4.5 Confidence intervals, hypothesis tests, and significance tests

The use of the symbol α in the development of both confidence intervals and hypothesis testing is not a coincidence.

In confidence intervals, the value of α determines the value of the stated coverage of the confidence interval, which is $1-\alpha$. If the confidence interval $L < \theta < U$ is an exact confidence interval, for instance, then $P(L < \theta < U) = 1-\alpha$.

In hypothesis testing, the value of the significance level α is the probability of committing a Type I error, that is, $\alpha = P(\text{Type I error}) = P(\text{rejecting } H_0 \,|\, H_0 \text{ true})$.

This subsection links the α that is used in constructing a confidence interval with the α used in hypothesis testing. Assume that the goal associated with the analysis of a data set is to conduct a test. As seen in Table 4.5, such a test associated with the null hypothesis $H_0 : \theta = \theta_0$ can be conducted in three different ways: using confidence intervals, using traditional hypothesis testing, and using significance testing. Each column in Table 4.5 is associated with one of the three methods of conducting the test. The rows in Table 4.5 show the sequential steps associated with the three methods. The first three steps in each of the three methods are identical. The methods differ in the last two steps. Regardless of the methods taken, the conclusion that is drawn, either reject H_0 or fail to reject H_0, will be identical for the same data set which is based on the same pivotal value/test statistic.

Confidence interval method	**Hypothesis testing method**	**Significance testing method**
Formulate H_0 and H_1	Formulate H_0 and H_1	Formulate H_0 and H_1
Determine significance level α	Determine significance level α	Determine significance level α
Determine sample size n	Determine sample size n	Determine sample size n
Construct confidence interval	Compute test statistic	Compute test statistic
Reject H_0 if θ_0 does not fall in the confidence interval	Reject H_0 if the test statistic is in the critical region	Reject H_0 if the p-value is less than α

Table 4.5: Steps associated with confidence intervals, hypothesis tests, and significance tests.

The next example illustrates these three methods for conducting a test based on a single data set. The three methods all result in the same conclusion being drawn.

Example 4.21 IQ scores are traditionally centered around a population mean of 100. An experimenter would like to test the simple null hypothesis

$$H_0 : \mu = 100$$

versus the two-sided alternative hypothesis

$$H_1 : \mu \neq 100$$

based on a random sample of IQ scores that is assumed to come from a normally distributed population. The significance level is set to $\alpha = 0.01$ prior to collecting the data. It is also decided that a sample of $n = 78$ IQ scores will be collected in order to conduct the test. More specifically, the $n = 78$ IQ scores of seventh-grade students at a rural school in the midwest of the United States from Example 3.9, and also listed below, will be used to conduct the test. This completes the first three common steps for the three methods given in Table 4.5.