



Statistics Terminology



A concise list of statistics terminology to take with you for easy reference.

AB Testing Statistics Terms

An easy-access guide to key terms and concepts

P-Value

The p-value is the probability of obtaining at least as extreme results given that the null hypothesis is true. In other words, the p-value is the expected fluctuation in a given sample, similar to the variance.

In AB testing, the p-value is the difference we would see across samples if we ran an A/A test.

Significance Level

The significance level is the probability of rejecting the null hypothesis given that it is true. It's a value we set based on the level of accuracy we deem acceptable.

In AB testing, the significance level is the industry standard 5% we use to give us results with 95% confidence.

Statistical Hypothesis Testing

An AB test is an example of statistical hypothesis testing, a process whereby a hypothesis is made about the relationship between two data sets and those data sets are then compared against each other to determine if there is a statistically significant relationship or not.

To put this in more practical terms, a prediction is made that Page Variation #B will perform better than Page Variation #A, and then data sets from both pages are observed and compared to determine if Page Variation #B is a statistically significant improvement over Page Variation #A.

Statistical Population

In statistics, a population is a set of similar items or events which is of interest for some question or experiment. It's the group we want information about.

When AB testing a webpage, the true population is every future individual who will visit that page.

Data Sample

A data sample is a set of data collected and/or selected from a statistical population by a defined procedure. It's a small portion of the larger population.

When AB testing, the sample is the number of visitors we display our new page variation too in order to collect data.

Mean

The mean is the central tendency of a probability distribution.

In AB testing, the mean is our page's conversion rate with the sample visitors.

Variance

The variance is a measure of variability across samples. It is officially the expectation of the squared deviation of a random variable from its mean.

In AB testing, the variance affects the sample size we need in order to have a chance of deriving statistically significant results.

Regression to the Mean

Regression to the mean is the phenomenon that if a variable is extreme on its first measurement, it will tend to be closer to the average on its second measurement.

In AB testing, this phenomenon ensures that as we continue increasing the sample size and the length of observation, the mean of our cumulative observations will get closer and closer to the true mean of the population.

Confidence Level & Interval

The confidence interval is an observed range in which a given percentage of test outcomes fall. The confidence interval is determined by the confidence level we manually select at the beginning of our test and influences the sample size required.

In AB testing, the confidence level is usually set to 95%, which gives us a range (the confidence interval) in which we know the mean will fall in at least 19 of 20 samples.

Margin of Error

The margin of error is a statistic expressing the amount of random sampling error in a sample's results. The margin for error is a function of the standard deviation, which is a function of the variance. Really all you need to know is that all of these terms are measures of variability across samples.

The margin of error is added to and subtracted from the mean to determine the confidence interval.

Null Hypothesis

The null hypothesis is a baseline assumption that there is no relationship between two data sets. When a statistical hypothesis test is run, the results either disprove the null hypothesis or they fail to disprove the null hypothesis.

In AB testing, the null hypothesis is the assumption that the original page and the new page variation have no statistical significant relationship.

Statistical Significance

Statistical significance is attained when the p-value is less than the significance level. And that is way too many new words in one sentence, so let's break down these terms real quick and then we'll summarize the entire concept in plain English.

In AB testing, statistical significance is how we verify that a new page truly outperforms the original.

Statistical Power

Statistical power is the probability that a test correctly rejects the null hypothesis.

In AB testing, it's the probability of correctly recognizing a page as a "winner" when that page does in fact convert at a higher rate.

Type I Error

A type I error occurs when we incorrectly reject the null hypothesis.

In AB testing, a type I error would occur if we concluded that Variation B was "better" than Variation A when it actually was not. These errors are avoided by achieving statistically significant results.

Type II Error

A type II error occurs when the null hypothesis is false, but we incorrectly fail to reject it.

To put this in AB testing terms, a type II error would occur if we concluded that Variation B was not "better" than Variation A when it actually was better. These errors are avoided by running tests with a high statistical power.

What can you learn from 30 minutes with a Conversion Scientist?

Joel Harvey and Brian Massey have over 36 years of combined experience.

Brian founded Conversion Sciences in 2007 and is the author of the book *Your Customer Creation Equation*.

Joel has worked with hundreds of IR 500 brands to increase the performance of their ecommerce web sites.

Click below to schedule a free strategy session with one of the web's foremost CRO experts!

[BOOK A FREE CONSULTATION](#)