

7.4: Null Hypothesis Significance Testing

Null Hypotheses and Research Hypotheses

So far, so good? We develop a directional research hypothesis that names our groups, the DV (the outcome that was measured), and indicates a direction (which group will be higher). And we have a null hypothesis that says that the groups will have similar means on the DV. It's at this point that things get somewhat counterintuitive. Because the null hypothesis seems to correspond to the opposite of what I want to believe, and then we focus exclusively on that, almost to the neglect of the thing I'm actually interested in (the research hypothesis). In our growth mindset example, the null hypothesis is that the sample of junior high students with high beliefs in growth mindset will have similar average study times compared to the population of all junior high students. But for Blackwell, Trzseniewski, and Dweck (2007), and, really, any teacher ever, we actually want to believe that the understanding that your intelligence and abilities can always improve (high belief in growth mindset) will result in working harder and spending more time on homework. So the *alternative* to this null hypothesis is that those junior high students with higher growth mindset scores will spend more time on their math homework than those from the population of junior high students. **The important thing to recognize is that the goal of a hypothesis test is *not* to show that the research hypothesis is (probably) true; the goal is to show that the null hypothesis is (probably) false.** Most people find this pretty weird.

The best way to think about it, in my experience, is to imagine that a hypothesis test is a criminal trial... *the trial of the null hypothesis*. The null hypothesis is the defendant, the researcher is the prosecutor, and the statistical test itself is the judge. Just like a criminal trial, there is a presumption of innocence: the null hypothesis is *deemed* to be true unless you, the researcher, can prove beyond a reasonable doubt that it is false. You are free to design your experiment however you like, and your goal when doing so is to maximize the chance that the data will yield a conviction... for the crime of being false. The catch is that the statistical test sets the rules of the trial, and those rules are designed to protect the null hypothesis – specifically to ensure that if the null hypothesis is actually true, the chances of a false conviction are guaranteed to be low. This is pretty important: after all, the null hypothesis doesn't get a lawyer. And given that the researcher is trying desperately to prove it to be false, *someone* has to protect it.

Notation

Okay, so the null hypothesis always states that there's no difference. In our examples so far, we've been saying that there's no difference between the sample mean and population mean. But we don't *really* expect that, or why would we be comparing the means? The purpose of null hypothesis significance testing is to be able to **reject** the *expectation* that the means of the two groups are the *same*.

- Reject the null hypothesis: The sample mean is different from the population mean.
 - Rejecting the null hypothesis means that $\bar{X} \neq \mu$.
 - Rejecting the null hypothesis doesn't automatically mean that the research hypothesis is supported.
- Retain the null hypothesis: The sample mean is similar to the population mean.
 - Retaining the null hypothesis means that $\bar{X} = \mu$.
 - This means that our research hypothesis cannot be true.

We only reject or retain the null hypothesis. If we reject the null hypothesis (which says that everything is similar), we are saying that some means are statistically different from some other means. We only support the research hypothesis if the means are in the direction that we said. For example, if we rejected the null hypothesis that junior highers with high growth mindset spend as much time on homework as all junior highers, we can't automatically say that junior high students with high growth mindset study more than the population of junior high students. Instead, we'd have to look at the actual means of each group, and then decide if the research hypothesis was supported or not.

I hope that it's obvious that you don't have to look at the group means if the null hypothesis is retained?

In sum, you reject or retain the null hypothesis, and your support or or don't support the research hypothesis.

Why predict that two things are similar?

Because each sample's mean will vary around the population mean (see the first few sections of this chapter to remind yourself of this), we can't tell if our sample's mean is within a "normal" variance. But we *can* gather data to show that this sample's mean is different (enough) from the population's mean. This is rejecting the null hypothesis.

We use statistics to determine the probability of the null hypothesis being true.

? Exercise 7.4.1

Does a true null hypothesis say the sample mean and the population mean are similar or different?

Answer

A null hypothesis always says that the means are similar (or that there is no relationship between the variables).

Why can't we prove that the mean of our sample is different from the mean of the population? Remember the first few sections of this chapter, that showed how different samples from the same population have different means and standard deviations. Researchers are a conservative bunch; we don't want to stake our reputation on a sample mean that could be fluke, one of the extreme handfuls of green gumballs even when the mean difference between hands was zero.

But what we can show is that our sample is so extreme that it is statistically unlikely to be similar to the population.

Null hypothesis significant testing is like how courts decide if defendants are Guilty or Not Guilty, not their Guilt v. Innocent. Similarly, we decide if the sample is similar to the population or not.

Summary

This is a tough concept to grasp, so we'll keep working on it. And if you never get it, that's okay, too, as long as you remember the pattern of rejecting or retaining the null hypothesis, and supporting or not supporting the research hypothesis.

This page titled [7.4: Null Hypothesis Significance Testing](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Michelle Oja](#).

- [11.1: A Menagerie of Hypotheses](#) by [Danielle Navarro](#) is licensed [CC BY-SA 4.0](#). Original source: <https://bookdown.org/ekothe/navarro26/>.
- [Current page](#) by [Michelle Oja](#) is licensed [CC BY-SA 4.0](#).