Chapter 10 Notes: Hypothesis Tests for two Population Parameters
(Tests involving data from Two Samples)

Chapter 10 investigates two populations in which the population parameter for the same variable is not known in both populations.
A hypothesis test is performed using two samples of data to determine if the values of the parameters in these populations are the same or different from each other.
We may be testing whether the difference is ≠ or we may have a theory about the difference that is directional (< or >).
The mathematics of this hypothesis test is that instead of testing two unknown parameters against each other, we test the difference between the two unknown values against a 0 difference (because if the difference is 0, then the two parameters are equal.)

A hint for recognizing these types of hypothesis tests is DATA FROM 2 SAMPLES

Test of two proportions:
- Samples are always independent in Math 10 for proportion problems.
- Example: Comparing proportions of male and female high school grads who attend college.
- Use 2 Prop Z Test.

Test of two means, independent samples
- Independent samples: samples are selected separately (independently) from each other.
- Example: Comparing the average ages of male and female community college students
- If one or both population standard deviation is not known use the t distribution and 2 sample T Test.
- If both population standard deviations are known, use 2 sample Z test – (this rarely occurs).

Test of means for dependent (matched, paired) samples
- Dependent (matched, paired) samples are when there is a correspondence between the items in each sample that pairs the data in the samples with each other.
- Clues for recognizing independent samples are listed below
  - Before and After measurements on the same items or individuals
  - TWO measurements on the same items or individuals
  - A description of a matching or pairing process used to select the samples
- Use the (regular) T Test with the data differences within each pair as the data.
- Many research studies can be set up as independent or as paired samples. An advantage of a paired test is a smaller sample size can be used to obtain reliable results because the paired test reduces the effect variation between individuals; a disadvantage is the difficulty obtaining or determining or availability of appropriate paired samples.

Test of two proportions	Test of two means independent samples	Test of two means paired, matched, or dependent samples
$H_0: p_1 = p_2$ $H_A: p_1 \neq p_2$	$H_0: \mu_1 = \mu_2$ $H_A: \mu_1 \neq \mu_2$	$H_0: \mu_d = 0$ $H_0: \neq 0$
OR $H_0: p_1 \geq p_2$ $H_A: p_1 < p_2$	*OR* $H_0: \mu_1 \geq \mu_2$ $H_A: \mu_1 < \mu_2$	*OR* $H_0: \mu_d \geq 0$ $H_A: \mu_d < 0$
OR $H_0: p_1 \leq p_2$ $H_A: p_1 > p_2$	*OR* $H_0: \mu_1 \leq \mu_2$ $H_A: \mu_1 > \mu_2$	*OR* $H_0: \mu_d \leq 0$ $H_A: \mu_d > 0$

Select the type of test appropriate for the situation described in each example:

Some examples on this page are from Introductory Statistics at OpenStaxcan be downloaded for free at https://openstaxcollege.org/textbooks/introductory-statistics

A. independent group means
B. matched or paired samples
C. single mean
D. two proportions
E. single proportion

____ **EXAMPLE 1 :** A dietician wants to determine if there average amount of salt per serving in hot dogs is more than that in canned soup.
For a sample of 10 brands of hot dogs, the average amount of salt per serving was 603 mg with a sample standard deviation of 41 mg.
For a sample of 10 types of canned soup the average amount of salt per serving was 542 mg with a sample standard deviation of 36 mg.

____ **EXAMPLE 2 :** We want to determine if the proportion of male students who commute to campus by bicycle is the same as the proportion of female students who commute by bicycle to campus.

____ **EXAMPLE 3 :** A study is conducted to investigate the effectiveness of hypnotism in reducing pain. For a sample of people who participated in the study, each person was tested to measure their pain perception (pain sensory measurement) before and after hypnotism.
Are the sensory pain measurements, on average, lower after hypnotism?

____ **EXAMPLE 4:** A hypothesis test is performed to determine if the average time that a pain medication lasts is more than 3 hours. A random sample of 40 patients is given this pain medication and the time in hours that the medication lasts is recorded for each

____ **EXAMPLE 5:** A hypothesis test is performed to determine if the average times that two pain medications A and B last are the same. A random sample of patients is given medication A; another random sample of patients is given medication B.

____ **EXAMPLE 6:** A hypothesis test is performed to determine if the average times that two pain medications A and B last are the same. For a sample of 20 patients, each patient in the sample is given medication A, and the next day each patient in the sample is given medication B.

____ **EXAMPLE 7:** A hypothesis test is performed to determine if recent female college graduates experience salary discrimination, earning less on average for similar work than recent male college graduates in similar jobs with similar qualifications. A random sample of female students is selected, and then a sample of male students is selected so that each male is matched by type of job, major, and GPA to a student in the sample of female students.

____ **EXAMPLE 8:** A hypothesis test is performed to determine if all female workers earn less on average than all male workers. Salary information is obtained for a random sample of female workers and for a random sample of male workers.

____ **EXAMPLE 9:** A study is done to determine if the proportions of residents of San Jose and San Francisco without health insurance is different by examining the proportions of samples of residents from each city who don't have health insurance.

____ **EXAMPLE 10:** Before the Affordable Care Act, 16% of Americans did not have health insurance. Now that the Affordable Care Act has been in existence for several years, we want to conduct a hypothesis test to determine whether the percent of California residents without health insurance under is different from that.

Chapter 10: Hypothesis Tests involving data from TWO SAMPLES
Hypothesis Tests comparing 2 unknown population parameters

Some but not all these examples will be used as class lecture examples.

Those examples with references noted for Introductory Statistics at OpenStax or for Collaborative Statistics by Illowsky and Dean at Connexions.org can be downloaded for free at https://openstaxcollege.org/textbooks/introductory-statistics or http://cnx.org/content/col10522/1.21/

Example A: Example 10.8 in OpenStax Introductory Statistics:
Two medications for hives are being tested to determine if there is a difference in the percentage of adult patient reactions. 20 people in a random sample of 200 adults given medication A still had hives thirty minutes after taking the medication. 12 people in another random sample of 200 adults given medication B still had hives thirty minutes after taking the medication. At a 1% level of significance, is there a difference in the "non-response" rate for medication A and medication B?

Example B: (*not in textbook***) "A/B" Testing:**
Companies and organizations collect data about how people visiting the site use the website. One way they use this data is to test different appearances or formats of the website to determine which gets better responses. Responses can be measured in various ways: time spent at the website, purchases made, donations made, or other metrics that are meaningful to the company or organization.
This is called A/B testing and is commonly to try to increase sales on shopping sites, increase advertising viewership on sites that have paid advertising, or increase donations to political campaigns made through candidates' websites. |
The example below gives one view of how A/B testing may sometimes be conducted; however the sites that conduct A/B testing usually have very large data sets.

To determine if changes in a social networking website's appearance makes a difference in the amount of time that people stay on that site, a hypothesis test is conducted to compare the average time spent at the site for two samples of users.
For a sample of 124 randomly selected users seeing interface A, the average time spent was 2.7 minutes with a standard deviation of 0.6 minutes.
For a sample of 82 randomly selected users seeing interface B, the average time spent was 2.4 minutes with a standard deviation of 0.5 minutes.
Conduct a hypothesis test of determine if there is a difference between the average times that users spend on the site with interface A vs with interface B.

Assume that the populations of times spent at the site by individual users for interface A and interface B are approximately normally distributed. (If this assumption were not true, then other statistical methods beyond the scope of Math 10 would be used to perform the testing.)

Example C (*not in textbook***):** A frozen pizza manufacturer wants to determine whether the average time needed to cook its low fat pizza is less than for its regular pizza.

SAMPLE DATA	Mean Cooking time In minutes	Standard Deviation	Number of Pizzas In sample
Low Fat Pizza	14.8	2.3	15
Regular Pizza	16.1	2.8	15

All pizzas were cooked in identical ovens at the same temperature.
Can we conclude that the true average cooking time is less for low fat pizzas?

Assume that the populations of individual cooking times approximately normally distributed. (If this assumption were not true, then other statistical methods beyond the scope of Math 10 would be used to perform the testing.)

Example D: Example 10.11 in OpenStax Introductory Statistics

A study was conducted to investigate the effectiveness of hypnotism in reducing pain.
Results for randomly selected subjects are shown in the table. The "before" value is matched to an "after" value.
Are the sensory pain measurements, on average, lower after hypnotism?
Perform a hypothesis test using a 5% significance level.

Subject	A	B	C	D	E	F	G	H
Before	6.6	6.5	9	10.3	11.3	8.1	6.3	11.6
After	6.8	2.5	7.4	8.5	8.1	6.1	3.4	2

Example E: (*not in textbook***) :**

A home health care service has ten nurse's aides in the company that visit patients' homes.
Under the old assignment system, appointments were scheduled on a first come first serve basis to fill available time. The company director wants to try a new scheduling system based on location.
The table shows the number of visits before and after the new system is implemented on randomly selected days for a sample of 10 aides.

Is there sufficient evidence to conclude that there is an average increase in the population number of visits per day made by the nurse's aides using the new schedule, as compared to the old schedule?
Perform a hypothesis test using a 5% level of significance.

	Ana	Binh	Cyd	Dina	Ed	Fran	Greg	Hal	Ido	Juna
Number of Visits/Day Old Schedule	6	7	8	6	8	7	11	9	10	12
Number of Visits/Day New Schedule	10	10	9	11	5	10	10	13	8	15

Example F (*not in textbook***):**

In a study of 15,600 patients, patients were randomly assigned to a treatment group receiving the medication plavix or to a control group receiving aspirin. Assume that the 15, 600 patients were equally divided between the two groups. (*data source: San Jose Mercury News 3/13/2006*)
In the treatment group, 6.8% suffered heart attack or stroke.
In the control group, 7.3% suffered a heart attack or stroke.
Perform a hypothesis test at the 2% level of significance to determine whether the treatment is effective at reducing the occurrence of heart attacks and strokes.

Example G (*not in textbook***):**

In 1998, the FDA approved the drug tamoxifen to prevent breast cancer in high risk women, stopping the study earlier than planned based on the strength of the data obtained thus far.

According to data contained in an article in the San Jose Mercury News opinion column on 11/16/98,
13,175 women were randomly assigned to the treatment (tamoxifen) or control (placebo) groups.
Of the 6,576 women in the tamoxifen group, 89 developed invasive breast cancer.
Of the 6,599 women in the placebo group, 175 developed invasive breast cancer.

Perform the appropriate hypothesis test to determine whether the sample data provides sufficient evidence that the incidence of invasive breast cancer is lower in the tamoxifen group than in the placebo group.
Use a 1% level of significance.

Example H: Example 10.6 in OpenStax Introductory Statistics:

The mean lasting times of two floor waxes is to be compared.
20 floors are randomly assigned to test each wax to see how long each wax lasts.
The data are given in the following table

	Sample Mean time in months	Population Standard Deviation
Wax 1	3	0.33
Wax 2	2.9	0.36

Do the data indicate that wax 1 is more effective than wax 2? Use a 5% level of significance.

Example I (*not in textbook*):

Strain A:	Strain B:
24	9
13	13
26	27
15	12
22	15
14	21
10	9
8	28
10	18
18	15
33	23
21	25
30	12
20	18
26	29
23	7
16	8
17	10
14	14
9	15
30	18
24	16
25	28
7	6
19	15

A biologist is studying the average germination times of two strains of seeds to determine whether the two strains of seeds have the same mean germination time.

The number of days until germination are given for a random sample of 25 seeds of strain A and a random sample of 25 seeds of strain B.

All seeds are grown in identical greenhouse conditions.

Assume that the underlying populations of germination times of individual plants is approximately normally distributed.

At a 2% level of significance, is there sufficient evidence of a difference in mean germination times for the two strains of seeds?

What type of hypothesis test is appropriate for this problem? Why?

Sample	Sample Mean	Sample Standard Deviation
A	18.96	7.35
B	16.44	6.98

SETTING UP HYPOTHESES

Test of two proportions	Test of two means independent samples	Test of two means paired, matched, or dependent samples
$H_0: p_1 = p_2$ $H_A: p_1 \neq p_2$	$H_0: \mu_1 = \mu_2$ $H_A: \mu_1 \neq \mu_2$	$H_0: \mu_d = 0$ $H_0: \neq 0$
OR $H_0: p_1 \geq p_2$ $H_A: p_1 < p_2$	*OR* $H_0: \mu_1 \geq \mu_2$ $H_A: \mu_1 < \mu_2$	*OR* $H_0: \mu_d \geq 0$ $H_A: \mu_d < 0$
OR $H_0: p_1 \leq p_2$ $H_A: p_1 > p_2$	*OR* $H_0: \mu_1 \leq \mu_2$ $H_A: \mu_1 > \mu_2$	*OR* $H_0: \mu_d \leq 0$ $H_A: \mu_d > 0$

HOW TO DO THE HYPOTHESIS TEST

Test of proportion p	2PropZTest	Parameters: p_1, p_2 Random variable: $p'_1 - p'_2$	Distribution : Normal
Test of means μ_1, μ_2 when σ_1 and σ_2 both are known independent samples	2 SamZTest	Parameters: μ_1, μ_2 Random variable is $\bar{X}_1 - \bar{X}_2$	Distribution : Normal
Test of means μ when σ_1 or σ_2 or both are NOT known	2 SamTTest *Use NO for "Pooled"*	Parameter: μ_1, μ_2 Random variable is $\bar{X}_1 - \bar{X}_2$	Distribution : t df is given by calculator output
Test of means with paired/ matched/dependent samples.	TTest using differences as data	Parameter: μ_d Random variable: $\bar{X}_d$	Distribution : t df = number of pairs–1

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

GRAPH: Put ZERO in the middle since we are testing if there is "no difference". *0 in the middle says the null hypothesis is that the means or proportions are equal to each other so their difference is 0.*

- For a one tailed test mark the value of the sample statistic $\bar{X}_1 - \bar{X}_2$ or $p'_1 - p'_2$ or $\bar{X}_d$ in the appropriate location on the horizontal axis. Be careful about signs.
  - 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 $\neq$
  - Mark the value of the sample statistic $\bar{X}_1 - \bar{X}_2$ or $p'_1 - p'_2$ or $\bar{X}_d$ in the appropriate location on the horizontal axis. Be careful about signs.
  - Also mark the value that is the same distance from the center on the other side.
  - Shade out to both sides.

DECISION RULE: If p value $< \alpha$, REJECT Ho ; If p value $\geq \alpha$, 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 Error: Deciding H_A is true when in reality H_0 is true
Type II Error: Deciding H_0 is true when in reality H_A is true