

# A step-by-step guide to hypothesis testing

Published on November 8, 2019 by [Rebecca Bevans](#). Revised on October 29, 2021.

Hypothesis testing is a formal procedure for investigating our ideas about the world using [statistics](#). It is most often used by scientists to test specific predictions, called hypotheses, that arise from theories.

There are 5 main steps in hypothesis testing:

1. State your [research hypothesis](#) as a null ( $H_0$ ) and alternate ( $H_a$ ) hypothesis.
2. [Collect data](#) in a way designed to test the hypothesis.
3. Perform an appropriate [statistical test](#).
4. Decide whether to reject or fail to reject your null hypothesis.
5. Present the findings in your [results](#) and [discussion](#) section.

Though the specific details might vary, the procedure you will use when testing a hypothesis will always follow some version of these steps.

## Step 1: State your null and alternate

## hypothesis

### ☰ Table of contents

that you want to investigate), it is important to restate it as a null ( $H_0$ ) and alternate ( $H_a$ ) hypothesis so that you can test it mathematically.

The **alternate hypothesis** is usually your initial hypothesis that predicts a relationship between [variables](#). The **null hypothesis** is a prediction of no relationship between the variables you are interested in.

You want to test whether there is a relationship between gender and height. Based on your knowledge of human physiology, you formulate a hypothesis that men are, on average, taller than women. To test this hypothesis, you restate it as:

$H_0$ : Men are, on average, not taller than women.

$H_a$ : Men are, on average, taller than women.

## Step 2: Collect data

For a statistical test to be [valid](#), it is important to perform [sampling](#) and collect data in a way that is designed to test your hypothesis. If your data are not representative, then you cannot make statistical inferences about the population you are interested in.

To test differences in average height between men and women, your sample should have an equal proportion of men and women, and cover a variety of socio-economic classes and any other variables that might influence average height.

You should also consider your scope (Worldwide? For one country?) A potential data source in this case might be

census data, since it includes data from a variety of regions

## ☰ Table of contents

the world.

### Receive feedback on language, structure and layout

Professional editors proofread and edit your paper by focusing on:

- ✓ Academic style
- ✓ Vague sentences
- ✓ Grammar
- ✓ Style consistency

[See an example](#)

## Step 3: Perform a statistical test

There are a variety of [statistical tests](#) available, but they are all based on the comparison of **within-group variance** (how spread out the data is within a category) versus **between-group variance** (how different the categories are from one another).

If the between-group [variance](#) is large enough that there is little or no overlap between groups, then your statistical test will reflect that by showing a low [p-value](#). This means it is unlikely that the differences between these groups came about by chance.

Alternatively, if there is high within-group variance and low between-group variance, then your statistical test will reflect that with a high [p-value](#). This means it is likely that any difference you measure between groups is due to chance.

Your choice of statistical test will be based on the [type of data](#) you

collected.

## ☰ Table of contents

Based on the type of data you collected, you perform a one-tailed **t-test** to test whether men are in fact taller than women. This test gives you:

- an estimate of the difference in average height between the two groups.
- a  $p$ -value showing how likely you are to see this difference if the null hypothesis of no difference is true.

Your t-test shows an average height of 175.4 cm for men and an average height of 161.7 cm for women, with an estimate of the true difference ranging from 10.2cm to infinity. The  $p$ -value is 0.002.

## Step 4: Decide whether to reject or fail to reject your null hypothesis

Based on the outcome of your statistical test, you will have to decide whether to reject or fail to reject your null hypothesis.

In most cases you will use the  $p$ -value generated by your statistical test to guide your decision. And in most cases, your cutoff for rejecting the null hypothesis will be 0.05 – that is, when there is a less than 5% chance that you would see these results if the null hypothesis were true.

In your analysis of the difference in average height between men and women, you find that the  $p$ -value of 0.002 is below your cutoff of 0.05, so you decide to reject your null hypothesis of no difference.

## Step 5: Present your findings

The results of hypothesis testing will be presented in the results

and discussion sections of your research paper.

## ☰ Table of contents

and a summary of the results of your statistical test (for example, the estimated difference between group means and associated  $p$ -value). In the [discussion](#), you can discuss whether your initial hypothesis was supported by your results or not.

In the formal language of hypothesis testing, we talk about rejecting or failing to reject the null hypothesis. You will probably be asked to do this in your statistics assignments.

### **Stating results in a statistics assignment**

In our comparison of mean height between men and women we found an average difference of 14.3cm and a  $p$ -value of 0.002; therefore, we can reject the null hypothesis that men are not taller than women and conclude that there is likely a difference in height between men and women.

However, when presenting research results in academic papers we rarely talk this way. Instead, we go back to our alternate hypothesis (in this case, the hypothesis that men are on average taller than women) and state whether the result of our test was consistent or inconsistent with the alternate hypothesis.

If your null hypothesis was rejected, this result is interpreted as being consistent with your alternate hypothesis.

### **Stating results in a research paper**

We found a difference in average height between men and women of 14.3cm, with a  $p$ -value of 0.002, consistent with our hypothesis that there is a difference in height between men and women.

These are superficial differences; you can see that they mean the

same thing.

### ☰ Table of contents

**the alternate hypothesis.** This is because hypothesis testing is not designed to prove or disprove anything. It is only designed to test whether a pattern we measure could have arisen spuriously, or by chance.

If we reject the null hypothesis based on our research (i.e., we find that it is unlikely that the pattern arose by chance), then we can say our test **lends support to our hypothesis**. But if the pattern does not pass our decision rule, meaning that it could have arisen by chance, then we say the test is **inconsistent with our hypothesis**.

Is this article helpful?

1115

178

### Rebecca Bevans

Rebecca is working on her PhD in soil ecology and spends her free time writing. She's very happy to be able to nerd out about statistics with all of you.

## Other students also liked

### Statistical tests: which one should you use?

1454

choice of statistical test depends on the types of variables you're dealing

with and whether your data meets certain assumptions.

## Table of contents

### Understanding and calculating variance

The variance reflects the variability of your dataset by taking the average of squared deviations from the mean.

**421**

### The p-value explained

The p-value shows the likelihood of your data occurring under the null hypothesis. P-values help determine statistical significance.

**324**

#### **Scribbr**

Our editors

Jobs

FAQ

Partners

#### **Our services**

Plagiarism Checker

Proofreading & Editing

Citation Checker

APA Citation Generator

MLA Citation Generator

Knowledge Base

#### **Contact**

[info@scribbr.com](mailto:info@scribbr.com)

 +1 (510) 822-8066

[☰ Table of contents](#)

---

## 4.9

[Terms of use](#)

[Privacy Policy](#)

[Happiness Guarantee](#)