Lecture 15:
Power and
Confidence Intervals/Interval Estimation
API-201Z

Maya Sen

Harvard Kennedy School
http://scholar.harvard.edu/msen
Announcements

▶ Second midterm just over two weeks, November 15
▶ Will be in-class, closed book and closed note (will provide formula sheet, probability tables)
▶ Have posted old exams and problem sets
▶ Review Session Tuesday 11/13, 4-5:15pm, Rubenstein 304
Announcements

- Second midterm just over two weeks, November 15
Announcements

- Second midterm just over two weeks, November 15
- Will be in-class, closed book and closed note (will provide formula sheet, probability tables)
Announcements

- Second midterm just over two weeks, November 15
- Will be in-class, closed book and closed note (will provide formula sheet, probability tables)
- Have posted old exams and problem sets
Announcements

- Second midterm just over two weeks, November 15
- Will be in-class, closed book and closed note (will provide formula sheet, probability tables)
- Have posted old exams and problem sets
- Review Session Tuesday 11/13, 4-5:15pm, Rubenstein 304
Roadmap

- Be comfortable with four kinds of hypothesis tests:
  - Single mean
  - Difference in means
  - Single proportion
  - Difference in proportions

- Proper interpretation of Hypothesis Tests

- Today:
  - Power for hypothesis testing
  - Interval estimation
    - Confidence Intervals
    - Proper Interpretation
Roadmap

- Be comfortable with four kinds of hypothesis tests:
Roadmap

- Be comfortable with four kinds of hypothesis tests:
  - Single mean
Roadmap

- Be comfortable with four kinds of hypothesis tests:
  - Single mean
  - Difference in means
Roadmap

Be comfortable with four kinds of hypothesis tests:
- Single mean
- Difference in means
- Single proportion
Roadmap

- Be comfortable with four kinds of hypothesis tests:
  - Single mean
  - Difference in means
  - Single proportion
  - Difference in proportions
Roadmap

- Be comfortable with four kinds of hypothesis tests:
  - Single mean
  - Difference in means
  - Single proportion
  - Difference in proportions

- Proper interpretation of Hypothesis Tests
Roadmap

- Be comfortable with four kinds of hypothesis tests:
  - Single mean
  - Difference in means
  - Single proportion
  - Difference in proportions
- Proper interpretation of Hypothesis Tests
- Today:
Roadmap

- Be comfortable with four kinds of hypothesis tests:
  - Single mean
  - Difference in means
  - Single proportion
  - Difference in proportions
- Proper interpretation of Hypothesis Tests
- Today:
  - Power for hypothesis testing
Roadmap

- Be comfortable with four kinds of hypothesis tests:
  - Single mean
  - Difference in means
  - Single proportion
  - Difference in proportions

- Proper interpretation of Hypothesis Tests

- Today:
  - Power for hypothesis testing
  - Interval estimation
Roadmap

- Be comfortable with four kinds of hypothesis tests:
  - Single mean
  - Difference in means
  - Single proportion
  - Difference in proportions

- Proper interpretation of Hypothesis Tests

- Today:
  - Power for hypothesis testing
  - Interval estimation
  - Confidence Intervals
Roadmap

- Be comfortable with four kinds of hypothesis tests:
  - Single mean
  - Difference in means
  - Single proportion
  - Difference in proportions

- Proper interpretation of Hypothesis Tests

- Today:
  - Power for hypothesis testing
  - Interval estimation
  - Confidence Intervals
  - Proper Interpretation
Type I versus Type II Errors
### Type I versus Type II Errors

<table>
<thead>
<tr>
<th></th>
<th>$H_0$ is true</th>
<th>$H_0$ is not true</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject $H_0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do not reject $H_0$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Type I versus Type II Errors

<table>
<thead>
<tr>
<th></th>
<th>$H_0$ is true</th>
<th>$H_0$ is not true</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject $H_0$</td>
<td></td>
<td>Good</td>
</tr>
<tr>
<td>Do not reject $H_0$</td>
<td>Good</td>
<td></td>
</tr>
</tbody>
</table>
## Type I versus Type II Errors

<table>
<thead>
<tr>
<th></th>
<th>$H_0$ is true</th>
<th>$H_0$ is not true</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject $H_0$</td>
<td>Bad!</td>
<td>Good</td>
</tr>
<tr>
<td>Do not reject $H_0$</td>
<td>Good</td>
<td>Bad!</td>
</tr>
</tbody>
</table>
### Type I versus Type II Errors

<table>
<thead>
<tr>
<th></th>
<th>$H_0$ is true</th>
<th>$H_0$ is not true</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject $H_0$</td>
<td>Type 1</td>
<td>Good</td>
</tr>
<tr>
<td>Do not reject $H_0$</td>
<td>Good</td>
<td>Type 2</td>
</tr>
</tbody>
</table>
Type I versus Type II Errors

Both types of errors can occur with a hypothesis test.

Good practice: Before study carried out, set limits on how much error we would tolerate.
Type I versus Type II Errors

- Both types of errors can occur with a hypothesis test.
Type I versus Type II Errors

- Both types of errors can occur with a hypothesis test
- Good practice: Before study carried out, set limits on how much error we would tolerate
Type I Error

- Null hypothesis rejected when in fact it is true
- Akin to false positive
- Mammogram analogy: Positive test, despite not having the disease

\[ P(\text{Type I Error}) = P(\text{Rejecting } H_0 | H_0 \text{ true}) = \alpha \]

- Also referred to as level of significance or the critical value
- Rejection region: Values of \( \bar{X} \) to left/right of values of \( \alpha \) for which the hypothesis test would reject the null
- Very common to set \( \alpha = 0.05, \alpha = 0.10, \) or \( \alpha = 0.01 \)

Question: Why would we want to avoid Type I Error?

Question: How do you interpret a level of significance of 0.05?
Type I Error

- **Type I**: Null hypothesis rejected when in fact it is true

  - Akin to false positive
  - Mammogram analogy: Positive test, despite not having the disease

\[ P(\text{Type I error}) = P(\text{Reject } H_0 | H_0 \text{ true}) = \alpha \]

- \( \alpha \): Also referred to as level of significance or the critical value
- Rejection region: Values of \( \bar{X} \) to left/right of values of \( \alpha \) for which the hypothesis test would reject the null
- Very common to set \( \alpha = 0.05, \alpha = 0.10, \text{ or } \alpha = 0.01 \)

**Question**: Why would we want to avoid Type I Error?

**Question**: How do you interpret a level of significance of 0.05?
Type I Error

- Type I: Null hypothesis rejected when in fact it is true
- Akin to false positive

\[ P(\text{Type I error}) = P(\text{Rejecting } H_0 | H_0 \text{ true}) = \alpha \]

\( \alpha \): Also referred to as level of significance or the critical value

Rejection region: Values of \( \bar{X} \) to left/right of values of \( \alpha \) for which the hypothesis test would reject the null

Very common to set \( \alpha = 0.05, \alpha = 0.10, \text{ or } \alpha = 0.01 \)

Question: Why would we want to avoid Type I Error?

Question: How do you interpret a level of significance of 0.05?
Type I Error

- **Type I**: Null hypothesis rejected when in fact it is true
- Akin to false positive
  - Mammogram analogy: Positive test, despite not having the disease
Type I Error

- **Type I**: Null hypothesis rejected when in fact it is true
- Akin to false positive
  - Mammogram analogy: Positive test, despite not having the disease
- \( P(\text{Type I error}) = P(\text{Rejecting } H_0 | H_0 \text{ true}) = \alpha \)

\( \alpha \): Also referred to as level of significance or the critical value

Rejection region: Values of \( \bar{X} \) to left/right of values of \( \alpha \) for which the hypothesis test would reject the null

Very common to set \( \alpha = 0.05, \alpha = 0.10, \) or \( \alpha = 0.01 \)

Question: Why would we want to avoid Type I Error?

Question: How do you interpret a level of significance of 0.05?
Type I Error

- **Type I**: Null hypothesis rejected when in fact it is true
- Akin to false positive
  - Mammogram analogy: Positive test, despite not having the disease
- \( P(\text{Type I error}) = P(\text{Rejecting } H_0|H_0 \text{ true}) = \alpha \)
- \( \alpha \): Also referred to as level of significance or the critical value
Type I Error

- **Type I**: Null hypothesis rejected when in fact it is true
- Akin to false positive
  - Mammogram analogy: Positive test, despite not having the disease
- \( P(\text{Type I error}) = P(\text{Rejecting } H_0|H_0 \text{ true}) = \alpha \)
- \( \alpha \): Also referred to as level of significance or the critical value
- **Rejection region**: Values of \( \bar{X} \) to left/right of values of \( \alpha \) for which the hypothesis test would reject the null
Type I Error

- **Type I**: Null hypothesis rejected when in fact it is true
- Akin to false positive
  - Mammogram analogy: Positive test, despite not having the disease
- $P(\text{Type I error}) = P(\text{Rejecting } H_0|H_0 \text{ true}) = \alpha$
- $\alpha$: Also referred to as level of significance or the critical value
- **Rejection region**: Values of $\bar{X}$ to left/right of values of $\alpha$ for which the hypothesis test would reject the null
- Very common to set $\alpha = 0.05$, $\alpha = 0.10$, or $\alpha = 0.01$
Type I Error

- **Type I**: Null hypothesis rejected when in fact it is true
- Akin to false positive
  - Mammogram analogy: Positive test, despite not having the disease
- \( P(\text{Type I error}) = P(\text{Rejecting } H_0|H_0 \text{ true}) = \alpha \)
- \( \alpha \): Also referred to as level of significance or the critical value
- **Rejection region**: Values of \( \bar{X} \) to left/right of values of \( \alpha \) for which the hypothesis test would reject the null
- Very common to set \( \alpha = 0.05 \), \( \alpha = 0.10 \), or \( \alpha = 0.01 \)
- Question: Why would we want to avoid Type I Error?
Type I Error

- **Type I**: Null hypothesis rejected when in fact it is true
- Akin to false positive
  - Mammogram analogy: Positive test, despite not having the disease
- \( P(\text{Type I error}) = P(\text{Rejecting } H_0|H_0 \text{ true}) = \alpha \)
- \( \alpha \): Also referred to as level of significance or the critical value
- Rejection region: Values of \( \bar{X} \) to left/right of values of \( \alpha \) for which the hypothesis test would reject the null
- Very common to set \( \alpha = 0.05 \), \( \alpha = 0.10 \), or \( \alpha = 0.01 \)
- Question: Why would we want to avoid Type I Error?
- Question: How do you interpret a level of significance of 0.05?
Type II Error

- Type II: Null hypothesis is not rejected when in fact it is false
- Akin to false negative
- Mammogram analogy: Negative test, despite having the disease

\[ P(\text{Type II error}) = P(\text{Not rejecting } H_0 | H_0 \text{ false}) \]

- Often set at \[ P(\text{Type II error}) = 0.20 = \beta \]
Type II Error

- **Type II**: Null hypothesis is not rejected when in fact it is false

- $P(\text{Type II error}) = P(\text{Not rejecting } H_0 | H_0 \text{ false})$

- Often set at $P(\text{Type II error}) = 0.20 = \beta$
Type II Error

- **Type II**: Null hypothesis is not rejected when in fact it is false
- Akin to false negative
Type II Error

- **Type II**: Null hypothesis is not rejected when in fact it is false
- Akin to false negative
  - Mammogram analogy: Negative test, despite having the disease
Type II Error

- **Type II**: Null hypothesis is not rejected when in fact it is false
- Akin to false negative
  - Mammogram analogy: Negative test, despite having the disease
- \[ P(\text{Type II error}) = P(\text{Not rejecting } H_0|H_0 \text{ false}) \]
Type II Error

- **Type II**: Null hypothesis is not rejected when in fact it is false
- Akin to false negative
  - Mammogram analogy: Negative test, despite having the disease
- \( P(\text{Type II error}) = P(\text{Not rejecting } H_0|H_0 \text{ false}) \)
- Often set at \( P(\text{Type II error}) = 0.20 = \beta \)
Type II Error and Power

\[ P(\text{Rejecting } H_0 | H_0 \text{ false}) \]

Probability of correctly rejecting the null

Known as the power of a test

Sometimes considered less important from substantive perspective

However: consider your specific problem → in some instances, either error may be more costly

Note: Exact power calculation will depend on your alternative \( H_a \)
Type II Error and Power

$1 - \beta = P(\text{Rejecting } H_0 \mid H_0 \text{ false})$
Type II Error and Power

- $1 - \beta = P(\text{Rejecting } H_0 | H_0 \text{ false})$
- Probability of correctly rejecting the null
Type II Error and Power

1 – \( \beta = P(Rejecting \ H_0 | H_0 \ false) \)

- Probability of correctly rejecting the null
- Known as the **power** of a test

Note: Exact power calculation will depend on your alternative \( H_a \).
Type II Error and Power

- $1 - \beta = P(\text{Rejecting } H_0 | H_0 \text{ false})$
- Probability of correctly rejecting the null
- Known as the power of a test
- Sometimes considered less important from substantive perspective
Type II Error and Power

- $1 - \beta = P(\text{Rejecting } H_0|H_0 \text{ false})$
- Probability of correctly rejecting the null
- Known as the power of a test
- Sometimes considered less important from substantive perspective
- However: consider your specific problem → in some instances, either error may be more costly

Note: Exact power calculation will depend on your alternative $H_a$
Type II Error and Power

- $1 - \beta = P(\text{Rejecting } H_0 | H_0 \text{ false})$
- Probability of correctly rejecting the null
- Known as the power of a test
- Sometimes considered less important from substantive perspective
- However: consider your specific problem → in some instances, either error may be more costly
- Note: Exact power calculation will depend on your alternative $H_a$
Power Analyses

Power Analysis: If a given alternative hypothesis were true, how good would our test be at (correctly) rejecting the null?

Somewhat similar to hypothesis test set up, but for purposes of calculating power, assume alternative true:

\[ P(\text{Rejecting } H_0 | H_a \text{ true}) \]

Usually approached in one of two ways:

1) You are asked to calculate the sample size you will need to detect an effect (difference between null and alternative) of a certain size
   - Need to state a precise alternative null
2) You are asked to calculate what is the smallest effect (difference) you could detect given a sample size

→ Both mean that power analyses usually done before data are collected (and $ spent!)
Power Analyses

- **Power Analysis**: If a given alternative hypothesis were true, how good would our test be at (correctly) rejecting the null?
Power Analyses

- **Power Analysis**: If a given alternative hypothesis were true, how good would our test be at (correctly) rejecting the null?
- Somewhat similar to hypothesis test set up, but for purposes of calculating power, assume alternative true:
  \[ P(\text{Rejecting } H_0 \mid H_a \text{ true}) \]

Both mean that power analyses usually done before data are collected (and $ spent!)
Power Analyses

- **Power Analysis**: If a given alternative hypothesis were true, how good would our test be at (correctly) rejecting the null?
- Somewhat similar to hypothesis test set up, but for purposes of calculating power, assume alternative true: $P(\text{Rejecting } H_0|H_a \text{ true})$
- Usually approached in one of two ways:
Power Analyses

- **Power Analysis**: If a given alternative hypothesis were true, how good would our test be at (correctly) rejecting the null?

- Somewhat similar to hypothesis test set up, but for purposes of calculating power, assume alternative true:
  \[ P(\text{Rejecting } H_0|H_a \text{ true}) \]

- Usually approached in one of two ways:

- 1) You are asked to calculate the sample size you will need to detect an effect (difference between null and alternative) of a certain size

- Both mean that power analyses usually done before data are collected (and $ spent!)


Power Analyses

- **Power Analysis**: If a given alternative hypothesis were true, how good would our test be at (correctly) rejecting the null?
- Somewhat similar to hypothesis test set up, but for purposes of calculating power, assume alternative true:
  \[ P(\text{Rejecting } H_0 | H_a \text{ true}) \]
- Usually approached in one of two ways:
- 1) You are asked to calculate the sample size you will need to detect an effect (difference between null and alternative) of a certain size
  - Need to state a precise alternative null
Power Analyses

- **Power Analysis:** If a given alternative hypothesis were true, how good would our test be at (correctly) rejecting the null?

- Somewhat similar to hypothesis test set up, but for purposes of calculating power, assume alternative true: $P(\text{Rejecting } H_0 | H_a \text{ true})$

- Usually approached in one of two ways:

  1) You are asked to calculate the sample size you will need to detect an effect (difference between null and alternative) of a certain size
     - Need to state a precise alternative null

  2) You are asked to calculate what is the smallest effect (difference) you could detect given a sample size

Both mean that power analyses usually done before data are collected (and $ spent!)
Power Analyses

- **Power Analysis**: If a given alternative hypothesis were true, how good would our test be at (correctly) rejecting the null?

- Somewhat similar to hypothesis test set up, but for purposes of calculating power, assume alternative true:
  \[ P(\text{Rejecting } H_0 | H_a \text{ true}) \]

- Usually approached in one of two ways:
  1. You are asked to calculate the sample size you will need to detect an effect (difference between null and alternative) of a certain size
     - Need to state a precise alternative null
  2. You are asked to calculate what is the smallest effect (difference) you could detect given a sample size

→ Both mean that power analyses usually done before data are collected (and $ spent!)
Power Analysis Example

You are an urban policy expert studying the effects of recent construction on commuting times.

Suppose known average time spent commuting is 110 minutes/week, with standard deviation of 30 minutes.

Take sample of 81 commuters.

Find power of a hypothesis test when alternative hypothesis is that commuting hours equals 120 minutes.

So alternative is that the effect of construction is 20 minutes.

So $H_0 = 110$, $H_A = 120$, $\sigma = 30$, $n = 81$. 

Power Analysis Example

- You are an urban policy expert studying the effects of recent construction on commuting times

- Suppose known average time spent commuting is 110 minutes/week, with standard deviation of 30 minutes

- Take sample of 81 commuters

- Find power of a hypothesis test when alternative hypothesis is that commuting hours equals 120 minutes

- So alternative is that the effect of construction is 20 minutes

  \[ H_0 = 110, \quad H_A = 120, \quad \sigma = 30, \quad n = 81 \]
Power Analysis Example

- You are an urban policy expert studying the effects of recent construction on commuting times.
- Suppose known average time spent commuting is 110 minutes/week, with standard deviation of 30 minutes.
Power Analysis Example

- You are an urban policy expert studying the effects of recent construction on commuting times.
- Suppose known average time spent commuting is 110 minutes/week, with standard deviation of 30 minutes.
- Take sample of 81 commuters.
Power Analysis Example

- You are an urban policy expert studying the effects of recent construction on commuting times.
- Suppose known average time spent commuting is 110 minutes/week, with standard deviation of 30 minutes.
- Take sample of 81 commuters.
- Find power of a hypothesis test when alternative hypothesis is that commuting hours equals 120 minutes.
Power Analysis Example

- You are an urban policy expert studying the effects of recent construction on commuting times.
- Suppose known average time spent commuting is 110 minutes/week, with standard deviation of 30 minutes.
- Take sample of 81 commuters.
- Find power of a hypothesis test when alternative hypothesis is that commuting hours equals 120 minutes.
  - So alternative is that the effect of construction is 20 minutes.
Power Analysis Example

- You are an urban policy expert studying the effects of recent construction on commuting times.
- Suppose known average time spent commuting is 110 minutes/week, with standard deviation of 30 minutes.
- Take sample of 81 commuters.
- Find power of a hypothesis test when alternative hypothesis is that commuting hours equals 120 minutes.
  - So alternative is that the effect of construction is 20 minutes.
- So $H_0 = 110$, $H_A = 120$, $\sigma = 30$, $n = 81$. 
Power Analysis Example

Steps:

▶ Because you'll eventually do a hypothesis test assuming the null, first calculate the rejection region under standard set-up (given your $\alpha$ values)
▶ Calculate critical values for this rejection region
▶ Assume alternative hypothesis is true, $\mu_A = 120$
▶ Calculate probability of not being in rejection region when this alternative is true (this is $\beta$)
▶ And then finally calculate $1 - \beta$ to get Power

Note: Can repeat for different values of $\mu_A$ (e.g., $\mu_A = 120, 130, 140$, etc) to plot a power curve or power function
Power Analysis Example

Steps:

Because you'll eventually do a hypothesis test assuming the null, first calculate the rejection region under standard set-up (given your $\alpha$ values)

Calculate critical values for this rejection region

Assume alternative hypothesis is true, $\mu_A = 120$

Calculate probability of not being in rejection region when this alternative is true (this is $\beta$)

And then finally calculate $1 - \beta$ to get Power

Note: Can repeat for different values of $\mu_A$ (e.g., $\mu_A = 120, 130, 140$, etc) to plot a power curve or power function
Power Analysis Example

Steps:

- Because you’ll eventually do a hypothesis test assuming the null, first calculate the rejection region under standard set-up (given your $\alpha$ values)

- Calculate critical values for this rejection region

- Assume alternative hypothesis is true, $\mu_A = 120$

- Calculate probability of not being in rejection region when this alternative is true (this is $\beta$)

- And then finally calculate $1 - \beta$ to get Power

- Note: Can repeat for different values of $\mu_A$ (e.g., $\mu_A = 120, 130, 140$, etc) to plot a power curve or power function
Power Analysis Example

Steps:

▶ Because you’ll eventually do a hypothesis test assuming the null, first calculate the rejection region under standard set-up (given your $\alpha$ values)

▶ Calculate critical values for this rejection region

▶ Assume alternative hypothesis is true, $\mu_A = 120$

▶ Calculate probability of not being in rejection region when this alternative is true (this is $\beta$)

▶ And then finally calculate $1 - \beta$ to get Power

▶ Note: Can repeat for different values of $\mu_A$ (e.g., $\mu_A = 120, 130, 140$, etc) to plot a power curve or power function
Power Analysis Example

Steps:
- Because you’ll eventually do a hypothesis test assuming the null, first calculate the rejection region under standard set-up (given your $\alpha$ values)
- Calculate critical values for this rejection region
- Assume alternative hypothesis is true, $\mu_A = 120$
- Calculate probability of not being in rejection region when this alternative is true (this is $\beta$)
- And then finally calculate $1 - \beta$ to get Power
- Note: Can repeat for different values of $\mu_A$ (e.g., $\mu_A = 120, 130, 140,$ etc) to plot a power curve or power function
Power Analysis Example

Steps:

▶ Because you’ll eventually do a hypothesis test assuming the null, first calculate the rejection region under standard set-up (given your $\alpha$ values)
▶ Calculate critical values for this rejection region
▶ Assume alternative hypothesis is true, $\mu_A = 120$
▶ Calculate probability of not being in rejection region when this alternative is true (this is $\beta$)

Note: Can repeat for different values of $\mu_A$ (e.g., $\mu_A = 120, 130, 140$, etc) to plot a power curve or power function
Power Analysis Example

Steps:

▶ Because you’ll eventually do a hypothesis test assuming the null, first calculate the rejection region under standard set-up (given your $\alpha$ values)
▶ Calculate critical values for this rejection region
▶ Assume alternative hypothesis is true, $\mu_A = 120$
▶ Calculate probability of not being in rejection region when this alternative is true (this is $\beta$)
▶ And then finally calculate $1 - \beta$ to get Power

Note: Can repeat for different values of $\mu_A$ (e.g., $\mu_A = 120, 130, 140$, etc) to plot a power curve or power function.
Power Analysis Example

Steps:

▶ Because you’ll eventually do a hypothesis test assuming the null, first calculate the rejection region under standard set-up (given your \( \alpha \) values)

▶ Calculate critical values for this rejection region

▶ Assume alternative hypothesis is true, \( \mu_A = 120 \)

▶ Calculate probability of not being in rejection region when this alternative is true (this is \( \beta \))

▶ And then finally calculate \( 1 - \beta \) to get Power

▶ Note: Can repeat for different values of \( \mu_A \) (e.g., \( \mu_A = 120, 130, 140, etc \)) to plot a power curve or power function
Power Analysis Intuition

Normal distribution under H0 and H1

Distribution under H0
Power Analysis Intuition
Power Analysis Intuition

Normal distribution under H0 and H1

Distribution under H0

Critical value

α
Power Analysis Intuition

Normal distribution under H0 and H1

Distribution under H0

Distribution under H1

Critical value

\[ \alpha \]
Power Analysis Intuition

Normal distribution under H0 and H1

- Distribution under H0
- Distribution under H1
- Critical value

\[ \alpha \] and \[ \beta \] represent the significance level and the probability of a Type II error, respectively.

\( x \) is the observed value.

Density is plotted on the y-axis, and the x-axis represents the range of values from 100 to 130.
Power Analysis Intuition

Normal distribution under H0 and H1

- Distribution under H0
- Distribution under H1
- Critical value
- Power
- $\alpha$
- $\beta$
Power Analysis Example

Step 1: Calculate rejection region under the null hypothesis being true

Remember that test statistic for single mean is $z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$

Assuming one-tailed test, we reject $H_0$ if $z > 1.645$ ($\alpha = 0.05$)

Step 2: Calculate critical value

$1.645 < \bar{X} - \frac{110}{30}/\sqrt{81} < \bar{X}$

That is, we would reject null for all $\bar{X} > 115.48$
Power Analysis Example

- Step 1: Calculate rejection region under the null hypothesis being true

\[ z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \]

Assuming one-tailed test, we reject \( H_0 \) if \( z > 1.645 \) (\( \alpha = 0.05 \))

\[ 1.645 < \frac{\bar{X} - 110}{30 / \sqrt{81}} < \bar{X} \]

That is, we would reject null for all \( \bar{X} > 115.48 \)
Power Analysis Example

- Step 1: Calculate rejection region under the null hypothesis being true
- Remember that test statistic for single mean is

\[ z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \]
Power Analysis Example

- Step 1: Calculate rejection region under the null hypothesis being true
- Remember that test statistic for single mean is

\[ z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \]

- Assuming one-tailed test, we reject \( H_0 \) if \( z > 1.645 \) \( (\alpha = 0.05) \)
Power Analysis Example

- Step 1: Calculate rejection region under the null hypothesis being true
  - Remember that test statistic for single mean is
    \[ z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \]
  - Assuming one-tailed test, we reject \( H_0 \) if \( z > 1.645 \) (\( \alpha = 0.05 \))

- Step 2: Calculate critical value
  - \( 1.645 < \frac{\bar{X} - 110}{30/\sqrt{81}} \)
  - \( 115.48 < \bar{X} \)
Power Analysis Example

- Step 1: Calculate rejection region under the null hypothesis being true
- Remember that test statistic for single mean is
  \[ z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \]
- Assuming one-tailed test, we reject \( H_0 \) if \( z > 1.645 \) (\( \alpha = 0.05 \))
- Step 2: Calculate critical value
  \[ 1.645 < \frac{\bar{X} - 110}{30 / \sqrt{81}} \]
  \[ 115.48 < \bar{X} \]
- That is, we would reject null for all \( \bar{X} > 115.48 \)
Power Analysis Example

Step 3: Assume alternative true, $\mu_A = 120$

Given alternative being true, how often would we (correctly) reject null?

That is, what is $P(\bar{X} > 115.48)$ if $\mu = 120$?

$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{115.48 - 120}{30/\sqrt{81}} = -1.36$

Step 4: Finally $\beta = P(Z < -1.36)$, which is 0.087

Step 5: Power $= 1 - P(Z < -1.36) = 0.913$
Power Analysis Example

- Step 3: Assume alternative true, \( \mu_A = 120 \)
Power Analysis Example

- Step 3: Assume alternative true, $\mu_A = 120$
- Given alternative being true, how often would we (correctly) reject null?

\[ z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{115.48 - 120}{30/\sqrt{81}} = -1.36 \]

- Step 4: Finally $\beta = P(Z < -1.36)$, which is 0.087

- Step 5: Power $= 1 - P(Z < -1.36) = 0.913$
Power Analysis Example

- Step 3: Assume alternative true, $\mu_A = 120$
- Given alternative being true, how often would we (correctly) reject null?
- That is, what is $P(\bar{X} > 115.48)$ if $\mu = 120$?
Step 3: Assume alternative true, $\mu_A = 120$

Given alternative being true, how often would we (correctly) reject null?

That is, what is $P(\bar{X} > 115.48)$ if $\mu = 120$?

$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

$$= \frac{115.48 - 120}{30/\sqrt{81}}$$

$$= -1.36$$
Power Analysis Example

- Step 3: Assume alternative true, $\mu_A = 120$
- Given alternative being true, how often would we (correctly) reject null?
- That is, what is $P(\bar{X} > 115.48)$ if $\mu = 120$?

$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

$$= \frac{115.48 - 120}{30/\sqrt{81}}$$

$$= -1.36$$

- Step 4: Finally $\beta = P(Z < -1.36)$, which is 0.087
Power Analysis Example

- Step 3: Assume alternative true, $\mu_A = 120$
- Given alternative being true, how often would we (correctly) reject null?
- That is, what is $P(\bar{X} > 115.48)$ if $\mu = 120$?

\[
z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{115.48 - 120}{30 / \sqrt{81}} = -1.36
\]

- Step 4: Finally $\beta = P(Z < -1.36)$, which is 0.087
- Step 5: $Power = 1 - P(Z < -1.36) = 0.913$
Power

Note: Power dependent on a) sample size, b) your $H_A$ (size of effect), and c) $\alpha$ level (and so type of test)

Larger sample size $\rightarrow$ more power

Bigger difference between null and alternative you are testing $\rightarrow$ requires less power

A two-sided hypothesis test has less power than the one-sided hypothesis, since it is more conservative

Rule of thumb: 80% power

Higher power may be better, but perhaps not if it comes in changing the type of test (b/c it affects Type 1 error)
Power

- Note: Power dependent on a) sample size, b) your $H_A$ (size of effect), and c) $\alpha$ level (and so type of test)
Power

- Note: Power dependent on a) sample size, b) your $H_A$ (size of effect), and c) $\alpha$ level (and so type of test)
- Larger sample size $\rightarrow$ more power
Note: Power dependent on a) sample size, b) your $H_A$ (size of effect), and c) $\alpha$ level (and so type of test)

- Larger sample size $\rightarrow$ more power
- Bigger difference between null and alternative you are testing $\rightarrow$ requires less power
Power

- Note: Power dependent on a) sample size, b) your $H_A$ (size of effect), and c) $\alpha$ level (and so type of test)
- Larger sample size $\rightarrow$ more power
- Bigger difference between null and alternative you are testing $\rightarrow$ requires less power
- A two-sided hypothesis test has less power than the one-sided hypothesis, since it is more conservative
Power

- Note: Power dependent on a) sample size, b) your $H_A$ (size of effect), and c) $\alpha$ level (and so type of test)
- Larger sample size $\rightarrow$ more power
- Bigger difference between null and alternative you are testing $\rightarrow$ requires less power
- A two-sided hypothesis test has less power than the one-sided hypothesis, since it is more conservative
- Rule of thumb: 80% power
Power

- Note: Power dependent on a) sample size, b) your $H_A$ (size of effect), and c) $\alpha$ level (and so type of test)
- Larger sample size $\rightarrow$ more power
- Bigger difference between null and alternative you are testing $\rightarrow$ requires less power
- A two-sided hypothesis test has less power than the one-sided hypothesis, since it is more conservative
- Rule of thumb: 80% power
- Higher power may be better, but perhaps not if it comes in changing the type of test (b/c it affects Type 1 error)
Interval Estimation

Hypothesis testing → Useful to compare sample/s against a null hypothesis

Interval estimation → Useful for calculating a range of possible values for the true population proportion/mean

Most commonly used are confidence intervals (CIs)

Takes into account not only the point estimate (for example, $\hat{\pi}$), but also variability and sample size

Caution: Frequently incorrectly interpreted!

Most basic example → Confidence interval for a mean
Interval Estimation

- **Hypothesis testing** → Useful to compare sample/s against a null hypothesis
Interval Estimation

- **Hypothesis testing** → Useful to compare sample/s against a null hypothesis
- **Interval estimation** → Useful for calculating a range of possible values for the true population proportion/mean
Interval Estimation

- **Hypothesis testing** → Useful to compare sample/s against a null hypothesis
- **Interval estimation** → Useful for calculating a range of possible values for the true population proportion/mean
- Most commonly used are confidence intervals (CIs)
Interval Estimation

- **Hypothesis testing** → Useful to compare sample/s against a null hypothesis
- **Interval estimation** → Useful for calculating a range of possible values for the true population proportion/mean
- Most commonly used are **confidence intervals (CIs)**
- Takes into account not only the point estimate (for example, \(\hat{\pi}\)), but also variability and sample size
Interval Estimation

- **Hypothesis testing** → Useful to compare sample/s against a null hypothesis
- **Interval estimation** → Useful for calculating a range of possible values for the true population proportion/mean
- Most commonly used are **confidence intervals** (CIs)
- Takes into account not only the point estimate (for example, $\hat{\pi}$), but also variability and sample size
- **Caution**: Frequently incorrectly interpreted!
Interval Estimation

- **Hypothesis testing** → Useful to compare sample/s against a null hypothesis
- **Interval estimation** → Useful for calculating a range of possible values for the true population proportion/mean
- Most commonly used are **confidence intervals** (CIs)
- Takes into account not only the **point estimate** (for example, $\hat{\pi}$), but also variability and sample size
- **Caution**: Frequently incorrectly interpreted!
- Most basic example → Confidence interval for a mean
Calculating Confidence Intervals for Means

Use our old friend, the CLT:

1. The sums and means of random samples of observations have an approximately normal distribution.
2. This distribution becomes more and more normal the more observations are included in the sum or the mean.
3. Via the large of large numbers, this will be centered around the true population mean/proportion.
4. CLT tells us about the underlying behavior of the sample proportion/mean across all different kinds of data.
Calculating Confidence Intervals for Means

Use our old friend, the CLT:
Calculating Confidence Intervals for Means

Use our old friend, the CLT:

1. The sums and means of random samples of observations have an approximately normal distribution

   - Via the large of large numbers, this will be centered around the true population mean/proportion

   - CLT tells us about the underlying behavior of the sample proportion/mean across all different kinds of data
Calculating Confidence Intervals for Means

Use our old friend, the CLT:

1. The sums and means of random samples of observations have an approximately normal distribution
2. This distribution becomes more and more normal the more observations are included in the sum or the mean
Calculating Confidence Intervals for Means

Use our old friend, the CLT:

1. The sums and means of random samples of observations have an approximately normal distribution.
2. This distribution becomes more and more normal the more observations are included in the sum or the mean.
3. Via the large of large numbers, this will be centered around the true population mean/proportion.
Calculating Confidence Intervals for Means

Use our old friend, the CLT:

1. The sums and means of random samples of observations have an approximately normal distribution.
2. This distribution becomes more and more normal the more observations are included in the sum or the mean.
3. Via the large of large numbers, this will be centered around the true population mean/proportion.
4. CLT tells us about the underlying behavior of the sample proportion/mean across all different kinds of data.
Calculating Confidence Intervals for Means

Stated more formally for the sample mean:

If $\bar{X}$ is the mean of $n$ measurements $x_1, x_2, \ldots, x_n$, then as $n$ goes up, $\bar{X}$ approaches:

$\bar{X} \sim N(\mu, \sigma^2/n)$

And b/c this is Normal, we can standardize

$\bar{X} - \mu \frac{s}{\sqrt{n}}$

We replace true standard error, $\sigma^2/n$ with the estimate from our sample, $s$: $\bar{X} - \mu \frac{s}{\sqrt{n}}$
Calculating Confidence Intervals for Means

- Stated more formally for the sample mean:

\[
\bar{X} \sim N(\mu, \sigma^2/n)
\]

And because this is Normal, we can standardize

\[
\bar{X} - \mu \frac{s}{\sqrt{n}}
\]

We replace true standard error, \(\sigma^2/n\) with the estimate from our sample, \(s\):
Calculating Confidence Intervals for Means

- Stated more formally for the sample mean:
- If $\bar{X}$ is the mean of $n$ measurements $x_1, x_2, \ldots, x_n$, then as $n$ goes up, $\bar{X}$ approaches:

$$\bar{X} \sim N(\mu, \sigma^2/n)$$

And because this is Normal, we can standardize

$$\bar{X} - \mu \frac{s}{\sqrt{n}}$$

We replace true standard error, $\sigma^2/n$ with the estimate from our sample, $s$: 

$$\bar{X} - \mu \frac{s}{\sqrt{n}}$$
Calculating Confidence Intervals for Means

- Stated more formally for the sample mean:
- If $\bar{X}$ is the mean of $n$ measurements $x_1, x_2, \ldots, x_n$, then as $n$ goes up, $\bar{X}$ approaches:

$$\bar{X} \sim N(\mu, \sigma^2/n)$$
Calculating Confidence Intervals for Means

- Stated more formally for the sample mean:
- If $\bar{X}$ is the mean of $n$ measurements $x_1, x_2, \ldots, x_n$, then as $n$ goes up, $\bar{X}$ approaches:

$$\bar{X} \sim N(\mu, \sigma^2/n)$$

- And b/c this is Normal, we can standardize
Calculating Confidence Intervals for Means

- Stated more formally for the sample mean:
- If $\bar{X}$ is the mean of $n$ measurements $x_1, x_2, \ldots, x_n$, then as $n$ goes up, $\bar{X}$ approaches:

$$\bar{X} \sim N(\mu, \sigma^2/n)$$

- And b/c this is Normal, we can standardize

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$
Calculating Confidence Intervals for Means

- Stated more formally for the sample mean:
- If \( \bar{X} \) is the mean of \( n \) measurements \( x_1, x_2, \ldots, x_n \), then as \( n \) goes up, \( \bar{X} \) approaches:

\[
\bar{X} \sim N(\mu, \sigma^2/n)
\]

- And b/c this is Normal, we can standardize

\[
\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}
\]

- We replace true standard error, \( \sigma^2/n \) with the estimate from our sample, \( s \):
Calculating Confidence Intervals for Means

- Stated more formally for the sample mean:
- If $\bar{X}$ is the mean of $n$ measurements $x_1, x_2, \ldots, x_n$, then as $n$ goes up, $\bar{X}$ approaches:

$$\bar{X} \sim N(\mu, \sigma^2/n)$$

- And b/c this is Normal, we can standardize

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

- We replace true standard error, $\sigma^2/n$ with the estimate from our sample, $s$:

$$\frac{\bar{X} - \mu}{s/\sqrt{n}}$$
Calculating Confidence Intervals for Means

Here's where confidence intervals differ from hypothesis tests:

- Leverage that we know:
  - Approx 68% of probability falling within 1 SD of mean
  - Approx 95% of probability falling within 2 SD of mean
  - Approx 99.7% of probability falling within 3 SD of mean

- Can be more exact than this
  - For example, know that 95% of probability mass of a standard Normal falls more precisely between -1.96 and 1.96

That means we know that:
\[ P(-1.96 \leq \bar{X} - \mu < 1.96) = 0.95 \]
Calculating Confidence Intervals for Means

Here’s where confidence intervals differ from hypothesis tests:

- Leverage fact that we know:
  - Approx 68% of probability falling within 1 SD of mean
  - Approx 95% of probability falling within 2 SD of mean
  - Approx 99.7% of probability falling within 3 SD of mean

- Can be more exact than this
- For example, know that 95% of probability mass of a standard Normal falls more precisely between -1.96 and 1.96
- That means we know that:
  \[ P(-1.96 \leq \bar{X} - \mu \leq 1.96) = 0.95 \]
Calculating Confidence Intervals for Means

- Here’s where confidence intervals differ from hypothesis tests:
- For CI’s: Leverage fact that we know:
Calculating Confidence Intervals for Means

- Here’s where confidence intervals differ from hypothesis tests:
- For CI’s: Leverage fact that we know:
  - Approx 68% of probability falling within 1 SD of mean
Calculating Confidence Intervals for Means

- Here’s where confidence intervals differ from hypothesis tests:
- For CI’s: Leverage fact that we know:
  - Approx 68% of probability falling within 1 SD of mean
  - Approx 95% of probability falling within 2 SD of mean
Calculating Confidence Intervals for Means

Here’s where confidence intervals differ from hypothesis tests:

For CI’s: Leverage fact that we know:
- Approx 68% of probability falling within 1 SD of mean
- Approx 95% of probability falling within 2 SD of mean
- Approx 99.7% of probability falling within 3 SD of mean
Calculating Confidence Intervals for Means

- Here’s where confidence intervals differ from hypothesis tests:
- For CI’s: Leverage fact that we know:
  - Approx 68% of probability falling within 1 SD of mean
  - Approx 95% of probability falling within 2 SD of mean
  - Approx 99.7% of probability falling within 3 SD of mean
- Can be more exact than this
Calculating Confidence Intervals for Means

- Here’s where confidence intervals differ from hypothesis tests:
  - For CI’s: Leverage fact that we know:
    - Approx 68% of probability falling within 1 SD of mean
    - Approx 95% of probability falling within 2 SD of mean
    - Approx 99.7% of probability falling within 3 SD of mean
  - Can be more exact than this
  - For example, know that 95% of probability mass of a standard Normal falls more precisely between -1.96 and 1.96
Calculating Confidence Intervals for Means

- Here’s where confidence intervals differ from hypothesis tests:
  - For CI’s: Leverage fact that we know:
    - Approx 68% of probability falling within 1 SD of mean
    - Approx 95% of probability falling within 2 SD of mean
    - Approx 99.7% of probability falling within 3 SD of mean
  - Can be more exact than this
  - For example, know that 95% of probability mass of a standard Normal falls more precisely between -1.96 and 1.96
  - That means we know that:
Calculating Confidence Intervals for Means

Here’s where confidence intervals differ from hypothesis tests:
- For CI’s: Leverage fact that we know:
  - Approx 68% of probability falling within 1 SD of mean
  - Approx 95% of probability falling within 2 SD of mean
  - Approx 99.7% of probability falling within 3 SD of mean
- Can be more exact than this
- For example, know that 95% of probability mass of a standard Normal falls more precisely between -1.96 and 1.96
- That means we know that:

\[ P(-1.96 \leq \frac{\bar{X} - \mu}{s/\sqrt{n}} \leq 1.96) = 0.95 \]
Calculating Confidence Intervals for Means

Take this and work backwards:

\[ P(-1.96 \leq \bar{X} - \mu \leq 1.96) = 0.95 \]

\[ P(-1.96 \frac{s}{\sqrt{n}} \leq \bar{X} - \mu \leq 1.96 \frac{s}{\sqrt{n}}) = 0.95 \]

\[ P(-\bar{X} - 1.96 \frac{s}{\sqrt{n}} \leq -\mu \leq -\bar{X} + 1.96 \frac{s}{\sqrt{n}}) = 0.95 \]

\[ P(\bar{X} + 1.96 \frac{s}{\sqrt{n}} \geq \mu \geq \bar{X} - 1.96 \frac{s}{\sqrt{n}}) = 0.95 \]
Calculating Confidence Intervals for Means

- Take this and work backwards:
Calculating Confidence Intervals for Means

- Take this and work backwards:

\[
P(-1.96 \leq \frac{\bar{X} - \mu}{s/\sqrt{n}} \leq 1.96) = 0.95
\]
Calculating Confidence Intervals for Means

- Take this and work backwards:

\[ P(-1.96 \leq \bar{X} - \mu \leq 1.96) = 0.95 \]

\[ P(-1.96 \frac{s}{\sqrt{n}} \leq \bar{X} - \mu \leq 1.96 \frac{s}{\sqrt{n}}) = 0.95 \]
Calculating Confidence Intervals for Means

- Take this and work backwards:

\[
P(-1.96 \leq \frac{\bar{X} - \mu}{s/\sqrt{n}} \leq 1.96) = 0.95
\]

\[
P(-1.96 \frac{s}{\sqrt{n}} \leq \bar{X} - \mu \leq 1.96 \frac{s}{\sqrt{n}}) = 0.95
\]

\[
P(-\bar{X} - 1.96 \frac{s}{\sqrt{n}} \leq -\mu \leq -\bar{X} + 1.96 \frac{s}{\sqrt{n}}) = 0.95
\]
Calculating Confidence Intervals for Means

- Take this and work backwards:

\[
P(-1.96 \leq \frac{\bar{X} - \mu}{s/\sqrt{n}} \leq 1.96) = 0.95
\]

\[
P(-1.96 \frac{s}{\sqrt{n}} \leq \bar{X} - \mu \leq 1.96 \frac{s}{\sqrt{n}}) = 0.95
\]

\[
P(-\bar{X} - 1.96 \frac{s}{\sqrt{n}} \leq -\mu \leq -\bar{X} + 1.96 \frac{s}{\sqrt{n}}) = 0.95
\]

\[
P(\bar{X} + 1.96 \frac{s}{\sqrt{n}} \geq \mu \geq \bar{X} - 1.96 \frac{s}{\sqrt{n}}) = 0.95
\]
Calculating Confidence Intervals for Means

This gives us the 95% confidence interval for $\bar{X} \pm 1.96 \times \hat{SE} [\bar{X}]$.

This is shorthand $[LB, UB]$:

$LB = \bar{X} - 1.96 \times \hat{SE} [\bar{X}]$

$UB = \bar{X} + 1.96 \times \hat{SE} [\bar{X}]$
Calculating Confidence Intervals for Means

This gives us the 95% confidence interval for $\bar{X}$

$\bar{X} \pm 1.96 \times \hat{SE} [\bar{X}]$

This is shorthand

$LB = \bar{X} - 1.96 \times \hat{SE} [\bar{X}]$

$UB = \bar{X} + 1.96 \times \hat{SE} [\bar{X}]$
This gives us the 95% confidence interval for $\bar{X}$

$$\bar{X} \pm 1.96 \times \hat{SE}[\bar{X}]$$
Calculating Confidence Intervals for Means

- This gives us the 95% confidence interval for $\bar{X}$

$$\bar{X} \pm 1.96 \times \hat{SE}[\bar{X}]$$

- This is shorthand $[LB, UB]$:
Calculating Confidence Intervals for Means

- This gives us the 95% confidence interval for \( \bar{X} \)

\[
\bar{X} \pm 1.96 \times \hat{SE}[\bar{X}]
\]

- This is shorthand \([LB, UB]\):

\[
LB = \bar{X} - 1.96 \times \hat{SE}[\bar{X}]
\]
\[
UB = \bar{X} + 1.96 \times \hat{SE}[\bar{X}]
\]
Calculating Confidence Intervals for Means

What about confidence intervals other than 95%?

Say we want a \((1 - \alpha)\)% confidence interval?

\[
P(-z_{\alpha/2} \leq \bar{X} - \mu \leq z_{\alpha/2} \hat{SE} \bar{X}) = (1 - \alpha)
\]

Gives general formula for a \((1 - \alpha)\)% confidence interval:

\[
\bar{X} \pm z_{\alpha/2} \hat{SE} \bar{X}
\]

Where we use \(z_{\alpha/2}\) from the standard normal (or Student's \(t\))
Calculating Confidence Intervals for Means

- What about confidence intervals other than 95%?

\[
\bar{X} \pm z_{\alpha/2} \times \hat{SE} \left[ \bar{X} \right]
\]

Gives general formula for a \((1 - \alpha)\)% confidence interval:

\[
\bar{X} - z_{\alpha/2} \times \hat{SE} \left[ \bar{X} \right] \leq \mu \leq \bar{X} + z_{\alpha/2} \times \hat{SE} \left[ \bar{X} \right]
\]

Where we use \(z_{\alpha/2}\) from the standard normal (or Student's t)
Calculating Confidence Intervals for Means

- What about confidence intervals other than 95%?
- Say we want an \((1 - \alpha)\)% confidence interval?
Calculating Confidence Intervals for Means

- What about confidence intervals other than 95%?
- Say we want an \((1 - \alpha)\)% confidence interval?

\[
P\left(-z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\hat{SE}[\bar{X}]} \leq z_{\alpha/2}\right) = (1 - \alpha)
\]

\[
P\left(\bar{X} - z_{\alpha/2} \times \hat{SE}[\bar{X}] \leq \mu \leq \bar{X} + z_{\alpha/2} \times \hat{SE}[\bar{X}]\right) = (1 - \alpha)
\]
Calculating Confidence Intervals for Means

▶ What about confidence intervals other than 95%?
▶ Say we want an \((1 - \alpha)\)% confidence interval?

\[
P\left(-z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\hat{SE}\left[\bar{X}\right]} \leq z_{\alpha/2}\right) = (1 - \alpha)
\]

\[
P\left(\bar{X} - z_{\alpha/2} \times \hat{SE}\left[\bar{X}\right] \leq \mu \leq \bar{X} + z_{\alpha/2} \times \hat{SE}\left[\bar{X}\right]\right) = (1 - \alpha)
\]

▶ Gives general formula for a \((1 - \alpha)\)% confidence interval:
Calculating Confidence Intervals for Means

What about confidence intervals other than 95%?

Say we want an \((1 - \alpha)\)% confidence interval?

\[
P\left( - z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\hat{SE}[\bar{X}]} \leq z_{\alpha/2} \right) = (1 - \alpha)
\]

\[
P\left( \bar{X} - z_{\alpha/2} \times \hat{SE}[\bar{X}] \leq \mu \leq \bar{X} + z_{\alpha/2} \times \hat{SE}[\bar{X}] \right) = (1 - \alpha)
\]

Gives general formula for a \((1 - \alpha)\)% confidence interval:

\[
\bar{X} \pm z_{\alpha/2} \times \hat{SE}[\bar{X}]
\]
Calculating Confidence Intervals for Means

▶ What about confidence intervals other than 95%?
▶ Say we want an $(1 - \alpha)\%$ confidence interval?

\[
P\left(-z_{\alpha/2} \leq \frac{\bar{X} - \mu}{SE[\bar{X}]} \leq z_{\alpha/2}\right) = (1 - \alpha)
\]

\[
P\left(\bar{X} - z_{\alpha/2} \times \hat{SE}[\bar{X}] \leq \mu \leq \bar{X} + z_{\alpha/2} \times \hat{SE}[\bar{X}]\right) = (1 - \alpha)
\]

▶ Gives general formula for a $(1 - \alpha)\%$ confidence interval:

\[
\bar{X} \pm z_{\alpha/2} \times \hat{SE}[\bar{X}]
\]

▶ Where we use $z_{\alpha/2}$ from the standard normal (or Student’s $t$)
Calculating Confidence Intervals for Means

- For $\alpha = 0.05 \rightarrow 95\%$ Confidence Interval: $\bar{X} \pm 1.96 \times SE [X]$
- For $\alpha = 0.1 \rightarrow 90\%$ Confidence Interval: $\bar{X} \pm 1.645 \times SE [X]$
- For $\alpha = 0.1 \rightarrow 99\%$ Confidence Interval: $\bar{X} \pm 2.58 \times SE [X]$. 
Calculating Confidence Intervals for Means

- Means we can easily calculate confidence intervals for commonly used $\alpha$ values

  $\alpha = 0.05 \rightarrow 95\%$ Confidence Interval $\bar{X} \pm 1.96 \text{SE} \left[ X \right]$

  $\alpha = 0.1 \rightarrow 90\%$ Confidence Interval $\bar{X} \pm 1.645 \text{SE} \left[ X \right]$

  $\alpha = 0.1 \rightarrow 99\%$ Confidence Interval $\bar{X} \pm 2.58 \text{SE} \left[ X \right]$
Calculating Confidence Intervals for Means

- Means we can easily calculate confidence intervals for commonly used $\alpha$ values
- $\alpha = 0.05 \rightarrow 95\%$ Confidence Interval
- $\alpha = 0.1 \rightarrow 90\%$ Confidence Interval
- $\alpha = 0.01 \rightarrow 99\%$ Confidence Interval
Calculating Confidence Intervals for Means

- Means we can easily calculate confidence intervals for commonly used $\alpha$ values
- $\alpha = 0.05 \rightarrow 95\%$ Confidence Interval
  \[ \bar{X} \pm 1.96SE[X] \]
- $\alpha = 0.1 \rightarrow 90\%$ Confidence Interval
- $\alpha = 0.1 \rightarrow 99\%$ Confidence Interval
Calculating Confidence Intervals for Means

- Means we can easily calculate confidence intervals for commonly used $\alpha$ values
- $\alpha = 0.05 \rightarrow 95\%$ Confidence Interval
  \[ \bar{X} \pm 1.96SE[X] \]
- $\alpha = 0.1 \rightarrow 90\%$ Confidence Interval
Calculating Confidence Intervals for Means

- Means we can easily calculate confidence intervals for commonly used $\alpha$ values
- $\alpha = 0.05 \rightarrow 95\%$ Confidence Interval
  $$\bar{X} \pm 1.96 SE[X]$$
- $\alpha = 0.1 \rightarrow 90\%$ Confidence Interval
  $$\bar{X} \pm 1.645 SE[X]$$
Calculating Confidence Intervals for Means

- Means we can easily calculate confidence intervals for commonly used $\alpha$ values
- $\alpha = 0.05 \rightarrow$ 95% Confidence Interval
  $$\bar{X} \pm 1.96SE[X]$$
- $\alpha = 0.1 \rightarrow$ 90% Confidence Interval
  $$\bar{X} \pm 1.645SE[X]$$
- $\alpha = 0.1 \rightarrow$ 99% Confidence Interval
Calculating Confidence Intervals for Means

- Means we can easily calculate confidence intervals for commonly used $\alpha$ values
- $\alpha = 0.05 \rightarrow 95\%$ Confidence Interval
  
  $$\bar{X} \pm 1.96SE[X]$$

- $\alpha = 0.1 \rightarrow 90\%$ Confidence Interval
  
  $$\bar{X} \pm 1.645SE[X]$$

- $\alpha = 0.1 \rightarrow 99\%$ Confidence Interval
  
  $$\bar{X} \pm 2.58SE[X]$$
Calculating Confidence Intervals for Other Quantities of Interest

CLT means we can calculate confidence intervals for any estimator that approximates the Normal distribution as \( n \) goes up.

- Means
- Difference in means
- Proportions
- Difference in proportions

All follow the general form of

\[
\text{Point Estimate} \pm z_{\alpha/2} \hat{SE}
\]

where \( z_{\alpha/2} \hat{SE} \) refers to the margin of error.

CI's most generally are:

\[
\text{Point Estimate} \pm \text{Margin of Error}
\]
Calculating Confidence Intervals for Other Quantities of Interest

- CLT means we can calculate confidence intervals for any estimator that approximates the Normal distribution as $n$ goes up.
Calculating Confidence Intervals for Other Quantities of Interest

- CLT means we can calculate confidence intervals for any estimator that approximates the Normal distribution as $n$ goes up
  - Means
Calculating Confidence Intervals for Other Quantities of Interest

- CLT means we can calculate confidence intervals for any estimator that approximates the Normal distribution as $n$ goes up
  - Means
  - Difference in means
Calculating Confidence Intervals for Other Quantities of Interest

- CLT means we can calculate confidence intervals for any estimator that approximates the Normal distribution as $n$ goes up
  - Means
  - Difference in means
  - Proportions
Calculating Confidence Intervals for Other Quantities of Interest

- CLT means we can calculate confidence intervals for any estimator that approximates the Normal distribution as $n$ goes up
  - Means
  - Difference in means
  - Proportions
  - Difference in proportions
Calculating Confidence Intervals for Other Quantities of Interest

- CLT means we can calculate confidence intervals for any estimator that approximates the Normal distribution as \( n \) goes up
  - Means
  - Difference in means
  - Proportions
  - Difference in proportions
- All follow general form of
Calculating Confidence Intervals for Other Quantities of Interest

- CLT means we can calculate confidence intervals for any estimator that approximates the Normal distribution as $n$ goes up
  - Means
  - Difference in means
  - Proportions
  - Difference in proportions
- All follow general form of

\[
\text{Point Estimate} \pm z_{\alpha/2} \hat{SE}
\]
Calculating Confidence Intervals for Other Quantities of Interest

- CLT means we can calculate confidence intervals for any estimator that approximates the Normal distribution as $n$ goes up
  - Means
  - Difference in means
  - Proportions
  - Difference in proportions
- All follow general form of
  
  \[ \text{Point Estimate} \pm z_{\alpha/2} \hat{SE} \]

- where $z_{\alpha/2} \hat{SE}$ refers to the margin of error
Calculating Confidence Intervals for Other Quantities of Interest

- CLT means we can calculate confidence intervals for any estimator that approximates the Normal distribution as $n$ goes up
  - Means
  - Difference in means
  - Proportions
  - Difference in proportions
- All follow general form of

\[
\text{Point Estimate} \pm z_{\alpha/2} \hat{SE}
\]

- where $z_{\alpha/2} \hat{SE}$ refers to the margin of error
- CI's most generally are:
Calculating Confidence Intervals for Other Quantities of Interest

- CLT means we can calculate confidence intervals for any estimator that approximates the Normal distribution as $n$ goes up
  - Means
  - Difference in means
  - Proportions
  - Difference in proportions
- All follow general form of

\[ \text{Point Estimate} \pm z_{\alpha/2} \hat{SE} \]

- where $z_{\alpha/2} \hat{SE}$ refers to the margin of error
- CI’s most generally are:

\[ \text{Point Estimate} \pm \text{Margin of Error} \]
Confidence Intervals for Proportions

Sample proportion, \( \hat\pi \), has a normal sampling distribution under CLT

\[ \hat\pi \sim N(\pi, \pi(1-\pi)/n) \]

Because this is normal, use the same general guidelines, where

\[ SE[\hat\pi] = \sqrt{\hat\pi(1-\hat\pi)/n} \]

E.g., 95% CI: \( \hat\pi \pm 1.96 \cdot SE[\hat\pi] \)

E.g., 90% CI: \( \hat\pi \pm 1.645 \cdot SE[\hat\pi] \)

E.g., 99% CI: \( \hat\pi \pm 2.58 \cdot SE[\hat\pi] \)
Confidence Intervals for Proportions

- Sample proportion, \( \hat{\pi} \), has a normal sampling distribution under CLT
Confidence Intervals for Proportions

- Sample proportion, \( \hat{\pi} \), has a normal sampling distribution under CLT

\[
\hat{\pi} \sim N(\pi, \frac{\pi(1 - \pi)}{n})
\]
Confidence Intervals for Proportions

- Sample proportion, \( \hat{\pi} \), has a normal sampling distribution under CLT

\[
\hat{\pi} \sim N(\pi, \frac{\pi(1-\pi)}{n})
\]

- Because this is normal, use the same general guidelines, where

\[
SE[\hat{\pi}] = \sqrt{\hat{\pi}(1-\hat{\pi})/n}
\]
Confidence Intervals for Proportions

- Sample proportion, $\hat{\pi}$, has a normal sampling distribution under CLT

\[
\hat{\pi} \sim N(\pi, \frac{\pi(1-\pi)}{n})
\]

- Because this is normal, use the same general guidelines, where

\[
SE[\hat{\pi}] = \sqrt{\hat{\pi}(1-\hat{\pi})/n}
\]

- E.g., 95% CI: $\hat{\pi} \pm 1.96 SE[\hat{\pi}]$
Confidence Intervals for Proportions

- Sample proportion, $\hat{\pi}$, has a normal sampling distribution under CLT:
  \[ \hat{\pi} \sim N(\pi, \frac{\pi(1-\pi)}{n}) \]

- Because this is normal, use the same general guidelines, where
  \[ SE[\hat{\pi}] = \sqrt{\hat{\pi}(1-\hat{\pi})/n} \]

- E.g., 95% CI: $\hat{\pi} \pm 1.96SE[\hat{\pi}]$
- E.g., 90% CI: $\hat{\pi} \pm 1.645SE[\hat{\pi}]$
Confidence Intervals for Proportions

- Sample proportion, $\hat{\pi}$, has a normal sampling distribution under CLT

$$\hat{\pi} \sim N(\pi, \frac{\pi(1-\pi)}{n})$$

- Because this is normal, use the same general guidelines, where

$$SE[\hat{\pi}] = \sqrt{\hat{\pi}(1-\hat{\pi})/n}$$

- E.g., 95% CI: $\hat{\pi} \pm 1.96 \times SE[\hat{\pi}]$
- E.g., 90% CI: $\hat{\pi} \pm 1.645 \times SE[\hat{\pi}]$
- E.g., 99% CI: $\hat{\pi} \pm 2.58 \times SE[\hat{\pi}]$
Confidence Intervals for Difference in Means

\[
\hat{\mu}_1 - \hat{\mu}_2 \sim N(\mu_1 - \mu_2, \sigma_1^2/n_1 + \sigma_2^2/n_2)
\]

E.g., 95% CI: \( \hat{\pi} \pm 1.96 \times SE[\hat{\mu}_1 - \hat{\mu}_2] \)

E.g., 90% CI: \( \hat{\pi} \pm 1.645 \times SE[\hat{\mu}_1 - \hat{\mu}_2] \)

E.g., 99% CI: \( \hat{\pi} \pm 2.58 \times SE[\hat{\mu}_1 - \hat{\mu}_2] \)
Confidence Intervals for Difference in Means

- Difference in two sample means, $\bar{X}_1 - \bar{X}_2$, has a normal sampling distribution under CLT.
Confidence Intervals for Difference in Means

- Difference in two sample means, $\bar{X}_1 - \bar{X}_2$, has a normal sampling distribution under CLT

$$\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$
Confidence Intervals for Difference in Means

- Difference in two sample means, $\bar{X}_1 - \bar{X}_2$, has a normal sampling distribution under CLT

$$\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

- Because this is normal, use the same general guidelines, where

$$SE[\bar{X}_1 - \bar{X}_2] = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- E.g., 95% CI: $\hat{\pi} \pm 1.96 \cdot SE[\bar{X}_1 - \bar{X}_2]$ 

- E.g., 90% CI: $\hat{\pi} \pm 1.645 \cdot SE[\bar{X}_1 - \bar{X}_2]$ 

- E.g., 99% CI: $\hat{\pi} \pm 2.58 \cdot SE[\bar{X}_1 - \bar{X}_2]$. 
Confidence Intervals for Difference in Means

- Difference in two sample means, $\bar{X}_1 - \bar{X}_2$, has a normal sampling distribution under CLT

$$\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

- Because this is normal, use the same general guidelines, where

$$SE[\bar{X}_1 - \bar{X}_2] = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- E.g., 95% CI: $\hat{\pi} \pm 1.96SE[\bar{X}_1 - \bar{X}_2]$
Confidence Intervals for Difference in Means

- Difference in two sample means, $\bar{X}_1 - \bar{X}_2$, has a normal sampling distribution under CLT

$$\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

- Because this is normal, use the same general guidelines, where

$$SE[\bar{X}_1 - \bar{X}_2] = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- E.g., 95% CI: $\hat{\pi} \pm 1.96SE[\bar{X}_1 - \bar{X}_2]$
- E.g., 90% CI: $\hat{\pi} \pm 1.645SE[\bar{X}_1 - \bar{X}_2]$
Confidence Intervals for Difference in Means

- Difference in two sample means, $\bar{X}_1 - \bar{X}_2$, has a normal sampling distribution under CLT

$$\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma^2_1}{n_1} + \frac{\sigma^2_2}{n_2})$$

- Because this is normal, use the same general guidelines, where

$$SE[\bar{X}_1 - \bar{X}_2] = \sqrt{\frac{s^2_1}{n_1} + \frac{s^2_2}{n_2}}$$

- E.g., 95% CI: $\hat{\pi} \pm 1.96SE[\bar{X}_1 - \bar{X}_2]$

- E.g., 90% CI: $\hat{\pi} \pm 1.645SE[\bar{X}_1 - \bar{X}_2]$

- E.g., 99% CI: $\hat{\pi} \pm 2.58SE[\bar{X}_1 - \bar{X}_2]$
Confidence Intervals for Difference in Proportions

The difference in two sample proportions, $\hat{\pi}_1 - \hat{\pi}_2$, has a normal sampling distribution under CLT:

$$\sim N(\pi_1 - \pi_2, \pi_1(1 - \pi_1)/n_1 + \pi_2(1 - \pi_2)/n_2).$$

Because this is normal, use the same general guidelines, where non-pooled SE:

$$\text{SE}[\hat{\pi}_1 - \hat{\pi}_2] = \sqrt{\hat{\pi}_1(1 - \hat{\pi}_1)/n_1 + \hat{\pi}_2(1 - \hat{\pi}_2)/n_2}.$$

Examples:
- 95% CI: $\hat{\pi} \pm 1.96 \times \text{SE}[\hat{\pi}_1 - \hat{\pi}_2]$
- 90% CI: $\hat{\pi} \pm 1.645 \times \text{SE}[\hat{\pi}_1 - \hat{\pi}_2]$
- 99% CI: $\hat{\pi} \pm 2.58 \times \text{SE}[\hat{\pi}_1 - \hat{\pi}_2]$
Confidence Intervals for Difference in Proportions

- Difference in two sample proportions, $\hat{\pi}_1 - \hat{\pi}_2$, has a normal sampling distribution under CLT.
Confidence Intervals for Difference in Proportions

- Difference in two sample proportions, \( \hat{\pi}_1 - \hat{\pi}_2 \), has a normal sampling distribution under CLT

\[
\sim N(\pi_1 - \pi_2, \frac{\pi_1 (1 - \pi_1)}{n_1} + \frac{\pi_2 (1 - \pi_2)}{n_2})
\]
Confidence Intervals for Difference in Proportions

- Difference in two sample proportions, $\hat{\pi}_1 - \hat{\pi}_2$, has a normal sampling distribution under CLT

$$\sim N(\pi_1 - \pi_2, \frac{\pi_1(1 - \pi_1)}{n_1} + \frac{\pi_2(1 - \pi_2)}{n_2})$$

- Because this is normal, use the same general guidelines, where non-pooled $SE[\hat{\pi}_1 - \hat{\pi}_2] = \sqrt{\frac{\hat{\pi}_1(1 - \hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1 - \hat{\pi}_2)}{n_2}}$
Confidence Intervals for Difference in Proportions

- Difference in two sample proportions, $\hat{\pi}_1 - \hat{\pi}_2$, has a normal sampling distribution under CLT

$$\sim N(\pi_1 - \pi_2, \frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2})$$

- Because this is normal, use the same general guidelines, where non-pooled $SE[\hat{\pi}_1 - \hat{\pi}_2] = \sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}$

- E.g., 95% CI: $\hat{\pi} \pm 1.96SE[\hat{\pi}_1 - \hat{\pi}_2]$
- E.g., 90% CI: $\hat{\pi} \pm 1.645SE[\hat{\pi}_1 - \hat{\pi}_2]$
- E.g., 99% CI: $\hat{\pi} \pm 2.58SE[\hat{\pi}_1 - \hat{\pi}_2]$
Question asked by Gallup:
“In general, are you satisfied or dissatisfied with the way things are going in the United States at this time?”

Suppose of \( n = 1017 \) respondents, 248 said “yes, satisfied”

Calculate 95% confidence interval for true \( \pi \) (true share of Americans who think country moving in right direction)
Confidence Interval Example

▶ Question asked by Gallup:

“In general, are you satisfied or dissatisfied with the way things are going in the United States at this time?”

Suppose of $n = 1017$ respondents, 248 said “yes, satisfied”

Calculate 95% confidence interval for true $\pi$ (true share of Americans who think country moving in right direction)
Confidence Interval Example

▶ Question asked by Gallup:
▶ “In general, are you satisfied or dissatisfied with the way things are going in the United States at this time?”
Confidence Interval Example

- Question asked by Gallup:
- “In general, are you satisfied or dissatisfied with the way things are going in the United States at this time?”
- Suppose of \( n = 1017 \) respondents, 248 said “yes, satisfied”
Confidence Interval Example

- Question asked by Gallup:
- “In general, are you satisfied or dissatisfied with the way things are going in the United States at this time?”
- Suppose of $n = 1017$ respondents, 248 said “yes, satisfied”
- Calculate 95% confidence interval for true $\pi$ (true share of Americans who think country moving in right direction)
Confidence Interval Example

A 95% CI for $\pi$ is $\hat{\pi} \pm 1.96 \text{SE}\left[\hat{\pi}\right]$.

This means we have $0.24 \pm 1.96 \sqrt{0.24 \left(1 - 0.24\right)} / 1017 \Rightarrow [0.214, 0.266]$.

How do we interpret this?
Confidence Interval Example

- A 95% CI for $\pi$ is

\[
\hat{\pi} \pm 1.96 \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}
\]

This means we have

\[
0.24 \pm 1.96 \sqrt{\frac{0.24(1-0.24)}{1017}}
\]

\[
0.24 \pm 0.026
\]

→ [0.214, 0.266]

How do we interpret this?
Confidence Interval Example

- A 95% CI for \( \pi \) is

\[
\hat{\pi} \pm 1.96 SE[\hat{\pi}]
\]

This means we have

\[
0.24 \pm 1.96 \sqrt{0.24 \left(1 - 0.24\right) / 1017}
\]

\[
0.24 \pm 0.026 \rightarrow [0.214, 0.266]
\]

How do we interpret this?
Confidence Interval Example

- A 95% CI for \( \pi \) is
  \[
  \hat{\pi} \pm 1.96 SE[\hat{\pi}]
  \]

- where \( SE[\hat{\pi}] = \sqrt{\hat{\pi}(1 - \hat{\pi})/n} \)
Confidence Interval Example

- A 95% CI for $\pi$ is
  \[
  \hat{\pi} \pm 1.96 SE[\hat{\pi}]
  \]

- where $SE[\hat{\pi}] = \sqrt{\hat{\pi}(1 - \hat{\pi})/n}$
- This means we have

\[
\hat{\pi} \pm 0.026 \Rightarrow [0.214, 0.266]
\]
Confidence Interval Example

- A 95% CI for $\pi$ is

$$\hat{\pi} \pm 1.96 SE[\hat{\pi}]$$

- where $SE[\hat{\pi}] = \sqrt{\hat{\pi}(1 - \hat{\pi})/n}$
- This means we have

$$0.24 \pm 1.96 \sqrt{0.24(1 - 0.24)/1017}$$

$$0.24 \pm 0.026$$

$\rightarrow [0.214, 0.266]$
Confidence Interval Example

- A 95% CI for $\pi$ is
  \[ \hat{\pi} \pm 1.96 \text{SE}[\hat{\pi}] \]

- where $\text{SE}[\hat{\pi}] = \sqrt{\hat{\pi}(1 - \hat{\pi})/n}$
- This means we have
  \[ 0.24 \pm 1.96 \sqrt{0.24(1 - 0.24)/1017} \]
  \[ 0.24 \pm 0.026 \]
  \[ \rightarrow [0.214, 0.266] \]

- How do we interpret this?
How to Interpret Confidence Intervals?

▶ CIs one of most misinterpreted estimators
▶ Remember: Calculation of confidence interval depends on the sampling distribution (from CLT)
▶ Different sample → different confidence interval
▶ With some samples, calculated CI would "capture" true % of Americans satisfied with country direction
▶ With some samples, calculated CI would not "capture" true % of Americans satisfied with country direction
▶ For all of the CIs we could calculated with repeat sampling, 95% of them would cover true population parameter
▶ 95% of the time, the CI we construct in this fashion would capture the % of Americans satisfied with country direction
How to Interpret Confidence Intervals?

- CIs one of most misinterpreted estimators

- Calculation of confidence interval depends on the sampling distribution (from CLT)

- Different sample → different confidence interval

- With some samples, calculated CI would "capture" true % of Americans satisfied with country direction

- With some samples, calculated CI would not "capture" true % of Americans satisfied with country direction

- For all of the CIs we could calculated with repeat sampling, 95% of them would cover true population parameter

- 95% of the time, the CI we construct in this fashion would capture the % of Americans satisfied with country direction
How to Interpret Confidence Intervals?

- CIs one of most misinterpreted estimators
- Remember: Calculation of confidence interval depends on the sampling distribution (from CLT)
How to Interpret Confidence Intervals?

- CIs one of most misinterpreted estimators
- Remember: Calculation of confidence interval depends on the sampling distribution (from CLT)
- Different sample $\rightarrow$ different confidence interval
How to Interpret Confidence Intervals?

- CIs one of most misinterpreted estimators
- Remember: Calculation of confidence interval depends on the sampling distribution (from CLT)
- Different sample $\rightarrow$ different confidence interval
- With some samples, calculated CI would “capture” true % of Americans satisfied with country direction
How to Interpret Confidence Intervals?

- CIs one of most misinterpreted estimators
- Remember: Calculation of confidence interval depends on the sampling distribution (from CLT)
- **Different sample** $\rightarrow$ **different confidence interval**
- With some samples, calculated CI would “capture” true % of Americans satisfied with country direction
- With some samples, calculated CI would **not** “capture” true % of Americans satisfied with country direction
How to Interpret Confidence Intervals?

- CIs one of most misinterpreted estimators
- Remember: Calculation of confidence interval depends on the sampling distribution (from CLT)
- Different sample → different confidence interval
- With some samples, calculated CI would “capture” true % of Americans satisfied with country direction
- With some samples, calculated CI would not “capture” true % of Americans satisfied with country direction
- For all of the CIs we could calculated with repeat sampling, 95% of them would cover true population parameter
How to Interpret Confidence Intervals?

- CIs one of most misinterpreted estimators
- Remember: Calculation of confidence interval depends on the sampling distribution (from CLT)
- Different sample $\rightarrow$ different confidence interval
- With some samples, calculated CI would “capture” true % of Americans satisfied with country direction
- With some samples, calculated CI would not “capture” true % of Americans satisfied with country direction
- For all of the CIs we could calculated with repeat sampling, 95% of them would cover true population parameter
- 95% of the time, the CI we construct in this fashion would capture the % of Americans satisfied with country direction
How to Interpret CIs?

Show this with simulation

For sake of simulation, assume observations come from Normal distribution with mean 1 and variance of 10

I sample 500 observations, 100 times

For each sample, calculate 95% CI

\[
\bar{X} \pm 1.96 \times \hat{SE}
\]
How to Interpret CIs?

- Show this with simulation

Assume observations come from a Normal distribution with mean 1 and variance of 10. I sample 500 observations, 100 times. For each sample, calculate a 95% CI as $\bar{X} \pm 1.96 \times \hat{SE}$.
Show this with simulation

For sake of simulation, assume observations come from Normal distribution w/ mean 1 and variance of 10
How to Interpret CIs?

- Show this with simulation
- For sake of simulation, assume observations come from Normal distribution w/ mean 1 and variance of 10
- I sample 500 observations, 100 times
How to Interpret CIs?

- Show this with simulation
- For sake of simulation, assume observations come from Normal distribution w/ mean 1 and variance of 10
- I sample 500 observations, 100 times
- For each sample, calculate 95% CI
How to Interpret CIs?

- Show this with simulation
- For sake of simulation, assume observations come from Normal distribution w/ mean 1 and variance of 10
- I sample 500 observations, 100 times
- For each sample, calculate 95% CI
  - $\bar{X} \pm 1.96 \times SE[\bar{X}]$
How to Interpret CIs?
How to Interpret CIs?

500 observations drawn from $N(1,10)$
How to Interpret CIs?

500 observations drawn from N(1,10)
How to Interpret CIs?

500 observations drawn from N(1,10)
How to Interpret CIs?

500 observations drawn from $N(1,10)$
How to Interpret CIs?

500 observations drawn from N(1,10)
How to Interpret CIs?

For all of the CIs we could calculated with repeat sampling, 95% of them would cover true population parameter.

Confidence intervals one of most frequently misinterpreted estimators.

“There is a 95% probability that this interval I've calculated contains the true population parameter.”

→ Not correct: Once you have calculated the CI, it either contains the true value or not.

“95% of the confidence intervals I calculate using this formula using repeated sampling will contain the true population parameter.”

→ Correct!
How to Interpret CIs?

▶ For all of the CIs we could calculate with repeat sampling, 95% of them would cover the true population parameter.

(True)

("There is a 95% probability that this interval I've calculated contains the true population parameter")

(False): Once you have calculated the CI, it either contains the true value or not.
How to Interpret CIs?

- For all of the CIs we could calculated with repeat sampling, 95% of them would cover true population parameter

- Confidence intervals one of most frequently misinterpreted estimators
How to Interpret CIs?

- For all of the CIs we could calculated with repeat sampling, 95% of them would cover true population parameter.
- Confidence intervals one of most frequently misinterpreted estimators.
- “There is a 95% probability that this interval I’ve calculated contains the true population parameter.”
How to Interpret CIs?

- For all of the CIs we could calculated with repeat sampling, 95% of them would cover true population parameter
- Confidence intervals one of most frequently misinterpreted estimators
- “There is a 95% probability that this interval I’ve calculated contains the true population parameter”
  - → Not correct: Once you have calculated the CI, it either contains the true value or not
How to Interpret CIs?

- For all of the CIs we could calculated with repeat sampling, 95% of them would cover true population parameter.

- Confidence intervals one of most frequently misinterpreted estimators.

- “There is a 95% probability that this interval I’ve calculated contains the true population parameter.”
  - → Not correct: Once you have calculated the CI, it either contains the true value or not.

- “95% of the confidence intervals I calculate using this formula using repeated sampling will contain the true population parameter.”
How to Interpret CIs?

- For all of the CIs we could calculated with repeat sampling, 95% of them would cover true population parameter
- Confidence intervals one of most frequently misinterpreted estimators
- “There is a 95% probability that this interval I’ve calculated contains the true population parameter”
  - → Not correct: Once you have calculated the CI, it either contains the true value or not
- “95% of the confidence intervals I calculate using this formula using repeated sampling will contain the true population parameter”
  - → Correct!
CIs and Sample Size

Notes on sample size:

▶ Small samples → Follow same rules as hypothesis tests for when to switch to Student’s t distribution

▶ Larger sample size → will shrink standard errors → will lead to smaller CIs

▶ Ex) Doubling sample size will reduce the width of the confidence interval for a sample mean by a half

\[ \bar{X} \pm \frac{\sigma}{\sqrt{n}} \]

\[ \bar{X} \pm \frac{\sigma}{\sqrt{4n}} \rightarrow \bar{X} \pm \frac{1}{2} \frac{\sigma}{\sqrt{n}} \]

▶ Problems may ask you to calculate minimum sample size, given \( \alpha \) and standard deviation
Notes on sample size:

- Small samples → Follow same rules as hypothesis tests for when to switch to Student's t distribution
- Larger sample size → will shrink standard errors → will lead to smaller CIs
  - Ex) Doubling sample size will reduce the width of the confidence interval for a sample mean by a half
  \[ \bar{X} \pm \frac{\sigma}{\sqrt{n}} \]
  \[ \bar{X} \pm \frac{1}{2} \frac{\sigma}{\sqrt{n}} \]
CIs and Sample Size

Notes on sample size:

- Small samples → Follow same rules as hypothesis tests for when to switch to Student’s $t$ distribution

\[
\bar{X} \pm \frac{\sigma}{\sqrt{n}}
\]

- Larger sample size → will shrink standard errors → will lead to smaller CIs

- Ex) Doubling sample size will reduce the width of the confidence interval for a sample mean by a half

- Problems may ask you to calculate minimum sample size, given $\alpha$ and standard deviation
CIs and Sample Size

Notes on sample size:

- Small samples → Follow same rules as hypothesis tests for when to switch to Student’s $t$ distribution
- Larger sample size → will shrink standard errors → will lead to smaller CIs
  - Ex) Doubling sample size will reduce the width of the confidence interval for a sample mean by a half
CIs and Sample Size

Notes on sample size:

- Small samples → Follow same rules as hypothesis tests for when to switch to Student’s $t$ distribution

- Larger sample size → will shrink standard errors → will lead to smaller CIs
  - Ex) Doubling sample size will reduce the width of the confidence interval for a sample mean by a half
  - $\bar{X} \pm \sigma / \sqrt{n}$
CIs and Sample Size

Notes on sample size:

- Small samples → Follow same rules as hypothesis tests for when to switch to Student’s $t$ distribution
- Larger sample size → will shrink standard errors → will lead to smaller CIs
  - Ex) Doubling sample size will reduce the width of the confidence interval for a sample mean by a half
  - $\bar{X} \pm \sigma/\sqrt{n}$
  - $\bar{X} \pm \sigma/\sqrt{4n}$

Problems may ask you to calculate minimum sample size, given $\alpha$ and standard deviation
Notes on sample size:

- Small samples → Follow same rules as hypothesis tests for when to switch to Student’s *t* distribution
- Larger sample size → will shrink standard errors → will lead to smaller CIs
  - Ex) Doubling sample size will reduce the width of the confidence interval for a sample mean by a half
  - $\bar{X} \pm \frac{\sigma}{\sqrt{n}}$
  - $\bar{X} \pm \frac{\sigma}{\sqrt{4n}} \rightarrow \bar{X} \pm \frac{1}{2} \frac{\sigma}{\sqrt{n}}$
CIs and Sample Size

Notes on sample size:

▶ Small samples → Follow same rules as hypothesis tests for when to switch to Student’s $t$ distribution
▶ Larger sample size → will shrink standard errors → will lead to smaller CIs
   ▶ Ex) Doubling sample size will reduce the width of the confidence interval for a sample mean by a half
      ▶ $\bar{X} \pm \sigma/\sqrt{n}$
      ▶ $\bar{X} \pm \sigma/\sqrt{4n} \rightarrow \bar{X} \pm \frac{1}{2} \sigma/\sqrt{n}$
▶ Problems may ask you to calculate minimum sample size, given $\alpha$ and standard deviation
CIs and Sample Size

What happens to CIs when \( n \) goes up?

- 500 observations drawn from \( N(1,10) \)
- 95% Confidence Interval
CIs and Sample Size

Show this graphically $\rightarrow$ What happens to CIs when $n$ goes up?
CIs and Sample Size

Show this graphically → What happens to CIs when $n$ goes up?
CIs and Sample Size

Show this graphically → What happens to CIs when $n$ goes up?

1500 observations drawn from N(1,10)
Cls versus Hypothesis Tests

- Close relationship between CIs and HTs
  - If value $A$ not in 95% CI → would be rejected by a two-sided hypothesis test at the 5% level
  - If value $A$ in 95% CI → would not be rejected by a two-sided hypothesis test at the 5% level
- A 95% confidence interval is all of null hypotheses that would not be rejected at the 0.05 level
- A $(1 - \alpha)$% confidence interval is all of null hypotheses that would not be rejected at the $\alpha$ level
CIs versus Hypothesis Tests

- Close relationship between CIs and HTs

  - If value $A$ not in 95% CI $\rightarrow$ would be rejected by a two-sided hypothesis test at the 5% level
  - If value $A$ in 95% CI $\rightarrow$ would not be rejected by a two-sided hypothesis test at the 5% level

- A 95% confidence interval is all of null hypotheses that would not be rejected at the 0.05 level

- $\rightarrow$ A $(1 - \alpha)$% confidence interval is all of null hypotheses that would not be rejected at the $\alpha$ level
Cls versus Hypothesis Tests

- Close relationship between CIs and HTs
- If value $A$ not in 95% CI → would be rejected by a two-sided hypothesis test at the 5% level
- If value $A$ in 95% CI → would not be rejected by a two-sided hypothesis test at the 5% level

A $(1-\alpha)$% confidence interval is all of null hypotheses that would not be rejected at the $\alpha$ level
CIs versus Hypothesis Tests

- Close relationship between CIs and HTs
- If value $A$ not in 95% CI $\rightarrow$ would be rejected by a two-sided hypothesis test at the 5% level
- If value $A$ in 95% CI $\rightarrow$ would not be rejected by a two-sided hypothesis test at the 5% level
CIs versus Hypothesis Tests

- Close relationship between CIs and HTs
- If value $A$ not in 95% CI $\rightarrow$ would be rejected by a two-sided hypothesis test at the 5% level
- If value $A$ in 95% CI $\rightarrow$ would not be rejected by a two-sided hypothesis test at the 5% level
- $\rightarrow$ A 95% confidence interval is all of null hypotheses that would not be rejected at the 0.05 level
CIs versus Hypothesis Tests

- Close relationship between CIs and HTs
  - If value $A$ not in 95% CI $\rightarrow$ would be rejected by a two-sided hypothesis test at the 5% level
  - If value $A$ in 95% CI $\rightarrow$ would not be rejected by a two-sided hypothesis test at the 5% level
- $\rightarrow$ A 95% confidence interval is all of null hypotheses that would not be rejected at the 0.05 level
- $\rightarrow$ A $(1 - \alpha)$% confidence interval is all of null hypotheses that would not be rejected at the $\alpha$ level
CIs versus Hypothesis Tests

Both CIs and HTs are useful tools in your inference toolkit.

HTs are useful when comparing groups or trying to test a theory.

CIs are useful for thinking about the range of possible values, providing additional information about a single sample.

Many people prefer CIs:

- They can give you information over all possible null hypotheses that would be rejected (as opposed to one), conditional on the $\alpha$ value.
- Many find the margin of error intuitive (although many incorrectly interpret it).
CIs versus Hypothesis Tests

- Both CIs and HTs useful tools in your inference toolkit
CIs versus Hypothesis Tests

- Both CIs and HTs useful tools in your inference toolkit
- HTs → useful when comparing groups or trying to test a theory

- Many people prefer CIs:
  - Can give you information over all possible null hypotheses that would be rejected (as opposed to one), conditional on $\alpha$ value
  - Many find margin of error intuitive (although many incorrectly interpret)
CIs versus Hypothesis Tests

- Both CIs and HTs useful tools in your inference toolkit
- HTs → useful when comparing groups or trying to test a theory
- CIs → useful for thinking about range of possible values, providing additional information about a single sample
Both CIs and HTs useful tools in your inference toolkit

HTs → useful when comparing groups or trying to test a theory

CIs → useful for thinking about range of possible values, providing additional information about a single sample

Many people prefer CIs:
CIs versus Hypothesis Tests

- Both CIs and HTs useful tools in your inference toolkit
- HTs → useful when comparing groups or trying to test a theory
- CIs → useful for thinking about range of possible values, providing additional information about a single sample
- Many people prefer CIs:
  - Can give you information over all possible null hypotheses that would be rejected (as opposed to one), conditional on $\alpha$ value
CIs versus Hypothesis Tests

- Both CIs and HTs useful tools in your inference toolkit
- HTs → useful when comparing groups or trying to test a theory
- CIs → useful for thinking about range of possible values, providing additional information about a single sample
- Many people prefer CIs:
  - Can give you information over all possible null hypotheses that would be rejected (as opposed to one), conditional on $\alpha$ value
  - Many find margin of error intuitive (although many incorrectly interpret)
Next time

- Comparing groups that have paired data