3

Testing Hypotheses

It's understandable that one might elect to specialize in the preparation of Chinese cuisine rather than Greek or vice versa. There's just so much time available. But to decide to eat only the one rather than the other is to act the fool.

Now substitute the words "parametric test" for "Chinese cuisine," "nonparametric" for "Greek,"and "use" for "eat" in the above paragraph and read it again. In this chapter, you learn how to approach and resolve a series of testing problems of increasing complexity, specifically, tests for location and scale parameters in one and two samples. You learn how to derive confidence intervals for the unknown parameters.

3.1 Testing a Simple Hypothesis

Our first challenge, that of finding the most powerful test of a simple hypothesis against a simple alternative, is chosen not for its practical applications but because its solution is fundamental to the solution of all other testing problems.

Suppose, first, we are trying to decide between two discrete probability distributions P_0 and P_1 such that $P_i\{X = k\} = P_i[k]$ for $i = 0, 1$; $k = 0, 1, \ldots$. We would like to designate a set of the possible values of X as our rejection region R such that, if k belongs to R, we will reject the simple hypothesis P_0 in favor of the simple alternative P_1. We specify a significance level α and require that $\sum_{k \in R} P_0[k] \leq \alpha$. Given this restriction, our objective is to choose the values of k to include in R so that the power, $\sum_{k \in R} P_1[k]$, is a maximum.

Let $r(k) = P_1[x]/P_0[x]$. To maximize the power against P_1, we need to include in R all the points with the highest values of r until we attain the desired significance level. That is, there exists a constant c such that, if $r(k) > c$, k is to be included in R. If $r[k] < c$, then we will accept the hypothesis.

What if $r[k] = c$? That depends. The answer according to theoreticians is that if $\sum_{k>c} P_0[k] = \alpha' < \alpha$, then, when $r[k] = c$, we pick a random number

from 0 to 1. If this random number is less than or equal to $(\alpha - \alpha')/P_0[c]$ we reject the hypothesis; otherwise we accept it.

In Section 15.2, we show that a similar result holds when $P_0[x]$ and $P_1[x]$ are continuous distributions (or a mixture of continuous and discrete). This latter result is the fundamental lemma of Neyman and Pearson [1928].

3.2 One-Sample Tests for a Location Parameter

One of the simplest of practical testing problems would appear to be that of testing for the value of the location parameter of a distribution $\mathrm{F}(\theta)$ using a series of observations $x_1, x_2, \ldots, x_n$ from that distribution.

3.2.1 A Permutation Test

This *semiparametric*[1] testing problem is a simple one *if* we can assume that the underlying distribution is symmetric about the unknown parameter θ, that is, if

$$\Pr\{X \leq \theta - x\} = F(\theta - x) = 1 - F(\theta + x) = \Pr\{X \geq \theta + x\}, \text{for all } x.$$

The normal distribution, with its familiar symmetric bell-shaped curve, the double exponential, Cauchy, and uniform distributions are examples of symmetric distributions.[2] The difference of two independent observations drawn from the same population also has a symmetric distribution, as you will see when we come to consider experiments involving matched pairs in Section 5.2.2.2.

Suppose now we wish to test the hypothesis that $\theta \leq \theta_0$ against the alternative that $\theta > \theta_0$. As in Chapter 1, we proceed in four steps:

First, we choose a test statistic that will discriminate between the hypothesis and the alternative. As one possibility, consider the sum of the deviations of θ about θ_0. Under the hypothesis, positive and negative deviations ought to cancel and this sum should be close to zero. Under the alternative, positive terms should predominate and this sum should be large. But how large should the sum be for us to reject the hypothesis?

We saw in Chapter 1 that we can use the permutation distribution to obtain the answer; but what should we permute? The principle of *sufficiency* can help us here.

Suppose we had lost track of the signs (plus or minus) of the deviations. We could attach new signs at random, selecting a plus or a minus with equal

[1] A problem is *parametric* if the form of the underlying distribution is known, and it is *nonparametric* if we have no knowledge concerning the distribution(s) from which the observations are drawn.

[2] These distributions are described in more detail in Chapter 5.