Probability: Random Variables and Large Samples

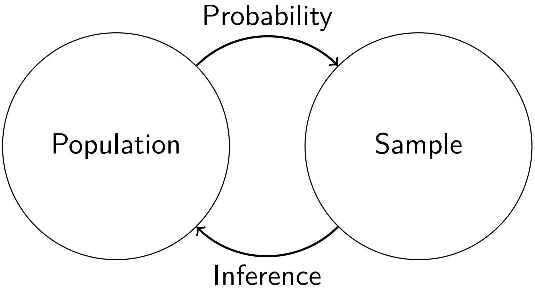
PSC4375: Week 10

Prof. Weldzius

Villanova University

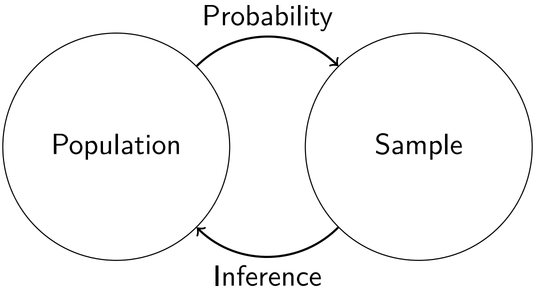
Slides Updated: 2025-03-28

Learning about populations



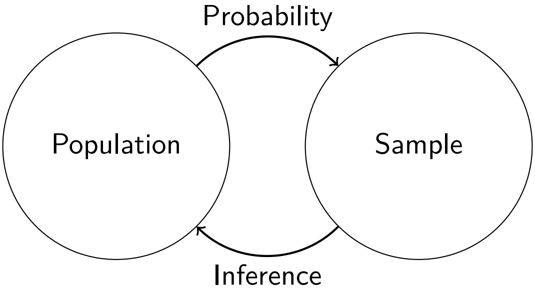
- We want to learn about the chance process that generated our data.

Learning about populations



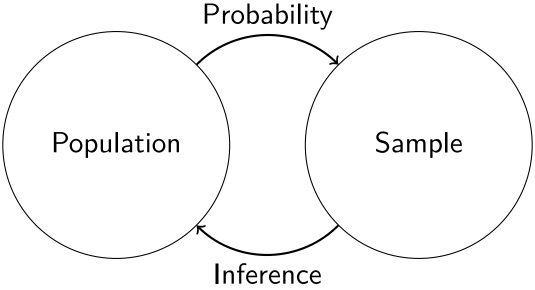
- We want to learn about the chance process that generated our data.
  - What's the true support for Trump in the population?

Learning about populations



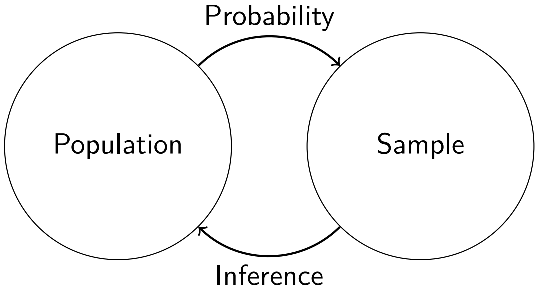
- We want to learn about the chance process that generated our data.
  - What's the true support for Trump in the population?
  - We only get to see a sample from the population.

Learning about populations



- We want to learn about the chance process that generated our data.
  - What's the true support for Trump in the population?
  - We only get to see a sample from the population.
  - Stare at the results of 1000 coin flips and determine if the coin was fair.

Learning about populations



- We want to learn about the chance process that generated our data.
  - What's the true support for Trump in the population?
  - We only get to see a sample from the population.
  - Stare at the results of 1000 coin flips and determine if the coin was fair.
- We have probability to help us, but. . .

What are random variables?

$$\{\text{draw a Trump supporter}\} \overset{???}{\longleftrightarrow} \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Randomly selecting senators, redux

	Democrats	*Republicans*	*Independents*	*Total*
Men	29	43	2	74
Women	16	10	0	26
Total	45	53	2	100

- Draw a Senator's name from a hat and define the random variable:

Randomly selecting senators, redux

	Democrats	*Republicans*	*Independents*	*Total*
Men	29	43	2	74
Women	16	10	0	26
Total	45	53	2	100

- Draw a Senator's name from a hat and define the random variable:
- A **random variable** is a mapping from the outcomes to numbers.

Randomly selecting senators, redux

	Democrats	*Republicans*	*Independents*	*Total*
Men	29	43	2	74
Women	16	10	0	26
Total	45	53	2	100

- Draw a Senator's name from a hat and define the random variable:
- A **random variable** is a mapping from the outcomes to numbers.
  - Example: $X = 1$ if selected Senator is a woman, $X = 0$ otherwise
- **Random**: before we draw, there is uncertainty about the value of X!
- Straightforward probability connection:

$$\mathbb{P}(X = 1) = \mathbb{P}(\text{draw a woman senator}) = \frac{26}{100}$$

Bernoulli r.v.

Bernoulli r.v.

- An r.v. is X is said to follow a Bernoulli distribution with probability p if:

Bernoulli r.v.

- An r.v. is X is said to follow a Bernoulli distribution with probability p if:
    - X takes on only two values, 0 and 1, and

Bernoulli r.v.

- An r.v. is X is said to follow a Bernoulli distribution with probability p if:
  - X takes on only two values, 0 and 1, and
  - $\mathbb{P}(X = 1) = p$ and $\mathbb{P}(X = 0) = 1 - p$

Bernoulli r.v.

- An r.v. is X is said to follow a Bernoulli distribution with probability p if:
  - X takes on only two values, 0 and 1, and
  - $\mathbb{P}(X = 1) = p$ and $\mathbb{P}(X = 0) = 1 - p$
- Simplest possible random variable: indicator/binary variable.

Bernoulli r.v.

- An r.v. is X is said to follow a Bernoulli distribution with probability p if:
  - X takes on only two values, 0 and 1, and
  - $\mathbb{P}(X = 1) = p$ and $\mathbb{P}(X = 0) = 1 - p$
- Simplest possible random variable: indicator/binary variable.
- Distribution of a Bernoulli r.v. entirely determined by p.

Bernoulli r.v.

- An r.v. is X is said to follow a Bernoulli distribution with probability p if:
  - X takes on only two values, 0 and 1, and
  - $\mathbb{P}(X = 1) = p$ and $\mathbb{P}(X = 0) = 1 - p$
- Simplest possible random variable: indicator/binary variable.
- Distribution of a Bernoulli r.v. entirely determined by p.
  - Infinite number of possible Bernoulli r.v.s: one for each value of p.

Why random variables?

- Why go through the trouble of defining random variables?

Why random variables?

- Why go through the trouble of defining random variables?
  - Allows us to think about the uncertainty of our estimates.

Why random variables?

- Why go through the trouble of defining random variables?
  - Allows us to think about the uncertainty of our estimates.
  - Before analyzing sample means useful to detour into sample sums.

Why random variables?

- Why go through the trouble of defining random variables?
  - Allows us to think about the uncertainty of our estimates.
  - Before analyzing sample means useful to detour into sample sums.
- Extremely small data example: sample two senators with replacement.

Why random variables?

- Why go through the trouble of defining random variables?
  - Allows us to think about the uncertainty of our estimates.
  - Before analyzing sample means useful to detour into sample sums.
- Extremely small data example: sample two senators with replacement.
  - $X_1 = 1$ if senator 1 is a woman, $X_1 = 0$ otherwise

Why random variables?

- Why go through the trouble of defining random variables?
  - Allows us to think about the uncertainty of our estimates.
  - Before analyzing sample means useful to detour into sample sums.
- Extremely small data example: sample two senators with replacement.
  - $X_1 = 1$ if senator 1 is a woman, $X_1 = 0$ otherwise
  - $X_2 = 1$ if senator 2 is a woman, $X_2 = 0$ otherwise

Why random variables?

- Why go through the trouble of defining random variables?
  - Allows us to think about the uncertainty of our estimates.
  - Before analyzing sample means useful to detour into sample sums.
- Extremely small data example: sample two senators with replacement.
  - $X_1 = 1$ if senator 1 is a woman, $X_1 = 0$ otherwise
  - $X_2 = 1$ if senator 2 is a woman, $X_2 = 0$ otherwise
- Define the sum of these: $S = X_1 + X_2$ (also an r.v.)

Why random variables?

- Why go through the trouble of defining random variables?
    - Allows us to think about the uncertainty of our estimates.
    - Before analyzing sample means useful to detour into sample sums.
- Extremely small data example: sample two senators with replacement.
    - $X_1 = 1$ if senator 1 is a woman, $X_1 = 0$ otherwise
    - $X_2 = 1$ if senator 2 is a woman, $X_2 = 0$ otherwise
- Define the sum of these: $S = X_1 + X_2$ (also an r.v.)
- What sums should we expect to see?

Why random variables?

- Why go through the trouble of defining random variables?
  - Allows us to think about the uncertainty of our estimates.
  - Before analyzing sample means useful to detour into sample sums.
- Extremely small data example: sample two senators with replacement.
  - $X_1 = 1$ if senator 1 is a woman, $X_1 = 0$ otherwise
  - $X_2 = 1$ if senator 2 is a woman, $X_2 = 0$ otherwise
- Define the sum of these: $S = X_1 + X_2$ (also an r.v.)
- What sums should we expect to see?
  - How surprised should we be if $S = 1$?

Why random variables?

- Why go through the trouble of defining random variables?
  - Allows us to think about the uncertainty of our estimates.
  - Before analyzing sample means useful to detour into sample sums.
- Extremely small data example: sample two senators with replacement.
  - $X_1 = 1$ if senator 1 is a woman, $X_1 = 0$ otherwise
  - $X_2 = 1$ if senator 2 is a woman, $X_2 = 0$ otherwise
- Define the sum of these: $S = X_1 + X_2$ (also an r.v.)
- What sums should we expect to see?
  - How surprised should we be if $S = 1$?
  - What is the probability of each possible value, $\mathbb{P}(S = k)$

Binomial distribution

Binomial distribution

- **Binomial r.v.**: X takes on any integer between 0 and n.

Binomial distribution

- **Binomial r.v.**: X takes on any integer between 0 and n.
    - Number of heads in n independent coin flips with probability p of heads.

Binomial distribution

- **Binomial r.v.**: X takes on any integer between 0 and n.
  - Number of heads in n independent coin flips with probability p of heads.
  - "Binomial with n trials and probability of success p''

Binomial distribution

- **Binomial r.v.**: X takes on any integer between 0 and n.
  - Number of heads in n independent coin flips with probability p of heads.
  - "Binomial with n trials and probability of success p''
- Example: random draws of two senators, Y is how many are women?

Binomial distribution

- **Binomial r.v.**: X takes on any integer between 0 and n.
  - Number of heads in n independent coin flips with probability p of heads.
  - "Binomial with n trials and probability of success p"
- Example: random draws of two senators, Y is how many are women?
  - Binomial with $n = 2$ and $p = 0.26$.

Binomial distribution

- **Binomial r.v.**: X takes on any integer between 0 and n.
  - Number of heads in n independent coin flips with probability p of heads.
  - "Binomial with n trials and probability of success p"
- Example: random draws of two senators, Y is how many are women?
  - Binomial with $n = 2$ and $p = 0.26$.
- **Probability mass function** gives the probability of any possible value:

Binomial distribution

- **Binomial r.v.**: X takes on any integer between 0 and n.
  - Number of heads in n independent coin flips with probability p of heads.
  - "Binomial with n trials and probability of success p''
- Example: random draws of two senators, Y is how many are women?
  - Binomial with $n = 2$ and $p = 0.26$.
- **Probability mass function** gives the probability of any possible value:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Binomial distribution

- **Binomial r.v.**: X takes on any integer between 0 and n.
  - Number of heads in n independent coin flips with probability p of heads.
  - "Binomial with n trials and probability of success p"
- Example: random draws of two senators, Y is how many are women?
  - Binomial with $n = 2$ and $p = 0.26$.
- **Probability mass function** gives the probability of any possible value:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Binomial distribution ($n = 10$, $p = 0.5$)

Binomial distribution (n = 10, p = 0.5)

More calls to senators

You work as a lobbyist and you've been asked to check to see the gender balance of the calls placed to Senate offices from your firm. The firm has placed 1000 calls over the last year. If the firm was randomly choosing senators (with replacement) each call, what numbers of women senators contacted would be more or less plausible?

More calls to senators

You work as a lobbyist and you've been asked to check to see the gender balance of the calls placed to Senate offices from your firm. The firm has placed 1000 calls over the last year. If the firm was randomly choosing senators (with replacement) each call, what numbers of women senators contacted would be more or less plausible?

- That math formula for $\mathbb{P}(X = k)$ looked not very fun…

More calls to senators

You work as a lobbyist and you've been asked to check to see the gender balance of the calls placed to Senate offices from your firm. The firm has placed 1000 calls over the last year. If the firm was randomly choosing senators (with replacement) each call, what numbers of women senators contacted would be more or less plausible?

- That math formula for $\mathbb{P}(X = k)$ looked not very fun...
- We can **simulate** data from this distribution using `rbinom()`.

More calls to senators

You work as a lobbyist and you've been asked to check to see the gender balance of the calls placed to Senate offices from your firm. The firm has placed 1000 calls over the last year. If the firm was randomly choosing senators (with replacement) each call, what numbers of women senators contacted would be more or less plausible?

- That math formula for $\mathbb{P}(X = k)$ looked not very fun. . .
- We can **simulate** data from this distribution using `rbinom()`.

```r
rbinom(n = 5, size = 1000, prob = 0.26)
```

More calls to senators

You work as a lobbyist and you've been asked to check to see the gender balance of the calls placed to Senate offices from your firm. The firm has placed 1000 calls over the last year. If the firm was randomly choosing senators (with replacement) each call, what numbers of women senators contacted would be more or less plausible?

- That math formula for $\mathbb{P}(X = k)$ looked not very fun...
- We can **simulate** data from this distribution using `rbinom()`.

```
rbinom(n = 5, size = 1000, prob = 0.26)
```

```
## [1] 269 247 259 268 266
```

Simulations

```r
sims <- 10000

draws <- rbinom(sims, size = 1000, prob = 0.26)
```

Simulations

```r
sims <- 10000

draws <- rbinom(sims, size = 1000, prob = 0.26)
```

```r
length(draws)
```

```
## [1] 10000
```

Simulations

```
sims <- 10000

draws <- rbinom(sims, size = 1000, prob = 0.26)
```

```
length(draws)
```

```
## [1] 10000
```

Simulations

```r
sims <- 10000

draws <- rbinom(sims, size = 1000, prob = 0.26)
```

```r
length(draws)
```

```
## [1] 10000
```

```r
mean(draws)
```

Simulations

```r
sims <- 10000

draws <- rbinom(sims, size = 1000, prob = 0.26)
```

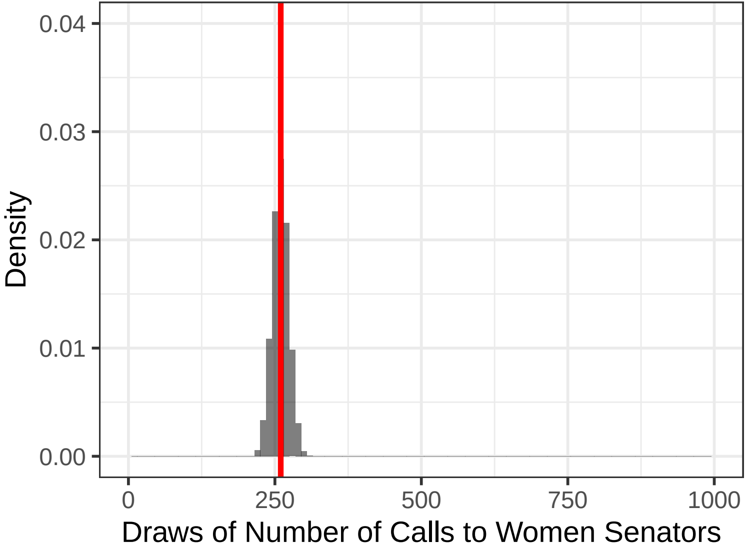
```r
length(draws)
```

```
## [1] 10000
```

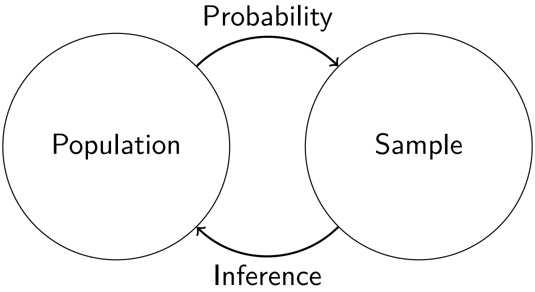
```r
mean(draws)
```

```
## [1] 260.0883
```

Simulations



Probability distributions



- We want to learn about the chance process that generated our data.
- More specifically: learn about the **distribution** of the r.v.s in our data.
  - What values of the r.v. are more or less likely?

Probability distribution

Probability distribution

- **Probability distributions** describe the uncertainty of a random variable

Probability distribution

- **Probability distributions** describe the uncertainty of a random variable
  - Functions that give the probability of different possible values of an r.v.

Probability distribution

- **Probability distributions** describe the uncertainty of a random variable
  - Functions that give the probability of different possible values of an r.v.
  - Why do we care? **Learning about populations from samples**

Probability distribution

- **Probability distributions** describe the uncertainty of a random variable
  - Functions that give the probability of different possible values of an r.v.
  - Why do we care? **Learning about populations from samples**
- Simple example: suppose we randomly sample a single US adult

Probability distribution

- **Probability distributions** describe the uncertainty of a random variable
  - Functions that give the probability of different possible values of an r.v.
  - Why do we care? **Learning about populations from samples**
- Simple example: suppose we randomly sample a single US adult
  - Let X be 1 if they support Trump, 0 otherwise.

Probability distribution

- **Probability distributions** describe the uncertainty of a random variable
  - Functions that give the probability of different possible values of an r.v.
  - Why do we care? **Learning about populations from samples**
- Simple example: suppose we randomly sample a single US adult
  - Let X be 1 if they support Trump, 0 otherwise.
  - X is Bernoulli with some probability p

Probability distribution

- **Probability distributions** describe the uncertainty of a random variable
  - Functions that give the probability of different possible values of an r.v.
  - Why do we care? **Learning about populations from samples**
- Simple example: suppose we randomly sample a single US adult
  - Let X be 1 if they support Trump, 0 otherwise.
  - X is Bernoulli with some probability p
  - Learning p would give us the probability a random adult supports Trump

Probability distribution

- **Probability distributions** describe the uncertainty of a random variable
  - Functions that give the probability of different possible values of an r.v.
  - Why do we care? **Learning about populations from samples**
- Simple example: suppose we randomly sample a single US adult
  - Let X be 1 if they support Trump, 0 otherwise.
  - X is Bernoulli with some probability p
  - Learning p would give us the probability a random adult supports Trump
- Multiple ways to represent the distribution

Probability distribution

- **Probability distributions** describe the uncertainty of a random variable
  - Functions that give the probability of different possible values of an r.v.
  - Why do we care? **Learning about populations from samples**
- Simple example: suppose we randomly sample a single US adult
  - Let X be 1 if they support Trump, 0 otherwise.
  - X is Bernoulli with some probability p
  - Learning p would give us the probability a random adult supports Trump
- Multiple ways to represent the distribution
  - Depends on what kind of r.v. we have.

Types of random variables

- **Discrete**: X can take a finite (countable infinite) number of values.

Types of random variables

- **Discrete**: X can take a finite (countable infinite) number of values.
  - Number of heads in 5 coin flips

Types of random variables

- **Discrete**: X can take a finite (countable infinite) number of values.
  - Number of heads in 5 coin flips
  - Sample senator is a woman ($X = 1$) or not ($X = 0$)

Types of random variables

- **Discrete**: X can take a finite (countable infinite) number of values.
  - Number of heads in 5 coin flips
  - Sample senator is a woman ($X = 1$) or not ($X = 0$)
  - Number of battle deaths in a civil war

Types of random variables

- **Discrete**: X can take a finite (countable infinite) number of values.
  - Number of heads in 5 coin flips
  - Sample senator is a woman ($X = 1$) or not ($X = 0$)
  - Number of battle deaths in a civil war
  - Number of journalists on a Signal chat with Executive branch officials

Types of random variables

- **Discrete**: X can take a finite (countable infinite) number of values.
  - Number of heads in 5 coin flips
  - Sample senator is a woman ($X = 1$) or not ($X = 0$)
  - Number of battle deaths in a civil war
  - Number of journalists on a Signal chat with Executive branch officials
- **Continuous**: X can take any real value (usually within an interval)

Types of random variables

- **Discrete**: X can take a finite (countable infinite) number of values.
  - Number of heads in 5 coin flips
  - Sample senator is a woman ($X = 1$) or not ($X = 0$)
  - Number of battle deaths in a civil war
  - Number of journalists on a Signal chat with Executive branch officials
- **Continuous**: X can take any real value (usually within an interval)
  - GDP per capita (average income) in a country

Types of random variables

- **Discrete**: X can take a finite (countable infinite) number of values.
  - Number of heads in 5 coin flips
  - Sample senator is a woman ($X = 1$) or not ($X = 0$)
  - Number of battle deaths in a civil war
  - Number of journalists on a Signal chat with Executive branch officials
- **Continuous**: X can take any real value (usually within an interval)
  - GDP per capita (average income) in a country
  - Share of population that approves of Trump

Types of random variables

- **Discrete**: X can take a finite (countable infinite) number of values.
  - Number of heads in 5 coin flips
  - Sample senator is a woman ($X = 1$) or not ($X = 0$)
  - Number of battle deaths in a civil war
  - Number of journalists on a Signal chat with Executive branch officials
- **Continuous**: X can take any real value (usually within an interval)
  - GDP per capita (average income) in a country
  - Share of population that approves of Trump
  - Amount of time spent on TikTok

Types of random variables

- **Discrete**: X can take a finite (countable infinite) number of values.
  - Number of heads in 5 coin flips
  - Sample senator is a woman ($X = 1$) or not ($X = 0$)
  - Number of battle deaths in a civil war
  - Number of journalists on a Signal chat with Executive branch officials
- **Continuous**: X can take any real value (usually within an interval)
  - GDP per capita (average income) in a country
  - Share of population that approves of Trump
  - Amount of time spent on TikTok
  - Amount of trade affected by US vs. Everyone Else trade war

Probability mass functions

- For discrete r.v.s: **probability mass function (PMF)**

Probability mass functions

- For discrete r.v.s: **probability mass function (PMF)**
  - Gives the probability of each possible value, $\mathbb{P}(X = k)$.

Probability mass functions

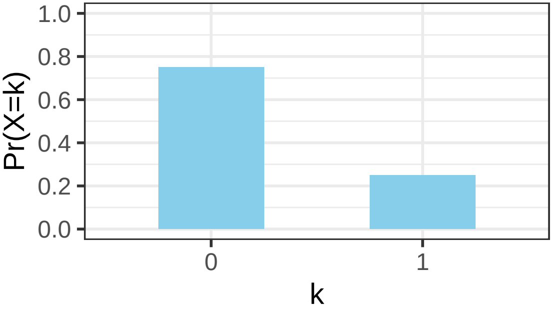
- For discrete r.v.s: **probability mass function (PMF)**
  - Gives the probability of each possible value, $\mathbb{P}(X = k)$.
  - Like a bar plot for the population shares of each value.

Probability mass functions

- For discrete r.v.s: **probability mass function (PMF)**
  - Gives the probability of each possible value, $\mathbb{P}(X = k)$.
  - Like a bar plot for the population shares of each value.
  - Here's the PMF for the Bernoulli of drawing a woman senator:

Probability mass functions

- For discrete r.v.s: **probability mass function (PMF)**
  - Gives the probability of each possible value, $\mathbb{P}(X = k)$.
  - Like a bar plot for the population shares of each value.
  - Here's the PMF for the Bernoulli of drawing a woman senator:



Binomial PMFs

Binomial PMFs

- PMFs expressed in mathematical formulas depending on **parameters**.

Binomial PMFs

- PMFs expressed in mathematical formulas depending on **parameters**.
  - Binomial with n draws and probability of "success" p:

Binomial PMFs

- PMFs expressed in mathematical formulas depending on **parameters**.
  - Binomial with n draws and probability of "success" p:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Binomial PMFs

- PMFs expressed in mathematical formulas depending on **parameters**.
  - Binomial with n draws and probability of "success" p:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- We'll almost always use R to calculate the PMF.

Binomial PMFs

- PMFs expressed in mathematical formulas depending on **parameters**.
  - Binomial with n draws and probability of "success" p:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- We'll almost always use R to calculate the PMF.
  - We can use the `dbinom()` function to calculate the PMF of a Binomial r.v.

Binomial PMFs

- PMFs expressed in mathematical formulas depending on **parameters**.
  - Binomial with n draws and probability of "success" p:

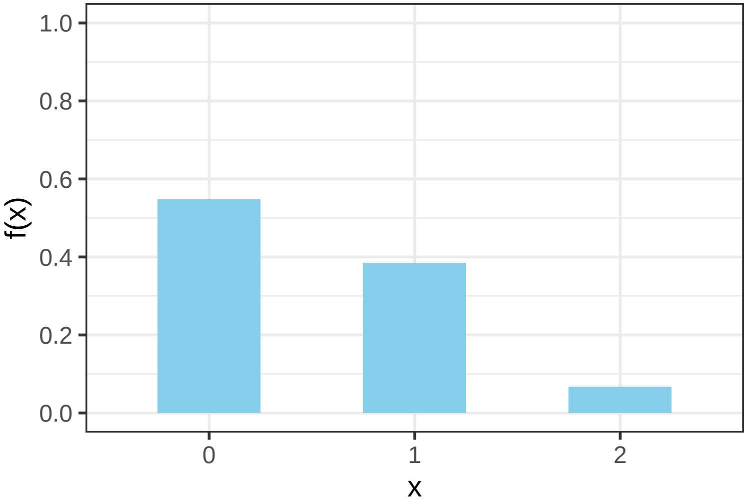
$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- We'll almost always use R to calculate the PMF.
  - We can use the `dbinom()` function to calculate the PMF of a Binomial r.v.

```r
dbinom(x = c(0, 1, 2), size = 2, prob = 26/100)
```

```
## [1] 0.5476 0.3848 0.0676
```

Binomial PMF plot



Probability density functions

- For continuous r.v.s, **probability density function (PDF)**

Probability density functions

- For continuous r.v.s, **probability density function (PDF)**
  - Gives density of probability around a given point

Probability density functions

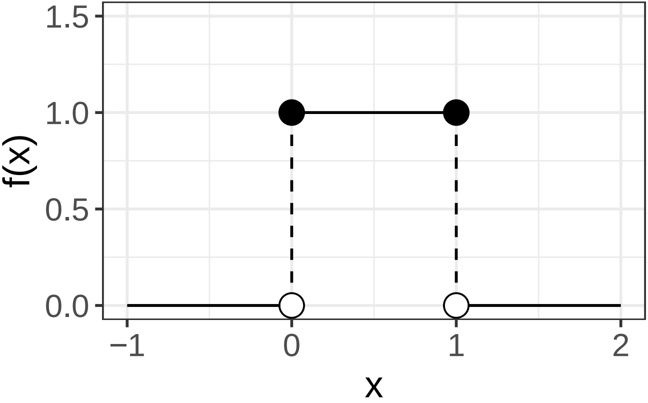
- For continuous r.v.s, **probability density function (PDF)**
  - Gives density of probability around a given point
  - Like an "infinite" histogram $\rightsquigarrow$ so many bins that things look smooth

Probability density functions

- For continuous r.v.s, **probability density function (PDF)**
  - Gives density of probability around a given point
  - Like an "infinite" histogram $\rightsquigarrow$ so many bins that things look smooth
  - Area under the curve $=$ prob. of some interval

Probability density functions

- For continuous r.v.s, **probability density function (PDF)**
  - Gives density of probability around a given point
  - Like an "infinite" histogram $\rightsquigarrow$ so many bins that things look smooth
  - Area under the curve = prob. of some interval



Cumulative distribution function

- **Cumulative distribution function (CDF)**: $F_X(k) = \mathbb{P}(X \leq k)$

Cumulative distribution function

- **Cumulative distribution function (CDF)**: $F_X(k) = \mathbb{P}(X \leq k)$
  - Returns the probability of X being at k or lower.

Cumulative distribution function

- **Cumulative distribution function (CDF)**: $F_X(k) = \mathbb{P}(X \leq k)$
  - Returns the probability of X being at k or lower.
  - Area under the density for a continuous r.v.

Cumulative distribution function

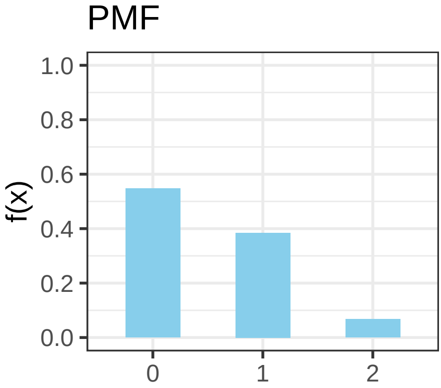
- **Cumulative distribution function (CDF)**: $F_X(k) = \mathbb{P}(X \leq k)$
  - Returns the probability of X being at k or lower.
  - Area under the density for a continuous r.v.
  - Never decreasing as k gets bigger.

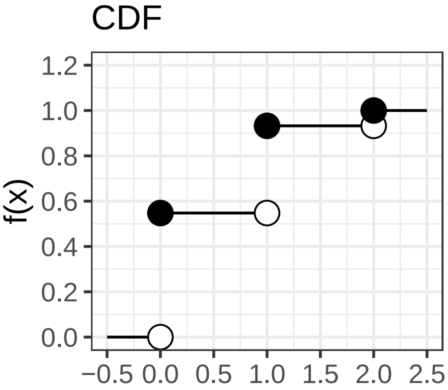
Cumulative distribution function

- **Cumulative distribution function (CDF)**: $F_X(k) = \mathbb{P}(X \leq k)$
  - Returns the probability of X being at k or lower.
  - Area under the density for a continuous r.v.
  - Never decreasing as k gets bigger.
  - Drawing two women senators example:

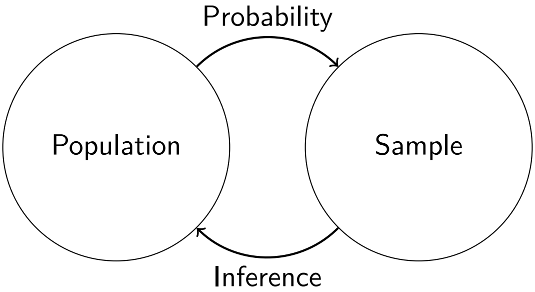
Cumulative distribution function

- **Cumulative distribution function (CDF)**: $F_X(k) = \mathbb{P}(X \leq k)$
  - Returns the probability of X being at k or lower.
  - Area under the density for a continuous r.v.
  - Never decreasing as k gets bigger.
  - Drawing two women senators example:





Let's recall our goal again:



- We want to learn about the chance process that generated our data.
- Last time: entire probability distributions. Is there something simpler?

Expectation, Variance, and Sample Means

How can we summarize distributions?

How can we summarize distributions?

- Two numerical summaries of the distribution are useful.

How can we summarize distributions?

- Two numerical summaries of the distribution are useful.
  - **Mean/expectation**: where the distribution is.

How can we summarize distributions?

- Two numerical summaries of the distribution are useful.
  - **Mean/expectation**: where the distribution is.
  - **Variance/standard deviation**: how spread out the distribution is around the center.

How can we summarize distributions?

- Two numerical summaries of the distribution are useful.
  - **Mean/expectation**: where the distribution is.
  - **Variance/standard deviation**: how spread out the distribution is around the center.
  - These are **population parameters** so we don't get to observe them.

How can we summarize distributions?

- Two numerical summaries of the distribution are useful.
  - **Mean/expectation**: where the distribution is.
  - **Variance/standard deviation**: how spread out the distribution is around the center.
  - These are **population parameters** so we don't get to observe them.
    - We won't get to observe them...

How can we summarize distributions?

- Two numerical summaries of the distribution are useful.
  - **Mean/expectation**: where the distribution is.
  - **Variance/standard deviation**: how spread out the distribution is around the center.
  - These are **population parameters** so we don't get to observe them.
    - We won't get to observe them...
    - but we'll use our sample to learn about them.

Two ways to calculate averages

- Calculate the average of: $\{1,1,1,3,4,5,5\}$

Two ways to calculate averages

- Calculate the average of: $\{1,1,1,3,4,5,5\}$

$$\frac{1+1+1+3+4+4+5+5}{8} = 3$$

Two ways to calculate averages

- Calculate the average of: $\{1,1,1,3,4,5,5\}$

$$\frac{1+1+1+3+4+4+5+5}{8} = 3$$

- Alternative way to calculate average based on **frequency weights**:

$$1 \times \frac{3}{8} \times 3 \times \frac{1}{8} \times 4 \times \frac{2}{8} \times 5 \times \frac{2}{8} = 3$$

Two ways to calculate averages

- Calculate the average of: $\{1,1,1,3,4,5,5\}$

$$\frac{1+1+1+3+4+4+5+5}{8} = 3$$

- Alternative way to calculate average based on **frequency weights**:

$$1 \times \frac{3}{8} \times 3 \times \frac{1}{8} \times 4 \times \frac{2}{8} \times 5 \times \frac{2}{8} = 3$$

- Each value times how often that value occurs in the data

Two ways to calculate averages

- Calculate the average of: $\{1,1,1,3,4,5,5\}$

$$\frac{1+1+1+3+4+4+5+5}{8} = 3$$

- Alternative way to calculate average based on **frequency weights**:

$$1 \times \frac{3}{8} \times 3 \times \frac{1}{8} \times 4 \times \frac{2}{8} \times 5 \times \frac{2}{8} = 3$$

- Each value times how often that value occurs in the data
- We'll use this intuition to create an average/mean for r.v.s.

Expectation

- We write $\mathbb{E}(X)$ for the **mean** of an r.v. X.

Expectation

- We write $\mathbb{E}(X)$ for the **mean** of an r.v. X.
- For discrete $X \in \{x_1, x_2, ..., x_k\}$ with k levels:

$$\mathbb{E}[X] = \sum_{j=1}^{k} x_j \mathbb{P}(X = x_j)$$

Expectation

- We write $\mathbb{E}(X)$ for the **mean** of an r.v. X.
- For discrete $X \in \{x_1, x_2, ..., x_k\}$ with k levels:

$$\mathbb{E}[X] = \sum_{j=1}^{k} x_j \mathbb{P}(X = x_j)$$

- Weighted average of the values of the r.v. weighted by the probability of each value occurring.

Expectation

- We write $\mathbb{E}(X)$ for the **mean** of an r.v. X.
- For discrete $X \in \{x_1, x_2, \ldots, x_k\}$ with k levels:

$$\mathbb{E}[X] = \sum_{j=1}^{k} x_j \mathbb{P}(X = x_j)$$

- Weighted average of the values of the r.v. weighted by the probability of each value occurring.
- If X is age of randomly selected registered voter, then $\mathbb{E}(X)$ is the average age in the population of registered voters.

Expectation

- We write $\mathbb{E}(X)$ for the **mean** of an r.v. X.
- For discrete $X \in \{x_1, x_2, ..., x_k\}$ with k levels:

$$\mathbb{E}[X] = \sum_{j=1}^{k} x_j \mathbb{P}(X = x_j)$$

- Weighted average of the values of the r.v. weighted by the probability of each value occurring.
- If X is age of randomly selected registered voter, then $\mathbb{E}(X)$ is the average age in the population of registered voters.
- Notation notes:

Expectation

- We write $\mathbb{E}(X)$ for the **mean** of an r.v. X.
- For discrete $X \in \{x_1, x_2, ..., x_k\}$ with k levels:

$$\mathbb{E}[X] = \sum_{j=1}^{k} x_j \mathbb{P}(X = x_j)$$

- Weighted average of the values of the r.v. weighted by the probability of each value occurring.
- If X is age of randomly selected registered voter, then $\mathbb{E}(X)$ is the average age in the population of registered voters.
- Notation notes:
  - Lots of other ways to refer to this: **expectation** or **expected value**

Expectation

- We write $\mathbb{E}(X)$ for the **mean** of an r.v. X.
- For discrete $X \in \{x_1, x_2, ..., x_k\}$ with k levels:

$$\mathbb{E}[X] = \sum_{j=1}^{k} x_j \mathbb{P}(X = x_j)$$

- Weighted average of the values of the r.v. weighted by the probability of each value occurring.
- If X is age of randomly selected registered voter, then $\mathbb{E}(X)$ is the average age in the population of registered voters.
- Notation notes:
  - Lots of other ways to refer to this: **expectation** or **expected value**
  - Often call the **population mean** to distinguish from the sample mean

Properties of the expected value

- We use properties of $\mathbb{E}(X)$ to avoid using the formula every time.

Properties of the expected value

- We use properties of $\mathbb{E}(X)$ to avoid using the formula every time.
- Let X and Y be r.v.s and a and b be constants

Properties of the expected value

- We use properties of $\mathbb{E}(X)$ to avoid using the formula every time.
- Let X and Y be r.v.s and a and b be constants

1. $\mathbb{E}(a) = a$

Properties of the expected value

- We use properties of $\mathbb{E}(X)$ to avoid using the formula every time.
- Let X and Y be r.v.s and a and b be constants

1. $\mathbb{E}(a) = a$
    - constants don't vary

Properties of the expected value

- We use properties of $\mathbb{E}(X)$ to avoid using the formula every time.
- Let X and Y be r.v.s and a and b be constants

1. $\mathbb{E}(a) = a$
   - constants don't vary
2. $\mathbb{E}(aX) = a\mathbb{E}(X)$

Properties of the expected value

- We use properties of $\mathbb{E}(X)$ to avoid using the formula every time.
- Let X and Y be r.v.s and a and b be constants

1. $\mathbb{E}(a) = a$
   - constants don't vary
2. $\mathbb{E}(aX) = a\mathbb{E}(X)$
   - Suppose X is income in dollars, income in \$10k is just: $X/1000$

Properties of the expected value

- We use properties of $\mathbb{E}(X)$ to avoid using the formula every time.
- Let X and Y be r.v.s and a and b be constants

1. $\mathbb{E}(a) = a$
   - constants don't vary
2. $\mathbb{E}(aX) = a\mathbb{E}(X)$
   - Suppose X is income in dollars, income in \$10k is just: $X/1000$
   - Mean of this new variable is mean of income in dollars divided by 10,000

Properties of the expected value

- We use properties of $\mathbb{E}(X)$ to avoid using the formula every time.
- Let X and Y be r.v.s and a and b be constants

1. $\mathbb{E}(a) = a$
   - constants don't vary
2. $\mathbb{E}(aX) = a\mathbb{E}(X)$
   - Suppose X is income in dollars, income in \$10k is just: $X/1000$
   - Mean of this new variable is mean of income in dollars divided by 10,000
3. $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$

Properties of the expected value

- We use properties of $\mathbb{E}(X)$ to avoid using the formula every time.
- Let X and Y be r.v.s and a and b be constants

1. $\mathbb{E}(a) = a$
   - constants don't vary
2. $\mathbb{E}(aX) = a\mathbb{E}(X)$
   - Suppose X is income in dollars, income in \$10k is just: $X/1000$
   - Mean of this new variable is mean of income in dollars divided by 10,000
3. $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$
   - Expectations can be distributed across sums

Properties of the expected value

- We use properties of $\mathbb{E}(X)$ to avoid using the formula every time.
- Let X and Y be r.v.s and a and b be constants

1. $\mathbb{E}(a) = a$
   - constants don't vary
2. $\mathbb{E}(aX) = a\mathbb{E}(X)$
   - Suppose X is income in dollars, income in \$10k is just: $X/1000$
   - Mean of this new variable is mean of income in dollars divided by 10,000
3. $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$
   - Expectations can be distributed across sums
   - X is partner 1's income, Y is partner 2's income

Properties of the expected value

- We use properties of $\mathbb{E}(X)$ to avoid using the formula every time.
- Let X and Y be r.v.s and a and b be constants

1. $\mathbb{E}(a) = a$
    - constants don't vary
2. $\mathbb{E}(aX) = a\mathbb{E}(X)$
    - Suppose X is income in dollars, income in \$10k is just: $X/1000$
    - Mean of this new variable is mean of income in dollars divided by 10,000
3. $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$
    - Expectations can be distributed across sums
    - X is partner 1's income, Y is partner 2's income
    - Mean household income is the sum of each partner's income

Variance

- The **variance** measures the spread of the distribution:

Variance

- The **variance** measures the spread of the distribution:

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

Variance

- The **variance** measures the spread of the distribution:

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

- Weighted average of the squared distances from the mean.

Variance

- The **variance** measures the spread of the distribution:

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

- Weighted average of the squared distances from the mean.
    - Larger deviations ($+$ or $-$) $\rightsquigarrow$ higher variance

Variance

- The **variance** measures the spread of the distribution:

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

- Weighted average of the squared distances from the mean.
  - Larger deviations ($+$ or $-$) $\rightsquigarrow$ higher variance
- If X is the age of a randomly selected registered voter, $\mathbb{V}[X]$ is the usual sample variance of age in the population

Variance

- The **variance** measures the spread of the distribution:

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

- Weighted average of the squared distances from the mean.
  - Larger deviations ($+$ or $-$) $\rightsquigarrow$ higher variance
- If X is the age of a randomly selected registered voter, $\mathbb{V}[X]$ is the usual sample variance of age in the population
  - Sometimes called **population variance** to contrast with sample variance.

Variance

- The **variance** measures the spread of the distribution:

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

- Weighted average of the squared distances from the mean.
  - Larger deviations ($+$ or $-$) $\rightsquigarrow$ higher variance
- If X is the age of a randomly selected registered voter, $\mathbb{V}[X]$ is the usual sample variance of age in the population
  - Sometimes called **population variance** to contrast with sample variance.
- **Standard deviation**: square root of the variance: $SD(X) = \sqrt{\mathbb{V}[X]}$.

Variance

- The **variance** measures the spread of the distribution:

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

- Weighted average of the squared distances from the mean.
  - Larger deviations ($+$ or $-$) $\rightsquigarrow$ higher variance
- If X is the age of a randomly selected registered voter, $\mathbb{V}[X]$ is the usual sample variance of age in the population
  - Sometimes called **population variance** to contrast with sample variance.
- **Standard deviation**: square root of the variance: $SD(X) = \sqrt{\mathbb{V}[X]}$.
  - Useful because it's on the scale of the original variable.

Properties of variances

Properties of variances

- Some properties of variance useful for calculation

Properties of variances

- Some properties of variance useful for calculation

1. If b is a constant, then $\mathbb{V}[b] = 0$.

Properties of variances

- Some properties of variance useful for calculation

1. If b is a constant, then $\mathbb{V}[b] = 0$.
2. If a and b are constants, $\mathbb{V}[aX + b] = a^2\mathbb{V}[X]$.

Properties of variances

- Some properties of variance useful for calculation

1. If b is a constant, then $\mathbb{V}[b] = 0$.
2. If a and b are constants, $\mathbb{V}[aX + b] = a^2\mathbb{V}[X]$.
3. In general, $\mathbb{V}[X + Y] \neq \mathbb{V}[X] + \mathbb{V}[Y]$

Properties of variances

- Some properties of variance useful for calculation

1. If b is a constant, then $\mathbb{V}[b] = 0$.
2. If a and b are constants, $\mathbb{V}[aX + b] = a^2\mathbb{V}[X]$.
3. In general, $\mathbb{V}[X + Y] \neq \mathbb{V}[X] + \mathbb{V}[Y]$
   - If X and Y are independent, then $\mathbb{V}[X + Y] = \mathbb{V}[X] + \mathbb{V}[Y]$

Sums and means are random variables

Sums and means are random variables

- If X_1 and X_2 are r.v.s, then $X_1 + X_2$ is an r.v.

Sums and means are random variables

- If X_1 and X_2 are r.v.s, then $X_1 + X_2$ is an r.v.
  - Has a mean $\mathbb{E}[X_1 + X_2]$ and a variance $\mathbb{V}[X_1 + X_2]$

Sums and means are random variables

- If X_1 and X_2 are r.v.s, then $X_1 + X_2$ is an r.v.
  - Has a mean $\mathbb{E}[X_1 + X_2]$ and a variance $\mathbb{V}[X_1 + X_2]$
- The **sample mean** is a function of sums and so it is an r.v. too:

Sums and means are random variables

- If X_1 and X_2 are r.v.s, then $X_1 + X_2$ is an r.v.
  - Has a mean $\mathbb{E}[X_1 + X_2]$ and a variance $\mathbb{V}[X_1 + X_2]$
- The **sample mean** is a function of sums and so it is an r.v. too:

$$\bar{X} = \frac{X_1 + X_2}{2}$$

Sums and means are random variables

- If X_1 and X_2 are r.v.s, then $X_1 + X_2$ is an r.v.
  - Has a mean $\mathbb{E}[X_1 + X_2]$ and a variance $\mathbb{V}[X_1 + X_2]$
- The **sample mean** is a function of sums and so it is an r.v. too:

$$\bar{X} = \frac{X_1 + X_2}{2}$$

- Example: the average of two randomly selected respondents.

Independent and identical r.v.s

Independent and identical r.v.s

- **Independent and identically distributed** r.v.s, $X_1, ..., X_n$

Independent and identical r.v.s

- **Independent and identically distributed** r.v.s, $X_1, ..., X_n$
  - Random sample of n respondents on a survey question
  - Written "i.i.d."
- **Independent**: value that X_i takes doesn't affect distribution of X_j

Independent and identical r.v.s

- **Independent and identically distributed** r.v.s, $X_1, ..., X_n$
  - Random sample of n respondents on a survey question
  - Written "i.i.d."
- **Independent**: value that X_i takes doesn't affect distribution of X_j
- **Identically distributed**: distribution of X_i is the same for all i

Independent and identical r.v.s

- **Independent and identically distributed** r.v.s, $X_1, ..., X_n$
  - Random sample of n respondents on a survey question
  - Written "i.i.d."
- **Independent**: value that X_i takes doesn't affect distribution of X_j
- **Identically distributed**: distribution of X_i is the same for all i
  - $\mathbb{E}(X_1) = \mathbb{E}(X_2) = ... = \mathbb{E}(X_n) = \mu$

Independent and identical r.v.s

- **Independent and identically distributed** r.v.s, $X_1, ..., X_n$
  - Random sample of n respondents on a survey question
  - Written "i.i.d."
- **Independent**: value that X_i takes doesn't affect distribution of X_j
- **Identically distributed**: distribution of X_i is the same for all i
  - $\mathbb{E}(X_1) = \mathbb{E}(X_2) = ... = \mathbb{E}(X_n) = \mu$
  - $\mathbb{V}(X_1) = \mathbb{V}(X_2) = ... = \mathbb{V}(X_n) = \sigma^2$

Distribution of the sample mean

- **Sample mean** of i.i.d. random variables:

Distribution of the sample mean

- **Sample mean** of i.i.d. random variables:

$$\bar{X}_n = \frac{X_1 + X_2 + ... + X_n}{n}$$

Distribution of the sample mean

- **Sample mean** of i.i.d. random variables:

$$\bar{X}_n = \frac{X_1 + X_2 + ... + X_n}{n}$$

- $\bar{X}_n$ is a random variable, what is its distibution?

Distribution of the sample mean

- **Sample mean** of i.i.d. random variables:

$$\bar{X}_n = \frac{X_1 + X_2 + ... + X_n}{n}$$

- $\bar{X}_n$ is a random variable, what is its distibution?
  - What is the expectation of this distribution, $\mathbb{E}[\bar{X}_n]$

Distribution of the sample mean

- **Sample mean** of i.i.d. random variables:

$$\bar{X}_n = \frac{X_1 + X_2 + ... + X_n}{n}$$

- $\bar{X}_n$ is a random variable, what is its distibution?
    - What is the expectation of this distribution, $\mathbb{E}[\bar{X}_n]$
    - What is the variance of this distribution, $\mathbb{V}[\bar{X}_n]$


MSS
fficer
Pete
Hegseth
Jeffrey
Goldberg
GRU
Officer
Michae
Waltz
JD
Vance

Properties of the sample mean

Properties of the sample mean

Definition (Mean and variance of the sample mean)

Suppose that $X_1, ..., X_n$ are i.i.d. r.v.s with $\mathbb{E}[X_i] = \mu$ and $\mathbb{V}[X_i] = \sigma^2$

$$\mathbb{E}[\bar{X}_n] = \mu \qquad \mathbb{V}[\bar{X}_n] = \frac{\sigma^2}{n}$$

Properties of the sample mean

Definition (Mean and variance of the sample mean)

Suppose that $X_1, \ldots, X_n$ are i.i.d. r.v.s with $\mathbb{E}[X_i] = \mu$ and $\mathbb{V}[X_i] = \sigma^2$

$$\mathbb{E}[\bar{X}_n] = \mu \qquad \mathbb{V}[\bar{X}_n] = \frac{\sigma^2}{n}$$

- Key insights:

Properties of the sample mean

Definition (Mean and variance of the sample mean)

Suppose that $X_1, ..., X_n$ are i.i.d. r.v.s with $\mathbb{E}[X_i] = \mu$ and $\mathbb{V}[X_i] = \sigma^2$

$$\mathbb{E}[\bar{X}_n] = \mu \qquad \mathbb{V}[\bar{X}_n] = \frac{\sigma^2}{n}$$

- Key insights:
  - Sample mean is on average equal to the population mean

Properties of the sample mean

Definition (Mean and variance of the sample mean)

Suppose that $X_1, ..., X_n$ are i.i.d. r.v.s with $\mathbb{E}[X_i] = \mu$ and $\mathbb{V}[X_i] = \sigma^2$

$$\mathbb{E}[\bar{X}_n] = \mu \qquad \mathbb{V}[\bar{X}_n] = \frac{\sigma^2}{n}$$

- Key insights:
  - Sample mean is on average equal to the population mean
  - Variance of $\bar{X}_n$ depends on the population variance of X_i and the sample size

Properties of the sample mean

Definition (Mean and variance of the sample mean)

Suppose that $X_1, \ldots, X_n$ are i.i.d. r.v.s with $\mathbb{E}[X_i] = \mu$ and $\mathbb{V}[X_i] = \sigma^2$

$$\mathbb{E}[\bar{X}_n] = \mu \qquad \mathbb{V}[\bar{X}_n] = \frac{\sigma^2}{n}$$

- Key insights:
  - Sample mean is on average equal to the population mean
  - Variance of $\bar{X}_n$ depends on the population variance of X_i and the sample size
- Standard deviation of the sample mean is called its **standard error**:

$$SE = \sqrt{\mathbb{V}[\bar{X}_n]} = \frac{\sigma}{\sqrt{n}}$$

Final probability lesson! Large Sample Theorems and the Normal Distribution

Final probability lesson! Large Sample Theorems and the Normal Distribution

Trump's tariff threats are hurting your job prospects

By Matt Egan, CNN

5 minute read · Published 7:00 AM EDT, Wed March 26, 2025

4 comments

New York (CNN) — One in four US businesses has scaled back their hiring plans because of the turmoil unleashed by President Donald Trump's trade war, according to a survey of chief financial officers released Wednesday.

The quarterly survey, conducted by Duke University and the Federal Reserve Banks of Richmond and Atlanta, found a significant drop in CFO economic optimism as they grapple with the fog of the trade war. Almost all of their post-election increase in optimism faded.

The tariff chaos has caused a deer-in-headlights moment for many firms. Executives don't know how high tariffs will go, what products will be affected, or how long they'll stay in place. Faced with deep uncertainty, some businesses are pulling back.

- Source: https://www.cnn.com/2025/03/26/economy/trump-tariffs-trade-war-jobs-economy/index.html

Savings Data

- See https://www.piie.com/blogs/realtime-economics/2025/lets-stop-trade-deficit-blame-game
- `savings.csv`: data on **all** countries domestic savings as a share of GDP (from World Development Indicators at the World Bank)

Name	Description
`cntry_cd`	3-character ISO code for country
`country`	country name
`year`	year
`save_gdp`	gross savings (the difference between disposable income and consumption) as a share of GDP

Load savings data

```
savings <- read_csv("../data/savings.csv")
head(savings)
```

```
## # A tibble: 6 x 5
##     ...1 cntry_cd  year save_gdp country
##    <dbl> <chr>    <dbl>    <dbl> <chr>
## 1      1 ABW       1960       NA Aruba
## 2      2 ABW       1961       NA Aruba
## 3      3 ABW       1962       NA Aruba
## 4      4 ABW       1963       NA Aruba
## 5      5 ABW       1964       NA Aruba
## 6      6 ABW       1965       NA Aruba
```

Large random samples

- In real data, we will have a set of n measurements on a variable:

$$X_1, X_2, ..., X_n$$

Large random samples

- In real data, we will have a set of n measurements on a variable:

$$X_1, X_2, ..., X_n$$

- X_1 is the savings/gdp of the randomly selected country

Large random samples

- In real data, we will have a set of n measurements on a variable:

$$X_1, X_2, ..., X_n$$

- X_1 is the savings/gdp of the randomly selected country
- X_2 is the savings/gdp of the second randomly selected country, etc.

Large random samples

- In real data, we will have a set of n measurements on a variable:

$$X_1, X_2, ..., X_n$$

- X_1 is the savings/gdp of the randomly selected country
- X_2 is the savings/gdp of the second randomly selected country, etc.
- What are the properties of the sample mean of these measurements?

Large random samples

- In real data, we will have a set of n measurements on a variable:

$$X_1, X_2, ..., X_n$$

- X_1 is the savings/gdp of the randomly selected country
- X_2 is the savings/gdp of the second randomly selected country, etc.
- What are the properties of the sample mean of these measurements?
  - Expectation: $\mathbb{E}(\bar{X}) = \mathbb{E}[X] = \mu$

Large random samples

- In real data, we will have a set of n measurements on a variable:

$$X_1, X_2, ..., X_n$$

- X_1 is the savings/gdp of the randomly selected country
- X_2 is the savings/gdp of the second randomly selected country, etc.
- What are the properties of the sample mean of these measurements?
  - Expectation: $\mathbb{E}(\bar{X}) = \mathbb{E}[X] = \mu$
  - Variance: $\mathbb{V}(\bar{X}) = \mathbb{V}(X_i)/n = \sigma_X^2/n$

Large random samples

- In real data, we will have a set of n measurements on a variable:

$$X_1, X_2, ..., X_n$$

- X_1 is the savings/gdp of the randomly selected country
- X_2 is the savings/gdp of the second randomly selected country, etc.
- What are the properties of the sample mean of these measurements?
  - Expectation: $\mathbb{E}(\bar{X}) = \mathbb{E}[X] = \mu$
  - Variance: $\mathbb{V}(\bar{X}) = \mathbb{V}(X_i)/n = \sigma_X^2/n$
  - Valid for any sample size!

Large random samples

- In real data, we will have a set of n measurements on a variable:

$$X_1, X_2, ..., X_n$$

- X_1 is the savings/gdp of the randomly selected country
- X_2 is the savings/gdp of the second randomly selected country, etc.
- What are the properties of the sample mean of these measurements?
  - Expectation: $\mathbb{E}(\bar{X}) = \mathbb{E}[X] = \mu$
  - Variance: $\mathbb{V}(\bar{X}) = \mathbb{V}(X_i)/n = \sigma_X^2/n$
  - Valid for any sample size!
- **Asymptotics**: what can we learn as n gets big?

Christian Schneider
@Schneider_CM

New phone, Houthis

Law of large numbers

Law of large numbers

Definition (Law of large numbers)

Let $X_1, ..., X_n$ be i.i.d. r.v.s with mean μ and finite variance σ^2. Then, $\bar{X}_n$ converges to μ as n gets large.

Law of large numbers

Definition (Law of large numbers)

Let $X_1, ..., X_n$ be i.i.d. r.v.s with mean μ and finite variance σ^2. Then, $\bar{X}_n$ converges to μ as n gets large.

- Probability of $\bar{X}_n$ being "far away" from μ goes to 0 as n gets big.

Law of large numbers

Definition (Law of large numbers)

Let $X_1, ..., X_n$ be i.i.d. r.v.s with mean μ and finite variance σ^2. Then, $\bar{X}_n$ converges to μ as n gets large.

- Probability of $\bar{X}_n$ being "far away" from μ goes to 0 as n gets big.
- The distribution of sample mean "collapses" to population mean.

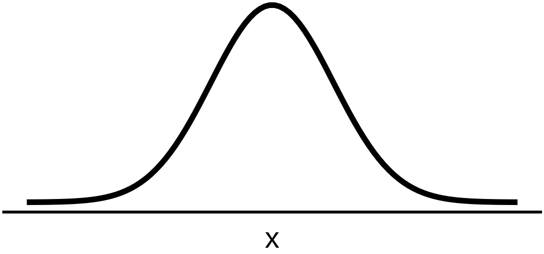
Law of large numbers

Definition (Law of large numbers)

Let $X_1, ..., X_n$ be i.i.d. r.v.s with mean μ and finite variance σ^2. Then, $\bar{X}_n$ converges to μ as n gets large.

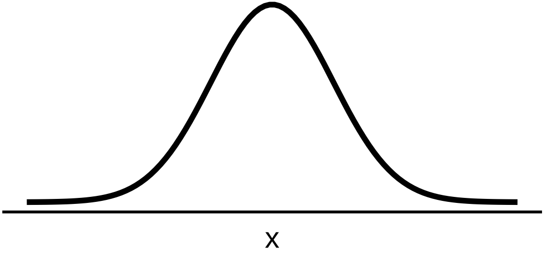
- Probability of $\bar{X}_n$ being "far away" from μ goes to 0 as n gets big.
- The distribution of sample mean "collapses" to population mean.
- Can see this from the variance of $\bar{X}_n$: $\mathbb{V}[X]/n$

Normal random variable



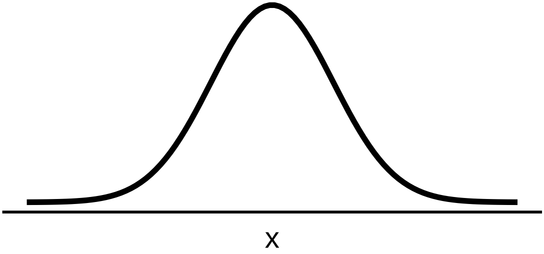
- A **normal distribution** has a PDF that is the class "bell-shaped" curve.

Normal random variable



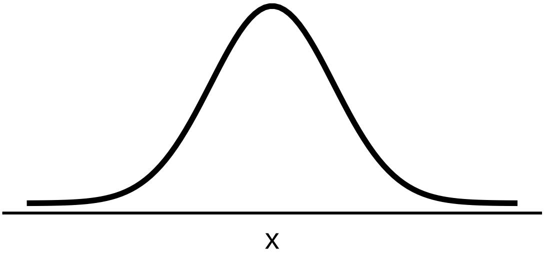
- A **normal distribution** has a PDF that is the class "bell-shaped" curve.
  - Extremely ubiquitous in statistics.

Normal random variable



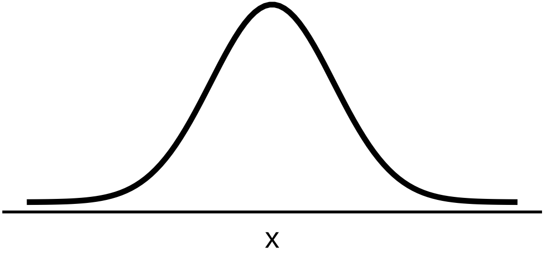
- A **normal distribution** has a PDF that is the class "bell-shaped" curve.
  - Extremely ubiquitous in statistics.
  - An r.v. is more likely to be in the center, rather than the tails.

Normal random variable



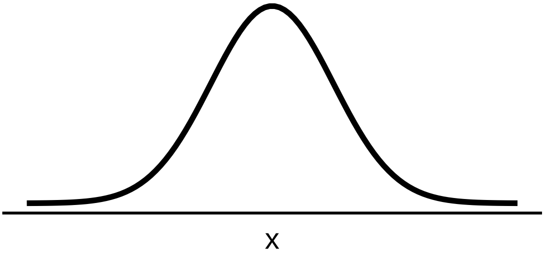
- A **normal distribution** has a PDF that is the class "bell-shaped" curve.
  - Extremely ubiquitous in statistics.
  - An r.v. is more likely to be in the center, rather than the tails.
- Three key properties of this PDF:

Normal random variable



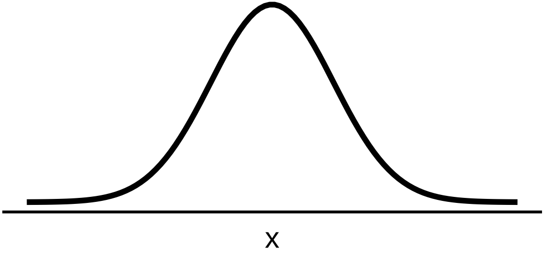
- A **normal distribution** has a PDF that is the class "bell-shaped" curve.
  - Extremely ubiquitous in statistics.
  - An r.v. is more likely to be in the center, rather than the tails.
- Three key properties of this PDF:
  - **Unimodal**: one peak at the mean.

Normal random variable



- A **normal distribution** has a PDF that is the class "bell-shaped" curve.
  - Extremely ubiquitous in statistics.
  - An r.v. is more likely to be in the center, rather than the tails.
- Three key properties of this PDF:
  - **Unimodal**: one peak at the mean.
  - **Symmetric** around the mean.

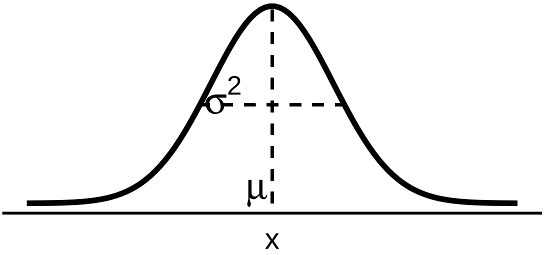
Normal random variable



- A **normal distribution** has a PDF that is the class "bell-shaped" curve.
  - Extremely ubiquitous in statistics.
  - An r.v. is more likely to be in the center, rather than the tails.
- Three key properties of this PDF:
  - **Unimodal**: one peak at the mean.
  - **Symmetric** around the mean.
  - **Everywhere positive**: any real value can possibly occur.

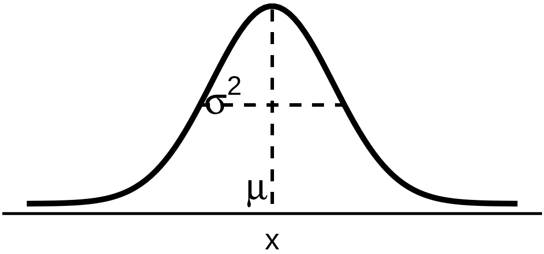

OPINION
MICHELLE COTTLE
Is the Cure to Male Loneliness
starting a group text to bomb the houthis?
July 19, 2023
Benjamin Marra
imgflip.com

Normal distribution



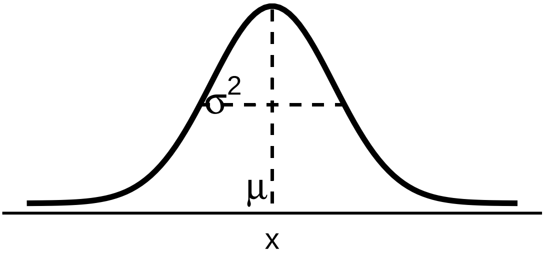
- A normal distribution can be affected by two values:

Normal distribution



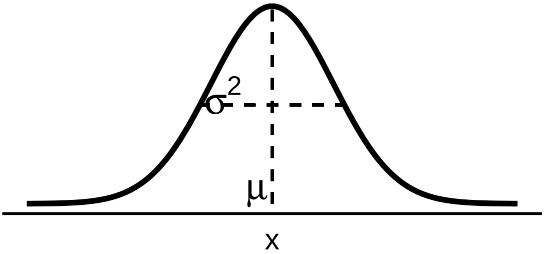
- A normal distribution can be affected by two values:
  - **mean/expected value** usually written as μ

Normal distribution



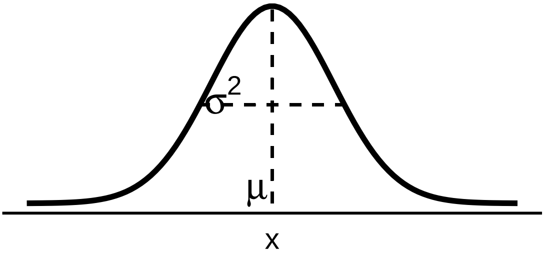
- A normal distribution can be affected by two values:
  - **mean/expected value** usually written as μ
  - **variance** written as σ^2 (standard deviation is σ)

Normal distribution



- A normal distribution can be affected by two values:
  - **mean/expected value** usually written as μ
  - **variance** written as σ^2 (standard deviation is σ)
  - Written $X \sim N(\mu, \sigma^2)$

Normal distribution



- A normal distribution can be affected by two values:
  - **mean/expected value** usually written as μ
  - **variance** written as σ^2 (standard deviation is σ)
  - Written $X \sim N(\mu, \sigma^2)$
- **Standard normal distribution**: mean 0 and standard deviation 1.

Recentering and scaling the normal

- How do transformations of a normal work?

Recentering and scaling the normal

- How do transformations of a normal work?
- Let $X \sim N(\mu, \sigma^2)$ and c be a constant.

Recentering and scaling the normal

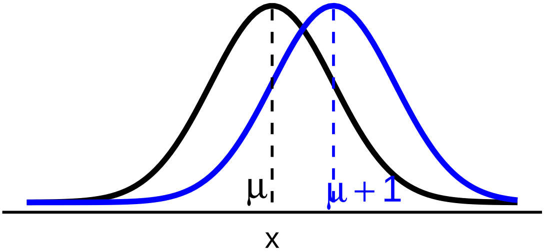
- How do transformations of a normal work?
- Let $X \sim N(\mu, \sigma^2)$ and c be a constant.
- If $Z = X + C$, then $Z \sim N(\mu + c, \sigma^2)$

Recentering and scaling the normal

- How do transformations of a normal work?
- Let $X \sim N(\mu, \sigma^2)$ and c be a constant.
- If $Z = X + C$, then $Z \sim N(\mu + c, \sigma^2)$
- Intuition: adding a constant to a normal shifts the distribution by that constant.

Recentering and scaling the normal

- How do transformations of a normal work?
- Let $X \sim N(\mu, \sigma^2)$ and c be a constant.
- If $Z = X + C$, then $Z \sim N(\mu + c, \sigma^2)$
- Intuition: adding a constant to a normal shifts the distribution by that constant.



Recentering and scaling the normal

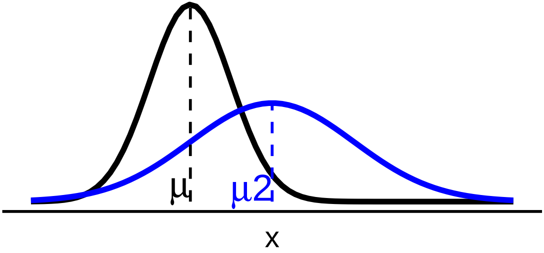
- Let $X \sim N(\mu, \sigma^2)$ and c be a constant

Recentering and scaling the normal

- Let $X \sim N(\mu, \sigma^2)$ and c be a constant
- If $Z = cX$, then $Z \sim N(c\mu, (c\sigma)^2)$

Recentering and scaling the normal

- Let $X \sim N(\mu, \sigma^2)$ and c be a constant
- If $Z = cX$, then $Z \sim N(c\mu, (c\sigma)^2)$
- Intuition: multiplying a normal by a constant scales the mean and the variance.



Z-scores of normals

- These facts imply the **z-score** of a normal variable is a standard normal:

Z-scores of normals

- These facts imply the **z-score** of a normal variable is a standard normal:

$$z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

- Subtract the mean and divide by the SD $\rightsquigarrow$ standard normal

Z-scores of normals

- These facts imply the **z-score** of a normal variable is a standard normal:

$$z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

- Subtract the mean and divide by the SD $\rightsquigarrow$ standard normal
- z-score measures how many SDs away from the mean a value of X is.

Matt Margolis
@ItsMattsLaw

Follow

dance like no one is watching

text like the editor of the Atlantic was accidentally added to the group chat and is reading everything

Central limit theorem

Definition (Central limit theorem)

Let $X_1, ..., X_n$ be i.i.d. r.v.s from a distribution with mean μ and variance σ^2. Then, $\bar{X}_n$ will be approximately distributed $N(\mu, \sigma^2/n)$ in large samples.

Central limit theorem

Definition (Central limit theorem)

Let $X_1, ..., X_n$ be i.i.d. r.v.s from a distribution with mean μ and variance σ^2. Then, $\bar{X}_n$ will be approximately distributed $N(\mu, \sigma^2/n)$ in large samples.

- "Sample means tend to be normally distributed as samples get large."

Central limit theorem

Definition (Central limit theorem)

Let $X_1, ..., X_n$ be i.i.d. r.v.s from a distribution with mean μ and variance σ^2. Then, $\bar{X}_n$ will be approximately distributed $N(\mu, \sigma^2/n)$ in large samples.

- "Sample means tend to be normally distributed as samples get large." -
$\rightsquigarrow$ we know (an approx. of) the entire probability distribution of $\bar{X}_n$

Central limit theorem

Definition (Central limit theorem)

Let $X_1, ..., X_n$ be i.i.d. r.v.s from a distribution with mean μ and variance σ^2. Then, $\bar{X}_n$ will be approximately distributed $N(\mu, \sigma^2/n)$ in large samples.

- "Sample means tend to be normally distributed as samples get large." - $\rightsquigarrow$ we know (an approx. of) the entire probability distribution of $\bar{X}_n$ - Approximation is better as n goes up.

Central limit theorem

Definition (Central limit theorem)

Let $X_1, ..., X_n$ be i.i.d. r.v.s from a distribution with mean μ and variance σ^2. Then, $\bar{X}_n$ will be approximately distributed $N(\mu, \sigma^2/n)$ in large samples.

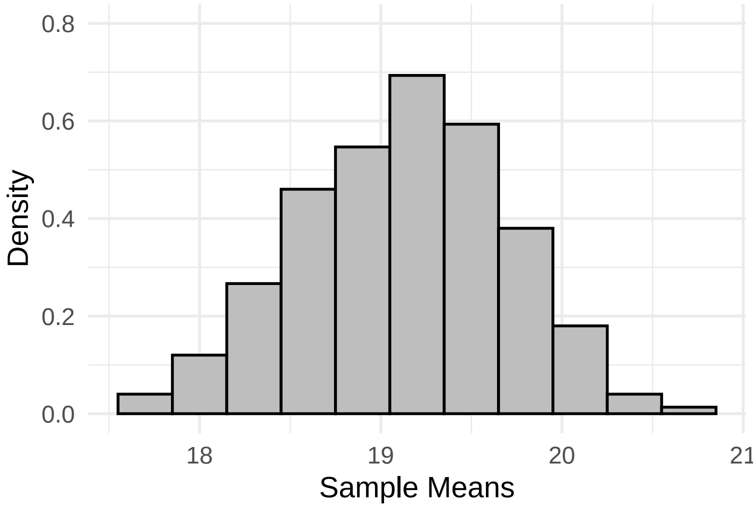
- "Sample means tend to be normally distributed as samples get large." - $\rightsquigarrow$ we know (an approx. of) the entire probability distribution of $\bar{X}_n$ - Approximation is better as n goes up. - Does not depend on the distribution of X_i!


THEY DON'T KNOW I'M THE EDITOR-IN-CHIEF OF THE ATLANTIC.
imgflip.com

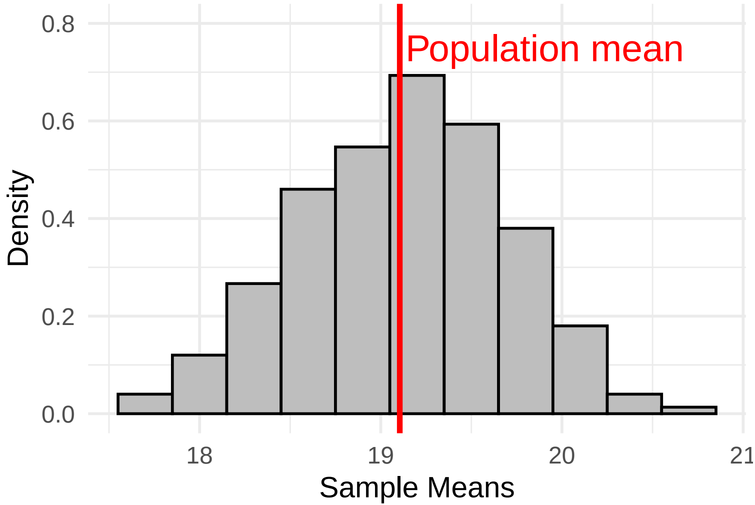
CLT simulation

1. Draw a sample size of 1,000 from the savings data
2. Calculate the sample mean of `save_gdp` for that sample
3. Save the sample mean
4. Repeat steps 1-3 a large number of times.

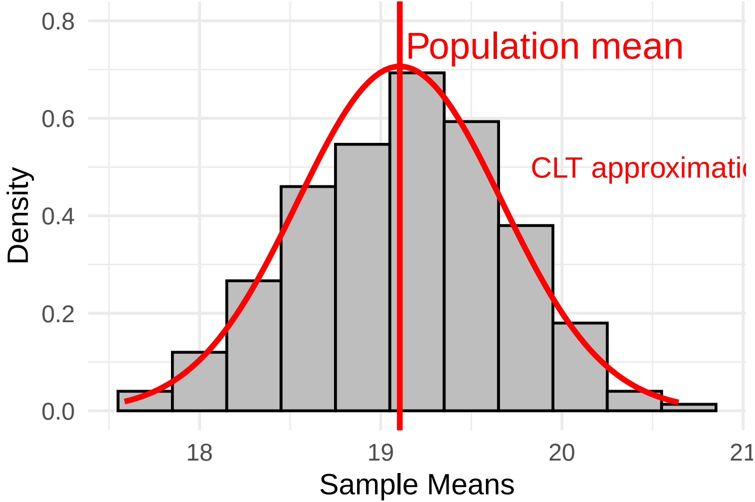
Histogram of sample means



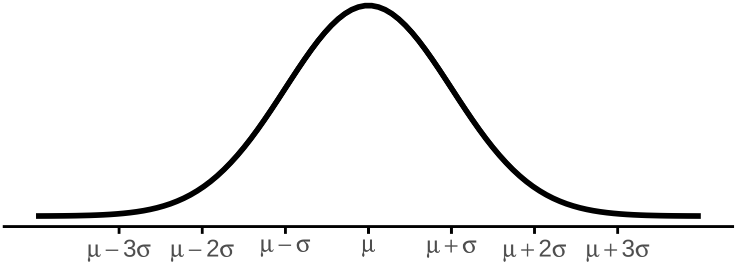
Histogram of sample means



Histogram of sample means

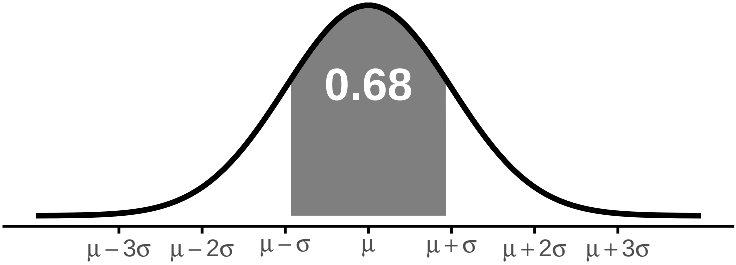


Empirical rule for the normal distribution



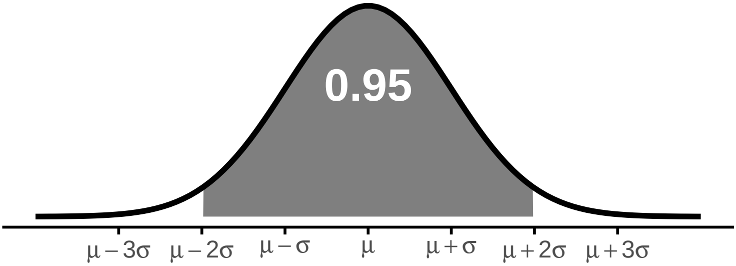
- If $X \sim N(\mu, \sigma^2)$, then:

Empirical rule for the normal distribution



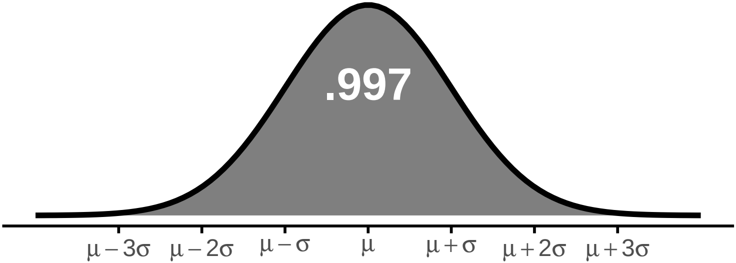
- If $X \sim N(\mu, \sigma^2)$, then:
  - $\approx$ 68% of the distribution of X is within 1 SD of the mean.

Empirical rule for the normal distribution



- If $X \sim N(\mu, \sigma^2)$, then:
  - $\approx$ 68% of the distribution of X is within 1 SD of the mean.
  - $\approx$ 95% of the distribution of X is within 2 SD of the mean.

Empirical rule for the normal distribution



- If $X \sim N(\mu, \sigma^2)$, then:
    - $\approx$ 68% of the distribution of X is within 1 SD of the mean.
    - $\approx$ 95% of the distribution of X is within 2 SD of the mean.
    - $\approx$ 99.7% of the distribution of X is within 3 SD of the mean.

Why the CLT?

- Why do we care about CLT?

Why the CLT?

- Why do we care about CLT?
  - We usually only sample once, so we'll only get 1 sample mean

Why the CLT?

- Why do we care about CLT?
    - We usually only sample once, so we'll only get 1 sample mean
    - Implies our 1 sample mean won't be too far from population mean.

Why the CLT?

- Why do we care about CLT?
  - We usually only sample once, so we'll only get 1 sample mean
  - Implies our 1 sample mean won't be too far from population mean.
- By CLT, sample mean $\approx$ normal with mean μ and SD $\frac{\sigma}{\sqrt{n}}$

Why the CLT?

- Why do we care about CLT?
  - We usually only sample once, so we'll only get 1 sample mean
  - Implies our 1 sample mean won't be too far from population mean.
- By CLT, sample mean $\approx$ normal with mean μ and SD $\frac{\sigma}{\sqrt{n}}$
- By empirical rule, sample mean will be...

Why the CLT?

- Why do we care about CLT?
  - We usually only sample once, so we'll only get 1 sample mean
  - Implies our 1 sample mean won't be too far from population mean.
- By CLT, sample mean $\approx$ normal with mean μ and SD $\frac{\sigma}{\sqrt{n}}$
- By empirical rule, sample mean will be...
  - Between $\mu - 2 \times \frac{\sigma}{\sqrt{n}}$ and $\mu + 2 \times \frac{\sigma}{\sqrt{n}}$ 95% of the time

Why the CLT?

- Why do we care about CLT?
  - We usually only sample once, so we'll only get 1 sample mean
  - Implies our 1 sample mean won't be too far from population mean.
- By CLT, sample mean $\approx$ normal with mean μ and SD $\frac{\sigma}{\sqrt{n}}$
- By empirical rule, sample mean will be...
  - Between $\mu - 2 \times \frac{\sigma}{\sqrt{n}}$ and $\mu + 2 \times \frac{\sigma}{\sqrt{n}}$ 95% of the time
- This will also help us create measure of uncertainty for our estimates