

T-Tests

by Sophia



WHAT'S COVERED

In this tutorial, you will learn about t-tests, and how to determine key characteristics of a t-distribution. Our discussion breaks down as follows:

1. Comparing Z-Tests and T-Tests

1. Comparing Z-Tests and T-Tests

In a z-test for means, the z-test statistic is equal to the sample mean minus the hypothesized population mean, over the standard deviation of the population divided by the square root of sample size.



FORMULA TO KNOW

Z-Statistic For Population Means

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

However, the z-statistic was based on the fact that the population standard deviation was known. If the population standard deviation is *not* known, we need a new statistic. We're going to use our sample standard deviation, s, instead.



FORMULA TO KNOW

T-Statistic For Population Means

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$



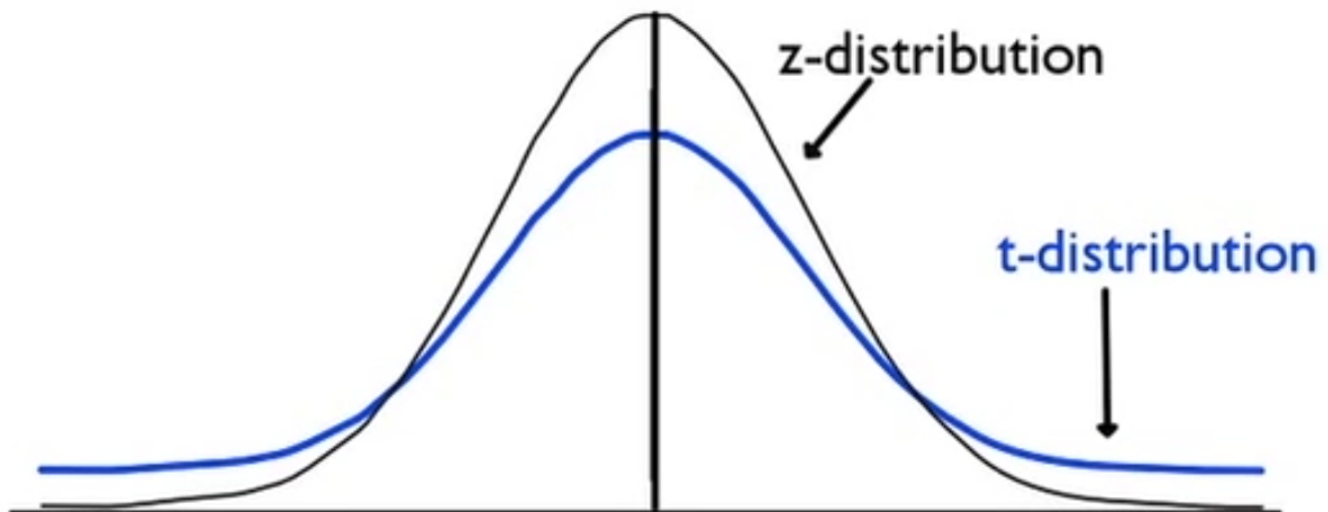
HINT

This "s" over the square root of n value, replacing the sigma over square root of n value, is called the standard error.

The only problem with using the sample standard deviation (s) as opposed to the population standard deviation (σ) is that the value of s can vary largely from sample to sample. Sigma (σ) is fixed, so we can base our normal distribution off of it.

The sample standard deviation is more variable than the population standard deviation and much more variable for small samples than for large samples. For large samples, s and sigma are very close, but with small samples particularly, the value of s can vary wildly.

Because s is so variable, it creates a new distribution of test statistics much like the normal distribution, but is known as the student's t-distribution, or sometimes just the **t-distribution**.



The only difference is the t-distribution is a more heavy-tailed distribution. If we used the normal distribution, it would underestimate the proportion of extreme values in the sampling distribution.

The t-distribution is actually a family of distributions. They all are a little bit shorter than the standard normal distribution and a little heavier on the tails. As the sample size gets larger, the t-distribution does get close to the normal distribution. It doesn't diminish as quickly in the tails when the sample size is small, but gets very close to the normal distribution when n is large (>30).



DID YOU KNOW

The t-test got its name from William Sealy Gosset, a chemist and statistician who worked at the Guinness Brewery in Dublin, Ireland, during the early 20th century. Gosset faced a common challenge: how to assess the quality of stout with small sample sizes.

When Gosset developed a novel statistical method, he did so under the pseudonym "Student". In 1908, he published his work in the esteemed scientific journal *Biometrika*, introducing what we now know as the Student's t-distribution and the associated t-test. The t-test became an economical way to monitor stout quality, but its impact extended far beyond the brewery.



TERM TO KNOW

T-Distribution

A family of distributions that are centered at zero and symmetric like the standard normal distribution, but heavier in the tails. Depending on the sample size, it does not diminish towards the tails as fast. If the sample size is large, the t-distribution approximates the normal distribution.



SUMMARY

In cases where the population standard deviation is not known--which is almost always--we should use the t-distribution to account for the additional variability introduced by using the sample standard deviation in the test statistic. A t-test means that the value will be a "t" statistic instead of a "z" statistic. The steps in the hypothesis test are the same as they are in a z-test: first stating the non-alternative hypotheses, stating and verifying the conclusions of the test, calculating the test statistic and the p-value, and then finally, comparing the p-value to alpha and making a decision about the null hypothesis.

Source: THIS TUTORIAL WAS AUTHORED BY JONATHAN OSTERS FOR SOPHIA LEARNING. PLEASE SEE OUR [TERMS OF USE](#).



TERMS TO KNOW

T-Distribution/Student's T-Distribution

A family of distributions that are centered at zero and symmetric like the standard normal distribution, but heavier in the tails. Depending on the sample size, it does not diminish towards the tails as fast. If the sample size is large, the t-distribution approximates the normal distribution.

T-test For Population Means

The type of hypothesis test used to test an assumed population mean when the population standard deviation is unknown. Due to the increased variability in using the sample standard deviation instead of the population standard deviation, the t-distribution is used in place of the z-distribution.



FORMULAS TO KNOW

T-Statistic For Population Means

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$