P-Values and the Null Hypothesis

Introduction

In this lesson, you'll learn about the relationship between p-values and the Null Hypothesis, and their role in designing an experiment.

Objectives

You will be able to:

- Describe what it means to "reject the null hypothesis" and how it is related to p-value

Understanding The Null Hypothesis

As stated previously, scientific experiments actually have 2 hypotheses:

Null Hypothesis: There is no relationship between A and B
Example: "There is no relationship between this flu medication and a reduced recovery time from the flu".

The *Null Hypothesis* is usually denoted as H_0

Alternative Hypothesis: The hypothesis traditionally thought of when creating a hypothesis for an experiment
Example: "This flu medication reduces recovery time for the flu."

The *Alternative Hypothesis* is usually denoted as H_1

An easy way to differentiate between the Null Hypothesis and the Alternative Hypothesis is that the Null Hypothesis is the more conservative choice. It always assumes that there is no difference between two different population means, and when it is represented mathematically, it should always contain an equals sign.

The Alternative Hypothesis is whatever claim you are trying to prove with an experiment.

P-Values and Alpha Values

No matter what you're experimenting on, good experiments come down to one question: Is your p-value less than your alpha value? Let's dive into what each of these values represents, and why they're so important to experimental design.

p-value: The probability of observing a test statistic at least as large as the one observed, by random chance, assuming that the null hypothesis is true.

If you calculate a p-value and it comes out to 0.03, you can interpret this as saying "There is a 3% chance of obtaining the results I'm seeing when the null hypothesis is true."

α **(alpha value)**: The marginal threshold at which you're okay with rejecting the null hypothesis.

An alpha value can be any value set between 0 and 1. However, the most common alpha value in science is 0.05 (although this is somewhat of a controversial topic in the scientific community, currently).

If you set an alpha value of $\alpha = 0.05$, you're essentially saying "I'm okay with accepting my alternative hypothesis as true if there is less than a 5% chance that the results that I'm seeing are actually due to randomness."

When you conduct an experiment, your goal is to calculate a p-value and compare it to the alpha value. If $p < \alpha$, then you **reject the null hypothesis** and accept that there is not "no relationship" between the dependent and independent variables. Note that any good scientist will admit that this doesn't prove that there is a *direct relationship* between the dependent and independent variables--just that they now have enough evidence to the contrary to show that they can no longer believe that there is no relationship between them.

In simple terms:

$p < \alpha$: Reject the *Null Hypothesis* and accept the *Alternative Hypothesis*

$p >= \alpha$: Fail to reject the *Null Hypothesis*.

There are many different ways that you can structure a hypothesis statement, but they always come down to this comparison in the end. In normally distributed data, you calculate p-values from t-statistics or (z-scores if the population parameters are known). This

is done a bit differently with discrete data. You may also have **One-Tail** and **Two-Tail** tests.

A **One-Tail Test** is when you want to know if a parameter from the treatment group is greater than (or less than) a corresponding parameter from the control group.

Example One-Tail Hypothesis

$H_1 : \mu_1 < \mu_2$ The treatment group given this weight loss drug will lose more weight on average than the control group that was given a competitor's weight loss drug

$H_0 : \mu 1 >= \mu_2$ The treatment group given this weight loss drug will not lose more weight on average than the control group that was given a competitor's weight loss drug".

A **Two-Tail Test** is for when you want to test if a parameter falls between (or outside of) a range of two given values.

Example Two-Tail Hypothesis

$H_1 : \mu_1 \neq \mu_2$ "People in the experimental group that are administered this drug will not lose the same amount of weight as the people in the control group. They will be heavier or lighter".

$H_0 : \mu_1 = \mu_2$ "People in the experimental group that are administered this drug will lose the same amount of weight as the people in the control group."

What Does an Experiment Really Prove?

You may be wondering why you need a **Null Hypothesis** at all. This is a good question. It has to do with being honest about what an experiment actually proves.

Scientists use the **Null Hypothesis** so that they can be very specific in their findings. This is because a successful experiment doesn't actually *prove a relationship* between a dependent and independent variable. Instead, it just proves that there is not enough evidence to convincingly believe there is *no relationship* between the dependent and the independent variable. There can always be a lurking variable behind the scenes that is actually responsible for the relationship between two variables--it's almost impossible to cover every possible angle. However, a successful experiment where a p-value is less than an alpha value (typically, $p < 0.05$) does give enough information to confidently allow someone to say that it's statistically unlikely that there is *no relationship* between the two, which is what would have to be true in order for the null hypothesis to be correct!

The Null Hypothesis Loves You and Wants You To Be Happy

You've covered a lot about the null hypothesis and how it's used in experiments in this lesson, but there's a lot more to learn about it!

Read the following article, The Null Hypothesis Loves You and Wants You To Be Happy. This does an excellent job of explaining why the concept of the *Null Hypothesis* is crucial to good science.

Summary

In this lesson, you learned about the relationship between p-values and the Null Hypothesis. Now you'll see how effect sizes affect your tests!