Name: Key Hour: ____ Date: ____

Learning Targets

- State appropriate hypotheses for a significance test about a population parameter.
- Interpret a P-value in context.
- Make an appropriate conclusion for a significance test.

Lesson 9.1: Day 1: Is this gender discrimination?

A local engineering firm had to conduct a series of lay offs recently. They will lay off 10 people. The company has 180 employees that could be laid off. All are equally qualified so the company decides to use a lottery system to be carried out by the manager to decide who will be laid off. The manager posts a list of the employees to be laid off. Five employees are women and 5 are men. One of the women claims this is gender discrimination and starts a lawsuit against the company.

60 females
120 males

$F: \frac{60}{180} = \frac{1}{3}$

$M: \frac{120}{180} = \frac{2}{3}$

1. The manager responds, "How could there be gender discrimination when half of the employees laid off were female and half were male?" What additional information do you need to evaluate this statement?

   We need to know how many males + females work at the company.

2. How can you investigate the gender discrimination claim? Detail a process that could be used.

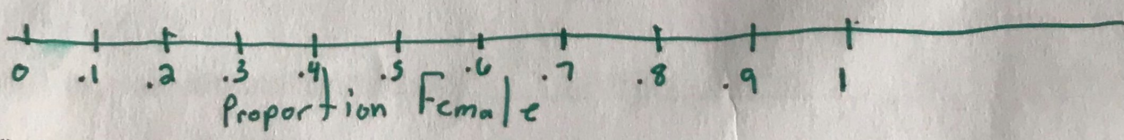
   Dice: 1 + 2 are women, 3 to 6 are men

   Spinner: 1/3 women, 2/3 men

   RNG: 1-60 women, 61-180 men. Random # Generator

3. Complete your investigation below.

   RNG: 1 → women, 2+3 men.



4. What percentage of the dots represent half or more females being laid off?

5. Interpret this percentage in context.

   Assuming the lottery was carried out fairly, there is a ____ probability of getting a sample proportion of 0.5 or higher.

6. Do you have convincing evidence of gender discrimination? Explain.

   No, a sample of 0.5 or higher happened fairly often.

Name: ______ Hour: ____ Date: ____

Lesson 9.1 Day 1– Significance Tests: The Basics

Significance test
A formal procedure for comparing observed data with a claim (hypothesis) 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 population
- Ask if sample data give good evidence against 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 not true!

Important ideas:

L.T. #1 Hypotheses

H_o: **Null Hypothesis: (What a person claims to be true)** will always use =, always use parameter

$H_o: \mu =$ null value or $H_o: p =$ null value

It is the claim about the parameter you are trying to find evidence against

H_a: **Alternative Hypothesis: (The claim we suspect is true instead of null hypothesis (H_o).**

One sided: $H_a: p >$ null value or $H_a: p <$ null value

$H_a: \mu >$ null value or $H_a: \mu <$ null value

Two sided: $H_a: p \neq$ null value or $H_a: \mu \neq$ null value

It is the claim about the parameter you are trying to find evidence for

L.T. #2 P- Value

The probability that we would get this sample result or one more extreme (in the direction specified by H_a) just by chance if H_o is true.

- **The smaller the p-value is,** the stronger the evidence against the null Hypothesis (H_o)
- **The p- value is a conditional probability:** P(sample data | H_o is true) given

L.T. #3 Conclusions

We do / do not **have convincing evidence** against H_o

Significant means likely not to happen by chance. alpha α = significance level

If p-value < α, then significant (meaning there is evidence against H_0.)

- **p-value < α** Reject H_o (evidence against H_o)
- **p – value > α** Fail to Reject H_o (not convincing evidence against H_o)

NEVER ACCEPT H_o as true!!!!!

In general, use $\alpha = 0.05$ unless otherwise noted concerned that teenagers are not getting enough calcium, on average. Is this true?

1. State appropriate hypotheses for performing a significance test. Be sure to define the parameter of interest. **(YOU MUST ALWAYS DEFINE PARAMETER OF INTEREST!)**

$H_0: \mu = 1300$ mg

$H_a: \mu < 1300$ mg

μ = true mean daily calcium intake of teens

Researchers decide to perform a test using the hypotheses stated in #1. They ask a random sample of 20 teens to record their food and drink consumption for 1 day. The researchers then compute the calcium intake for each student. Data analysis reveals that $\bar{x} = 1198$ mg and $s_x = 411$ mg. Researchers performed a significance test and obtained a P-value of 0.1404.

2. Explain what it would mean for the null hypothesis to be true in this setting.

If $H_0: \mu = 1300$mg is true, then the mean daily calcium intake in the population of teens is 1300 mg.

3. Interpret the P-value. P(sample data | H_0 is true)

given

There is a 0.1404 probability of getting a sample mean of 11.98 mg or less purely by chance given that the true mean daily calcium intake of teens is 1300mg.

4. What conclusion would you make at the $\alpha = 0.05$ level? → significance level

Because .1404 > 0.05, we fail to reject H_0. We do not have convincing evidence that true mean calcium daily intake of teen pop. is 1300mg.

–NOW TRY ON YOUR OWN

5. When Mike was testing a new 7-iron, the hypotheses were $H_0: \sigma = 15$ yards, $H_a: \sigma < 15$ yards

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

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

Always use context *

There is a 0.002 probability of getting a sample s.d. of 10.9 yards or less purely by chance given that the true s.d. of the distances Mike hits golf balls with new 7-iron is 15 yards.

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

.002 < 0.01, so we reject H_0.

We have convincing evidence that the s.d. with new 7irons is less than 15 yards.

TheStatsMedic