Midterm 2 Review

March 27, 2019

Coming up…

- **Problem Set 5** due today 11:55pm
- **Performance Assessment 5** due Sunday 11:55pm.
- **Lab Assignment** due next Thursday 4/4.

- **Midterm 1 Review** tomorrow in labs!
- **Midterm 1** Monday 4/1

Feedback + Help
- *Problem Set 4 already returned.*
- *Problem Set 5 TAs can give feedback in office hours.*

Midterm 2

- **NEW Material Covered:**
  - Units 3.3, Unit 4, Unit 5
  - Quasi-cumulative!
- **What to bring:**
  - Cheat sheet
    - 1 page (8.5'' by 11'')
    - Front/back ok
    - CAN be typed
  - Calculator (no phones)
- **Provided:**
  - Z-tables
  - T-tables
  - Chi-Squared tables
- **Exam**
  - 3 written questions (like AEs)
  - 5 T/F
  - 10 MC

1

Midterm 2 Review Suggestions

- **Short answer *review*:**
  - Make sure you understand how to do the application exercises.
  - Review Problem Sets (graded)
- **Short answer *practice*:**
  - Tomorrow's lab practice problem review
  - Practice test
  - Suggested practice problems in the book

1

Midterm 2 Short Answer Structure

Full Analysis

•Determine which graphs might be appropriate for visualizing this original raw data
•Determine which test is appropriate.
•Set up hypotheses.
•Check conditions.
•Calculate test statistic
•Calculate p-value
•Interpret p-value/make conclusion in the context of the data.
•Perhaps also calculate confidence intervals, and make other interpretations about things we do in the analysis.

1

Final Review Suggestions

- **Concept review:**
  - Lecture slides **(has material not in the videos/book)**
  - Video notes
  - Readiness Assessments+Performance Assessments
    - Why are all the other options wrong?
  - What to think about (among other things):
    - Interpretations of analyses (WORDING IS IMPORTANT)
    - Conclusions we would make (WORDING IS IMPORTANT)
    - Equations
    - Know exact definitions
    - FOCUS ON THE WHY BEHIND ANALYSES
    - If there's an equation/analysis, make sure you know how to put that equation/analysis into words in the context of the problem.
    - Conditions
    - Common misconceptions (lecture notes)
    - What test to use under certain conditions

1

How do MT1 material and MT2 material relate? Will MT1 material be on MT2?

MT1
Foundations for Inference

- **What does a p-value represent? How to put it in the context of the data.**
- **How do we calculate a p-value with:**
  - The sampling distribution?
  - A randomization distribution?
  - A bootstrap distribution?

1

How do MT1 material and MT2 material relate? Will MT1 material be on MT2?

MT1
Foundations for Inference

- **What does a p-value represent? How to put it in the context of the data.**
- **How do we calculate a p-value with:**
  - The sampling distribution?
  - A randomization distribution?
  - A bootstrap distribution?
- **When calculating a p-value with the following, what should they be centered at? What do they all need to assume?**
  - The sampling distribution?
  - A randomization distribution?
  - A bootstrap distribution?

1

How do MT1 material and MT2 material relate? Will MT1 material be on MT2?

MT1
Foundations for Inference

- **What does a p-value represent? How to put it in the context of the data.**
- **How do we calculate a p-value with:**
  - The sampling distribution?
  - A randomization distribution?
  - A bootstrap distribution?
- **When calculating a p-value with the following, what should they be centered at? What do they all need to assume?**
  - The sampling distribution?
  - A randomization distribution?
  - A bootstrap distribution?

- **What does a confidence interval represent?**
  - <u>How do we put it into words</u>. "We are XX% confident that the true pop. Parameter is in the interval."
  - XX% of random samples of size n will have confidence interval that contains the pop. Parameter.

1

How do MT1 material and MT2 material relate? Will MT1 material be on MT2?

MT1
Foundations for Inference

- **What does a p-value represent? How to put it in the context of the data.**
- **How do we calculate a p-value with:**
  - The sampling distribution?
  - A randomization distribution?
  - A bootstrap distribution?
- **When calculating a p-value with the following, what should they be centered at? What do they all need to assume?**
  - The sampling distribution?
  - A randomization distribution?
  - A bootstrap distribution?

- **What does a confidence interval represent?**
  - <u>How do we put it into words</u>. "We are XX% confident that the true pop. Parameter is in the interval."
  - XX% of random samples of size n will have confidence interval that contains the pop. Parameter.
- **Why is the Central Limit Theorem important?**
  - It tells us when a sampling distribution is normal.
  - Remember properties of normal distributions!

1

How do MT1 material and MT2 material relate? Will MT1 material be on MT2? –Yes!

MT1
Foundations for Inference

Material after MT1
How do we *specifically* conduct inference for various combinations of numerical/categorical(2 vs. >2 levels) variables?

- **What does a p-value represent? How to put it in the context of the data.**
- **How do we calculate a p-value with:**
  - The sampling distribution?
  - A randomization distribution?
  - A bootstrap distribution?
- **When calculating a p-value with the following, what should they be centered at? What do they all need to assume?**
  - The sampling distribution?
  - A randomization distribution?
  - A bootstrap distribution?

- **What does a confidence interval represent?**
  - <u>How do we put it into words</u>. "We are XX% confident that the true pop. Parameter is in the interval."
  - XX% of random samples of size n will have confidence interval that contains the pop. Parameter.
- **Why is the Central Limit Theorem important?**
  - It tells us when a sampling distribution is normal.
  - Remember properties of normal distributions!

1

Conducting an Analysis… where to start?

Conducting an Analysis… where to start?

- How many numerical and categorical variables are involved?
- If categorical variables, how many levels do they have?
- What visualization would be appropriate for visualizing this data?

Conducting an Analysis… where to start?

- How many numerical and categorical variables are involved?
- If categorical variables, how many levels do they have?
- What visualization would be appropriate for visualizing this data?

- Is it a combo of variables that **has one specified Population Parameter of Interest** where
  - We can create a **confidence interval** for the population parameter.
  - We can conduct a **hypothesis test** for the population parameter using:
    - $Ho: Pop.Param = \#$
    - $Ha: Pop.Param\ (\neq\ or\ <\ or\ >)\ \#?$

- Is it a combo of variables that **doesn't have one specified Population Parameter of Interest** and
  - We can't make a **confidence interval**.
  - Has a different **hypothesis test** set up?

Conducting an Analysis… where to start?

- How many numerical and categorical variables are involved?
- If categorical variables, how many levels do they have?
- What visualization would be appropriate for visualizing this data?

- Is it a combo of variables that **has one specified Population Parameter of Interest** where
  - We can create a **confidence interval** for the population parameter.
  - We can conduct a **hypothesis test** for the population parameter using:
    - $Ho: Pop.Param = \#$
    - $Ha: Pop.Param\ (\neq\ or\ <\ or\ >)\ \#?$

- If we are making a confidence interval or hypothesis test for this population parameter, when should we use each of the following **methods** to do this?
  - CLT methods
  - Randomization testing methods
  - Bootstrap methods

- Is it a combo of variables that **doesn't have one specified Population Parameter of Interest** and
  - We can't make a **confidence interval**.
  - Has a different **hypothesis test** set up?

Conducting an Analysis… where to start?

- How many numerical and categorical variables are involved?
- If categorical variables, how many levels do they have?
- What visualization would be appropriate for visualizing this data?

- Is it a combo of variables that **has one specified Population Parameter of Interest** where
  - We can create a **confidence interval** for the population parameter.
  - We can conduct a **hypothesis test** for the population parameter using:
    - $Ho: Pop.Param = \#$
    - $Ha: Pop.Param\ (\neq\ or\ <\ or\ >)\ \#?$

- If we are making a confidence interval or hypothesis test for this population parameter, when should we use each of the following **methods** to do this?
  - CLT methods
  - Randomization testing methods
  - Bootstrap methods

		Analyses and Ways to Conduct Them (that we know so far)	
Types of Variable(s) Involved	**Population Parameter**	**Confidence Interval for the Population Parameter**	**Hypothesis Test for the Population Parameter** Ho: Pop. Param = # Ha: Pop. Param (≠ or < or >) #
Single Numerical Variable	μ	CLT Confidence Interval (Unit 3+4) Bootstrap Confidence Interval (Unit 4)	CLT Hypothesis Test (Unit 3+4) Bootstrap Hypothesis Test (Unit 4)
	μdiff	CLT Confidence Interval (Unit 4) Bootstrap Confidence Interval (Unit 4)	CLT Hypothesis Test (Unit 4) Bootstrap Hypothesis Test (Unit 4)
	Median	Bootstrap Confidence Interval (Unit 4)	Bootstrap Hypothesis Test (Unit 4)
Single Categorical Variable (2 levels)	p	CLT Confidence Interval (Unit 5) Bootstrap Confidence Interval (Unit 4)	CLT Hypothesis Test (Unit 5) Bootstrap Hypothesis Test (Unit 4) Randomization Testing (Unit 5-Selecting ball/chips out of bag, rolling dice)
Numerical Response Variable Categorical Explanatory Variable (2 levels)	μ1-μ2	CLT Confidence Interval (Unit 4) Bootstrap Confidence Interval (Unit 5)	CLT Hypothesis Test (Unit 4) Randomization Testing (Unit 1-Shuffling Cards)
	Median1-Median2	Bootstrap Confidence Interval (Unit 5)	Randomization Testing (Unit 1-Shuffling Cards)
Categorical Response Variable Categorical Explanatory Variable (both have 2 levels)	p1-p2	CLT Confidence Interval (Unit 5) Bootstrap Confidence Interval (Unit 5)	CLT Hypothesis Test (Unit 5) Randomization Testing (Unit 1-Shuffling Cards)

- Is it a combo of variables that **doesn't have one specified Population Parameter of Interest** and
  - We can't make a **confidence interval**.
  - Has a different **hypothesis test** set up?

Types of Variables	Analysis	Hypotheses
Numerical Response Variable Categorical Explanatory Variable (>2 levels)	**ANOVA**	Ho: μ1=μ2=…=μk Ha: at least one pair of groups has means that are different
Single Categorical Variable (>2 levels)	**Chi-Squared Goodness of Fit Test**	Ho: The data follows the specified distribution. Ha: The data does not follow the specified distribution.
Categorical Response Variable Categorical Explanatory Variable (at least one has >2 levels)	**Chi-Squared Independence Test**	Ho: The two variables are independent/not associated. Ha: The two variables are dependent/associated.

Conducting an Analysis… where to start?

- How many numerical and categorical variables are involved?
- If categorical variables, how many levels do they have?
- What visualization would be appropriate for visualizing this data?

- Is it a combo of variables that **has one specified Population Parameter of Interest** where
  - We can create a **confidence interval** for the population parameter.
  - We can conduct a **hypothesis test** for the population parameter using:
    - $Ho: Pop.Param = \#$
    - $Ha: Pop.Param\ (\neq \text{ or } < \text{ or } >)\ \#?$

- If we are making a confidence interval or hypothesis test for this population parameter, when should we use each of the following **methods** to do this?
  - CLT methods (if CLT conditions met)
  - Randomization testing methods (just for HT)
  - Bootstrap methods

		Analyses and Ways to Conduct Them (that we know so far)	
Types of Variable(s) Involved	Population Parameter	Confidence Interval for the Population Parameter	Hypothesis Test for the Population Parameter Ho: Pop. Param = # Ha: Pop. Param (≠ or < or >) #
Single Numerical Variable	μ	CLT Confidence Interval (Unit 3+4) Bootstrap Confidence Interval (Unit 4)	CLT Hypothesis Test (Unit 3+4) Bootstrap Hypothesis Test (Unit 4)
	μdiff	CLT Confidence Interval (Unit 4) Bootstrap Confidence Interval (Unit 4)	CLT Hypothesis Test (Unit 4) Bootstrap Hypothesis Test (Unit 4)
	Median	Bootstrap Confidence Interval (Unit 4)	Bootstrap Hypothesis Test (Unit 4)
Single Categorical Variable (2 levels)	p	CLT Confidence Interval (Unit 5) Bootstrap Confidence Interval (Unit 4)	CLT Hypothesis Test (Unit 5) Bootstrap Hypothesis Test (Unit 4) Randomization Testing (Unit 5-Selecting ball/chips out of bag, rolling dice)
Numerical Response Variable Categorical Explanatory Variable (2 levels)	μ1-μ2	CLT Confidence Interval (Unit 4) Bootstrap Confidence Interval (Unit 5)	CLT Hypothesis Test (Unit 4) Randomization Testing (Unit 1-Shuffling Cards)
	Median1-Median2	Bootstrap Confidence Interval (Unit 5)	Randomization Testing (Unit 1-Shuffling Cards)
Categorical Response Variable Categorical Explanatory Variable (both have 2 levels)	p1-p2	CLT Confidence Interval (Unit 5) Bootstrap Confidence Interval (Unit 5)	CLT Hypothesis Test (Unit 5) Randomization Testing (Unit 1-Shuffling Cards)

- Is it a combo of variables that **doesn't have one specified Population Parameter of Interest** and
  - We can't make a **confidence interval**.
  - Has a different **hypothesis test** set up?

Types of Variables	Analysis	Hypotheses
Numerical Response Variable Categorical Explanatory Variable (>2 levels)	**ANOVA**	Ho: μ1=μ2=…=μk Ha: at least one pair of groups has means that are different
Single Categorical Variable (>2 levels)	**Chi-Squared Goodness of Fit Test**	Ho: The data follows the specified distribution. Ha: The data does not follow the specified distribution.
Categorical Response Variable Categorical Explanatory Variable (at least one has >2 levels)	**Chi-Squared Independence Test**	Ho: The two variables are independent/not associated. Ha: The two variables are dependent/associated.

One Population Parameter of Interest

		Analyses and Ways to Conduct Them (that we know so far)	
Types of Variable(s) Involved	Population Parameter	Confidence Interval for the Population Parameter	Hypothesis Test for the Population Parameter Ho: Pop. Param = # Ha: Pop. Param (≠ or < or >) #
Single Numerical Variable	μ	CLT Confidence Interval (Unit 3+4) Bootstrap Confidence Interval (Unit 4)	CLT Hypothesis Test (Unit 3+4) Bootstrap Hypothesis Test (Unit 4)
	μdiff	CLT Confidence Interval (Unit 4) Bootstrap Confidence Interval (Unit 4)	CLT Hypothesis Test (Unit 4) Bootstrap Hypothesis Test (Unit 4)
	Median	Bootstrap Confidence Interval (Unit 4)	Bootstrap Hypothesis Test (Unit 4)
Single Categorical Variable (2 levels)	p	CLT Confidence Interval (Unit 5) Bootstrap Confidence Interval (Unit 4)	CLT Hypothesis Test (Unit 5) Bootstrap Hypothesis Test (Unit 4) Randomization Testing (Unit 5-Selecting ball/chips out of bag, rolling dice)
Numerical Response Variable Categorical Explanatory Variable (2 levels)	μ1-μ2	CLT Confidence Interval (Unit 4) Bootstrap Confidence Interval (Unit 5)	CLT Hypothesis Test (Unit 4) Randomization Testing (Unit 1-Shuffling Cards)
	Median1-Median2	Bootstrap Confidence Interval (Unit 5)	Randomization Testing (Unit 1-Shuffling Cards)
Categorical Response Variable Categorical Explanatory Variable (both have 2 levels)	p1-p2	CLT Confidence Interval (Unit 5) Bootstrap Confidence Interval (Unit 5)	CLT Hypothesis Test (Unit 5) Randomization Testing (Unit 1-Shuffling Cards)

What are these CLT conditions?

When other certain conditions are met…

$$\bar{x} \sim N(mean = \mu, standard\ dev./error = \frac{\sigma}{\sqrt{n}})$$

and we can make CLT CIs and HTs for μ

When other certain conditions are met…

$$\bar{x}_1 - \bar{x}_2 \sim N(mean = \mu_1 - \mu_2, standard\ dev./error = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$$

and we can make CLT CIs and HTs for μ1-μ2

When other certain conditions are met…

$$\hat{p} \sim N(mean = p, standard\ dev./error = \sqrt{\frac{p(1-p)}{n}})$$

and we can make CLT CIs and HTs for p

When other certain conditions are met…

$$\hat{p}_1 - \hat{p}_2 \sim N(mean = p_1 - p_2, standard\ dev./error = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}})$$

and we can make CLT CIs and HTs for p1-p2

When other certain conditions are met…

$$\bar{x} \sim N(mean = \mu, standard\ dev./error = \frac{\sigma}{\sqrt{n}})$$

and we can make CLT CIs and HTs for μ

When other certain conditions are met…

$$\bar{x}_1 - \bar{x}_2 \sim N(mean = \mu_1 - \mu_2, standard\ dev./error = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$$

and we can make CLT CIs and HTs for μ1-μ2

When other certain conditions are met…

$$\hat{p} \sim N(mean = p, standard\ dev./error = \sqrt{\frac{p(1-p)}{n}})$$

and we can make CLT CIs and HTs for p

When other certain conditions are met…

$$\hat{p}_1 - \hat{p}_2 \sim N(mean = p_1 - p_2, standard\ dev./error = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}})$$

and we can make CLT CIs and HTs for p1-p2

Be sure to refresh on properties of Normal Distributions!

(Deck 2.3)

What are some properties of normal distributions that others don't have?

Central Limit Theorem

Confidence Interval for Population Parameter

$$(point\ estimate) \pm (crit.value)SE$$

When CLT conditions are met.

Types of Variable(s) Involved	Population Parameter	Point Estimate	Standard Error	Distribution: (1) To get Critical Values From (CI) (2) That the Test Statistic Follows (HT)
Single **Numerical** Variable	μ	$\bar{x}$	$\frac{\sigma}{\sqrt{n}}$	Z (if you know σ) T_{n-1} (if you don't know σ)
	μdiff	$\bar{x}_{diff}$	$\frac{\sigma_{diff}}{\sqrt{n}}$	Z (if you know σdiff) T_{n-1} (if you don't know σdiff)
Single **Categorical** Variable *(2 levels)*	p	$\hat{p}$	For Confidence Intervals: $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ For Hypothesis Tests: $\sqrt{\frac{p_0(1-p_0)}{n}}$	Z
Numerical Response Variable **Categorical** Explanatory Variable *(2 levels)*	μ1-μ2	$\bar{x}_1 - \bar{x}_2$	$\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$	Z (if you know σ1 and σ2) $T_{min(n1-1,n2-1)}$ (if you don't know σ1 or σ2)
Categorical Response Variable **Categorical** Explanatory Variable *(both have 2 levels)*	p1-p2	$\hat{p}_1 - \hat{p}_2$	For Confidence Intervals: $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ For Hypothesis Tests with Ho:p1-p2=0: $\sqrt{\frac{\hat{p}_{pool}(1-\hat{p}_{pool})}{n_1}+\frac{\hat{p}_{pool}(1-\hat{p}_{pool})}{n_2}}$	Z

Central Limit Theorem

Hypothesis Testing for Population Parameter

$Ho: pop.param = null\ value$
$Ha: pop.param (\neq or > or <) null\ value$

$$Test-Stat = \frac{(point\ estimate) - null\ value}{SE}$$

When CLT conditions are met.

Types of Variable(s) Involved	Population Parameter	Point Estimate	Standard Error	Distribution: (1) To get Critical Values From (CI) (2) That the Test Statistic Follows (HT)
Single **Numerical** Variable	μ	$\bar{x}$	$\frac{\sigma}{\sqrt{n}}$	Z (if you know σ) T_{n-1} (if you don't know σ)
	μdiff	$\bar{x}_{diff}$	$\frac{\sigma_{diff}}{\sqrt{n}}$	Z (if you know σdiff) T_{n-1} (if you don't know σdiff)
Single **Categorical** Variable *(2 levels)*	p	$\hat{p}$	For Confidence Intervals: $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ For Hypothesis Tests: $\sqrt{\frac{p_0(1-p_0)}{n}}$	Z
Numerical Response Variable **Categorical** Explanatory Variable *(2 levels)*	μ1-μ2	$\bar{x}_1 - \bar{x}_2$	$\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$	Z (if you know σ1 and σ2) $T_{min(n1-1,n2-1)}$ (if you don't know σ1 or σ2)
Categorical Response Variable **Categorical** Explanatory Variable *(both have 2 levels)*	p1-p2	$\hat{p}_1 - \hat{p}_2$	For Confidence Intervals: $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ For Hypothesis Tests with Ho:p1-p2=0: $\sqrt{\frac{\hat{p}_{pool}(1-\hat{p}_{pool})}{n_1}+\frac{\hat{p}_{pool}(1-\hat{p}_{pool})}{n_2}}$	Z

Which test to use? Calculating p-values appropriately

Clicker question

A group of researchers want to see if there is a difference in the average SAT score between the oldest and youngest child in a family. 20 families with just two children were sampled and the SAT score was recorded for each child. The test statistic for this test is *T = -1.8.* What is the p-value?

(a) 0.01 < p-value < 0.025
(b) 0.025 < p-value < 0.05
(c) 0.05 < p-value < 0.10
(d) 0.10 < p-value < 0.20

Clicker question

A group of researchers want to see if there is a difference in the average SAT score between the oldest and youngest child in a family. 20 families with just two children were sampled and the SAT score was recorded for each child. The test statistic for this test is *T = -1.8.* What is the p-value?

(a) 0.01 < p-value < 0.025
(b) 0.025 < p-value < 0.05
(c) 0.05 < p-value < 0.10
(d) 0.10 < p-value < 0.20

Population Parameter of Interest:

- **μ_{diff} (Paired Means Test) OR**
- **$\mu_{oldest} - \mu_{youngest}$ (Independent Means Test)?**

Clicker question

A group of researchers want to see if there is a difference in the average SAT score between the oldest and youngest child in a family. 20 families with just two children were sampled and the SAT score was recorded for each child. The test statistic for this test is *T = -1.8.* What is the p-value?

(a) 0.01 < p-value < 0.025
(b) 0.025 < p-value < 0.05
(c) 0.05 < p-value < 0.10
(d) 0.10 < p-value < 0.20

Population Parameter of Interest:

- **μ_{diff} (Paired Means Test)**

$Ho: \mu_{diff} = 0$

$Ha: \mu_{diff} \neq 0$

T = -1.8

Degrees of Freedom = n-1 = 20-1 = 19

Clicker question

A group of researchers want to see if there is a difference in the average SAT score between the oldest and youngest child in a family. 20 families with just two children were sampled and the SAT score was recorded for each child. The test statistic for this test is *T = -1.8.* What is the p-value?

(a) 0.01 < p-value < 0.025
(b) 0.025 < p-value < 0.05
(c) 0.05 < p-value < 0.10
(d) 0.10 < p-value < 0.20

Population Parameter of Interest:

- **μ_{diff} (Paired Means Test)**

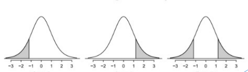
$Ho: \mu_{diff} = 0$

$Ha: \mu_{diff} \neq 0$

T = -1.8

Degrees of Freedom = n-1 = 20-1 = 19

t distribution probability table



one tail	0.100	0.050	0.025	0.010	0.005
two tails	0.200	0.100 *	0.050	0.020	0.010
df 1	3.08	6.31	12.71	31.82	63.66
19	1.33	1.73	2.09	2.54	2.86

Take absolute value.

Clicker question

A group of researchers want to see if there is a difference in the average SAT score between the oldest and youngest child in a family. 20 families with just two children were sampled and the SAT score was recorded for each child. The test statistic for this test is *T = -1.8.* What is the p-value?

(a) 0.01 < p-value < 0.025
(b) 0.025 < p-value < 0.05
(c) 0.05 < p-value < 0.10
(d) 0.10 < p-value < 0.20

Population Parameter of Interest:

- **μ_{diff} (Paired Means Test)**

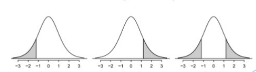
$Ho: \mu_{diff} = 0$

$Ha: \mu_{diff} \neq 0$

T = -1.8

Degrees of Freedom = n-1 = 20-1 = 19

t distribution probability table



one tail	0.100	0.050	0.025	0.010	0.005
two tails	0.200	0.100 *	0.050	0.020	0.010
df 1	3.08	6.31	12.71	31.82	63.66
19	1.33	1.73	2.09	2.54	2.86

Take absolute value.

Clicker question

A group of researchers want to see if there is a difference in the average SAT score between the oldest and youngest child in a family. 20 families with just two children were sampled and the SAT score was recorded for each child. The test statistic for this test is *T = -1.8*. What is the p-value?

(a) 0.01 < p-value < 0.025
(b) 0.025 < p-value < 0.05
(c) 0.05 < p-value < 0.10
(d) 0.10 < p-value < 0.20

Population Parameter of Interest:

- **μ_{diff} (Paired Means Test)**

$Ho: \mu_{diff} = 0$

$Ha: \mu_{diff} \neq 0$

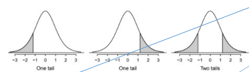
Two-tailed test = two tails in t-table

T = -1.8

Degrees of Freedom = n-1 = 20-1 = 19

Take absolute value.

t distribution probability table



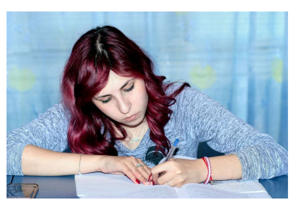
one tail		0.100	0.050	0.025	0.010	0.005
two tails		0.200	0.100 *	0.050	0.020	0.010
df	1	3.08	6.31	12.71	31.82	63.66
	19	1.33	1.73	2.09	2.54	2.86

Which test to use? Calculating p-values appropriately

Clicker question

A group of researchers want to see if the average SAT score of first-born children is higher than that of last-born children. 10 first-born children were sampled and 10 last-born children were sampled. The test statistic for this test is *T = -1.8*. What is the p-value?

(a) 0.01 < p-value < 0.025
(b) 0.025 < p-value < 0.05
(c) 0.05 < p-value < 0.10
(d) 0.10 < p-value < 0.20



Clicker question

A group of researchers want to see if the average SAT score of first-born children is higher than that of last-born children. 10 first-born children were sampled and 10 last-born children were sampled. The test statistic for this test is *T = -1.8*. What is the p-value?

(a) 0.01 < p-value < 0.025
(b) 0.025 < p-value < 0.05
(c) 0.05 < p-value < 0.10
(d) 0.10 < p-value < 0.20

Population Parameter of Interest:

- **μ_{diff} (Paired Means Test) OR**
- **$\mu_{oldest} - \mu_{youngest}$ (Independent Means Test)?**

Clicker question

A group of researchers want to see if the average SAT score of first-born children is higher than that of last-born children. 10 first-born children were sampled and 10 last-born children were sampled. The test statistic for this test is *T = -1.8.* What is the p-value?

(a) 0.01 < p-value < 0.025
(b) 0.025 < p-value < 0.05
(c) 0.05 < p-value < 0.10
(d) 0.10 < p-value < 0.20

Population Parameter of Interest:

- **$\mu_{oldest} - \mu_{youngest}$ (Independent Means Test)**

$Ho: \mu_{oldest} - \mu_{youngest} = 0$

$Ha: \mu_{oldest} - \mu_{youngest} > 0$

T = -1.8

Degrees of Freedom = min(n_{oldest}-1, $n_{youngest}$-1)

= min(10-1,10-1) = 9

Clicker question

A group of researchers want to see if the average SAT score of first-born children is higher than that of last-born children. 10 first-born children were sampled and 10 last-born children were sampled. The test statistic for this test is *T = -1.8.* What is the p-value?

(a) 0.01 < p-value < 0.025
(b) 0.025 < p-value < 0.05
(c) 0.05 < p-value < 0.10
(d) 0.10 < p-value < 0.20

Population Parameter of Interest:

- **$\mu_{oldest} - \mu_{youngest}$ (Independent Means Test)**

$Ho: \mu_{oldest} - \mu_{youngest} = 0$

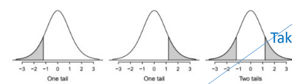
$Ha: \mu_{oldest} - \mu_{youngest} > 0$

T = -1.8

Degrees of Freedom = min(n_{oldest}-1, $n_{youngest}$-1)

= min(10-1,10-1) = 9

t distribution probability table



Take absolute value.

one tail	0.100	0.050	0.025	0.010	0.005
two tails	0.200	0.100	0.050	0.020	0.010
df 1	3.08	6.31	12.71	31.82	63.66
9	1.38	1.83	2.26	2.82	3.25

Clicker question

A group of researchers want to see if the average SAT score of first-born children is higher than that of last-born children. 10 first-born children were sampled and 10 last-born children were sampled. The test statistic for this test is *T = -1.8.* What is the p-value?

(a) 0.01 < p-value < 0.025
(b) 0.025 < p-value < 0.05
(c) 0.05 < p-value < 0.10
(d) 0.10 < p-value < 0.20

Population Parameter of Interest:

- **$\mu_{oldest} - \mu_{youngest}$ (Independent Means Test)**

$Ho: \mu_{oldest} - \mu_{youngest} = 0$

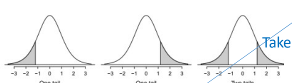
$Ha: \mu_{oldest} - \mu_{youngest} > 0$

T = -1.8

Degrees of Freedom = min(n_{oldest}-1, $n_{youngest}$-1)

= min(10-1,10-1) = 9

t distribution probability table



Take absolute value.

one tail	0.100	*0.050	0.025	0.010	0.005
two tails	0.200	0.100	0.050	0.020	0.010
df 1	3.08	6.31	12.71	31.82	63.66
9	1.38	1.83	2.26	2.82	3.25

Clicker question

A group of researchers want to see if the average SAT score of first-born children is higher than that of last-born children. 10 first-born children were sampled and 10 last-born children were sampled. The test statistic for this test is *T = -1.8.* What is the p-value?

(a) 0.01 < p-value < 0.025
(b) 0.025 < p-value < 0.05
(c) 0.05 < p-value < 0.10
(d) 0.10 < p-value < 0.20

Population Parameter of Interest:

- **$\mu_{oldest} - \mu_{youngest}$ (Independent Means Test)**

$Ho: \mu_{oldest} - \mu_{youngest} = 0$

$Ha: \mu_{oldest} - \mu_{youngest} > 0$

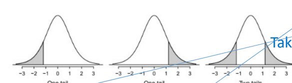
One-tailed test = one tail in t-table

T = -1.8

Degrees of Freedom = min(n_{oldest}-1, $n_{youngest}$-1)

= min(10-1,10-1) = 9

t distribution probability table



Take absolute value.

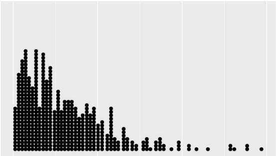
one tail	0.100	*0.050	0.025	0.010	0.005
two tails	0.200	0.100	0.050	0.020	0.010
df 1	3.08	6.31	12.71	31.82	63.66
9	1.38	1.83	2.26	2.82	3.25

What should you try if CLT conditions not met for hypothesis test for *p*?

What should you try if CLT conditions not met for hypothesis test for *p*?

Randomization testing!

When designing a simulation that creates a randomization distribution, what is the main property the randomization distribution should have?

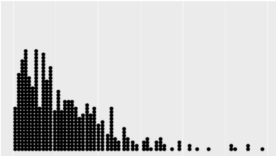


When designing a simulation that creates a randomization distribution, what is the main property the randomization distribution should have?

Should be roughly centered at the null value.

Assumes Ho.

We would expect the null value.



Clicker question

A researcher has set up the following randomization test for *p*. What were her hypotheses, observed sample proportion, and sample size?

Randomization Test for p
1. Place 100 chips in a bag, 20 red and 80 green.
2. Draw 10 chips from the bag with replacement.
3. Note the proportion of the 10 chips that were red and plot it in the randomization distribution.
4. Repeat (2) and (3) many times.
5. In the randomization distribution find the proportion of simulations where the sample proportion was 17% or less.

a) Ho: p = 0.17; Ha: p < 0.17 $\hat{p} = 0.20$ n=10
b) Ho: p = 0.20; Ha: p ≠ 0.20 $\hat{p} = 0.17$ n=100
c) Ho: p = 0.17; Ha: p < 0.17 $\hat{p} = 0.20$ n=20
d) Ho: p = 0.20; Ha: p < 0.20 $\hat{p} = 0.17$ n=10
e) Ho: p = 0.20; Ha: p < 0.20 $\hat{p} = 0.20$ n=100

Clicker question

A researcher has set up the following randomization test for *p*. What were her hypotheses, observed sample proportion, and sample size?

Randomization Test for p
1. Place 100 chips in a bag, 20 red and 80 green.
2. Draw 10 chips from the bag with replacement.
3. Note the proportion of the 10 chips that were red and plot it in the randomization distribution.
4. Repeat (2) and (3) many times.
5. In the randomization distribution find the proportion of simulations where the sample proportion was 17% or less.

a) Ho: p = 0.17; Ha: p < 0.17 $\hat{p} = 0.20$ n=10
b) Ho: p = 0.20; Ha: p ≠ 0.20 $\hat{p} = 0.17$ n=100
c) Ho: p = 0.17; Ha: p < 0.17 $\hat{p} = 0.20$ n=20
d) Ho: p = 0.20; Ha: p < 0.20 $\boldsymbol{\hat{p} = 0.17}$ ***n=10***
e) Ho: p = 0.20; Ha: p < 0.20 $\hat{p} = 0.20$ n=100

Clicker question

A researcher has set up the following randomization test for *p*. What were her hypotheses, observed sample proportion, and sample size?

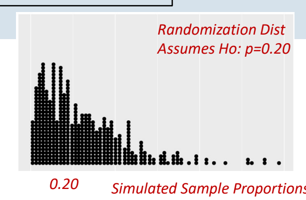
Randomization Test for p
1. Place 100 chips in a bag, 20 red and 80 green.
2. Draw 10 chips from the bag with replacement.
3. Note the proportion of the 10 chips that were red and plot it in the randomization distribution.
4. Repeat (2) and (3) many times.
5. In the randomization distribution find the proportion of simulations where the sample proportion was 17% or less.

Ho: p = 0.20; *Ha: p* < *0.20* $\hat{p} = 0.17$ *n=10*

Clicker question

A researcher has set up the following randomization test for *p*. What were her hypotheses, observed sample proportion, and sample size?

Randomization Test for p
1. Place 100 chips in a bag, 20 red and 80 green.
2. Draw 10 chips from the bag with replacement.
3. Note the proportion of the 10 chips that were red and plot it in the randomization distribution.
4. Repeat (2) and (3) many times.
5. In the randomization distribution find the proportion of simulations where the sample proportion was 17% or less.



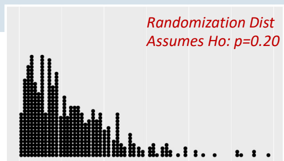
Ho: p = 0.20; *Ha: p* < *0.20* $\hat{p} = 0.17$ *n=10*

Clicker question

A researcher has set up the following randomization test for *p*. What were her hypotheses, observed sample proportion, and sample size?

Randomization Test for p

1. Place 100 chips in a bag, 20 red and 80 green.
2. Draw 10 chips from the bag with replacement.
3. Note the proportion of the 10 chips that were red and plot it in the randomization distribution.
4. Repeat (2) and (3) many times.
5. In the randomization distribution find the proportion of simulations where the sample proportion was 17% or less.



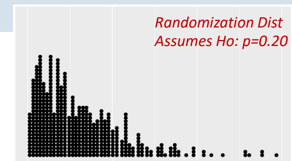
Ho: p = 0.20; Ha: p < 0.20 $\hat{p} = 0.17$ *n=10*

Clicker question

A researcher has set up the following randomization test for *p*. What were her hypotheses, observed sample proportion, and sample size?

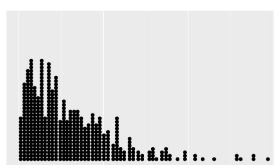
Randomization Test for p

1. Place 100 chips in a bag, 20 red and 80 green.
2. Draw 10 chips from the bag with replacement.
3. Note the proportion of the 10 chips that were red and plot it in the randomization distribution.
4. Repeat (2) and (3) many times.
5. In the randomization distribution find the proportion of simulations where the sample proportion was 17% or less.



Ho: p = 0.20; Ha: p < 0.20 $\hat{p} = 0.17$ *n=10*

What can we say and do with the sampling distribution of $\hat{p}$ when CLT conditions are/aren't met?



Clicker question

We would like to test if the proportion of Duke students that are only-children is larger than 0.10. We collected a random sample of 60 Duke students and found that 0.25 are only-children. Which of the following is true?

I. If we assume Ho, the sampling distribution for $\hat{p}$ is unimodal and symmetric.
II. If we assume Ho, the sampling distribution for $\hat{p}$ is left-skewed.
III. If we assume Ho, the sampling distribution for $\hat{p}$ is right skewed.
IV. We can conduct this hypothesis test with Central Limit Theorem methods.
V. We can create a confidence interval with Central Limit Theorem methods.

a) I, IV, and V
b) II
c) III
d) III and V
e) I and V

Clicker question

We would like to test if the proportion of Duke students that are only-children is larger than 0.10. We collected a random sample of 60 Duke students and found that 0.25 are only-children. Which of the following is true?

I. If we assume Ho, the sampling distribution for $\hat{p}$ is unimodal and symmetric.
II. If we assume Ho, the sampling distribution for $\hat{p}$ is left-skewed.
III. If we assume Ho, the sampling distribution for $\hat{p}$ is right skewed.
IV. We can conduct this hypothesis test with Central Limit Theorem methods.
V. We can create a confidence interval with Central Limit Theorem methods.

a) I, IV, and V
b) II
c) III
d) III and V
e) I and V

Hypothesis Test
$Ho: p = 0.1$
$Ha: p > 0.1$

Problem: In Confidence intervals and Hypothesis Testing, we don't know what ***p*** is but we need to plug it in in two parts of the analyses.

What part of the analyses do we need to plug in *p*?	Confidence Interval for *p*	Hypothesis Test for *p*
Checking SF Conditions: $np \geq 10$ $n(1-p) \geq 10$	Solution: Plug in the **observed** $\hat{p}$ for *p*.	Solution: Plug in the **null value** **p0** for *p*.
Calculating Standard Error: $SE = \sqrt{\frac{p(1-p)}{n}}$	Solution: Plug in the **observed** $\hat{p}$ for *p*.	Solution: Plug in the **null value** **p0** for *p*.

Clicker question

We would like to test if the proportion of Duke students that are only-children is larger than 0.10. We collected a random sample of 60 Duke students and found that 0.25 are only-children. Which of the following is true?

I. If we assume Ho, the sampling distribution for $\hat{p}$ is unimodal and symmetric.
II. If we assume Ho, the sampling distribution for $\hat{p}$ is left-skewed.
III. If we assume Ho, the sampling distribution for $\hat{p}$ is right skewed.
IV. We can conduct this hypothesis test with Central Limit Theorem methods.
V. We can create a confidence interval with Central Limit Theorem methods.

a) I, IV, and V
b) II
c) III
d) III and V
e) I and V

Hypothesis Test
$Ho: p = 0.1$
$Ha: p > 0.1$

CLT Conditions for HT:
1. **Independence:**
   - Random sampling
   - N=60<10% of Duke Students
2. **SF-Conditions:**
   - ~~$np_0 = 60(0.10) < 10$~~
   - $n(1 - p_0) = 60(1 - 0.10) \geq 10$

Clicker question

We would like to test if the proportion of Duke students that are only-children is larger than 0.10. We collected a random sample of 60 Duke students and found that 0.25 are only-children. Which of the following is true?

~~I. If we assume Ho, the sampling distribution for $\hat{p}$ is unimodal and symmetric.~~
II. If we assume Ho, the sampling distribution for $\hat{p}$ is left-skewed.
III. If we assume Ho, the sampling distribution for $\hat{p}$ is right skewed.
~~IV. We can conduct this hypothesis test with Central Limit Theorem methods.~~
V. We can create a confidence interval with Central Limit Theorem methods.

a) I, IV, and V
b) II
c) III
d) III and V
e) I and V

Hypothesis Test
$Ho: p = 0.1$
$Ha: p > 0.1$

CLT Conditions for HT:
1. **Independence:**
   - Random sampling
   - N=60<10% of Duke Students
2. **SF-Conditions:**
   - ~~$np_0 = 60(0.10) < 10$~~
   - $n(1 - p_0) = 60(1 - 0.10) \geq 10$

SF conditions not met ***for a hypothesis test.***
- *…can't use CLT hypothesis test*
- *…sampling distribution (that assumes Ho) is not normal (symmetric/unimodal)*

Clicker question

We would like to test if the proportion of Duke students that are only-children is larger than 0.10. We collected a random sample of 60 Duke students and found that 0.25 are only-children. Which of the following is true?

I. ~~If we assume Ho, the sampling distribution for $\hat{p}$ is unimodal and symmetric.~~
II. If we assume Ho, the sampling distribution for $\hat{p}$ is left-skewed.
III. If we assume Ho, the sampling distribution for $\hat{p}$ is right skewed.
IV. ~~We can conduct this hypothesis test with Central Limit Theorem methods.~~
V. We can create a confidence interval with Central Limit Theorem methods.

a) I, IV, and V
b) II
c) III
d) III and V
e) I and V

Hypothesis Test
$Ho: p = 0.1$
$Ha: p > 0.1$

CLT Conditions for HT:
1. **Independence:**
   - Random sampling
   - N=60<10% of Duke Students
2. **SF-Conditions:**
   - ~~$np_o = 60(0.10) < 10$~~
   - $n(1-p_0) = 60(1-0.10) \geq 10$

SF conditions not met ***for a hypothesis test***.
- *…can't use CLT hypothesis test*
- *…sampling distribution (that assumes Ho) is not normal (symmetric/unimodal)*

Clicker question

We would like to test if the proportion of Duke students that are only-children is larger than 0.10. We collected a random sample of 60 Duke students and found that 0.25 are only-children. Which of the following is true?

I. ~~If we assume Ho, the sampling distribution for $\hat{p}$ is unimodal and symmetric.~~
II. ~~If we assume Ho, the sampling distribution for $\hat{p}$ is left-skewed.~~
III. If we assume Ho, the sampling distribution for $\hat{p}$ is right skewed.
IV. ~~We can conduct this hypothesis test with Central Limit Theorem methods.~~
V. We can create a confidence interval with Central Limit Theorem methods.

a) I, IV, and V
b) II
c) III
d) III and V
e) I and V

Hypothesis Test
$Ho: p = 0.1$
$Ha: p > 0.1$

CLT Conditions for HT:
1. **Independence:**
   - Random sampling
   - N=60<10% of Duke Students
2. **SF-Conditions:**
   - ~~$np_o = 60(0.10) < 10$~~
   - $n(1-p_0) = 60(1-0.10) \geq 10$

- Peak of the sampling distribution (that assumes Ho) is centered at p=0.1
- Not symmetric/unimodal
- Natural boundaries are [0,1]… peak close to 0.

Clicker question

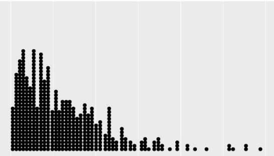
We would like to test if the proportion of Duke students that are only-children is larger than 0.10. We collected a random sample of 60 Duke students and found that 0.25 are only-children. Which of the following is true?

I. ~~If we assume Ho, the sampling distribution for $\hat{p}$ is unimodal and symmetric.~~
II. ~~If we assume Ho, the sampling distribution for $\hat{p}$ is left-skewed.~~
III. If we assume Ho, the sampling distribution for $\hat{p}$ is right skewed.
IV. ~~We can conduct this hypothesis test with Central Limit Theorem methods.~~
V. We can create a confidence interval with Central Limit Theorem methods.

a) I, IV, and V
b) II
c) III
d) III and V
e) I and V

CLT Conditions for CI:
1. **Independence:**
   - Random sampling
   - N=60<10% of Duke Students
2. **SF-Conditions:**
   - $n\hat{p} = 60(0.25) \geq 10$
   - $n(1-\hat{p}) = 60(1-0.25) \geq 10$

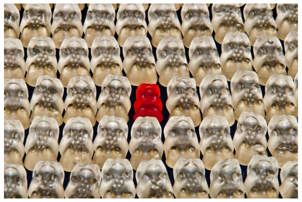
What are the two ways we can make an error in a HT?
What are the two ways we can make the right decision in a HT?



Probability of Errors in a HT
Probability of Correct Decisions in a HT

What are the names for these errors/correct decisions?

How are they affected by:
- **Sample size**
- **Effect Size**
- **Significance Level**



Clicker question

Suppose we were to decrease the sample size and increase the effect size *at the same time*. What happens to the following?

I. Probability of rejecting the null hypothesis given that the null hypothesis is actually true.
II. Probability of rejecting the null hypothesis given that the alternative hypothesis is actually true.

a) We don't know what happens to either I and II.
b) I and II stay both stay same.
c) I and II both decrease.
d) I stays the same, II decreases.
e) I stays the same, and we don't know what happens to II.

Clicker question

Suppose we were to decrease the sample size and increase the effect size at the same time. What happens to the following?

I. Probability of rejecting the null hypothesis given that the null hypothesis is actually true. α=P(Type 1 Error) = Significance Level
II. Probability of rejecting the null hypothesis given that the alternative hypothesis is actually true. *Power*

a) We don't know what happens to either I and II.
b) I and II stay both stay same.
c) I and II both decrease.
d) I stays the same, II decreases.
e) I stays the same, and we don't know what happens to II.

Clicker question

Suppose we were to decrease the sample size and increase the effect size at the same time. What happens to the following?

I. Probability of rejecting the null hypothesis given that the null hypothesis is actually true. α = P(Type 1 Error)=Significance Level (Stays the same… the researcher pre-selects α, not affected by sample size or effect size changes.)
II. Probability of rejecting the null hypothesis given that the alternative hypothesis is actually true. *Power*

a) We don't know what happens to either I and II.
b) I and II stay both stay same.
c) I and II both decrease.
d) I stays the same, II decreases.
e) I stays the same, and we don't know what happens to II.

Clicker question

Suppose we were to decrease the sample size and increase the effect size at the same time. What happens to the following?

I. Probability of rejecting the null hypothesis given that the null hypothesis is actually true. α = P(Type 1 Error)= Significance Level (Stays the same… the researcher pre-selects α, not affected by sample size or effect size changes.)

II. Probability of rejecting the null hypothesis given that the alternative hypothesis is actually true. *Power (Power decreases when sample size decreases. Power increases when sample size increases…. so they have counterbalancing effects…. Not sure*

a) We don't know what happens to either I and II.
b) I and II stay both stay same.
c) I and II both decrease.
d) I stays the same, II decreases.
e) I stays the same, and we don't know what happens to II.

Clicker question

Suppose we were to increase the sample size. What happens to the following?

I. Probability of failing to reject the null hypothesis given that the alternative hypothesis is actually true.

II. The complement of I.

a) Both increase.
b) I increases, II decreases.
c) I decreases, II increases.
d) We don't know what happens to either.

Clicker question

Suppose we were to increase the sample size. What happens to the following?

I. Probability of failing to reject the null hypothesis given that the alternative hypothesis is actually true. ***β = P(Type 2 Error Rate)***

II. The complement of I. ***1- β =Power***

a) Both increase.
b) I increases, II decreases.
c) I decreases, II increases.
d) We don't know what happens to either.

		Decision	
		fail to reject H_0	reject H_0
Truth	H_0 true	✓	*Type 1 Error*, α
	H_A true	*Type 2 Error*, β	*Power*, $1-\beta$

- A *Type 1 Error* is rejecting the null hypothesis when H_0 is true.
  - P(Type 1 Error) = α = P(reject Ho|Ho is true)
- A *Type 2 Error* is failing to reject the null hypothesis when H_A is true.
  - P(Type 2 Error) = β = P(fail to reject Ho|Ha is true)
- *Power* is the probability of *correctly* rejecting H_0, and hence the complement of the probability of a Type 2 Error
  - Power = 1 – β = P(reject Ho|Ha is true)

There are several ways to **increase power** (and hence **decrease P(Type 2 error)**, **β**:

1. **Increase the sample size**.
2. **Decrease the standard deviation of the sample**, which is equivalent to increasing the sample size (it will decrease the standard error). With a smaller s we have a better chance of distinguishing the null value from the observed point estimate. This is difficult to ensure but cautious measurement process and limiting the population so that it is more homogenous may help.
3. **Increase** α, which will make it more likely to reject H_0 (but note that this has the side effect of increasing the Type 1 error rate).
4. **Consider a larger effect size**. If the true mean of the population is in the alternative hypothesis but close to the null value, it will be harder to detect a difference.

Completing a Chi-Squared Test of Independence

Clicker question

Assuming that region and opinion on whether the country is headed on the right track are independent, what is the expected number of Southerners that think the country is on the wrong track?

a) 131
b) 131/500 = 0.26
c) (329·193)/500 = 127
d) 193/500 = 0.39
e) (131-127)^2 / 127

	Right Direction	Wrong Track	Total
Northeast	29	54	83
North Central	44	77	121
South	*62*	131	193
West	36	67	103
Total	171	329	500

Clicker question

Assuming that region and opinion on whether the country is headed on the right track are independent, what is the expected number of Southerners that think the country is on the wrong track?

a) 131
b) 131/500 = 0.26
c) (329·193)/500 = 127
d) 193/500 = 0.39
e) (131-127)^2 / 127

Expected Count for a Cell
*in **Chi-Squared Independence Test**=*
(row total)(column total)/(sample total)*

	Right Direction	Wrong Track	Total
Northeast	29	54	83
North Central	44	77	121
South	*62*	131	193
West	36	67	103
Total	171	329	500

What is the test statistic for:

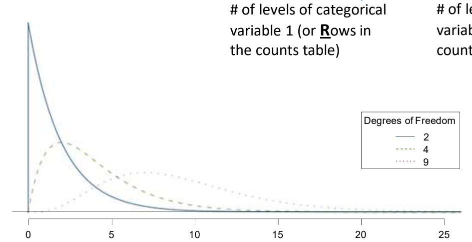
- **Chi-Squared Goodness of Fit Test?**
- **Chi-Squared Independence Test?**

It's the same calculation!

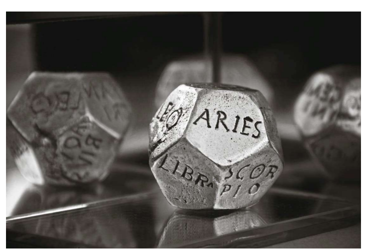
$$\chi^2 = \sum_{i=1}^{k} \frac{(O - E)^2}{E}$$ where k = total number of cells

The χ^2 distribution has just one parameter, *degrees of freedom (df)*, which influences the shape, center, and spread of the distribution.

- **For χ^2 GOF test**: $df = k - 1$ (# of levels of the one categorical variable)
- **For χ^2 independence test**: $df = (R - 1) \times (C - 1)$ (R: # of levels of categorical variable 1 (or Rows in the counts table); C: # of levels of categorical variable 2 (or Columns in the counts table))



Figuring out which test to use.



Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies. We are interested to know if you are equally likely to be a CEO for all zodiac signs. What test would we use?

a) ANOVA
b) Single Proportion Hypothesis Test
c) Two Proportion Hypothesis Test
d) Chi-Squared Goodness of Fit Test
e) Chi-Squared Independence Test

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies. We are interested to know if you are equally likely to be a CEO for all zodiac signs. What test would we use?

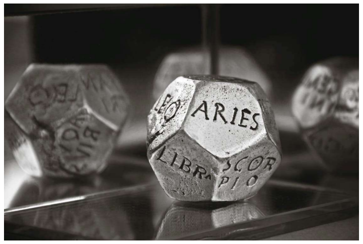
a) ANOVA
b) Single Proportion Hypothesis Test
c) Two Proportion Hypothesis Test
d) Chi-Squared Goodness of Fit Test
e) Chi-Squared Independence Test

	What Zodiac Sign?
CEO 1	Saggitarius
CEO 2	Picses
CEO 3	Taurus
CEO 4	Taurus
	…

- **One Categorical Variable** with **>2** (12 **levels)**

Ho: The CEO zodiac sign data follows the distribution where all zodiac signs are equally likely.
Ha: The CEO zodiac sign data doesn't follow the distribution where all zodiac signs are equally likely.

Figuring out which test to use.



Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as the zodiac signs for all non-CEOs. We are interested to know if there's a difference in the proportions of Sagittarius' and Aquarius' that are CEOs.

a) ANOVA
b) Single Proportion Hypothesis Test
c) Two Proportion Hypothesis Test
d) Chi-Squared Goodness of Fit Test
e) Chi-Squared Independence Test

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as the zodiac signs for all non-CEOs. We are interested to know if there's a difference in the proportions of Sagittarius' and Aquarius' that are CEOs.

a) ANOVA
b) Single Proportion Hypothesis Test
c) Two Proportion Hypothesis Test
d) Chi-Squared Goodness of Fit Test
e) Chi-Squared Independence Test

- Categorical Variable (2 levels): CEO/Not CEO
- Categorical Variable (2 levels): Sagittarius/Aquarius (ignore the other signs)

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as the zodiac signs for all non-CEOs. We are interested to know if there's a difference in the proportions of Sagittarius' and Aquarius' that are CEOs.(Assume random sampling used). We want to make a confidence interval and hypothesis test. Which of the following methods are most appropriate?

	CEO	Not-CEO	Total
Sagittarius	5	45	50
Aquarius	25	40	50
Total	30	85	100

a) CI – CLT methods, HT – Randomization Testing
b) CI – Randomization Testing, HT – CLT Methods
c) CI – CLT methods, HT – CLT Methods
d) CI – Bootstrap methods, HT – CLT Methods
e) CI – Bootstrap methods, HT – Randomization Testing

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as the zodiac signs for all non-CEOs. We are interested to know if there's a difference in the proportions of Sagittarius' and Aquarius' that are CEOs.(Assume random sampling used). We want to make a confidence interval and hypothesis test. Which of the following methods are most appropriate?

	CEO	Not-CEO	Total
Sagittarius	5	45	50
Aquarius	25	40	50
Total	30	85	100

a) CI – CLT methods, HT – Randomization Testing
b) CI – Randomization Testing, HT – CLT Methods
c) CI – CLT methods, HT – CLT Methods
d) CI – Bootstrap methods, HT – CLT Methods
e) CI – Bootstrap methods, HT – Randomization Testing

What part of the analyses do we need to plug in *p1 and p2*?	Confidence Interval for *p1-p2*	Hypothesis Test with Ho:*p1-p2=(non-zero #)*	Hypothesis Test with Ho:*p1-p2=0*
Checking SF Conditions: $n_1 p_1 \geq 10$, $n_1(1-p_1) \geq 10$, $n_2 p_2 \geq 10$, $n_2(1-p_2) \geq 10$	Solution: Plug in the **observed** $\hat{p}_1$ for *p1*. Plug in the **observed** $\hat{p}_2$ for *p2*.	Solution: Plug in the **observed** $\hat{p}_1$ for *p1*. Plug in the **observed** $\hat{p}_2$ for *p2*.	Solution: Plug in the **pooled proportion** for *p1 and p2*.
Calculating Standard Error: $SE = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$	Solution: Plug in the **observed** $\hat{p}_1$ for *p1*. Plug in the **observed** $\hat{p}_2$ for *p2*.	Solution: Plug in the **observed** $\hat{p}_1$ for *p1*. Plug in the **observed** $\hat{p}_2$ for *p2*.	Solution: Plug in the **pooled proportion** for *p1 and p2*.

Expected proportion of success for both groups when $H_0: p_1 = p_2$ is defined as the *pooled proportion*:

$$\hat{p}_{pool} = \frac{total\ successes}{total\ sample\ size} = \frac{suc_1 + suc_2}{n_1 + n_2}$$

3

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as the zodiac signs for all non-CEOs. We are interested to know if there's a difference in the proportions of Sagittarius' and Aquarius' that are CEOs.(Assume random sampling used). We want to make a confidence interval and hypothesis test. Which of the following methods are most appropriate?

	CEO	Not-CEO	Total
Sagittarius	5	45	50
Aquarius	25	40	50
Total	30	85	100

a) CI – CLT methods, HT – Randomization Testing
b) CI – Randomization Testing, HT – CLT Methods
c) CI – CLT methods, HT – CLT Methods
d) CI – Bootstrap methods, HT – CLT Methods
e) CI – Bootstrap methods, HT – Randomization Testing

$$\hat{p}_{pool} = \frac{\#\ of\ success\ (CEOS)}{total\ sample} = \frac{5+25}{50+50} = 0.3$$

CLT Conditions for HT:

1. **Independence:**
   - RS of Sag, RS of Aquarius
   - N_{sag}=50<10% of all Sags, $N_{aquarius}$=50<10% of all Aquarius',
2. **SF-Conditions:**
   - $n_{sag}\hat{p}_{pool} = 50(0.3) \geq 10$
   - $n_{sag}(1-\hat{p}_{pool}) = 50(1-0.30) \geq 10$
   - $n_{aqua}\hat{p}_{pool} = 50(0.3) \geq 10$
   - $n_{aqua}(1-\hat{p}_{pool}) = 50(1-0.30) \geq 10$

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as the zodiac signs for all non-CEOs. We are interested to know if there's a difference in the proportions of Sagittarius' and Aquarius' that are CEOs.(Assume random sampling used). We want to make a confidence interval and hypothesis test. Which of the following methods are most appropriate?

	CEO	Not-CEO	Total
Sagittarius	5	45	50
Aquarius	25	40	50
Total	30	85	100

a) CI – CLT methods, HT – Randomization Testing
b) CI – Randomization Testing, HT – CLT Methods
c) CI – CLT methods, HT – CLT Methods
d) CI – Bootstrap methods, HT – CLT Methods
e) CI – Bootstrap methods, HT – Randomization Testing

$\hat{p}_{sag} = \frac{5}{50} = 0.1$,
$\hat{p}_{aqua} = \frac{25}{50} = 0.5$,

CLT Conditions for CI… not met!

1. **Independence:**
   - RS of Sag, RS of Aquarius
   - N_{sag}=50<10% of all Sags, $N_{aquarius}$=50<10% of all Aquarius',
2. **SF-Conditions:**
   - ~~$n_{sag}\hat{p}_{sag} = 50(0.1) < 10$~~
   - $n_{sag}(1 - \hat{p}_{sag}) = 50(1 - 0.10) \geq 10$
   - $n_{aqua}\hat{p}_{sag} = 50(0.5) \geq 10$
   - $n_{aqua}(1 - \hat{p}_{sag}) = 50(1 - 0.50) \geq 10$

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as their annual salary. We want to know if there's an association between a CEO's zodiac sign and their salary.

a) ANOVA
b) Single Proportion Hypothesis Test
c) Two Proportion Hypothesis Test
d) Chi-Squared Goodness of Fit Test
e) Chi-Squared Independence Test

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as their annual salary. We want to know if there's an <u>association</u> between a CEO's zodiac sign and their annual salary.

a) ANOVA
b) Single Proportion Hypothesis Test
c) Two Proportion Hypothesis Test
d) Chi-Squared Goodness of Fit Test
e) Chi-Squared Independence Test

- **Categorical Explanatory Variable** with **>2** (**12 levels**)
- **One Numerical Response Variable**

Ho: $\mu_{saggittarius} = \mu_{libra} = \cdots = \mu_{capricorn}$
Ha: ________________________________

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as their annual salary. We want to know if there's an <u>association</u> between a CEO's zodiac sign and their annual salary.

a) ANOVA
b) Single Proportion Hypothesis Test
c) Two Proportion Hypothesis Test
d) Chi-Squared Goodness of Fit Test
e) Chi-Squared Independence Test

- **Categorical Explanatory Variable** with **>2** (**12 levels**)
- **One Numerical Response Variable**

Ho: $\mu_{saggittarius} = \mu_{libra} = \cdots = \mu_{capricorn}$
Ha: At least one pair of zodiac signs have mean annual salaries that are different from one another.

Step 1a: Fill out the ANOVA table.

Test Statistic p-value

	Df	Sum Sq	Mean Sq	F Value	Pr(>F)
Between groups	k-1	$SSG = \sum_{j=1}^{k} n_j(\bar{y}_j - \bar{y})^2$	MSG = SSG/(k-1)	MSG/MSE	P(F > MSG/MSE)
Within Groups	n-k	$SSE = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(y_{ij} - \bar{y}_j)^2$	MSE = SSE/(n-k)		
Total	n-1	$SST = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(y_{ij} - \bar{y})^2$			

k: # of groups; n: # of obs.

Step 1b: Evaluate the p-value to make a conclusion.

- p-value<α: Reject Ho. There is sufficient evidence to suggest Ha.
- p-value≥α: Fail to reject Ho. There is not sufficient evidence to suggest Ha.

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as their annual salary. We conducted an ANOVA and found there *was* sufficient evidence to suggest at least one of the pairs of zodiac signs had mean annual salaries that are different from one another. How many post-hoc tests would we need run to determine *which* of these pairs were different?

a) 66
b) 12
c) 132
d) 2

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as their annual salary. We conducted an ANOVA and found there *was* sufficient evidence to suggest at least one of the pairs of zodiac signs had mean annual salaries that are different from one another. How many post-hoc tests would we need run to determine *which* of these pairs were different?

a) 66 *= total number of pairs you can make out of k (12) levels = k(k-1)/2 = 12(11)/2*
b) 12
c) 132
d) 2

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as their annual salary. We conducted an ANOVA and found there *was* sufficient evidence to suggest at least one of the pairs of zodiac signs had mean annual salaries that are different from one another (at α=0.1). The test statistic and p-value for the post-hoc multiple comparison test that compares the salaries of Sagittarius and Leo CEOS is below. How many errors did we make in these calculations and conclusion?

a) 0
b) 1
c) 2
d) 3
e) 4 or more

$$T_{14} = \frac{(300 - 350) - 0}{\sqrt{\frac{100^2}{30} + \frac{50^2}{15}}} = 2.24$$

$$0.01 < p - value < 0.025$$

Because p-value < 0.1, reject Ho.

Zodiac	Sample Mean	Sample Std Dev	Sample Size
Sagittarius	$300K	$100K	30
Leo	$350K	$50K	15
...	...	...	...

one tail	0.100	0.050	0.025	0.010	0.005
two tails	0.200	0.100	0.050	0.020	0.010
df 1	3.08	6.31	12.71	31.82	63.66
14	1.35	1.76	2.14	2.62	2.98

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as their annual salary. We conducted an ANOVA and found there *was* sufficient evidence to suggest at least one of the pairs of zodiac signs had mean annual salaries that are different from one another (at α=0.1). The test statistic and p-value for the post-hoc multiple comparison test that compares the salaries of Sagittarius and Leo CEOS is below. **How many errors did we make in these calculations and conclusion?**

a) 0
b) 1
c) 2
d) 3
e) 4 or more

$$T_{14} = \frac{(300-350)-0}{\sqrt{\frac{100^2}{30}+\frac{50^2}{15}}} = 2.24$$

$0.01 < p-value < 0.025$

Because p-value < 0.1, reject Ho.

Zodiac	Sample Mean	Sample Std Dev	Sample Size
Sagittarius	$300K	$100K	30
Leo	$350K	$50K	15
…	…	…	…

one tail	0.100	0.050	0.025	0.010	0.005
two tails	0.200	0.100	0.050	0.020	0.010
df 1	3.08	6.31	12.71	31.82	63.66
14	1.35	1.76	2.14	2.62	2.98

Things that are different for conducting one of the post-hoc multiple comparison tests (after ANOVA).

Ho: $\mu_a - \mu_b = 0$
Ha: $\mu_a - \mu_b \neq 0$

Things that are Different	Regular Independent Means Test (with null value =0)	Post-hoc (ANOVA Step 2) Multiple Pairwise Comparisons Tests (do this for each one)
Significance Level	α	$\alpha^* = \frac{\alpha}{k(k-1)/2}$
T-Statistics	$T = \frac{(\bar{x}_a - \bar{x}_b) - 0}{\sqrt{\frac{s_a^2}{n_a}+\frac{s_b^2}{n_b}}}$	$T = \frac{(\bar{x}_a - \bar{x}_b) - 0}{\sqrt{\frac{MSE}{n_a}+\frac{MSE}{n_b}}}$
Degrees of Freedom to use in T-distribution	$df = \min(n_a - 1, n_b - 1)$	$df = df_E = n - k$ • n=total number of observations • k=# of groups

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as their annual salary. We conducted an ANOVA and found there *was* sufficient evidence to suggest at least one of the pairs of zodiac signs had mean annual salaries that are different from one another (at α=0.1). The test statistic and p-value for the post-hoc multiple comparison test that compares the salaries of Sagittarius and Leo CEOS is below. **How many errors did we make in these calculations and conclusion?**

a) 0
b) 1
c) 2
d) 3
e) 4 or more

$$T_{14} = \frac{(300-350)-0}{\sqrt{\frac{MSE}{30}+\frac{MSE}{15}}} = 2.24$$

$0.01 < p-value < 0.025$

Came from the one-tailed in the t-table. Multiple comparison tests are two-tailed tests.

Ho: $\mu_{saggittarius} - \mu_{leo} = 0$
Ha: $\mu_{saggittarius} - \mu_{leo} \neq 0$

Because p-value < 0.1, reject Ho.

Zodiac	Sample Mean	Sample Std Dev	Sample Size
Sagittarius	$300K	$100K	30
Leo	$350K	$50K	15
…	…	…	…

one tail	0.100	0.050	0.025	0.010	0.005
two tails	0.200	0.100	0.050	0.020	0.010
df 1	3.08	6.31	12.71	31.82	63.66
14	1.35	1.76	2.14	2.62	2.98

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as their annual salary. We conducted an ANOVA and found there *was* sufficient evidence to suggest at least one of the pairs of zodiac signs had mean annual salaries that are different from one another (at α=0.1). The test statistic and p-value for the post-hoc multiple comparison test that compares the salaries of Sagittarius and Leo CEOS is below. **How many errors did we make in these calculations and conclusion?**

a) 0
b) 1
c) 2
d) 3
e) 4 or more

$$T_{n-k} = \frac{(300-350)-0}{\sqrt{\frac{MSE}{30}+\frac{MSE}{15}}} = 1.97$$

p-value=0.05

Because p-value ≥ 0.1/[k(k-1)/2], reject Ho.

Zodiac	Sample Mean	Sample Std Dev	Sample Size
Sagittarius	$300K	$100K	30
Leo	$350K	$50K	15
…	…	…	…

Ho: $\mu_{saggittarius} - \mu_{leo} = 0$
Ha: $\mu_{saggittarius} - \mu_{leo} \neq 0$

Use 2-tails in the t-table.

Clicker question

Fortune magazine collected the zodiac signs of 256 CEOs of the largest 400 companies as well as their annual salary. We conducted an ANOVA and found there *was* sufficient evidence to suggest at least one of the pairs of zodiac signs had mean annual salaries that are different from one another (at α=0.1). The test statistic and p-value for the post-hoc multiple comparison test that compares the salaries of Sagittarius and Leo CEOS is below. How many errors did we make in these calculations and conclusion?

$$T_{256-12} = \frac{(300-350)-0}{\sqrt{\frac{MSE}{30}+\frac{MSE}{15}}} = 1.97$$

p-value=0.05

Zodiac	Sample Mean	Sample Std Dev	Sample Size
Sagittarius	\$300K	\$100K	30
Leo	\$350K	\$50K	15
…	…	…	…

Ho: $\mu_{saggittarius} - \mu_{leo} = 0$

Ha: $\mu_{saggittarius} - \mu_{leo} \neq 0$

Use 2-tails in the t-table.

a) 0
b) 1
c) 2
d) 3
e) 4 or more

Because p-value ≥ 0.1/[12(12-1)/2], reject Ho.

Why/when would we want to make a bootstrap confidence interval?

Clicker question

The general structure for a CLT confidence interval for μ and a bootstrap confidence interval (SE method) for μ has three components (see below). How many of these components are the same for both a CLT confidence interval for μ (assume we don't know σ) and a bootstrap confidence interval for μ (SE method) ?

( ) ± ( )( )
1 2 3

a) 0
b) 1
c) 2
d) 3

Clicker question

The general structure for a CLT confidence interval for μ and a bootstrap confidence interval (SE method) for μ has three components (see below). How many of these components are the same for both a CLT confidence interval for μ (assume we don't know σ) and a bootstrap confidence interval for μ (SE method) ?

( ) ± ( )( )
1 2 3

a) 0
b) 1
c) 2
d) 3

CLT confidence interval for μ (don't know σ)

$$\bar{x} \pm t_{n-1}\frac{s}{\sqrt{n}}$$

Bootstrap confidence interval for μ (SE Method)

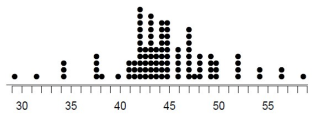
$$\bar{x} \pm t_{n-1}\sigma_{boot}$$

Bootstrap Confidence Intervals

Step 1: Create the bootstrap distribution

	Creating a sampling distribution (approximation)	Creating a bootstrap distribution
Do the following many many times..		
Step 1:	with replacement, take a random sample of size n from the…	
	population.	sample (of size n). aka:bootstrap sample
Step 2:	Calculate the sample statistic (ie: sample mean/median/proportion) of this random sample from (step 1). aka:bootstrap statistic	
Step 3:	Plot this sample statistic in a histogram/dotplot. This histogram/dotplot is your…	
	sampling distribution.	bootstrap distribution.

For comparison



6

Bootstrap Confidence Intervals

*Step 2: Use the **bootstrap distribution** to create a confidence interval.*

- **Percentile Method**
  - XX% bootstrap confidence interval = the cutoff values for the middle XX% of the bootstrap distribution
- **Standard Error Method**
  - XX% bootstrap confidence interval = $point\ estimate \pm t^{*}SE_{boot}$

Standard deviation of all the bootstrap statistics you calculated

When the point estimate is: median, μ, μ_{diff}, p
- ***degrees of freedom** = n-1 (n = size of the original sample = each bootstrap sample)*
- ***area under t-distribution curve** =(1-XX/100)% TWO TAILS.*

6

Interpreting confidence intervals

Clicker question

A study examining the relationship between weights of school children and absences found a 95% confidence interval for the difference between the average number of days missed by overweight and non-overweight children ($\mu_{overweight} - \mu_{non-overweight}$) to be 1.3 days to 2.8 days. According to this interval, we are 95% confident that overweight children on average miss

1. 1.3 days fewer to 2.8 days more
2. 1.3 to 2.8 days more
3. 1.3 to 2.8 days fewer
4. 1.3 days more to 2.8 days fewer

than non-overweight children.

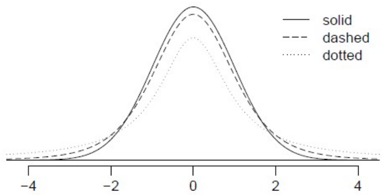
Clicker question

A study examining the relationship between weights of school children and absences found a 95% confidence interval for the difference between the average number of days missed by overweight and non-overweight children ($\mu_{overweight} - \mu_{non-overweight}$) to be 1.3 days to 2.8 days. According to this interval, we are 95% confident that overweight children on average miss

1. 1.3 days fewer to 2.8 days more
2. 1.3 to 2.8 days more
3. 1.3 to 2.8 days fewer
4. 1.3 days more to 2.8 days fewer

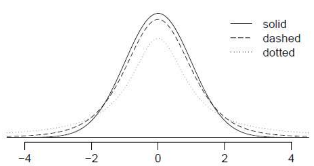
than non-overweight children.

Properties of Normal, T, Chi-Squared, and F distribution



Clicker question

The figure below shows three unimodal and symmetric curves. Which of the following is most plausible?



(a) solid: normal, dashed: $t_{df=5}$, dotted: $t_{df=1}$

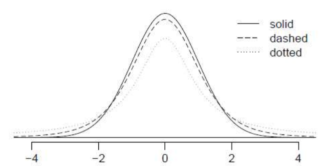
(b) solid: $t_{df=1}$, dashed: $t_{df=5}$, dotted: normal

(c) solid: $t_{df=5}$, dashed: $t_{df=1}$, dotted: normal

(d) solid: normal, dashed: $t_{df=1}$, dotted: $t_{df=5}$

Clicker question

The figure below shows three unimodal and symmetric curves. Which of the following is most plausible?



(a) solid: normal, dashed: $t_{df=5}$, dotted: $t_{df=1}$

(b) solid: $t_{df=1}$, dashed: $t_{df=5}$, dotted: normal

(c) solid: $t_{df=5}$, dashed: $t_{df=1}$, dotted: normal

(d) solid: normal, dashed: $t_{df=1}$, dotted: $t_{df=5}$

- **Calculating the necessary sample size for a CI of *p* with a given margin of error:**
  - Option 1: If there is a previous study, use $\hat{p}$ from that study
  - Option 2: If not, use $\hat{p} = 0.5$:
    - if you don't know any better, 50-50 is a good guess
    - $\hat{p} = 0.5$ gives the most conservative estimate of the standard error, which gives the highest possible sample size