What Is a Z-Test?

A z-test is a statistical test used to determine whether two population means are different when the variances are known, and the sample size is large.

The test statistic is assumed to have a normal distribution, and nuisance parameters such as standard deviation should be known in order for an accurate z-test to be performed.

KEY TAKEAWAYS

- A z-test is a statistical test to determine whether two population means are different when the variances are known and the sample size is large.
- A z-test is a hypothesis test in which the z-statistic follows a normal distribution.
- A z-statistic, or z-score, is a number representing the result from the z-test.
- Z-tests are closely related to t-tests, but t-tests are best performed when an experiment has a small sample size.
- Z-tests assume the standard deviation is known, while t-tests assume it is unknown.

Understanding Z-Tests

The z-test is also a hypothesis test in which the z-statistic follows a normal distribution. The z-test is best used for greater-than-30 samples because, under the central limit theorem, as the number of samples gets larger, the samples are considered to be approximately normally distributed.

When conducting a z-test, the null and alternative hypotheses, alpha and z-score should be stated. Next, the test statistic should be calculated, and the results and conclusion stated. A z-statistic, or z-score, is a number representing how many standard deviations above or below the mean population a score derived from a z-test is.

Examples of tests that can be conducted as z-tests include a one-sample location test, a two-sample location test, a paired difference test, and a maximum likelihood estimate. Z-tests are closely related to t-tests, but t-tests are best performed when an experiment has a small sample size. Also, t-tests assume the standard deviation is unknown, while z-tests assume it is

known. If the standard deviation of the population is unknown, the assumption of the sample variance equaling the population variance is made.

One-Sample Z-Test Example

Assume an investor wishes to test whether the average daily return of a stock is greater than 3%. A simple random sample of 50 returns is calculated and has an average of 2%. Assume the standard deviation of the returns is 2.5%. Therefore, the null hypothesis is when the average, or mean, is equal to 3%.

Conversely, the alternative hypothesis is whether the mean return is greater or less than 3%. Assume an alpha of 0.05% is selected with a two-tailed test. Consequently, there is 0.025% of the samples in each tail, and the alpha has a critical value of 1.96 or -1.96. If the value of z is greater than 1.96 or less than -1.96, the null hypothesis is rejected.

The value for z is calculated by subtracting the value of the average daily return selected for the test, or 1% in this case, from the observed average of the samples. Next, divide the resulting value by the standard deviation divided by the square root of the number of observed values.

Therefore, the test statistic is:

$$(0.02 - 0.01) \div (0.025 \div \sqrt{50}) = 2.83$$

The investor rejects the null hypothesis since z is greater than 1.96 and concludes that the average daily return is greater than 1%.

What's the Difference Between a T-Test and Z-Test?

Z-tests are closely related to t-tests, but t-tests are best performed when the data consists of a small sample size, i.e., less than 30. Also, t-tests assume the standard deviation is unknown, while z-tests assume it is known.

When Should You Use a Z-Test?

If the standard deviation of the population is unknown and the sample size is greater than or equal to 30, then the assumption of the sample variance equaling the population variance should be made using the z-test. Regardless of the sample size, if the population standard deviation for a variable remains unknown, a t-test should be used instead.

What Is a Z-Score?

A z-score, or z-statistic, is a number representing how many standard deviations above or below the mean population the score derived from a z-test is. Essentially, it is a numerical measurement that describes a value's relationship to the mean of a group of values. If a z-score is 0, it indicates that the data point's score is identical to the mean score. A z-score of 1.0 would indicate a value that is one standard deviation from the mean. Z-scores may be positive or negative, with a positive value indicating the score is above the mean and a negative score indicating it is below the mean.

What Is Central Limit Theorem (CLT)?

In the study of probability theory, the central limit theorem (CLT) states that the distribution of sample approximates a normal distribution (also known as a “bell curve”) as the sample size becomes larger, assuming that all samples are identical in size, and regardless of the population distribution shape. Sample sizes equal to or greater than 30 are considered sufficient for the CLT to predict the characteristics of a population accurately. The z-test's fidelity relies on the CLT holding.

The Bottom Line

A z-test is used in hypothesis testing to evaluate whether a finding or association is statistically significant or not. In particular, it tests whether two means are the same (the null hypothesis). A z-test can only be used if the population standard deviation is known and the sample size is 30 data points or larger. Otherwise, a t-test should be employed.