

This class is being conducted over Zoom. As the instructor, I will be **recording** this session. I have disabled the recording feature for others so that no one else will be able to record this session. I will be posting this session to the course's website.

If you have privacy concerns and **do not wish to appear in the recording**, you may turn video off (click "**stop video**") so that Zoom does not record you.

The chat box is always open for discussion and questions to the entire class. You may also send messages privately to the instructor or the TAs. Please note that Zoom saves all chat transcripts.

I create a live transcription of each session using **Otter.ai**. This means that Otter.ai will transcribe anything spoken over the Zoom audio. The transcript will be posted with the session video on the course website.

# One Sample Inference

Stats 7

Mary Ryan

Aug. 27, 2020



Course website:

<https://canvas.eee.uci.edu/courses/28451>



Slides can be found at:

<https://maryryan.github.io/stats7-SS2-2020-slides/stats7-SS2-2020-oneSampleInference/stats7-SS2-2020-oneSampleInference>

# Learning Objectives

By the end of today's lecture, you should be able to:

- calculate confidence intervals and test statistics
- understand testing error
- interpret what confidence intervals tell us about our estimate of the population parameter
- draw an appropriate conclusion from a hypothesis test in the context of our question of interest

# What is One-Sample Inference?

- When we perform **one-sample inference**, we are using our sample data to infer or make assumptions about one population of interest
  - What does our sample mean tell us about the population?
- One-sample inference is how we turn data into knowledge we can use
- Several ways to perform inference...

# Capturing the Population Mean

- Recall: our sample mean is only one estimate of the population mean
  - If we took other samples, we would get other sample means with different values
  - We can throw a spear where we see a fish but we will probably miss

# Capturing the Population Mean

- Recall: our sample mean is only one estimate of the population mean
  - If we took other samples, we would get other sample means with different values
  - We can throw a spear where we see a fish but we will probably miss
- It would be nicer if there were a **range of plausible values** between which we might think the true population mean might lie, with a certain degree of certainty
  - If we toss a net in the general area we saw a fish, we have a good chance of catching it

# 95% Confidence Intervals

- Maybe we would think a plausible range of values might be the middle 95% of the sampling distribution of the sample mean  $\bar{x}$

$$P(a < \bar{X} < b) = 0.95$$

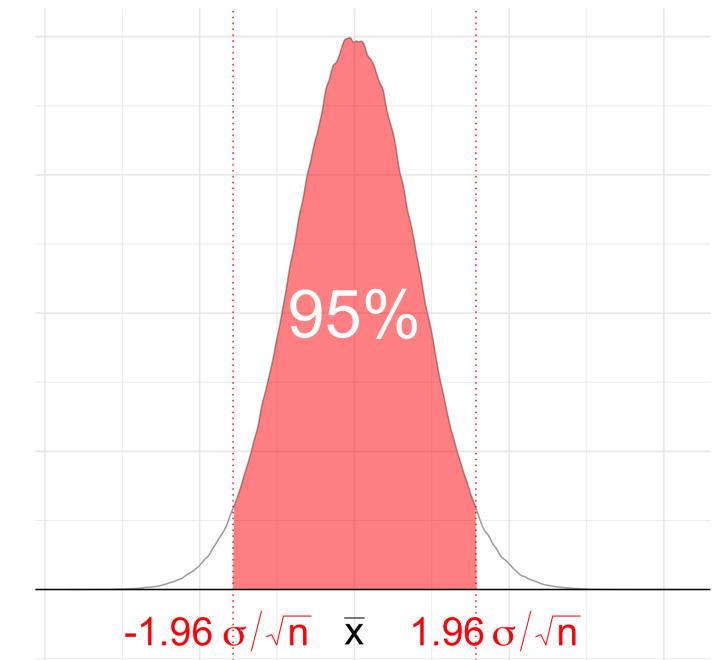
- But what are  $a$  and  $b$ ?

# 95% Confidence Intervals

- Maybe we would think a plausible range of values might be the middle 95% of the sampling distribution of the sample mean  $\bar{x}$

$$P(a < \bar{X} < b) = 0.95$$

- But what are  $a$  and  $b$ ?
- We know that the middle 95% of a sampling distribution is the range of values within 1.96 standard deviations above and below the mean
  - Standard deviation for the sampling distribution is  $\frac{\sigma}{\sqrt{n}}$  -- here we call it the **standard error**



# 95% Confidence Intervals

- So:

$$P\left(-1.96 \frac{\sigma}{\sqrt{n}} < \bar{X} < 1.96 \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

# 95% Confidence Intervals

- So:

$$P\left(-1.96 \frac{\sigma}{\sqrt{n}} < \bar{X} < 1.96 \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

We call this a 95% confidence interval

$$\left(\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}\right)$$

- We can shorthand the math to:

$$\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}$$

- We call the red portion of the equation the **margin of error**

# 95% Confidence Intervals

"We are 95% confident that the interval

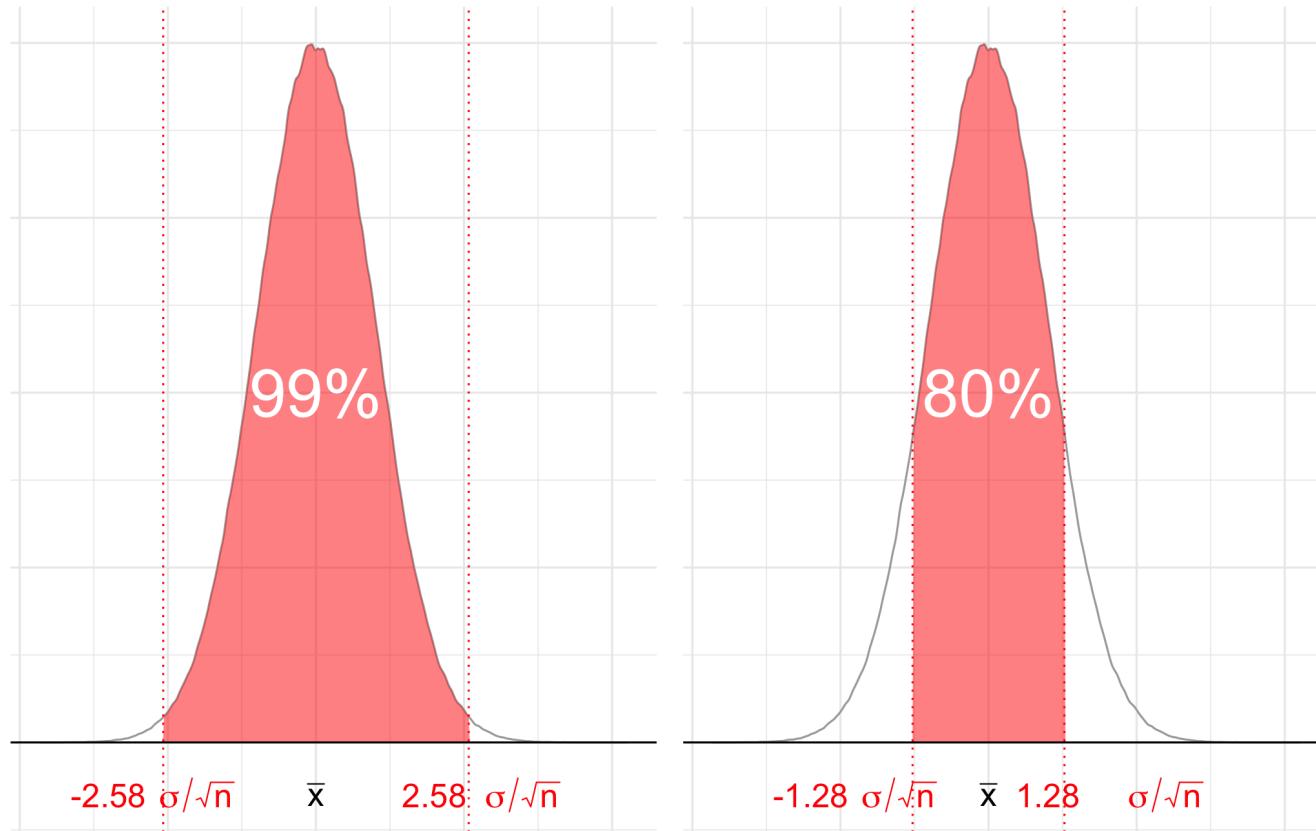
$$\left( \bar{x} - 1.96 \frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}} \right)$$

captures/contains the true population mean."

- We call 95% the **confidence level**
- Helps tell us how certain we are of our estimate of the mean
- The true population mean is **constant**
  - The confidence interval is **random**: it changes every time we take a new sample (because it's centered over the sample mean)

# General Confidence Intervals

- What if we think an interval other than the middle 95% is appropriate?
  - We change the number of standard errors we're adding and subtracting to  $\bar{x}$



# General Confidence Intervals

- Can write confidence interval formula generally as:

$$\bar{x} \pm Z_{1 - \frac{1}{2}\text{confidence}}^* \frac{\sigma}{\sqrt{n}}$$

# General Confidence Intervals

- Can write confidence interval formula generally as:

$$\bar{x} \pm Z_{1 - \frac{1}{2}\text{confidence}}^* \frac{\sigma}{\sqrt{n}}$$

- Width of confidence interval depends on:
  - confidence level
  - variance of observations in population
  - sample size

# General Confidence Intervals

- Can write confidence interval formula generally as:

$$\bar{x} \pm Z_{1-\frac{1}{2}\text{confidence}}^* \frac{\sigma}{\sqrt{n}}$$

- Width of confidence interval depends on:

- confidence level
- variance of observations in population
- sample size

- **Wider** confidence interval:

- **increase** confidence
- **increase** population variance
- **decrease** sample size

- **Narrower** confidence interval:

- **decrease** confidence
- **decrease** population variance
- **increase** sample size

# Calculator: ZInterval

To get to the calculator function on a TI-84:

- STAT > TESTS > 7: ZInterval

To calculate confidence interval:

- input:
  - population standard deviation ( $\sigma$ )
  - sample mean ( $\bar{x}$ )
  - sample size (n)
  - confidence level (C-Level)
- then arrow down to CALCULATE and press ENTER

# Example: SAT

The Scholastic Assessment Test (SAT) is a standardized test that's (was? 🤔) widely used by colleges and universities to help make admission decisions. Scores can range 400 to 1600. SAT scores have a population standard deviation of 200. We take a sample of 100 scores and get an average of 956.

- Calculate a 95% confidence interval.

# Example: SAT

The Scholastic Assessment Test (SAT) is a standardized test that's (was? 🤔) widely used by colleges and universities to help make admission decisions. Scores can range 400 to 1600. SAT scores have a population standard deviation of 200. We take a sample of 100 scores and get an average of 956.

- Calculate a 90% confidence interval.

# Example: Solar Energy

A Pew Research poll found that 88.7% of a random sample of 1000 American adults supported expanding the role of solar power.

- Calculate a 95% confidence interval.

# Example: Solar Energy

A Pew Research poll found that 88.7% of a random sample of 1000 American adults supported expanding the role of solar power.

- Calculate a 99% confidence interval.

# Confidence Intervals with Sample Standard Deviation

- We can also create confidence intervals if we're estimating the population standard deviation with the sample standard deviation
  - Instead of using the Normal distribution to figure out how many standard deviations above and below the mean, **we use the T distribution**

$$\bar{x} \pm T^*_{1 - \frac{1}{2}\text{confidence}, df=n-1} \frac{s}{\sqrt{n}}$$

# Calculator: TInterval

To get to the calculator function on a TI-84:

- STAT > TESTS > 8: TInterval

To calculate confidence interval:

- input:
  - sample mean ( $\bar{x}$ )
  - sample standard deviation ( $S_x$ )
  - sample size (n)
  - confidence level (C-Level)
- then arrow down to CALCULATE and press ENTER

# Example: Mercury in Dolphins

Elevated mercury concentrations are an important problem for both dolphins and other animals. We randomly sample 19 dolphins and test their muscle for mercury levels. We get a sample mean of 4.4, and a sample standard deviation of 2.3.

- Calculate a 95% confidence interval.

# Example: Mercury in Dolphins

Elevated mercury concentrations are an important problem for both dolphins and other animals. We randomly sample 19 dolphins and test their muscle for mercury levels. We get a sample mean of 4.4, and a sample standard deviation of 2.3.

- Calculate a 99% confidence interval.

# Testing our Notions of the Population

- We might be gathering our sample because we want to see if a **preconceived notion** we have about the population is true
- If we believe the population mean is one value and the data gives us an estimate, how can we tell if the sample mean is **confirming or challenging** our assumption?

# How to Perform a Hypothesis Test

1. State your hypotheses
2. Calculate your Z-statistic
3. Find your critical value/calculate your p-value
4. Draw a conclusion
5. Interpret your conclusion in context

# Step 1: State your hypotheses

- Null hypothesis ( $H_0$ ) represents "status quo", or what we assume to be true without more evidence
- Question will usually give you the alternative hypothesis ( $H_A$ )
- Two types of tests:
  - Two-sided test:
    - $H_0$ : We assume the population mean is  $= \mu_0$
    - $H_A$ : We assume the population mean is  $\neq \mu_0$
  - One-sided test:
    - $H_0$ : We assume the population mean is  $\geq / \leq \mu_0$
    - $H_A$ : We assume the population mean is  $> / < \mu_0$
- Always write hypotheses in terms of the parameter of interest (not the sample estimate)

# Step 2: Calculate your Z-statistic

- If we know the population standard deviation:

$$z = \frac{\bar{x} - \mu_0}{\sqrt{\sigma/n}}$$

$$z = \frac{\hat{p} - p_0}{\sqrt{p(1-p)/n}}$$

- If we're estimating the population standard deviation with the sample standard deviation:

$$t = \frac{\bar{x} - \mu_0}{\sqrt{s/n}}$$

"If the null hypothesis was true, this is how many standard deviations away from the null hypothesis mean our sample mean is"

# Step 3: Find Your Critical Value/Calculate Your P-value

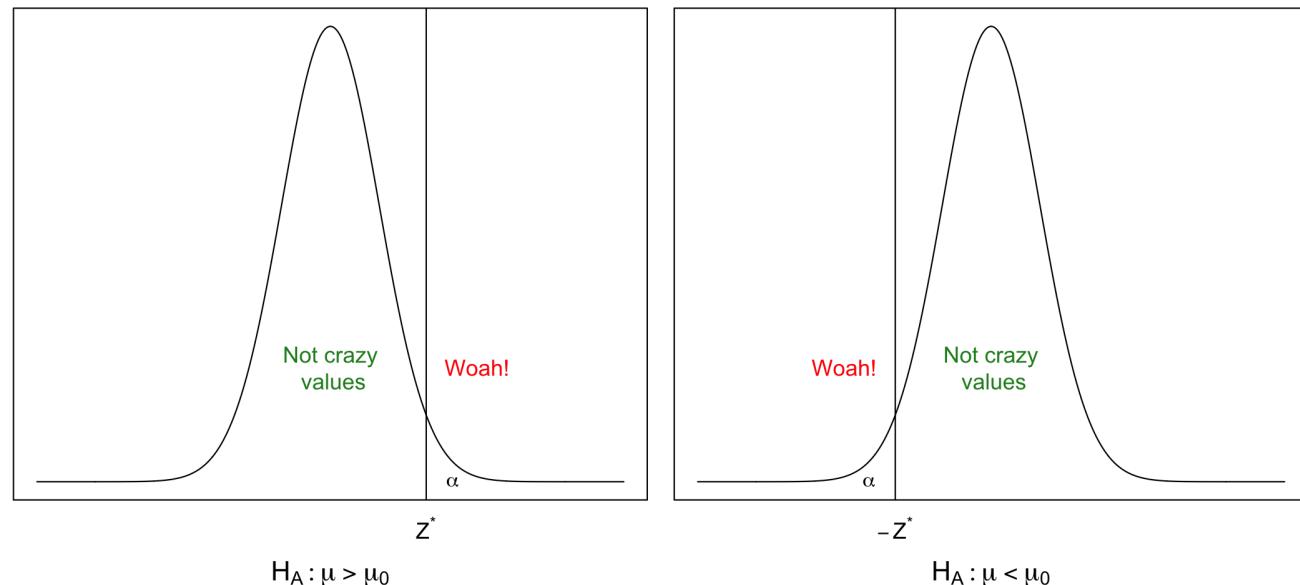
- Two ways to see if our sample mean is "too different" than the null hypothesis population mean

- Find "critical value" to compare test statistic against

- If the test statistic is **more extreme** than this value, we know our sample mean is **very different** than our null mean

- Determine before seeing data

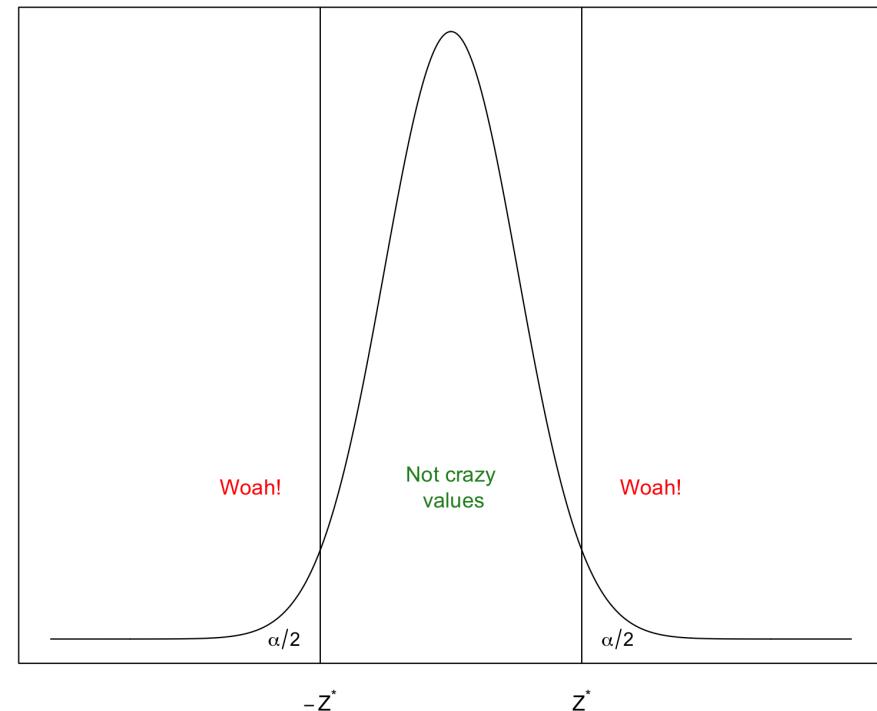
- For one-sided tests, we can determine critical value by the probability we would see sample means more extreme than it



# Step 3: Find Your Critical Value/Calculate Your P-value

1. Find "critical value" to compare test statistic against

- For two-sided tests, we can determine critical value by the probability we would see sample means more extreme than it
  - Then split that probability between both the tails



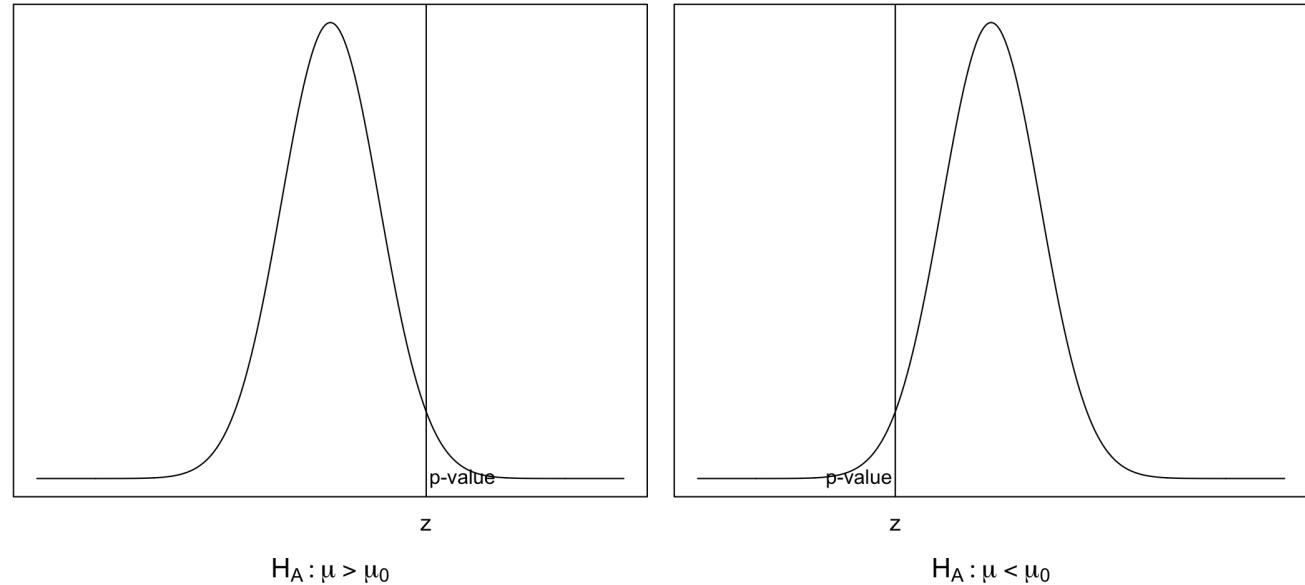
$$H_A: \mu \neq \mu_0$$

# Step 3: Find Your Critical Value/Calculate Your P-value

1. Calculate probability we see sample mean **as or more extreme** as your sample mean, **if the null hypothesis mean was true**

- Call this the **p-value**

- For a one-sided test:



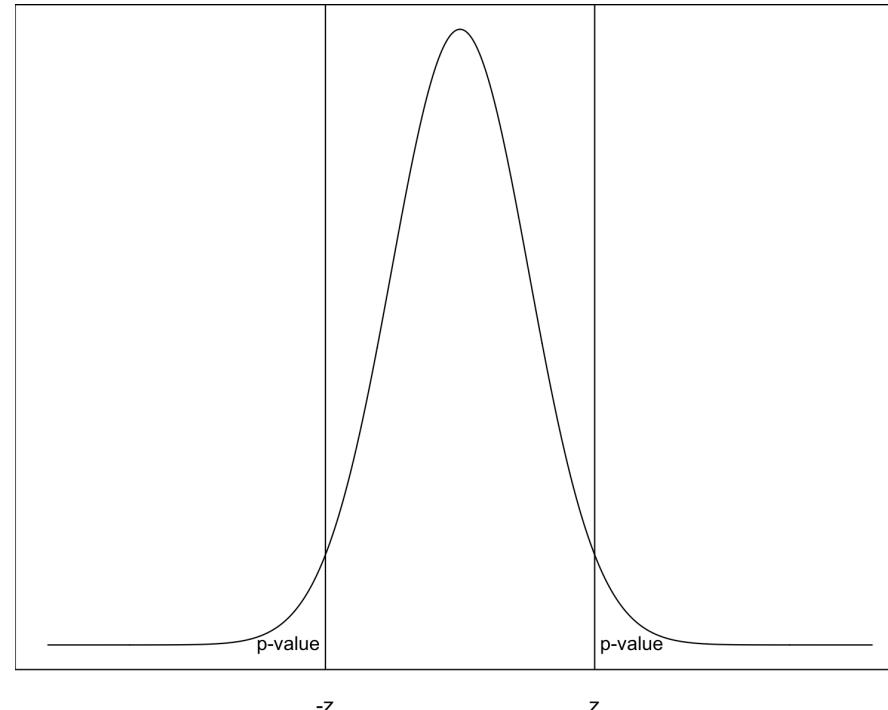
# Step 3: Find Your Critical Value/Calculate Your P-value

1. Calculate probability we see sample mean **as or more extreme** as your sample mean, **if the null hypothesis mean was true**

- Call this the **p-value**

- For a two-sided test, you need to look at the probability on **both sides** of the distribution

- Since Normal/T distributions symmetric, can find the probability on one side and **double it**



$$H_A : \mu \neq \mu_0$$

# Step 4: Draw a conclusion

- For two-sided test:
  - If  $|Z| > Z^*$ , **reject** the null
  - If  $|Z| < Z^*$ , **fail to reject** the null
- For one-sided test:
  - Greater-than test:
    - If  $Z > Z^*$ , **reject** the null
    - If  $Z < Z^*$ , **fail to reject** the null
  - Less-than test:
    - If  $Z < -Z^*$ , **reject** the null
    - If  $Z > -Z^*$ , **fail to reject** the null

# Step 4: Draw a conclusion

- For two-sided test:
  - If  $|Z| > Z^*$ , **reject** the null
  - If  $|Z| < Z^*$ , **fail to reject** the null
- For one-sided test:
  - Greater-than test:
    - If  $Z > Z^*$ , **reject** the null
    - If  $Z < Z^*$ , **fail to reject** the null
  - Less-than test:
    - If  $Z < -Z^*$ , **reject** the null
    - If  $Z > -Z^*$ , **fail to reject** the null

We never "accept" the null hypothesis!

- We are gathering evidence against the null hypothesis
  - If we don't have evidence against, we just can't assume the null is false
- Like in an American legal trial, where defendant assumed innocent until proven guilty
  - Judge gives verdict of "guilty" or "not guilty"
  - If found "not guilty", we aren't saying they are innocent - just that we don't have enough evidence to say they're guilty

# Step 5: Interpret your conclusion in context

- For a rejection
  - Why are we rejecting?
  - What strength of evidence do we have?
  - What does a rejection of the null hypothesis mean for our knowledge about the population?
- For a fail-to-reject
  - Why are we not rejecting?
  - Was our evidence not very strong?
  - What does a failure to reject the null hypothesis mean for our knowledge about the population?

# Calculator: Z-Test

To get to the calculator function on a TI-84:

- STAT > TESTS > 1: Z-Test

To calculate hypothesis test statistic and p-value:

- input:
  - population mean under null hypothesis ( $\mu_0$ )
  - population standard deviation ( $\sigma$ )
  - sample mean ( $\bar{x}$ )
  - sample size (n)
  - direction of test: ( $\mu : \neq \mu_0 < \mu_0 < \mu_0$ )
- then arrow down to CALCULATE and press ENTER

# Calculator: T-Test

To get to the calculator function on a TI-84:

- STAT > TESTS > 2: T-Test

To calculate hypothesis test statistic and p-value:

- input:
  - population mean under null hypothesis ( $\mu_0$ )
  - sample mean ( $\bar{x}$ )
  - sample standard deviation ( $S_x$ )
  - sample size (n)
  - direction of test: ( $\mu : \neq \mu_0 < \mu_0 < \mu_0$ )
- then arrow down to CALCULATE and press ENTER

# Example: Mercury in Dolphins

Say that mercury levels above 5.2 micrograms of mercury per wet gram of muscle in dolphins indicates that there may be pollution in the area.

As part of an ocean pollution watchdog group, we randomly sample 19 dolphins and test their muscle for mercury levels. We get a sample mean of 4.4, and a sample standard deviation of 2.3.

- What are our null and alternative hypotheses?

# Example: Mercury in Dolphins

Say that mercury levels above 5.2 micrograms of mercury per wet gram of muscle in dolphins indicates that there may be pollution in the area.

We randomly sample 19 dolphins and test their muscle for mercury levels. We get a sample mean of 4.4, and a sample standard deviation of 2.3.

- What is our test statistic? What is our critical value/p-value?

# Example: Mercury in Dolphins

Say that mercury levels above 5.2 micrograms of mercury per wet gram of muscle in dolphins indicates that there may be pollution in the area.

We randomly sample 19 dolphins and test their muscle for mercury levels. We get a sample mean of 4.4, and a sample standard deviation of 2.3.

- What's our conclusion? What does this mean?

# Testing Error

- Type I errors happen when we **reject** the null hypothesis when actually it is **true**
- Type II errors happen when we **fail to reject** the null hypothesis when actually it is **false**

		Truth: $H_0$ is...	
		True	False
Test Result: $H_0$ is...	Not Rejected	<b>Correct Decision</b>	Type II <b>Error</b> (False Negative)
	Rejected	Type I <b>Error</b> (False Positive)	<b>Correct Decision</b>

# What Else Can We Do Inference On?

- Much of the time we're focused on trying to make inference about the population mean, but that's not the only parameter we can work with!

# What Else Can We Do Inference On?

- Much of the time we're focused on trying to make inference about the population mean, but that's not the only parameter we can work with!
- We saw that we can also perform one-sample inference on sample proportions

# What Else Can We Do Inference On?

- Much of the time we're focused on trying to make inference about the population mean, but that's not the only parameter we can work with!
- We saw that we can also perform one-sample inference on sample proportions
- We can also perform one-sample inference on the slope of regressions!
  - Regression outputs will often give you both the estimated slope of a regression,  $\hat{\beta}$ , and the standard error of that slope
  - If we perform a hypothesis test on the regression slope, we often want to know if it is different from 0 -- many outputs will also calculate a t-statistic for you and calculate the two-sided p-value
- 95% confidence interval for regression slope:  
$$\text{slope} \pm t_{1-\frac{1}{2}\text{confidence}, df=n - \# of variables - 1} \times SE$$
- Hypothesis test for regression slope:
  - $H_0$ : True slope = /  $\geq$  /  $\leq$  0
  - $H_A$ : True slope  $\neq$  /  $>$  /  $<$  0
  - Test statistic for regression slope:  $t = \frac{\text{slope} - 0}{SE}$

# Example: MLB 2011

Is the number of times a team has a player go up to bat throughout a season (called "at-bats") related to the number of runs that teams scores? There are 30 teams in the dataset ( $n=30$ ).

$$\text{Runs} = \hat{\beta}_0 + \hat{\beta}_1 (\text{at-bats})$$

Regression output:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-2789.24289	853.6957215	-3.267256	0.0028706
at_bats	0.63055	0.1545414	4.080136	0.0003388

$$\text{Runs} = -2789.2429 + 0.6305(\text{at-bats})$$

- Hypotheses:
  - $H_0$ : slope = 0 (no association)
  - $H_A$  : slope  $\neq 0$  (some kind of association)
- Degrees of freedom:  $n - \# \text{ of variables} - 1 = 30 - 1 - 1 = 28$
- p-value: 0.0003
- Reject  $H_0$  : strong evidence to suggest the number of at-bats a team has is related to the number of runs the team scores