

## CHAPTER OVERVIEW

### 7: Testing the Significance of Data

A confidence interval is a useful way to report the result of an analysis because it sets limits on the expected result. In the absence of determinate error, or bias, a confidence interval based on a sample's mean indicates the range of values in which we expect to find the population's mean. When we report a 95% confidence interval for the mass of a penny as  $3.117 \text{ g} \pm 0.047 \text{ g}$ , for example, we are stating that there is only a 5% probability that the penny's expected mass is less than 3.070 g or more than 3.164 g.

Because a confidence interval is a statement of probability, it allows us to consider comparative questions, such as these:

“Are the results for a newly developed method to determine cholesterol in blood significantly different from those obtained using a standard method?”

“Is there a significant variation in the composition of rainwater collected at different sites downwind from a coal-burning utility plant?”

In this chapter we introduce a general approach that uses experimental data to ask and answer such questions, an approach we call significance testing.

The reliability of significance testing recently has received much attention—see Nuzzo, R. “Scientific Method: Statistical Errors,” *Nature*, **2014**, 506, 150–152 for a general discussion of the issues—so it is appropriate to begin this chapter by noting the need to ensure that our data and our research question are compatible so that we do not read more into a statistical analysis than our data allows; see Leek, J. T.; Peng, R. D. “What is the Question? *Science*, **2015**, 347, 1314–1315 for a useful discussion of six common research questions.

In the context of analytical chemistry, significance testing often accompanies an exploratory data analysis

“Is there a reason to suspect that there is a difference between these two analytical methods when applied to a common sample?”  
or an inferential data analysis.

“Is there a reason to suspect that there is a relationship between these two independent measurements?”

A statistically significant result for these types of analytical research questions generally leads to the design of additional experiments that are better suited to making predictions or to explaining an underlying causal relationship. A significance test is the first step toward building a greater understanding of an analytical problem, not the final answer to that problem!

[7.1: Significance Testing](#)

[7.2: Significance Tests for Normal Distributions](#)

[7.3: Analysis of Variance](#)

[7.4: Non-Parametric Significance Tests](#)

[7.5: Using R for Significance Testing and Analysis of Variance](#)

[7.6: Exercises](#)

---

This page titled [7: Testing the Significance of Data](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by David Harvey.