How to Understand P-Values and Statistical Significance in Statistical Analysis

Authored by

stats writer

March 3, 2026

ARABPSYCHOLOGY.COM

RECOMMENDED CITATION

stats writer (2026). *How to Understand P-Values and Statistical Significance in Statistical Analysis*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=133594

P-values and statistical significance play a crucial role in statistical analysis. P-values represent the probability that the observed results of a study occurred by chance alone. A lower P-value indicates a higher level of significance, meaning that the results are less likely to have occurred by chance. This helps researchers determine whether the results of a study are statistically significant or not.

Statistical significance is a measure of the strength of evidence in support of a hypothesis. It is determined by comparing the observed results to the expected results, and if the difference is unlikely to have occurred by chance, the results are considered statistically significant. This allows researchers to draw conclusions and make decisions based on the data collected.

The significance of P-values and statistical significance lies in their ability to provide objective and quantitative evidence for the validity of a study's results. They help researchers determine the reliability of their findings and make informed decisions based on the data. Ultimately, these measures are essential in ensuring the accuracy and credibility of statistical analysis.

ARABPSYCHOLOGY.COM

An Explanation of P-Values and Statistical Significance

In statistics, p-values are commonly used in hypothesis testing for t-tests, chi-square tests, regression analysis, ANOVAs, and a variety of other statistical methods.

Despite being so common, people often interpret p-values incorrectly, which can lead to errors when interpreting the findings from an analysis or a study.

This post explains how to understand and interpret p-values in a clear, practical way.

Hypothesis Testing

To understand p-values, we first need to understand the

concept of .

A hypothesis test is a formal statistical test we use to reject or fail to reject some hypothesis. For example, we may hypothesize that a new drug, method, or procedure provides some benefit over a current drug, method, or procedure.

To test this, we can conduct a hypothesis test where we use a null and alternative hypothesis:

Null hypothesis - There is no effect or difference between the new method and the old method.

Alternative hypothesis - There does exist some effect or difference between the new method and the old method.

A p-value indicates how believable the null hypothesis is, given the sample data. Specifically, assuming the null hypothesis is true, the p-value tells us the probability of obtaining an effect at least as large as the one we actually observed in the sample data.

If the p-value of a hypothesis test is sufficiently low, we can reject the null hypothesis. Specifically, when we conduct a hypothesis test, we must choose a

ARABPSYCHOLOGY.COM

significance level at the outset. Common choices for significance levels are 0.01, 0.05, and 0.10.

If the p-values is *less than* our significance level, then we can reject the null hypothesis.

Otherwise, if the p-value is *equal to or greater than* our significance level, then we fail to reject the null hypothesis.

How to Interpret a P-Value

The textbook definition of a p-value is:

A p-value is the probability of observing a sample statistic that is at least as extreme as your sample statistic, given that the null hypothesis is true.

For example, suppose a factory claims that they produce tires that have a mean weight of 200 pounds. An auditor hypothesizes that the true mean weight of tires produced at this factory is different from 200 pounds so he runs a hypothesis test and finds that the p-value of the test is 0.04. Here is how to interpret this p-value:

ARABPSYCHOLOGY.COM

Depending on the significance level used in this hypothesis test, the auditor would likely reject the null hypothesis that the true mean weight of tires produced at this factory is indeed 200 pounds. The sample data that he obtained from the audit is not very consistent with the null hypothesis.

How *Not* to Interpret a P-Value

The biggest misconception about p-values is that they are equivalent to the probability of making a mistake by rejecting a true null hypothesis (known as a Type I error).

There are two primary reasons that p-values can't be the error rate:

1. P-values are calculated based on the assumption that the null hypothesis is true and that the difference between the sample data and the null hypothesis is simple caused by random chance. Thus, p-values can't tell you the probability that the null is true or false since it is 100% true based on the perspective of the calculations.

2. Although a low p-value indicates that your sample

ARABPSYCHOLOGY.COM

data are unlikely assuming the null is true, a p-value still can't tell you which of the following cases is more likely:

The null is falseThe null is true but you obtained an odd sample

In regards to the previous example, here is a correct and incorrect way to interpret the p-value:

Correct Interpretation: Assuming the factory does produce tires with a mean weight of 200 pounds, you would obtain the observed difference that you *did* obtain in your sample or a more extreme difference in 4% of audits due to random sampling error.Incorrect Interpretation: If you reject the null hypothesis, there is a 4% chance that you are making a mistake.

Examples of Interpreting P-Values

The following examples illustrate correct ways to interpret p-values in the context of hypothesis testing.

Example 1

A phone company claims that 90% of its customers are satisfied with their service. To test this claim, an

ARABPSYCHOLOGY.COM

independent researcher gathered a of 200 customers and asked them if they are satisfied with their service, to which 85% responded yes. The p-value associated with this sample data turned out to be 0.018.

Correct interpretation of p-value: Assuming that 90% of the customers actually are satisfied with their service, the researcher would obtain the observed difference that he *did* obtain in his sample or a more extreme difference in 1.8% of audits due to random sampling error.

Example 2

A company invents a new battery for phones. The company claims that this new battery will work for at least 10 minutes longer than the old battery. To test this claim, a researcher takes a simple random sample of 80 new batteries and 80 old batteries. The new batteries run for an average of 120 minutes with a standard deviation of 12 minutes and the old batteries run for an average of 115 minutes with a standard deviation of 15 minutes. The p-value that results from the test for a difference in population means is 0.011.

ARABPSYCHOLOGY.COM

Correct interpretation of p-value: Assuming that the new battery works for the same amount of time or less than the old battery, the researcher would obtain the observed difference or a more extreme difference in 1.1% of studies due to random sampling error.

ARABPSYCHOLOGY.COM