CHAPTER 6

# Confidence Intervals and Hypothesis Tests



## 6.0 What We Need to Know When We Finish This Chapter

This chapter reviews the topics of *confidence intervals* and *hypothesis tests*. Confidence intervals give us ranges that contain the parameters with prespecified degrees of certainty. They are more useful if they are narrower. Hypothesis tests evaluate whether the data at hand are consistent or inconsistent with prespecified beliefs about parameter values. They are more useful if they are unlikely to contradict these beliefs when the beliefs are really true, and if they are unlikely to be consistent with these beliefs when the beliefs are really false. Here are the essentials.

1. **Equation (6.7), section 6.2:** The fundamental equation of this chapter is

$$1-\alpha=\mathrm{P}\left(-t_{\alpha/2}^{(df)}\leq\frac{d-\delta}{\mathrm{SD}(d)}\leq t_{\alpha/2}^{(df)}\right).$$

2. **Equation (6.8), section 6.3:** Confidence intervals consist of *known* boundaries with a fixed probability of containing the *unknown value* of the parameter of interest. The general expression is

$$1-\alpha=\mathrm{P}\left(d-t_{\alpha/2}^{(df)}\mathrm{SD}(d)\leq\delta\leq d+t_{\alpha/2}^{(df)}\mathrm{SD}(d)\right).$$

   Confidence intervals ask the data for *instruction*.

3. **Section 6.4:** Hypothesis tests ask the data for *validation*. The null hypothesis is the opposite of what we expect to find. Estimates in the *acceptance region* validate the null hypothesis. Estimates in the *rejection region* contradict it.
4. **Equation (6.14), section 6.4.1:** The two-sided hypothesis test is

$$1-\alpha=\mathrm{P}\left(\delta_0-t_{\alpha/2}^{(df)}\mathrm{SD}(d)<d<\delta_0+t_{\alpha/2}^{(df)}\mathrm{SD}(d)\right).$$

5. **Section 6.4.1:** Reject the null hypothesis when the estimate falls in the rejection region, the *test statistic* is greater than or equal to the *critical value*, or the *p-value* is less than or equal to the *significance level*. These decision rules are all equivalent.
6. **Equation (6.30), section 6.4.2:** The one-sided, upper-tailed hypothesis test is

$$1-\alpha=\mathrm{P}\left(d<\delta_0+t_{\alpha}^{(df)}\mathrm{SD}(d)\right).$$

7. **Section 6.4.3:** The *size* of the test is its *significance level*, the *probability of a Type I error*. A Type I error occurs when the null hypothesis is rejected even though it is true. It is the statistical equivalent of convicting an innocent person.
8. **Equation (6.36), section 6.4.3:** A Type II error occurs when the null hypothesis is accepted even though it is false. It is the statistical equivalent of acquitting a guilty person. The *power* of the test is the

probability of rejecting the null hypothesis when it is false, or one minus the probability of a Type II error.

$$\begin{aligned}\mathrm{P}\left(\text{Type II error}\right) &= \mathrm{P}\left(\frac{d-\delta_1}{\mathrm{SD}\left(d\right)} \le \frac{\delta_0 + t_\alpha^{(df)}\mathrm{SD}\left(d\right)-\delta_1}{\mathrm{SD}\left(d\right)}\right)\\ &= \mathrm{P}\left(t^{(df)} \le \frac{\delta_0 + t_\alpha^{(df)}\mathrm{SD}\left(d\right)-\delta_1}{\mathrm{SD}\left(d\right)}\right).\end{aligned}$$

9. **Section 6.4.3:** All else equal, reducing the probability of either a Type I or a Type II error increases the probability of the other.
10. **Section 6.4.3:** *Statistical distance* is what matters, not *algebraic distance*. The standard deviation of the estimator is the *metric* for the statistical distance.
11. **Section 6.5:** Any value within the confidence interval constructed at the $(1-\alpha)\%$ confidence level would, if chosen as the null hypothesis, not be rejected by a two-sided hypothesis test at the $\alpha\%$ significance level.