Chapter 9 Testing Hypotheses

- Overview
- 5 Steps for testing hypotheses
- One and two-tailed tests
- Type 1 and Type 2 Errors
- Z tests and t tests

Testing Hypotheses (and Null Hypotheses)

Testing Hypotheses is a procedure that allows us to evaluate hypotheses that are typically drawn from a theory and based on sample statistics.

Example of a theory: UNT sociology majors are brighter than the average UNT student.

Example of an hypothesis: UNT Sociology majors have a higher GPA than the UNT student population.

Example of a theory: Commitment to an organization (Organizational Commitment) is affected by the characteristics of the organization.

Example of an hypothesis: Organizational commitment is affected by the procedures used to do the work.

Example of a null-hypothesis:

Organizational commitment is not affected by the procedures used to do the work.

Research Hypothesis

A research hypothesis is a statement typically reflecting a relationship between two variables that can be statistically tested.

By examining the truth of an hypothesis, we are able to draw conclusions about the broader theory from which the hypothesis was derived.

That is, if the hypothesis is found to be true, this lends support for the theory from which the hypothesis was drawn. If the hypothesis is found to be false, this provides evidence that the theory is false.

Null Hypothesis

A null hypothesis is a statement of "no difference." That is, rather than suggesting a relationship exists between two variables, the null hypothesis suggests there is "no relationship" between the variables.

Example of null hypothesis: Organizational commitment is not related to the procedures used to do the work.

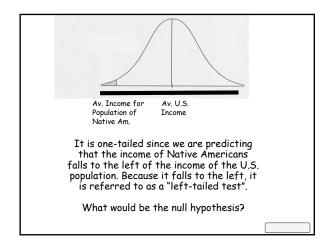
If the null hypothesis is found to be false (i.e., we reject the null hypothesis) then we have support for the research hypothesis and the broader theory. Statisticians typically test the null hypothesis rather than the research hypothesis. The reasons are statistically based.

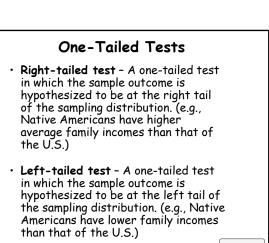
One-Tailed Tests

One-tailed hypothesis test - A hypothesis to be tested where the sample statistic is believed to be either higher or lower when compared to another group.

Example of one-tailed test: The average family income of Native Americans is less than that of the average U.S. family.

In this example the sample statistic is "average family income of Native Americans" and the comparison group is the average income of a family in the U.S.



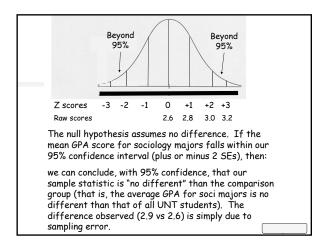


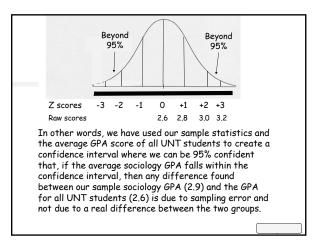
Two-Tailed TestsTwo-tailed hypothesis test - A hypothesis
test in which a sample statistic might fall
within either tail of the sampling
distribution.The nu
for ex
is no disciple
socioleWe are not sure which tail of the curve the
statistic is likely to fall.(fictit
GPA for
Stand
GPA for
Stand
GPA for
Stand
US the average
family income of Native Ams. is not the
same as that of the average U.S. family
(we haven't specified greater or lesser
than)—what would be the null hypothesis?In classes

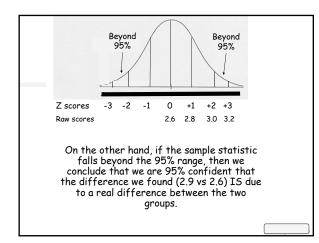
Conducting a one-tailed test The null hypothesis is typically tested: for example, we might test the null hypothesis: there is no difference between the average GPA score of sociology majors and that of all UNT students.

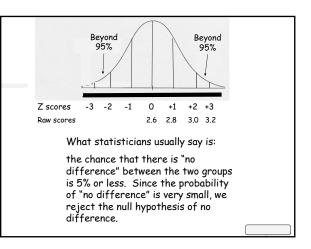
(fictitious data:) GPA for a sample of sociology majors = 2.9 Sample size (N) = 50 Standard Error = .2 GPA for all UNT students = 2.6

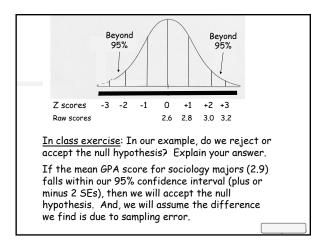
In class: plot the information above on a normal curve. Use the GPA for all UNT students as the mean of the normal curve. Show Z scores and raw scores for three standard errors.

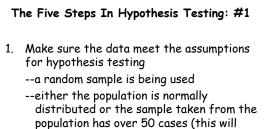












allow us to apply the Central Limit Theorem) --knowing the level of measurement of the

data.

The Five Steps In Hypotheses Testing: #2 State the (a) research hypothesis, (b) the null hypotheses and select (c) alpha. 2a. What is a <u>research hypothesis</u> (H₁) - A statement typically reflecting a relationship between two variables that can be statistically tested. It is typically drawn from a theory. If the research hypothesis is supported by the data then this supports the broader theory.

2b. What is a <u>null hypothesis</u> (Ho) –

- A statement of "no difference," which contradicts the research hypothesis.
- Example: The average GPA of UNT sociology majors is no different than that for all UNT students.
- If we reject the null hypothesis, this provides support for the research hypothesis.

2c. What is an alpha?

Alpha is the probability level we have chosen at which the <u>null hypothesis</u> will be rejected.

In our example above we selected an "alpha" of .05. At .05 alpha, there is only a 5% probability that the null hypothesis is true, i.e., that there is no difference between the sample statistic and comparison group. If the probability is at or smaller than alpha (in this case .05), we will reject the null hypothesis of no difference and assume that the difference found is a true difference (not simply due to sampling error). In other words, we ask the question: If there is actually no difference between the two groups, what is the probability that we would have randomly selected a sample with a statistic (e.g., average sociology GPA) this much larger or smaller than that of the comparison parameter (e.g.,whole UNT student population)?

We then calculate the probability of getting the difference we found (assuming no difference). If we find that there is very little chance (5% or less) of drawing a sample with this much difference, then we will conclude that the difference is a real difference and not due to sampling error. We will reject the null hypothesis and accept the research hypothesis that the two groups are different.

The Five Steps In Hypotheses Testing: #3

3. Select the sampling distribution and specify the test statistic.

Our sampling distribution will be either the Z distribution or t distribution. We will use either the Z or t distribution to assist us in testing our null hypothesis (so far we have used only the Z distribution—what we have referred to as the normal curve).

Our test statistic will be either the Z statistic or the t statistic. We will use either the Z or t statistic to assist us in testing our null hypothesis (so far we have used only Z statistics such as 1.96, and 2.58)

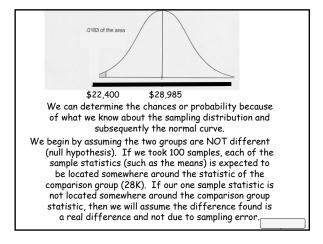


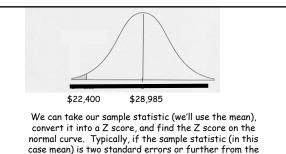
4. Calculate the test statistic, e.g., convert the sample mean to a Z statistic or t statistic.

Let's use an example to calculate a Z statistic: If we assume that our sample mean family income for Native Americans is \$22,400 and that our mean family income for the population as a whole is \$28,985:

then we would ask:

What are the chances that we would have randomly selected a sample of Native Americans with an average family income of \$22,400 if there is really no difference between their income and that of the population as a whole (\$28,985)?





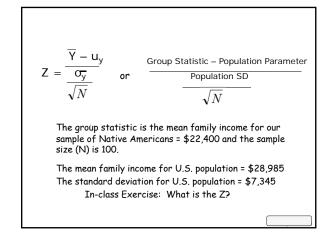
mean, we will reject the null hypothesis. That is, if the "p value" of the sample statistic is .05 or less, we will reject the null hypothesis (in some cases we may decide to use three standard errors as the level at which to reject the null hypothesis in which case the p value of the sample statistic must be .01 or less).

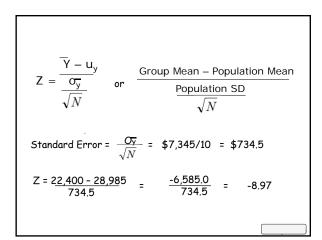


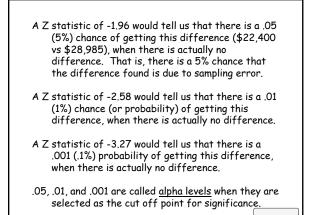
The mean family income for our sample of Native Americans = \$22,400 with a sample size of 100

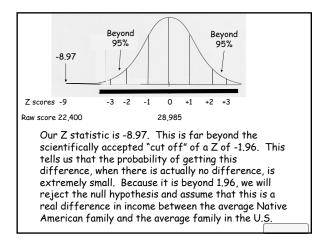
The mean family income for U.S. population = \$28,985 The standard deviation for U.S. population = \$7,345

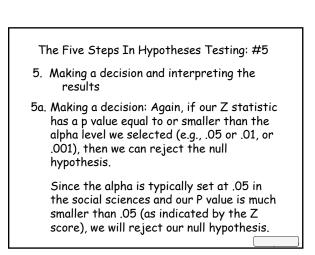
We want to test the null hypothesis that Native American income is no different than the U.S. population income.











5b. Interpreting the results:

The "Z statistic" of -8.97 is <u>statistically significant</u>, that is, it is very unlikely that the difference we found occurred by chance or sampling error.

We can say that the difference between the family income of Native Americans and the U.S. population is significant beyond the .001 level.

We can also say that the average family income of Native Americans was substantially less than the average income of families in the U.S. when the survey was taken. We say "substantially" because there is a \$6,400 difference and we consider this to be a very big (substantive) difference. Of course, this is only our opinion of the difference.

Summary: Steps in Testing an Hypothesis

- 1. Verify that assumptions are met
- State research and null hypotheses and alpha (our example hypothesized that the average income of Native Americans was less than the U.S. population as a whole)
- 3. Select sampling distribution and test statistic to be used (Z or t statistic)
- 4. Compute test statistic
- 5. Make a decision and interpret results

Errors that we try to avoid

Type I error: if the null hypothesis is rejected when it is actually true. (this is most common since researchers want to reject the null hypotheses so they can accept their hypotheses).

Type II error: if the null hypothesis is accepted when it is actually false.

Type I and Type II Errors and their relationship to alpha

- When selecting our alpha, we need to be aware that if we set alpha too large (e.g. p value of.10) we may create a Type I error that is, we might reject the null hypothesis when it is actually true.
- Or, if we set the **alpha too small** (e.g., .001) we may create a **Type II** error by failing to reject a false null hypothesis.

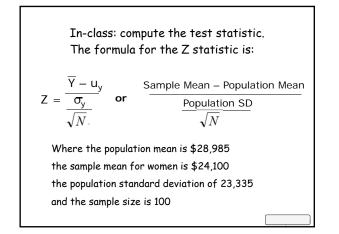
Let's take another example:

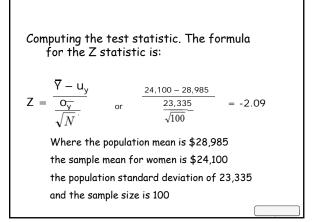
--Our research hypothesis is that the salary for women is less than that for the U.S. population.

--We will test the null hypothesis that the salary for women is no different than that for the U.S. population as a whole.

--We sample at least 50 women so that our theoretical sampling distribution will be normally distributed (a required assumption). --The test statistic used is either the Z statistic or t statistic and since we have

the population standard deviation we will use the Z statistic.



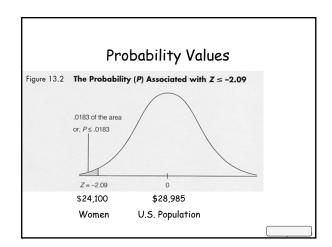


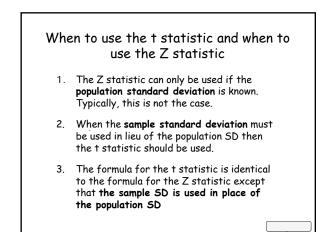
Make a Decision and Interpret the results. In our example:

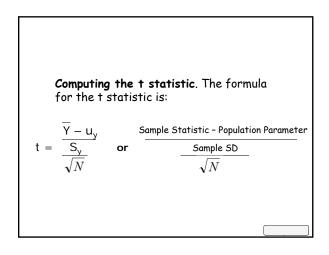
--we confirm that the Z is on the left tail of the distribution (-2.09)

--the Z score is smaller than our selected alpha of .05. We know this because a Z score of -1.96 is at the .05 level and our Z score of -2.09 is further out than -1.96.

--thus, we can reject the null hypothesis of no difference and can conclude that the average income for women is less than that of the general population.







Steps for interpreting the t statistic

- Unfortunately, locating where the t statistic falls on the normal curve is not as easy as when using the Z statistic.
- Once the t statistic is calculated it is compared to the t value needed to reject the null hypothesis. The t value needed can be found on a t distribution table and will vary depending on whether the researcher has chosen an alpha of .05, .01, etc.

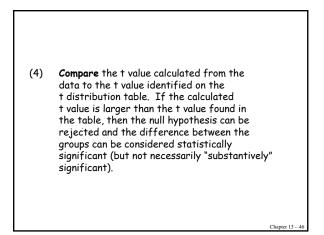
Chapter 13 - 4

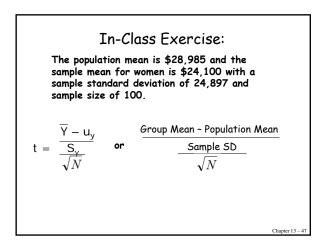
How to use the t distribution table to determine significance

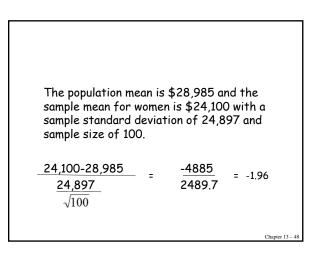
- Determine the degrees of freedom your sample provides (this is typically: N-1) and then locate the DF on the t-distribution table (table is on page 484-5).
- (2) Find on the table: the alpha which you selected at the start of the statistical analysis (an alpha of .05 and a two-tailed test are typically used by researchers)
- (3) Find the intersecting point where the DF and the alpha cross. At the intersecting point you will find the t value needed to reject the null hypothesis.

Chapter 13 - 44

	Jes of the t Distribution							
	.10	.05	.025	.01	.005	.0005		
	Level of Significance for Two-Tailed Test							
df	.20	.10	.05	.02	.01	.001		
1	3.078	6.314	12.706	31.821	63.657	636.619		
2	1.886	2.920	4.303	6.965	9.925	31.598		
3	1.638	2.353	3.182	4.541	5.841	12.94		
4	1.533	2.132	2.776	3.747	4.604	8.610		
5	1.476	2.015	2.571	3.365	4.032	6.85		
10	1.372	1.812	2.228	2.764	3.169	4.58		
15	1.341	1.753	2.131	2.602	2.947	4.073		
20	1.325	1.725	2.086	2.528	2.845	3.850		
25	1.316	1.708	2.060	2.485	2.787	3.72		
30	1.310	1.697	2.042	2.457	2.750	3.640		
40	1.303	1.684	2.021	2.423	2.704	3.55		
60	1.296	1.671	2.000	2.390	2.660	3.460		
120	1.289	1.658	1.980	2.358	2.617	3.37		
00	1.282	1.645	1.960	2.326	2.576	3.29		







Finding the t statistic in the t distribution table

Our **degrees of freedom** for this example is N - 1 or 99 and our t statistic is -1.96 (the larger the t statistic the more likely it will be significant).

On page 497-98 of your book we can find the **t** distribution table. It displays the degrees of freedom for 60 and for 120. Since ours is 99 it is less than 120. Therefore, to be conservative we will use 60 DF.

We can assume a **one-tailed test** since existing knowledge indicates that women make less than the population as a whole and certainly not more (the mean will fall on the left side of the curve).

		Level of Significance for One-Tailed Test								
		.10	.05	.025	.01	.005	.0005			
		Level of Significance for Two-Tailed Test								
	df	.20	.10	.05	.02	.01	.001			
	1	3.078	6.314	12.706	31.821	63.657	636.619			
	2	1.886	2.920	4.303	6.965	9.925	31.598			
	3	1.638	2.353	3.182	4.541	5.841	12.941			
	4	1.533	2.132	2.776	3.747	4.604	8.610			
	5	1.476	2.015	2.571	3.365	4.032	6.859			
	10	1.372	1.812	2.228	2.764	3.169	4.587			
	15	1.341	1.753	2.131	2.602	2.947	4.073			
	20	1.325	1.725	2.086	2.528	2.845	3.850			
	25	1.316	1.708	2.060	2.485	2.787	3.725			
	30	1.310	1.697	2.042	2.457	2.750	3.646			
	40	1.303	1.684	2.021	2.423	2.704	3.551			
	60	1.296	1.671	2.000	2.390	2.660	3.460			
	120	1.289	1.658	1.980	2.358	2.617	3.373			
	00	1.282	1.645	1.960	2.326	2.576	3.291			

Determining Statistical Significance

Since our t statistic is -1.96 we can conclude statistical significance at the .05 level.

Would our findings be significant if we had chosen an alpha of .01?

Would our findings be significant using a 2-tailed test (women's income is different from the general population)?

Chapter 13 - 51

Chapter 13 -

Comparing the Sample Statistics of Two Groups (Presented above is a comparison of a group's sample statistic to a population parameter)

Example for comparing two groups:

Comparing the mean salary of new sociology professors (group 1) to the mean salary for new engineering professors (group 2). Previously we were comparing the group statistic (such as the mean salary of sociology professors) to the population parameter (such as the mean salary of the whole U.S. population).

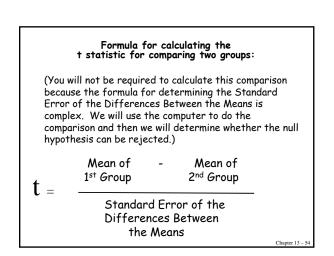
Steps for comparing the sample statistics of two groups are the same as that for comparing a sample statistic to the population parameter with three exceptions:

(1) the formula for calculating the t statistic is different

(2) calculating the degrees of freedom is different, and

(3) determining whether the two groups have equal or unequal variances. (If the Levene's test is significant then their variances are unequal)

Chapter 13 - 53



Example: Use "Class Survey Data" and compare two groups by calculating the t statistic. 1. Click Analyze 2. Click Compare Means 3. Independent Sample t test 4. Move Variable 13 (experiments w/ drugs) into the "Grouping Variable" 5. Click Define Commence and enter 1 for Comm 1 and 2

- 5. Click Define Groups and enter 1 for Group 1 and 2 for Group 2
- 6. Click Continue
- 7. Move V2, V11, V12, V17, V19, V23, V24, V26, to "Test Variables"

Chapter 13 - 55

8. Click Okay

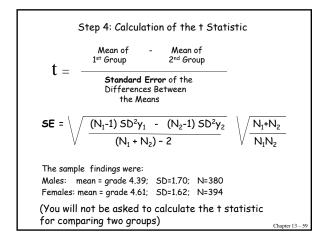


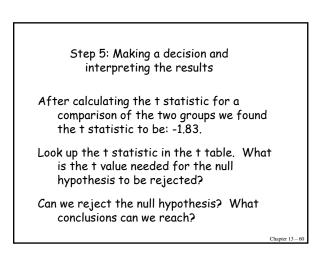
Step 1: What assumptions must be met?

- 1. Independent random samples
- 2. Years of education measured as interval level
- 3. normal population or sample size at least 50 for each group

Step 2: What are the research and null hypotheses and what alpha will we use (we suspect males are different from females but we don't know if they started smoking in a higher or lower grade than did females)?

Step 3: What test statistic will be used?





Example: Comparing Two Sample Means (male and female job burnout) In SPSS:

1.Analyze

2.compare means

3.independent sample t test

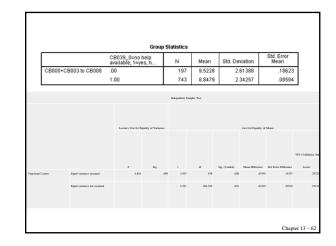
4.move V101 (sex) to "group variable" box

5.click "define group"

6.in group 1 put "1" (female) and in group 2 put "2" (male)

7.click continue

8.move "burnout", V102 (age), V100 (education), V114 (# residents assigned) to "test variables" box and then click Okay



To determine the probability of observing a Z value of -8.97, we look up the value in Appendix B to find the area beyond a Z of -8.97 (Column C).

Course Review

- 1. Levels of Measurement
- 2. Measures of central tendency
- 3. Measures of variability
- 4. Normal Distribution
- 5. Sampling
- 6. Estimation of population
- 7. Testing hypotheses

Chapter 13 - 63

Course Review

- 1. Levels of Measurement (nominal, ordinal, interval/ratio)
- 2. Measures of central tendency (mean, median, mode)

Chapter 13

- 3. Measures of variability
- 4. Normal Distribution
- 5. Sampling
- 6. Estimation of population
- 7. Testing hypotheses

1. Levels of Measurement (nominal, ordinal,

Course Review

- interval/ratio) 2. Measures of central tendency
- 3. Measures of variability
- 4. Normal Distribution
- 5. Sampling
- 6. Estimation of population
- 7. Testing hypotheses

Course Review

- 1. Levels of Measurement (nominal, ordinal, interval/ratio)
- 2. Measures of central tendency (mean, median, mode)
- 3. Measures of variability (range, variance, standard deviation)
- 4. Normal Distribution
- 5. Sampling
- 6. Estimation of population
- 7. Testing hypotheses

Course Review

- 1. Levels of Measurement (nominal, ordinal, interval/ratio)
- 2. Measures of central tendency (mean, median, mode)
- 3. Measures of variability (range, variance, standard deviation)
- 4. Normal Distribution (normal curve)
- 5. Sampling

Chapter 13 - 6

Chapter

- 6. Estimation of population
- 7. Testing hypotheses

Course Review

- 1. Levels of Measurement (nominal, ordinal, interval/ratio)
- 2. Measures of central tendency (mean, median, mode)
- 3. Measures of variability (range, variance, standard
- deviation) 4. Normal Distribution (normal curve)
- 5. Sampling (random, sampling distribution)
- Estimation of population
- 7. Testing hypotheses

Course Review

- 1. Levels of Measurement (nominal, ordinal, interval/ratio)
- 2. Measures of central tendency (mean, median, mode)
- 3. Measures of variability (range, variance, standard deviation)
- 4. Normal Distribution (normal curve)
- 5. Sampling (random, sampling distribution)
- 6. Estimation of population (confidence intervals, confidence levels)
- 7. Testing hypotheses

Course Review

- 1. Levels of Measurement (nominal, ordinal, interval/ratio)
- 2. Measures of central tendency (mean, median, mode)
- 3. Measures of variability (range, variance, standard deviation)
- 4. Normal Distribution (normal curve)
- 5. Sampling (random, sampling distribution)
- 6. Estimation of population (confidence intervals, confidence levels)
- 7. Testing hypotheses (z test, t test)

Once identifying the t score, it can be found (or a number close to it) in the t distribution table.

This requires knowing the degrees of freedom. If the variances of the two groups are relatively equal then the formula for the degrees of freedom is simple:

$$df = (N_1 + N_2) - 2$$

 N_1 = number of cases in first group N_2 = number of cases in second group

Chapter 13 - 72

Chapter 13 - 68