Chapter Nine - Hypothesis testing

Simona Simona

5/1/2020

The confidence interval is appropriate when our goal is to estimate population parameters. The second type of inference is directed at assessing the evidence provided by the data in favour of some claim about the population parameter. This is called Hypothesis Testing.

In social sciences, we select samples in order to learn more about populations of interest to us. In terms of the mean, we measure a sample mean to learn more about the population mean.

Hypothesis testing involves collecting the sample from the population and comparing the mean of the sample to that of assumed mean of the population. If the sample mean deviates substantially from that of the population mean, we reject the assumed population mean.

The logic here is that on average the value of the sample mean will equal the population mean. The larger the difference or discrepancy between the sample mean and the population mean, the less likely, it is that we could have selected that sample mean, if the value of the population mean is correct.

For example Note: Hypotheses are from theory and literature and they are claims about the population.

Suppose we read in the Post that children in Zambia watch an average of 3 hours of television per week. To test whether this claim is true we record the time (in hour) that a group of 50 Zambian children (the sample), among all children in Zambia (population) watch television. The mean we measure for these 50 samples in a sample mean. We can then compare the sample mean we select to the population mean (3 hours) stated in the article.

Hypothesis testing or significance testing is a method for testing a claim or hypothesis about a parameter in a population, using data measured in a sample. In this method, we test some hypothesis by determining the likelihood that a sample mean could have been selected, if the hypothesis regarding the population parameter were true.

The method of hypothesis testing can be summarised in four steps and these are:

Step 1: State the hypothesis
Step 2: Set the criteria for a decision
Step 3: Compute the test statistic
Step 4: Make a decision

Step 1: State the Hypothesis

We begin by stating the value of a population mean in a null hypothesis, which we presume is true. The null hypothesis, often represented as (H0) is the statement about a population parameter, such as the population mean, that is assumed to be true. In our watching television example, our null hypothesis would be H0 = 3 hours. We also need to state the alternative hypothesis. An alternative hypothesis (H1) is a statement that directly contradicts a null hypothesis by stating that, the actual value of a population parameter is les than, greater than, or not equal to the value stated in the null hypothesis. The alternative hypothesis states what we think is wrong about the null hypothesis.

H1 x 3 hours
The courtroom analogy.

Step 2: Set the Criteria for a Decision

To state the criteria for a decision, we state the level of significance, for the test. The level of significance refers to criteria of judgement upon which a decision is made regarding the value stated in a null hypothesis. The criterion is based on the probability of obtaining a statistics measured in a sample if the value stated in the null hypothesis were true.

In the social and behavioural science, the criterion or level of significance is typically set at 5%. When the probability of obtaining a sample mean is less than 5% if the null hypothesis were true, then we reject the value stated on the null hypothesis.

Step 3: Compute the Test Statistic

We use a test statistic to determine how likely the sample outcome is, if the population mean stated by the null hypothesis is true. Specifically the test statistic tells us how far or how many standard deviations, a sample mean is from the population mean. The larger the value of the test statistics, the further the distance, or number of standard deviations, sample mean is from the population mean stated in the null hypothesis. The value of the test statistic is used to value a deviation.

Step 4: Make a Decision

We use the value of the test statistic to make a decision about the null hypothesis. The decision is based on the probability of obtaining a sample mean, given the value stated in the null hypothesis is true. If the probability of obtaining a sample mean is less than 5% when the null hypothesis is true, then the decision is to reject the null hypothesis. If the probability is greater than 5%, then the decision is retain the null hypothesis.

Making a Decision: Types of Error

Types of Error

In step 4, we decide whether to retain or reject the null hypothesis. Because we are observing a sample and not an entire population, it is possible that a conclusion may be wrong. Resignation

When we decide to reject the null hypothesis when infact it is true, we are committing a type I error. Type 1 error is the probability of rejecting a null hypothesis that is actually true. An alpha(d)level is the level of significance or criterion for a hypothesis test. It is the largest probability of committing a type I error that we will allow and still decide to reject the null hypothesis. This criterion is usually set at 0.05 ($\alpha = 0.05$). When the probability of a type I error is less than 5% ($p < 0.05$) we decide to reject the null hypothesis, otherwise, we retain the null hypothesis.

Type II error on the other hand is the probability of retaining a null hypothesis that is actually false. The relationship between type I and type II error is inverse i.e. if we change.

Directional vs Non Directional Tests

You can recall that when formulating the alternative hypothesis, that it can be formulated to say the population mean is greater than ($>$), less than ($<$) or not equal ($\neq$) to the value stated in the null hypothesis.

When the alternative hypothesis indicates the direction of the hypothesis (showing whether the population mean is greater or less than a given value) this is a directional test, and it is called a non-tailed hypothesis test. We use the lower or upper tail of the sampling distribution to place the level of significance.

When the alternative hypothesis on the other hand does not show the direction (the population mean is not equal $\neq$) to a given value, then that is a non- directional hypothesis test. We use both tails of the sampling distribution to place the level of significance.

Example

Suppose you are studying the expenditure patterns of students of UNZA and you say students spend an average of K10 per day, you therefore want to see if this is true. You select a sample of 144 students, you compute the mean and find that it is X = 13 and standard deviation is K11.

Step 2: The significance 0.05

We reject the null hypothesis if only if the significance in p< 0.05

Step 3: Compute the test statistic. This is basically the Z – Scores

Step 4: Decision significance <0.05 = we reject the null hypothesis.
Meaning that the claim that students spend K10 per day is not statistically significant.