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.

Performing a Hypothesis Test is a multi-step process.
Our calculator's built in statistics tests will help us with the calculations in Step 3.

Step 1: Planning the test:
- **Formulate questions as a pair of hypotheses** Write null and alternate hypotheses
- **Set criteria for how to draw a conclusion from the data** Determine significance level α

Step 2: Select sample(s) and collect data.
- Examine the data
- Determine how to perform test (which distribution and type of test to use)

Step 3: Analyze sample data to find:
- "test statistic" indicating how far away your sample is from the null hypothesis
- "p value" indicating how likely or unusual your sample would be under the null hypothesis.

Step 4: Decide which hypothesis is more appropriate based on the data
"reject null hypothesis" or "not reject null hypothesis" based on p value

Step 5: 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 opposite statements about the possible value of the parameters.

Ho: Null hypothesis: The assumption about the population parameter that will be believed unless it can be shown to be wrong beyond a reasonable doubt

Ha: Alternate hypothesis: The claim about the population parameter that must be shown correct "beyond a reasonable doubt" to believe that it is true.

Statisticians design the hypothesis test so that the outcome that needs to be proved is the alternate hypothesis.

Read problem *carefully*. Read it more than once if you need to.
Be very careful to interpret inequality language correctly.

- **Is the parameter a MEAN or a PROPORTION?**
- **Hypotheses always refer to the population parameter** p or μ
  Hypotheses *never the refer to the sample statistic* $\bar{X}$ *or p′*
- **Describe the parameter, p or μ, being tested in a sentence**
  p = *description* or μ = *description*
- **Write both null hypothesis Ho and alternate hypothesis Ha using mathematical symbols**
  - $>$ or $<$ or $\neq$ in the words gives the ALTERNATE hypothesis
    and the opposite of it is the null hypothesis
  - $\leq$ or $\geq$ or $=$ in the words gives the NULL hypothesis
    and the opposite of it is the alternative hypothesis

Math 10 RULE FOR HYPOTHESES

Null hypothesis Ho must contain equality *of some type*: $=$ $\leq$ or $\geq$
Alternate hypothesis Ha must contain a pure inequality. $\neq$ $>$ or $<$

Math 10 RULE FOR HYPOTHESES

Null hypothesis Ho must contain equality *of some type*: $=$ $\leq$ or $\geq$

Alternate hypothesis Ha must contain a pure inequality. $\neq$ $>$ or $<$

Ho and Ha are opposite of each other

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_o: ____________________

H_a: ____________________

Example B: 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_o: Null hypothesis: ______________________________________.

H_a: Alternate hypothesis: ______________________________________

Some of the following examples will be used in class. Link to answers is posted on instructor's website.

1. Engineering students at a college compare the price of their calculus book at their bookstore to the price at other college bookstores. The calculus book costs $150 at their bookstore. They will collect data about the price of this calculus book from 30 other bookstores nationwide and will use the sample data to decide whether the average price of this book at all other colleges in the nation is different from the price of $150 at their bookstore.

_____ = ____

____

H_0: ________ H_A: ________

____

2. 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: __________

3. 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 determin if the average desired workout time is longer than the current 30 minutes.

_____ = ____

____

H_0: _________ H_A: __________

4. 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. A city wants to determine if the percent of city residents with no savings is different from the national percent.

_____ = ____

____

H_0: _________ H_A: __________

5. 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: __________

6. The average price per cup for coffee at the airport is at least $2.50

_____ = ____

____

H_0: _________ H_A: __________

7. The average number of credit units that students take is 10.5 credits per quarter.

_____ = ____

____

H_0: _________ H_A: __________

8. 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.

It is similar 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).

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

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 will let the hospital perform this surgery in a clinical trial. After studying the medical considerations, 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

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

H_o: Null hypothesis: $p \leq 0.60$ 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.

***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-β 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.*

Type I Error: *Rejecting the null hypothesis when in reality the null hypothesis is true*

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

Type II Error: *Failing to reject the null hypothesis when in reality the null hypothesis is false*

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

Guidelines:

- The interpretation has 2 parts:
  - state the conclusion ("we conclude, or decide, ______")
  - state what is true in reality ("when in reality _______")
- Each part of the interpretation should be stated in context of the problem.
- The decision and reality should NOT agree – otherwise it's a good decision and not an error
- Each part of the interpretation should clearly and accurately state a hypothesis (Ho or Ha) in words.
- Be extremely careful to reflect equalities and inequalities accurately in your sentences

Example B: 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.

μ = the true average amount of gluten per serving

Null hypothesis: H_o: $\mu \geq 20$ parts per million of gluten

Alternate hypothesis: H_a: $\mu < 20$ parts per million of gluten

A Type I Error is concluding that____

when in reality____

A Type II Error is concluding that____

when in reality____

In this example, what would be a consequence of a Type I error?

In this example, what would be a consequence of a Type I error?

PRACTICE: Use Examples 1 through 8 on pages 2 & 3 to interpret errors.

Example #1:

A Type I Error is concluding that____

when in reality____

A Type II Error is concluding that____

when in reality____

PRACTICE: We will use some of Examples 2 through 8 on page 3 to interpret errors.

Example ___:

A Type I Error is concluding that____

____

when in reality____

____

A Type II Error is concluding that____

____

when in reality____

____

Example ___:

A Type I Error is concluding that____

____

when in reality____

____

A Type II Error is concluding that____

____

when in reality____

____

Example ___:

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 problems on page 3.

RARE EVENTS

We make an assumption about a property of a population (null hypothesis).
We select a sample.

- If our sample data has properties that are extremely unlikely to occur based on our assumption, then we would conclude that the assumption is not correct.
- If our sample data has properties that are 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

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 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.

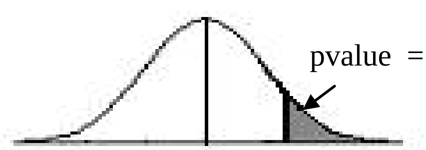
Criteria for "what is a small probability?"

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: $\neq p_0$ $<p_0$ $>p_0$** **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'≥.65 if p=.60) = P(p'≥.65 | p=.60)
pvalue is calclulated by 1Prop Z Test as

$$\text{normalcdf}\left(0.65, 10\text{^}99, 0.60, \sqrt{\frac{0.60*0.40}{300}}\right)$$

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

Decision: ______________________________________

Conclusion: ____
____
____

DECISION RULE: If pvalue < α , REJECT Ho
If pvalue ≥ α , DO NOT REJECT Ho

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

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

EXAMPLE C: Hypothesis Test of Population Mean μ when Population Standard Deviation σ is KNOWN

A truck manufacturer sells delivery vans to package delivery services.
From it 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 this redesigned part has an 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 D: Hypothesis Test of Population Mean μ when Population Standard Deviation σ is NOT KNOWN

OptiFiber Network advertises that their internet service has an average connection speed of 200 mb per second.
Their competitor CableNet claims that OptiFiber's average connection speed is slower than that.
If CableNet can prove that, they will file a lawsuit against OptiFiber for misleading advertising; otherwise their lawyers tell them they will not be able to win the lawsuit.

CableNet selects a sample of 16 different internet connections on OptiFiber's network and measures the speed of each connection in the sample.
For the sample, the average speed is 180 mb per second with standard deviation is 64 mb per second.

Does CableNet have sufficient evidence in its sample data to support its claim that OptiFiber's average connection speed is less than 200 mb per second? Use a 5% significance level.

Sample data: internet connection speeds in mb per second n = 16							
128	161	297	100	259	226	124	141
237	133	111	111	115	151	267	219

(Assume the underlying population of individual internet connections speeds is normally distributed.)

EXAMPLE E: Hypothesis Test of Population Mean μ when Population Standard Deviation σ is NOT KNOWN

At Dina's Dress shop, 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 an average height of 64 inches with standard deviation 2.7 inches.
(Assume that the population of individual customer heights is approximately 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

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

EXAMPLE F: **<u>Hypothesis Test of Population Proportion p</u>**

A city zoning board is studying whether the parking regulations for residential neighborhoods are still appropriate. Their current regulations assume that 27% of households are "car-free". They want to know if the percent of households that are "car-free" has changed, so they decide to conduct a survey and perform a hypothesis test using a 3% level of significance. The survey of 1200 households shows that 372 are "car-free".

EXAMPLE G: The library board of directors believes that at most half of all library customers borrow ebooks, and are not willing to increase funds for ebook borrowing. A random sample of 100 customers shows that 56 of them borrowed ebooks. At a 5% level of significance, perform the hypothesis test.

EXAMPLE H: 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 an average dosage of 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 I: The management of SuperSaver Grocery Outlet is consider adopting Apple Pay. It would require some investment; they will do this only if they strongly believe that more than 10% of their customers would use Apple Pay if available. They 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 J: 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 K: *<u>http://www.multifamilyexecutive.com/property-management/demographics/gauging-student-living-preferences_o</u>*
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.

EXAMPLE L: A college instructor assigns online homework using MyWebHW.edu. MyWebHW.edu has records for all students using the system and knows 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:

Math 10 RULE FOR HYPOTHESES: Hypothesis must contain symbol μ or p (never $\bar{X}$ *or* p′)
Null hypothesis Ho must contain equality *of some type*: = ≤ **or** ≥
Alternate hypothesis Ha must contain a pure inequality. ≠ > **or** <
Ho and Ha are opposite of each other

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′**	**Distribution is Normal** $N(p, \sqrt{\frac{pq}{n}})$

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

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 Ha is < : shade to the left from the sample statistic
- If Ha is > : shade to the right from the sample statistic

♦ For a two tailed test where Ha is ≠
- 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 (*fill in the word* "**average**" *or* "**proportion**") **of** (*state value of* $\bar{X}$ *or* p′)
(*pick one choice : use* "**or less**" *if Ha has* < *OR use* "**or more**" *if Ha has* >
OR use "**or further away from Ho**" or **"more extreme"** *if Ha has* ≠

DECISION RULE: If p value < α , REJECT Ho ; If p value ≥ α , DO NOT REJECT Ho

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

If you reject Ho, then the result is "significant"

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

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

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

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