Statistics: The Null and Alternate Hypotheses
A Student Academic Learning Services Guide
The Null and Alternate Hypotheses: before we begin

- The Null and Alternate Hypothesis statements are important parts of the analytical methods collectively known as inferential statistics.
- Inferential statistics are methods used to determine something about a population, based on the observation of a sample.
- Information about a population will be presented in one of two forms, as a mean ($\mu$) or as a proportion ($p$).
- Use the population mean ($\mu$) in the hypothesis statements when the question gives you information about the population in the form of an average.
  - e.g. “the average travel time was 40 minutes…”, $\mu = 40$ minutes
- Use the population proportion ($p$) in the hypothesis statements when the question gives you information about the population in the form of a fraction, percentage, or decimal.
  - e.g. “4 out of 5 dentists agree…”, $p = \frac{4}{5}$ or $p = 80\%$ or $p = .80$

The Null Hypothesis: $H_0$

- Stating the Null Hypothesis is the starting point of any hypothesis testing question solution.
- When solving a problem, it is written as “$H_0$.”
- The Null Hypothesis is the stated or assumed value of a population parameter (the mean or proportion that is being analyzed).
  - What the question says the population is doing.
  - The current or reported condition.
- The necessary information tends to be in the first sentence of the problem.
- When trying to identify the population parameter needed for your solution, look for the following phrases:
  - “It is known that…”
  - “Previous research shows…”
  - “The company claims that…”
  - “A survey showed that…”
- When writing the Null Hypothesis, make sure it includes an “=” symbol. It may look like one of the following:
  - e.g. $H_0: \mu = 40$ minutes
  - e.g. $H_0: \mu \leq 40$ minutes
  - e.g. $H_0: \mu \geq 40$ minutes

---

The Alternate Hypothesis: $H_1$

- The Alternate Hypothesis accompanies the Null Hypothesis as the starting point to answering hypothesis testing questions.
- When solving a problem, it is written as “$H_1$.”
- The Alternate Hypothesis is the stated or assumed value of a population parameter if the Null Hypothesis ($H_0$) is rejected (through testing).
- The necessary information tends to be found in the last sentence of the problem (or the sentence ending in a “?”).
- When trying to identify the information needed for your Alternate Hypothesis statement, look for the following phrases:
  - “Is it reasonable to conclude…”
  - “Is there enough evidence to substantiate…”
  - “Does the evidence suggest…”
  - “Has there been a significant…”
- There are three possible symbols to use in the Alternate Hypotheses, depending on the wording of the question.
- Use “$\neq$” when the question uses words/phrases such as:
  - “is there a difference…?”
  - “is there a change…?”
- Use “$<$” when the question uses words/phrases such as:
  - “is there a decrease…?”
  - “is there less…?”
  - “are there fewer…?”
- Use “$>$” when the question uses words/phrases such as:
  - “is there a increase…?”
  - “is there more…?”
- When writing the Alternate Hypothesis, make sure it never includes an “=” symbol. It should look similar to one of the following:
  - e.g. $H_1$: $\mu < 40$ minutes
  - e.g. $H_1$: $\mu > 40$ minutes
  - e.g. $H_1$: $\mu \neq 40$ minutes
Reading the Question

Here is an example problem to demonstrate the process of creating Null and Alternate Hypothesis statements.

Example

A recent survey of college campuses across Ontario claims that students spend an average of 2.7 hours a day using their cell phones. A random sample of 35 Durham College students showed an average use of 2.9 hours a day, with a standard deviation of 0.4 hours. Do Durham College students use their cell phones more than the typical Ontario college student?

Step 1: Find the population information

- Read the question carefully and try and find information that is being presented as, or claims to be, fact.
- In the first sentence we see the phrases “A recent survey…” and “claims that…” (both are good indicators that the information we need is in that sentence)
- Next, determine if you are working with a population average (µ) or population proportion (p)
- The information is given to us in the form of an average (2.7 hours) so we know we will use µ in the Null and Alternate Hypothesis statements
- So far the Null and Alternate Hypothesis statements look like this:

\[
H_0: \mu = 2.7 \text{ hours} \\
H_1: \mu \neq 2.7 \text{ hours}
\]

Step 2: Determine the operators (math symbols)

- Read the question carefully and find the sentence that ends in “?”. It is often (but not always) the last sentence of the problem
- Examine the wording of the question sentence, looking for words/phrases that indicate which operator to use
- The example question asks, “Do Durham College students use their cell phones more than the typical Ontario college student?”
- Because the phrase “more than” is used in the question, we will use the greater than symbol (>)
- The Null and Alternate Hypothesis statements now look like this:

\[
H_0: \mu = 2.7 \text{ hours} \\
H_1: \mu > 2.7 \text{ hours}
\]
The Null and Alternate Hypothesis statements must oppose each other.

So if the Alternate Hypothesis uses the greater than symbol (>), the Null Hypothesis statement must use the less than symbol (<).

However, we know that the Null Hypothesis statement must also include an equals symbol (=).

Combining the less than symbol and the equals symbol we have the “less than or equals to” symbol (≤).

The Null and Alternate Hypothesis statements now look like this:

\[ H_0: \mu \leq 2.7 \text{ hours} \]
\[ H_1: \mu > 2.7 \text{ hours} \]

**Problem Types**

During your course you will be asked to analyze and solve a variety of different hypothesis testing questions. Listed below are some of the problem types you may encounter, and what the Null and Alternate Hypothesis statements might look like.

**One Sample**

- A comparison of sample data to the stated population information to determine if the stated population information is still true.
- Use the population symbols in your hypothesis statements (μ and p).
  - e.g. \( H_0: \mu = 40 \text{ hours} \)
  - \( H_1: \mu \neq 40 \text{ hours} \)
  - e.g. \( H_0: p = .20 \)
  - \( H_1: p \neq .20 \)

**Two Samples**

- A comparison of data from one sample to data from a different sample to determine if the two populations they came from are the same.
- Use the population symbols in your hypothesis statements (μ and p).
- When comparing multiple samples, the Null and Alternate Hypothesis statements can be written two ways. Be sure to check with your instructor as to which is preferred.
  - e.g. \( H_0: \mu_A = \mu_B \) or \( H_0: \mu_A - \mu_B = 0 \)
  - \( H_1: \mu_A \neq \mu_B \) or \( H_1: \mu_A - \mu_B \neq 0 \)
  - e.g. \( H_0: p_A = p_B \) or \( H_0: p_A - p_B = 0 \)
  - \( H_1: p_A \neq p_B \) or \( H_1: p_A - p_B \neq 0 \)
More than Two Samples (ANOVA)

- A comparison of sample data across more than two samples or “treatments” to determine if the populations are the same
- When performing an ANOVA, you may be asked to comment on the variation/variance of the samples or the means of the samples. Be sure to look for the following statements to determine what symbols to use in your hypothesis statements
  - When...“Is there (more/less/difference) variation...”
    - Use the population variance symbol ($\sigma^2$) in the hypothesis statements
    - e.g. $H_0: \sigma_A^2 = \sigma_B^2$ or $H_1: \sigma_A^2 \neq \sigma_B^2$ or $H_0: \sigma_A^2 - \sigma_B^2 = 0$ or $H_1: \sigma_A^2 - \sigma_B^2 \neq 0$
  - When...“Is there difference in the mean/average...”
    - Use the population mean symbol ($\mu$) in the hypothesis statements
    - e.g. $H_0: \mu_A = \mu_B = \mu_C$
      - $H_1$: the means are not equal

Linear Regression

- An analysis of the relationship between two variables within a sample to determine the affect changing one of them (the independent variable) has on the other (dependent variable)
- The Null and Alternate Hypothesis statements use the population correlation coefficient ($\rho$) instead of the population mean, proportion, or variance
  - Note that this symbol is called “rho” (sounds like “row”). Although it looks like the letter “p” it is not, and has a very different meaning
  - e.g. $H_0: \rho = 0$
    - $H_1: \rho \neq 0$
Multiple Regression

- An analysis of the relationship between multiple variables within a sample to determine the relationship (strength and nature) of those variables. In multiple regression there is a single dependent variable, but multiple independent variables.
- There are usually more than one set of hypothesis statements needed to complete the problem when performing a multiple regression analysis.
- The first set is used when performing a “global test” to see if there is a relationship between any of the independent variables and the dependent variable.
- The Null and Alternate Hypothesis statements use the symbol $\beta$ to represent the net regression coefficients in the population. There is a $\beta$ for each independent variable in the problem.
  - e.g. $H_0: \beta_1 = \beta_2 = \beta_3 = 0$
  - $H_1$: not all $\beta$ are 0
- Later in a problem you may be asked to test each independent variable’s regression coefficient on its own. In this case, you will create Null and Alternate Hypothesis statements for each independent variable.
  - e.g. $H_0: \beta_1 = 0$
  - $H_1: \beta_1 \neq 0$

Chi-Squared ($\chi^2$)

- A comparison to observed data to expected data.
- Symbols are not used in the Null and Alternate Hypothesis statements.
  - e.g. $H_0$: there is NO difference between the (observed frequency) and the (expected frequency)
  - $H_1$: there IS a difference between the (observed frequency) and the (expected frequency)