Search...

Statistics

Statistical analysis step by step
- Data collection
  - Data collection guide
  - Experimental design
  - + Populations and samples
  - + Types of variables
  - + Data cleansing
- Descriptive statistics
  - Overview descriptive statistics
  - + Measures of central tendency
  - + Measures of variability
  - + Frequency distribution
- **Probability (distributions)**
  - **Probability distributions**
    - Overview
    - Normal distribution
    - Standard normal distribution
    - Poisson distribution
    - Chi-square distribution
    - Chi-square table
    - *t* distribution
    - *t* table
- Inferential statistics
  - Overview inferential statistics
  - Degrees of freedom
  - Central limit theorem
  - + Parameters & test statistics
  - + Estimation
  - + Hypothesis testing
- Statistical tests
  - Choosing the right test
  - + Assumptions for hypothesis testing
  - + Correlation
  - + Regression analysis
  - *t* tests
  - + ANOVAs
  - + Chi-square
  - + Effect size
  - + Model selection

Reporting statistics in APA

Interesting topics

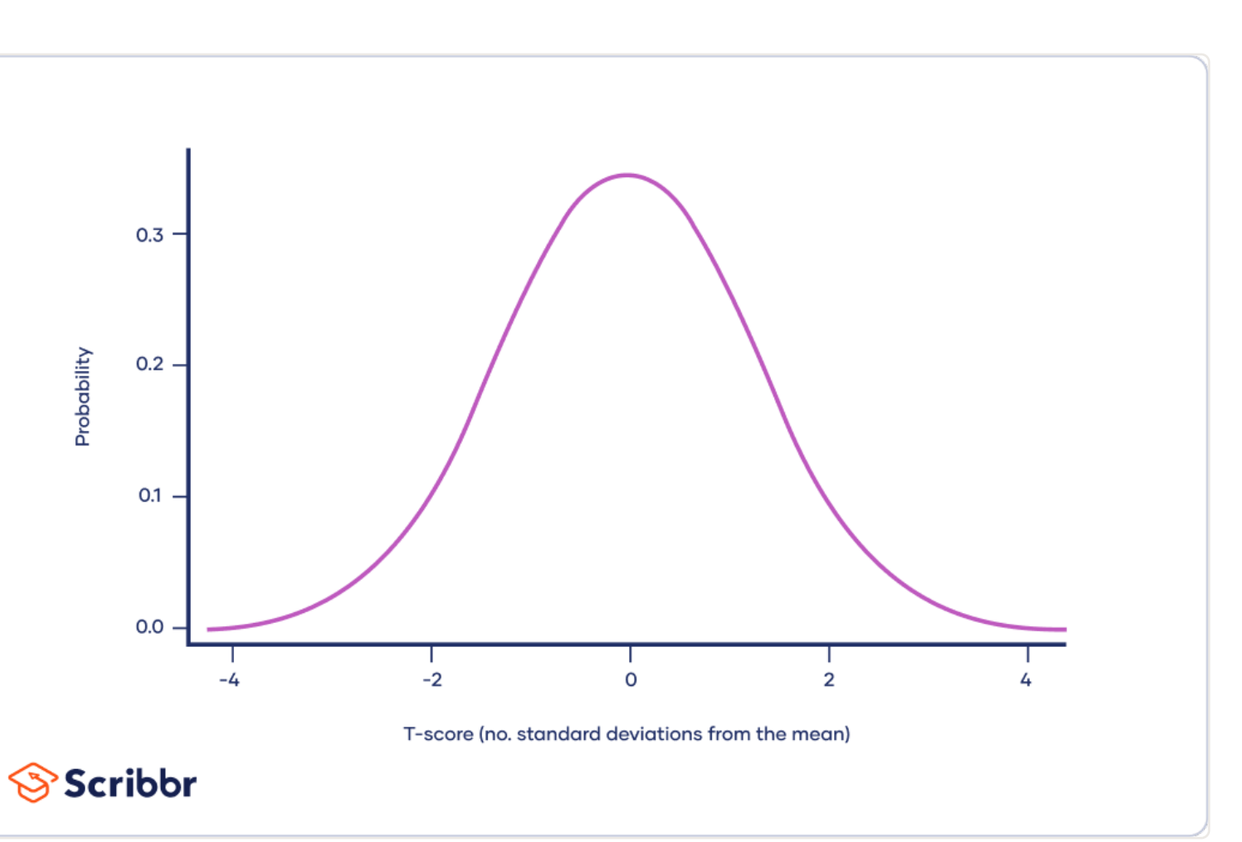
AMA style
Working with sources
IEEE

T-Distribution | What It Is and How To Use It (With Examples)

Published on August 28, 2020 by Rebecca Bevans. Revised on June 21, 2023.

The *t*-distribution, also known as Student's *t*-distribution, is a way of describing data that follow a bell curve when plotted on a graph, with the greatest number of observations close to the mean and fewer observations in the tails.

It is a type of normal distribution used for smaller sample sizes, where the variance in the data is unknown.



In statistics, the *t*-distribution is most often used to:

- Find the critical values for a confidence interval when the data is approximately normally distributed.
- Find the corresponding *p*-value from a statistical test that uses the *t*-distribution (*t*-tests, regression analysis).

Table of contents

What is a *t*-distribution?

The *t*-distribution is a type of normal distribution that is used for smaller sample sizes. Normally-distributed data form a bell shape when plotted on a graph, with more observations near the mean and fewer observations in the tails.

The *t*-distribution is used when data are *approximately* normally distributed, which means the data follow a bell shape but the population variance is unknown. The variance in a *t*-distribution is estimated based on the degrees of freedom of the data set (total number of observations minus 1).

It is a more conservative form of the **standard normal distribution**, also known as the *z*-distribution. This means that it gives a lower probability to the center and a higher probability to the tails than the standard normal distribution.

Example: *t*-distribution vs *z*-distribution

If you measure the average test score from a sample of only 20 students, you should use the *t*-distribution to estimate the confidence interval around the mean. If you use the *z*-distribution, your confidence interval will be artificially precise.

T-distribution and the standard normal distribution

As the **degrees of freedom** (total number of observations minus 1) increases, the *t*-distribution will get closer and closer to matching the standard normal distribution, a.k.a. the *z*-distribution, until they are almost identical.

Above 30 degrees of freedom, the *t*-distribution roughly matches the *z*-distribution. Therefore, the *z*-distribution can be used in place of the *t*-distribution with large sample sizes.

The *z*-distribution is preferable over the *t*-distribution when it comes to making statistical estimates because it has a known variance. It can make more precise estimates than the *t*-distribution, whose variance is approximated using the degrees of freedom of the data.

T-distribution and *t*-scores

A *t*-score is the number of standard deviations from the mean in a *t*-distribution. You can typically look up a *t*-score in a *t*-table, or by using an online *t*-score calculator.

In statistics, *t*-scores are primarily used to find two things:

1. The upper and lower bounds of a confidence interval when the data are approximately normally distributed.
2. The *p*-value of the test statistic for *t*-tests and regression tests.

T-scores and confidence intervals

Confidence intervals use *t*-scores to calculate the upper and lower bounds of the prediction interval. The *t*-score used to generate the upper and lower bounds is also known as the **critical value** of *t*, or *t**.

Example of a confidence interval

You have sampled 20 students from two different classes to estimate the mean standardized test scores and want to know if there is a difference between the two groups.

Using a two-tailed *t*-test, you generate an estimate of the difference between the two classes and a confidence interval around that estimate. From the *t*-test you find the difference in average score between class 1 and class 2 is 4.61, with a 95% confidence interval of 3.87 to 5.35.

Because the confidence interval does not cross zero, and is in fact quite far from zero, it is unlikely that this difference in test scores could have occurred under the null hypothesis of no difference between groups.

T-scores and *p*-values

Statistical tests generate a test statistic showing how far from the null hypothesis of the statistical test your data is. They then calculate a *p*-value that describes the likelihood of your data occurring if the null hypothesis were true.

The test statistic for *t*-tests and regression tests is the *t*-score. While most statistical programs will automatically calculate the corresponding *p*-value for the *t*-score, you can also look up the values in a *t*-table, using your degrees of freedom and *t*-score to find the *p*-value.

The *t*-score which generates a *p*-value below your threshold for statistical significance is known as the critical value of *t*, or *t**.

Example of a p-value

The two-tailed *t*-test of the difference in test scores generates a *t*-value of 12.79. This means that the difference in group means is 12.79 standard deviations away from the mean of the distribution of the null hypothesis.

The degrees of freedom is 38 (n–1 for each group). Looking this up in a *t*-table (or calculating it in your favorite stats program) you find a *p*-value < 0.001.

This finding, like the finding from the confidence interval, suggests that you are not likely to find a difference this large if the true difference in average test scores is zero.

Other interesting articles

If you want to know more about statistics, methodology, or research bias, make sure to check out some of our other articles with explanations and examples.

Statistics

- Student's *t* table
- Kurtosis
- Descriptive statistics
- Measures of central tendency
- Correlation coefficient
- *p* value

Methodology

- Cluster sampling
- Stratified sampling
- Types of interviews
- Case study
- Cohort study
- Thematic analysis

Research bias

- Implicit bias
- Cognitive bias
- Survivorship bias
- Availability heuristic
- Nonresponse bias
- Regression to the mean

Frequently asked questions about the t-distribution

What is a t-distribution?

What is the difference between the t-distribution and the standard normal distribution?

What is a t-score?

What is a test statistic?

What is a critical value?

Cite this Scribbr article

If you want to cite this source, you can copy and paste the citation or click the "Cite this Scribbr article" button to automatically add the citation to our free Citation Generator.

Bevans, R. (2023, June 21). *T-Distribution | What It Is and How To Use It (With Examples).* Scribbr. Retrieved September 9, 2024, from https://www.scribbr.com/statistics/t-distribution/

Cite this article

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.