6

Experimental Designs

In this chapter and the next, you learn to analyze the results of complex experimental designs that may involve multiple control variables, covariates, and restricted randomization.

The material is advanced and the discussion presupposes you have already completed Chapters 1–3.

6.1 Invariance

In Section 2.1.6, we discussed the importance of impartiality in a test. We shall apply the principle of impartiality repeatedly in subsequent sections.

Let X be distributed according to a probability distribution $P_\theta, \theta \in \Omega$ and let G be a group of transformations of the sample space, such as transformations of scale or zero point. Denote by gX the random variable that takes on the value gx when $X = x$ and suppose the distribution of gX is P_{θ^*}. Suppose that $\theta * \in \Omega$; that is, the transformation g on the sample space induces a transformation g^* on the parameter space Ω with $P_{g^*\theta}(gA) = P_\theta(A)$ for all events A belonging to the sample space. If the distributions P_θ corresponding to different values of θ are distinct, then a group of transformations G on the sample space induces a group of transformations G^* on the parameter space (Exercise 6.1).

We shall write that the problem of testing H: $\theta \in \Omega_H$ against H: $\theta \in \Omega_K$ remains invariant with respect to a group of transformations G, providing the induced group of transformations G^* leaves Ω_H and Ω_K distinct. In such a case, any UMP invariant test will also be most stringent (Exercise 6.2).

A function M will be said to be *maximal invariant* with respect to G if it is invariant with respect to G and if $M(x) = M(y)$ implies $y = gx$ for some $g \in G$.

Theorem 6.1. *Let X be distributed according to a probability distribution P_θ and let G be a group of transformations of the sample space. If $M(x)$ is*

invariant under G and if $v(\theta)$ is maximal invariant under the induced group G^*, then the distribution of $M(x)$ depends only on $v(\theta)$.

Proof. Let $v(\theta_1) = v(\theta_2)$. Then $\theta_2 = g^*\theta_1$ so

$$\begin{aligned} P\{M(x) \in A|\theta_2\} &= P\{M(x) \in A|g^*\theta_1\} = P\{M(gx) \in A|\theta_1\} \\ &= P\{M(x) \in A|\theta_1\}. \end{aligned}$$

□

We make use of this theorem in what follows to reduce the number of potential statistics to those that involve only a single maximal invariant with respect to one or more groups of transformations.

6.1.1 Some Examples

If $X_1, \ldots, X_n$ is a sample from $N(\mu, \sigma^2)$, the hypothesis $\sigma \geq \sigma_0$ remains invariant under transformations of the zero point $X_i' = X_i + c$. $U = \sum_{i=1}^n (x_i - \bar{x})^2$ is a maximal invariant (Exercise 6.3) and in accordance with the arguments presented in Section 3.5.2, the test that rejects when $U \leq C$ is UMP among tests that remain invariant under transformations of the zero-point. The power of this test is a constant on each of the sets $\{(\mu, \sigma)\colon -\infty < \mu < \infty, \sigma = \sigma'\}$. As it is the most powerful test on each such set, it also is most stringent.

Suppose now that we conduct a series of side-by-side comparisons under varying conditions, that is, the members of each pair differ only in the treatment that is applied, but the various pairs may be handled quite differently. Let the probability that the first or control member of each pair prove superior be $1 - p_i$. We often wish to test the hypothesis that the treatment has no effect, that is, $p_i = \frac{1}{2}$, against the alternative that the treatment is superior, that is, $p_i > \frac{1}{2}$ for all i.

This problem remains invariant under all permutations of the n observations, and a maximal invariant is the total number of instances S in which the treatment is superior. The distribution of S is $P\{S = k\} = \prod_i q_i \sum_{\pi_j} \left(\prod_i (p_{i_j}/q_{i_j})\right)$ where the summation extends over all $\binom{n}{k}$ choices of subscripts $i_1, \ldots, i_n$. The most powerful invariant test against the specific alternative $(p_1', \ldots, p_n')$ rejects the hypothesis when

$$f(k) = \sum_{\pi_j} \left(\prod_i (p_{i_j}'/q_{i_j}') \right) \Big/ \binom{n}{k} > C.$$

f is an increasing function of k (Exercise 6.4) so that regardless of the alternative, the test rejects when $k > C'$ and is UMP among all tests invariant under permutations of the subscripts.

Of course, a much more powerful test would be in a matched pairs comparison which utilizes the actual results of each experiment as in Section 6.4.2.2.