MOD 18 INTRODUCTION TO HYPOTHESIS TESTING

Upon successful completion of this activity and associated homework, students will be able to:

- Test a hypothesis (i.e. a claim) about a population proportion, and state the conclusion in context
- State correct hypotheses for a significance test about a population proportion or mean.
- Interpret P-values in context
- Interpret a Type I error and a Type II error in context and give the consequences of each.

An Introduction to Hypothesis Testing

1) Here is a research question.

   Has the percentage of U.S. adults who believe illegal drug use is a serious problem in their local area decreased from 32% in 2016?

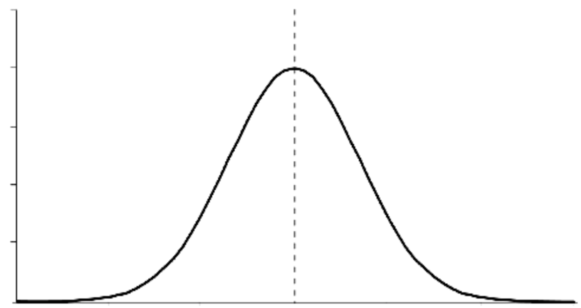
   A Gallup poll in October 2017 found that 29% of respondents agreed that illegal drug use is a serious problem in their local area. The results are based on telephone interviews with a random sample of 1,018 adults.

   a) Obviously, 29% (the sample proportion) is less than 32%. But we don't know whether the population proportion is likely to be less than 32%. Explain why.

   b) Hmmm … perhaps we need to determine whether it is likely or unlikely to get a sample proportion of 29% from a population with a proportion of 32%. In other words, if we assume the population proportion is 32%, we need to determine whether this sample is usual or unusual. Start by assuming $P = 0.32$. Then determine whether the sampling proportions are normally distributed (show your work). If the sampling proportions are <u>not</u> normally distributed, go to number 2. If the sampling proportion are normally distributed, construct the normal density curve on the next page (round the standard error to four decimal places, and the boundaries for the middle 95% to three decimal places).

Continued on the next page …

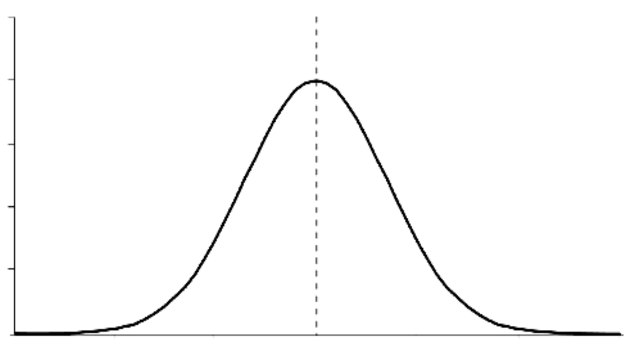
c) Mark the sample proportion from 2017 on the number line. Is the random sample from 2017 usual or unusual? How do you know?



d) If the sample proportion is usual, it is likely that the sample came from a population with a proportion of 32%. So we **cannot conclude** that the actual population proportion in 2017 is less than 32%. However, if the sample proportion is unusual, it is unlikely that the sample came from a population with a proportion of 32%. So we **can conclude** that the population proportion is less than 32%. What can you conclude from your answer to part b)? Explain, and try to write like a statistician.

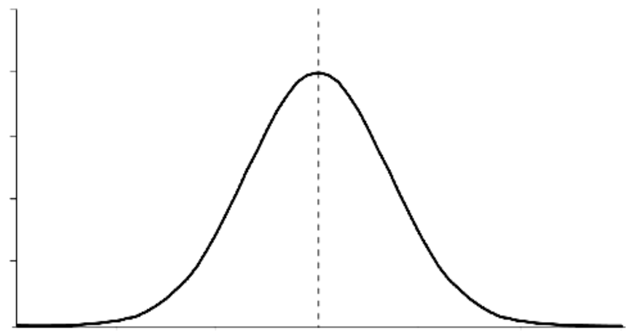
2) According to a September 2017 Gallup poll, 52% of Americans now trust the U.S. government to handle international problems, up from 49% in 2016. The results are based on telephone interviews with a random sample of 1,022 adults.

Has the population proportion actually increased? In other words, has the proportion of all American's who trust the U.S. government to handle international problems increased from 49% in 2016? Hint: you must determine whether the sample is usual or unusual (i.e. whether it is likely or unlikely to be drawn from the given population).



3) The CDC's National Health and Nutrition Survey indicates that 40% of US adults were obese in 2015-16. Suppose that 37% of a random sample of 200 US adults are obese in 2017-18.

Has the proportion of obese adults in the US decreased? Hint: you must determine whether the sample is usual or unusual (i.e. whether it is likely or unlikely to be drawn from the given population).



4) Look back at the *middle 95% intervals* you found in number 1, 2, and 3. Use your work in each of these problems to complete the table below.

Problem #	**n**	**SE**	**Interval for Middle 95%**	**Interval width**	**$\hat{p}$ usual or unusual**
1					
2					
3					

Which problem number has the largest *middle 95%* interval? Why do you suppose this interval is bigger than the others?

Activity: I'm a Great Free-throw Shooter!

(This Activity should be completed in the computer lab or in class using cell phones.)

A basketball player claims to make 80% of the free throws that he attempts. We think he might be exaggerating. To test the claim, we'll ask him to shoot some free throws virtually using the *Test of Significance* applet.

- Go to http://digitalfirst.bfwpub.com/stats_applet/stats_applet_15_reasoning.html
- Click "Show null hypothesis" so that the player's claim is visible in the graph.

1) The applet is set to take 10 shots. Click "Shoot". How many of the 10 shots did the player make? Do you have enough data to decide whether the player's claim is valid?

2) Click "Shoot" again for 10 more shots. Keep doing this until you are convinced that the player makes less than 80% of his shots or that the player's claim is true. To make your decision, how large did you need the sample to be (how many shots)?

3) Click "Show true probability" to reveal the truth. Was your conclusion correct?

4) Choose a new shooter and repeat numbers 5 through 7. Is it easier to tell that the player is exaggerating when his actual proportion of free throws made is closer to 0.8 or farther from 0.8?

5) A ___________________ is a formal procedure for comparing observed data with a claim (typically called the________________) whose truth we want to assess. The claim is a statement about a parameter, like the population proportion p or the population mean μ. We express the results of a significance test in terms of a probability that measures how well the data and the claim agree. Now let's take a closer look at the underlying logic of significance tests (a.k.a. hypothesis testes).

I'm a Great Free-Throw Shooter! *The logic of statistical tests.*

6) Our virtual basketball player claimed to be an 80% free-throw shooter. Suppose that he shoots 50 free throws and makes 32 of them.

$\hat{p}$ =

Let's the sample data and probability to draw a conclusion about the player's claim. Is there anything that we have previously learned that could be used to calculate the probability that the shooter will make 32 or fewer shots in 50 attempts? If so, go ahead and calculate the probability now.

Did your work provide convincing evidence? Explain.

7) There are two possible explanations of the fact that our virtual player made only $\hat{p}$ = 32/50 of his free throws even though he claims to make 80%. What are they?

First possible reason he only made 64% of his free-throws:

Second possible reason he only make 64% of his free-throws:

Which one of these two explanations is more likely to be correct?

8) The basic idea for a statistical test is:

Stating Hypotheses

9) A significant test starts with a careful statement of the claims we want to compare. Complete the following definitions:

Null hypotheses, H_o:

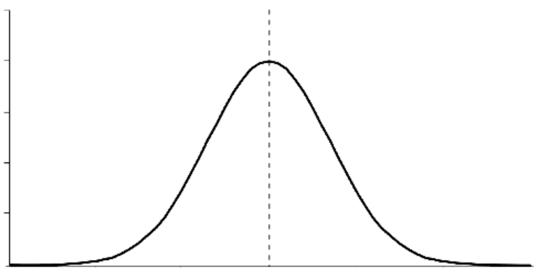
Alternative hypotheses, H_a:

10) Remember in our free-throw shooter example we think our friend is exaggerating when he claims to make at least 80% of his shots. What would our null and alternative hypotheses be in this case?

Identifying the Nature of a Hypothesis Test

For each scenario in numbers 15 through 19 below, write the claim and then write its opposite. Just ignore the bell curves for now.

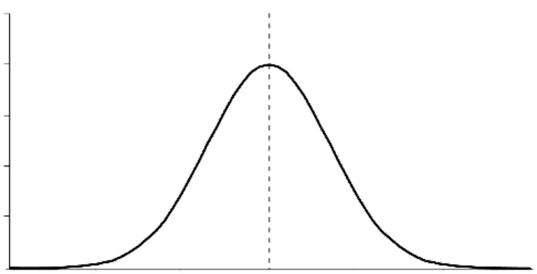
11) A university claims that the proportion of its student who graduate in 4 years is at least 82%.



Claim:

Opposite:

12) A water faucet manufacturer claims that the mean flow rate of a certain type of faucet is less than 2.5 gallons per minute.



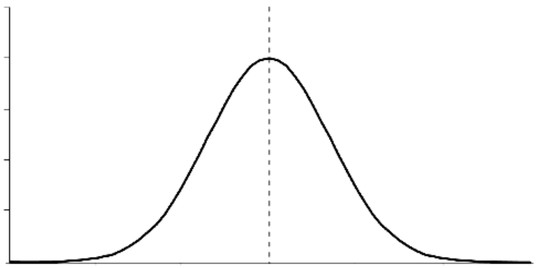
Claim:

Opposite:

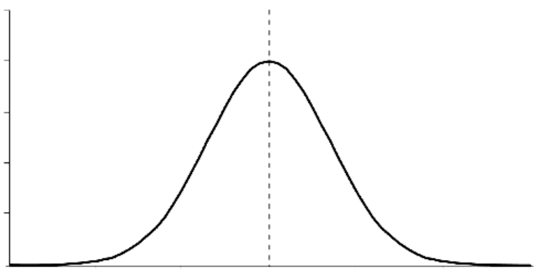
13) A cereal company claims that the mean weight of the contents of its 20-ounce size cereal boxes is more than 20-ounces.

Claim:

Opposite:



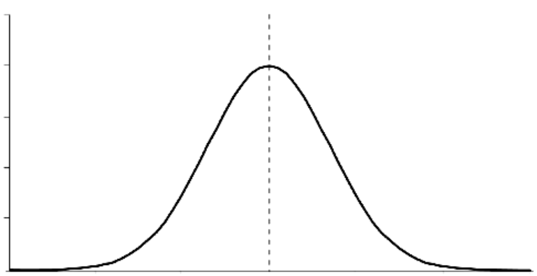
14) An automobile battery manufacturer claims that the mean life of a certain type of battery is 74 months.



Claim:

Opposite:

15) A radio station claims that its proportion of the local listening audience is greater than 39%.



Claim:

Opposite:

The null hypothesis, H_o, always includes an equal sign ($=$, $\leq$, or $\geq$).

And the alternative hypothesis, H_a, never includes an equal sign ().

Go back through numbers 15 through 19 and complete the following three tasks for each scenario.

a) **State the hypotheses (H_o and H_a) in words and in symbols.**
b) **Determine whether the hypothesis test is a left-tailed test, right-tailed test, or two-tailed test.**
c) **Sketch a normal sampling distribution and shade the area for the P-value.**

Decision Rule Based on *P*-value

16) To use a *P*-value to make a conclusion in a hypothesis test, compare the *P*-value to α.

- If $P \leq \alpha$, then we ____________________the H_o
- If $P > \alpha$, then we ____________________ the H_o

Types of Errors and Level of Significance

No matter which hypothesis represents the claim, you always begin a hypothesis test by assuming that the equality condition in the null hypothesis is true. So when you perform a hypothesis test, you make one of two decisions:

Reject the null hypothesis OR

Fail to reject the null hypothesis

A **Type I error** occurs if the null hypothesis is rejected when it is actually true.

A **Type II error** occurs if the null hypothesis is not rejected when it is actually false

Identifying Type I and Type II Errors

17) The USDA limit for salmonella contamination for chicken is 20%. A meat-packing company claims that its chicken falls within the limit. You perform a hypothesis test to determine whether the company's claim is true. In this scenario, when will a type I or type II error occur? Which is more serious?

18) A company specializing in a parachute assembly claims that its main parachute failure rate is not more that 1%. You perform a hypothesis test to determine whether the company's claim is true. In this scenario, when will a type I or type II error occur? Which is more serious?

19) Slow response times by paramedics, firefighters, and policemen can have serious consequences for accident victims. Several cities have begun to monitor emergency response times. In one such city, the mean response time to all accidents involving life=threatening injuries last year was μ = 6.7 minutes. At the end of this year, the city manager selects an SRS of 400 calls involving life-threatening injuries and examines the response times. The city manager wants to test for a decrease in the average response time. When will a type I or type II error occur? Which error would be more serious?

20) In a hypothesis test, the **level of significance** is your maximum allowable probability of making a type I error. It is denoted by α, the lowercase Greek letter alpha. Interpret this definition.

21) In a hypothesis test, the probability of a type II error is denoted by β, the lowercase Greek letter beta.

    a) How is the significance level, α, related to β?

    b) What happens when a researcher chooses a smaller α to reduce the probability of a Type I error?

Interpreting a Decision

22) You perform a hypothesis test for each of the following claims. How should you interpret your decision if you reject H_0? If you fail to reject the H_0?

a) H_0 (claim): A university claims that the proportion of its students who graduate in four years is 82%.

Reject H_0:

Fail to reject H_0:

b) H_a (claim): A government safety administration claims that the mean stopping distance (on a dry surface) for a Chevrolet Malibu is less than 148 feet.

Reject H_0:

Fail to reject H_0:

c) H_a (claim): A radio station claims that its proportion of the local listening audience is greater than 39%.

Reject H_0:

Fail to reject H_0: