Chapter 8: Hypothesis Testing

Section 8.1: Basics of Hypothesis Testing

Stat 50

GOAL: Make a decision about p or μ based on $\hat{p}$ or $\bar{x}$ using probability $\hat{p}$. μ

pop. sample

Chapter 7: Find a sample then **estimate** whether the population fits within a certain interval. CI	**Chapter 8**: Given a past claim of the parameter, we will test whether or not it has changed. Δ

STRUCTURE OF A HYPOTHESIS TEST

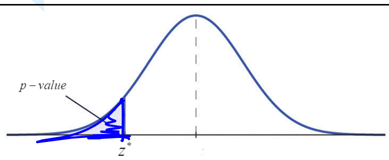
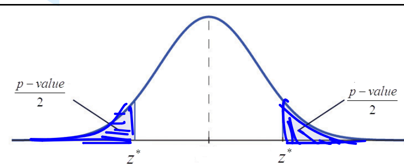
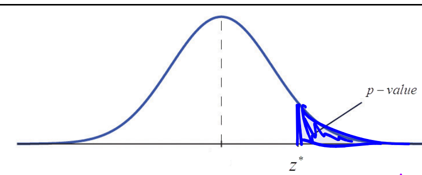
1) Make an assumption about reality	• make a hypothesis about population parameter
2) Look at a sample evidence	• find a point estimate to test pop. parameter
3) Determine whether it contradicts our assumption.	• make a decision if p or μ has changed.

We won't be 100% certain, we will just be able to tell if sample data substantiates/supports a statement or not.

HYPOTHESES STATEMENTS (GTK)

NULL HYPOTHESIS (H_0) nothing = 0	ALTERNATIVE HYPOTHESIS (H_1) or H_A
A statement of no change, no effect, no difference and is assumed true until statistical evidence indicates otherwise.	A statement that we are trying to find evidence to support ~~instead of~~ that Null Hyp (H_0) is "not true".

THREE TYPES OF HYPOTHESIS TESTS

LEFT-TAILED	TWO-TAILED	RIGHT-TAILED
H_0: *parameter* $= \#$ H_1: *parameter* $< \#$	H_0: *parameter* $= \#$ H_1: *parameter* $\neq \#$	H_0: *parameter* $= \#$ H_1: *parameter* $> \#$
		

(H_A: para changes smaller) H_A: Parameter changes H_A: parameter changes bigger

GTK!

Ex: **1)** What's the parameter? p μ **2)** What do "they say"? "past claim" **3)** What do we think? "H_A" **4)** What type of test? tail

The packaging on a light bulb says it should last 500 hours. Consumer Reports wants to know if the mean lifetime is actually less than that.	The standard deviation of the rate of return for some mutual funds is 0.08%. A manager believes the standard deviation might be higher than that.	According to a Gallup poll in 2008, 80% of Americans felt satisfied with the way things are going in their lives. A researcher wonders if the percentage is different now.
1) parameter: μ 2) and 3) past claim: $H_0: \mu = 500$ hrs/life think: $H_A: \mu < 500$ hrs/life 4) Test: left tailed Test	1) parameter: σ 2) and 3) past $H_0: \sigma = 0.0008$ think $H_1: \sigma > 0.0008$ 4) Test: Right-Tailed Test	1) parameter: proportion 2) and 3) past: $H_0: p = 0.8$ think $H_A: p \neq 0.8$ 4) Two tailed test

Two possibly correct conclusions:

ANOTE: we are not saying "accept H_0" or not saying we "proved" is H_0 true

1) We decide there is evidence to support H_1	2) We decide there is NOT enough evidence to support H_1
"reject the null hypothesis" or "reject H_0"	"failed to reject the null hypothesis" ↳ "failed to reject H_0"

Ex: Historically, Jimbo's pizza had a mean (μ) delivery time of 48 minutes. After getting a new pizza oven, he takes a sample of 50 orders and finds that the mean delivery time is now 45 minutes, which makes Jimbo think that the mean delivery time has been reduced. Note is $\mu = 45$ statistically less or is it simply due to "sample variability"

State Jimbo's hypotheses in statistical notation:

$H_0: \mu = 48$ min/deliv.

$H_A: \mu < 48$ min/deliv.

State the conclusion if the **null is rejected:**

"There is enough statistical evidence to support that the true (pop. para.) mean pizza delivery time is less than 48 min/deliv."

State the conclusion if the **null is *not* rejected:**

"There is not enough statistical evidence to support that the true mean delivery time is less than 48 min/deliv."

NOT SAYING "we proved that mean is 48."

Four possible outcomes (2 errors)

Example: In a court case
H_0: the defendant is innocent
H_1: the defendant is guilty

		Truth about the Population (**Reality**): H_0 is true (innocent)	Truth about the Population (**Reality**): H_0 is false (guilty)
Decision Based On Sample **(Our Conclusion)**	Fail to Reject H_0 "keep"	Conclude innocent when innocent. Good!	Conclude innocent when guilty. Bad Type II Error
	Reject H_0	Conclude guilty when innocent. Bad Type I Error	Conclude guilty when guilty. Good

Worse type of Error

***NOTE:** The defendant is NEVER declared INNOCENT!!

Type I and Type II Errors

Type I error: The mistake of rejecting the null hypothesis when it is actually true.
The symbol α (alpha) is used to represent the probability of such an error.

Type II error: The mistake of failing to reject the null hypothesis when it is actually false.
The symbol β (beta) is used to represent the probability of such an error

Ex: On average (μ), it used to take 30 minutes to find parking, but we think we have sufficient evidence to say that the time has decreased. But, in fact, the true parking time is still 30 minutes. What kind of error did we make?

$H_0: \mu = 30$ min

$H_A: \mu < 30$ min

conclude statistically: reject H_0 (ie time decreased support H_A)

in reality: H_0 is true

↳ this is Type I Error

Note: The majority is a number or percentage equaling more than 50% of a total. To test a claim about a majority, your null hypothesis will be $p =$ 0.5. So $H_A: p > 0.5$. "Right Tailed Test"

<u>Ex</u>: A Gallup survey reports that 57% of 504 randomly selected gun owners support stricter gun laws. Test the claim that a majority of gun owners favor stricter gun laws. Write out the hypotheses for this example. What would a Type II error be in this scenario?

Hypotheses

$H_0: p = 0.5$

$H_A: p > 0.5$

when testing majority

	H_0 T	H_0 F
FTR H_0		II
R H_0	I	

• H_0 F: it is not true that the proportion of gun owners support stricter gun laws is 0.5.

• FTR H_0: FTR that proportion of gun owners support stricter gun laws is 0.5

Type II Error "We conclude that there's no majority of gun owners that supp. str. gun laws when in fact there is a majority support!"

<u>Ex</u>: Your company markets a computerized device to test a patient's mean resting heart rate. Based on the sample results, the device determines whether there is significant evidence that the patient's mean resting heart rate is greater than 100 beats per minute. If so, your company recommends that the person seeks medical attention.

a. State appropriate null and alternative hypotheses in this setting.

H_0:

H_1:

b. Which error is worse for your company?

	H_0 is true	H_0 is false
Fail to Reject H_0		**Type II Error**
Reject H_0	**Type I Error**	

Type II Error → Seek Med Attention? YES NO
Did They Need It? YES NO

Type I Error → Seek Med Attention? YES NO
Did They Need It? YES NO

We will NOT know 100% if our conclusion of our Hypothesis Test is ___________, but we can assign ______________________________ to making Type I and Type II Errors when we complete a hypothesis test.

Level of Significance	The probability of making a Type I Error. In other words, we take a sample that makes H_0 look WRONG when it's actually TRUE.

Note: As you ________________ the probability of one type of error, then the probability of the other type ________________.

Choosing a Significance Level

Typically the significance level, α is given to be greater than ______________ and less than ______________.

When a Type I error is…