

Big Picture: Different statistical tests

<http://www.wellcome.ac.uk/Education-resources/Education-and-learning/Big-Picture/All-issues/Number-Crunching/WTS057456.htm>

There are several different statistical tests you can use, and the best one to pick will depend on the type of data you are dealing with. Read about some common tests below, and follow the links to more information and videos on each.

Chi-squared test

The chi-squared test is used with categorical data to see whether any difference in frequencies between your sets of results is due to chance. For example, a ladybird lays a clutch of eggs. You expect that all of the clutch will hatch, but only three-quarters of them do. Is the failure of some of the clutch to hatch statistically significant, and if it is, what could be the reason for it?

In a chi-squared test, you draw a table of your observed frequencies and your predicted frequencies and calculate the chi-squared value. You compare this to the critical value to see whether the difference between them is likely to have occurred by chance. If your calculated value is bigger than the critical value, you reject your null hypothesis.

- This [worked example from the Big Picture team](#) is about vitamin C and getting colds.
- This [worked example from the Field Studies Council](#) is on ecology.
- This [video on chi-squared](#) from the Big Picture team investigates fingerprint types.
- This [video on chi-squared](#) is from Paul Andersen.

T-test

The t-test enables you to see whether two samples are different when you have data that are continuous and normally distributed.

The test allows you to compare the means and standard deviations of the two groups to see whether there is a statistically significant difference between them. For example, you could test the heights of the members of two different biology classes.

- This [video from StatsCast explains the purpose of t-tests, how they work, and how to interpret the results](#).

Mann-Whitney U-test

The Mann-Whitney U-test is similar to the t-test. It is used when comparing ordinal data (i.e. data that can be ranked or has some sort of rating scale) that are not normally distributed. Measurements must be categorical - for instance, yes or no - and independent of each other (e.g. a single person cannot be represented twice).

For example, the Mann-Whitney U-test could be used to test the effectiveness of an antihistamine tablet compared to a spray in a group of people with hay fever. To do this, you would split the group in half, then give each half a different treatment and ask each person how effective they thought it was. The test could be used to see whether there is a difference in the perceived efficacy of the two treatments.

- This [PDF from Salters Nuffield Advanced Biology includes a worked example of the test](#).

Standard error and 95 per cent confidence limits

The standard error and 95 per cent confidence limits allow us to gauge how representative of the real world population the data are.

- A [video tutorial explaining what the standard error is and how to work it out](#), with Paul Andersen.
- This [site from the University of Glasgow outlines why we use confidence intervals and has questions to test your understanding](#).

Spearman's rank correlation coefficient

The Spearman's rank correlation coefficient tests the relationship between two variables in a dataset; for example, is a person's weight related to their height? If there is a statistically significant relationship, you can reject the null hypothesis, which may be that there is no link between the two variables.

- This [site from the Barcelona Field Studies Centre explains the theories behind the test, outlines the potential pitfalls and includes a worked example](#).

Wilcoxon matched pairs test

Like the Mann-Whitney U-test, this test is used for discontinuous data that are not normally distributed but do have a link between the two datasets. For example, when asking people to rank how hungry they feel before a meal and doing so again after they have eaten - because the same person is providing both answers, the datasets are not independent.

- This site from the University of Central Missouri includes [a worked example on a biological topic](#).