AP statistics Chapter 9 Note-Taking Guide
Testing a Claim

Name ________________________
Per ___ Date __________________

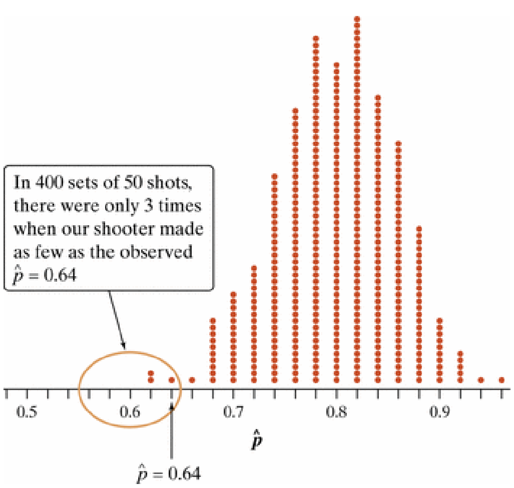
9.1A

A basketball player claims to make 80% of the free throws he attempts. You have seen him play basketball before and you don't believe him. So you ask him to shoot some free throws. Suppose he shoots 50 free throws and makes 32 of them. Is his claim true but he just had a "bad streak" or is his claim false?

A) What is his sample proportion (p)?

B) What can we conclude about the player's claim based on the sample data?

When you are an 80% free throw shooter you don't always make 80% of your free throws. Instead we average all performances of shooting free throws to get the 80%.
So how do we determine the likelihood (probability) that he had a "bad streak"?
We can perform a simulation to find out. We used Fathom software to simulate 400 sets of 50 shots assuming that the player **really** is an 80% free throw shooter. The dotplot of the results is below. Each dot is the proportion of shots made for each group of 50 attempts.



C) How many of the 400 sets of shots were at our sample proportion (_________) or lower?

You can say **how strong the evidence against the player's claim** is by giving the ____________________ that he would make as few as 32 out of 50 free throws if he really makes 80% in the long run.

D) What is the probability that he made only 32 shots out of 50 tries JUST BY CHANCE ("bad streak")?
Based on the simulation, our estimate of this probability is

________________. The observed statistic, $p = 0.64$, is ____________ that it gives convincing evidence that the player's claim is ________________

Be sure that you understand why this evidence is convincing. There are two possible explanations of the fact that our virtual player made only $p = 0.64$ of his free throws:

1. The player's claim is correct (p = 0.8), and by bad luck, a very unlikely outcome occurred.

2. The population proportion is actually less than 0.8, so the sample result is not an unlikely outcome.

9.1A

Significance test
A formal procedure for comparing _______________________ with a ___________________ whose truth we want to assess. We express the results of a significance test in terms of a probability that measures how well the data and the claim agree.
Significance tests

- Deal with claims about a _______________________
- Ask if sample data give good evidence ___________________ a claim
- "If we took many random samples and the claim were true, we would get a result like this ___% of the time"
- BASIC IDEA: An outcome that would rarely happen if a claim were true is good evidence that the claim is ________________________________!

9.1A	**Stating a hypotheses**
	• **Null hypothesis** (H_o) a statement of ___________________(What the person claims to be true) Will always be in the form H_0: p = ___ or H_0: μ = ___
	• **Alternative hypothesis** (H_a) the claim that we hope or _____________________ instead of the null hypothesis.
	  • **One sided** alternative hypothesis will claim that the actual parameter is ___________________ ___________________ than the null hypothesis. Will always be in the form H_a: p > ___ or H_a: μ > ___ or H_a: p < ___ or H_a: μ < ___
	  • **Two sided** alternative hypothesis will claim that the actual parameter is _______ the parameter stated in the null hypothesis. Will always be in the form H_a: p ≠ ___ or H_a: μ ≠ ___
	****Note the number in the null and alternative hypothesis will always be the same!
	
	Sample: Mike is an avid golfer who would like to improve his play. A friend suggests getting new clubs and lets Mike try out his 7-iron. Based on years of experience, Mike has established that the mean distance that balls travel when hit with his old 7-iron is μ = 175 yards with a standard deviation of σ = 15 yards. He is hoping that this new club will make his shots with a 7-iron more consistent (less spread), so he goes to the driving range and hits 50 shots with the new 7-iron.
	Problem:
	(a) State the appropriate null hypotheses
	
	(b) State the appropriate alternative hypotheses

9.1A	***Sample:*** VitaBlend, a new brand of vitamins, claims in a radio spot that people who take VitaBlend vitamins wake refreshed and ready for the new day within five minutes of their alarm sounding. VitaBlend's internal research department conducts a test to support the claim heard on the radio. a.) Null hypothesis: H_0: b.) Alternative hypothesis: H_a:
9.1A	***Sample:*** The school principal is not pleased with the first semester exam scores of the two first-period US history classes. The mean test score of both classes was 72. He believes that the student's final exam score average will increase because the two teachers will teach the second semester by combining the classes and using a "team teaching" approach. a.) Null hypothesis: H_0: b.) Alternative hypothesis: H_a:
9.1A	***Sample:*** Heading into the election, Jack needs 50% of the vote (or higher) to win the election. Jack is feeling pretty confident that he will win the election. Is his confidence warranted? a.) Null hypothesis: H_0: b.) Alternative hypothesis: H_a:

9.1A **<u>Notes about hypotheses:</u>**

- The alternative hypothesis should express the hopes or suspicions we have ___________ we see the data. **<u>It is cheating to look at the data first and then frame the hypotheses to fit what the data show.</u>**
- Hypotheses always refer to a ____________________, not to a __________________.
- If unsure, use a _____________________ alternative hypothesis!

9.1A **<u>Interpreting P-values</u>**

The probability that measures the ____________________________ against a null hypothesis is called a ___________________.

- The _____________________ the p-value, the stronger the evidence **against** the null hypothesis provided by the data.
- The p-value is the conditional probability ___________________

9.1A **<u>Significance Level</u>**

- To determine what P-value is considered small (___________) or large (_________________) we compare it to a _____________________________ (α).
- Significance level requires evidence against Ho to be ________________________ that is would happen ___________________ than _____% of the time by chance when H_o is true.
- When our P-value is ________________________ our chosen α, we say that the result is **statistically significant**.
- Significant doesn't mean important, it means not likely to _______________________________.
- If you are going to draw a conclusion based on statistical significance, the significance level should be determined __________________ the data are produced.
- Significance levels should range from ___________________. Courts will accept 0.05 or lower.
- When in doubt or if not explicitly stated in a problem, use __________________

9.1A **<u>Statistical significance</u>**

The final step in performing a significance test is to draw a conclusion about the competing claims you were testing. We will make one of ______________________ based on the strength of the evidence:

- ___________________________
- ___________________________ – Does not guarantee that H_o is ___________ just that there is not enough evidence to reject.

*****<u>DO NOT ACCEPT H_o</u>**, you will lose credit on AP exam!

9.1A In the **free-throw shooter example**, the estimated p-value of 0.0075 is ____________________ against the null hypotheses H_o: $p = 0.80$. For that reason, we would _________________ in favor of the alternative hypotheses H_a: $p < 0.80$. It appears that the virtual player makes fewer than 80% of his free throws.

9.1A ***Sample:*** When Mike was testing a new 7-iron, the hypotheses were

$$H_0: \sigma = 15$$

$$H_a: \sigma < 15$$

where σ = the true standard deviation of the distances Mike hits golf balls using the new 7-iron. Based on 50 shots with the new 7-iron, the standard deviation was s_x = 10.9 yards.

Problem: A significance test using the sample data produced a *P*-value of 0.002.

(a) **Interpret the *P*-value** in this context.

(b) Our significance level is 0.01, what should our conclusion be?

9.1A

Sample: For his second semester project in AP Statistics, Zenon decided to investigate if students at his school prefer name-brand potato chips to generic potato chips. He randomly selected 50 students and had each student try both types of chips, in random order. Overall, 34 of the 50 students preferred the name-brand chips. Zenon performed a significance test using the hypotheses:

$H_0 : p = 0.5$

$H_a : p > 0.5$

where p = the true proportion of students at his school that prefer name-brand chips.
The **resulting *P*-value was 0.0055**.
Problem: What conclusion would you make at each of the following significance levels? State the conclusion in context.
(a) $\alpha = 0.01$

(b) $\alpha = 0.001$

9.2

Significance tests for population proportions

9.2A

Significance tests for population proportions
Conditions must be met:

1. **Random**: Data should come from a well-designed ________________ or ______________ ___________________. Otherwise we can't infer to the population or establish cause and effect.

2. **Normal:** sampling distribution of the statistic is _______________________
   - Normal condition for proportions: _________________________________

   P will be replaced with $\mathbf{p_o}$ which is the _________________________.

3. **Independent:** sampling with replacement for the population allows us to use standard deviation formulas, or if sampling without replacement, we meet the 10% condition for independence _____________.

9.2A

Calculations: Test statistic and P-value
A significance test uses sample data to measure the strength of evidence against H_o. Here are some principles that apply to most tests:

- The test compares a _________________ calculated from sample data with the value of the parameter stated by the __________________________.
- Values of the statistic far from the ____________________ in the direction specified by the alternative hypothesis give ___________________________.
- To assess how far the statistic is from the parameter, ____________________ the statistic.
- This value is called the **test statistic**:

- The test statistic measures how far the sample result is from the ___________________, in what ________________________, on a standardized scale.
- You can use the test statistic to find the ______________________ of the test

9.2A

Four step process for significance testing for proportions

State: What hypotheses do you want to test, and at what significance level?

- Test the hypothesis that true proportion of __(context)___ is _(p_0)__ at a ____ significance level.
- $H_0 : p = p_0$
- $H_a : p > p_0$ or $p < p_0$ or $p \neq p_0$
- Where p is the true proportion of ___(context)____.

Plan: Name procedure you are using. Check conditions.

- Use a one-sample z-test for proportions
- Random condition? Random Sample or Random Assignment
- Normal condition? $np_0 \geq 10$ and $n(1-p_0) \geq 10$
- Independence condition? independent or 10% rule

Do: If conditions are met, perform calculations

- Compute p
- Compute the test statistic z (show work) $z = \dfrac{p - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$
- Find the P-value = normal cdf(lower bound, upper bound, μ = 0, σ = 1)
  If Ha is > then lower bound is test statistic, upper bound is 10
  If Ha is < then lower bound is -10, upper bound is test statistic
  If Ha is ≠ then use one of the above and times by 2.

Conclude: Interpret the results of your test in the context of the problem.

- Since our P-value ____ is greater than our significance level ____, we fail to reject H_0. We do not have sufficient evidence _____(H_a in context)______.
- Since our P-value ____ is less than our significance level ____, we reject H_0. We have sufficient evidence _____(H_a in context)______.

9.2A

What happens when the data don't support Ha?

There is ________________to continue with the significance test. The conclusions is clear,__________ __________________. If you aren't paying attention, you may _____________________ the test. The test will give you the _______________________, fail to reject H_o. A lot more work with the _____________________!

<table>
<tr><td>9.2A</td><td>Sample: On shows like American Idol, contestants often wonder if there is an advantage to performing last. To investigate this, a random sample of 600 American Idol fans is selected and they are shown the audition tapes of 12 never-before-seen contestants. For each fan, the order of the 12 videos is randomly determined. Thus, if the order of performance doesn't matter, we would expect approximately 1/12 of the fans to prefer the last contestant they view. In this study, 59 of the 600 fans preferred the last contestant they viewed. Does this data provide convincing evidence that there is an advantage to going last?
State: Test the hypothesis that true proportion of ____ is ___ at a ____ significance level.
H_0:
H_a:
where p = the true proportion of ____
Plan: If conditions are met, we will perform a _______________________________________.
• Random:

• Normal:

• Independent:

Do:

Conclude: Since the P-value is __________ than _________________, we ____________________ the null hypothesis. There ___________ convincing evidence to conclude that there is an advantage to performing last in American Idol.</td></tr>
<tr><td>9.2A</td><td>Sample: According to the National Campaign to Prevent Teen and Unplanned Pregnancy, 20% of teens aged 13 to 19 say that they have electronically sent or posted sexually suggestive images of themselves. The counselor at a large high school worries that the actual figure might be higher at her school. To find out, she gives an anonymous survey to a random sample of 250 of the school's 2800 students. All 250 respond, and 63 admit to sending or posting sexual images. Carry out a significance test at the $\alpha = 0.05$ significance level. What conclusion should the counselor draw?

State: Test the hypothesis that the true proportion of teens aged 13 to 19 that have electronically sent or posted sexually suggestive images of themselves is 0.20
H_0: p = 0.2 H_a: p > 0.2 where p is the proportion of students "sexting"
Plan: We will perform a one-sided z-test for proportions
Random?

Normal?

Independent?

Do: z =

P-value

Conclude: Since the P-value is __________ than _________________, we ____________________ the null hypothesis. There ___________ convincing evidence to conclude that more than 20% of teens at the school has sent sexually explicit photos of themselves.</td></tr>
</table>

9.2B	**<u>More on Two sided tests</u>** We perform a two sided test when looking for convincing evidence that the true parameter is different from the hypothesized value of the parameter, ____ **<u>Sample:</u>** According to the Centers for Disease Control and Prevention (CDC) Web site, 50% of high school students have never smoked a cigarette. Taya wonders whether this national result hold true in her large, urban high school. For her statistics class, Taya takes an SRS of 150 students from her school. She gets responses from all 150 students, and 90 say that they have never smoked a cigarette. What should Taya conclude? Give appropriate evidencd to support your answer. **State:** Test the hypothesis that true proportion of ____ is ___ at a ____ significance level. H_0 : H_a : where *p* = the true proportion of ____ **Plan:** If conditions are met, we will perform a _______________________________________. • Random: • Normal: • Independent: **Do:** **Conclude:** Since the *P*-value is __________ than _________________, we ____________________ the null hypothesis. There ___________ convincing evidence to conclude that __________________ ____.
9.2B	**<u>Confidence intervals and significance tests</u>** **Significance tests** can tell us if the smaller specific population proportion ___________ from the entire population proportion. **Confidence intervals** give us an idea of what the _______________________ may be. Therefore, a confidence interval can be more ____________________________. ***Sample:*** Taya found that 90 of an SRS of 150 students said that they had never smoked a cigarette. We checked the conditions for performing the significance test earlier. Before we construct a confidence interval for the population proportion p, we should check that both $np \geq 10$ and $n(1-p) \geq 10$. Since the number of successes and the number of failures in the sample are 90 and 60, we can proceed with our calculations. Calculate the 95% confidence interval for the true proportion of students at the school that never smoked a cigarette. **State:** Estimate the true proportion of students at the high school that have ______________________ ____________________ with ____ confidence. **Plan:** Use a 1 sample z-interval for proportions. Conditions were checked in the significance test above. **Do:** z* = Interval: **Conclude:** We are ___% confident that the interval from _______ to _______ captures the true proportion of high school students have ___________________________________ at Taya's School.

9.2B

<u>Connection between confidence intervals and two sided significance tests</u>

There is a link between confidence intervals and two-sided significance tests:

- The confidence interval gives an approximate range of values p_o that would ________________ by a two sided test at the ________________________________ significance level.
- However, with proportions the link is ___________________________________ because of the standard error used for the confidence interval is based on the sample proportion p , while the denominator statistic is based on the value ___________________________________

Test statistic: $z = \dfrac{p - p_o}{\sqrt{\dfrac{p_o\left(1 - p_o\right)}{n}}}$ **Confidence interval:** $p \pm z^* \sqrt{\dfrac{p\left(1 - p\right)}{n}}$

- So on the AP exam it is acceptable to use a confidence interval rather than the test statistic to address a ________________ alternative hypotheses but ___________ alternative hypotheses.

9.1B

Errors in a Significance Test

9.1B

<u>Making Conclusions based on the sample</u>

When we draw conclusions from a significance test, we _________ our conclusions will be _________.

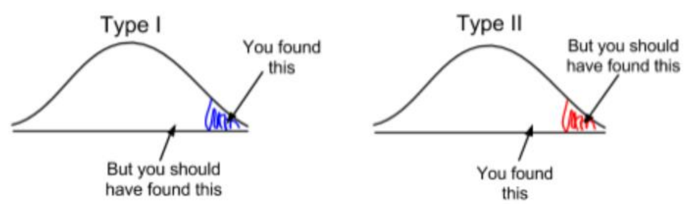
Truth about the population

		H_0 is true	H_0 is false (H_a true)
Your conclusion based on the sample	Reject H_o		
	Fail to reject H_o		

What happens when our conclusions are _________. There are two types of mistakes we can make:

1. Type 1 error: ___________________________________.
   (Determined they are guilty when really they are innocent)

2. Type 2 error: ____.
   (Determined they are innocent when really they are guilty)



9.1B

Sample: Heading into the election, Jack needs 50% of the vote (or higher) to win the election. Jack is feeling pretty confident that he will win the election. Is his confidence warranted?

$H_0 : p = 0.5$

$H_a : p > 0.5$

a.) Describe in context a type 1 error and its consequences.

b.) Describe in context a type 2 error and its consequences.

9.1B

Error probabilities

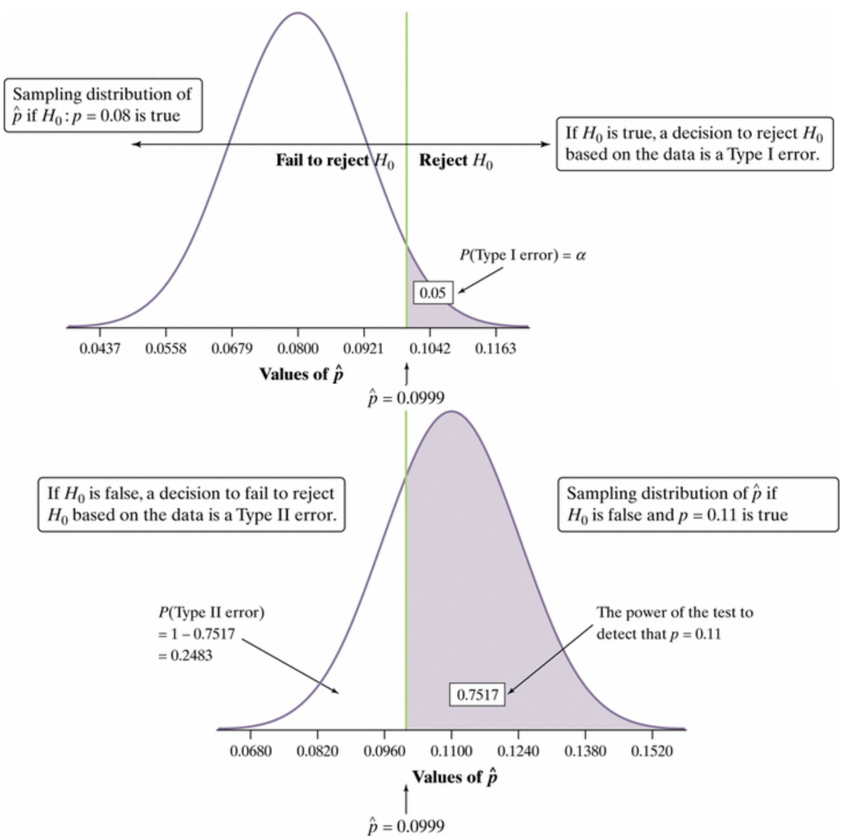
We can assess the ______________________________ of a significance test by looking at the probabilities of making these two errors.

Truth about the population

Your conclusion based on the sample		H_0 is true	H_0 is false (H_a true)
	Reject H_o	WRONG!!!! Type 1 error P(type 1 error) = α	CORRECT!!!! Power The power (probability of correctly rejecting H_0) = 1 - β
	Fail to reject H_o	CORRECT!!!! Inconclusive test . . . Didn't really learn anything	WRONG!!!!!! Type 2 error P(type 2 error) = β

Why is power important?

- The probability of making a Type II error is ___________________. So the higher the power, the ______ the chance of making a type 2 error.
- Low power means the null hypothesis (H_0) is unlikely to be rejected and the experiment will be ____________________ (Did all that work for _____________).
- Unfortunately, calculating the power of a test can only be calculated if you know the true population value (which you rarely know).



Test: $H_0: p = 0.8$, $H_a: p > 0.8$ at a 0.05 significance level, the sampling distribution if p = 0.08 for n = 500 is shown.

The green line is at invNorm (area 0.95, μ = 0.08, $\sigma = \sqrt{\frac{0.08(0.92)}{500}} \approx 0.0121$) =

0.0999563

What if the true p = 0.11?

The sampling distribution of the true p = 0.11 is drawn

How to calculate power:

Calculate the probability of getting the 0.0999 or higher for the sampling distribution for p = 0.11

Power = normalcdf(0.0999, 100, 0.11, 0.014) ≈ 0.7517

What does the power mean? Interpreting Power

75.17% chance of correctly rejecting the null hypothesis when the true proportion is 0.11

9.1B	A certain cigarette brand advertises that the mean nicotine content of their cigarettes is 1.5 mg, but you are suspicious and plan to investigate the advertised claim by testing the hypotheses $H_0: \mu = 1.5$ versus $H_a: \mu > 1.5$ at the $\alpha = 0.05$ significance level. You will do so by measuring the nicotine content of 30 randomly selected cigarettes of this brand. (a) Describe what a Type I error would be in this context. (b) Describe what a Type II error would be in this context. (c) From the perspective of public health, which error—Type I or Type II—is more serious? Explain. (d) You have determined that at the α = 0.05 significance level, the power of the test against the alternative $\mu = 1.75$ is 0.88. Explain what the power of the test means in the context of the problem.
9.1B	**How is power affected by different parts of your study?** http://digitalfirst.bfwpub.com/stats_applet/stats_applet_9_power.html

Higher power	Lower power
• ____________ α (significance level) • ____________ sample size • ____________ standard deviation (variance/spread) • ____________ the difference between H_0 value and the true population value • Use a _____-sided test	• ____________ α (significance level) • ____________ sample size • ____________ standard deviation (variance/spread) • ____________ the difference between H_0 value and the true population value • Use a _____-sided test

9.1B	***Sample:*** The manager of a fast-food restaurant wants to reduce the proportion of drive-through customers who have to wait more than 2 minutes to receive their food once their order is placed. Based on store records, the proportion of customers who had to wait at least 2 minutes was p = 0.63. To reduce this proportion, the manager assigns an additional employee to assist with drive-through orders. During the next month the manager will collect a random sample of drive-through times and test the following hypotheses: $H_0: p = 0.63$ where p = the true proportion of drive-through customers who have $H_a: p < 0.63$ to wait more than 2 minutes after their order is placed to receive their food. Suppose that the manager of the fast food restaurant wants to change some aspects of his study. How will these changes **affect the power** of the test? 1. To reduce the possibility of a Type I error and avoid the possibility of unnecessarily paying an extra employee, the manager reduces the significance level from 0.10 to 0.01. 2. To justify the additional cost of the extra employee, the manager decides that the true proportion must be reduced to at most 0.53. 3. To get faster results, the manager reduces the sample size from 250 to 100.

9.1B	**Planning studies: Determining Sample Size** Here are the questions we must answer to decide how many observations we need: 1. **Significance level**? If you insist on a _______________ significance level (such as 1% rather than 5%), you have to take a _____________sample. A smaller significance level requires ______________________ evidence to reject the null hypothesis. 2. **Practical Importance**? At any significance level and desired power, detecting a ___________ _____________requires a ___________________ than detecting a large difference. 3. **Power**? If you insist on ______________ power (such as 99% rather than 90%), you will need a _________________ sample. Higher power gives a better chance of detecting a __________ ________________________________.
9.3	# Significance tests for population means
9.3A	• When working with population ____________________ we use _____________ since they are based from the z-values from a standard normal distribution. • When working with population _______________, we use __________ from a t distribution with n – 1 degrees of freedom. • To keep it straight we have "**zap tax**" for ___________________________________
9.3A	**Conditions for population means using a t distribution** • **Random:** The data come from a ______________________ from a sample of size n from the population of interest or a __________________________. This condition is very important. • **Normal**: **Sample size less than 15**: use t procedures if the data appear close to normal. If it is **skewed** or has **outliers** ________________ use t procedures. **Sample size of at least 15**: the t procedures can be used except if the data has __________ _______________________ **Large samples**: The t procedures can be used even for clearly skewed distributions when the sample is _____________________________ • **Independent**: individual observations are independent so if we sample without replacement the 10% must be met __________________________
9.3A	**Calculations: Test statistic and P-value** • Perform calculations assuming the H_o _______________. • The test statistic measures how ________________ from the parameter value specified by H_o in standardized units. • Due to our t distribution our test statistic is When the normal condition is met, this statistic has a _________________ with __________ degrees of freedom. • Once we have calculated the test statistic we can use table B or a calculator to find the _______ for a significance test about μ.
9.3A	**Using to calculator to compute P-values from a t distribution** • Go the the distribution menu (2^{nd} vars) • Choose tcdf(

9.3A

Four step process for significance testing for means

State: What hypotheses do you want to test, and at what significance level?

- Test the hypothesis that true mean of __(context)___ is _(μ_0)__ at a ____ significance level.
- $H_0 : \mu = \mu_0$
- $H_a : \mu > \mu_0$ or $\mu < \mu_0$ or $\mu \neq \mu_0$
- Where μ is the true mean of ___(context).

Plan: Name procedure you are using. Check conditions.

- Use a one-sample t-test for means
- Random condition? Random Sample or Random Assignment
- Normal condition? n<15, IF close to normal, NO outliers or skewness
  15≤n<30, IF NO outliers or strong skewness
  n≥30 -meets normal condition no matter what-
- Independence condition? independent or 10% rule

Do: If conditions are met, perform calculations

- Compute $\bar{x}$ and S_x
- Compute the test statistic t (show work) with df = n -1, $t = \dfrac{\bar{x} - \mu_0}{S_x / \sqrt{n}}$
- Find the P-value = tcdf (lower bound, upper bound, df)
  If Ha is > then lower bound is test statistic, upper bound is 100
  If Ha is < then lower bound is -100, upper bound is test statistic
  If Ha is ≠ then use one of the above and times by 2.

Conclude: Interpret the results of your test in the context of the problem.

- Since our P-value ____ is greater than our significance level ____, we fail to reject H_0. We do not have sufficient evidence _____(H_a in context)______.
- Since our P-value ____ is less than our significance level ____, we reject H_0. We have sufficient evidence _____(H_a in context)______.

9.3A	***Sample:*** Every road has one at some point—construction zones that have much lower speed limits. To see if drivers obey these lower speed limits, a police officer used a radar gun to measure the speed (in miles per hour, or mph) of a random sample of 10 drivers in a 25 mph construction zone. Here are the results: 27 33 32 21 30 30 29 25 27 34 **Problem:** Can we conclude that the average speed of drivers in this construction zone is greater than the posted 25 mph speed limit? ***State:*** We will test the hypothesis that the true mean speed of drivers is 25mph in this construction zone. H_0: μ = H_a: μ Where μ is the ____ ***Plan:*** We will perform a ____ Random: Normal: Independent: ***Do:*** ***Conclude:*** Since the *P*-value is __________ than _________________, we ____________________ the null hypothesis. There ___________ convincing evidence to conclude that __________________ ____. (b) Given your conclusion in part (a), which kind of mistake—a Type I or a Type II error—could you have made? Explain what this mistake means in this context.

9.3A	***Sample:*** A college professor suspects that students at his school are getting less than 8 hours of sleep a night, on average. To test his belief, the professor asks a random sample of 28 students, "how much sleep did you get last night?" Here are the data (in hours): 9 6 8 6 8 8 6 6.5 6 7 9 4 3 4 5 6 11 6 3 6 6 10 7 8 4.5 9 7 7 Do these data provide convincing evidence in support of the professors suspicion? Carry out a significance test at the $\alpha = 0.05$ level to help answer this question. ***State:*** Test the hypothesis that the true mean hours of sleep college student gets is 8 hours. H_0: μ = ___ H_a: μ Where μ is the true mean hours of sleep a college student gets. ***Plan:*** We will perform a one-sample t-test for means Random? SRS of 28 students at the college Normal? The normal probability plot shows Independent? ***Do:*** ***Conclude:*** Since the *P*-value is __________ than ________________, we ___________________ the null hypothesis. There ___________ convincing evidence to conclude that __________________ ____.

9.3A

Two sided significance tests

In the children's game Don't Break the Ice, small plastic ice cubes are squeezed into a square frame. Each child takes turns tapping out a cube of "ice" with a plastic hammer hoping that the remaining cubes don't collapse. For the game to work correctly, the cubes must be big enough so that they hold each other in place in the plastic frame but not so big that they are too difficult to tap out. The machine that produces the plastic ice cubes is designed to make cubes that are 29.5 millimeters (mm) wide, but the actual width varies a little. To make sure the machine is working well, a supervisor inspects a random sample of 50 cubes every hour and measures their width. The output summarizes the data from a sample taken during one hour.

variable	N	Mean	SEmean	stdDev	Min	Q1	Median	Q3	Max
Width	50	29.4874	0.0132	0.0935	29.2717	29.4225	29.4821	29.5544	29.7148

Do these data give convincing evidence that the mean width of cubes produced this hour is **not** 29.5 mm?

State: Test the hypothesis that the true mean width of "ice cubes" is 29.5mm.

H_0: μ = 29.5mm H_a: μ ≠ 29.5mm

Where μ is the true mean width of "ice cubes".

Plan: We will perform a one-sample t-test for means

Random? SRS of ________ "ice cubes"

Normal?

Independent? More than _________ "ice cubes" produced.

Do:

Conclude: Since the *P*-value is __________ than _________________, we ____________________ the null hypothesis. There ___________ convincing evidence to conclude that __________________ ____.

9.3A

Two sided significance tests and confidence intervals

- Unfortunately the significance test doesn't tell us the _______________________, for that we need a confidence interval.
- The connection between significance tests and confidence intervals is even ________________ __________ than it was for proportions. That is because both the test statistic and the confidence interval use the standard error of $\bar{x}$ in the calculations.
- So when the two sided significance test at a level α, ___________________, the 100(1 – α)% confidence interval for μ _____________________ the hypothesized value μ_o
- And when the test _________________________ Ho, the confidence interval ___________ μ_o.

Estimate of Collection 1 Estimate Mean

Attribute (numeric): Width

Interval estimate for population mean of **Width**

```
Count:              50
Mean:               29.4874 mm
Std dev:            0.0934676 mm
Std error:          0.0132183 mm
Confidence level:   95.0 %
Estimate:           29.4874 mm +/- 0.0265632 mm
Range:              29.4609 mm to 29.514 mm
```

Sample: Don't break the ice

Here is Fathom output for a 95% confidence interval for the true mean width of plastic ice cubes produced this hour.

(a) Interpret the confidence interval. Would you make the same conclusion with the confidence interval as you did with the significance test in the previous example?

(b) Interpret the confidence level.

9.3B **<u>Inference for means: Paired data</u>**

- Comparative studies are more convincing than ______________________ investigations.
- Therefore, one-sample inference ______________________ than comparative inference.
- Study designs that involve making ______________________ on the same individual, or one observation on each of ______________________ individuals, result in **paired data**. (this experiment design is called matched pairs)
- When paired data result from measuring the same variable twice, we can make comparisons by analyzing ______________________ in each pair. If the conditions for inferences are met, we can use one-sample t procedures to perform inference about the ______________________ ______________________. These methods are called **paired t procedures**.

9.3B ***Sample:*** For their second semester project in AP Statistics, Libby and Kathryn decided to investigate which line was faster in the supermarket, the express lane or the regular lane. To collect their data, they randomly selected 15 times during a week, went to the same store, and bought the same item. However, one of them used the express lane and the other used a regular lane. To decide which lane each of them would use, they flipped a coin. If it was heads, Libby used the express lane and Kathryn used the regular lane. If it was tails, Libby used the regular lane and Kathryn used the express lane. They entered their randomly assigned lanes at the same time and each recorded the time in seconds it took them to complete the transaction.

Time in express lane (seconds)	Time in regular lane (seconds)	
337	342	
226	472	
502	456	
408	529	
151	181	
284	339	
150	229	
357	263	
349	332	
257	352	
321	341	
383	397	
565	694	
363	324	
85	127	

Carry out a test to see if there is convincing evidence that the express lane is faster.

State:

Plan:

	Do: **Conclude:**
9.3B	**<u>About paired data</u>** • Individual scores are ____________________ • Differences in scores are not dependent (____________________________) • Be sure to report _____________________________ when using calculator • If subjects in an experiment were not randomly chosen, we can't generalize our findings to _________________________. • If subjects in an experiment were not randomly assigned a treatment, we can't make an _______________________________________. • A confidence interval gives _____________________________ than a significance test.
9.3B	**<u>Using tests wisely</u>** Significance tests are widely used in reporting the _______________________ in many fields. New drugs require significant evidence of ____. Courts ask about _________________________________ in hearing discrimination cases. In all cases, statistical significance is valued because it points to an effect that is ________________ ________________________. • Statistical significance is not the same thing as ___________________. Pay attention to the actual data (________________________) along with the statistical significance to avoid this. • The foolish user of statistics who feeds the data to a calculator or computer without exploratory analysis will often be embarassed. If you were to actually ________________, would it be worthwhile to calculate? • Don't ignore the lack of significance. Usually ______________ will be needed to help us make some conclusions. • When planning a study, verify that the test you plan to use has a _______________________ of detecting a difference of the size you hope to find. • Statistical inference is not valid for ______________________. Badly designed surveys or experiments will yield invalid results. Always ask how the data was _____________________. • Running tests multiple times to get the significance you want will have ___________________.