HYPOTHESIS TEST CLASS NOTES

Hypothesis Test: Procedure that allows us to ask a question about an unknown population parameter
Uses sample data to draw a conclusion about the unknown population parameter.
A multi-step process is needed to set up and perform the hypothesis test and draw a conclusion about the outcome.

	Overview	Specific Steps to complete Hypothesis Test
Step 1:	Planning the test: • Formulate questions as a pair of hypotheses • Set criteria for how to draw a conclusion from the data	•Write null and alternate hypotheses •Determine significance level α
Step 2:	Select sample(s) and collect data.	Examine the data Determine how to perform test (distribution and type of test)
Step 3:	Analyze sample data. using calculator's built in statistics tests to do calculations	•Find "test statistic" indicating how far away our sample statistic is from the null hypothesis • Find "p value" indicating how likely or unusual our sample would be under the null hypothesis.
Step 4	Decide which hypothesis is more appropriate based on the analysis of the data	Decide to "reject null hypothesis" or to "not reject null hypothesis" based on p value and significance level
Step 5	Interpret the decision in the context of the problem.	Interpret the decision in the context of the problem.

Step 1: Set up hypotheses that ask a question about the population by setting up two opposing statements about the possible value of the parameters.

The two opposing statements are called the "Null Hypothesis" and the "Alternate Hypothesis"
In setting up a hypothesis test, statisticians must very carefully design the hypothesis test so that:

H_0: Null hypothesis: This is the assumption about the population parameter that will be assumed of believed unless it can be shown to be wrong beyond a reasonable doubt

H_A : Alternate hypothesis: This is the claim about the population parameter that must be shown correct "beyond a reasonable doubt" in order for us to believe that it is true.

- the hypotheses always refer to the population parameter p or μ (*never the sample statistic* $\bar{X}$ *or p′*)
- the outcome that needs to be "proved" is the alternate hypothesis
- the alternate hypothesis always contains a strict inequality: $\neq$ or $<$ or $>$
- the null hypothesis contains equality of some kind: $=$ or $\geq$ or $\leq$

Describe the parameter, p or μ , being tested in a sentence: p = *description* or μ = *description*

Write both null hypothesis H_0 and alternate hypothesis H_A using mathematical symbols

- $>$ or $<$ or $\neq$ in the words of the problem gives **H_A** , the ALTERNATE hypothesis and the opposite of it is the null hypothesis
- $\leq$ or $\geq$ or $=$ in the words of the problem gives the **H_0 ,** the NULL hypothesis and the opposite of it is the alternative hypothesis
- In the null hypothesis you can use $=$ instead of $\leq$ or $\geq$

Math 10 RULE FOR FROMULATING HYPOTHESES

Null hypothesis H_0 must contain equality *of some type*: = ≤ or ≥

Alternate hypothesis H_A must contain a pure inequality. ≠ > or <

H_0 and H_A are usually the opposite of each other

But its also acceptable to use strict equality = in H_0 no matter what inequality is in H_A

Example A: A hospital is testing a new surgery for a type of knee injury. Many patients with knee injuries recover with non-surgical treatment, and surgery has risks. The surgery review board has decided that the hospital can perform this surgery as a clinical trial. They discuss and study the medical considerations and they decide that they will approve this type of surgery for future use if the clinical trial shows that the surgery would cure more than 60% of all such knee injuries

The hypothesis test should be set up so that the surgery must be proven effective.

Null hypothesis: H_O: A new surgery is as effective as non-surgical treatment

Alternate hypothesis: H_a: A new surgery is more effective than non-surgical treatment

We need to write this mathematically.

p = ____

H_0: _________ H_A: __________

Examples 1 – 9: Some of the following examples will be done in class for setting up hypotheses. Those not done in class should be done for practice.

Example 1: FDA guidelines require that to be considered "gluten-free", a serving of food must contain less than 20 parts per million of gluten . A food manufacturer should be able to document, if asked, that its food satisfies these quidelines in order to put the words " gluten-free" on its label. Several batches of food are tested to determine if the average amount of gluten per serving meets these guidelines.

μ = ____

H_0: _________ H_A: __________

Example 2: https://online.drexel.edu/online-degrees/**nursing**-degrees/**nursing-salary**.../index.aspx Nov 12. 2016

Its been cited that the average starting salary for registered nurses in the US is $66,000. A nurses professional organization in a particular city wants to conduct a study to determine if the average starting salary for nurses in that city is different from the national average.

_____ = ____

____

H_0: _________ H_A: __________

Example 3. A soda bottler wants to determine whether the 12 ounce soda cans filled at their plant are underfilled, containing less than 12 ounces, on average.

_____ = ____

____

H_0: _________ H_A: __________

Example 4. Circuit Fitness advertises a 30 minute workout rotating clients exercising though fitness stations. Some clients complain that they want longer workouts; others prefer a 30 minute workout. A survey is done to determine if the average desired workout time is longer than the current 30 minutes.

_____ = ____

____

H_0: _________ H_A: __________

Example 5. http://www.nbcnews.com/business/personal-finance/three-10-americans-have-no-savings-says-survey-n379946 Jun 23 2015
It has been estimated that nationally, 30% of US residents have no savings.
The mayor wants to determine if the percent of residents of his city who have no savings is different from the national percent.

_____ = ____

H_0: _________ H_A: __________

Example 6. The Center for Disease Control reports that only 14% of California adults smoke. A study is conducted to determine if the percent of De Anza college students who smoke is higher than that.

_____ = ____

H_0: _________ H_A: __________

Example 7. Uber rides from De Anza College to SJ Airport cost $25 on average.

_____ = ____

H_0: _________ H_A: __________

Example 8. Students take an average of at least 10.5 credits per quarter.

_____ = ____

H_0: _________ H_A: __________

Example 9. At most half of all library customers borrow ebooks.

_____ = ____

H_0: _________ H_A: __________

Hypothesis Tests: Correct Decisions and Errors in Decisions

In a hypothesis test, we decide about the hypotheses based on the strength of evidence in sample data. The sample data may lead us to make a correct decision or sometimes to make a wrong decision.

A hypothesis test can be compared to a trial where we assume a person is innocent (null hypothesis) unless "proven" guilty beyond a reasonable doubt (alternate hypothesis) based on the strength of evidence (sample data). A jury's decision based on evidence may be correct, or incorrect for the person's innocence or guilt.

Null Hypothesis: Person on Trial is Innocent		**Alternate Hypothesis: Person on Trial is Not Innocent**	
Person is innocent AND jury decides innocent Good decision	**Person is innocent** **BUT jury decides guilty** **Wrong Decision :** Innocent person goes to jail for a crime he did not do **TYPE I ERROR: Deciding in favor of the Alternate Hypothesis when in reality the Null Hypothesis is true**	**Person is NOT innocent** **BUT jury decides innocent** **Wrong Decision** Guilty person does not go to jail for a crime he did **TYPE II ERROR: Deciding in favor of the Null Hypothesis when in reality Alternate Hypothesis is true**	Person is NOT innocent AND jury decides guilty Good decision

α: the probability of making a Type I error is called the SIGNIFICANCE LEVEL α (alpha)

We want α to be small: usually want 5% or 3% or 2% or 1%, but could be even smaller

α is the risk we are willing to take of making a wrong decision in the form of a Type I error.

β: the probability of making a Type II error. We want this to be small also.

$1-\beta$ is called the POWER of the test. *It's the probability of making good decision if HA is true: We want this probability to be big. Statisticians consider this when planning sample size in designing the test.*

Example A: A hospital is testing a new surgery for a type of knee injury. They will approve this type of surgery for future use if the clinical trial shows that the surgery would cure more than 60% of all such knee injuries

p = the true (population) proportion of all knee injuries that would be cured by this surgery

H_0: Null hypothesis: $p \leq 0.60$ $\quad$ H_A: Alternate hypothesis: $p > 0.60$

A Type I error would be to decide that the surgery cures more than 60% of injuries

when in reality the surgery cures at most 60% (60% or less).

A consequence of a Type I error would be that the surgery is approved and patients might get a surgery that is not effective.

A Type II error would be to decide that the surgery cures at most 60% of injuries

when in reality the surgery cures more than 60% of injuries

A consequence of a Type II error would be that the we think surgery is not effective so it is not approved and patients can't have a surgery that is effective at curing their injuries.

<u>Type I Error:</u> *Rejecting the null hypothesis Ho when in reality Ho is true*

- **<u>concluding</u> (based on sample data) in favor of the alternate hypothesis**
- **<u>when in reality</u> the null hypothesis is true**

<u>Type II Error:</u> *Failing to reject the null hypothesis Ho when in reality Ho is false*

- **<u>concluding</u> (based on sample data) in favor of the null hypothesis**
- **<u>when in reality</u> the alternate hypothesis is true**

Guidelines : Interpretation has 2 parts:

- **clearly state the conclusion ("we conclude, or decide that , ______")**
- **clearly state what is true in reality ("when in reality ________")**

- State each part of the interpretation in context of the problem.
- Each part of the interpretation should clearly and accurately state a hypothesis (H_0 or H_A) in words.
- The decision and reality should NOT agree – otherwise it's a good decision and not an error
- **Be extremely careful to reflect both equalities and inequalities accurately in your sentences**

<u>Example 1:</u> FDA guidelines require that to be considered "gluten-free", a serving of food must contain less than 20 parts per million of gluten .

μ = the true average amount of gluten per serving

H_0: $\mu \geq 20$ parts per million of gluten H_A: $\mu < 20$ parts per million of gluten

A Type I Error is concluding that____

____

when in reality____

____

Consequence of a Type I error?

A Type II Error is concluding that____

____

when in reality____

____

Consequence of a Type II error?

<u>Example 2</u>: A nurses professional organization conduces a study in a certain city to determine if the average starting salary for registered nurses in that city is different from the US average of $66,000.

μ = the true average starting salary for all starting registered nurses in this city

H_0: $\mu = \$66000$ H_A: $\mu \neq \$66000$

A Type I Error is concluding that____

____

when in reality____

____

A Type II Error is concluding that____

____

when in reality____

____

PRACTICE: *Write the Type I and Type II errors for the rest of the Examples 1-9 from pages 2&3.*

RARE EVENTS

The null hypothesis is an assumption or a theory about a property of a population; it is not a known fact. We select a sample. The sample is real data.

- If our sample is extremely unlikely to occur based on our assumption, then we would conclude that the assumption is not correct.
- If our sample data is reasonably likely to occur based on our assumption, then this would not give us any reason to doubt the assumption.

Example A: A hospital is testing a new surgery for a type of knee injury The surgery review board has decided that they will approve this surgery for future use if a clinical trial shows that the true population cure rate for this surgery would be more than 60%. Otherwise they will not approve it.

Population parameter: p = true population cure rate for this surgery
Random Variable: P ' = cure rate for a sample of patients having this surgery

Ho: $p \le .60$ Ha: $p > .60$

Suppose the new surgery is tested on 200 patients.

- Suppose the sample proportion of people who are cured is p'=0.90, a 90% cure rate.

  Would this strongly support Ha or would we believe Ho might be true?

  ____

  ____

- Suppose the sample proportion of people who are cured is p'=0.46, a 46% cure rate.

  Would this strongly support Ha or would we believe Ho might be true?

  ____

  ____

- Suppose the sample proportion of people who are cured is p'=0.605, a 60.5% cure rate.

  Would this strongly support Ha or would we believe Ho might be true?

  ____

  ____

Where do we draw the line between "far" and "close"? What if p' = 0.62 or 0.65 or 0.70?

2 calculations to help us decide this:

- **We calculate a test statistic (a z-score or t-score in chapter 9) that tells us if our sample is close to or far from the null hypothesis**
- **We findthe probability "p value" of getting a sample that "looks like ours" if the null hypothesis is true.**
  - If that probability is small, then the sample is not consistent with the null hypothesis. The sample data seem to strongly contradict the null hypothesis. It gives strong evidence to support the alternate hypothesis, so we reject the null hypothesis.
  - If that probability is large, then the sample is reasonably likely to occur if the null hypothesis is true; the sample is consistent with the null hypothesis. So we don't have strong enough evidence in the sample data to decide to reject the null hypothesis

<u>Example A</u>: A hospital is testing a new surgery for a type of knee injury The surgery review board has decided that they will approve this surgery for future use if a clinical trial shows that the true population cure rate for this surgery would be more than 60%. Otherwise they will not approve it.

Population parameter: p = true population cure rate for this surgery
Random Variable: P ' = cure rate for a sample of patients having this surgery

Ho: $p \le .60$ **Ha: $p > .60$**

Suppose that in a sample of 200 patients having this surgery, 130 of them are cured:
p' = 130/200 = 0.65
Is p' = .65 close to or far from the null hypothesis that p = .60 ?

Find the test statistic that tells us how far our sample is from the null hypothesis.
Find the probability of getting a sample that "looks like ours" if the null hypothesis is true.

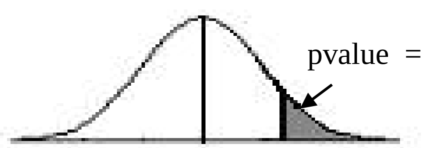
Criteria for "what is a small probability?"

The significance level α is our criteria for "what is a small probability?"
It is the risk we are willing to accept of making a Type I error if we reject the null hypothesis.

What risk of making a Type I error (allowing surgery that is not effective) are we willing to accept for this situation?_____________

Calculate the p value : probability of getting a sample at least as far from the null hypothesis as our sample is. The inequality in the alternate hypothesis Ha tells us how to calculate the probability (direction of shading). Since our alternate hypothesis says >, we use the right tail of the distribution

1-PropZTest **p_0 : 0.60** **x: 130** **n: 200** **prop: $\ne p_0$ $<p_0$ $\boxed{>p_0}$** **CALCULATE**	**1-PropZTest** **Prop > .6** **z = 1.443** **p = .0745** **$\hat{p}$ = .65**	**z** = is the test statistic **p** = is the pvalue $\hat{p}$ is sample proportion. Calculator uses $\hat{p}$ rather than p'



pvalue =

pvalue = $P(p' \ge .65$ if $p=.60) = P(p' \ge .65 \mid p=.60)$
pvalue is calclulated by 1Prop Z Test as

$$\text{normalcdf}\left(0.65, 10\text{^}99, 0.60, \sqrt{\frac{0.60*0.40}{200}}\right)$$

Compare p value to signficiance level α Is the p value smaller than the significance level? _______

<u>Decision</u>: ______________________________________

<u>Conclusion</u>: ____

____

____

DECISION RULE: If p value < α , REJECT Ho
If p value $\ge$ α , DO NOT REJECT Ho

CONCLUSION: At a (<u>*state α as %*</u>) level of significance, the sample data <u>*DO / DO NOT*</u> provide strong enough evidence to conclude that (<u>*state in words what the alternate hypothesis says in context of the problem*</u>)

Take notes in class on a Chapter 9 Solution Sheet as we do *SOME* of these examples in class.

EXAMPLE B: <u>Hypothesis Test of Population Mean μ</u>
when Population Standard Deviation <u>σ is KNOWN</u>

A truck manufacturer sells delivery vans to package delivery services.
From its records, management knows that the average fuel efficiency for these vans is 19.8 miles per gallon (mpg) with standard deviation 2.9 miles per gallon.

Engineers redesigned an engine part to increase average fuel efficiency.
The engineering team believes that the known population standard deviation for fuel efficiency of 2.9 miles per gallon will continue to apply to the new engine redesign.

At a 5% level of significance, perform a hypothesis test to determine if the redesigned engine part is effective in increasing the average fuel efficiency.

A sample of 36 vans with the redesigned part has average fuel efficiency of 21.2 mpg.

20.1	15.7	24.5	13.8	20.7	24.2	22.4	22.2	24.5	18.4	21.1	19.4
24.6	24.8	21.5	19.9	16.6	21.7	17.7	20.5	21.1	22.2	19.9	23.1
18.1	25.1	26.9	22.2	24.7	20.1	20.5	23.4	18.3	21.2	18.9	23.5

EXAMPLE C: <u>Hypothesis Test of Population Mean μ</u>
when Population Standard Deviation <u>σ is NOT KNOWN</u>

At Dina's Dress shop, some of her customers complain that fashion designers assume that all women are tall when they design clothes. Dina read that the designers whose clothes she carries tend to design for women who are 5 feet 6 inches (66 inches) tall. Because of her customers' comments, Dina wants to determine whether the average height of her customers is shorter.

A sample of 25 customers has average height of 64 inches with standard deviation 2.5 inches. *Assume the population of individual customer heights is normally distributed.*

Perform a hypothesis test to determine if the average height of all Dina's customers is less than 66 inches. Use a 2% level of significance

61	63	65	62	63	61	64	63	61	63	68	66	63
60	65	66	66	68	67	66	66	60	65	61	67	

NOTE: *Then find a confidence interval to estimate the true average height of all Dina's customers, so that she can decide which designers to buy dresses from for the shop.*

EXAMPLE D: <u>Hypothesis Test of Population Proportion</u>

A commercial cookie bakery needs to produce "perfect" cookies. Defective cookies (broken or are misshapen) can not be sold at full price, but only at a reduced price or for use in ice cream flavoring. The bakery's past experience shows a 6% defective rate. They recently purchased some new equipment and want to determine if the defective rate has changed. A random sample of 250 cookies contained 21 "defective" cookies. Perform a hypothesis test to determine if the percent of cookies that are defective has changed when using the new equipment. Use 4% significance level.

EXAMPLE E: The library board of directors are not willing to increase funds for ebook borrowing because they believe that at most half of all library customers borrow ebooks, and. A random sample of 150 customers shows that 80 of them borrowed ebooks. At a 5% level of significance, perform the hypothesis test.

EXAMPLE F: A certain medication is supposed to contain a dosage of 250 mg per pill. A lab is quality testing the drug to determine if it contains the correct dosage, on average. A sample of 50 pills has average dosage 232 mg with standard deviation 25 mg. At a 2% level of significance, perform the hypothesis test.
(Assume the underlying population of amount of drug in individual pills is approximately normally distributed.)

EXAMPLE G: The management of SuperSaver Grocery Outlet is considering adopting Apple Pay. It would require investment so they will adopt it only if they strongly believe that more than 10% of their customers would use Apple Pay if available. In a survey 300 customers, 45 customers (15%) state they would use Apple Pay. Perform a hypothesis test to determine whether they should adopt Apple Pay.

EXAMPLE H: Does the label "GLUTEN FREE" guarantee that a food has absolutely NO gluten? FDA guidelines state a food must contain less than 20 parts per million of gluten to qualify as gluten free. (So some foods containing some gluten may legally qualify as gluten free.) A brand of crackers claiming to be gluten free was tested ; for a sample of 7 batches of these crackers, the gluten levels in parts per million were: **9 20 14 22 14 12 15**

Perform a hypothesis test to determine if the average gluten content is less than 20 parts per million so that the manufacturer can support its claim that the crackers can be represented as gluten free. Use a 3% level of significance.
(Assume the underlying population of gluten content in individual batches is approximately normally distributed.)

EXAMPLE I: A graduate student believes that less than 15% of graduate students live alone because they are not able to afford rent on their own. Suppose that a survey of 500 graduate students shows that 17.4% live alone. Conduct a hypothesis test.
http://www.multifamilyexecutive.com/property-management/demographics/gauging-student-living-preferences_o

EXAMPLE J: A college instructor assigns online homework using MyWebHW.edu. MyWebHW.edu records for all students using the system shows that their standard homework assignments take an average of 79 minutes with a population standard deviation of 34 minutes. The instructor suspects that the time needed for these assignments may be different for students at her college than for the population of all students using MyWebHW.edu. She collects data for a random sample of 40 students at her college. The sample average completion time is 91 minutes. Conduct a hypothesis test using a 2% level of significance.

CHAPTER 9: SUMMARY OF SKILLS for Hypothesis Tests

Math 10 RULE FOR HYPOTHESES: Hypothesis must contain symbol μ or p (*never* $\bar{X}$ *or* p′)
Null hypothesis H_0 must contain equality *of some type*: **=** $\leq$ **or** $\geq$
Alternate hypothesis H_A must contain a pure inequality. $\neq$ > **or** <
Ho and Ha are usually opposite of each other, but Ho can also use "=", even when Ha uses "<" or ">".

Test of mean μ when σ is known	**ZTest**	**Parameter is μ** **Random variable is** $\bar{X}$	**Distribution is Normal** $N(\mu, \sigma/\sqrt{n})$
Test of mean μ when σ is not known	**TTest**	**Parameter is μ** **Random variable is** $\bar{X}$	**Distribution is t with df = n–1**
Test of proportion p	**1PropZTest**	**Parameter is p** **Random variable is P′** ***or*** $\hat{P}$	**Distribution is Normal** $N(p, \sqrt{\frac{pq}{n}})$

Calculator output: check that the alternate hypothesis at top of output screen is correct
z = or **t**= gives Test Statistic **p=** gives pvalue

Graph: Put the number from the null hypothesis in the middle

♦ For a one tailed test mark the sample statistic on the horizontal axis.
- If H_A is < : shade to the left from the sample statistic
- If H_A is > : shade to the right from the sample statistic

♦ For a two tailed test where H_A is $\neq$
- Mark the sample statistic on the horizontal axis.
- Also mark the value that is the same distance from the center on the other side.
- Shade out to both sides.

Intepreting the pvalue:

If the null hypothesis is true, then there is a probability of (*fill in the pvalue***) of getting a sample average** $\bar{X}$ **of** (*state value of* $\bar{X}$) (*pick one: or less, or more, or more extreme*).

If the null hypothesis is true, then there is a probability of (*fill in the pvalue***) of getting a sample proportion** p′ **of** (*state value of* p′) (*pick one: or less, or more, or more extreme*).

To pick one choice : use "**or less**" *if H_A has <* *OR* *use* "**or more**" *if H_A has >*
OR *use* "**or further away from Ho**" or **"more extreme"** *if H_A has $\neq$*

DECISION RULE: If p value < α , REJECT H_0 ; If p value $\geq$ α , DO NOT REJECT H_0

CONCLUSION:

If you reject H_0: At a (*state α as %*) level of significance, the sample data provide sufficient evidence to conclude that (*state alternate hypothesis Ha in words in context of the problem*).

If you do not reject H_0: At a (*state α as %*) level of significance, the sample data do NOT provide sufficient evidence to conclude that (*state alternate hypothesis H_A in words in context of the problem*). Therefore we continue to assume that (*state null hypothesis Ho in words in context of the problem*).

If you reject Ho, then the result is "statistically significant", or just "significant"
If you do not reject Ho, then the result is "not statistically significant"or "not significant"

Type I and Type II Error: State interpretations in the context of the problem

TYPE I ERROR: **concluding** based on sample data **in favor of the alternate hypothesis when in reality the null hypothesis is true**

TYPE II ERROR: **concluding** based on sample data **in favor of the null hypothesis when in reality the alternate hypothesis is true**