# Statistically Significant Results

Statistically significant results are those that are understood as not likely to have occurred purely by chance and thereby have other underlying causes for their occurrence - hopefully, the underlying causes you are trying to investigate!

Whenever a statistical analysis is performed and the results interpreted, there is always a possibility that the results are purely by chance (random error). This is an inherent limitation of any statistical analysis and cannot be done away with. In addition, mistakes such as measurement errors may cause the experimenter to misinterpret the results (systematic error).

Fortunately, the probability that the process was simply a chance encounter can be calculated, and a minimum threshold of statistical significance can be set. If the results are obtained such that the probability that they are simply a chance process is less than this threshold of significance, then we can say the results have a high probability of not being due to chance. Note that the probability is never zero; *statistical tests are never 100% certain*. Threshold levels merely indicate the risk we are willing to take when it comes to accepting or rejecting a particular hypothesis.

> Common statistically significant levels are 5%, 1% and 0.1% depending on the analysis and the field of study.

In terms of null hypothesis, the concept of statistical significance can be understood to be the minimum level at which the null hypothesis can be rejected. This means if the experimenter sets his statistical significance level at 5% and the probability that the results are a chance process (i.e. the p-value) is 3%, then the experimenter can claim that the null hypothesis can be rejected.

In this case, the experimenter will claim his results statistically significant. Some research disciplines are stricter than others however, and the research design itself may warrant a more or less stringent threshold. In any case, the lower the significance level, the higher the confidence you can have in the result.

# What's a Good Significance Level?

Statistically significant results [1] are required for many practical cases of experimentation [2] in various branches of research. The choice of the statistical significance level is influenced by a number of parameters and depends on the experiment in question.

In most cases, the data follows a normal distribution, which is thankfully also the simplest case. With standard normal distribution, you can use a threshold level of 0.05 confidently. However, care should always be taken to account for other distributions within the given population.

Although 5%, 1% and 0.1% are common significance levels, it is not clear cut which level to use in an actual study - it depends on the norms of the field, previous studies, and the amount of evidence needed. However, it is not recommended to have a significance level higher than 5% because it too often leads to type 1-errors [3].

# Be Cautious When Interpreting the P-Value!

It can be very satisfying to work out the p-value after a long experiment, see that it's below the threshold, reject the null hypothesis and assume the experiment is done and dusted. But the truth is that researchers still need to use care when deciding how to interpret the p-value.

- It is an error to assume that a very small p-value means a strong result or large effect. The p-value gives *probabilistic* information about our result, nothing more. You will need to conduct further investigation to understand the magnitude of any relationships in your experiment.

- While determining significant results statistically, remember that it's impossible to use statistics to prove that the difference in levels of two parameters is zero. The only thing that the statistical analysis can state is that the experiment failed to find any difference.

-

It's commonly said that "statistical significance is not the same as biological significance." What this means is that statistics is only a tool that can attempt to describe complex, concrete biological systems, but can't capture every extra factor that may come into play.

- Lastly, a result may be statistically significant but still not *practically* significant, meaning it's still not large enough to warrant any change or action in the "real world." Again, the p-value is not the final say, but needs to be interpreted carefully.

---

**Source URL:** https://explorable.com/statistically-significant-results

**Links**
[1] http://www.numberwatch.co.uk/significance.htm
[2] https://explorable.com/experimental-research
[3] https://explorable.com/type-i-error