



EXPLORABLE
Think Outside The Box

Published on *Explorable.com* (<https://explorable.com>)

[Home](#) > Significance Test

Significance Test

Martyn Shuttleworth, Lyndsay T Wilson 222.1K reads

Significance tests play a key role in experiments: they allow researchers to determine whether their data supports or rejects the null hypothesis, and consequently whether they can accept their alternative hypothesis.

In everyday language, "significance" means that something is meaningful or important, but in statistical language, the definition is more precise. Furthermore, significance here does not imply theoretical, practical or research importance. A result can be statistically significant but a rather unimportant finding considering the bigger picture! A result is statistically significant if it satisfies certain statistical criteria.

The banner features the Explorable logo and the text "Quiz Time!". Below this are three quiz cards:

- Quiz: Psychology 101 Part 2** (Image: A pair of red roller skates on a wooden deck)
- Quiz: Psychology 101 Part 2** (Image: A fan of colorful pencils)
- Quiz: Flags in Europe** (Image: A Ferris wheel at sunset)

[See all quizzes =>](#)

The P-Value and the Significance Level

Significance comes down to the relationship between two crucial quantities, the p-value and the significance level (α). We can call a result statistically significant when $P < \alpha$. Let's consider what each of these quantities represents.

- **p-value:** This is calculated after you obtain your results. It is the probability of observing an extreme effect even with the null hypothesis still being true. Importantly, it does not

measure the *size* of an effect.

- **alpha:** This is decided on before gathering data. It is the probability of the study rejecting the null hypothesis despite it being true (i.e. the chance of committing a Type 1 error). It is essential an error rate and usually set at or below 5%.

A Comfortable Confidence Level

It's important to remember that there is nothing inherent about a 5% confidence level; it is merely a common convention. Where exactly the threshold is set is largely determined by the data in question and what the researchers are trying to achieve.

- Sciences where random error and natural variation are likely to play a part (for example investigative biology) will likely be content with alpha set to 5%.
- If you can expect a high level of precision and accuracy with the measurements and instruments employed, alpha can be set lower.

P-values are between 0 and 1. If P is less than the cut-off you've pre-chosen, you should reject the null hypothesis in favor of the alternative. Alternatively, if P is greater than the cut-off, say 0.05, you should not reject the null.

A note about falsifiability: though you could be forgiven for thinking otherwise, any piece of research is technically setting out to prove or disprove the null hypothesis, and nothing more. The alternative hypothesis is correctly named – it is only a position that is (provisionally) accepted as an alternative after the null hypothesis has been ruled out. All a significant result tells us is that there is “something going on” as opposed to nothing.

A Word of Caution

If you are studying statistics for a university course, the above may well be sufficient when it comes to writing up a term paper or understanding the general concepts behind statistical testing. However, the fact is that statistics is a complex and evolving science, and nowhere near the panacea that many students believe it to be.

Interestingly, the ASA (American Statistical Association) has published some [guidelines](#) ^[1] about the proper use of the p-value, which will be of interest to those publishing more serious research. Some of these recommendations are:

- p-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
- Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.
- A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.
- By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis.

In other words, good research goes well beyond the simple yes/no mechanisms many

students of statistics are first taught. A depth understanding of the limits of significance testing is beyond the scope of most students' curricula, however it does confirm the fact that research is seldom black and white!

Source URL: <https://explorable.com/significance-test?gid=1590>

Links

[1] <https://www.amstat.org/asa/files/pdfs/P-ValueStatement.pdf>