One-tailed and two-tailed tests

Earlier, you learned that a hypothesis test can be either one-tailed or two-tailed. A tail in hypothesis testing refers to the tail at either end of a distribution curve.

In this reading, we'll go over the main differences between one-tailed and two-tailed tests, and discuss the procedure for conducting each test.

One-tailed and two-tailed tests

First, let's discuss the differences between one-tailed and two-tailed tests.

A **one-tailed test** results when the alternative hypothesis states that the actual value of a population parameter is either less than or greater than the value in the null hypothesis.

A one-tailed test may be either left-tailed or right-tailed. A left-tailed test results when the alternative hypothesis states that the actual value of the parameter is less than the value in the null hypothesis. A right-tailed test results when the alternative hypothesis states that the actual value of the parameter is greater than the value in the null hypothesis.

A **two-tailed test** results when the alternative hypothesis states that the actual value of the parameter does not equal the value in the null hypothesis.

For example, imagine a test in which the null hypothesis states that the mean weight of a penguin population equals 30 lbs.

- In a left-tailed test, the alternative hypothesis might state that the mean weight of the penguin population is less than ("<") 30 lbs.
- In a right-tailed test, the alternative hypothesis might state that the mean weight of the penguin population is greater than (">") 30 lbs.
- In a two-tailed test, the alternative hypothesis might state that the mean weight of the penguin population is not equal ("≠") to 30 lbs.

Let's explore a more detailed example to get a better understanding of the difference between one-tailed and two-tailed tests.

Example: One-tailed tests

Imagine you're a data professional working for an online retail company. The company claims that *at least* 80% of its customers are satisfied with their shopping experience. You survey a random sample of 100 customers. According to the survey, 73% of customers say they are satisfied. Based on the survey data, you conduct a z-test to evaluate the claim that *at least* 80% of customers are satisfied.

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

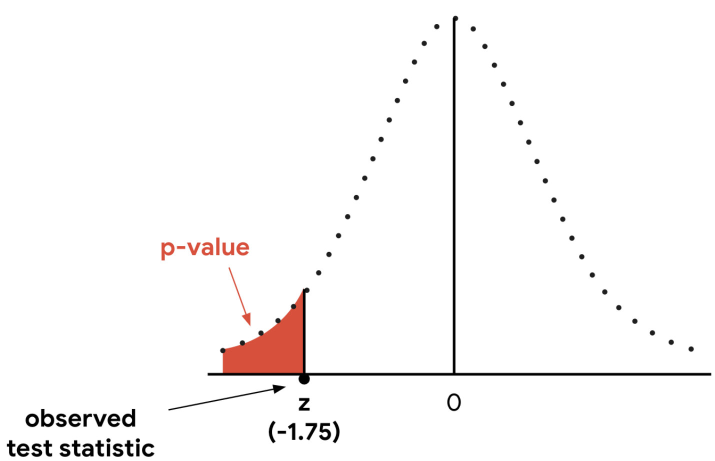
1. State the null hypothesis and the alternative hypothesis.
2. Choose a significance level.
3. Find the p-value.
4. Reject or fail to reject the null hypothesis.

First, you state the null and alternative hypotheses:

- **H_0**: P >= 0.80 (the proportion of satisfied customers is greater than or equal to 80%)
- **H_a**: P < 0.80 (the proportion of satisfied customers is less than 80%)

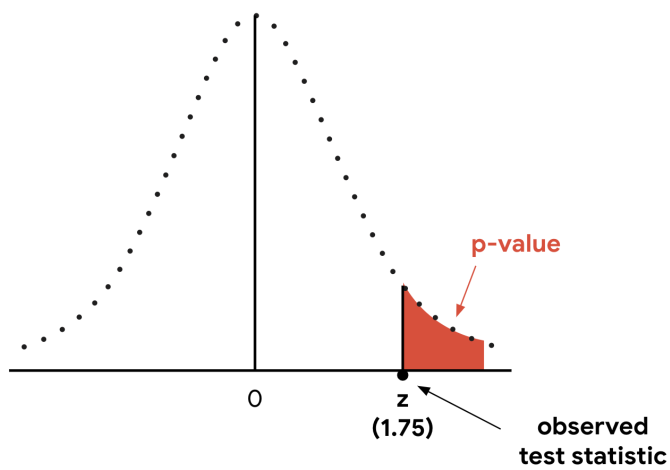
Note: This is a one-tailed test as the alternative hypothesis contains the less than sign ("<").
Next, you choose a significance level of 0.05, or 5%.
Then, you calculate your p-value based on your test statistic. Recall that p-value is the probability of observing results as or more extreme than those observed when the null hypothesis is true**.** In the context of hypothesis testing, "extreme" means extreme in the direction(s) of the alternative hypothesis.
Your test statistic is a z-score of 1.75 and your p-value is 0.04.
Since this is a left-tailed test, the p-value is the probability that the z-score is less than 1.75 standard units away from the mean to the left. In other words, it's the probability that the z-score is less than -1.75. The probability of getting a value less than your z-score of -1.75 is calculated by taking the area under the distribution curve to the left of the z-score. This is called a left-tailed test, because your p-value is located on the left tail of the distribution. The area under this part of the curve is the same as your p-value: 0.04.

Left-tailed test



Finally, you draw a conclusion. Since your p-value of 0.04 is less than your significance level of 0.05, you *reject* the null hypothesis.
Note: In a different testing scenario, your test statistic might be positive 1.75, and you might be interested in values as great or greater than the z-score 1.75. In that case, your p-value would be located on the right tail of the distribution, and you'd be conducting a right-tailed test.

Right-tailed test



Example: Two-tailed tests

Now, imagine our previous example has a slightly different set up. Suppose the company claims that 80% of its customers are satisfied with their shopping experience. To test this claim, you survey a random sample of 100 customers. According to the survey, 73% of customers say they are satisfied. Based on the survey data, you conduct a z-test to evaluate the claim that 80% of customers are satisfied.

First, you state the null and alternative hypotheses:

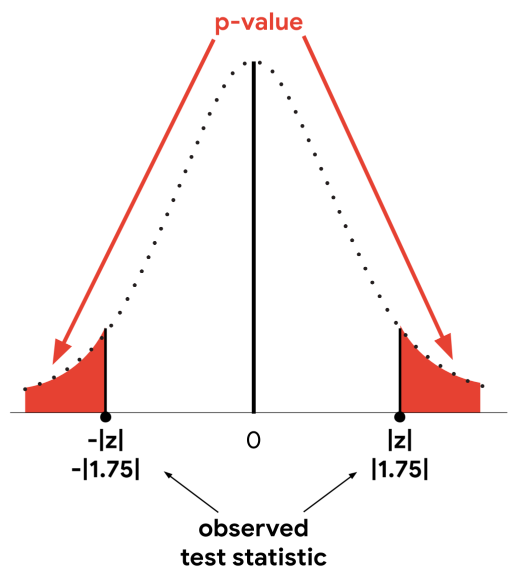
- **H**$_0$: P = 0.80 (the proportion of satisfied customers equals 80%)
- **H**$_a$: P ≠ 0.80 (the proportion of satisfied customers does not equal 80%)

Note: This is a two-tailed test as the alternative hypothesis contains the not equal sign ("≠").

Next, you choose a significance level of 0.05, or 5%.

Then, you calculate your p-value based on your test statistic. Your test statistic is a z-score of 1.75. Since this is a two-tailed test, the p-value is the probability that the z-score is less than -1.75 *or* greater than 1.75. Note that the *p*-value for a two-tailed test is always two times the *p*-value for a one-tailed test. So, in this case, your p-value = 0.04 + 0.04 = 0.08. In a two-tailed test, your p-value corresponds to the area under the curve on *both* the left tail and right tail of the distribution.

Two-tailed test



Finally, you draw a conclusion. Since your p-value of 0.08 is greater than your significance level of 0.05, you **fail to reject** the null hypothesis.

One-tailed versus two-tailed

You can use one-tailed and two-tailed tests to examine different effects.

In general, a one-tailed test may provide more power to detect an effect in a single direction. However, before conducting a one-tailed test, you should consider the consequences of missing an effect in the other direction. For example, imagine a pharmaceutical company develops a new medication they believe is more effective than an existing medication. As a data professional analyzing the results of the clinical trial, you may wish to choose a one-tailed test to maximize your ability to detect the improvement. In doing so, you fail to test for the possibility that the new medication is less effective than the existing medication. And, of course, the company doesn't want to release a less effective medication to the public.

A one-tailed test may be appropriate if the negative consequences of missing an effect in the untested direction are minimal. For example, imagine that the company develops a new, less expensive medication that they believe is at least as effective as the existing medication. The lower price gives the new medication an advantage in the market. So, they just want to make sure the new medication is not *less* effective than the existing medication. Testing whether it's *more* effective is not a priority. In this case, a one-tailed test may be appropriate.