P-value

What Is P-Value?

In statistics, the p-value is the probability of obtaining results at least as extreme as the observed results of a statistical hypothesis test, assuming that the null hypothesis is correct. The p-value serves as an alternative to rejection points to provide the smallest level of significance at which the null hypothesis would be rejected. A smaller p-value means that there is stronger evidence in favor of the alternative hypothesis.

P-value is often used to promote credibility for studies or reports by government agencies. For example, the U.S. Census Bureau stipulates that any analysis with a p-value greater than 0.10 must be accompanied by a statement that the difference is not statistically different from zero. The Census Bureau also has standards in place stipulating which p-values are acceptable for various publications.

How Is P-Value Calculated?

P-values are usually found using p-value tables or spreadsheets/statistical software. These calculations are based on the assumed or known probability distribution of the specific statistic tested. P-values are calculated from the deviation between the observed value and a chosen reference value, given the probability distribution of the statistic, with a greater difference between the two values corresponding to a lower p-value.

Mathematically, the p-value is calculated using integral calculus from the area under the probability distribution curve for all values of statistics that are at least as far from the reference value as the observed value is, relative to the total area under the probability distribution curve.

The calculation for a p-value varies based on the type of test performed. The three test types describe the location on the probability distribution curve: lower-tailed test, upper-tailed test, or two-tailed test.

The calculation for a p-value varies based on the type of test performed. The three test types describe the location on the probability distribution curve: lower-tailed test, upper-tailed test, or two-tailed test.

The P-Value Approach to Hypothesis Testing

The p-value approach to hypothesis testing uses the calculated probability to determine whether there is evidence to reject the null hypothesis. The null hypothesis, also known as the conjecture, is the initial claim about a population (or data-generating process). The alternative hypothesis states whether the population parameter differs from the value of the population parameter stated in the conjecture.

In practice, the significance level is stated in advance to determine how small the p-value must be to reject the null hypothesis. Because different researchers use different levels of significance when examining a question, a reader may sometimes have difficulty comparing results from two different tests. P-values provide a solution to this problem.

For example, suppose a study comparing returns from two particular assets was undertaken by different researchers who used the same data but different significance levels. The researchers might come to opposite conclusions regarding whether the assets differ.

If one researcher used a confidence level of 90% and the other required a confidence level of 95% to reject the null hypothesis, and if the p-value of the observed difference between the two returns was 0.08 (corresponding to a confidence level of 92%), then the first researcher would find that the two assets have a difference that is statistically significant, while the second would find no statistically significant difference between the returns.

To avoid this problem, the researchers could report the p-value of the hypothesis test and allow readers to interpret the statistical significance themselves. This is called a p-value approach to hypothesis testing. Independent observers could note the p-value and decide for themselves whether that represents a statistically significant difference or not.