7.10: Different Significance Level

Finally, let's take a look at an example phrased in generic terms, rather than in the context of a specific research question, to see the individual pieces one more time. This time, however, we will use a stricter significance level, α = 0.01, to test the hypothesis.

Step 1: State the Hypotheses We will use 60 as an arbitrary null hypothesis value:

H_0: The average score does not differ from the population

$H_0 : \mu = 50$

We will assume a two-tailed test:

H_A: The average score does differ

$H_A : \mu \neq 50$

Step 2: Find the Critical Values We have seen the critical values for z-tests at α = 0.05 levels of significance several times. To find the values for α = 0.01, we will go to the standard normal table and find the z-score cutting of 0.005 (0.01 divided by 2 for a two-tailed test) of the area in the tail, which is $z*$ = ±2.575. Notice that this cutoff is much higher than it was for α = 0.05. This is because we need much less of the area in the tail, so we need to go very far out to find the cutoff. As a result, this will require a much larger effect or much larger sample size in order to reject the null hypothesis.

Step 3: Calculate the Test Statistic We can now calculate our test statistic. We will use σ = 10 as our known population standard deviation and the following data to calculate our sample mean:

61	62
65	61
58	59
54	61
60	63

The average of these scores is $\overline{\mathrm{X}}$= 60.40. From this we calculate our z-statistic as:

$$z = \frac{60.40 - 60.00}{10.00/\sqrt{10}} = \frac{0.40}{3.16} = 0.13$$

Step 4: Make the Decision Our obtained z-statistic, z = 0.13, is very small. It is much less than our critical value of 2.575. Thus, this time, we fail to reject the null hypothesis. Our conclusion would look something like:

Based on the sample of 10 scores, we cannot conclude that there is no effect causing the mean ($\overline{\mathrm{X}}$= 60.40) to be statistically significantly different from 60.00, z = 0.13, p > 0.01.

Notice two things about the end of the conclusion. First, we wrote that p is greater than instead of p is less than, like we did in the previous two examples. This is because we failed to reject the null hypothesis. We don't know exactly what the p-value is, but we know it must be larger than the α level we used to test our hypothesis. Second, we used 0.01 instead of the usual 0.05, because this time we tested at a different level. The number you compare to the p-value should always be the significance level you test at.

Finally, because we did not detect a statistically significant effect, we do not need to calculate an effect size.