 1) + (-c)P(Y = 0) \\ &= B \times p(Y = 1) - c(P(Y = 1) + P(Y = 0)) \end{aligned}$$

Example: Winning a War

Suppose country 1 is engaged in a conflict and can either win or lose. Define $Y = 1$ if the country wins and $Y = 0$ otherwise.

Then,

$$Y \sim \text{Bernoulli}(\pi)$$

Suppose country 1 is deciding whether to fight a war.

Engaging in the war will cost the country c .

If they win, country 1 receives B .

What is 1's expected utility from fighting a war?

$$\begin{aligned} E[U(\text{war})] &= (\text{Utility}|\text{win}) \times P(\text{win}) + (\text{Utility}|\text{lose}) \times P(\text{lose}) \\ &= (B - c)P(Y = 1) + (-c)P(Y = 0) \\ &= B \times p(Y = 1) - c(P(Y = 1) + P(Y = 0)) \\ &= B \times \pi - c \end{aligned}$$

Binomial Random Variable

- A model to count the number of successes across N trials

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent
 - Each Bernoulli trial i is

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent
 - Each Bernoulli trial i is

$$Y_i \sim \text{Bernoulli}(\pi)$$

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent
 - Each Bernoulli trial i is

$$Y_i \sim \text{Bernoulli}(\pi)$$

Independent and identically distributed.

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent
 - Each Bernoulli trial i is

$$Y_i \sim \text{Bernoulli}(\pi)$$

Independent and identically distributed.

- Z = number of successful trials

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent
 - Each Bernoulli trial i is

$$Y_i \sim \text{Bernoulli}(\pi)$$

Independent and identically distributed.

- Z = number of successful trials
- Derive probability mass function $P(Z = M) = p(M)$

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent
 - Each Bernoulli trial i is

$$Y_i \sim \text{Bernoulli}(\pi)$$

Independent and identically distributed.

- Z = number of successful trials
- Derive probability mass function $P(Z = M) = p(M)$
- One way to obtain M successful trials:

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent
 - Each Bernoulli trial i is

$$Y_i \sim \text{Bernoulli}(\pi)$$

Independent and identically distributed.

- Z = number of successful trials
- Derive probability mass function $P(Z = M) = p(M)$
- One way to obtain M successful trials:

$$P(Y_1 = 1, Y_2 = 0, Y_3 = 1, \dots, Y_N = 1)$$

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent
 - Each Bernoulli trial i is

$$Y_i \sim \text{Bernoulli}(\pi)$$

Independent and identically distributed.

- Z = number of successful trials
- Derive probability mass function $P(Z = M) = p(M)$
- One way to obtain M successful trials:

$$\begin{aligned} P(Y_1 = 1, Y_2 = 0, Y_3 = 1, \dots, Y_N = 1) \\ = P(Y_1 = 1)P(Y_2 = 0) \cdots P(Y_N = 1) \end{aligned}$$

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent
 - Each Bernoulli trial i is

$$Y_i \sim \text{Bernoulli}(\pi)$$

Independent and identically distributed.

- Z = number of successful trials
- Derive probability mass function $P(Z = M) = p(M)$
- One way to obtain M successful trials:

$$\begin{aligned} &P(Y_1 = 1, Y_2 = 0, Y_3 = 1, \dots, Y_N = 1) \\ &= P(Y_1 = 1)P(Y_2 = 0) \cdots P(Y_N = 1) \\ &= \underbrace{P(Y_1 = 1)P(Y_3 = 1) \cdots P(Y_M = 1)}_M \times \underbrace{P(Y_2 = 0) \cdots P(Y_N = 0)}_{N-M} \end{aligned}$$

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent
 - Each Bernoulli trial i is

$$Y_i \sim \text{Bernoulli}(\pi)$$

Independent and identically distributed.

- Z = number of successful trials
- Derive probability mass function $P(Z = M) = p(M)$
- One way to obtain M successful trials:

$$\begin{aligned} &P(Y_1 = 1, Y_2 = 0, Y_3 = 1, \dots, Y_N = 1) \\ &= P(Y_1 = 1)P(Y_2 = 0) \cdots P(Y_N = 1) \\ &= \underbrace{P(Y_1 = 1)P(Y_3 = 1) \cdots P(Y_M = 1)}_M \times \underbrace{P(Y_2 = 0) \cdots P(Y_N = 0)}_{N-M} \\ &= \underbrace{\pi \pi \cdots \pi}_M \times \underbrace{(1 - \pi)(1 - \pi) \cdots (1 - \pi)}_{N-M} \end{aligned}$$

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent
 - Each Bernoulli trial i is

$$Y_i \sim \text{Bernoulli}(\pi)$$

Independent and identically distributed.

- Z = number of successful trials
- Derive probability mass function $P(Z = M) = p(M)$
- One way to obtain M successful trials:

$$\begin{aligned} &P(Y_1 = 1, Y_2 = 0, Y_3 = 1, \dots, Y_N = 1) \\ &= P(Y_1 = 1)P(Y_2 = 0) \cdots P(Y_N = 1) \\ &= \underbrace{P(Y_1 = 1)P(Y_3 = 1) \cdots P(Y_M = 1)}_M \times \underbrace{P(Y_2 = 0) \cdots P(Y_N = 0)}_{N-M} \\ &= \underbrace{\pi \pi \cdots \pi}_M \times \underbrace{(1 - \pi)(1 - \pi) \cdots (1 - \pi)}_{N-M} \\ &= \pi^M (1 - \pi)^{N-M} \end{aligned}$$

Binomial Random Variable

- A model to count the number of successes across N trials
 - Assume the Bernoulli trials are independent
 - Each Bernoulli trial i is

$$Y_i \sim \text{Bernoulli}(\pi)$$

Independent and identically distributed.

- Z = number of successful trials
- Derive probability mass function $P(Z = M) = p(M)$
- One way to obtain M successful trials:

$$\begin{aligned} &P(Y_1 = 1, Y_2 = 0, Y_3 = 1, \dots, Y_N = 1) \\ &= P(Y_1 = 1)P(Y_2 = 0) \cdots P(Y_N = 1) \\ &= \underbrace{P(Y_1 = 1)P(Y_3 = 1) \cdots P(Y_M = 1)}_M \times \underbrace{P(Y_2 = 0) \cdots P(Y_N = 0)}_{N-M} \\ &= \underbrace{\pi \pi \cdots \pi}_M \times \underbrace{(1 - \pi)(1 - \pi) \cdots (1 - \pi)}_{N-M} \\ &= \pi^M (1 - \pi)^{N-M} \end{aligned}$$

Are we done?

Are we done? No

Are we done? **No**

- This is just one instance of M successes

Are we done? **No**

- This is just one instance of M successes
- How many total instances?

Are we done? **No**

- This is just one instance of M successes
- How many total instances?
 - N total trials

Are we done? **No**

- This is just one instance of M successes
- How many total instances?
 - N total trials
 - We want to select M

Are we done? **No**

- This is just one instance of M successes
- How many total instances?
 - N total trials
 - We want to select M
- $\binom{N}{M} = \frac{N!}{(N-M)!M!}$

Are we done? **No**

- This is just one instance of M successes
- How many total instances?
 - N total trials
 - We want to select M
- $\binom{N}{M} = \frac{N!}{(N-M)!M!}$

Then,

Are we done? **No**

- This is just one instance of M successes
- How many total instances?
 - N total trials
 - We want to select M
- $\binom{N}{M} = \frac{N!}{(N-M)!M!}$

Then,

$$P(Z = M) = p(M) = \binom{N}{M} \pi^M (1 - \pi)^{N-M}$$

Are we done? **No**

- This is just one instance of M successes
- How many total instances?
 - N total trials
 - We want to select M
- $\binom{N}{M} = \frac{N!}{(N-M)!M!}$

Then,

$$P(Z = M) = p(M) = \binom{N}{M} \pi^M (1 - \pi)^{N-M}$$

Definition

Suppose Y is a random variable that counts the number of successes in N independent and identically distributed Bernoulli trials. Then Y is a **Binomial** random variable,

$$p(k) = \binom{N}{k} \pi^k (1 - \pi)^{N-k}$$

for $k \in \{0, 1, 2, \dots, N\}$ and $p(k) = 0$ otherwise.

Equivalently,

$$Y \sim \text{Binomial}(N, \pi)$$

Binomial Example

Recall our experiment example:

Binomial Example

Recall our experiment example:

$$P(T) = P(C) = 1/2.$$

Binomial Example

Recall our experiment example:

$$P(T) = P(C) = 1/2.$$

Z = number of units assigned to treatment

Binomial Example

Recall our experiment example:

$$P(T) = P(C) = 1/2.$$

Z = number of units assigned to treatment

$$Z \sim \text{Binomial}(1/2)$$

Binomial Example

Recall our experiment example:

$$P(T) = P(C) = 1/2.$$

Z = number of units assigned to treatment

$$Z \sim \text{Binomial}(1/2)$$

$$p(0) = \binom{3}{0} (1/2)^0 (1 - 1/2)^{3-0} = 1 \times \frac{1}{8}$$

Binomial Example

Recall our experiment example:

$$P(T) = P(C) = 1/2.$$

Z = number of units assigned to treatment

$$Z \sim \text{Binomial}(1/2)$$

$$p(0) = \binom{3}{0} (1/2)^0 (1 - 1/2)^{3-0} = 1 \times \frac{1}{8}$$

$$p(1) = \binom{3}{1} (1/2)^1 (1 - 1/2)^2 = 3 \times \frac{1}{8}$$

Binomial Example

Recall our experiment example:

$$P(T) = P(C) = 1/2.$$

Z = number of units assigned to treatment

$$Z \sim \text{Binomial}(1/2)$$

$$p(0) = \binom{3}{0} (1/2)^0 (1 - 1/2)^{3-0} = 1 \times \frac{1}{8}$$

$$p(1) = \binom{3}{1} (1/2)^1 (1 - 1/2)^2 = 3 \times \frac{1}{8}$$

$$p(2) = \binom{3}{2} (1/2)^2 (1 - 1/2)^1 = 3 \times \frac{1}{8}$$

Binomial Example

Recall our experiment example:

$$P(T) = P(C) = 1/2.$$

Z = number of units assigned to treatment

$$Z \sim \text{Binomial}(1/2)$$

$$p(0) = \binom{3}{0} (1/2)^0 (1 - 1/2)^{3-0} = 1 \times \frac{1}{8}$$

$$p(1) = \binom{3}{1} (1/2)^1 (1 - 1/2)^2 = 3 \times \frac{1}{8}$$

$$p(2) = \binom{3}{2} (1/2)^2 (1 - 1/2)^1 = 3 \times \frac{1}{8}$$

$$p(3) = \binom{3}{3} (1/2)^3 (1 - 1/2)^0 = 1 \times \frac{1}{8}$$

Binomial Example

Recall our experiment example:

$$P(T) = P(C) = 1/2.$$

Z = number of units assigned to treatment

$$Z \sim \text{Binomial}(1/2)$$

$$p(0) = \binom{3}{0} (1/2)^0 (1 - 1/2)^{3-0} = 1 \times \frac{1}{8}$$

$$p(1) = \binom{3}{1} (1/2)^1 (1 - 1/2)^2 = 3 \times \frac{1}{8}$$

$$p(2) = \binom{3}{2} (1/2)^2 (1 - 1/2)^1 = 3 \times \frac{1}{8}$$

$$p(3) = \binom{3}{3} (1/2)^3 (1 - 1/2)^0 = 1 \times \frac{1}{8}$$

Binomial Random Variable Moments

$$Z = \sum_{i=1}^N Y_i \text{ where } Y_i \sim \text{Bernoulli}(\pi)$$

Binomial Random Variable Moments

$$Z = \sum_{i=1}^N Y_i \text{ where } Y_i \sim \text{Bernoulli}(\pi)$$

$$E[Z] = E[Y_1 + Y_2 + Y_3 + \dots + Y_N]$$

Binomial Random Variable Moments

$$Z = \sum_{i=1}^N Y_i \text{ where } Y_i \sim \text{Bernoulli}(\pi)$$

$$\begin{aligned} E[Z] &= E[Y_1 + Y_2 + Y_3 + \dots + Y_N] \\ &= \sum_{i=1}^N E[Y_i] \end{aligned}$$

Binomial Random Variable Moments

$$Z = \sum_{i=1}^N Y_i \text{ where } Y_i \sim \text{Bernoulli}(\pi)$$

$$\begin{aligned} E[Z] &= E[Y_1 + Y_2 + Y_3 + \dots + Y_N] \\ &= \sum_{i=1}^N E[Y_i] \\ &= N\pi \end{aligned}$$

$$E[Z] = N\pi$$

Binomial Random Variable Moments

$$Z = \sum_{i=1}^N Y_i \text{ where } Y_i \sim \text{Bernoulli}(\pi)$$

$$\begin{aligned} E[Z] &= E[Y_1 + Y_2 + Y_3 + \dots + Y_N] \\ &= \sum_{i=1}^N E[Y_i] \\ &= N\pi \\ \text{var}(Z) &= \sum_{i=1}^N \text{var}(Y_i) \end{aligned}$$

$$E[Z] = N\pi$$

Binomial Random Variable Moments

$$Z = \sum_{i=1}^N Y_i \text{ where } Y_i \sim \text{Bernoulli}(\pi)$$

$$E[Z] = E[Y_1 + Y_2 + Y_3 + \dots + Y_N]$$

$$= \sum_{i=1}^N E[Y_i]$$

$$= N\pi$$

$$\text{var}(Z) = \sum_{i=1}^N \text{var}(Y_i)$$

$$= N\pi(1 - \pi)$$

$$E[Z] = N\pi$$

$$\text{var}(Z) = N\pi(1 - \pi)$$

Voter Turnout

Suppose we have a set N voters, with iid turnout decisions

$$Y_i \sim \text{Bernoulli}(\pi)$$

Voter Turnout

Suppose we have a set N voters, with iid turnout decisions

$$Y_i \sim \text{Bernoulli}(\pi)$$

What is the probability that at least M voters turnout?

Voter Turnout

Suppose we have a set N voters, with iid turnout decisions

$Y_i \sim \text{Bernoulli}(\pi)$

What is the probability that at least M voters turnout?

$$P(k \geq M) = \sum_{k=M}^N \binom{N}{k} \pi^k (1 - \pi)^{N-k}$$

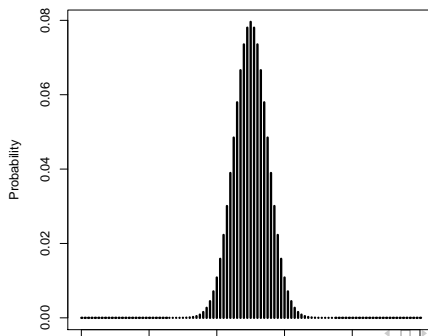
Voter Turnout

Suppose we have a set N voters, with iid turnout decisions

$Y_i \sim \text{Bernoulli}(\pi)$

What is the probability that at least M voters turnout?

$$P(k \geq M) = \sum_{k=M}^N \binom{N}{k} \pi^k (1 - \pi)^{N-k}$$



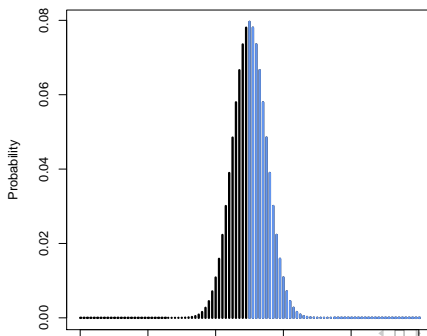
Voter Turnout

Suppose we have a set N voters, with iid turnout decisions

$Y_i \sim \text{Bernoulli}(\pi)$

What is the probability that at least M voters turnout?

$$P(k \geq M) = \sum_{k=M}^N \binom{N}{k} \pi^k (1 - \pi)^{N-k}$$



Voter Turnout

Suppose we have a set N voters, with iid turnout decisions

$Y_i \sim \text{Bernoulli}(\pi)$

What is the probability that at least M voters turnout?

$$P(k \geq M) = \sum_{k=M}^N \binom{N}{k} \pi^k (1 - \pi)^{N-k}$$

R Code!

Voter Turnout, with Spillovers

Suppose we have the same set of N voters.

Voter Turnout, with Spillovers

Suppose we have the same set of N voters.

Now, $N/2$ are leaders, who turnout with probability $(1/2)$

Voter Turnout, with Spillovers

Suppose we have the same set of N voters.

Now, $N/2$ are leaders, who turnout with probability $(1/2)$

But, $N/2$ are followers, whose turnout depends on a specific leader

Voter Turnout, with Spillovers

Suppose we have the same set of N voters.

Now, $N/2$ are leaders, who turnout with probability $(1/2)$

But, $N/2$ are followers, whose turnout depends on a specific leader

Suppose follower i depends on only one leader j (and each follower has their own leader)

Voter Turnout, with Spillovers

Suppose we have the same set of N voters.

Now, $N/2$ are leaders, who turnout with probability $(1/2)$

But, $N/2$ are followers, whose turnout depends on a specific leader

Suppose follower i depends on only one leader j (and each follower has their own leader)

$$Y_i \sim \text{Bernoulli}(0.9) \text{ if } j \text{ votes}$$

Voter Turnout, with Spillovers

Suppose we have the same set of N voters.

Now, $N/2$ are leaders, who turnout with probability $(1/2)$

But, $N/2$ are followers, whose turnout depends on a specific leader

Suppose follower i depends on only one leader j (and each follower has their own leader)

$Y_i \sim \text{Bernoulli}(0.9)$ if j votes

$Y_i \sim \text{Bernoulli}(0.1)$ if j does not

Voter Turnout, with Spillovers

Suppose we have the same set of N voters.

Now, $N/2$ are leaders, who turnout with probability $(1/2)$

But, $N/2$ are followers, whose turnout depends on a specific leader

Suppose follower i depends on only one leader j (and each follower has their own leader)

$$Y_i \sim \text{Bernoulli}(0.9) \text{ if } j \text{ votes}$$
$$Y_i \sim \text{Bernoulli}(0.1) \text{ if } j \text{ does not}$$

Let Z be the number of voters who turnout.

Voter Turnout, with Spillovers

Suppose we have the same set of N voters.

Now, $N/2$ are leaders, who turnout with probability $(1/2)$

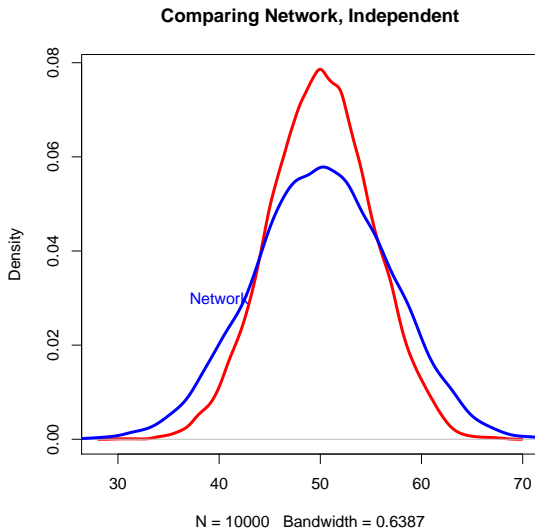
But, $N/2$ are followers, whose turnout depends on a specific leader

Suppose follower i depends on only one leader j (and each follower has their own leader)

$$Y_i \sim \text{Bernoulli}(0.9) \text{ if } j \text{ votes}$$
$$Y_i \sim \text{Bernoulli}(0.1) \text{ if } j \text{ does not}$$

Let Z be the number of voters who turnout.

Voter Turnout, with Spillovers



Trials with More than Two Outcomes

Definition

Suppose we observe a trial, which might result in J outcomes.

And that $P(\text{outcome} = i) = \pi_i$

$\mathbf{Y} = (Y_1, Y_2, \dots, Y_J)$ where $Y_j = 1$ if outcome j occurred and 0 otherwise.

*Then \mathbf{Y} follows a **multinomial** distribution, with*

$$p(\mathbf{y}) = \pi_1^{y_1} \pi_2^{y_2} \dots \pi_k^{y_J}$$

if $\sum_{i=1}^J y_i = 1$ and the pmf is 0 otherwise.

Equivalently, we'll write

$$\mathbf{Y} \sim \text{Multinomial}(1, \boldsymbol{\pi})$$

$$\mathbf{Y} \sim \text{Categorical}(\boldsymbol{\pi})$$

Multinomial Properties + Notes

Computer scientists: commonly call Multinomial($1, \pi$) **Discrete**(π).

$$\begin{aligned} E[X_i] &= N\pi_i \\ \text{var}(X_i) &= N\pi_i(1 - \pi_i) \end{aligned}$$

Investigate Further in Homework!

Counting the Number of Events

Often interested in counting number of events that occur:

- 1) Number of wars started
- 2) Number of speeches made
- 3) Number of bribes offered
- 4) Number of people waiting for license

Generally referred to as **event counts**

Stochastic processes: a course provide introduction to many processes
(**Queing Theory**)

Poisson Distribution

Definition

Suppose X is a random variable that takes on values $X \in \{0, 1, 2, \dots\}$ and that $P(X = k) = p(k)$ is,

$$p(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

for $k \in \{0, 1, \dots\}$ and 0 otherwise. Then we will say that X follows a *Poisson* distribution with *rate* parameter λ .

$$X \sim \text{Poisson}(\lambda)$$

Example: Poisson Distribution

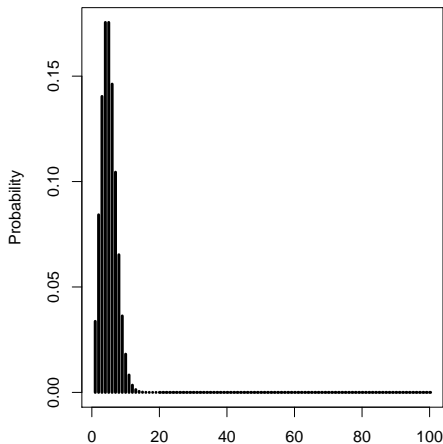
Suppose the number of threats a president makes in a term is given by $X \sim \text{Poisson}(5)$.

Example: Poisson Distribution

Suppose the number of threats a president makes in a term is given by $X \sim \text{Poisson}(5)$. What is the probability the president will make ten or more threats?

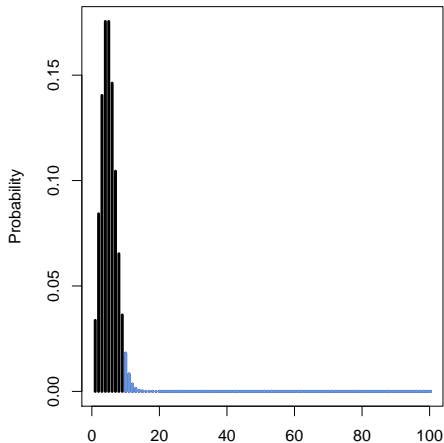
Example: Poisson Distribution

Suppose the number of threats a president makes in a term is given by $X \sim \text{Poisson}(5)$. What is the probability the president will make ten or more threats?



Example: Poisson Distribution

Suppose the number of threats a president makes in a term is given by $X \sim \text{Poisson}(5)$. What is the probability the president will make ten or more threats?



Example: Poisson Distribution

Suppose the number of threats a president makes in a term is given by $X \sim \text{Poisson}(5)$. What is the probability the president will make ten or more threats?

$$P(X \geq 10) = e^{-\lambda} \sum_{k=10}^{\infty} \frac{5^k}{k!}$$

Example: Poisson Distribution

Suppose the number of threats a president makes in a term is given by $X \sim \text{Poisson}(5)$. What is the probability the president will make ten or more threats?

$$\begin{aligned} P(X \geq 10) &= e^{-\lambda} \sum_{k=10}^{\infty} \frac{5^k}{k!} \\ &= 1 - P(X < 10) \end{aligned}$$

Example: Poisson Distribution

Suppose the number of threats a president makes in a term is given by $X \sim \text{Poisson}(5)$. What is the probability the president will make ten or more threats?

$$\begin{aligned} P(X \geq 10) &= e^{-\lambda} \sum_{k=10}^{\infty} \frac{5^k}{k!} \\ &= 1 - P(X < 10) \end{aligned}$$

R code!

Poisson Distribution

Properties:

- 1) It is a probability distribution.

Poisson Distribution

Properties:

- 1) It is a probability distribution.

Recall the **Taylor expansion** of e^x

Poisson Distribution

Properties:

- 1) It is a probability distribution.

Recall the **Taylor expansion** of e^x

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Poisson Distribution

Properties:

1) It is a probability distribution.

Recall the **Taylor expansion** of e^x

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$
$$e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} (1 + \lambda + \frac{\lambda^2}{2!} + \dots)$$

Poisson Distribution

Properties:

- 1) It is a probability distribution.

Recall the **Taylor expansion** of e^x

$$\begin{aligned}e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \\e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} &= e^{-\lambda} \left(1 + \lambda + \frac{\lambda^2}{2!} + \dots \right) \\&= e^{-\lambda} (e^{\lambda}) = 1\end{aligned}$$

Poisson Distribution

Properties:

Poisson Distribution

Properties:

$$2) E[X] = \lambda$$

Poisson Distribution

Properties:

$$2) E[X] = \lambda$$

$$E[X] = e^{-\lambda} \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!}$$

Poisson Distribution

Properties:

$$2) E[X] = \lambda$$

$$\begin{aligned} E[X] &= e^{-\lambda} \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \end{aligned}$$

Poisson Distribution

Properties:

$$2) E[X] = \lambda$$

$$\begin{aligned} E[X] &= e^{-\lambda} \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \end{aligned}$$

Define $j = k - 1$, then

Poisson Distribution

Properties:

$$2) E[X] = \lambda$$

$$\begin{aligned} E[X] &= e^{-\lambda} \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \end{aligned}$$

Define $j = k - 1$, then

$$E[X] = e^{-\lambda} \lambda \sum_{j=0}^{\infty} \frac{\lambda^j}{j!}$$

Poisson Distribution

Properties:

$$2) E[X] = \lambda$$

$$\begin{aligned} E[X] &= e^{-\lambda} \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \end{aligned}$$

Define $j = k - 1$, then

$$\begin{aligned} E[X] &= e^{-\lambda} \lambda \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \\ &= e^{-\lambda} \lambda e^{\lambda} \end{aligned}$$

Poisson Distribution

Properties:

$$2) E[X] = \lambda$$

$$\begin{aligned} E[X] &= e^{-\lambda} \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \end{aligned}$$

Define $j = k - 1$, then

$$\begin{aligned} E[X] &= e^{-\lambda} \lambda \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \\ &= e^{-\lambda} \lambda e^{\lambda} \\ &= \lambda \end{aligned}$$

Poisson Distribution

Properties:

Poisson Distribution

Properties:

$$3) \text{ var}(X) = \lambda$$

Poisson Distribution

Properties:

$$3) \text{ var}(X) = \lambda$$

$$E[X^2] = \sum_{k=0}^{\infty} \frac{k^2 e^{-\lambda} \lambda^k}{k!}$$

Poisson Distribution

Properties:

$$3) \text{ var}(X) = \lambda$$

$$\begin{aligned} E[X^2] &= \sum_{k=0}^{\infty} \frac{k^2 e^{-\lambda} \lambda^k}{k!} \\ &= \lambda e^{-\lambda} \left(\sum_{k=1}^{\infty} \frac{k \lambda^{k-1}}{(k-1)!} \right) \end{aligned}$$

Poisson Distribution

Properties:

$$3) \text{ var}(X) = \lambda$$

$$\begin{aligned} E[X^2] &= \sum_{k=0}^{\infty} \frac{k^2 e^{-\lambda} \lambda^k}{k!} \\ &= \lambda e^{-\lambda} \left(\sum_{k=1}^{\infty} \frac{k \lambda^{k-1}}{(k-1)!} \right) \end{aligned}$$

Let $j = k - 1$,

Poisson Distribution

Properties:

$$3) \text{ var}(X) = \lambda$$

$$\begin{aligned} E[X^2] &= \sum_{k=0}^{\infty} \frac{k^2 e^{-\lambda} \lambda^k}{k!} \\ &= \lambda e^{-\lambda} \left(\sum_{k=1}^{\infty} \frac{k \lambda^{k-1}}{(k-1)!} \right) \end{aligned}$$

Let $j = k - 1$,

$$E[X^2] = \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{(j+1) \lambda^j}{j!}$$

Poisson Distribution

Properties:

$$3) \text{ var}(X) = \lambda$$

$$\begin{aligned} E[X^2] &= \sum_{k=0}^{\infty} \frac{k^2 e^{-\lambda} \lambda^k}{k!} \\ &= \lambda e^{-\lambda} \left(\sum_{k=1}^{\infty} \frac{k \lambda^{k-1}}{(k-1)!} \right) \end{aligned}$$

Let $j = k - 1$,

$$\begin{aligned} E[X^2] &= \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{(j+1) \lambda^j}{j!} \\ &= \lambda e^{-\lambda} \left(\sum_{j=0}^{\infty} \frac{(j) \lambda^j}{j!} + \sum_{j=0}^{\infty} \frac{(1) \lambda^j}{j!} \right) \end{aligned}$$

Poisson Distribution

Properties:

$$3) \text{ var}(X) = \lambda$$

$$\begin{aligned} E[X^2] &= \sum_{k=0}^{\infty} \frac{k^2 e^{-\lambda} \lambda^k}{k!} \\ &= \lambda e^{-\lambda} \left(\sum_{k=1}^{\infty} \frac{k \lambda^{k-1}}{(k-1)!} \right) \end{aligned}$$

Let $j = k - 1$,

$$\begin{aligned} E[X^2] &= \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{(j+1) \lambda^j}{j!} \\ &= \lambda e^{-\lambda} \left(\sum_{j=0}^{\infty} \frac{(j) \lambda^j}{j!} + \sum_{j=0}^{\infty} \frac{(1) \lambda^j}{j!} \right) \\ &= \lambda e^{-\lambda} (\lambda e^{\lambda} + e^{\lambda}) \end{aligned}$$

Poisson Distribution

Properties

$$3) \text{ var}(X) = \lambda$$

$$E[X^2] = \lambda e^{-\lambda}(\lambda e^{\lambda} + e^{\lambda})$$

Poisson Distribution

Properties

$$3) \text{ var}(X) = \lambda$$

$$\begin{aligned} E[X^2] &= \lambda e^{-\lambda}(\lambda e^{\lambda} + e^{\lambda}) \\ &= \lambda(\lambda + 1) \end{aligned}$$

Poisson Distribution

Properties

$$3) \text{ var}(X) = \lambda$$

$$\begin{aligned} E[X^2] &= \lambda e^{-\lambda}(\lambda e^{\lambda} + e^{\lambda}) \\ &= \lambda(\lambda + 1) \end{aligned}$$

$$\text{var}(X) = E[X^2] - E[X]^2$$

Poisson Distribution

Properties

$$3) \text{ var}(X) = \lambda$$

$$\begin{aligned} E[X^2] &= \lambda e^{-\lambda}(\lambda e^{\lambda} + e^{\lambda}) \\ &= \lambda(\lambda + 1) \end{aligned}$$

$$\text{var}(X) = E[X^2] - E[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Poisson Distribution

Properties

$$3) \text{ var}(X) = \lambda$$

$$\begin{aligned} E[X^2] &= \lambda e^{-\lambda}(\lambda e^{\lambda} + e^{\lambda}) \\ &= \lambda(\lambda + 1) \end{aligned}$$

$$\text{var}(X) = E[X^2] - E[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Very useful distribution, with strong assumptions. We'll explore in homework!

Often interested in how processes evolve over time

- Given voting history, probability of voting in the future
- Given history of candidate support, probability of future support
- Given prior conflicts, probability of future war
- Given previous words in a sentence, probability of next word

Potentially complex history

Often interested in how processes evolve over time

- Given voting history, probability of voting in the future
- Given history of candidate support, probability of future support
- Given prior conflicts, probability of future war
- Given previous words in a sentence, probability of next word

Potentially complex history

Stochastic Process

Definition

Suppose we have a sequence of random variables

$\{X\}_{i=0}^M = X_0, X_1, X_2, \dots, X_M$ that take on the countable values of S . We will call $\{X\}_{i=0}^M$ a stochastic process with state space S .

If index gives time, then we might condition on history to obtain probability

$$\text{PMF } X_t, \text{ given history} = P(X_t | X_{t-1}, X_{t-2}, \dots, X_1, X_0)$$

Still Complex

Markov Chain

Definition

Suppose we have a stochastic process $\{X\}_{i=0}^M$ with countable state space S . Then $\{X\}_{i=0}^M$ is a markov chain if:

$$P(X_t | X_{t-1}, X_{t-2}, \dots, X_1, X_0) = P(X_t | X_{t-1})$$

A Markov chain's future depends only on its current state

Transition Matrix

Habitual turnout?

$$\mathbf{T} = \begin{pmatrix} & \text{Vote}_t & \text{Not Vote}_t \\ \text{Vote}_{t-1} & 0.8 & 0.2 \\ \text{Not Vote}_{t-1} & 0.3 & 0.7 \end{pmatrix}$$

- Suppose someone starts as a voter—what is their behavior after
- 1 iteration?
- 2 iterations?
- The long run?

R Code!

Tomorrow: Continuous Random Variables!