Choosing the Right Statistical Test | Types & Examples

Published on January 28, 2020 by Rebecca Bevans. Revised on July 6, 2022.

Statistical tests are used in hypothesis testing. They can be used to:

- determine whether a predictor variable has a statistically significant relationship with an outcome variable.
- estimate the difference between two or more groups.

Statistical tests assume a null hypothesis of no relationship or no difference between groups. Then they determine whether the observed data fall outside of the range of values predicted by the null hypothesis.

If you already know what types of variables you're dealing with, you can use the flowchart to choose the right statistical test for your data.

Statistical tests flowchart

Table of contents

What does a statistical test do?

Statistical tests work by calculating a **test statistic** – a number that describes how much the relationship between variables in your test differs from the null hypothesis of no relatior

It then calculates a ***p*-value** (probability value). The *p*-value estimates how likely it is that you would see the difference described by the test statistic if the null hypothesis of no relationship were true.

If the value of the test statistic is more extreme than the statistic calculated from the null hypothesis, then you can infer a **statistically significant relationship** between the predictor and outcome variables.

If the value of the test statistic is less extreme than the one calculated from the null hypothesis, then you can infer **no statistically significant relationship** between the predictor and outcome variables.

When to perform a statistical test

You can perform statistical tests on data that have been collected in a statistically valid manner – either through an experiment, or through observations made using probability sampling methods.

For a statistical test to be valid, your sample size needs to be large enough to approximate the true distribution of the population being studied.

To determine which statistical test to use, you need to know:

- whether your data meets certain assumptions.
- the types of variables that you're dealing with.

Statistical assumptions

Statistical tests make some common assumptions about the data they are testing:

1. **Independence of observations** (a.k.a. no autocorrelation): The observations/variables you include in your test are not related (for example, multiple measurements of a single test subject are not independent, while measurements of multiple different test subjects are independent).

2. **Homogeneity of variance**: the variance within each group being compared is similar among all groups. If one group has much more variation than others, it will limit the effectiveness.
3. **Normality of data**: the data follows a normal distribution (a.k.a. a bell curve). This assumption applies only to quantitative data.

If your data do not meet the assumptions of normality or homogeneity of variance, you may be able to perform a **nonparametric statistical test**, which allows you to make comparisons without any assumptions about the data distribution.

If your data do not meet the assumption of independence of observations, you may be able to use a test that accounts for structure in your data (repeated-measures tests or tests that include blocking variables).

Types of variables

The types of variables you have usually determine what type of statistical test you can use.

Quantitative variables represent amounts of things (e.g. the number of trees in a forest). Types of quantitative variables include:

- **Continuous** (a.k.a ratio variables): represent measures and can usually be divided into units smaller than one (e.g. 0.75 grams).
- **Discrete** (a.k.a integer variables): represent counts and usually can't be divided into units smaller than one (e.g. 1 tree).

Categorical variables represent groupings of things (e.g. the different tree species in a forest). Types of categorical variables include:

- **Ordinal**: represent data with an order (e.g. rankings).
- **Nominal**: represent group names (e.g. brands or species names).
- **Binary**: represent data with a yes/no or 1/0 outcome (e.g. win or lose).

Choose the test that fits the types of predictor and outcome variables you have collected (if you are doing an experiment, these are the independent and dependent variables). Consult the tables below to see which test best matches your variables.

Choosing a parametric test: regression, comparison, or correlation

Parametric tests usually have stricter requirements than nonparametric tests, and are able to make stronger inferences from the data. They can only be conducted with data that adheres to the common assumptions of statistical tests.

The most common types of parametric test include regression tests, comparison tests, and correlation tests.

Regression tests

Regression tests look for **cause-and-effect relationships**. They can be used to estimate the effect of one or more continuous variables on another variable.

Simple linear regression

Predictor variable:

- Continuous
- 1 predictor

Outcome variable:

- Continuous
- 1 outcome

Research question example: What is the effect of income on longevity?

Multiple linear regression

Predictor variable:

- Continuous
- 2 or more predictors

Outcome variable:

- Continuous
- 1 outcome

Research question example: What is the effect of income and minutes of exercise per day on longevity?

Logistic regression

Predictor variable:

- Continuous

Outcome variable:

- Binary

Research question example: What is the effect of drug dosage on the survival of a test subject?

Comparison tests

Comparison tests look for **differences among group means**. They can be used to test the effect of a categorical variable on the mean value of some other characteristic.

T-tests are used when comparing the means of precisely two groups (e.g. the average heights of men and women). ANOVA and MANOVA tests are used when comparing the i of more than two groups (e.g. the average heights of children, teenagers, and adults).

Paired t-test
Predictor variable: • Categorical • 1 predictor
Outcome variable: • Quantitative • groups come from the same population
Research question example: What is the effect of two different test prep programs on the average exam scores for students from the same class?

Independent t-test
Predictor variable: • Categorical • 1 predictor
Outcome variable: • Quantitative • groups come from different populations
Research question example: What is the difference in average exam scores for students from two different schools?

ANOVA

Predictor variable:

- Categorical
- 1 or more predictor

Outcome variable:

- Quantitative
- 1 outcome

Research question example: What is the difference in average pain levels among post-surgical patients given three different painkillers?

MANOVA

Predictor variable:

- Categorical
- 1 or more predictor

Outcome variable:

- Quantitative
- 2 or more outcome

Research question example: What is the effect of flower species on petal length, petal width, and stem length?

Correlation tests

Correlation tests **check whether variables are related** without hypothesizing a cause-and-effect relationship.

These can be used to test whether two variables you want to use in (for example) a multiple regression test are autocorrelated.

Pearson's *r*

Variables:

- 2 continuous variables

Research question example: How are latitude and temperature related?

Choosing a nonparametric test

Non-parametric tests don't make as many assumptions about the data, and are useful when one or more of the common statistical assumptions are violated. However, the inferences they make aren't as strong as with parametric tests.

Spearman's *r*

Predictor variable:

- Quantitative

Outcome variable:

- Quantitative

Use in place of...: Pearson's *r*

Chi square test of independence

Predictor variable:

- Categorical

Outcome variable:

- Categorical

Use in place of...: Pearson's *r*

Sign test

Predictor variable:

- Categorical

Outcome variable:

- Quantitative

Use in place of...: One-sample *t*-test

Kruskal–Wallis *H*

Predictor variable:

- Categorical
- 3 or more groups

Outcome variable:

- Quantitative

Use in place of...: ANOVA

ANOSIM

Predictor variable:

- Categorical
- 3 or more groups

Outcome variable:

- Quantitative
- 2 or more outcome variables

Use in place of...: MANOVA

Wilcoxon Rank-Sum test

Predictor variable:

- Categorical
- 2 groups

Outcome variable:

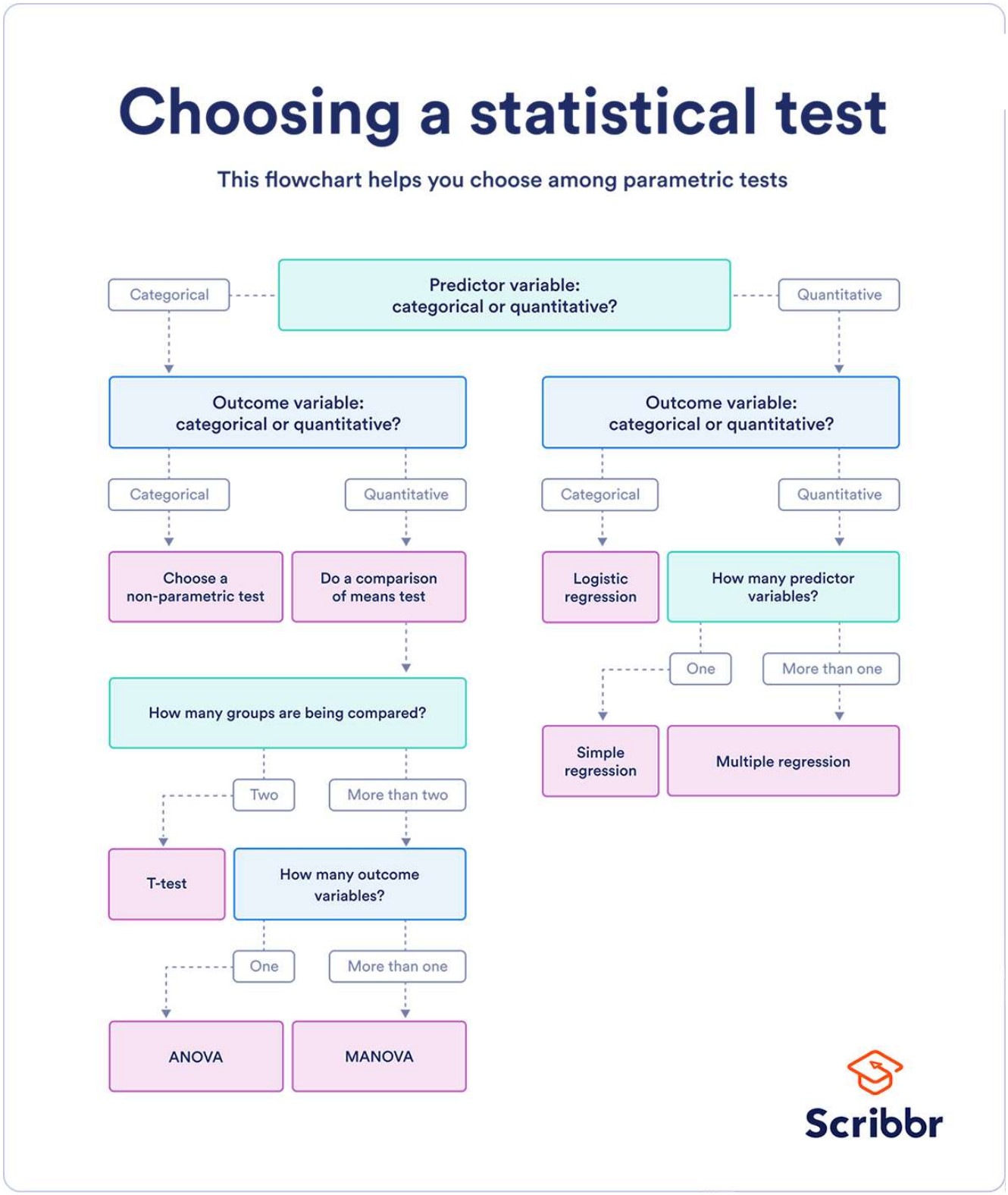
- Quantitative
- groups come from different populations

Use in place of...: Independent t-test

Wilcoxon Signed-rank test
Predictor variable: • Categorical • 2 groups
Outcome variable: • Quantitative • groups come from the same population
Use in place of...: Paired t-test

Flowchart: choosing a statistical test

This flowchart helps you choose among parametric tests. For nonparametric alternatives, check the table above.



Frequently asked questions about statistical tests

What are the main assumptions of statistical tests?

Statistical tests commonly assume that:

1. the data are normally distributed
2. the groups that are being compared have similar variance
3. the data are independent

If your data does not meet these assumptions you might still be able to use a nonparametric statistical test, which have fewer requirements but also make weaker inferences.

What is a test statistic?

A test statistic is a number calculated by a statistical test. It describes how far your observed data is from the null hypothesis of no relationship between variables or no difference among sample groups.

The test statistic tells you how different two or more groups are from the overall population mean, or how different a linear slope is from the slope predicted by a null hypothesis. Different test statistics are used in different statistical tests.

What is statistical significance?

Statistical significance is a term used by researchers to state that it is unlikely their observations could have occurred under the null hypothesis of a statistical test. Significance is usually denoted by a *p*-value, or probability value.

Statistical significance is arbitrary – it depends on the threshold, or alpha value, chosen by the researcher. The most common threshold is $p < 0.05$, which means that the data is likely to occur less than 5% of the time under the null hypothesis.

When the *p*-value falls below the chosen alpha value, then we say the result of the test is statistically significant.

What is the difference between quantitative and categorical variables?

Quantitative variables are any variables where the data represent amounts (e.g. height, weight, or age).

Categorical variables are any variables where the data represent groups. This includes rankings (e.g. finishing places in a race), classifications (e.g. brands of cereal), and binary outcomes (e.g. coin flips).

You need to know what type of variables you are working with to choose the right statistical test for your data and interpret your results.

What is the difference between discrete and continuous variables?

Discrete and continuous variables are two types of quantitative variables:

- **Discrete variables** represent counts (e.g. the number of objects in a collection).
- **Continuous variables** represent measurable amounts (e.g. water volume or weight).

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.

Measurement type	Continuous, parametric	Nominal/ordinal/ nonparametric	Dichotomous (two possible outcomes)	Survival (time to event) (*not clear*)
Describe one group	Mean, SD	Median, percentiles	Proportion	Kaplan–Meier survival curve, median survival
Compare one group to a hypothetical value	One sample *t*-test	Wilcoxon test	Chi-squared or binominal test	
Compare two unpaired groups	Unpaired *t*-test	Mann–Whitney	Fisher's or Chi-squared	Log rank or Mantel–Haenszel
Compare two paired groups	Paired *t*-test	Wilcoxon	McNamara's	Conditional proportional hazards regression
Compare three or more unmatched groups	One way ANOVA	Kruskal–Wallis	Chi-squared	Cox proportional hazards regression
Compare three or more matched groups	Repeated-measured ANOVA	Friedman	Cochrane Q	Conditional proportional hazards regression
Quantify relationship between two variables	Pearson correlation	Spearman correlation	Contingency coefficients	
Predict value from another variable	Linear (or nonlinear) regression	Nonparametric regression	Simple logistic regression	Cox proportional hazards regression
Predict values from several measured or binominal variables	Multiple linear (or nonlinear) regression		Multiple logistic regression	Cox proportional hazards regression

Source: This table is derived from Mikulski, H. (1995). *Intuitive Statistics*. Oxford Press. New York.

Nonparametric Test	What it does	Parametric Counterpart
Wilcoxon Signed Rank	Compares 1 Median to a specified value	z-test, 1-Sample t-test
	Compares 2 Dependent (Paired) Medians	Paired (Dependent) Samples t-test
Mann-Whitney	Compares 2 Independent Medians	2 (Independent) Samples t-test
Kruskal-Wallis	Compares 3 or more Medians, 1 Variable	1-way ANOVA
Friedman	Compares 3 or more Medians, 2 Variables	2-way ANOVA
Chi-Square Test of Independence	Tests 2 Categorical Variables for Independence (lack of Association)	none

	Parametric	Non-parametric
Distribution	Normal/Gaussian	Any
Assumed variance	Homogeneous	Any
Type of data	Ratio or Interval	Ordinal or Nominal
Data set relationships	Independent	Any
Measure of central tendency	Mean	Median
Correlation test	Pearson	Spearman
Independent measures, 2 groups	Independent-measures t-test	Mann-Whitney test
Independent measures, >2 groups	One-way, independent-measures ANOVA	Kruskal-Wallis test
Repeated measures, 2 conditions	Matched-pair t-test	Wilcoxon test
Repeated measures, >2 conditions	One-way, repeated measures ANOVA	Friedman's test