Chapter 3: Introduction to Hypothesis Testing

Vocabulary

Population: The collection of all people or objects to be studied.

Sample: Collecting data from a small subgroup of the population.

Random Sample: Sample data collected in such a way that everyone in the population has an equal chance to be included.

Statistic: A number calculated from sample data in order to understand the characteristics of the data. For example, a sample mean average, a sample standard deviation, or a sample percentage.

Parameter: A number that describes the characteristics of a population like a population mean or a population percentage. Can be calculated from an unbiased census, but is often just a guess about the population.

Hypothesis Test: A procedure for testing a claim about a population.

Null Hypothesis (H_0): A statement about the population that involves equality. It is often a statement about "no change", "no relationship" or "no effect".

Alternative Hypothesis $(H_A \text{ or } H_1)$: A statement about the population that does not involve equality. It is often a statement about a "significant difference", "significant change", "relationship" or "effect".

Population Claim: What someone thinks is true about a population.

Test Statistic: A number calculated in order to determine if the sample data significantly disagrees with the null hypothesis. There are a variety of different test statistics depending on the type of data.

One-Population Proportion Test Statistic (z): The sample proportion is this many standard errors above or below the population proportion in the null hypothesis.

One-Population Mean Test Statistic (t): The sample mean is this many standard errors above or below the population mean in the null hypothesis.

Critical Value: A number we compare our test statistic to in order to determine significance. In a sampling distribution or a theoretical distribution approximating the sampling distribution, the critical value shows us where the tail or tails are. The test statistic must fall in the tail to be significant.

Sampling Variability: Also called "random chance". The principle that random samples from the same population will usually be different and give very different statistics. The random samples will usually be different than the population parameter.

P-value: The probability of getting the sample data or more extreme because of sampling variability (by random chance) if the null hypothesis is true.

Significance Level (α): Also called the Alpha Level. This is the probability of making a type 1 error. The P-value is compared to this number to determine significance and sampling variability. If the P-value is lower than the significance level, then the sample data significantly disagrees with the null hypothesis and is unlikely to have happened because of sampling variability.

Randomized Simulation: A technique for visualizing sampling variability in a hypothesis test. The computer assumes the null hypothesis is true, and then generates random samples. If the sample data or test statistic falls in the tail, then the sample data significantly disagrees with the null hypothesis. This technique can also calculate the P-value and standard error without a formula.

Type 1 Error: When biased sample data leads you to support the alternative hypothesis when the alternative hypothesis is actually wrong in the population.

Type 2 Error: When biased sample data leads you fail to reject the null hypothesis when the null hypothesis is actually wrong in the population.

Beta Level (β): The probability of making a type 2 error.

Conclusion: A final statement in a hypothesis test that addresses the claim and evidence.

Introduction: The goal of statistics is to learn about the world around us. While we may collect sample data, our goal is not just to analyze sample data. We want to know what is happening in the population. Sometimes, people make guesses about what they think is happening in the population. These guesses are often called "claims". In this chapter, we will discuss the process of hypothesis testing. A hypothesis test is a scientific procedure for using representative random sample data to investigate claims about populations. A hypothesis test involves many difficult concepts and has many steps. It is difficult to learn all of hypothesis testing at once. For this reason, we will be learning the steps for hypothesis testing one at a time. Eventually we will put all of the parts together and complete the hypothesis test from start to finish.

Section 3A – Null and Alternative Hypothesis

Vocabulary

Population: The collection of all people or objects to be studied.

Census: Collecting data from everyone in a population.

Sample: Collecting data from a small subgroup of the population.

Statistic: A number calculated from sample data in order to understand the characteristics of the data. For example, a sample mean average, a sample standard deviation, or a sample percentage.

Parameter: A number that describes the characteristics of a population like a population mean or a population percentage. Can be calculated from an unbiased census, but is often just a guess about the population.

Hypothesis Test: A procedure for testing a claim about a population.

Null Hypothesis (H_0): A statement about the population that involves equality. It is often a statement about "no change", "no relationship" or "no effect".

Alternative Hypothesis (H_A *or* H_1): A statement about the population that does not involve equality. It is often a statement about a "significant difference", "significant change", "relationship" or "effect".

Population Claim: What someone thinks is true about a population.

Introduction:

Remember, a hypothesis test is a procedure for checking what someone has said about the population. This is often called the population "claim". From this claim, we will need to identify two opposing views. The claim and the rival hypothesis. If the claim is not true, then what would be true? These two opposing views are referred to as the null hypothesis (H_0) and the alternative hypothesis (H_A *or* H_1). The symbol often used for "null hypothesis" is "H_0". The symbol used to represent the alternative hypothesis is "H_A" or "H_1". I prefer "H_A" for alternative hypothesis.

Null Hypothesis (H_0): A statement about the population that involves equality. It is often a statement about "no change", "no relationship" or "no effect".

Alternative Hypothesis (H_A): A statement about the population that does not involve equality. It is often a statement about a "significant difference", "significant change", "relationship" or "effect".

Important Notes:

- In the last chapter, we talked about symbols that represent population parameters like the population mean (μ), the population proportion (π) or the population standard deviation (σ). The null and alternative hypothesis are competing ideas about the population and only involve population parameters like μ, π or σ. The sample data and sample statistics like the sample proportion ($\hat{p}$), the sample mean ($\bar{x}$) or the sample standard deviation (*s*) are never part of the null or alternative hypotheses.

A hypothesis test is a procedure for deciding between two opposing views about the population. Sometimes the person will tell you the two opposing views, but there will be one view that the person thinks is true or wants you as the data scientist to give evidence toward. This is called the "claim".

Claim: What the person now thinks is true about the population. The claim can be a question that someone needs to figure out. It can also be an opinion about the population that they want you to investigate.

Let's go over some basic steps to writing the null and alternative hypothesis.

Steps for finding the Null and Alternative Hypothesis

1. *Write down the two competing views about the population in symbolic language. Make sure to determine if it is one-population, or two-population, or more and the correct letter (parameter) to use. If the person only gives you his or her claim, we will often use the opposite as the rival hypothesis or competing view.*
2. *Write the word "claim" next to what the person thinks is true or what they are asking you to provide evidence toward.*
3. *The statement that has "=" or "≥" or "≤" is the null hypothesis. It is usually "=". Put an "H_0" next to it. Remember, the null hypothesis is usually a statement about no change, no effect, or not related. That is why the null hypothesis is often given with "=".*
4. *The statement that has "≠" or "<" or ">" is the alternative hypothesis. Put an "H_A" next to it. This is usually a statement about something changing or being related or having a significant effect.*
5. *Determine what type of test you are dealing with. Is it right-tailed, left-tailed or two-tailed? In a hypothesis test, we are often attempting to "reject the null hypothesis". This would happen if we believed the alternative hypothesis is correct. For this reason, the alternative hypothesis determines what type of test you are doing. In a one-population or two-population test, if H_A is less than (<), it is a left tailed test. (Notice "<" points to the left). If H_A is greater than (>), it is a right tailed test. (Notice ">" points to the right). If H_A is not equal (≠), it is a two-tailed test. (not equal means less than or greater than.) Things get more complicated if we are dealing with 3 or more populations. We will deal with these cases in the next chapter.*

    - *Left-tailed test: H_A is less than (<)*
    - *Right-tailed test: H_A is greater than (>)*
    - *Two-tailed test: H_A is less than (≠)*

Symbols for population parameters:

μ *(population mean)*
π or p *(population proportion/percentage)*
σ *(population standard deviation)*
σ^2 *(population variance)*

Important Notes

- *Never use a sample statistic* $(\bar{x}, \hat{p}, s)$. *Remember a null and alternative hypothesis are statements about the population parameters* (μ, π, σ).
- *The symbol "<" means less than. Notice the symbol for "less than" points to the left. The opposite of less than is "≥" (greater than or equal to).*
- *The symbol ">" means less than. Notice the symbol for "greater than" points to the right. The opposite of less than is "≤" (less than or equal to).*
- *For one-population, always write the letter on the left side and the number on the right side. μ < 50 (never as 50 > μ)*
- *For two-population, always put population 1 on the left side and population 2 on the right side.* $\pi_1 > \pi_2$ *(never as* $\pi_2 < \pi_1$*).*

Three Types of Hypothesis Tests

- Hypothesis Tests are designated as one of three types. It is important to know what type of test you are doing.
- Note that the alternative hypothesis H_A decides the type of test. In a hypothesis test, we are often attempting to "reject the null hypothesis". This would happen if we believed the alternative hypothesis is correct. For this reason, the alternative hypothesis determines what type of test you are doing. In a one-population or two-population test, if H_A is less than (<), it is a left tailed test. (Notice "<" points to the left). If H_A is greater than (>), it is a right tailed test. (Notice ">" points to the right). If H_A is not equal (≠), it is a two-tailed test. (not equal means less than or greater than.)

Left-tailed test: H_A is less than (<) (points to the left)

Right-tailed test: H_A is greater than (>) (points to the right)

Two-tailed test: H_A is less than (≠)

Note: If you have more than two populations, things become more complicated. For example, ten populations will often be condensed into one right-tailed test. We will explain more about these tests in the next chapter.

Example 1: Auto Magazine Article

"The population mean average weight of car transmissions used to be about 300 pounds. With more and more small car options, we think the population mean average weight of car transmissions has decreased."

Step 1: The population views are about one population mea (μ). *Notice there are two opposing views are the population mean given. Write down the two opposing views in symbolic language. Notice again that "decreased" means less than "<" which points to the left. Make sure to put the symbol on the left and the number on the right. ("μ < 300" NOT "300 > μ"). Notice also that the opposing view is that it used to be exactly 300 pounds. They never believed that the population mean average was more than 300 pounds. It seems more appropriate to use "=" as the opposing view instead of "≥".*

μ = 300
μ < 300

Step 2: Let's identify the claim. It seems they think that it used to be 300 pounds, but now they think the population mean is really less than 300 pounds. This would be the claim.

μ = 300
μ < 300 (Claim)

Step 3. The statement that has "=" or "≥" or "≤" is the null hypothesis. Put an "H_0" next to it. Notice that "claim" has nothing to do with a statement being a null or alternative hypothesis. It is the symbol that decides H_0.

H_0: μ = 300
μ < 300 (Claim)

Step 4: The statement that has "≠" or "<" or ">" is the alternative hypothesis. Put an "H_A" next to it. Notice that "claim" has nothing to do with a statement being a null or alternative hypothesis. It is the symbol that decides H_A.

H_0: μ = 300
H_A: μ < 300 (Claim)

Step 5: Determine the type of test you are doing. Remember, the alternative hypothesis determines the type of test. If H_A is less than (<), it is a left tailed test. (Notice "<" points to the left). If H_A is greater than (>), it is a right tailed test. (Notice ">" points to the right). If H_A is not equal (≠), it is a two-tailed test. (not equal means less than or greater than.)

In this problem, the alternative hypothesis is less than (<) which points to the left. So this is a left tailed test.

H_0: μ = 300
H_A: μ < 300 (Claim)

Left Tailed Test

Example 2: Medication Side Effects

"The FDA says that about 2.5% of people that take this medicine will have serious side effects".

Step 1: The population views are about one population proportion (π or p). *Notice that only one view about the population proportion is given. In this case we will have to think about opposites. The FDA seems to be making a claim about equaling 2.5% (0.025). They did not specify higher or lower. The opposite of equal is not equal* (≠) *so we will use that as our opposing view. Write down the two opposing views in symbolic language. Population percentage claims are usually written as a decimal proportion. Make sure to put the symbol on the left and the number on the right.("π = 0.025" NOT "0.025 = π").*

π = 0.025
π ≠ 0.025

OR

p = 0.025
p ≠ 0.025

Step 2: Let's identify the claim. It seems that the FDA thinks it is 2.5%. Don't let the word "about" change your mind about this. Remember, we never know anything definite about millions of people. Language like "about" or "around" is often used in population claims. They did not specify higher or lower, so it is still an equal to claim.

π = 0.025 (Claim)
π ≠ 0.025

OR

p = 0.025 (Claim)
p ≠ 0.025

Step 3. The statement that has "=" or "≥" or "≤" is the null hypothesis. Put an "H_0" next to it. Notice that "claim" has nothing to do with a statement being a null or alternative hypothesis. It is the symbol that decides H_0.

H_0: π = 0.025 (Claim)
π ≠ 0.025

OR

H_0: p = 0.025 (Claim)
p ≠ 0.025

Step 4: The statement that has "≠" or "<" or ">" is the alternative hypothesis. Put an "H_A" next to it. Notice that "claim" has nothing to do with a statement being a null or alternative hypothesis. It is the symbol that decides H_A.

H_0: π = 0.025 (Claim)
H_A: π ≠ 0.025

OR

H_0: p = 0.025 (Claim)
H_A: p ≠ 0.025

Step 5: Determine the type of test you are doing. Remember, the alternative hypothesis determines the type of test. If H_A is less than (<), it is a left tailed test. (Notice "<" points to the left). If H_A is greater than (>), it is a right tailed test. (Notice ">" points to the right). If H_A is not equal (≠), it is a two-tailed test. (not equal means less than or greater than.)

In this problem, the alternative hypothesis is not equal (≠), so this is a two-tailed test.

H_0: π = 0.025 (Claim)
H_A: π ≠ 0.025

OR

H_0: p = 0.025 (Claim)
H_A: p ≠ 0.025

Two-Tailed Test

Example 3: Comparing female and male SAT scores

The school board claims that the average SAT score for female high school students is greater than the average SAT score for male high school students. If gender is not related to SAT scores, then the SAT scores should be the same.

Step 1: Since this is a two-population mean average problem we will need to decide what is population 1 and what is population 2 and the correct letter to use. I tend to write statements as they are claimed. In this problem, they said that the population mean average SAT for females is higher than males. The most straight forward way is to make females population 1 and males population 2 and then say females is higher than males. We could reverse it and put males as population 1, but then we would have to say that population 1 is lower than population 2. Remember that the population 1 parameter should always go on the left.

μ_1*: Female*
μ_2*: Male*

Write down the two opposing views in symbolic language. Notice that the two opposing views indicate a gender / SAT relationship (the mean average for females is higher than males) verses no relationship (mean averages are the same.) Notice "equal to" goes with not related. This is why not related is always the null hypothesis. Even though the opposite of > is "≤", this symbol does not seem appropriate for this test. If the female SAT's were lower, that also would indicate a gender/SAT relationship.

$\mu_1 > \mu_2$

$\mu_1 = \mu_2$

Step 2: Decide the claim. It seems that while the school board longs for the day when the genders are the same, they do not believe that is true right now. They believe that gender does matter. They believe that the population mean average for females is higher than for males.

$\mu_1 > \mu_2$ (Claim)

$\mu_1 = \mu_2$

Step 3: Decide which statement is the null hypothesis. Remember, the statement that has "=" or "≥" or "≤" is the null hypothesis. Put an "H_0" next to it. Remember, claim has nothing to do with it.

$\mu_1 > \mu_2$ (Claim)

H_0: $\mu_1 = \mu_2$

Step 4: Decide which statement is the alternative hypothesis. Remember, the statement that has "≠" or "<" or ">" is always the alternative hypothesis. Put an "H_A" next to it. Claim has nothing to do with a statement being the alternative hypothesis. It is the symbol.

H_A: $\mu_1 > \mu_2$ (Claim)

H_0: $\mu_1 = \mu_2$

Step 5: Determine the type of test you are doing. Remember, the alternative hypothesis determines the type of test. If H_A is less than (<), it is a left tailed test. (Notice "<" points to the left). If H_A is greater than (>), it is a right tailed test. (Notice ">" points to the right). If H_A is not equal (≠), it is a two-tailed test. (not equal means less than or greater than.)

In this problem, the alternative hypothesis is greater than (>), which points to the right. So this is a right-tailed test.

Null Hypothesis Confusion

Some students get confused over the sign in the null hypothesis. It is important to pay attention to the language. Look at the following two examples.

Claim: "The population standard deviation (σ) used to be 2 inches but now we think it has increased."

Since we have two opposing views, we can write them both and then chose the null and alternative hypothesis by the sign.

H_0: σ = 2
H_A: σ > 2 (Claim)

Claim: "We think that the population standard deviation is more than 2 inches."

This time, we do not have the opposing view, so we use opposites. The opposite of ">" is "≤".

H_0: σ ≤ 2
H_A: σ > 2 (Claim)

These two examples illustrate a point of confusion. Sometimes you may see the null hypothesis of the same hypothesis test written as "=" and sometimes it may be written with "≤" or "≥".

H_0: σ = 2
H_A: σ > 2 (Claim)

OR

H_0: σ ≤ 2
H_A: σ > 2 (Claim)

Either answer is ok. Notice the parameter is still 2 inches and they are both right-tailed tests. In all practicality, they are the same test.

Many scientists prefer to write the null almost always as "=". Remember the null in an experiment is usually "no change" or "no effect". Change of any kind is usually denoted by the alternative hypothesis.

Confusion about "At Least" or "At Most".

When we say we have at least $20, we mean the amount of money is greater than or equal to $20 (≥).
When we say we have at most $20, we mean the amount of money is less than or equal to $20 (≤).

If we stick to the language of "at least" or "at most" we would have to make them the null hypothesis.

If the claim was that the population mean amount of money is at least $20, we would write the following null and alternative hypothesis. Remember at least means greater than or equal to "≥". The opposing view would be less than "<". Since the alternative is less than, this would be a left tailed test.

H_0: μ ≥ 20 (Claim)
H_A: μ < 20

This can create confusion. Does the person really want to check "at least" or do they really mean "more than"? It has been my experience that when people want to check an "at least" claim, they are better off changing the claim to "more than" or "increased" and doing a traditional right tailed test.

Similarly if someone wants to test an "at most" claim, they are better off changing the claim to "less than" or "decreased" and doing a traditional left tailed test.
