Contents

Understanding the Z-Score Distribution in Hypothesis Testing

The fallacy in educating people in statistics is always failing to articulate the way graphs diagrams or data are derived whilst assuming that the student already understands this. A deep dive into the complications of a Z-score distribution curve is outlined in this short paper.

Formula: How to calculate Z-Score for several sample means

$$z = \frac{\bar{x} - \mu}{SEM}$$

Case Study For Visualization purposes

Imagine one is presented with the problem of performing a hypothesis test to determine if the systolic arterial pressure from 36 patients (considered a sample) differed from the population mean of 120. A good demonstration will be done using 5 tuples.

Systolic Pressure	Population Systolic Pressure	Technical Explanation
112	120	Each pressure point is now considered a sample mean; 112 represents an unknown sample
113	120	Each pressure point is now considered a sample mean; 113 represents an unknown sample
114 (Mean=Median)	120	Each pressure point is now considered a sample mean; 114 represents an unknown sample
115	120	Each pressure point is now considered a sample mean; 115 represents an unknown sample
116	120	Each pressure point is now considered a sample mean; 116 represents an unknown sample

The Z-score distribution is derived from each data point and each data point is the mean of an assumed sample. In reality, the assumed sample means (pseudo-means) are just individual data points assumed to be group means. The 0.05 alpha level then represents 5% of the data points or pseudo-means.

This is complicated even more because the mean of the sample means or pseudo-means (114) is then used to calculate a Z-score that will be used in hypothesis decisions. This seems like an easy realization but it is not the case.

Conclusion

A rather simple paper that expounds on what happens behind the scenes when performing a t-test. A t-test can be performed against 0 or against a pre-determined population mean. This paper forms the basics of hypothesis testing when comparing sample means to a population mean. In short when a hypothesis test is performed based on the scenario above, one should be able to decide whether samples (sample means) when averaged differ significantly or non-significantly from the population mean.