12

Solving the Unsolved and the Insolvable

In this chapter, we consider the problem of developing optimal solutions for yet-to-be-encountered problems, problems for which test statistics are not immediately to be found in the pages of this or any other text. First, we review the criteria for test statistics that we first encountered in Chapters 2 and 3. Then, we consider some permutation test statistics that have been developed in some highly specialized situations. Last, to be used when all else fails, we consider bootstrap confidence intervals.

12.1 Key Criteria

In virtually all the instances we have studied to this point in our text, the "obvious" test statistic is one that tends to be very large under the alternative (or very small), while under a null hypothesis no value is more likely than any other. The formal justification for this approach comes from the fundamental lemma of Neyman and Pearson, and if our statistic is sufficient for the parameters we are testing, then we can be almost certain we've made the correct choice.

12.1.1 Sufficient Statistics

Recall that a statistic $T(X)$ is *sufficient* for a parameter θ if the conditional distribution of X given T is independent of θ. Once we have calculated the value of a sufficient statistic or statistics, we may be able to throw away the original observations, for frequently, a sufficient statistic(s) can provide us with all the information a sample has to offer.

An example we have encountered many time in the derivation of permutation tests is that of the order statistics $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$. If we know these order statistics, we know as much about the unknown distribution as we would if we had the original observations in hand.

Another commonly encountered example is that of the mean of a sample of independent, identically Poisson-distributed random variables; this statistic is sufficient to represent the mean of the underlying Poisson distribution. Likewise the mean of a sample of normally distributed random variables is sufficient to represent the mean of the underlying normally distributed population. But there is a distinction: In the first example, the Poisson, the sample mean possesses all the information the sample has to offer with regard to the underlying single-parameter distribution. In the second case, a normal distribution depends on two parameters, the population mean and the population variance. We need to compute both the sample mean and the sample variance to obtain all the information a sample from a normal distribution has to offer.

Even in the case of a normal distribution, as it is a member of an exponential family, we were able to derive in Chapter 3 a test of the population mean, conditional on the value of a sufficient statistic.

In selecting a statistic to test a hypothesis about a population parameter θ, we look first at those statistics that are sufficient for θ.

12.1.2 Three Stratagems

Occasionally—the k-sample comparison of means when $k > 2$ is an excellent example—the use of sufficient statistics alone will not reduce consideration to a single statistic. Three stratagems may help us. We may

- restrict the alternatives;
- consider the loss function;
- invoke impartiality.

12.1.3 Restrict the Alternatives

In the k-sample case, by restricting attention to ordered alternatives, we were able to obtain a UMP-unbiased test (Theorem 6.2). In the next example, that of an $r \times 1$ contingency table, we cannot derive a most powerful test that will protect us against all alternatives, but we can use the likelihood ratio to derive a most powerful test against those alternatives that are of immediate interest. The approach lends itself to any set of data for which we have knowledge of an underlying model.

Suppose the hypothesis to be tested is that certain events (births, deaths, accidents) occur randomly over a given time interval. If we divide this time interval into m equal parts and p_i denotes the probability of an event in the ith subinterval, the null hypothesis becomes H: $p_i = 1/m$ for $i = 1, \ldots, m$. Our test statistic is

$$\chi^2 = mn \sum_{i=1}^{m} \left(v_i - \frac{1}{m} \right)^2 ,$$

where v_i is the relative frequency of occurrence in the ith interval.