Chapter 11: Hypothesis Testing

Lesson 11.1: The Alpha Value

The alpha value is the degree of risk we are willing to take when making a decision. The alpha value, often abbreviated using the Greek letter \(\alpha\), is sometimes called the “level of significance”.

The use of alpha value is probably best explained using an example:

Suppose we are interested in determining if the average final exam score for the students in our Management class is significantly different from the average score of other Management students throughout the university. In order to determine this, we would have to test the following null hypothesis:

*There is no significant difference in exam scores between our Management students and other Management students throughout the university.*

Now, let’s imagine we have computed an average score of 92 for our students and found the population average to be 80. Because the difference between these mean scores are fairly large, it appears we should reject our null hypothesis (i.e., there appears to be a significant difference between the two groups).

**Question:** How do we know when the sample mean and the population mean are different due to chance and when they are significantly different from one another?

**Bad Note:** Since we are dealing with a sample of data, we are never 100% sure if the differences are due to chance or represent a significant difference.

**Good Note:** By using our alpha value we can control the risk of making the wrong decision.

**Four Things Alpha Value Can Do**
1. We can use it to indicate the percentage of time we are willing to make an incorrect decision based on the sample data we have collected.
2. We can control the level of risk we are willing to take by using different alpha values (e.g., .10 and .01), although it is generally set at .05 or 5% for the types of studies we are most likely to undertake.
3. The alpha value can help us create a range called a “confidence level”.
4. The alpha value can help us test a hypothesis about the population using data drawn from a sample.

Lesson 11.2: The Type I Error and the Type II Error

**Type I Error Rate**
It is also refer to as the alpha. This is also the probability of making a Type I Error.

**Type I Error**
This error happens if we reject a null hypothesis when we should not have (i.e., due to chance). We can control for Type I Error using our alpha value.
Type II Error Rate
It is also refer to as the beta. This is also the probability of making a Type II Error.

Type II Error
This error occurs when, because of sampling error, we fail to reject a null hypothesis when we actually should reject it (i.e., we say the values are not significantly different when they really are).

Criticality of Type I and Type II Errors
Suppose we have a research hypothesis that reads:

The defendant is guilty.

The null hypothesis would read:

The defendant is not guilty.

What will happen to the defendant if you will commit a Type I Error?
In this case, a Type I Error would mean that an innocent person would be sentenced to jail because you would be supporting the research hypothesis.

What will happen to the defendant if you will commit a Type II Error?
Failing to reject the null and not supporting the research hypothesis due to sampling error would mean a criminal would be found not guilty and be allowed to go free.

Question: Which is worse, an innocent man in jail or a criminal on the street?

<table>
<thead>
<tr>
<th>Table 11.1 Type I and Type II Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>If we do not reject the null hypothesis</td>
</tr>
<tr>
<td>If null hypothesis is <strong>TRUE</strong></td>
</tr>
<tr>
<td>If null hypothesis is <strong>FALSE</strong></td>
</tr>
</tbody>
</table>

Lesson 11.3: The Confidence Interval

Predicting a Population Parameter Based on a Sample Statistics Using Confidence Intervals
In this section, we are only going to look at making estimations where one sample taken from a population. In doing so, we will learn to estimate a population parameter (e.g., a population mean) based on a sample statistics (e.g., a sample mean).

Let’s go back to our example using the students in the Management class. This time we want to use a sample statistic to predict a population parameter. In this case, we are going to use our class average to estimate the average for all of the other Management students.

Confidence Interval
The confidence interval is a range of numbers around a sample statistic within which the true value of the population is likely to fall. Here, we can build a range around the sample statistic, called the confidence interval, within which we predict the population parameter will fall.
How to compute for the confidence interval?

To compute the confidence interval you need two things:
1. a sample from the population; and
2. an alpha value.

Example 1:
Let’s say the Sample size of the Management class we randomly selected is 50. For this Management class, the average test score is 92. We are going to set an alpha value of .05 – this means we are going to have 95% confidence interval. In this case, we will use a population standard deviation of 5.

Formula in Computing the Confidence Interval:

\[
\text{Confidence Interval (CI)} = \bar{x} \pm \left( \frac{z_{\alpha/2}}{\sqrt{n}} \right) \left( \frac{\sigma}{\sqrt{n}} \right)
\]

where:
- \( \bar{x} \) stands for the mean of the sample we are dealing with.
- \( z \) is the z value for the alpha level we chose.
- \( \alpha \) is our alpha value
- \( \sqrt{n} \) is the square root of the sample size we are dealing with.
- \( \sigma \) is our sigma value and represents the population standard deviation.

Question: Why do you think in our formula for finding the confidence interval the z score for alpha value is divided by 2?

Did you know?
Many beginning statisticians think a 95% confidence interval means we are 95% sure the confidence interval contains the population parameter of interest. This isn’t true. The confidence interval you compute around the sample statistic either contains the population parameter or it doesn’t.

Example 2:
Let’s say the sample size of the Management class we randomly selected is 100. For this Management class, the average test score is 70. We are going to set an alpha value of .05 – this means we are going to have 95% confidence interval. In this case, we will use a population standard deviation of 10. (Solution on the Board)

General Conclusion: Here, we’re 95% confident that our population average is somewhere between 68.04 and 71.96. Again, we’re not saying that 95 out of 100 values fall in this range; this is just the probability that any given value does.

Confidence Interval for Alpha .01 and Alpha .10
How can we determine the z values for .01 alpha value and for .10 alpha value? Unfortunately, it is not quite so easy to determine the z scores for the other common alpha values.

Task 1: Compute for the z values of .01 alpha value and .10 alpha value on the board.

Task 2: Using Example 1, determine the confidence interval using .10 alpha value.
Question: What can you observe to the confidence interval we have using .05 alpha value? Compare it to .10 alpha value confidence interval.

Task 3: Using Example 1 determine the confidence interval using .01 alpha value.

Table 10.2. z Scores for Common Confidence Intervals

<table>
<thead>
<tr>
<th>Confidence Interval (CI)</th>
<th>( \alpha )</th>
<th>( \alpha/2 )</th>
<th>Area between mean and tails</th>
<th>( z_{\alpha/2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>.90</td>
<td>.10</td>
<td>.05</td>
<td>45%</td>
<td>1.645</td>
</tr>
<tr>
<td>.95</td>
<td>.05</td>
<td>.025</td>
<td>47.5%</td>
<td>1.96</td>
</tr>
<tr>
<td>.99</td>
<td>.01</td>
<td>.005</td>
<td>49.5%</td>
<td>2.575</td>
</tr>
</tbody>
</table>

Example 3:
Suppose you are hired as a school psychologist and are interested in finding the average IQ in your school. You do not have time to test all 500 students in your school and you want to work with a representative sample of 30. Let’s assume that you have a sample size of 30 with a mean IQ of 110; assume further you have a population standard deviation of 15. Now, determine the confidence intervals using the common alpha values.

Be Careful When Changing Your Alpha Values
The size of your alpha value is inversely related to the width of your confidence interval; larger alpha values lead to smaller confidence intervals and vice versa. Of course, this creates havoc when you’re trying to estimate a population parameter. Do you try to “cast as large a net” as possible with a small alpha value, or do you try to be as accurate as possible with a large alpha value? Unfortunately, this is the dilemma that statisticians face when dealing with uncertainty. How do we strike a happy medium? Easy – set alpha equal to .05.

Let’s Practice
Direction: In the table below you are supplied with the sample mean, the alpha value, the number of values in the sample, and the population standard deviations. Use these values to compute the lower and upper limits of the confidence interval. Kindly show your solution and final answers on yellow paper.

<table>
<thead>
<tr>
<th>( \bar{x} )</th>
<th>( \alpha )</th>
<th>( n )</th>
<th>( \sigma )</th>
<th>Lower Limit of CI</th>
<th>Upper Limit of CI</th>
<th>Width of CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>.10</td>
<td>25</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>.05</td>
<td>50</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>.01</td>
<td>15</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>.05</td>
<td>20</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>.01</td>
<td>22</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>220</td>
<td>.10</td>
<td>40</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>