Statistical significance

In statistical hypothesis testing, a result has statistical significance when it is very unlikely to have occurred given the null hypothesis.

More precisely, the significance level defined for a study, α, is the probability of the study rejecting the null hypothesis, given that it were true; and the p-value of a result, p, is the probability of obtaining a result at least as extreme, given that the null hypothesis were true. The result is statistically significant, by the standards of the study, when $p < \alpha$.

The significance level for a study is chosen before data collection, and typically set to 5% or much lower, depending on the field of study.

In any experiment or observation that involves drawing a sample from a population, there is always the possibility that an observed effect would have occurred due to sampling error alone.

But if the p-value of an observed effect is less than the significance level, an investigator may conclude that the effect reflects the characteristics of the whole population, thereby rejecting the null hypothesis.

This technique for testing the significance of results was developed in the early 20th century.