**Notation & Definitions**

In Chapter 10, we learn how to make inferences about two populations – either about the difference between two population means, or the difference between two population proportions. We draw those inferences by taking a sample from each population, and then using the sample data to do hypothesis testing or to calculate confidence intervals. In each problem, we will treat one population as Population 1, and the other as Population 2, and those numbers become subscripts in our notation (see below). Sample 1 is the sample taken from Population 1, and Sample 2 is the sample taken from Population 2.

|  |  |  |  |
| --- | --- | --- | --- |
| Population Parameters | | Sample Statistics | |
|  | mean of population 1 |  | mean of sample 1 |
|  | mean of population 2 |  | mean of sample 2 |
|  | standard deviation of population 1 |  | standard deviation of sample 1 |
|  | standard deviation of population 2 |  | standard deviation of sample 2 |
|  | proportion of population 1 |  | proportion of sample 1 |
|  | proportion of population 2 |  | proportion of sample 2 |
|  | the difference between two population means |  | the size of sample 1 |
|  | the difference between two population proportions |  | the size of sample 2 |

**Part One: Hypothesis Tests about the Difference between Two**

**Population Means,**

1. **Formulating the Hypotheses**: There are three possible forms of hypotheses. They each correspond to a different question you might want to ask about the difference between two population means,

should be set to nless you are testing for a **specific numerical difference** between the means (like a difference of 5 or 12 or 178). If then there is no difference between the means – they are equal.

|  |  |  |
| --- | --- | --- |
| Hypotheses for Hypothesis Tests about the Difference between Two Population Means, | | |
| **Lower Tail Test** | **Upper Tail Test** | **Two-Tailed Test** |
|  |  |  |
| Answers Questions About: | | |
| **If** then this test determines whether the mean of population 1, is less than the mean of population 2,  **If**  then this test determines whether the difference between the mean of population 1 and the mean of population 2 is less than | **If** then this test determines whether the mean of population 1, is greater than the mean of population 2,  **If**  then this test determines whether the difference between the mean of population 1 and the mean of population 2 is greater than | **If** then this test determines whether the mean of population 1, is different from the mean of population 2,  **If**  then this test determines whether the difference between the mean of population 1 and the mean of population 2 is not equal to |

|  |
| --- |
| *Example 1:* Suppose you have two populations and you want to know whether the mean of population 1 is lower than the mean of population 2. This corresponds to a lower tail test. No specific numerical difference between the two populations is asked for, so we set to 0 and the hypotheses would be: |
| *Example 2:* Suppose you have two populations, and you want to know whether the difference between their means is 25 or not. (Of course, a hypothesis test can only prove that the difference is NOT 25, not that it IS 25). This question matches with a two-tailed test. It asks for a specific numerical difference between the means of 25, so we set to 25 and the hypotheses are: |

NOTE: The procedure explained in this handout is based on drawing two random, independent samples – one from each population – and taking the difference between the means of those samples, .

1. **Choosing and Calculating the Test Statistic:**

The sample statistic of interest in hypothesis tests about the difference between two population means is the difference between the two sample means, We will need to determine how probable a given difference between two sample means is in order to decide whether or not to reject . To do that, we will calculate a test statistic using information from both samples, and sometimes from both populations.

For hypothesis testing about the difference between two population means, the correct choice of test statistic depends on the sampling distribution of , which in turn depends on what information you have about the variability in the underlying populations. The *population standard deviations,*  are sometimes known from prior knowledge, historical information, or past experience. On the other hand, if are unknown, then we use the information about variability of the samples: the standard deviations measured on the samples themselves, .

**When are KNOWN:**

Under the assumption that is true as an equality, the sampling distribution of the difference between the two sample means, follows the distribution. Therefore, we standardize the samples against the distribution by calculating the following test statistic:

where

* + **When are UNKNOWN:**

Under the assumption that is true as an equality, the sampling distribution of the difference between the two sample means, follows the distribution with given by the equation below. Therefore, we standardize the samples against the distribution by calculating the following test statistic:

where

and the degrees of freedom are:

NOTE: if the degrees of freedom has decimal places, **ALWAYS round DOWN to the nearest whole number.**

1. **Deciding whether or not to Reject :**

There are two approaches to deciding whether or not to reject the Null:

* In the ***p-value approach***, we **compare** the **p-value** of our test statistic to the  **significance level** and reject if the p-value is less than or equal to the significance level.
* In the ***critical value approach***, we **compare** the **test statistic** to **a critical value** – this can be done with a diagram in which we use the critical value(s) to construct a rejection region or regions; if the test statistic is in a rejection region, we reject and accept Or, this step can be accomplished by following the mathematical rules given in the tables below.

|  |  |  |  |
| --- | --- | --- | --- |
| **When to Reject for test statistics:** | | | |
|  | **For a Lower Tail Test:** | **For an Upper Tail Test:** | **For a Two-Tailed Test:** |
| **p-value approach:** | Look up the lower tail of  If the then reject and accept  If the then do not reject is unsupported. | Look up the upper tail of  If the then reject and accept  If the then do not reject is unsupported. | The two-tailed is two times the one tailed p-value of  If the then  reject and accept  If the then do not reject is unsupported. |
| **Critical Value: Approach** | If then reject and accept  If then do not reject is unsupported. | If then reject and accept  If then do not reject is unsupported. | If  , then reject and accept .  If then do not reject is unsupported. |
| NOTES:   1. is the Test Statistic 2) and are, respectively, UT, LT, and 2T Critical Values | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **When to Reject for test statistics:** | | | |
|  | **For a Lower Tail Test:** | **For an Upper Tail Test:** | **For a Two-Tailed Test:** |
| **p-value approach:** | Look up the lower tail of  If the then reject and accept  If the then do not reject is unsupported. | Look up the upper tail of  If the then reject and accept  If the then do not reject is unsupported. | The two-tailed is two times the one tailed p-value of  If the then  reject and accept  If the then do not reject is unsupported. |
| **Critical Value: Approach** | If then reject and accept  If then do not reject is unsupported. | If then reject and accept  If then do not reject is unsupported. | If  , then reject and accept .  If then do not reject is unsupported. |
| NOTES:   1. is a Test Statistic 2. and are, respectively, UT, LT, and 2T Critical Values 3. The correct t distribution to use when determining the p-value or the Critical Value depends on the degrees of freedom. | | | |

1. **Interpreting the test:**

(Note: This explanation of interpretation holds for ALL hypothesis tests.) We start every hypothesis test with a question about the parameter of interest, so we must end every hypothesis test with the answer to that question. In other words, we must *interpret* the conclusion of our test in terms of the original question.

Remember: in hypothesis testing you can never prove the null hypothesis. You can only prove the alternative hypothesis: when you reject the null and accept the alternative, then at your given level of significance you may conclude that is true. If you do not reject then you must conclude that is unsupported by the evidence. This gives us a clear guideline for how to *interpret* hypothesis tests: *always look at the alternative hypothesis!*

In all that follows, you would substitute the actual words and numbers from your hypothesis test for the symbols. Notice that each interpretation simply states the alternative hypothesis in words, and says either that it is true or that it is unsupported by the evidence.

(*continued on next page)*

|  |  |  |  |
| --- | --- | --- | --- |
| **How to Interpret a Hypothesis Test about the Difference between Two Population Means, :** | | | |
| **When you:** | **For a Lower Tail Test:** | **For an Upper Tail Test:** | **For a Two-Tailed Test:** |
| **Reject** | **If** At the significance level, we can conclude that the mean of population 1, is less than the mean of population 2,  **If**  At the significance level, we can conclude that the difference between the mean of population 1, and the mean of population 2, is less than | **If** At the significance level, we can conclude that the mean of population 1, is greater than the mean of population 2,  **If**  At the significance level, we can conclude that the difference between the mean of population 1, and the mean of population 2, is greater than | **If** At the significance level, we can conclude that the mean of population 1, is different from the mean of population 2,  **If**  At the significance level, we can conclude that the difference between the mean of population 1, and the mean of population 2, is not equal to |
| **Do not reject** | **If** At the significance level, we cannot conclude that the mean of population 1, is less than the mean of population 2,  **If**  At the significance level, we cannot conclude that the difference between the mean of population 1, and the mean of population 2, is less than | **If** At the significance level, we cannot conclude that the mean of population 1, is greater than the mean of  population 2,  **If**  At the significance level, we cannot conclude that the difference between the mean of population 1, and the mean of population 2, is greater than | **If** At the significance level, we cannot conclude that the mean of population 1, is different from the mean of population 2,  **If**  At the significance level, we cannot conclude that the difference between the mean of population 1, and the mean of population 2, is different from |
| NOTES:   1. is the true value of the mean of population 1   is the true value of the mean of population 2   1. is the hypothesized difference between the two population means, | | | |

**Assumptions Underlying Hypothesis Tests about the Difference between Two Population Means,**

All hypothesis tests use sampling distributions to determine the probability of sample statistics – that is how we determine whether or not to reject the null hypothesis. In order for us to be confident that our choice of sampling distribution for any given test really is the way the sample statistic is distributed, certain assumptions must be met. If the assumptions are not met – that is, if any given assumption is not true – then we cannot rely on the results of the hypothesis tests. They may mislead us, give us the wrong answers, and cause us to draw the wrong conclusions.

* For the hypothesis tests presented above about the difference between two population means,
  + Both samples must be random and independent of one another
  + If both populations are normally distributed, then the sample sizes can be small
  + If both populations are not normally distributed, or if one or more of the distributions are unknown, then the sample sizes must both be greater than or equal to 30
  + When and are unknown (that is, when you are using the t distribution):
    1. It is recommended to have equal sample sizes, although this is not strictly necessary. If the sample sizes are nearly equal, then having sample sizes that satisfy are adequate.
    2. if the populations have skewed distributions or have outliers, then the sample sizes must be larger

**Part Two: Hypothesis Tests about the Difference between Two Population Proportions,**

1. **Formulating the Hypotheses**: There are three possible forms of hypotheses. They each correspond to a different question you might want to ask about the true value of the difference between two population proportions,

|  |  |  |
| --- | --- | --- |
| Hypotheses for Hypothesis Tests about the Difference between Two Population  Proportions, | | |
| **Lower Tail Test** | **Upper Tail Test** | **Two-Tailed Test** |
|  |  |  |
| Answers Questions About: | | |
| If the true proportion of population 1, is less than the true proportion of population 2, | If the true proportion of population 1, is greater than the true proportion of population 2, | If the true proportion of population 1, is different from the true proportion of population 2, |

NOTE: The procedure explained in this handout is based on drawing two random, independent samples – one from each population – and taking the difference between the proportions of those samples, .

1. **Calculating the Test Statistic:**

NOTE: The logic is the same as all previous hypothesis tests: we are looking to see whether our sample values contradict the null hypothesis.

The sample statistic of interest in hypothesis tests about the difference between two population proportions is the difference between the two sample proportions, In order to decide whether or not to reject we need to use a sampling distribution of to determine the probability of a given difference between two sample proportions, under the assumption that is true as an equality – that is, assuming that the two proportions are equal. If the sample statistic is less probable than the significance level, then it contradicts the null hypothesis () and we will reject .

*(continued on next page)*

Under the assumption that is true as an equality, the sampling distribution of follows the distribution. Therefore, we have to standardize the sample information against the distribution by calculating the test statistic as follows:

* **Calculating the Test Statistic for Hypothesis Tests about :**
  1. First, calculate the **pooled estimator, :**
  2. Second, calculate the test statistic:

where

1. **Deciding whether or not to Reject :**

The rules for z test statistics remain the same regardless of the context in which the z test statistic is used. See the rejection rules for z test statistics given on p5 above.

1. **Interpreting the test:**

(Note: This explanation of interpretation holds for ALL hypothesis tests.) We start every hypothesis test with a question about the parameter of interest, so we must end every hypothesis test with the answer to that question. In other words, we must *interpret* the conclusion of our test in terms of the original question.

Remember: in hypothesis testing you can never prove the null hypothesis. You can only prove the alternative hypothesis: when you reject the null and accept the alternative, then at your given level of significance you may conclude that is true. If you do not reject then you must conclude that is unsupported by the evidence. This gives us a clear guideline for how to *interpret* hypothesis tests: *always look at the alternative hypothesis!*

In all that follows, you would substitute the actual words and numbers from your hypothesis test for the symbols. Notice that each interpretation simply states the alternative hypothesis in words, and says either that it is true or that it is unsupported by the evidence.

|  |  |  |  |
| --- | --- | --- | --- |
| **How to Interpret a Hypothesis Test about the Difference Between Two Population Proportions, :** | | | |
| **When you:** | **For a Lower Tail Test:** | **For an Upper Tail Test:** | **For a Two-Tailed Test:** |
| **Reject** | At the significance level, we can conclude that is less than . | At the significance level, we can conclude that is greater than . | At the significance level, we can conclude that is different from . |
| **Do not reject** | At the significance level, we cannot conclude that is less than . | At the significance level, we cannot conclude that is greater than . | At the significance level, we cannot conclude that is different from . |
| NOTE:  is the true value of the proportion of population 1;  is the true value of the proportion of population 2 | | | |

**Assumptions Underlying Hypothesis Tests About**

In order to use this hypothesis test:

* The samples must be random and independent
* The following conditions must hold true:

where

**Formulating Hypotheses: What is the difference?**

*Example 1:*

An article suggests that Business School students take more credit hours on average per semester than students in the liberal arts and sciences. An analyst wishes to see whether this is true at Wayne State. How should the null and alternative hypotheses be formulated? Use Business School students as Population 1.

*Example 2:*

Is there any difference between Mac users and PC users when it comes to mean customer satisfaction? Formulate the null and alternative hypotheses that could be used to test this question. Use Mac users as Population 1.

*Example 3:*

Sprint advertises that its prices for cell phone service are lower than Verizon. Formulate the null and alternative hypotheses to test Sprint’s claim. Use Sprint customers as Population 1.

*Example 4*:

An analyst suspects that, on average, iPhone users download 2 apps per month more than Android users. If the analyst wants to test for this difference, what should the null and alternative hypotheses be? (HINT: the test must determine whether the difference between the mean apps downloaded by iPhone users and Android users is 2 or not). Use iPhone users as Population 1.

**Hypothesis Tests about the Difference between Two Population Means: Exercises**

*Exercise 1:*

*Using data from Chapter 10, Exercise 7 (although changing the question asked)*

*Consumer Reports* uses a survey of readers to obtain customer satisfaction ratings for the nation’s largest retailers (*Consumer Reports*, March 2012). Each survey respondent is asked to rate a specified retailer in terms of six factors: quality of products, selection, value, checkout efficiency, service, and store layout. An overall satisfaction score summarizes the rating for each respondent with 100 meaning the respondent is completely satisfied in terms of all six factors. From past experience, *Consumer Reports* customer satisfaction scores have had a population standard deviation, regardless of store, of 12. We assume that customer satisfaction scores are normally distributed.

Two independent, random samples were taken. A random sample of 25 Target customers had a mean customer satisfaction score of 79. A random sample of 30 Walmart customers had a mean customer satisfaction score of 71.

Conduct and interpret a hypothesis test to determine whether the population mean customer satisfaction score for Target is higher than the population mean customer satisfaction score for Walmart at the significance level. Use Target customers as Population 1.

*Exercise 2:*

A winery was interested in the difference between white wine and red wine drinkers when it comes to average weekly wine consumption. Two random, independent samples were taken. A random sample of 35 white wine drinkers drank a mean of 17.6 ounces of wine per week with a standard deviation of 4.2. A random sample of 38 red wine drinkers drank a mean of 14.7 ounces of wine per week with a standard deviation of 3.64.

Conduct and interpret a hypothesis test to determine if the mean weekly wine consumption of white wine drinkers differs from the mean weekly wine consumption of red wine drinkers at the significance level. Consider the white wine drinkers to be Population 1.

**Hypothesis Testing for the Difference between Two Population Proportions: Exercise**

A heart rate monitor manufacturer offers a service for users to upload their workout data and analyze it using a web-based app. The manufacturer is interested in different groups of athletes and how they compare in using this upload service. The question is: do triathletes and pure runners differ in terms of the proportion that uses the upload service?

Random, independent samples of triathletes and pure runners were taken. In a sample of 292 triathletes, 79.45% used the upload service. In a sample of 315 pure runners, 72.38% used the upload service.

Conduct and interpret a hypothesis test to determine whether triathletes and pure runners differ in the proportion that uses the upload service for their workout data. Use an significance level. Assign triathletes as Population 1 and pure runners as Population 2.

**Confidence Intervals for the Differences between Two Populations**

*Hypothesis tests* are one way to draw inferences about population parameters from sample statistics. Hypothesis tests can answer questions about whether parameters are equal to or not equal to one other, greater than one another, or less than one another.

*Confidence interval estimates* are another way to draw inferences about population parameters from sample statistics. Confidence intervals give us information about what the difference between two population parameters actually is, to a specified degree of certainty. In the case of difference in means, the **confidence interval estimate** gives **a range that contains the true difference between two population means.**

Confidence intervals are always calculated at a given level of confidence, which is a percentage that is related to the significance level.

This table shows the conversions between some typical significance levels and the corresponding confidence levels:

|  |  |
| --- | --- |
| significance level | Confidence level (%) |
| 0.10 | 90% |
| 0.05 | 95% |
| 0.01 | 99% |

You can calculate these conversions yourself. For example, if the confidence level is

A **confidence interval** consists of two parts: a ***point estimate*** plus or minus a ***margin of error***. A point estimate is a single number that estimates a population parameter. The margin of error determines the range around the point estimate. In the equations that follow, the point estimate is the part before the and the margin of error is the part following it.

The point estimate minus the margin of error is the ***lower bound*** **(LB)** of the confidence interval, and the point estimate plus the margin of error is the ***upper bound* (UB)**. Confidence intervals are usually reported in square brackets, with the bounds separated by a comma, as in **[***LB***,** *UB***].**

**Part I: Confidence Intervals for the Difference between Two Population Means,**

For the difference between two population means, choosing the correct formula for the confidence interval estimate depends on whether and – the population standard deviations for population 1 and 2 – are KNOWN or UNKNOWN.

1. **If are KNOWN**, the  *confidence interval for the difference between two population means,*  is given by:

where

1. **If are UNKNOWN**, the sample standard deviations s1 and s2 must be used instead, and the  *confidence interval for the difference between two population means,*

is given by:

where has the degrees of freedom:

(NOTE: Always round the degrees of freedom DOWN to the nearest whole number, no

matter what.)

and

The template for the **interpretation** of a confidence interval for is:

* *With* ***\_\_\_****% confidence*, *we can conclude that the true difference between* *and is between* ***[lower bound]*** *and* ***[upper bound]****.*
  + IMPORTANT: If the interval **contains zero**, then there is **no statistically significant difference** between the two means.
  + As usual, you should fill in the meaning for the notation above, referring to the specific interval you are interpreting.

**Part Two: Confidence Intervals for the Difference Between Two Proportions,**

A confidence interval estimate for the difference between two population proportions, is a range of numbers that contains the true difference between the two population proportions, to a given level of confidence.

BE CAREFUL: confidence interval calculations for the difference between two population proportions DO NOT use the pooled estimator, that is used in hypothesis testing for difference in proportions.

The  *confidence interval estimate for the difference between two population proportions,* is given by:

where

Remember, the point estimate minus the margin of error is the ***lower bound*** **(LB)** of the confidence interval, and the point estimate plus the margin of error is the ***upper bound* (UB)**. Confidence intervals are usually reported in square brackets, with the bounds separated by a comma, as in **[***LB***,** *UB***].**

The template for the **interpretation** of a confidence interval for is:

* *With* ***\_\_\_****% confidence*, *we can conclude that the true difference between* *and is between* ***[lower bound]*** *and* ***[upper bound]****.*
  + IMPORTANT: If the interval **contains zero**, then there is **no statistically significant difference** between the two proportions.

**Confidence Intervals for Differences: Exercises**

*Exercise 1:*

Do Android or iPhone users differ in how many apps they download per month? What is the difference between them? An analyst takes two independent, random samples: one of Android users and one of iPhone users. The sample of 50 Android users downloaded a mean of 2.4 apps per month. The sample of 55 iPhone users downloaded a mean of 3.7 apps per month. Historical data suggest that the population standard deviation for Android users is 1 app per month, and the population standard deviation for iPhone users is 1.2 apps per month.

Calculate and interpret the 95% confidence interval estimate for the difference in mean apps downloaded per month between Android and iPhone users. Assign population 1 to be iPhone users.

*Exercise 2*:

An HR manager would like to know how much the mean number of overtime hours differs between employees that have children at home and those who don’t have children at home. The manager takes a random sample of 35 employees with children at home and finds that these employees work a mean of 4.2 overtime hours and have a standard deviation of 4 hours. An independent random sample of 40 employees without children at home has a mean of 6.5 overtime hours and a standard deviation of 5.9 hours.

Calculate and interpret a 99% confidence interval estimate for the difference in mean overtime hours between these two groups. Use employees with children at home as Population 1. The t-distribution appropriate for use in this instance has 68 degrees of freedom. (On your own time, calculate out the degrees of freedom – remember to *always round down* when calculating degrees of freedom this way!!)

*Exercise 3:*

A heart rate monitor manufacturer offers a service for users to upload their workout data and analyze it using a web-based app. The manufacturer is interested in different groups of athletes and how they compare in using this upload service. The company would like to know: what is the difference between the proportions of triathletes and cyclists who upload their workout data? Random, independent samples of triathletes and cyclists were taken. In a sample of 301 triathletes, 239 used the upload service. In a sample of 278 cyclists, 191 used the upload service.

Construct and interpret a 90% confidence interval estimate for the difference between the proportions of triathletes and cyclists that use the upload service. Use Triathletes as Population 1. Can you confirm that there is indeed a difference between the proportions, at the 90% confidence level? If so, which group has a higher proportion who uploads their data?