Determine if data has statistical significance

Recently, you learned that **statistical significance** is the claim that the results of a test or experiment are not explainable by chance alone. A hypothesis test can help you determine whether your observed data is statistically significant, or likely due to chance. For example, in a clinical trial of a new medication, a hypothesis test can help determine if the medication's positive effect on a sample group is statistically significant, or due to chance.

In this reading, you'll learn more about the concept of statistical significance and its role in hypothesis testing.

Statistical significance in hypothesis testing

Data professionals use hypothesis testing to determine whether a relationship between variables or a difference between groups is statistically significant.

Let's explore an example to get a better understanding of the role of statistical significance in hypothesis testing.

Example: Mean battery life

Let's review the steps for conducting a hypothesis test:

1. State the null hypothesis and the alternative hypothesis.
2. Choose a significance level.
3. Find the p-value.
4. Reject or fail to reject the null hypothesis.

Imagine you're a data professional working for a computer company. The company claims the mean battery life for their best selling laptop is 8.5 hours with a standard deviation of 0.5 hours. Recently, the engineering team redesigned the laptop to increase the battery life. The team takes a random sample of 40 redesigned laptops. The sample mean is 8.7 hours.

The team asks you to determine if the increase in mean battery life is statistically significant, or if it's due to random chance. You decide to conduct a z-test to find out.

Step 1: State the null hypothesis and alternative hypothesis

The null hypothesis typically assumes that your observed data occurs by chance, and it is not statistically significant. In this case, your null hypothesis says that there is no actual effect on mean battery life in the population of laptops.

The alternative hypothesis typically assumes that your observed data does *not* occur by chance, and is statistically significant. In this case, your alternative hypothesis says that there is an effect on mean battery life in the population of laptops.

In this example, you formulate the following hypotheses:

- **H_0**: $\mu = 8.5$ (the mean battery life of all redesigned laptops is equal to 8.5 hours)
- **H_a:** $\mu > 8.5$ (the mean battery life of all redesigned laptops is greater than 8.5 hours)

Step 2: Choose a significance level

The **significance level**, or alpha (α), is the threshold at which you will consider a result statistically significant. The significance level is also the probability of rejecting the null hypothesis when it is true.

Typically, data professionals set the significance level at 0.05, or 5%. That means results at least as extreme as yours only have a 5% chance (or less) of occurring when the null hypothesis is true.

Note: 5% is a conventional choice, and not a magical number. It's based on tradition in statistical research and education. Other common choices are 1% and 10%. You can adjust the significance level to meet the specific requirements of your analysis. A lower significance level means an effect has to be larger to be considered statistically significant.

Pro tip: As a best practice, you should set a significance level before you begin your test. Otherwise, you might end up in a situation where you are manipulating the results to suit your convenience.

In this example, you choose a significance level of 5%, which is the company's standard for research.

Step 3: Find the p-value

P-value refers to the probability of observing results as or more extreme than those observed when the null hypothesis is true.

Your p-value helps you determine whether a result is statistically significant. A low p-value indicates high statistical significance, while a high p-value indicates low or no statistical significance.

Every hypothesis test features:

- A test statistic that indicates how closely your data match the null hypothesis. For a z-test, your test statistic is a z-score; for a t-test, it's a t-score.
- A corresponding p-value that tells you the probability of obtaining a result at least as extreme as the observed result if the null hypothesis is true.

As a data professional, you'll almost always calculate p-value on your computer, using a programming language like Python or other statistical software. In this example, you're conducting a z-test, so your test statistic is a z-score of 2.53. Based on this test statistic, you calculate a p-value of 0.0057, or 0.57%.

Step 4: Reject or fail to reject the null hypothesis

In a hypothesis test, you compare your p-value to your significance level to decide whether your results are statistically significant.

There are two main rules for drawing a conclusion about a hypothesis test:

- If your p-value is less than your significance level, you reject the null hypothesis.

- If your p-value is greater than your significance level, you fail to reject the null hypothesis.

Note: Data professionals and statisticians always say “fail to reject” rather than “accept.” This is because hypothesis tests are based on probability, not certainty—acceptance implies certainty. In general, data professionals avoid claiming certainty about results based on statistical methods.
In this example, your p-value of 0.57% is less than your significance level of 5%. Your test provides sufficient evidence to conclude that the mean battery life of all redesigned laptops has increased from 8.5 hours. You reject the null hypothesis. You determine that your results are statistically significant.