

I have two groups of data values. I want to test for a difference between the two groups. What Test Should I Use?

The comments below are meant to provide a relatively quick answer to this common question. Please read Chapters 3 and 4 of *The Statistical Sleuth* to get a more complete answer.

First of all, you must identify whether the two groups of data values are paired. The data are paired if there is a natural way to match each value in one group with a value in the other group because of the way the data were collected.

If the data are paired, proceed as follows:

1. Compute a difference for each pair of data.
2. If the distribution of differences appears roughly normal with no extreme outliers, conduct a t -test to determine if the population of differences has a mean of zero.
3. If you notice that the differences between pairs of values get larger as the data values in a pair get larger and/or that the values in one of the groups seems to be approximately a common multiple of the values in the other group, consider taking logs of all the individual values prior to taking differences. (See Sections 3.6.2 and 3.6.3.) Then return to step 2.
4. If outliers are present in the differences or there is clear skewness in the distribution of differences, use either the signed-rank test or the sign test. The sign test is appropriate if only a quick assessment of significance is needed or if only the sign of the difference is available or important. If the sizes of the differences in addition to their signs convey scientific information, then the signed-rank test should be used.

If the data are not paired, proceed as follows:

1. If the distributions of both samples are right skewed and the standard deviation of the sample with the larger mean is greater than the standard deviation of the sample with the smaller mean, take logarithms of all the data values and conduct a two-sample t -test on the logged data. If these conditions are not satisfied, proceed to 2.
2. If the distributions of the two samples are similarly shaped with no major outliers and similar standard deviations, conduct a two-sample t -test. Even if the standard deviations differ, the two-sample t -test will perform reasonably well if there are approximately the same number of observations in both groups. If these conditions are not satisfied, proceed to 3.
3. Conduct a rank-sum test or a permutation test. The normal approximation used to obtain a p -value for the rank-sum test will work well if there are not many ties in the data values. The permutation test will produce a valid result regardless of the number of ties, but obtaining a p -value for the permutation test can be more difficult.

Note that for both paired and unpaired data, the use of a t -test is the recommended strategy unless features of the data require a different course of action. The t -test is recommended because it gives the best chance of detecting differences between the groups when real differences exist and the data are "well behaved." Also, the confidence intervals that correspond to t -tests are very useful, and obtaining confidence intervals that correspond to other tests is more difficult.

Sometimes students are concerned that by considering transformations and alternative test procedures, we are trying to "make results turn out the way we would like them to." Note that in the instructions above, the decision procedure does not depend in any way on the results of any test. When the data are not well-behaved, the t -tests will often give p -values that do not accurately reflect the information in the data about whether the two groups differ. Thus we seek alternative procedures that will accurately reflect information in the data when necessary. The point of statistical analysis is to uncover the truth, regardless of whether the truth is that differences exist or the truth is that differences do not exist.