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.**

Steps to perform a Hypothesis Test

Step 1: Planning the test: Formulate questions as hypotheses
AND set criteria for how to draw a conclusion from the data

Step 2: *In real life: Gather Data:* Select random sample(s) and collect data.
In HW or exam problem: Examine the given data and information, *and* determine how to perform test (which distribution and type of test to use)

Step 3: Analyze sample data

Step 4: Decide what the data analysis shows

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.

In real world studies, statisticians design the hypothesis test so that the outcome that needs to be proved is the alternate hypothesis.

Example A: A new drug is being tested for safety.

The hypothesis test should be set up so that the drug must be proven safe in order for it to be used.

Ho: Null hypothesis: ____

Ha: Alternate hypothesis: ____

Example B: To be considered "fat-free", regulations require that a serving of salad dressing must contain less than ½ gram of fat. A salad dressing manufacturer must be able to show that its salad dressing satisfies these guidelines in order to put the words "fat-free" on its label. (*one "serving" is 2 tablespoons of salad dressing*)

The hypothesis test should be set up so that the manufacturer must prove that the salad dressing meets satisfies the regulations in order to be able to call it "fat-free"

μ = ____

Ho: Null hypothesis: ____.

Ha: Alternate hypothesis: ____

<u>Math 10 RULE FOR HYPOTHESES</u>

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

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

Example C: Suppose that a hospital is testing a new surgery for a certain type of knee injury.
The hospital's surgery review board needs to decide whether to approve this type of surgery.
In general , many patients with this type of knee injury recover on their own without treatment, and surgery may have risks. Therefore, the surgery review board has decided that the hospital can perform this surgery as a clinical trial.
After much discussion of the medical considerations involved in this surgery, they decide to approve this type of surgery for future use if the clinical trial shows that the new surgery would cure more than 60% of all such injuries; otherwise they will not approve it.

p = true population proportion of all people with this knee injury who would be cured if they had this surgery

Ho: $p \leq .60$
Ha: $p > .60$ one tailed test to the right (right tailed test)

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 want to determine how the price of their calculus book at their bookstore compares 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.

   _____ = ____

   Ho: _________ Ha: __________ ____________________________

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.

   _____ = ____

   Ho: _________ Ha: __________ ____________________________

3. Exercise Circuit, a fitness center, advertises a 30 minute workout that rotates clients exercising though all the fitness stations in 30 minutes. Clients who want longer workouts have complained, but other clients prefer the relatively quick 30 minute workout. The center conducts a survey to determine whether the average desired workout time for its clients is longer than the current 30 minute circuit.

   _____ = ____

   Ho: _________ Ha: __________ ____________________________

4. It has been estimated that nationally, 16% of US residents lack health insurance coverage. Suppose that a city wanted to determine whether the percent of city residents without health insurance is different from the national percent.

   _____ = ____

   Ho: _________ Ha: __________ ____________________________

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.

   _____ = ____

   Ho: _________ Ha: __________ ____________________________

6. The average price of a cup of coffee in the airport is at least $2.50

   _____ = ____

   Ho: _________ Ha: __________ ____________________________

7. The average amount of time that students do statistics homework each night is 2 hours.

_____ = ____

Ho: _________ Ha: __________ ____________________________

8. At most half of all customers at Ace Auto Repair drive foreign cars.

_____ = ____

Ho: _________ Ha: __________ ____________________________

PROCEDURE FOR FORMULATING HYPOTHESES

- **Is the problem is asking about a MEAN or a PROPORTION?**
- **Hypothesis always refers to the population parameter** (*never the sample statistic*)
- **Describe the parameter,** p or μ , being tested in a sentence that starts p =*description* or μ =*description*
- **Read problem *carefully***: Look for words indicating

EQUALITY OF SOME TYPE: = ≤ ≥ (indicates Ho: NULL Hypothesis)

STRICT INEQUALITY: ≠ > < (indicates Ha: ALTERNATE Hypothesis)

Ho	Ha	
≤ or =	>	One tailed test to the right
≥ or =	<	One tailed test to the left
=	≠	Two tailed test

Hypothesis Tests: Correct Decisions and Errors in Decisions

In a hypothesis test, we decide about the hypotheses, based on the strength of the evidence in the sample data.
The sample data may lead us to make a correct decision or sometimes make a wrong decision.
We want to make decisions in a way that minimizes the chance of making a wrong decision.

IN REALITY: / **DECISION:**	Null Hypothesis is TRUE	Null Hypothesis is FALSE Alternate Hypothesis is true
Based on sample data we DECIDE NOT TO REJECT null hypothesis	**Correct Decision**	**Type II Error: Wrong Decision** Probability is β
Based on sample data we DECIDE TO REJECT null hypothesis	**Type I Error: Wrong Decision** Acceptable probability (risk) is α α is called **significance level**	**Correct Decision** Probability is $\mathbf{1-\beta}$ $\mathbf{1-\beta}$ is called the **power**

Type I Error:

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

Rejecting the null hypothesis when in reality the null hypothesis is true

The probability you are willing to risk of making a Type I error is denoted by α (Greek letter alpha).
α is called the (pre-determined) **significance level** of the hypothesis test**.**
The significance level α should be small: a low risk of incorrectly rejecting the null hypothesis if it is really true

Type II Error:

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

Failing to reject the null hypothesis when in reality the null hypothesis is false

The probability of Type II error is denoted by β (Greek letter beta).
1 − β is called the **power** of the hypothesis test.
The power of a hypothesis test should be large: large probability of correctly rejecting a false null hypothesis

Example C Revisited: (see page 2 for the statement of this problem)

p = true population proportion of all people with this knee injury who would be cured if they had this knee surgery

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

A Type I error would be to conclude that this surgery cures more than 60% of all knee injuries of this type, when in reality it cures 60% or fewer of all such injuries.

A Type II error would be to conclude that this surgery cures at most 60% of all knee injuries of this type, when in reality it cures more than 60%of all such injuries.

Some of the examples on page 2 will be used in class to interpret errors; Catalyst website has a link to all answers.

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____

Example #____:

A Type I Error is concluding that____

when in reality____

A Type II Error is concluding that____

when in reality____

Rare Events

We have an assumption about a property of a population.
We select a sample.

- Suppose our sample data has properties that are extremely unlikely to occur if the population actually has the properties we assumed. We would then conclude that the assumption is not correct.
- Suppose our sample data has properties that are reasonably likely to occur if the population actually has the properties we assumed. This would not give us any reason to doubt the assumption.

Example C Revisited: Suppose that a hospital is testing a new surgery for a certain type of knee injury. The hospital's surgery review board needs to decide whether to approve this type of surgery. The surgery review board has decided that they will approve this surgery 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.

  What does this lead you to conclude about Ho? Explain:

  ____

  ____

  What does this lead you to conclude about Ha? Explain:

  ____

  ____

- Suppose the sample proportion of people who are cured is p'=0.61, a 61% cure rate.

  What does this lead you to conclude about Ho? Explain:

  ____

  ____

  What does this lead you to conclude about Ha? Explain:

  ____

  ____

But what if p' = 0.75 or 0.70 or 0.65? How far away from the null hypothesis should the sample value be in order to reject the null hypothesis and assume the alternate hypothesis is true?
We will use probability calculations to determine where to draw the "dividing line" to decide if a sample is close to or far from the null hypothesis.

Example C continued: 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, 130 of them are cured by this surgery: **p' = 130/200 = 0.65**

Is p' = **.65 close to or far from the null hypothesis that p = .60 ?**
We need to determine whether our sample proportion p' = 0.65 is "close to" (consistent with) or "far from" (not consistent with) the hypothesized proportion, 0.60, in Ho
To do this we examine probability of getting a sample that "looks like ours" if the null hypothesis is true.
If that probability is small, it indicates that our sample is not consistent with the null hypothesis, but is strong evidence in support of the alternate hypothesis, leading us to reject the null hypothesis.

We need to set a criteria for "what is a small probability?"
A Type I Error would be to decide to approve the surgery, believing it has a cure rate of more than 60%, (p > .60) if in reality surgery is not effective, having a cure rate that is not more than 60% ($p \le .60$)

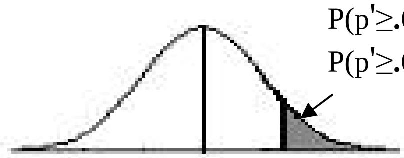
What risk of making a Type I error are we willing to accept?_____________

The risk we are willing to accept of making a Type I error is the significance level α.

The significance level α is our criteria for "what is a small probability?"

Calculate the probability of getting a sample at least as far from the null hypothesis as our sample is. This is called the p value. Since our alternate hypothesis says >, we use the right tail of the distribution

If Ho is true, our random variable P' ~ ____ (_____ , ___________)



P(p'≥.65 if p=.60)
P(p'≥.65 | p=.60) = ____

$\hat{p}$ is another symbol used for the sample proportion
1-PropZTest on our calculator uses $\hat{p}$ rather than p'

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

Decision: ____

Conclusion: ____

____

____

____

____

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 says in context of the problem*)

TRY IT: Suppose in a sample of 200 patients, 138 of them are cured by this surgery: **p' = 138/200 = 0.69**

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

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

A computer company has a new "power saver feature" that it claims will extend the average working time for its laptop batteries to longer than 240 minutes (4 hours) before they need to be charged.
The company management is willing to invest in producing and marketing the power-saver laptop if its engineers can show strong evidence in the sample data to conclude that that the true average working time for all batteries used in power-saver laptops is longer than 240 minutes.
(Testing is done under conditions conforming to industry standards for brightness and typical laptop workload.)

Data for sample of 30 power-saver laptops: time, in minutes, that batteries work until needing recharging

244	223	243	299	212	190	271	254	295	264	258	255	258	281	234
237	230	281	250	257	263	282	264	259	256	265	269	191	225	250

The sample average working time is 252 minutes. The engineers believe that the known population standard deviation of 28 minutes also applies to the new power saver laptop.
Perform a hypothesis test of the claim that the true average working time for all batteries in power saver laptops is longer than 240 minutes. Use a 5% level of significance.

NOTE: *Then find a confidence interval estimate of the true average working time for all batteries in power saver laptops, so that the marketing department will know what claims they can make on packaging and in advertising..*

EXAMPLE F: 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 all her customers is shorter.

A random 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 G: Hypothesis Test of Population Proportion p

A city government needs to determine whether the proportion (percent) of city residents without health insurance is different than the percent nationally. Nationally, it is believed that 16% of US residents lack health insurance. In a survey of 1600 city residents, 244 residents lack health insurance.
Perform a hypothesis test using a 5% level of significance

***TRY IT*:** It is believed that 20% of all residents in that state lack health insurance. Can we conclude that the true proportion (percent) of city residents without health insurance is different than the statewide percent?

EXAMPLE H: Hypothesis Test of Population Proportion p

The owners of Ace Auto Repair believe that at most half of all their customers drive foreign cars.
One of their mechanics is taking Math 10 and as his student project performs a hypothesis test.
He selects a random sample of 50 cars brought in for repair and find that 21 of them are foreign cars.
At a 5% level of significance, perform the hypothesis test.