Chapter 6
Invariance

6.1 Symmetry and Invariance

Many statistical problems exhibit symmetries, which provide natural restrictions to impose on the statistical procedures that are to be employed. Suppose, for example, that $X_1, \ldots, X_n$ are independently distributed with probability densities $p_{\theta_1}(x_1), \ldots, p_{\theta_n}(x_n)$. For testing the hypothesis $H : \theta_1 = \cdots = \theta_n$ against the alternative that the θ's are not all equal, the test should be symmetric in $x_1, \ldots, x_n$, since otherwise the acceptance or rejection of the hypothesis would depend on the (presumably quite irrelevant) numbering of these variables.

As another example consider a circular target with center O, on which are marked the impacts of a number of shots. Suppose that the points of impact are independent observations on a bivariate normal distribution centered on O. In testing this distribution for circular symmetry with respect to O, it seems reasonable to require that the test itself exhibit such symmetry. For if it lacks this feature, a two-dimensional (for example, Cartesian) coordinate system is required to describe the test, and acceptance or rejection will depend on the choice of this system, which under the assumptions made is quite arbitrary and has no bearing on the problem.

The mathematical expression of symmetry is invariance under a suitable group of transformations. In the first of the two examples above the group is that of all permutations of the variables $x_1, \ldots, x_n$ since a function of n variables is symmetric if and only if it remains invariant under all permutations of these variables. In the second example, circular symmetry with respect to the center O is equivalent to invariance under all rotations about O. A third example is the following.

Example 6.1.1 (**Testing a Fair Coin**) Suppose X is the number of successes in n i.i.d. Bernoulli trials, each with success probability p. The problem is to test H : $p = 1/2$ against $K : \theta \neq 1/2$. Erich is given the data X and he seeks a test function $\phi = \phi(X)$ to test H. Meanwhile, Julie is given the number of failures $X' = n - X$. Then, X' is also binomial with parameters n and $\theta' = 1 - \theta$. From Julie's point of view, she faces the identical problem Erich faces, testing that a binomial distribution with n trials has success probability $1/2$ versus not $1/2$. It would be inconsistent

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, Springer Texts in Statistics, https://doi.org/10.1007/978-3-030-70578-7_6

for Erich and Julie to reach different conclusions, and thus invariance considerations would then require that

$$\phi(X) = \phi(n - X). \tag{6.1}$$

Tests satisfying (6.1) are said to be invariant, and such tests represent a restriction on the class of tests. The goal is to find a UMP level α test among such invariant tests. ■

In general, let X be distributed according to a probability distribution $P_\theta, \theta \in \Omega$, and let g be a transformation of the sample space $\mathcal{X}$. All such transformations considered in connection with invariance will be assumed to be $1:1$ transformations of $\mathcal{X}$ onto itself. Denote by gX the random variable that takes on the value gx when $X = x$, and suppose that when the distribution of X is $P_\theta, \theta \in \Omega$, the distribution of gX is $P_{\theta'}$ with θ' also in Ω. The element θ' of Ω which is associated with θ in this manner will be denoted by $\bar{g}\theta$, so that

$$P_\theta\{gX \in A\} = P_{\bar{g}\theta}\{X \in A\}. \tag{6.2}$$

Here the subscript θ on the left member indicates the distribution of X, not that of gX. Equation (6.2) can also be written as $P_\theta(g^{-1}A) = P_{\bar{g}\theta}(A)$ and hence as

$$P_{\bar{g}\theta}(gA) = P_\theta(A). \tag{6.3}$$

The parameter set Ω remains invariant under g (or is preserved by g) if $\bar{g}\theta \in \Omega$ for all $\theta \in \Omega$, and if in addition for any $\theta' \in \Omega$ there exists $\theta \in \Omega$ such that $\bar{g}\theta = \theta'$. These two conditions can be expressed by the equation

$$\bar{g}\Omega = \Omega. \tag{6.4}$$

The transformation $\bar{g}$ of Ω onto itself defined in this way is $1:1$ provided the distributions P_θ corresponding to different values of θ are distinct. To see this let $\bar{g}\theta_1 = \bar{g}\theta_2$. Then $P_{\bar{g}\theta_1}(gA) = P_{\bar{g}\theta_2}(gA)$ and therefore $P_{\theta_1}(A) = P_{\theta_2}(A)$ for all A, so that $\theta_1 = \theta_2$.

Lemma 6.1.1 *Let g, g' be two transformations preserving Ω. Then the transformations $g'g$ and g^{-1} defined by*

$$(g'g)x = g'(gx) \quad \text{and} \quad g(g^{-1}x) = x \quad \text{for all} \quad x \in \mathcal{X}$$

also preserve Ω and satisfy

$$\overline{g'g} = \overline{g'} \cdot \bar{g} \qquad \text{and} \qquad (\overline{g^{-1}}) = (\bar{g})^{-1}. \tag{6.5}$$

Proof. If the distribution of X is P_θ then that of gX is $P_{\bar{g}\theta}$ and that of $g'gX = g'(gX)$ is therefore $P_{\bar{g}'\bar{g}\theta}$. This establishes the first equation of (6.5); the proof of the second one is analogous. ■

We shall say that *the problem of testing* $H : \theta \in \Omega_H$ against $K : \theta \in \Omega_K$ *remains invariant* under a transformation g if $\bar{g}$ preserves both Ω_H and Ω_K, so that the equation

$$\bar{g}\Omega_H = \Omega_H \tag{6.6}$$

holds in addition to (6.4). Let $\mathcal{C}$ be a class of transformations satisfying these two conditions, and let G be the smallest class of transformations containing $\mathcal{C}$ such that $g, g' \in G$ implies that $g'g$ and g^{-1} belong to G. Then G is a group of transformations, all of which by Lemma 6.1.1 preserve both Ω and Ω_H. Any class $\mathcal{C}$ of transformations leaving the problem invariant can therefore be extended to a group G. It follows further from Lemma 6.1.1 that the class of induced transformations $\bar{g}$ form a group $\bar{G}$. The two equations (6.5) express the fact that $\bar{G}$ is a homomorphism of G.

In the presence of symmetries in both the sample and parameter spaces represented by the groups G and $\bar{G}$, it is natural to restrict attention to tests ϕ which are also symmetric, that is, which satisfy

$$\phi(gx) = \phi(x) \qquad \text{for all} \quad x \in X \quad \text{and } g \in G. \tag{6.7}$$

A test ϕ satisfying (6.7) is said to be *invariant under* G. The restriction to invariant tests is a particular case of the principle of invariance formulated in Section 1.5. As was indicated there and in the examples above, a transformation g can be interpreted as a change of coordinates. From this point of view, a test is invariant if it is independent of the particular coordinate system in which the data are expressed.[1]

A transformation g, in order to leave a problem invariant, must in particular preserve the class $\mathcal{A}$ of measurable sets over which the distributions P_θ are defined. This means that any set $A \in \mathcal{A}$ is transformed into a set of $\mathcal{A}$ and is the image of such a set, so that gA and $g^{-1}A$ both belong to $\mathcal{A}$. Any transformation satisfying this condition is said to be *bimeasurable*. Since a group with each element g also contains g^{-1} its elements are automatically bimeasurable if all of them are measurable. If g' and g are bimeasurable, so are $g'g$ and g^{-1}. The transformations of the group G above generated by a class $\mathcal{C}$ are therefore all bimeasurable provided this is the case for the transformations of $\mathcal{C}$.

6.2 Maximal Invariants

If a problem is invariant under a group of transformations, the *principle of invariance* restricts attention to invariant tests. In order to obtain the best of these, it is convenient first to characterize the totality of invariant tests.

---

[1] The relationship between this concept of invariance under reparametrization and that considered in differential geometry is discussed in Barndorff–Nielson, Cox and Reid (1986).

Let two points x_1, x_2 be considered equivalent under G,

$$x_1 \sim x_2 (\bmod G),$$

if there exists a transformation $g \in G$ for which $x_2 = gx_1$. This is a true equivalence relation, since G is a group and the sets of equivalent points, the *orbits* of G, therefore constitute a partition of the sample space. (Cf. Appendix, Section A.1.) A point x traces out an orbit as all transformations g of G are applied to it; this means that the orbit containing x consists of the totality of points gx with $g \in G$. It follows from the definition of invariance that a function is invariant if and only if it is constant on each orbit.

A function M is said to be *maximal invariant* if it is invariant and if

$$M(x_1) = M(x_2) \quad \text{implies} \quad x_2 = gx_1 \quad \text{for some } g \in G, \tag{6.8}$$

that is, if it is constant on the orbits but for each orbit takes on a different value. All maximal invariants are equivalent in the sense that their sets of constancy coincide.

Theorem 6.2.1 *Let $M(x)$ be a maximal invariant with respect to G. Then, a necessary and sufficient condition for ϕ to be invariant is that it depends on x only through $M(x)$; that is, that there exists a function h for which $\phi(x) = h[M(x)]$ for all x.*

PROOF. If $\phi(x) = h[M(x)]$ for all x, then $\phi(gx) = h[M(gx)] = h[M(x)] = \phi(x)$ so that ϕ is invariant. On the other hand, if ϕ is invariant and if $M(x_1) = M(x_2)$, then $x_2 = gx_1$ for some g and therefore $\phi(x_2) = \phi(x_1)$. ■

Example 6.2.1 (i) Let $x = (x_1, \ldots, x_n)$, and let G be the group of translations

$$gx = (x_1 + c, \ldots, x_n + c), \qquad -\infty < c < \infty.$$

Then the set of differences $y = (x_1 - x_n, \ldots, x_{n-1} - x_n)$ is invariant under G. To see that it is maximal invariant suppose that $x_i - x_n = x_i' - x_n'$ for $i = 1, \ldots, n-1$. Putting $x_n' - x_n = c$, one has $x_i' = x_i + c$ for all i, as was to be shown. The function y is of course only one representation of the maximal invariant. Others are for example $(x_1 - x_2, x_2 - x_3, \ldots, x_{n-1} - x_n)$ or the redundant $(x_1 - \bar{x}, \ldots, x_n - \bar{x})$. In the particular case that $n = 1$, there are no invariants. The whole space is a single orbit, so that for any two points there exists a transformation of G taking one into the other. In such a case the transformation group G is said to be *transitive*. The only invariant functions are then the constant functions $\phi(x) \equiv c$.

(ii) if G is the group of transformations

$$gx = (cx_1, \ldots, cx_n), \qquad c \neq 0,$$

a special role is played by any zero coordinates. However, in statistical applications the set of points for which none of the coordinates is zero typically has probability 1; attention can then be restricted to this part of the sample space, and the set of ratios $x_1/x_n, \dots, x_{n-1}/x_n$ is a maximal invariant. Without this restriction, two points x, x' are equivalent with respect to the maximal invariant partition if among their coordinates there are the same number of zeros (if any), if these occur at the same places, and if for any two nonzero coordinates x_i, x_j the ratios x_j/x_i and x'_j/x'_i are equal.

(iii) Let $x = (x_1, \dots, x_n)^\top$, and let G be the group of all orthogonal transformations $x' = Ox$ of n-space. Then $\sum x_i^2$ is maximal invariant, that is, two points x and x' can be transformed into each other by an orthogonal transformation if and only if they have the same distance from the origin. The proof of this is immediate if one restricts attention to the plane containing the points x, x' and the origin. ■

Example 6.2.2 (i) Let $x = (x_1, \dots, x_n)$, and let G be the set of $n!$ permutations of the coordinates of x. Then the set of ordered coordinates (*order statistics*) $x_{(1)} \le \cdots \le x_{(n)}$ is maximal invariant. A permutation of the x_i obviously does not change the set of values of the coordinates and therefore not the $x_{(i)}$. On the other hand, two points with the same set of ordered coordinates can be obtained from each other through a permutation of coordinates.

(ii) Let G be the totality of transformations $x'_i = f(x_i)$, $i = 1, \dots, n$, such that f is continuous and strictly increasing, and suppose that attention can be restricted to the points that have n distinct coordinates. If the x_i are considered as n points on the real line, any such transformation preserves their order. Conversely, if $x_1, \dots, x_n$ and $x'_1, \dots, x'_n$ are two sets of points in the same order, say $x_{i_1} < \cdots < x_{i_n}$ and $x'_{i_1} < \cdots < x'_{i_n}$, there exists a transformation f satisfying the required conditions and such that $x'_i = f(x_i)$ for all i. It can be defined, for example, as $f(x) = x + (x'_{i_1} - x_{i_1})$ for $x \le x_{i_1}$, $f(x) = x + (x'_{i_n} - x_{i_n})$ for $x \ge x_{i_n}$, and to be linear between x_{i_k} and $x_{i_{k+1}}$ for $k = 1, \dots, n-1$. A formal expression for the maximal invariant in this case is the set of *ranks* $(r_1, \dots, r_n)$ of $(x_1, \dots, x_n)$. Here the rank r_i of x_i is defined through

$$x_i = x_{(r_i)}$$

so that r_i is the number of x's $\le x_i$. In particular, $r_i = 1$ if x_i is the smallest x, $r_i = 2$ if it is the second smallest, and so on. ■

Example 6.2.3 Let x be an $n \times s$ matrix ($s \le n$) of rank s, and let G be the group of linear transformations $gx = xB$, where B is any nonsingular $s \times s$ matrix. Then a maximal invariant under G is the matrix $t(x) = x(x^\top x)^{-1}x^\top$, where $x^\top$ denotes the transpose of x. Here $(x^\top x)^{-1}$ is meaningful because the $s \times s$ matrix $x^\top x$ is nonsingular; see Problem 6.3. That $t(x)$ is invariant is clear, since

$$t(gx) = xB(B^\top x^\top xB)^{-1}B^\top x^\top = x(x^\top x)^{-1}x^\top = t(x).$$

To see that $t(x)$ is maximal invariant, suppose that

$$x_1(x_1^\top x_1)^{-1}x_1^\top = x_2(x_2^\top x_2)^{-1}x_2.$$

Since $(x_i^\top x_i)^{-1}$ is positive definite, there exist nonsingular matrices C_i such that $(x_i^\top x_i)^{-1} = C_i C_i^\top$ and hence

$$(x_1 C_1)(x_1 C_1)^\top = (x_2 C_2)(x_2 C_2)^\top.$$

This implies the existence of an orthogonal matrix Q such that $x_2 C_2 = x_1 C_1 Q$ and thus $x_2 = x_1 B$ with $B = C_1 Q C_2^{-1}$, as was to be shown.

In the special case $s = n$, we have $t(x) = I$, so that there are no nontrivial invariants. This corresponds to the fact that in this case G is transitive, since any two nonsingular $n \times n$ matrices x_1 and x_2 satisfy $x_2 = x_1 B$ with $B = x_1^{-1} x_2$. This result can be made more intuitive through a geometric interpretation. Consider the s-dimensional subspace S of R^n spanned by the s columns of x. Then $P = x(x^\top x)^{-1} x^\top$ has the property that for any y in R^n, the vector Py is the projection of y onto S. (This will be proved in Section 7.2.) The invariance of P expresses the fact that the projection of y onto S is independent of the choice of vectors spanning S. To see that it is maximal invariant, suppose that the projection of every y onto the spaces S_1 and S_2 spanned by two different sets of s vectors is the same. Then $S_1 = S_2$, so that the two sets of vectors span the same space. There then exists a nonsingular transformation taking one of these sets into the other. ■

A somewhat more systematic way of determining maximal invariants is obtained by selecting, by means of a specified rule, a unique point $M(x)$ on each orbit. Then clearly $M(X)$ is maximal invariant. To illustrate this method, consider once more two of the earlier examples.

Example 6.2.1 (i) (continued). The orbit containing the point $(a_1, \dots, a_n)$ under the group of translations is the set $\{(a_1 + c, \dots, a_n + c), -\infty < c < \infty\}$, which is a line in E_n.

(a) As representative point $M(x)$ on this line, take its intersection with the hyperplane $x_n = 0$. Since then $a_n + c = 0$, this point corresponds to the value $c = -a_n$ and thus has coordinates $(a_1 - a_n, \dots, a_{n-1} - a_n, 0)$. This leads to the maximal invariant $(x_1 - x_n, \dots, x_{n-1} - x_n)$.
(b) An alternative point on the line is its intersection with the hyperplane $\sum x_i = 0$. Then $c = -\bar{a}$, and $M(a) = (a_1 - \bar{a}, \dots, a_n - \bar{a})$.
(c) The point need not be specified by an intersection property. It can, for instance, be taken as the point on the line that is closest to the origin. Since the value of c minimizing $\sum (a_i + c)^2$ is $c = -\bar{a}$, this leads to the same point as (b). ■

Example 6.2.1(iii) (continued). The orbit containing the point $(a_1, \dots, a_n)$ under the group of orthogonal transformations is the hypersphere containing $(a_1, \dots, a_n)$ and with center at the origin. As representative point on this sphere, take its north pole, i.e., the point with $a_1 = \cdots = a_{n-1} = 0$. The coordinates of this point are $(0, \dots, 0, \sqrt{\sum a_i^2})$ and hence lead to the maximal invariant $\sum x_i^2$. (Note that in this example, the determination of the orbit is essentially equivalent to the determination of the maximal invariant.) ■

Frequently, it is convenient to obtain a maximal invariant in a number of steps, each corresponding to a subgroup of G. To illustrate the process and a difficulty that may arise in its application, let $x = (x_1, \dots, x_n)$, suppose that the coordinates are distinct, and consider the group of transformations

$$gx = (ax_1 + b, \dots, ax_n + b), \qquad a \neq 0, \qquad -\infty < b < \infty.$$

Applying first the subgroup of translations $x'_i = x_i + b$, a maximal invariant is $y = (y_1, \dots, y_{n-1})$ with $y_i = x_i - x_n$. Another subgroup consists of the scale changes $x''_i = ax_i$. This induces a corresponding change of scale in the y's: $y''_i = ay_i$, and a maximal invariant with respect to this group acting on the y-space is $z = (z_1, \dots, z_{n-2})$ with $z_i = y_i/y_{n-1}$. Expressing this in terms of the x's, we get $z_i = (x_i - x_n)/(x_{n-1} - x_n)$, which is maximal invariant with respect to G.

Suppose now the process is carried out in the reverse order. Application first of the subgroup $x''_i = ax_i$ yields as maximal invariant $u = (u_1, \dots, u_{n-1})$ with $u_i = x_i/x_n$. However, the translations $x'_i = x_i + b$ do not induce transformations in u-space, since $(x_i + b)/(x_n + b)$ is not a function of x_i/x_n.

Quite generally, let a transformation group G be *generated* by two subgroups D and E in the sense that it is the smallest group containing D and E. Then G consists of the totality of products $e_m d_m \dots e_1 d_1$ for $m = 1, 2, \dots,$ with $d_i \in D$, $e_i \in E$ $(i = 1, \dots, m)$.[2] The following theorem shows that whenever the process of determining a maximal invariant in steps can be carried out at all, it leads to a maximal invariant with respect to G.

Theorem 6.2.2 *Let G be a group of transformations, and let D and E be two subgroups generating G. Suppose that $y = s(x)$ is maximal invariant with respect to D, and that for any $e \in E$*

$$s(x_1) = s(x_2) \quad \textit{implies} \quad s(ex_1) = s(ex_2). \tag{6.9}$$

If $z = t(y)$ is maximal invariant under the group E^* of transformations e^* defined by

$$e^* y = s(ex) \quad \textit{when} \quad y = s(x),$$

then $z = t[s(x)]$ is maximal invariant with respect to G.

Proof. To show that $t[s(x)]$ is invariant, let $x' = gx$, $g = e_m d_m \cdots e_1 d_1$. Then

$$\begin{aligned} t[s(x')] = t[s(e_m d_m \cdots e_1 d_1 x)] &= t[e_m^* s(d_m \cdots e_1 d_1 x)] \\ &= t[s(e_{m-1} d_{m-1} \cdots e_1 d_1 x)], \end{aligned}$$

and the last expression can be reduced by induction to $t[s(x)]$. To see that $t[s(x)]$ is in fact maximal invariant, suppose that $t[s(x')] = t[s(x)]$. Setting $y' = s(x')$, $y = s(x)$,

[2] See Section A.1 of the Appendix.

one has $t(y') = t(y)$, and since $t(y)$ is maximal invariant with respect to E^*, there exists e^* such that $y' = e^* y$. Then $s(x') = e^* s(x) = s(ex)$, and by the maximal invariance of $s(x)$ with respect to D there exists $d \in D$ such that $x' = dex$. Since de is an element of G this completes the proof. ■

Techniques for obtaining the distribution of maximal invariants are discussed by Andersson (1982), Eaton (1983, 1989), Farrell (1985b), Wijsman (1990) and Anderson (2003).

6.3 Uniformly Most Powerful Invariant Tests

In the presence of symmetries, one may wish to restrict attention to invariant tests, and it then becomes of interest to determine the uniformly most powerful invariant (UMPI) test. The following is a simple example.

Example 6.3.1 Let $X_1, \ldots, X_n$ be i.i.d. on (0, 1) and consider testing the hypothesis H_0 that the common distribution of the X's is uniform on (0, 1) against the two alternatives H_1:

$$p_1(x_1, \ldots, x_n) = f(x_1) \cdots f(x_n)$$

and

$$p_2(x_1, \ldots, x_n) = f(1 - x_1) \cdots f(1 - x_n) ,$$

where f is a fixed (known) density.
(i) This problem remains invariant under the 2 element group G consisting of the transformations

$$g: \; x_i' = 1 - x_i \; , \quad i = 1, \ldots, n$$

and the identity transformation $x_i' = x_i$ for $i = 1, \ldots, n$.
(ii) The induced transformation $\bar{g}$ in the space of alternatives takes p_1 into p_2 and p_2 into p_1.
(iii) A test $\phi(x_1, \ldots, x_n)$ remains invariant under G if and only if

$$\phi(x_1, \ldots, x_n) = \phi(1 - x_1, \ldots, 1 - x_n) .$$

(iv) There exists a UMP invariant test (i.e., an invariant test which is simultaneously most powerful against both p_1 and p_2), and it rejects H_0 when the average

$$\bar{p}(x_1, \ldots, x_n) = \frac{1}{2} \left[p_1(x_1, \ldots, x_n) + p_2(x_1, \ldots, x_n) \right]$$

is sufficiently large.

We leave the proof of (i)–(iii) to Problem 6.5. To prove (iv), note that any invariant test satisfies

$$E_{p_1}[\phi(X_1, \dots, X_n)] = E_{p_2}[\phi(X_1, \dots, X_n)] = E_{\bar{p}}[\phi(X_1, \dots, X_n)] \, .$$

Therefore, maximizing the power against p_1 or p_2 is equivalent to maximizing the power under $\bar{p}$, and the result follows from the Neyman–Pearson Lemma. ■

This example is a special case of the following result.

Theorem 6.3.1 *Suppose the problem of testing Ω_0 against Ω_1 remains invariant under a finite group $G = \{g_1, \dots, g_N\}$ and that $\bar{G}$ is transitive over Ω_0 and over Ω_1. Then there exists a UMP invariant test of Ω_0 against Ω_1, and it rejects Ω_0 when*

$$\frac{\sum_{i=1}^N p_{\bar{g}_i \theta_1}(x)/N}{\sum_{i=1}^N p_{\bar{g}_i \theta_0}(x)/N} \tag{6.10}$$

is sufficiently large, where θ_0 and θ_1 are any elements of Ω_0 and Ω_1, respectively.

The proof is exactly analogous to that of the preceding example; see Problem 6.6.

The results of the previous section provide an alternative approach to the determination of most powerful invariant tests. By Theorem 6.2.1, the class of all invariant functions can be obtained as the totality of functions of a maximal invariant $M(x)$. Therefore, in particular the class of all invariant tests is the totality of tests depending only on the maximal invariant statistic M. The latter statement, while correct for all the usual situations, actually requires certain qualifications regarding the class of measurable sets in M-space. These conditions will be discussed at the end of the section; they are satisfied in the examples below.

Example 6.3.2 Let $X = (X_1, \dots, X_n)$, and suppose that the density of X is $f_i(x_1 - \theta, \dots, x_n - \theta)$ under H_i $(i = 0, 1)$, where θ ranges from $-\infty$ to ∞. The problem of testing H_0 against H_1 is invariant under the group G of transformations

$$gx = (x_1 + c, \dots, x_n + c), \qquad -\infty < c < \infty$$

which in the parameter space induces the transformations

$$\bar{g}\theta = \theta + c.$$

By Example 6.2.1, a maximal invariant under G is $Y = (X_1 - X_n, \dots, X_{n-1} - X_n)$. The distribution of Y is independent of θ and under H_i has the density

$$\int_{-\infty}^{\infty} f_i(y_1 + z, \dots, y_{n-1} + z, z)\, dz.$$

When referred to Y, the problem of testing H_0 against H_1 therefore becomes one of testing a simple hypothesis against a simple alternative. The most powerful test is then independent of θ, and therefore UMP among all invariant tests. Its rejection region by the Neyman–Pearson Lemma is

$$\frac{\int_{-\infty}^{\infty} f_1(y_1 + z, \ldots, y_{n-1} + z, z)\, dz}{\int_{-\infty}^{\infty} f_0(y_1 + z, \ldots, y_{n-1} + z, z)\, dz} = \frac{\int_{-\infty}^{\infty} f_1(x_1 + u, \ldots, x_n + u)\, du}{\int_{-\infty}^{\infty} f_0(x_1 + u, \ldots, x_n + u)\, du} > C. \tag{6.11}$$

A general theory of *separate families of hypotheses* (in which the family K of alternatives does not adjoin the hypothesis H but, as above, is separated from it) was initiated by Cox (1961, 1962). A bibliography of the subject is given in Pereira (1977); see also Loh (1985), Pace and Salvan (1990) and Rukhin (1993). ■

Example 6.3.2 illustrates the fact, also utilized in Theorem 6.3.1, that if the group $\bar{G}$ is transitive over both Ω_0 and Ω_1, then the problem reduces to one of testing a simple hypothesis against a simple alternative, and a UMP invariant test is then obtained by the Neyman–Pearson Lemma. Note also the close similarity between Theorem 6.3.1 and Example 6.3.2 shown by a comparison of (6.10) and the right side of (6.11), where the summation in (6.10) is replaced by integration with respect to Lebesgue measure.

In $\bar{G}$ is not transitive, the existence of a UMPI test is not guaranteed. The problem then is to determine whether or not there exists a UMP test based on a maximal invariant T. If the family of distributions of T has monotone likelihood ratio, then a UMPI test may exist, as illustrated in the following example.

Example 6.3.3 Testing many normal means) Assume $X_1, \ldots, X_n$ are independent with $X_i \sim N(\xi_i, \sigma^2)$, where we assume σ is known and equal to 1. The parameter space Ω is n-dimensional Euclidean space as $\xi_1, \ldots, \xi_n$ vary freely. The problem is to test

$$H_0 : \xi_1 = \xi_2 = \cdots = \xi_n = 0$$

against the alternative where not all ξ_i are 0. Note there does not exist even a UMPU test unless $n = 1$. Let $X = (X_1, \ldots, X_n)^\top$ (where the superscript T denotes transpose), and consider an orthogonal matrix O. If $X' = OX$, then X' consists of independent normals, each with variance one and possibly different means. Moreover, X' has mean 0 iff X has mean 0. (Here the prime in X' just denotes that it is a transformation of X, so not $X^\top$.) So, the problem of testing the mean vector is 0 based on X' is identical to the problem based on X. For this reason, invariance or symmetry requires restricting to tests ϕ satisfying

$$\phi(X) = \phi(OX) \quad \text{for all orthogonal matrices } O\,. \tag{6.12}$$

If G is the group of orthogonal transformations, then from Example 6.2.1(iii), $T = \sum_{i=1}^n X_i^2$ is a maximal invariant. The distribution of T is noncentral Chi-squared with n degrees of freedom and noncentrality parameter $\psi^2 = \sum_{i=1}^n \xi_i^2$. Its density function is

$$p_{\psi^2}(t) = \exp(-\psi^2/2) \sum_{k=0}^{\infty} \frac{(\psi^2/2)^k}{k!} \cdot \frac{t^{\frac{n}{2}-1+k} \exp(-t/2)}{2^{2k+n}\Gamma(k+\frac{n}{2})} \ .$$

The central Chi-squared just corresponds to the $k = 0$ term. Then,

$$\frac{p_{\psi^2}(t)}{p_0(t)} = \exp(-\psi^2/2) \sum_{k=0}^{\infty} c_k (\psi^2/2)^k t^k \ ,$$

for constants c_k. Since each term is increasing in t, the whole ratio is increasing in t. Therefore, the family of distributions of T has monotone likelihood ratio. Hence, the UMPI test rejects when $T > c_n(1-\alpha)$, where $c_n(1-\alpha)$ is the $1-\alpha$ quantile of the Chi-squared distribution with n degrees of freedom. ■

Before applying invariance, it is frequently convenient first to reduce the data to a sufficient statistic T. If there exists a test $\phi_0(T)$ that is UMP among all invariant tests depending only on T, one would like to be able to conclude that $\phi_0(T)$ is also UMP among all invariant tests based on the original X. Unfortunately, this does not follow, since it is not clear that for any invariant test based on X there exists an equivalent test based on T, which is also invariant. Sufficient conditions for $\phi_0(T)$ to have this property are provided by Hall et al. (1965) and Hooper (1982a), and a simple version of such a result (applicable to Examples 6.3.4 and 6.3.5 below) will be given by Theorem 6.5.3 in Section 6.5. For a review and clarification of this and later work on invariance and sufficiency see Berk et al. (1996), Nogales and Oyola (1996) and Nogales et al. (2000).

Example 6.3.4 If $X_1, \ldots, X_n$ is a sample from $N(\xi, \sigma^2)$, the hypothesis $H : \sigma \geq \sigma_0$ remains invariant under the transformations $X'_i = X_i + c$, $-\infty < c < \infty$. In terms of the sufficient statistics $Y = \bar{X}$, $S^2 = \Sigma(X_i - \bar{X})^2$ these transformations become $Y' = Y + c$, $(S^2)' = S^2$, and a maximal invariant is S^2. The class of invariant tests is therefore the class of tests depending on S^2. It follows from Theorem 3.4.1 that there exists a UMP invariant test, with rejection region $\Sigma(X_i - \bar{X})^2 \leq C$. This coincides with the UMP unbiased test (5.9). ■

Example 6.3.5 If $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ are samples from $N(\xi, \sigma^2)$ and $N(\eta, \tau^2)$, a set of sufficient statistics is $T_1 = \bar{X}$, $T_2 = \bar{Y}$, $T_3 = \sqrt{\Sigma(X_i - \bar{X})^2}$, and $T_4 = \sqrt{\Sigma(Y_j - \bar{Y})^2}$. The problem of testing $H : \tau^2/\sigma^2 \leq \Delta_0$ remains invariant under the transformations $T'_1 = T_1 + c_1, T'_2 = T_2 + c_2, T'_3 = T_3, T'_4 = T_4, -\infty < c_1$, $c_2 < \infty$, and also under a common change of scale of all four variables. A maximal invariant with respect to the first group is (T_3, T_4). In the space of this maximal invariant, the group of scale changes induces the transformations $T''_3 = cT_3$, $T''_4 = cT_4$, $0 < c$, which has as maximal invariant the ratio T_4/T_3. The statistic $Z = [T_4^2/(n-1)] \div [T_3^2/(m-1)]$ on division by $\Delta = \tau^2/\sigma^2$ has an F-distribution with density given by (5.21), so that the density of Z is

$$\frac{c(\Delta)z^{\frac{1}{2}(n-3)}}{\left(\Delta + \frac{n-1}{m-1}z\right)^{\frac{1}{2}(m+n-2)}}, \qquad z > 0.$$

For varying Δ, these densities constitute a family with monotone likelihood ratio, so that among all tests of H based on Z, and therefore among all invariant tests, there exists a UMP one given by the rejection region $Z > C$. This coincides with the UMP unbiased test (5.20). ■

Example 6.3.6 In the method of *paired comparisons* for testing whether a treatment has a beneficial effect, the experimental material consists of n pairs of subjects. From each pair, a subject is selected at random for treatment while the other serves as control. Let X_i be 1 or 0 as for the ith pair the experiment turns out in favor of the treated subject or the control, and let $p_i = P\{X_i = 1\}$. The hypothesis of no effect, $H : p_i = \frac{1}{2}$ for $i = 1, \dots, n$, is to be tested against the alternatives that $p_i > \frac{1}{2}$ for all i.

The problem remains invariant under all permutations of the n variables $X_1, \dots, X_n$, and a maximal invariant under this group is the total number of successes $X = X_1 + \cdots + X_n$. The distribution of X is

$$P\{X = k\} = q_1 \cdots q_n \sum \frac{p_{i_1}}{q_{i_1}} \cdots \frac{p_{i_k}}{q_{i_k}},$$

where $q_i = 1 - p_i$ and where the summation extends over all $\binom{n}{k}$ choices of subscripts $i_1 < \cdots < i_k$. The most powerful invariant test against an alternative $(p'_1, \dots, p'_n)$ rejects H when

$$f(k) = \frac{1}{\binom{n}{k}} \sum \frac{p'_{i_1}}{q'_{i_1}} \cdots \frac{p'_{i_k}}{q'_{i_k}} > C.$$

To see that f is an increasing function of k, note that $a_i = p'_i/q'_i > 1$, and that

$$\sum_j \sum a_j a_{i_1} \cdots a_{i_k} = (k+1) \sum a_{i_1} \cdots a_{i_{k+1}}$$

and

$$\sum_j \sum a_{i_1} \cdots a_{i_k} = (n-k) \sum a_{i_1} \cdots a_{i_{k_1}}.$$

Here, in both equations, the second summation on the left-hand side extends over all subscripts $i_1 < \cdots < i_k$ of which none is equal to j, and the summation on the right-hand side extends over all subscripts $i_1 < \cdots < i_{k+1}$ and $i_1 < \cdots < i_k$ respectively without restriction. Then

$$f(k+1) = \frac{1}{\binom{n}{k+1}} \sum a_{i_1} \cdots a_{i_{k+1}} = \frac{1}{(n-k)\binom{n}{k}} \sum_j \sum a_j a_{i_1} \cdots a_{i_k}$$
$$> \frac{1}{\binom{n}{k}} \sum a_{i_1} \cdots a_{i_k} = f(k),$$

as was to be shown. Regardless of the alternative chosen, the test therefore rejects when $X > C$, and hence is UMP invariant. If the ith comparison is considered plus or minus as X_i is 1 or 0, this is seen to be another example of the sign test. (Cf. Example 3.8.1 and Section 4.9.) ■

Sufficient statistics provide a simplification of a problem by reducing the sample space; this process involves no change in the parameter space. Invariance, on the other hand, by reducing the data to a maximal invariant statistic M, whose distribution may depend only on a function of the parameter, typically also shrinks the parameter space. The details are given in the following theorem.

Theorem 6.3.2 *If $M(x)$ is invariant under G, and if $v(\theta)$ maximal invariant under the induced group $\bar{G}$, then the distribution of $M(X)$ depends only on $v(\theta)$.*

PROOF. Let $v(\theta_1) = v(\theta_2)$. Then $\theta_2 = \bar{g}\theta_1$, and hence

$$P_{\theta_2}\{M(X) \in B\} = P_{\bar{g}\theta_1}\{M(X) \in B\} = P_{\theta_1}\{M(gX) \in B\}$$
$$= P_{\theta_1}\{M(X) \in B\}.$$

This result can be paraphrased by saying that the principle of invariance identifies all parameter points that are equivalent with respect to $\bar{G}$. ■

In applications, for instance, in Examples 6.3.4 and 6.3.5, the maximal invariants $M(x)$ and $\delta = v(\theta)$ under G and $\bar{G}$ are frequently real-valued, and the family of probability densities $p_\delta(m)$ of M has monotone likelihood ratio. For testing the hypothesis $H : \delta \le \delta_0$ there exists then a UMP test among those depending only on M, and hence a UMP invariant test. Its rejection region is $M \ge C$, where

$$\int_C^\infty P_{\delta_0}(m)\, dm = \alpha. \tag{6.13}$$

Consider this problem now as a two-decision problem with decisions d_0 and d_1 of accepting or rejecting H, and a loss function $L(\theta, d_i) = L_i(\theta)$. Suppose that $L_i(\theta)$ depends only on the parameter δ, $L_i(\theta) = L'_i(\delta)$ say, and satisfies

$$L'_1(\delta) - L'_0(\delta) \gtrless 0 \quad \text{as} \quad \delta \lessgtr \delta_0. \tag{6.14}$$

It then follows from Theorem 3.4.2 that the family of rejection regions $M \ge C(\alpha)$, as α varies from 0 to 1, forms a complete family of decision procedures among those depending only on M, and hence a complete family of invariant procedures.

As before, the choice of a particular significance level α can be considered as a convenient way of specifying a test from this family.

At the beginning of the section it was stated that the class of invariant tests coincides with the class of tests based on a maximal invariant statistic $M = M(X)$. However, a statistic is not completely specified by a function, but requires also specification of a class $\mathcal{B}$ of measurable sets. If in the present case $\mathcal{B}$ is the class of all sets B for which $M^{-1}(B) \in \mathcal{A}$, the desired statement is correct. For let $\phi(x) = \psi[M(x)]$ and ϕ by $\mathcal{A}$-measurable, and let C be a Borel set on the line. Then $\phi^{-1}(C) = M^{-1}[\psi^{-1}(C)] \in \mathcal{A}$ and hence $\psi^{-1}(C) \in \mathcal{B}$, so that ψ is $\mathcal{B}$-measurable and $\phi(x) = \psi[M(x)]$ is a test based on the statistic M.

In most applications, $M(x)$ is a measurable function taking on values in a Euclidean space and it is convenient to take $\mathcal{B}$ as the class of Borel sets. If $\phi(x) = \psi[M(x)]$ is then an arbitrary measurable function depending only on $M(x)$, it is not clear that $\psi(m)$ is necessarily $\mathcal{B}$-measurable. This measurability can be concluded if $\mathcal{X}$ is also Euclidean with $\mathcal{A}$ the class of Borel sets, and if the range of M is a Borel set. We shall prove it here only under the additional assumption (which in applications is usually obvious, and which will not be verified explicitly in each case) that there exists a vector-valued Borel-measurable function $Y(x)$ such that $[M(x), Y(x)]$ maps $\mathcal{X}$ onto a Borel subset of the product space $\mathcal{M} \times \mathcal{Y}$, that this mapping is $1:1$, and that the inverse mapping is also Borel-measurable. Given any measurable function ϕ of x, there exists then a measurable function ϕ' of (m, y) such that $\phi(x) \equiv \phi'[M(x), Y(x)]$. If ϕ depends only on $M(x)$, then ϕ' depends only on m, so that $\phi'(m, y) = \psi(m)$ say, and ψ is a measurable function of m.[3] In Example 6.2.1(i) for instance, where $x = (x_1, \dots x_n)$ and $M(x) = (x_1 - x_n, \dots, x_{n-1} - x_n)$, the function $Y(x)$ can be taken as $Y(x) = x_n$.

6.4 Sample Inspection by Variables

A sample is drawn from a lot of some manufactured product in order to decide whether the lot is of acceptable quality. In the simplest case, each sample item is classified directly as satisfactory or defective (*inspection by attributes*), and the decision is based on the total number of defectives. More generally, the quality of an item is characterized by a variable Y (*inspection by variables*), and an item is considered satisfactory if Y exceeds a given constant u. The probability of a defective is then

$$p = P\{Y \le u\}$$

and the problem becomes that of testing the hypothesis $H : p \ge p_0$.

As was seen in Example 3.8.1, no use can be made of the actual value of Y unless something is known concerning the distribution of Y. In the absence of such information, the decision will be based, as before, simply on the number of defectives

[3] The last statement follows, for example, from Theorem 18.1 of Billingsley (1995).

in the sample. We shall consider the problem now under the assumption that the measurements $Y_1, \dots, Y_n$ constitute a sample from $N(\eta, \sigma^2)$. Then

$$p = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}(y-\eta)^2\right] dy = \Phi\left(\frac{u-\eta}{\sigma}\right),$$

where

$$\Phi(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}t^2\right) dt$$

denotes the cumulative distribution function of a standard normal distribution, and the hypothesis H becomes $(u-\eta)/\sigma \geq \Phi^{-1}(p_0)$. In terms of the variables $X_i = Y_i - u$, which have mean $\xi = \eta - u$ and variance σ^2, this reduces to

$$H : \frac{\xi}{\sigma} \leq \theta_0$$

with $\theta_0 = -\Phi^{-1}(p_0)$. This hypothesis, which was considered in Section 5.2, for $\theta_0 = 0$, occurs also in other contexts. It is appropriate when one is interested in the mean ξ of a normal distribution, expressed in σ units rather than on a fixed scale.

For testing H, attention can be restricted to the pair of variables $\bar{X}$ and $S = \sqrt{\sum(X_i - \bar{X})^2}$, since they form a set of sufficient statistics for (ξ, σ), which satisfy the conditions of Theorem 6.5.3 of the next section. These variables are independent, the distribution of $\bar{X}$ being $N(\xi, \sigma^2/n)$ and that of S/σ being χ_{n-1}. Multiplication of $\bar{X}$ and S by a common constant $c > 0$ transforms the parameters into $\xi' = c\xi$, $\sigma' = c\sigma$, so that ξ/σ and hence the problem of testing H remain invariant. A maximal invariant under these transformations is $\bar{x}/s$ or

$$t = \frac{\sqrt{n}\bar{x}}{s/\sqrt{n-1}},$$

the distribution of which depends only on the maximal invariant in the parameter space $\theta = \xi/\sigma$ (cf. Section 5.2). Thus, the invariant tests are those depending only on t, and it remains to find the most powerful test of $H : \theta \leq \theta_0$ within this class.

The probability density of t is (Problem 5.3)

$$p_\delta(t) = C \int_0^\infty \exp\left[-\frac{1}{2}\left(t\sqrt{\frac{w}{n-1}} - \delta\right)^2\right] w^{\frac{1}{2}(n-2)} \exp\left(-\tfrac{1}{2}w\right) dw,$$

where $\delta = \sqrt{n}\theta/\sigma$ is the noncentrality parameter, and this will now be shown to constitute a family with monotone likelihood ratio. To see that the ratio

$$r(t) = \frac{\int_0^\infty \exp\left[-\frac{1}{2}\left(t\sqrt{\frac{w}{n-1}} - \delta_1\right)^2\right] w^{\frac{1}{2}(n-2)} \exp(-\frac{1}{2}w)\,dw}{\int_0^\infty \exp\left[-\frac{1}{2}\left(t\sqrt{\frac{w}{n-1}} - \delta_0\right)^2\right] w^{\frac{1}{2}(n-2)} \exp(-\frac{1}{2}w)\,dw}$$

is an increasing function of t for $\delta_0 < \delta_1$, suppose first that $t < 0$ and let $v = -t\sqrt{w/(n-1)}$. The ratio then becomes proportional to

$$\frac{\int_0^\infty f(v) \exp\left[-(\delta_1 - \delta_0)v - \frac{(n-1)v^2}{2t^2}\right] dv}{\int_0^\infty f(v) \exp\left[-\frac{(n-1)v^2}{2t^2}\right] dv}$$
$$= \int \exp[-(\delta_1 - \delta_0)v] g_{t^2}(v)\,dv,$$

where

$$f(v) = \exp(-\delta_0 v) v^{n-1} \exp(-v^2/2)$$

and

$$g_{t^2}(v) = \frac{f(v) \exp\left[-\frac{(n-1)v^2}{2t^2}\right]}{\int_0^\infty f(z) \exp\left[-\frac{(n-1)z^2}{2t^2}\right] dz}.$$

Since the family of probability densities $g_{t^2}(v)$ is a family with monotone likelihood ratio, the integral of $\exp[-(\delta_1 - \delta_0)v]$ with respect to this density is a decreasing function of t^2 (Problem 3.41), and hence an increasing function of t for $t < 0$. Similarly one finds that $r(t)$ is an increasing function of t for $t > 0$ by making the transformation $v = t\sqrt{w/(n-1)}$. By continuity it is then an increasing function of t for all t.

There exists therefore a UMP invariant test of $H : \xi/\sigma \leq \theta_0$, which rejects when $t > C$, where C is determined by (6.13). In terms of the original variables Y_i the rejection region of the UMP invariant test of $H : p \geq p_0$ becomes

$$\frac{\sqrt{n}(\bar{y} - u)}{\sqrt{\sum(y_i - \bar{y})^2/(n-1)}} > C. \tag{6.15}$$

If the problem is considered as a two-decision problem with losses $L_0(p)$ and $L_1(p)$ for accepting or rejecting $p \geq p_0$, which depend only on p and satisfy the condition corresponding to (6.14), the class of tests (6.15) constitutes a complete family of invariant procedures as C varies from $-\infty$ to ∞.

Consider next the comparison of two probabilities on the basis of samples $X_1, \ldots, X_m; Y_1, \ldots, Y_n$ from $N(\xi, \sigma^2)$ and $N(\eta, \sigma^2)$. If

$$p = \Phi\left(\frac{u - \xi}{\sigma}\right), \qquad \pi = \Phi\left(\frac{u - \eta}{\sigma}\right),$$

one wishes to test the hypothesis $p \leq \pi$, which is equivalent to

$$H : \eta \leq \xi.$$

The statistics $\bar{X}, \bar{Y}$, and $S = \sqrt{\sum(X_i - \bar{X})^2 + \sum(Y_j - \bar{Y})^2}$ are a set of sufficient statistics for ξ, η, σ. The problem remains invariant under the addition of an arbitrary common constant to $\bar{X}$ and $\bar{Y}$, which leaves $\bar{Y} - \bar{X}$ and S as maximal invariants. It is also invariant under multiplication of $\bar{X}, \bar{Y}$, and S, and hence of $\bar{Y} - \bar{X}$ and S, by a common positive constant, which reduces the data to the maximal invariant $(\bar{Y} - \bar{X})/S$. Since

$$t = \frac{(\bar{y} - \bar{x})/\sqrt{\frac{1}{m} + \frac{1}{n}}}{s/\sqrt{m + n - 2}}$$

has a noncentral t-distribution with noncentrality parameter $\delta = \sqrt{mn}(\eta - \xi)/\sqrt{(m + n)}\sigma$, the UMP invariant test of $H : \eta - \xi \leq 0$ rejects when $t > C$. This coincides with the UMP unbiased test (5.27). Analogously, the corresponding two-sided test (5.30), with rejection region $|t| \geq C$, is UMP invariant for testing the hypothesis $p = \pi$ against the alternatives $p \neq \pi$ (Problem 6.19).

6.5 Almost Invariance

Let G be a group of transformations leaving a family $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ of distributions of X invariant. A test ϕ is said to be *equivalent to an invariant test* if there exists an invariant test ψ such that $\phi(x) = \psi(x)$ for all x except possibly on a $\mathcal{P}$-null set N; ϕ is said to be *almost invariant with respect to G if*

$$\phi(gx) = \phi(x) \qquad \text{for all} \quad x \in \mathcal{X} - N_g, \quad g \in G, \tag{6.16}$$

where the exceptional null set N_g is permitted to depend on g. This concept is required for investigating the relationship of invariance to unbiasedness and to certain other desirable properties. In this connection it is important to know whether a UMP invariant test is also UMP among almost invariant tests. This turns out to be the case under assumptions which are made precise in Theorem 6.5.1 below and which are satisfied in all the usual applications.

If ϕ is equivalent to an invariant test, then $\phi(gx) = \phi(x)$ for all $x \notin N \cup g^{-1}N$. Since $P_\theta(g^{-1}N) = P_{\bar{g}\theta}(N) = 0$, it follows that ϕ is then almost invariant. The following theorem gives conditions under which conversely any almost invariant test is equivalent to an invariant one.

Theorem 6.5.1 *Let G be a group of transformations of $\mathcal{X}$, and let $\mathcal{A}$ and $\mathcal{B}$ be σ-fields of subsets of $\mathcal{X}$ and G such that for any set $A \in \mathcal{A}$ the set of pairs (x, g) for which $gx \in A$ is measurable $\mathcal{A} \times \mathcal{B}$. Suppose further that there exists a σ-finite* *measure ν over G such that $\nu(B) = 0$ implies $\nu(Bg) = 0$ for all $g \in G$. Then any measurable function that is almost invariant under G (where "almost" refers to some σ-finite measure μ) is equivalent to an invariant function.*

PROOF. Because of the measurability assumptions, the function $\phi(gx)$ considered as a function of the two variables x and g is measurable $\mathcal{A} \times \mathcal{B}$. It follows that $\phi(gx) - \phi(x)$ is measurable $\mathcal{A} \times \mathcal{B}$, and so therefore is the set S of points (x, g) with $\phi(gx) \neq \phi(x)$. If ϕ is almost invariant, any section of S with fixed g is a μ-null set. By Fubini's Theorem (Theorem 2.2.4), there exists therefore a μ-null set N such that for all $x \in \mathcal{X} - N$

$$\phi(gx) = \phi(x) \quad \text{a.e. } \nu.$$

Without loss of generality suppose that $\nu(G) = 1$, and let A be the set of points x for which

$$\int \phi(g'x)\, d\nu(g') = \phi(gx) \quad \text{a.e. } \nu.$$

If

$$f(x, g) = \left| \int \phi(g'x)\, d\nu(g') - \phi(gx) \right|,$$

then A is the set of points x for which

$$\int f(x, g)\, d\nu(g) = 0.$$

Since this integral is a measurable function of x, it follows that A is measurable. Let

$$\psi(x) = \begin{cases} \int \phi(gx) d\nu(g) & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Then ψ is measurable and $\psi(x) = \phi(x)$ for $x \notin N$, since $\phi(gx) = \phi(x)$ a.e. ν implies that $\int \phi(g'x)\, d\nu(g') = \phi(x)$ and that $x \in A$. To show that ψ is invariant it is enough to prove that the set A is invariant. For any point $x \in A$, the function $\phi(gx)$ is constant except on a null subset N_x of G. Then $\phi(ghx)$ has the same constant value for all $g \notin N_x h^{-1}$, which by assumption is again a ν-null set. Hence $hx \in A$, which completes the proof. ■

Additional results concerning the relation of invariance and almost invariance are given by Berk and Bickel (1968) and Berk (1970). In particular, the basic idea of the following example is due to Berk (1970).

Example 6.5.1 (Counterexample) Let $Z, Y_1, \ldots, Y_n$ be independently distributed as $N(\theta, 1)$, and consider the 1 : 1 transformations $y_i' = y_i (i = 1, \ldots, n)$ and

$z' = z$ except for a finite number of points $a_1, \ldots, a_k$ for which $a_i' = a_{j_i}$, for some permutation $(j_1, \ldots, j_k)$ of $(1, \ldots, k)$.

If the group G is generated by taking for $(a_1, \dots, a_k)$, $k = 1, 2, \dots$, all finite sets and for $(j_1, \dots, j_k)$ all permutations of $(1, \dots, k)$, then $(z, y_1, \dots, y_n)$ is almost invariant. It is however not equivalent to an invariant function, since $(y_1, \dots, y_n)$ is maximal invariant. ■

Corollary 6.5.1 *Suppose that the problem of testing $H : \theta \in \omega$ against $K : \theta \in \Omega - \omega$ remains invariant under G and that the assumptions of Theorem 6.5.1 hold. Then if ϕ_0 is UMP invariant, it is also UMP within the class of almost invariant tests.*

Proof. If ϕ is almost invariant, it is equivalent to an invariant test ψ by Theorem 6.5.1. The tests ϕ and ψ have the same power function, and hence ϕ_0 is uniformly at least as powerful as ϕ. ■

In applications, $\mathcal{P}$ is usually a dominated family, and μ any σ-finite measure equivalent to $\mathcal{P}$ (which exists by Theorem A.4.2 of the Appendix). If ϕ is almost invariant with respect to $\mathcal{P}$, it is then almost invariant with respect to μ and hence equivalent to an invariant test. Typically, the sample space $\mathcal{X}$ is n-dimensional Euclidean space, $\mathcal{A}$ is the class of Borel sets, and the elements of G are transformations of the form $y = f(x, \tau)$, where τ ranges over a set of positive measure in an m-dimensional space and f is a Borel-measurable vector-valued function of $m + n$ variables. If $\mathcal{B}$ is taken as the class of Borel sets in m-space the measurability conditions of the theorem are satisfied.

The requirement that for all $g \in G$ and $B \in \mathcal{B}$

$$\nu(B) = 0 \quad \text{implies} \quad \nu(Bg) = 0 \tag{6.17}$$

is satisfied in particular when

$$\nu(Bg) = \nu(B) \qquad \text{for all} \quad g \in G, \quad B \in \mathcal{B}. \tag{6.18}$$

The existence of such a *right invariant measure* is guaranteed for a large class of groups by the theory of Haar measure. (See, for example, Eaton 1989.) Alternatively, it is usually not difficult to check Condition (6.17) directly.

Example 6.5.2 Let G be the group of all nonsingular linear transformations of n-space. Relative to a fixed coordinate system the elements of G can be represented by nonsingular $n \times n$ matrices $A = (a_{ij})$, $A' = (a'_{ij})$, ... with the matrix product serving as the group product of two such elements. The σ-field $\mathcal{B}$ can be taken to be the class of Borel sets in the space of the n^2 elements of the matrices, and the measure ν can be taken as Lebesgue measure over $\mathcal{B}$. Consider now a set S of matrices with $\nu(S) = 0$, and the set S^* of matrices $A'A$ with $A' \in S$ and A fixed. If $a = \max |a_{ij}|$, $C' = A'A$, and $C'' = A''A$, the inequalities $|a''_{ij} - a'_{ij}| \le \epsilon$ for all i, j imply $|c''_{ij} - c'_{ij}| \le na\epsilon$. Since a set has ν-measure zero if and only if it can be covered by a union of rectangles whose total measure does not exceed any given $\epsilon > 0$, it follows that $\nu(S^*) = 0$, as was to be proved. ■

In the preceding chapters, tests were compared purely in terms of their power functions (possibly weighted according to the seriousness of the losses involved). Since the restriction to invariant tests is a departure from this point of view, it is of interest to consider the implications of applying invariance to the power functions rather than to the tests themselves. Any test that is invariant or almost invariant under a group G has a power function which is invariant under the group $\bar{G}$ induced by G in the parameter space.

To see that the converse is in general not true, let X_1, X_2, X_3 be independently, normally distributed with mean ξ and variance σ^2, and consider the hypothesis $\sigma \geq \sigma_0$. The test with rejection region

$$\begin{aligned} |X_2 - X_1| > k &\qquad \text{when} \quad \bar{X} < 0, \\ |X_3 - X_2| > k &\qquad \text{when} \quad \bar{X} \geq 0 \end{aligned}$$

is not invariant under the group G of transformations $X_i' = X_i + c$, but its power function is invariant under the associated group $\bar{G}$.

The two properties, almost invariance of a test ϕ and invariance of its power function, become equivalent if before the application of invariance considerations the problem is reduced to a sufficient statistic whose distributions constitute a boundedly complete family.

Lemma 6.5.1 *Let the family $\mathcal{P}^T = \{P_\theta^T, \theta \in \Omega\}$ of distributions of T be boundedly complete, and let the problem of testing $H : \theta \in \Omega_H$ remain invariant under a group G of transformations of T. Then a necessary and sufficient condition for the power function of a test $\psi(t)$ to be invariant under the induced group $\bar{G}$ over Ω is that $\psi(t)$ is almost invariant under G.*

PROOF. For all $\theta \in \Omega$ we have $E_{\bar{g}\theta}\psi(T) = E_\theta\psi(gT)$. If ψ is almost invariant, $E_\theta\psi(T) = E_\theta\psi(gT)$ and hence $E_{\bar{g}\theta}\psi(T) = E_\theta\psi(T)$, so that the power function of ψ is invariant. Conversely, if $E_\theta\psi(T) = E_{\bar{g}\theta}\psi(T)$, then $E_\theta\psi(T) = E_\theta\psi(gT)$, and by the bounded completeness of $\mathcal{P}^T$, we have $\psi(gt) = \psi(t)$ a.e. $\mathcal{P}^T$. ■

As a consequence, it is seen that UMP almost invariant tests also possess the following optimum property.

Theorem 6.5.2 *Under the assumptions of Lemma 6.5.1, let $v(\theta)$ be maximal invariant with respect to $\bar{G}$, and suppose that among the tests of H based on the sufficient statistic T there exists a UMP almost invariant one, say $\psi_0(t)$. Then $\psi_0(t)$ is UMP in the class of all tests based on the original observations X, whose power function depends only on $v(\theta)$.*

PROOF. Let $\phi(x)$ be any such test, and let $\psi(t) = E[\phi(X)|t]$. The power function of $\psi(t)$, being identical with that of $\phi(x)$, depends then only on $v(\theta)$, and hence is invariant under $\bar{G}$. It follows from Lemma 6.5.1 that $\psi(t)$ is almost invariant under G, and $\psi_0(t)$ is uniformly at least as powerful as $\psi(t)$ and therefore as $\phi(x)$. ■

Example 6.5.3 For the hypothesis $\tau^2 \le \sigma^2$ concerning the variances of two normal distributions, the statistics $(\bar{X}, \bar{Y}, S_x^2, S_Y^2)$ constitute a complete set of sufficient statistics. It was shown in Example 6.3.5 that there exists a UMP invariant test with respect to a suitable group G, which has rejection region $S_Y^2/S_X^2 > C_0$. Since in the present case almost invariance of a test with respect to G implies that it is equivalent to an invariant one (Problem 6.23), Theorem 6.5.2 is applicable with $v(\theta) = \Delta = \tau^2/\sigma^2$, and the test is therefore UMP among all tests whose power function depends only on Δ. ■

Theorem 6.5.1 makes it possible to establish a simple condition under which reduction to sufficiency before the application of invariance is legitimate.

Theorem 6.5.3 *Let X be distributed according to P_θ, $\theta \in \Omega$, and let T be sufficient for θ. Suppose G leaves invariant the problem of testing $H : \theta \in \Omega_H$, and that T satisfies*

$$T(x_1) = T(x_2) \quad \textit{implies} \quad T(gx_1) = T(gx_2) \quad \textit{for all} \quad g \in G,$$

so that G induces a group $\tilde{G}$ of transformations of T-space through

$$\tilde{g}T(x) = T(gx).$$

(i) *If $\varphi(x)$ is any invariant test of H, there exists an almost invariant test ψ based on T, which has the same power function as φ.*

(ii) *If in addition the assumptions of Theorem 6.5.1 are satisfied, the test ψ of (i) can be taken to be invariant.*

(iii) *If there exists a test $\psi_0(T)$ which is UMP among all $\tilde{G}$-invariant tests based on T, then under the assumptions of (ii), ψ_0, is also UMP among all G-invariant tests based on X.*

This theorem justifies the derivation of the UMP invariant tests of Examples 6.3.4 and 6.3.5.

Proof. (i): Let $\psi(t) = E[\varphi(X)|t]$. Then ψ has the same power function as φ. To complete the proof, it suffices to show that $\psi(t)$ is almost invariant, i.e., that

$$\psi(\tilde{g}t) = \psi(t) \qquad (\text{a.e. } \mathcal{P}^{\mathcal{T}}).$$

It follows from (6.2) that

$$E_\theta[\varphi(gX)|\tilde{g}t] = E_{\bar{g}\theta}[\varphi(X)|t] \qquad (\text{a.e. } P_\theta).$$

Since T is sufficient, both sides of this equation are independent of θ. Furthermore $\varphi(gx) = \varphi(x)$ for all x and g, and this completes the proof.

Part (ii) follows immediately from (i) and Theorem 6.5.1, and part (iii) from (ii). ■

6.6 Unbiasedness and Invariance

The principles of unbiasedness and invariance complement each other in that each is successful in cases where the other is not. For example, there exist UMP unbiased tests for the comparison of two binomial or Poisson distributions, problems to which invariance considerations are not applicable. UMP unbiased tests also exist for testing the hypothesis $\sigma = \sigma_0$ against $\sigma \neq \sigma_0$ in a normal distribution, while invariance does not reduce this problem sufficiently far. Conversely, there exist UMP invariant tests of hypotheses specifying the values of more than one parameter (to be considered in Chapter 7) but for which the class of unbiased tests has no UMP member. There are also hypotheses, for example, the one-sided hypothesis $\xi/\sigma \leq \theta_0$ in a univariate normal distribution or $\rho \leq \rho_0$ in a bivariate one (Problem 6.20) with $\theta_0, \rho_0 \neq 0$, where a UMP invariant test exists but the existence of a UMP unbiased test does not follow by the methods of Chapter 5 and is an open question.

On the other hand, to some problems both principles have been applied successfully. These include Student's hypotheses $\xi \leq \xi_0$ and $\xi = \xi_0$ concerning the mean of a normal distribution, and the corresponding two-sample problems $\eta - \xi \leq \Delta_0$ and $\eta - \xi = \Delta_0$ when the variances of the two samples are assumed equal. Other examples are the one-sided hypotheses $\sigma^2 \geq \sigma_0^2$ and $\tau^2/\sigma^2 \geq \Delta_0$ concerning the variances of one or two normal distributions. The hypothesis of independence $\rho = 0$ in a bivariate normal distribution is still another case in point (Problem 6.20). In all these examples the two optimum procedures coincide. We shall now show that this is not accidental but is the case whenever the UMP invariant test is UMP also among all almost invariant tests and the UMP unbiased test is unique. In this sense, the principles of unbiasedness and of almost invariance are consistent.

Theorem 6.6.1 *Suppose that for a given testing problem there exists a UMP unbiased test ϕ^* which is unique (up to sets of measure zero), and that there also exists a UMP almost invariant test with respect to some group G. Then the latter is also unique (up to sets of measure zero), and the two tests coincide a.e.*

PROOF. If $U(\alpha)$ is the class of unbiased level-α tests, and if $g \in G$, then $\phi \in U(\alpha)$ if and only if $\phi g \in U(\alpha)$.[4] Denoting the power function of the test ϕ by $\beta_\phi(\theta)$, we thus have

$$\beta_{\phi^* g}(\theta) = \beta_{\phi^*}(\bar{g}\theta) = \sup_{\phi \in U(\alpha)} \beta_\phi(\bar{g}\theta) = \sup_{\phi \in U(\alpha)} \beta_{\phi g}(\theta)$$
$$= \sup_{\phi g \in U(\alpha)} \beta_{\phi g}(\theta) = \beta_{\phi^*}(\theta).$$

It follows that ϕ^* and $\phi^* g$ have the same power function, and, because of the uniqueness assumption, that ϕ^* is almost invariant. Therefore, if ϕ' is UMP almost invariant, we have $\beta_{\phi'}(\theta) \geq \beta_{\phi^*}(\theta)$ for all θ. On the other hand, ϕ' is unbiased, as is seen by comparing it with the invariant test $\phi(x) \equiv \alpha$, and hence $\beta_{\phi'}(\theta) \leq \beta_{\phi^*}(\theta)$

---

[4] ϕg denotes the critical function which assigns to x the value $\phi(gx)$.

for all θ. Since ϕ' and ϕ^* therefore have the same power function, they are equal a.e. because of the uniqueness of ϕ^*, as was to be proved. ■

This theorem provides an alternative derivation for some of the tests of Chapter 5. In Theorem 4.4.1, the existence of UMP unbiased tests was established for one- and two-sided hypotheses concerning the parameter θ of the exponential family (4.10). For this family, the statistics (U, T) are sufficient and complete, and in terms of these statistics the UMP unbiased test is therefore unique. Convenient explicit expressions for some of these tests, which were derived in Chapter 5, can instead be obtained by noting that when a UMP almost invariant test exists, the same test by Theorem 6.6.1 must also be UMP unbiased. This proves, for example, that the tests of Examples 6.3.4 and 6.3.5 are UMP unbiased.

The principles of unbiasedness and invariance can be used to supplement each other in cases where neither principle alone leads to a solution but where they do so when applied in conjunction. As an example consider a sample $X_1, \dots, X_n$ from $N(\xi, \sigma^2)$ and the problem of testing $H : \xi/\sigma = \theta_0 \neq 0$ against the two-sided alternatives that $\xi/\sigma \neq \theta_0$. Here sufficiency and invariance reduce the problem to the consideration of $t = \sqrt{n}\bar{x}/\sqrt{\sum(x_i - \bar{x})^2/(n-1)}$. The distribution of this statistic is the noncentral t-distribution with noncentrality parameter $\delta = \sqrt{n}\xi/\sigma$ and $n-1$ degrees of freedom. For varying δ, the family of these distributions can be shown to be STP_∞. [Karlin (1968, pp. 118–119; see Problem 3.55] and hence in particular STP_3. It follows by Problem 3.57 that among all tests of H based on t, there exists a UMP unbiased one with acceptance region $C_1 \le t \le C_2$, where C_1, C_2 are determined by the conditions

$$P_{\delta_0}\{C_1 \le t \le C_2\} = 1 - \alpha \quad \text{and} \quad \left.\frac{\partial P_\delta\{C_1 \le t \le C_2\}}{\partial \delta}\right|_{\delta=\delta_0} = 0.$$

In terms of the original observations, this test then has the property of being UMP among all tests that are unbiased and invariant. Whether it is also UMP unbiased without the restriction to invariant tests is an open problem.

An analogous example occurs in the testing of the hypotheses $H : \rho = \rho_0$ and $H' : \rho_1 \le \rho \le \rho_2$ against two-sided alternatives on the basis of a sample from a bivariate normal distribution with correlation coefficient ρ. (The testing of $\rho \le \rho_0$ against $\rho > \rho_0$ is treated in Problem 6.20.) The distribution of the sample correlation coefficient has not only monotone likelihood ratio as shown in Problem 6.20, but is in fact STP_∞. [Karlin (1968, Section 3.4)]. Hence there exist tests of both H and H' which are UMP among all tests that are both invariant and unbiased.

Another case in which the combination of invariance and unbiasedness appears to offer a promising approach is the *Behrens–Fisher problem*. Let $X_1, \dots, X_m$ and $Y_1, \dots, Y_n$ be samples from normal distributions $N(\xi, \sigma^2)$ and $N(\eta, \tau^2)$, respectively. The problem is that of testing $H : \eta \le \xi$ (or $\eta = \xi$) without assuming equality of the variances σ^2 and τ^2. A set of sufficient statistics for $(\xi, \eta, \sigma, \tau)$ is then $(\bar{X}, \bar{Y}, S_X^2, S_Y^2)$, where $S_X^2 = \sum(X_i - \bar{X})^2/(m-1)$ and $S_Y^2 = \sum(Y_j - \bar{Y})^2/(n-1)$. Adding the same constant to $\bar{X}$ and $\bar{Y}$ reduces the problem to $\bar{Y} - \bar{X}, S_X^2, S_Y^2$, and multiplication of all variables by a common positive constant to $(\bar{Y} - \bar{X})/\sqrt{S_X^2 + S_Y^2}$ and S_Y^2/S_X^2. One would expect any reasonable invariant rejection region to be of the form

$$\frac{\bar{Y} - \bar{X}}{\sqrt{S_X^2 + S_Y^2}} \geq g\left(\frac{S_Y^2}{S_X^2}\right) \tag{6.19}$$

for some suitable function g. If this test is also to be unbiased, the probability of (6.19) must equal α when $\eta = \xi$ for all values of τ/σ. It has been shown by Linnik and others that only pathological functions g with this property can exist. [This work is reviewed by Pfanzagl (1974).] However, approximate solutions are available which provide tests that are satisfactory for all practical purposes. These are the Welch approximate t-solution described in Section 13.2, and the Welch–Aspin test. Both are discussed, and evaluated, in Scheffé (1970) and Wang (1971); see also Chernoff (1949), Wallace (1958), Davenport and Webster (1975) and Robinson (1982). The Behrens–Fisher problem will be revisited in Examples 15.5.4 and 18.5.4 and Section 17.3.

The property of a test ϕ_1 being UMP invariant is relative to a particular group G_1, and does not exclude the possibility that there might exist another test ϕ_2 which is UMP invariant with respect to a different group G_2. Simple instances can be obtained from Examples 6.5.1 and 6.6.2.

Example 6.6.1 If G_1 is the group G of Example 6.5.1, a UMP invariant test of $H : \theta \leq \theta_0$ against $\theta > \theta_0$ rejects when $Y_1 + \cdots + Y_n > C$. Let G_2 be the group obtained by interchanging the role of Z and Y_1. Then a UMP invariant test with respect to G_2 rejects when $Z + Y_2 + \cdots + Y_n > C$. Analogous UMP invariant tests are obtained by interchanging the role of Z and any one of the other Y's and further examples by applying the transformations of G in Example 6.5.1 to more than one variable. In particular, if it is applied independently to all $n + 1$ variables, only the constants remain invariant, and the test $\phi \equiv \alpha$ is UMP invariant. ■

Example 6.6.2 For another example (due to Charles Stein), let (X_{11}, X_{12}) and (X_{21}, X_{22}) be independent and have bivariate normal distributions with zero means and covariance matrices

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \Delta\sigma_1^2 & \Delta\rho\sigma_1\sigma_2 \\ \Delta\rho\sigma_1\sigma_2 & \Delta\sigma_2^2 \end{pmatrix}.$$

Suppose that these matrices are nonsingular, or equivalently that $|\rho| \neq 1$, but that all σ_1, σ_2, ρ, and Δ are otherwise unknown. The problem of testing $\Delta = 1$ against $\Delta > 1$ remains invariant under the group G_1 of all nonsingular transformations

$$\begin{aligned} X'_{i1} &= bX_{i1} \\ X'_{i2} &= a_1 X_{i1} + a_2 X_{i2}, \end{aligned} \qquad (a_2, b > 0).$$

Since the probability is 0 that $X_{11}X_{22} = X_{12}X_{21}$, the 2×2 matrix (X_{ij}) is nonsingular with probability 1, and the sample space can therefore be restricted to be the set of all nonsingular such matrices. A maximal invariant under the subgroup corresponding to $b = 1$ is the pair (X_{11}, X_{21}). The argument of Example 6.3.5 then shows that there exists a UMP invariant test under G_1 which rejects when $X_{21}^2/X_{11}^2 > C$.

By interchanging 1 and 2 in the second subscript of the X's one sees that under the corresponding group G_2 the UMP invariant test rejects when $X_{22}^2/X_{12}^2 > C$.

A third group leaving the problem invariant is the smallest group containing both G_1 and G_2, namely, the group G of all common nonsingular transformations

$$\begin{aligned} X'_{i1} &= a_{i1}X_{i1} + a_{12}X_{i2} \\ X'_{i2} &= a_{21}X_{i1} + a_{22}X_{i2} \end{aligned}, \qquad (i = 1, 2).$$

Given any two nonsingular sample points $Z = (X_{ij})$ and $Z' = (X'_{ij})$, there exists a nonsingular linear transformation A such that $Z' = AZ$. There are therefore no invariants under G, and the only invariant size-α test is $\phi \equiv \alpha$. It follows vacuously that this is UMP invariant under G. ■

6.7 Admissibility

Any UMP unbiased test has the important property of admissibility (Problem 4.1), in the sense that there cannot exist another test which is uniformly at least as powerful and against some alternatives actually more powerful than the given one. The corresponding property does not necessarily hold for UMP invariant tests, as is shown by the following example.

Example 6.7.1 (continued) Under the assumptions of Example 6.6.2 it was seen that the UMP invariant test under G is the test $\varphi \equiv \alpha$ which has power $\beta(\Delta) \equiv \alpha$. On the other hand, X_{11} and X_{21} are independently distributed as $N(0, \sigma_1^2)$ and $N(0, \Delta\sigma_1^2)$. On the basis of these observations there exists a UMP test for testing $\Delta = 1$ against $\Delta > 1$ with rejection region $X_{21}^2/X_{11}^2 > C$ (Problem 3.67). The power function of this test is strictly increasing in Δ and hence $> \alpha$ for all $\Delta > 1$. ■

Admissibility of optimum invariant tests therefore cannot be taken for granted but must be established separately for each case.

We shall distinguish two slightly different concepts of admissibility. A test φ_0 will be called *α-admissible* for testing $H : \theta \in \Omega_H$ against a class of alternatives $\theta \in \Omega'$ if for any other level-α test φ

$$E_\theta\varphi(X) \geq E_\theta\varphi_0(X) \qquad \text{for all} \quad \theta \in \Omega' \tag{6.20}$$

implies $E_\theta\varphi(X) = E_\theta\varphi_0(X)$ for all $\theta \in \Omega'$. This definition takes no account of the relationship of $E_\theta\varphi(X)$ and $E_\theta\varphi_0(X)$ for $\theta \in \Omega_H$ beyond the requirement that both tests are of level α. For some unexpected, and possibly undesirable consequences of α-admissibility, see Example 8.7.1 and Perlman and Wu (1999). A concept closer to the decision-theoretic notion of admissibility discussed in Section 1.8, defines φ_0 to be *d-admissible* for testing H against Ω' if (6.20) and

$$E_\theta\varphi(X) \le E_\theta\varphi_0(X) \qquad \text{for all} \quad \theta \in \Omega_H \tag{6.21}$$

jointly imply $E_\theta\varphi(X) = E_\theta\varphi_0(X)$ for all $\theta \in \Omega_H \cup \Omega'$ (see Problem 6.34).

Any level-α test φ_0 that is α-admissible is also d-admissible provided no other test φ exists with $E_\theta\varphi(X) = E_\theta\varphi_0(X)$ for all $\theta \in \Omega'$ but $E_\theta\varphi(X) \neq E_\theta\varphi_0(X)$ for some $\theta \in \Omega_H$. That the converse does not hold is shown by the following example.

Example 6.7.2 Let X be normally distributed with mean ξ and known variance σ^2. For testing $H : \xi \le -1$ or ≥ 1 against $\Omega' : \xi = 0$, there exists a level-α test φ_0, which rejects when $C_1 \le X \le C_2$ and accepts otherwise, such that (Problem 6.35)

$$E_\xi\varphi_0(X) \le E_{\xi=-1}\varphi_0(X) = \alpha \qquad \text{for} \quad \xi \le -1$$

and

$$E_\xi\varphi_0(X) \le E_{\xi=+1}\varphi_0(X) = \alpha' < \alpha \qquad \text{for} \quad \xi \ge +1.$$

A slight modification of the proof of Theorem 3.7.1 shows that φ_0 is the unique test maximizing the power at $\xi = 0$ subject to

$$E_\xi\varphi(X) \le \alpha \quad \text{for} \quad \xi \le -1 \quad \text{and} \quad E_\xi\varphi(X) \le \alpha' \quad \text{for} \quad \xi \ge 1,$$

and hence that φ_0 is d-admissible.

On the other hand, the test φ with rejection region $|X| \le C$, where $E_{\xi=-1}\varphi(X) = E_{\xi=1}\varphi(X) = \alpha$, is the unique test maximizing the power at $\xi = 0$ subject to $E_\xi\varphi(X) \le \alpha$ for $\xi \le -1$ or ≥ 1, and hence is more powerful against Ω' than φ_0, so that φ_0 is not α-admissible. ■

A test that is admissible under either definition against Ω' is also admissible against any Ω'' containing Ω' and hence in particular against the class of all alternatives $\Omega_K = \Omega - \Omega_H$. The terms α- and d-admissible without qualification will be reserved for admissibility against Ω_K. Unless a UMP test exists, any α-admissible test will be admissible against some $\Omega' \subset \Omega_K$ and inadmissible against others. Both the strength of an admissibility result and the method of proof will depend on the set Ω'.

Consider in particular the admissibility of a UMP unbiased test mentioned at the beginning of the section. This does not rule out the existence of a test with greater power for all alternatives of practical importance and smaller power only for alternatives so close to H that the value of the power there is immaterial. In the present section, we shall discuss two methods for proving admissibility against various classes of alternatives.

Theorem 6.7.1 *Let X be distributed according to an exponential family with density*

$$p_\theta(x) = C(\theta)\exp\left(\sum_{j=1}^{s}\theta_j T_j(x)\right)$$

with respect to a σ-finite measure μ over a Euclidean sample space $(\mathcal{X}, \mathcal{A})$, and let Ω be the natural parameter space of this family. Let Ω_H and Ω' be disjoint nonempty subsets of Ω, and suppose that φ_0 is a test of $H : \theta \in \Omega_H$ based on $T = (T_1, \ldots, T_s)$ with acceptance region A_0 which is a closed convex subset of R^s possessing the following property: If $A_0 \cap \{\sum a_i t_i > c\}$ is empty for some c, there exists a point $\theta^* \in \Omega$ and a sequence $\lambda_n \to \infty$ such that $\theta^* + \lambda_n a \in \Omega'$ [where λ_n is a scalar and $a = (a_1, \ldots, a_s)$]. Then if A is any other acceptance region for H satisfying

$$P_\theta(X \in A) \le P_\theta(X \in A_0) \quad \textit{for all} \quad \theta \in \Omega',$$

A is contained in A_0, except for a subset of measure 0, i.e., $\mu(A \cap A_0^c) = 0$.

PROOF. Suppose to the contrary that $\mu(A \cap A_0^c) > 0$. Then it follows from the closure and convexity of A_0, that there exist $a \in R^s$ and a real number c such that

$$A_0 \cap \left\{t : \sum a_i t_i > c\right\} \text{ is empty} \tag{6.22}$$

and

$$A \cap \left\{t : \sum a_i t_i > c\right\} \text{ has positive } \mu\text{-measure,} \tag{6.23}$$

that is, the set A protrudes in some direction from the convex set A_0. We shall show that this fact and the exponential nature of the densities imply that

$$P_\theta(A) > P_\theta(A_0) \quad \text{for some} \quad \theta \in \Omega', \tag{6.24}$$

which provides the required contradiction. Let φ_0 and φ denote the indicators of A_0^c and A^c, respectively, so that (6.24) is equivalent to

$$\int [\varphi_0(t) - \varphi(t)]\, dP_\theta(t) > 0 \quad \text{for some} \quad \theta \in \Omega'.$$

If $\theta = \theta^* + \lambda_n a \in \Omega'$, the left side becomes

$$\frac{C(\theta^* + \lambda_n a)}{C(\theta^*)} e^{c\lambda_n} \int [\varphi_0(t) - \varphi(t)] e^{\lambda_n(\sum a_i t_i - c)}\, dP_{\theta^*}(t).$$

Let this integral be $I_n^+ + I_n^-$, where I_n^+ and I_n^- denote the contributions over the regions of integration $\{t : \sum a_i t_i > c\}$ and $\{t : \sum a_i t_i \le c\}$, respectively. Since I_n^- is bounded, it is enough to show that $I_n^+ \to \infty$ as $n \to \infty$. By (6.22), $\varphi_0(t) = 1$ and hence $\varphi_0(t) - \varphi(t) \ge 0$ when $\sum a_i t_i > c$, and by (6.23)

$$\mu\left\{\varphi_0(t) - \varphi(t) > 0 \quad \text{and} \quad \sum a_i t_i > c\right\} > 0.$$

This shows that $I_n^+ \to \infty$ as $\lambda_n \to \infty$ and therefore completes the proof. ∎

Corollary 6.7.1 *Under the assumptions of Theorem 6.7.1, the test with acceptance region A_0 is d-admissible. If its size is α and there exists a finite point θ_0 in the closure $\bar{\Omega}_H$ of Ω_H for which $E_{\theta_0}\varphi_0(X) = \alpha$, then φ_0 is also α-admissible.*

PROOF.

(i) Suppose φ satisfies (6.20). Then by Theorem 6.7.1, $\varphi_0(x) \le \varphi(x)$ (a.e. μ). If $\varphi_0(x) < \varphi(x)$ on a set of positive measure, then $E_\theta\varphi_0(X) < E_\theta\varphi(X)$ for all θ and hence (6.21) cannot hold.

(ii) By the argument of part (i), (6.20) implies $\alpha = E_{\theta_0}\varphi_0(X) < E_{\theta_0}\varphi(X)$, and hence by the continuity of $E_\theta\varphi(X)$ there exists a point $\theta \in \Omega_H$ for which $\alpha < E_\theta\varphi(X)$. Thus φ is not a level-α test. ∎

Theorem 6.7.1 and the corollary easily extend to the case where the competitors φ of φ_0 are permitted to be randomized but the assumption that φ_0 is nonrandomized is essential. Thus, the main applications of these results are to the case that μ is absolutely continuous with respect to Lebesgue measure. The boundary of A_0 will then typically have measure zero, so that the closure requirement for A_0 can be dropped.

Example 6.7.3 (Normal mean) If $X_1, \dots, X_n$ is a sample from the normal distribution $N(\xi, \sigma^2)$, the family of distributions is exponential with $T_1 = \bar{X}$, $T_2 = \sum X_i^2$, $\theta_1 = n\xi/\sigma^2$, $\theta_2 = -1/2\sigma^2$. Consider first the one-sided problem $H : \theta_1 \le 0$, $K : \theta_1 > 0$ with $\alpha < \frac{1}{2}$. Then the acceptance region of the t-test is $A : T_1/\sqrt{T_2} \le C$ ($C > 0$), which is convex (Problem 6.36(i)). The alternatives $\theta \in \Omega' \subseteq K$ will satisfy the conditions of Theorem 6.7.1 if for any half plane $a_1t_1 + a_2t_2 > c$ that does not intersect the set $t_1 \le C\sqrt{t_2}$ there exists a ray $(\theta_1^* + \lambda a_1, \theta_2^* + \lambda a_2)$ in the direction of the vector (a_1, a_2) for which $(\theta_1^* + \lambda a_1, \theta_2^* + \lambda a_2) \in \Omega'$ for all sufficiently large λ. In the present case, this condition must hold for all $a_1 > 0 > a_2$. Examples of sets Ω' satisfying this requirement (and against which the t-test is therefore admissible) are

$$\Omega_1' : \theta_1 > k_1 \text{ or } \frac{\xi}{\sigma^2} > k_1'$$

and

$$\Omega_2' : \frac{\theta_1}{\sqrt{-\theta_2}} > k_2 \text{ or } \frac{\xi}{\sigma} > k_2'.$$

On the other hand, the condition is not satisfied for $\Omega' : \xi > k$ (Problem 6.36).

Analogously, the acceptance region $A : T_1^2 \le CT_2$ of the two-sided t-test for testing $H : \theta_1 = 0$ against $\theta_1 \ne 0$ is convex, and the test is admissible against $\Omega_1' : |\xi/\sigma^2| > k_1$ and $\Omega_2' : |\xi/\sigma| > k_2$. ∎

In decision theory, a quite general method for proving admissibility consists in exhibiting a procedure as a unique Bayes solution. In the present case, this is justified by the following result, which is closely related to Theorem 3.8.1.

Theorem 6.7.2 *Assume the set* $\{x : f_\theta(x) > 0\}$ *is independent of* θ*, and let a* σ*-field be defined over the parameter space* Ω*, containing both* Ω_H *and* Ω_K *and such that the densities* $f_\theta(x)$ *(with respect to* μ*) of* X *are jointly measurable in* θ *and* x*. Let* Λ_0 *and* Λ_1 *be probability distributions over this* σ*-field with* $\Lambda_0(\Omega_H) = \Lambda_1(\Omega_K) = 1$*, and let*

$$h_i(x) = \int f_\theta(x)\, d\Lambda_i(\theta).$$

Suppose φ_0 *is a nonrandomized test of* H *against* K *defined by*

$$\varphi_0(x) = \begin{cases} 1 \ when\ h_1(x) \geq k h_0(x) \\ 0 \ when\ h_1(x) < k h_0(x) \end{cases} \tag{6.25}$$

and that $\mu\{x : h_1(x)/h_0(x) = k\} = 0$*.*

(i) *Then,* φ_0 *is* d*-admissible for testing* H *against* K*.*

(ii) *Let* $\sup_{\Omega_H} E_\theta \varphi_0(X) = \alpha$ *and*

$$\omega = \{\theta :\in \Omega_H : E_\theta \varphi_0(X) = \alpha\}\ .$$

If $\Lambda_0(\omega) = 1$*, then* φ_0 *is also* α*-admissible.*

(iii) *If* Λ_1 *assigns probability 1 to* $\Omega' \subset \Omega_K$*, then the conclusions of (i) and (ii) apply with* Ω' *in place of* Ω_K*.*

PROOF. (i): Suppose φ is any other test, satisfying (6.20) and (6.21) with $\Omega' = \Omega_K$. Then also

$$\int E_\theta \varphi(X)\, d\Lambda_0(\theta) \leq \int E_\theta \varphi_0(X)\, d\Lambda_0(\theta)$$

and

$$\int E_\theta \varphi(X)\, d\Lambda_1(\theta) \geq \int E_\theta \varphi_0(X)\, d\Lambda_1(\theta).$$

By the argument of Theorem 3.8.1, these inequalities are equivalent to

$$\int \varphi(x) h_0(x)\, d\mu(x) \leq \int \varphi_0(x) h_0(x)\, d\mu(x)$$

and

$$\int \varphi(x) h_1(x)\, d\mu(x) \geq \int \varphi_0(x) h_1(x)\, d\mu(x),$$

and the $h_i(x)$ $(i = 0, 1)$ are probability densities with respect to μ. This contradicts the uniqueness of the most powerful test of h_0 against h_1 at level $\int \varphi(x) h_0(x)\, d\mu(x)$.

(ii): By assumption, $\int E_\theta \varphi_0(x)\, d\Lambda_0(\theta) = \alpha$, so that φ_0 is a level-α test of h_0. If φ is any other level-α test of H satisfying (6.20) with $\Omega' = \Omega_K$, it is also a level-α test of h_0 and the argument of part (i) can be applied as before.

(iii): This follows immediately from the proofs of (i) and (ii). ■

Example 6.7.4 In the two-sided normal problem of Example 6.7.3 with $H : \xi = 0$, $K : \xi \neq 0$ consider the class $\Omega'_{a,b}$ of alternatives (ξ, σ) satisfying

$$\sigma^2 = \frac{1}{a + \eta^2}, \quad \xi = \frac{b\eta}{a + \eta^2}, \qquad -\infty < \eta < \infty \tag{6.26}$$

for some fixed $a, b > 0$, and the subset ω, of Ω_H of points $(0, \sigma^2)$ with $\sigma^2 < 1/a$. Let Λ_0, Λ_1 be distributions over ω and $\Omega'_{a,b}$ defined by the densities (Problem 6.37(i))

$$\lambda_0(\eta) = \frac{C_0}{(a + \eta^2)^{n/2}}$$

and

$$\lambda_1(\eta) = \frac{C_1 e^{(n/2)b^2\eta^2/(a+\eta^2)}}{(a + \eta^2)^{n/2}}.$$

Straightforward calculation then shows (Problem 6.37(ii)) that the densities h_0 and h_1 of Theorem 6.7.2 become

$$h_0(x) = \frac{C_0 e^{-(a/2)\sum x_i^2}}{\sqrt{\sum x_i^2}}$$

and

$$h_1(x) = \frac{C_1 \exp\left(-\frac{a}{2}\sum x_i^2 + \frac{b^2\left(\sum x_i\right)^2}{2\sum x_i^2}\right)}{\sqrt{\sum x_i^2}},$$

so that the Bayes test φ_0 of Theorem 6.7.2 rejects when $\bar{x}^2 / \sum x_i^2 > k$ and hence reduces to the two-sided t-test.

The condition of part (ii) of the theorem is clearly satisfied so that the t-test is both d- and α-admissible against $\Omega'_{a,b}$.

When dealing with invariant tests, it is of particular interest to consider admissibility against invariant classes of alternatives. In the case of the two-sided test φ_0, this means sets Ω' depending only on $|\xi/\sigma|$. It was seen in Example 6.7.4 that φ_0 is admissible against $\Omega' : |\xi/\sigma| \geq B$ for any B, that is, against distant alternatives, and it follows from the test being UMP unbiased or from Example 6.7.4 (continued) that φ_0, is admissible against $\Omega' : |\xi/\sigma| \leq A$ for any $A > 0$, that is, against alternatives close to H. This leaves open the question whether φ_0 is admissible against sets $\Omega' : 0 < A < |\xi/\sigma| < B < \infty$, which include neither nearby nor distant alternatives. It was in fact shown by Lehmann and Stein (1953) that φ_0 is admissible for testing H against $|\xi|/\sigma = \delta$ for any $\delta > 0$ and hence that it is admissible against any invariant Ω'. It was also shown there that the one-sided t-test of $H : \xi = 0$ is admissible against $\xi/\sigma = \delta'$ for any $\delta' > 0$. These results will not be proved here. The proof is based on assigning to log σ the uniform density on $(-N, N)$ and letting $N \to \infty$, thereby approximating the "improper" prior distribution which assigns to log a the uniform distribution on $(-\infty, \infty)$, that is, Lebesgue measure.

That the one-sided t-test φ_1 of $H : \xi < 0$ is not admissible against all Ω' is shown by Brown and Sackrowitz (1984), who exhibit a test φ satisfying

$$E_{\xi,\sigma}\varphi(X) < E_{\xi,\sigma}\varphi_1(X) \quad \text{for all} \quad \xi < 0,\ 0 < \sigma < \infty$$

and ■

$$E_{\xi,\sigma}\varphi(X) > E_{\xi,\sigma}\varphi_1(X) \quad \text{for all} \quad 0 < \xi_1 < \xi < \xi_2 < \infty,\ 0 < \sigma < \infty.$$

Example 6.7.5 **(Normal variance)** For testing the variance σ^2 of a normal distribution on the basis of a sample $X_1, \dots, X_n$ from $N(\xi, \sigma^2)$, the Bayes approach of Theorem 6.7.2 easily proves α-admissibility of the standard test against any location invariant set of alternatives Ω', that is, any set Ω' depending only on σ^2. Consider first the one-sided hypothesis $H : \sigma \le \sigma_0$ and the alternatives $\Omega' : \sigma = \sigma_1$ for any $\sigma_1 > \sigma_0$. Admissibility of the UMP invariant (and unbiased) rejection region $\sum(X_i - \bar{X})^2 > C$ follows immediately from Section 3.9, where it was shown that this test is Bayes for a pair of prior distributions (Λ_0, Λ_1), namely, Λ_1 assigning probability 1 to any point (ξ_1, σ_1), and Λ_0 putting $\sigma = \sigma_0$ and assigning to ξ the normal distribution $N(\xi_1, (\sigma_1^2 - \sigma_0^2)/n)$. Admissibility of $\sum(X_i - \bar{X})^2 \le C$ when the hypothesis is $H : \sigma \ge \sigma_0$ and $\Omega' = \{(\xi, \sigma) : \sigma = \sigma_1\}$, $\sigma_1 < \sigma_0$, is seen by interchanging Λ_0 and Λ_1, σ_0 and σ_1.

A similar approach proves α-admissibility of any size-α rejection region

$$\sum(X_i - \bar{X})^2 \le C_1 \text{ or } \ge C_2 \tag{6.27}$$

for testing $H : \sigma = \sigma_0$ against $\Omega' : \{\sigma = \sigma_1\} \cup \{\sigma = \sigma_2\}$ $(\sigma_1 < \sigma_0 < \sigma_2)$. On Ω_H, where the only variable is ξ, the distribution Λ_0 for ξ can be taken as the normal distribution with an arbitrary mean ξ_1 and variance $(\sigma_2^2 - \sigma_0^2)/n$. On Ω', let the conditional distribution of ξ given $\sigma = \sigma_2$ assign probability 1 to the value ξ_1, and let the conditional distribution of ξ given $\sigma = \sigma_1$ be $N(\xi_1, (\sigma_2^2 - \sigma_1^2)/n)$. Finally, let Λ_1 assign probabilities p and $1 - p$ to $\sigma = \sigma_1$ and $\sigma = \sigma_2$, respectively. Then the rejection region satisfies (6.27), and any constants C_1 and C_2 for which the test has size a can be attained by proper choice of p (Problem 6.38(i)). ■

The results of Examples 6.7.4 and 6.7.5 can be used as the basis for proving admissibility results in many other situations involving normal distributions. The main new difficulty tends to be the presence of additional (nuisance) means. These can often be eliminated by use of the following lemma.

Lemma 6.7.1 *For any given* σ^2 *and* $M^2 > \sigma^2$ *there exists a distribution* Λ_σ *such that*

$$I(z) = \int \frac{1}{\sqrt{2\pi}\sigma} e^{-(1/2\sigma^2)(z-\zeta)^2} d\Lambda_\sigma(\zeta)$$

is the normal density with mean zero and variance M^2.

PROOF. Let $\theta = \zeta/\sigma$, and let θ be normally distributed with zero mean and variance τ^2. Then it is seen (Problem 6.38(ii)) that

$$I(z) = \frac{1}{\sqrt{2\pi}\sigma\sqrt{1+\tau^2}} \exp\left[-\frac{1}{2\sigma^2(1+\tau^2)} z^2\right].$$

The result now follows by letting $\tau^2 = (M^2/\sigma^2) - 1$, so that $\sigma^2(1+\tau^2) = M^2$. ■

Example 6.7.6 Let $X_1, \dots, X_m$; $Y_1, \dots, Y_n$ be samples from $N(\xi, \sigma^2)$ and $N(\eta, \tau^2)$, respectively, and consider the problem of testing $H : \tau/\sigma = 1$ against $\tau/\sigma = \Delta > 1$.

(i) Suppose first that $\xi = \eta = 0$. If Λ_0 and Λ_1 assign probability 1 to the points $(\sigma_0, \tau_0 = \sigma_0)$ and $(\sigma_1, \tau_1 = \Delta\sigma_1)$, respectively, the ratio h_1/h_0 of Theorem 6.7.2 is proportional to

$$\exp\left\{-\frac{1}{2}\left[\left(\frac{1}{\Delta^2\sigma_1^2} - \frac{1}{\sigma_0^2}\right)\sum y_j^2 - \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)\sum x_i^2\right]\right\},$$

and for suitable choice of critical value and $\sigma_1 < \sigma_0$, the rejection region of the Bayes test reduces to

$$\frac{\sum y_j^2}{\sum x_i^2} > \frac{\Delta^2\sigma_1^2 - \sigma_0^2}{\sigma_0^2 - \sigma_1^2}.$$

The values σ_0^2 and σ_1^2 can then be chosen to give this test any preassigned size α.

(ii) If ξ and η are unknown, then $\bar{X}$, $\bar{Y}$, $S_X^2 = \sum(X_i - \bar{X})^2$, $S_Y^2 = \sum(Y_j - \bar{Y})^2$ are sufficient statistics, and S_X^2 and S_Y^2 can be represented as $S_X^2 = \sum_{i=1}^{m-1} U_i^2$, $S_Y^2 = \sum_{j=1}^{n-1} V_j^2$, with the U_i, V_j independent normal with means 0 and variances σ^2 and τ^2 respectively.

To σ and τ assign the distributions Λ_0 and Λ_1 of part (i) and conditionally, given σ and τ, let ξ and η be independently distributed according to $\Lambda_{0\sigma}$, $\Lambda_{0\tau}$, over Ω_H and $\Lambda_{1\sigma}$, $\Lambda_{1\tau}$ over Ω_K, with these four conditional distributions determined from Lemma 6.7.1 in such a way that

$$\int \frac{\sqrt{m}}{\sqrt{2\pi}\sigma_0} e^{-(m/2\sigma_0^2)(\bar{x}-\xi)^2} d\Lambda_{0\sigma_0}(\xi) = \int \frac{\sqrt{m}}{\sqrt{2\pi}\sigma_1} e^{-(m/2\sigma_1^2)(\bar{x}-\xi)^2} d\Lambda_{0\sigma_1}(\xi),$$

and analogously for η. This is possible by choosing the constant M^2 of Lemma 6.7.1 greater than both σ_0^2 and σ_1^2. With this choice of priors, the contribution from $\bar{x}$ and $\bar{y}$ to the ratio h_1/h_0 of Theorem 6.7.2 disappears, so that h_1/h_0 reduces to the expression for this ratio in part (i), with $\sum x_i^2$ and $\sum y_j^2$ replaced by $\sum(x_i - \bar{x})^2$ and $\sum(y_j - \bar{y})^2$, respectively. ■

This approach applies quite generally in normal problems with nuisance means, provided the prior distribution of the variances σ^2, τ^2, ... assigns probability 1 to a bounded set, so that M^2 can be chosen to exceed all possible values of these variances.

Admissibility questions have been considered not only for tests but also for confidence sets. These will not be treated here (but see Example 8.5.4); convenient entries to the literature are Cohen and Strawderman (1973) and Joshi (1982). For additional results, see Hooper (1982a) and Arnold (1984).

6.8 Rank Tests

One of the basic problems of statistics is the two-sample problem of testing the equality of two distributions. A typical example is the comparison of a treatment with a control, where the hypothesis of no treatment effect is tested against the alternatives of a beneficial effect. This was considered in Chapter 5 under the assumption of normality, and the appropriate test was seen to be based on Student's t. It was also shown that when approximate normality is suspected but the assumption cannot be trusted, one is led to replacing the t-test by its permutation analogue, which in turn can be approximated by the original t-test. For further details, see Chapter 17.

We shall consider the same problem below without, at least for the moment, making any assumptions concerning even the approximate form of the underlying distributions, assuming only that they are continuous. The observations then consist of samples $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ from two distributions with continuous cumulative distribution functions F and G, and the problem becomes that of testing the hypothesis

$$H_1 : G = F.$$

If the treatment effect is assumed to be additive, the alternatives are $G(y) = F(y - \Delta)$. We shall here consider the more general possibility that the size of the effect may depend on the value of y (so that Δ becomes a nonnegative function of y) and therefore test H_1 against the one-sided alternatives that the Y's are stochastically larger than the X's,

$$K_1 : G(z) \leq F(z) \quad \text{for all } z, \quad \text{and} \quad G \neq F.$$

An alternative experiment that can be performed to test the effect of a treatment consists of the comparison of N pairs of subjects, which have been matched so as to eliminate as far as possible any differences not due to the treatment. One member of each pair is chosen at random to receive the treatment while the other serves as control. If the normality assumption of Section 5.10 is dropped and the pairs of subjects can be considered to constitute a sample, the observations $(X_1, Y_1), \ldots, (X_N, Y_N)$ are a sample from a continuous bivariate distribution F. The hypothesis of no effect is then equivalent to the assumption that F is symmetric with respect to the line $y = x$:

$$H_2 : F(x, y) = F(y, x).$$

Another basic problem, which occurs in many different contexts, concerns the dependence or independence of two variables. In particular, if $(X_1, Y_1), \ldots, (X_N, Y_N)$ is a sample from a bivariate distribution F, one will be interested in the hypothesis

$$H_3 : F(x, y) = G_1(x)G_2(y)$$

that X and Y are independent, which was considered for normal distributions in Section 5.13. The alternatives of interest may, for example, be that X and Y are positively dependent. An alternative formulation results when x, instead of being random, can be selected for the experiment. If the chosen values are $x_1 < \cdots < x_N$ and F_i denotes the distribution of Y given x_i, the Y's are independently distributed with continuous cumulative distribution functions $F_1, \ldots, F_N$. The hypothesis of independence of Y from x becomes

$$H_4 : F_1 = \cdots = F_N,$$

while under the alternatives of positive regression dependence the variables Y_i are stochastically increasing with i.

In these and other similar problems, invariance reduces the data so completely that the actual values of the observations are discarded and only certain order relations between different groups of variables are retained. It is nevertheless possible on this basis to test the various hypotheses in question, and the resulting tests frequently are nearly as powerful as the standard normal tests. We shall now carry out this reduction for the four problems above.

The two-sample problem of testing H_1 against K_1 remains invariant under the group G of all transformations

$$x_i' = \rho(x_i), \quad y_j' = \rho(y_j) \qquad (i = 1, \ldots, m, \quad j = 1, \ldots, n)$$

such that ρ is continuous and strictly increasing. This follows from the fact that these transformations preserve both the continuity of a distribution and the property of two variables being either identically distributed or one being stochastically larger than the other. As was seen (with a different notation) in Example 6.2.3, a maximal invariant under G is the set of ranks

$$(R'; S') = (R_1', \ldots, R_m'; S_1', \ldots, S_n')$$

of $X_1, \dots, X_m; Y_1, \dots, Y_n$ in the combined sample. Since the distribution of $(R'_1, \dots, R'_m; S'_1, \dots, S'_n)$ is symmetric in the first m and in the last n variables for all distributions F and G, a set of sufficient statistics for (R', S') is the set of the X-ranks and that of the Y-ranks without regard to the subscripts of the X's and Y's. This can be represented by the ordered X-ranks and Y-ranks

$$R_1 < \cdots < R_m \quad \text{and} \quad S_1 < \cdots < S_n,$$

and therefore by one of these sets alone since each of them determines the other. Any invariant test is thus a *rank test*, that is, it depends only on the ranks of the observations, for example, on $(S_1, \dots, S_n)$.

That almost invariant tests are equivalent to invariant ones in the present context was shown first by Bell (1964). A streamlined and generalized version of his approach is given by Berk and Bickel (1968) and Berk (1970), who also show that the conclusion of Theorem 6.5.3 remains valid in this case.

To obtain a similar reduction for H_2, it is convenient first to make the transformation $Z_i = Y_i - X_i$, $W_i = X_i + Y_i$. The pairs of variables (Z_i, W_i) are then again a sample from a continuous bivariate distribution. Under the hypothesis this distribution is symmetric with respect to the w-axis, while under the alternatives the distribution is shifted in the direction of the positive z-axis. The problem is unchanged if all the w's are subjected to the same transformation $w'_i = \lambda(w_i)$, where λ is 1 : 1 and has at most a finite number of discontinuities, and $(Z_1, \dots, Z_N)$ constitutes a maximal invariant under this group. [Cf. Problem 6.2(ii).]

The Z's are a sample from a continuous univariate distribution D, for which the hypothesis of symmetry with respect to the origin,

$$H'_2 : D(z) + D(-z) = 1 \quad \text{for all } z,$$

is to be tested against the alternatives that the distribution is shifted toward positive z-values. This problem is invariant under the group G of all transformations

$$z'_i = \rho(z_i) \qquad (i = 1, \dots, N)$$

such that ρ is continuous, odd, and strictly increasing. If $z_{i_1}, \dots, z_{i_m} < 0 < z_{j_1}, \dots, z_{j_n}$, where $i_1 < \cdots < i_m$ and $j_1 < \cdots < j_n$, let $s'_1, \dots, s'_n$ denote the ranks of $z_{j_i}, \dots, z_{j_n}$, among the absolute values $|z_1|, \dots, |z_N|$, and $r'_1, \dots, r'_m$ the ranks of $|z_{i_1}|, \dots, |z_{i_m}|$ among $|z_1|, \dots, |z_N|$. The transformations ρ preserve the sign of each observation, and hence in particular also the numbers m and n. Since ρ is a continuous, strictly increasing function of $|z|$, it leaves the order of the absolute values invariant and therefore the ranks r'_i and s'_j. To see that the latter are maximal invariant, let $(z_1, \dots, z_N)$ and $(z'_1, \dots, z'_N)$ be two sets of points with $m' = m$, $n' = n$, and the same r'_i and s'_j. There exists a continuous, strictly increasing function on the positive real axis such that $|z'_i| = \rho(|z_i|)$ and $\rho(0) = 0$. If ρ is defined for negative z by $\rho(-z) = -\rho(z)$, it belongs to G and $z'_i = \rho(z_i)$ for all i, as was to be proved. As in the preceding problem, sufficiency permits the further reduction to the ordered ranks $r_1 < \cdots < r_m$ and $s_1 < \cdots < s_n$. This retains the information for the rank of each absolute value whether it belongs to a positive or negative observation, but not with which positive or negative observation it is associated.

The situation is very similar for the hypotheses H_3 and H_4. The problem of testing for independence in a bivariate distribution against the alternatives of positive dependence is unchanged if the X_i and Y_i are subjected to transformations $X_i' = \rho(X_i)$, $Y_i' = \lambda(Y_i)$ such that ρ and λ are continuous and strictly increasing. This leaves as maximal invariant the ranks $(R_1', \ldots, R_N')$ of $(X_1, \ldots, X_N)$ among the X's and the ranks $(S_1', \ldots, S_N')$ of $(Y_1, \ldots, Y_N)$ among the Y's. The distribution of $(R_1', S_1'), \ldots, (R_N', S_N')$ is symmetric in these N pairs for all distributions of (X, Y). It follows that a sufficient statistic is $(S_1, \ldots, S_N)$ where $(1, S_1), \ldots, (N, S_N)$ is a permutation of $(R_1', S_1'), \ldots, (R_N', S_N')$ and where therefore S_i is the rank of the variable Y associated with the ith smallest X.

The hypothesis H_4 that $Y_1, \ldots, Y_n$ constitutes a sample is to be tested against the alternatives K_4 that the Y_i are stochastically increasing with i. This problem is invariant under the group of transformations $y_i' = \rho(y_i)$ where ρ is continuous and strictly increasing. A maximal invariant under this group is the set of ranks $S_1, \ldots, S_N$ of $Y_1, \ldots, Y_N$.

Some invariant tests of the hypotheses H_1 and H_2 will be considered in the next two sections. Corresponding results concerning H_3 and H_4 are given in Problems 6.62–6.64.

6.9 The Two-Sample Problem

The problem of testing the two-sample hypothesis $H : G = F$ against the one-sided alternatives K that the Y's are stochastically larger than the X's is reduced by the principle of invariance to the consideration of tests based on the ranks $S_1 < \cdots < S_n$ of the Y's. The specification of the S_i is equivalent to specifying for each of the $N = m + n$ positions within the combined sample (the smallest, the next smallest, etc.) whether it is occupied by an x or a y. Since for any set of observations n of the N positions are occupied by y's and since the $\binom{N}{n}$ possible assignments of n positions to the y's are all equally likely when $G = F$, the joint distribution of the S_i under H is

$$P\{S_1 = s_1, \ldots, S_n = s_n\} = 1 \Big/ \binom{N}{n} \tag{6.28}$$

for each set $1 \le s_1 < s_2 < \cdots < s_n \le N$. Any rank test of H of size

$$\alpha = k \Big/ \binom{N}{n}$$

therefore has a rejection region consisting of exactly k points $(s_1, \ldots, s_n)$.

For testing H against K there exists no UMP rank test, and hence no UMP invariant test. This follows, for example, from a consideration of two of the standard tests for this problem, since each is most powerful among all rank tests against some alternative. The two tests in question have rejection regions of the form

$$h(s_1) + \cdots + h(s_n) > C. \tag{6.29}$$

One, the Wilcoxon *two-sample test*, is obtained from (6.29) by letting $h(s) = s$, so that it rejects H when the sum of the y-ranks is too large. We shall show below that for sufficiently small Δ, this is most powerful against the alternatives that F is the logistic distribution $F(x) = 1/(1 + e^{-x})$, and that $G(y) = F(y - \Delta)$. The other test, the *normal scores test*, has the rejection region (6.29) with $h(s) = E(W_{(s)})$, where $W_{(1)} < \cdots < W_{(N)}$, is an ordered sample of size N from a standard normal distribution.[5] This is most powerful against the alternatives that F and G are normal distributions with common variance and means ξ and $\eta = \xi + \Delta$, when Δ is sufficiently small.

To prove that these tests have the stated properties it is necessary to know the distribution of $(S_1, \ldots, S_n)$ under the alternatives. If F and G have densities f and g such that f is positive whenever g is, the joint distribution of the S_i is given by

$$P\{S_1 = s_1, \ldots, S_n = s_n\} = E\left[\frac{g(V_{(s_1)})}{f(V_{(s_1)})} \cdots \frac{g(V_{(s_n)})}{f(V_{(s_n)})}\right] \Big/ \binom{N}{n}, \tag{6.30}$$

where $V_{(1)} < \cdots < V_{(N)}$ is an ordered sample of size N from the distribution F. (See Problem 6.44.) Consider in particular the translation (or shift) alternatives

$$g(y) = f(y - \Delta),$$

and the problem of maximizing the power for small values of Δ. Suppose that f is differentiable and that the probability (6.30), which is now a function of Δ, can be differentiated with respect to Δ under the expectation sign. The derivative of (6.30) at $\Delta = 0$ is then

$$\frac{\partial}{\partial \Delta} P_\Delta\{S_1 = s_1, \ldots, S_n = S_n\}\Big|_{\Delta=0} = -\sum_{i=1}^{n} E\left[\frac{f'(V_{(s_i)})}{f(V_{(s_i)})}\right] \Big/ \binom{N}{n}.$$

Since under the hypothesis the probability of any ranking is given by (6.28), it follows from the Neyman–Pearson Lemma in the extended form of Theorem 3.6.1, that the derivative of the power function at $\Delta = 0$ is maximized by the rejection region

$$-\sum_{i=1}^{n} E\left[\frac{f'(V_{(s_i)})}{f(V_{(s_i)})}\right] > C. \tag{6.31}$$

[5] Tables of the expected order statistics from a normal distribution are given in *Biometrika Tables for Statisticians*, Vol. 2, Cambridge U. P., 1972, Table 9. For additional references, see David (1981, Appendix, Section 3.2).

The same test maximizes the power itself for sufficiently small Δ. To see this let s denote a general rank point $(s_1, \dots, s_n)$, and denote by $s^{(j)}$ the rank point giving the jth largest value to the left-hand side of (6.31). If

$$\alpha = k \Big/ \binom{N}{n},$$

the power of the test is then

$$\beta(\Delta) = \sum_{j=1}^{k} P_\Delta(s^{(j)}) = \sum_{j=1}^{k} \left[\frac{1}{\binom{N}{n}} + \Delta \frac{\partial}{\partial \Delta} P_\Delta(s^{(j)}) \Big|_{\Delta=0} + \cdots \right].$$

Since there is only a finite number of points s, there exists for each j a number $\Delta_j > 0$ such that the point $s^{(j)}$ also gives the jth largest value to $P_\Delta(s)$ for all $\Delta < \Delta_j$. If Δ is less than the smallest of the numbers

$$\Delta_j, \qquad j = 1, \dots, \binom{N}{n},$$

the test also maximizes $\beta(\Delta)$.

If $f(x)$ is the normal density $N(\xi, \sigma^2)$, then

$$-\frac{f'(x)}{f(x)} = -\frac{d}{dx} \log f(x) = \frac{x - \xi}{\sigma^2},$$

and the left-hand side of (6.31) becomes

$$\sum E \frac{V_{(s_i)} - \xi}{\sigma^2} = \frac{1}{\sigma} \sum E(W_{(s_i)}),$$

where $W_{(1)} < \cdots < W_{(N)}$ is an ordered sample from $N(0, 1)$. The test that maximizes the power against these alternatives (for sufficiently small Δ) is therefore the normal scores test.

In the case of the logistic distribution,

$$F(x) = \frac{1}{1 + e^{-x}}, \qquad f(x) = \frac{e^{-x}}{(1 + e^{-x})^2},$$

and hence

$$-\frac{f'(x)}{f(x)} = 2F(x) - 1.$$

The locally most powerful rank test therefore rejects when $\sum E[F(V_{(x_i)})] > C$. If V has the distribution F, then $U = F(V)$ is uniformly distributed over $(0, 1)$ (Problem 3.22). The rejection region can therefore be written as $\sum E(U_{(s_i)}) > C$, where $U_{(1)} < \cdots < U_{(N)}$ is an ordered sample of size N from the uniform distribution $U(0, 1)$. Since $E(U_{(s_i)}) = s_i/(N+1)$, the test is seen to be the Wilcoxon test.

Both the normal scores test and the Wilcoxon test are unbiased against the one-sided alternatives K. In fact, let ϕ be the critical function of any test determined by (6.29) with h nondecreasing. Then ϕ is nondecreasing in the y's and the probability of rejection is α for all $F = G$. By Lemma 5.9.1 the test is therefore unbiased against all alternatives of K.

It follows from the unbiasedness properties of these tests that the most powerful invariant tests in the two cases considered are also most powerful against their respective alternatives among all tests that are invariant and unbiased. The nonexistence of a UMP test is thus not relieved by restricting the tests to be unbiased as well as invariant. Nor does the application of the unbiasedness principle alone lead to a solution, as was seen in the discussion of permutation tests in Section 5.9. With the failure of these two principles, both singly and in conjunction, the problem is left not only without a solution but even without a formulation. A possible formulation (stringency) will be discussed in Chapter 8. However, the determination of a most stringent test for the two-sample hypothesis is an open problem.

For testing $H : G = F$ against the two-sided alternatives that the Y's are either stochastically smaller or larger than the X's, two-sided versions of the rank tests of this section can be used. In particular, suppose that h is increasing and that $h(s) + h(N+1-s)$ is independent of s, as is the case for the Wilcoxon and normal scores statistics. Then under H, the statistic $\Sigma h(s_j)$ is symmetrically distributed about $n\Sigma_{i=1}^N h(i)/N = \mu$, and (6.29) suggests the rejection region

$$\left|\sum h(s_j) - \mu\right| = \frac{1}{N}\left|m\sum_{j=1}^n h(s_j) - n\sum_{i=1}^m h(r_i)\right| > C.$$

The theory here is still less satisfactory than in the one-sided case. These tests need not even be unbiased [Sugiura (1965)], and it is not known whether they are admissible within the class of all rank tests. On the other hand, the relative asymptotic efficiencies are the same as in the one-sided case.

The two-sample hypothesis $G = F$ can also be tested against the general alternatives $G \neq F$. This problem arises in deciding whether two products, two sets of data, or the like can be pooled when nothing is known about the underlying distributions. Since the alternatives are now unrestricted, the problem remains invariant under all transformations $x_i' = f(x_i)$, $y_j' = f(y_j)$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, such that f has only a finite number of discontinuities. There are no invariants under this group, so that the only invariant test is $\phi(x, y) \equiv \alpha$. This is however not admissible, since there do exist tests of H that are strictly unbiased against all alternatives $G \neq F$ (Problem 6.56). One of the tests most commonly employed for this problem is the *Kalmogorov-Smirnov test*. Let the *empirical distribution functions* of the two samples be defined by

$$S_{x_1,\ldots,x_m}(z) = \frac{a}{m}, \qquad S_{y_1,\ldots,y_n}(z) = \frac{b}{n},$$

where a and b are the numbers of x's and y's less or equal to z respectively. Then H is rejected according to this test when

$$\sup_z \left| S_{x_1,\ldots,x_m}(z) - S_{y_1,\ldots,y_n}(z) \right| > C.$$

Accounts of the theory of this and related tests are given, for example, in Durbin (1973), Serfling (1980), Gibbons and Chakraborti (1992) and Hájek, Sidák, and Sen (1999).

Two-sample rank tests are distribution-free for testing $H : G = F$ but not for the nonparametric Behrens–Fisher situation of testing $H : \eta = \xi$ when the X's and Y's are samples from $F((x - \xi)/\sigma)$ and $F((y - \eta)/\tau)$ with σ, τ unknown. A detailed study of the effect of the difference in scales on the levels of the Wilcoxon and normal scores tests is provided by Pratt (1964).

6.10 The Hypothesis of Symmetry

When the method of paired comparisons is used to test the hypothesis of no treatment effect, the problem was seen in Section 6.8 to reduce through invariance to that of testing the hypothesis

$$H_2' : D(z) + D(-z) = 1 \text{ for all } z,$$

which states that the distribution D of the differences $Z_i = Y_i - X_i$ $(i = 1, \ldots, N)$ is symmetric with respect to the origin. The distribution D can be specified by the triple (ρ, F, G) where

$$\rho = P\{Z \leq 0\}, \qquad F(z) = P\{|Z| \leq z \mid Z > 0\},$$
$$G(z) = P\{Z \leq z \mid Z > 0\},$$

and the hypothesis of symmetry with respect to the origin then becomes

$$H : p = \tfrac{1}{2}, G = F.$$

Invariance and sufficiency were shown to reduce the data to the ranks $S_1 < \cdots < S_n$ of the positive Z's among the absolute values $|Z_1|, \ldots, |Z_N|$. The probability of $S_1 = s_1, \ldots, S_n = s_n$ is the probability of this event given that there are n positive observations multiplied by the probability that the number of positive observations is n. Hence

$$P\{S_1 = s_1, \ldots, S_n = s_n\}$$
$$= \binom{N}{n} (1 - \rho)^n \rho^{N-n} P_{F,G}\{S_1 = s_1, \ldots, S_n = s_n \mid n\},$$

where the second factor is given by (6.30). Under H, this becomes

$$P\{S_1 = s_1, \ldots, S_n = s_n\} = \frac{1}{2^N}$$

for each of the

$$\sum_{n=0}^{N} \binom{N}{n} = 2^N$$

n-tuples $(s_1, \ldots, s_n)$ satisfying $1 \leq s_1 < \cdots < s_n \leq N$. Any rank test of size $\alpha = k/2^N$ therefore has a rejection region containing exactly k such points $(s_1, \ldots, s_n)$.

The alternatives K of a beneficial treatment effect are characterized by the fact that the variable Z being sampled is stochastically larger than some random variable which is symmetrically distributed about 0. It is again suggestive to use rejection regions of the form $h(s_1) + \cdots + h(s_n) > C$, where however n is no longer a constant as it was in the two-sample problem, but depends on the observations. Two particular cases are the *Wilcoxon one-sample test*, which is obtained by putting $h(s) = s$, and the analogue of the normal scores test with $h(s) = E(W_{(s)})$ where $W_{(1)} < \cdots < W_{(N)}$ are the ordered values of $|V_1|, \ldots, |V_N|$, the V's being a sample from $N(0, 1)$. The W's are therefore an ordered sample of size N from a distribution with density $\sqrt{2/\pi} e^{-w^2/2}$ for $w \geq 0$.

As in the two-sample problem, it can be shown that each of these tests is most powerful (among all invariant tests) against certain alternatives, and that they are both unbiased against the class K. Their asymptotic efficiencies relative to the t-test for testing that the mean of Z is zero have the same values $3/\pi$ and 1 as the corresponding two-sample tests, when the distribution of Z is normal.

In certain applications, for example, when the various comparisons are made under different experimental conditions or by different methods, it may be unrealistic to assume that the variables $Z_1, \ldots, Z_N$ have a common distribution. Suppose instead that the Z_i are still independently distributed but with arbitrary continuous distributions D_i. The hypothesis to be tested is that each of these distributions is symmetric with respect to the origin.

This problem remains invariant under all transformations $z_i' = f_i(z_i)$ $i = 1, \ldots, N$, such that each f_i is continuous, odd, and strictly increasing. A maximal invariant is then the number n of positive observations, and it follows from Example 6.5.1 that there exists a UMP invariant test, the *sign test*, which rejects when n is too large. This test reflects the fact that the magnitude of the observations or of their absolute values can be explained entirely in terms of the spread of the distributions D_i, so that only the signs of the Z's are relevant.

Frequently, it seems reasonable to assume that the Z's are identically distributed, but the assumption cannot be trusted. One would then prefer to use the information provided by the ranks s_i but require a test which controls the probability of false rejection even when the assumption fails. As is shown by the following lemma, this requirement is in fact satisfied for every (symmetric) rank test. Actually, the lemma will not require even the independence of the Z's; it will show that any symmetric rank test continues to correspond to the stated level of significance provided only the treatment is assigned at random within each pair.

Lemma 6.10.1 *Let $\phi(z_1, \dots, z_N)$ be symmetric in its N variables and such that*

$$E_D\phi(Z_1, \dots, Z_N) = \alpha \tag{6.32}$$

when the Z's are a sample from any continuous distribution D which is symmetric with respect to the origin. Then

$$E\phi(Z_1, \dots, Z_N) = \alpha \tag{6.33}$$

if the joint distribution of the Z's is unchanged under the 2^N transformations $Z_1' = \pm Z_1, \dots, Z_N' = \pm Z_N$.

PROOF. Condition (6.32) implies

$$\sum_{(j_1,\dots,j_N)} \sum \frac{\phi(\pm z_{j_1}, \dots, \pm z_{j_N})}{2^N \cdot N!} = \alpha \quad \text{a.e.,} \tag{6.34}$$

where the outer summation extends over all $N!$ permutations $(j_1, \dots, j_N)$ and the inner one over all 2^N possible choices of the signs $+$ and $-$. This is proved exactly as was Theorem 5.8.1. If in addition ϕ is symmetric, (6.34) implies

$$\sum \frac{\phi(\pm z_1, \dots, \pm z_N)}{2^N} = \alpha. \tag{6.35}$$

Suppose that the distribution of the Z's is invariant under the 2^N transformations in question. Then the conditional probability of any sign combination of $Z_1, \dots, Z_N$ given $|Z_1|, \dots, |Z_N|$ is $1/2^N$. Hence (6.35) is equivalent to

$$E[\phi(Z_1, \dots, Z_N) \mid |Z_1|, \dots, |Z_N|] = \alpha \quad \text{a.e.,} \tag{6.36}$$

and this implies (6.33) which was to be proved. ■

The tests discussed above can be used to test symmetry about any known value θ_0 by applying them to the variables $Z_i - \theta_0$. The more difficult problem of testing for symmetry about an unknown point θ will not be considered here. Tests of this hypothesis are discussed, among others, by Antille et al. (1982), Bhattacharya et al. (1982), Boos (1982), and Koziol (1983).

As will be seen in Section 13.2.1, the one-sample t-test is not robust against dependence. Unfortunately, this is also true-although to a somewhat lesser extent—of the sign and one-sample Wilcoxon tests [Gastwirth and Rubin (1971)].

6.11 Equivariant Confidence Sets

Confidence sets for a parameter θ in the presence of nuisance parameters ϑ were discussed in Chapter 5 (Sections 5.4 and 5.5) under the assumption that θ is real-valued. The correspondence between acceptance regions $A(\theta_0)$ of the hypotheses $H(\theta_0) : \theta = \theta_0$ and confidence sets $S(x)$ for θ given by (5.33) and (5.34) is, however, independent of this assumption; it is valid regardless of whether θ is real-valued, vector-valued, or possibly a label for a completely unknown distribution function (in the latter case, confidence intervals become confidence bands for the distribution function). This duality, which can be summarized by the relationship

$$\theta \in S(x) \quad \text{if and only if} \quad x \in A(\theta), \tag{6.37}$$

was the basis for deriving uniformly most accurate and uniformly most accurate unbiased confidence sets. In the present section, it will be used to obtain uniformly most accurate equivariant confidence sets.

We begin by defining equivariance for confidence sets. Let G be a group of transformations of the variable X preserving the family of distributions $\{P_{\theta,\vartheta}, (\theta, \vartheta) \in \Omega\}$ and let $\bar{G}$ be the induced group of transformations of Ω. If $\bar{g}(\theta, \vartheta) = (\theta', \vartheta')$, we shall suppose that θ' depends only on $\bar{g}$ and θ and not on ϑ, so that $\bar{g}$ induces a transformation in the space of θ. In order to keep the notation from becoming unnecessarily complex, it will then be convenient to write also $\theta' = \bar{g}\theta$. For each transformation $g \in G$, denote by g^* the transformation acting on sets S in θ-space and defined by

$$g^* S = \{\bar{g}\theta : \theta \in S\}, \tag{6.38}$$

so that $g^* S$ is the set obtained by applying the transformation $\bar{g}$ to each point θ of S. The invariance argument of Section 1.5 then suggests restricting consideration to confidence sets satisfying

$$g^* S(x) = S(gx) \qquad \text{for all} \quad x \in \mathcal{X}, \quad g \in G. \tag{6.39}$$

We shall say that such confidence sets are *equivariant* under G. This terminology is preferable to the older term invariance which creates the impression that the confidence sets remain unchanged under the transformation $X' = gX$. If the transformation g is interpreted as a change of coordinates, (6.39) means that the confidence statement does not depend on the coordinate system used to express the data. The statement that the transformed parameter $\bar{g}\theta$ lies in $S(gx)$ is equivalent to stating that $\theta \in g^{*-1} S(gx)$, which is equivalent to the original statement $\theta \in S(x)$ provided (6.39) holds.

Example 6.11.1 Let X, Y be independently normally distributed with means ξ, η and unit variance, and let G be the group of all rigid motions of the plane, which is generated by all translations and orthogonal transformations. Here $\bar{g} = g$ for all $g \in G$. An example of an equivariant class of confidence sets is given by

$$S(x, y) = \left\{(\xi, \eta) : (x - \xi)^2 + (y - \eta)^2 \le C\right\},$$

the class of circles with radius $\sqrt{C}$ and center (x, y). The set $g^*S(x, y)$ is the set of all points $g(\xi, \eta)$ with $(\xi, \eta) \in S(x, y)$ and hence is obtained by subjecting $S(x, y)$ to the rigid motion g. The result is the circle with radius $\sqrt{C}$ and center $g(x, y)$, and (6.39) is therefore satisfied. ■

In accordance with the definitions given in Chapters 3 and 5, a class of confidence sets for θ will be said to be *uniformly most accurate equivariant* at confidence level $1 - \alpha$ if among all equivariant classes of sets $S(x)$ at that level it minimizes the probability

$$P_{\theta,\vartheta}\{\theta' \in S(X)\} \qquad \text{for all} \quad \theta' \neq \theta.$$

In order to derive confidence sets with this property from families of UMP invariant tests, we shall now investigate the relationship between equivariance of confidence sets and invariance of the associated tests.

Suppose that for each θ_0 there exists a group of transformations G_{θ_0} which leaves invariant the problem of testing $H(\theta_0) : \theta = \theta_0$, and denote by G the group of transformations generated by the totality of groups G_θ.

Lemma 6.11.1 (i) *Let $S(x)$ be any class of confidence sets that is equivariant under G, and let $A(\theta) = \{x : \theta \in S(x)\}$; then the acceptance region $A(\theta)$ is invariant under G_θ for each θ.*

(ii) *If in addition, for each θ_0 the acceptance region $A(\theta_0)$ is UMP invariant for testing $H(\theta_0)$ at level α, the class of confidence sets $S(x)$ is uniformly most accurate among all equivariant confidence sets at confidence level $1 - \alpha$.*

PROOF. (i): Consider any fixed θ, and let $g \in G_\theta$. Then

$$\begin{aligned} gA(\theta) &= \{gx : \theta \in S(x)\} = \{x : \theta \in S(g^{-1}x)\} = \{x : \theta \in g^{*-1}S(x)\} \\ &= \{x : \bar{g}\theta \in S(x)\} = \{x : \theta \in S(x)\} = A(\theta). \end{aligned}$$

Here the third equality holds because $S(x)$ is equivariant, and the fifth one because $g \in G_\theta$ and therefore $\bar{g}\theta = \theta$.
(ii): If $S'(x)$ is any other equivariant class of confidence sets at the prescribed level, the associated acceptance regions $A'(\theta)$ by (i) define invariant tests of the hypotheses $H(\theta)$. It follows that these tests are uniformly at most as powerful as those with acceptance regions $A(\theta)$ and hence that

$$P_{\theta,\vartheta}\{\theta' \in S(X)\} \le P_{\theta,\vartheta}\{\theta' \in S'(X)\} \qquad \text{for all} \quad \theta' \neq \theta$$

as was to be proved. ■

It is an immediate consequence of the lemma that if UMP invariant acceptance regions $A(\theta)$ have been found for each hypothesis $H(\theta)$ (invariant with respect to G_θ), and if the confidence sets $S(x) = \{\theta : x \in A(\theta)\}$ are equivariant under G, then they are uniformly most accurate equivariant.

Example 6.11.2 Under the assumptions of Example 6.11.1, the problem of testing $\xi = \xi_0$, $\eta = \eta_0$ is invariant under the group G_{ξ_0,η_0} of orthogonal transformations about the point (ξ_0, η_0):

$$X' - \xi_0 = a_{11}(X - \xi_0) + a_{12}(Y - \eta_0),$$
$$Y' - \eta_0 = a_{21}(X - \xi_0) + a_{22}(Y - \eta_0),$$

where the matrix (a_{ij}) is orthogonal. There exists under this group a UMP invariant test, which has acceptance region (Problem 7.8)

$$(X - \xi_0)^2 + (Y - \eta_0)^2 \le C.$$

Let G_0 be the smallest group containing the groups $G_{\xi,\eta}$, for all ξ, η. Since this is a subgroup of the group G of Example 6.11.1 (the two groups actually coincide, but this is immaterial for the argument), the confidence sets $(X - \xi)^2 + (Y - \eta)^2 \le C$ are equivariant under G_0 and hence uniformly most accurate equivariant. ■

Example 6.11.3 Let $X_1, \dots, X_n$ be independently normally distributed with mean ξ and variance σ^2. Confidence intervals for ξ are based on the hypotheses $H(\xi_0)$: $\xi = \xi_0$, which are invariant under the groups G_{ξ_0} of transformations $X_i' = a(X_i - \xi_0) + \xi_0$ $(a \ne 0)$. The UMP invariant test of $H(\xi_0)$ has acceptance region

$$\frac{\sqrt{(n-1)n}|\bar{X} - \xi_0|}{\sqrt{\sum(X_i - \bar{X})^2}} \le C,$$

and the associated confidence intervals are

$$\bar{X} - \frac{C}{\sqrt{n(n-1)}}\sqrt{\sum(X_i - \bar{X})^2} \le \xi \le \bar{X} + \frac{C}{\sqrt{n(n-1)}}\sqrt{\sum(X_i - \bar{X})^2}. \quad (6.40)$$

The group G in the present case consists of all transformations $g : X_i' = aX_i + b$ $(a \ne 0)$, which on ξ induces the transformation $\bar{g} : \xi' = a\xi + b$. Application of the associated transformation g^* to the interval (6.40) takes it into the set of points $a\xi + b$ for which ξ satisfies (6.40), that is, into the interval with end points

$$a\bar{X} + b - \frac{|a|C}{\sqrt{n(n-1)}}\sqrt{\sum(X_i - \bar{X})^2}, \qquad a\bar{X} + b + \frac{|a|C}{\sqrt{n(n-1)}}\sqrt{\sum(X_i - \bar{X})^2}.$$

Since this coincides with the interval obtained by replacing X_i in (6.40) with $aX_i + b$, the confidence intervals (6.40) are equivariant under G_0 and hence uniformly most accurate equivariant. ■

Example 6.11.4 In the two-sample problem of Section 6.9, assume the shift model in which the X's and Y's have densities $f(x)$ and $g(y) = f(y - \Delta)$, respectively, and consider the problem of obtaining confidence intervals for the shift parameter Δ

which are distribution-free in the sense that the coverage probability is independent of the true f. The hypothesis $H(\Delta_0) : \Delta = \Delta_0$ can be tested, for example, by means of the Wilcoxon test applied to the observations X_i, $Y_j - \Delta_0$, and confidence sets for Δ can then be obtained by the usual inversion process. The resulting confidence intervals are of the form $D_{(k)} < \Delta < D_{(mn+1-k)}$ where $D_{(1)} < \cdots < D_{(mn)}$ are the mn ordered differences $Y_j - X_i$. [For details see Problem 6.54 and for fuller accounts nonparametric books such as Randles and Wolfe (1979), Gibbons and Chakraborti (1992) and Lehmann (1998).] By their construction, these intervals have coverage probability $1 - \alpha$, which is independent of f. However, the invariance considerations of Sections 6.8 and 6.9 do not apply. The hypothesis $H(\Delta_0)$ is invariant under the transformations $X_i' = \rho(X_i)$, $Y_j' = \rho(Y_j - \Delta_0) + \Delta_0$ with ρ continuous and strictly increasing, but the shift model, and hence the problem under consideration, is not invariant under these transformations. ■

6.12 Average Smallest Equivariant Confidence Sets

In the examples considered so far, the invariance and equivariance properties of the confidence sets corresponded to invariant properties of the associated tests. In the following examples this is no longer the case.

Example 6.12.1 Let $X_1, \ldots, X_n$, be a sample from $N(\xi, \sigma^2)$, and consider the problem of estimating σ^2.

The model is invariant under translations $X_i' = X_i + a$, and sufficiency and invariance reduce the data to $S^2 = \sum(X_i - \bar{X})^2$. The problem of estimating σ^2 by confidence sets also remains invariant under scale changes $X_i' = bX_i$, $S' = bS$, $\sigma' = b\sigma$ $(0 < b)$, although these do not leave the corresponding problem of testing the hypothesis $\sigma = \sigma_0$ invariant. (Instead, they leave invariant the *family* of these testing problems, in the sense that they transform one such hypothesis into another.) The totality of equivariant confidence sets based on S is given by

$$\frac{\sigma^2}{S^2} \in A, \tag{6.41}$$

where A is any fixed set on the line satisfying

$$P_{\sigma=1}\left(\frac{1}{S^2} \in A\right) = 1 - \alpha. \tag{6.42}$$

That any set $\sigma^2 \in S^2 \cdot A$ is equivariant is obvious. Conversely, suppose that $\sigma^2 \in C(S^2)$ is an equivariant family of confidence sets for σ^2. Then $C(S^2)$ must satisfy $b^2 C(S^2) = C(b^2 S^2)$ and hence

$$\sigma^2 \in C(S^2) \quad \text{if and only if} \quad \frac{\sigma^2}{S^2} \in \frac{1}{S^2} C(S^2) = C(1),$$

which establishes (6.41) with $A = C(1)$.

Among the confidence sets (6.41) with A satisfying (6.42) there does not exist one that uniformly minimizes the probability of covering false values (Problem 6.75). Consider instead the problem of determining the confidence sets that are physically smallest in the sense of having minimum Lebesgue measure. This requires minimizing $\int_A dv$ subject to (6.42). It follows from the Neyman–Pearson Lemma that the minimizing A^* is

$$A^* = \{v : p(v) > C\}, \tag{6.43}$$

where $p(v)$ is the density of $V = 1/S^2$ when $\sigma = 1$, and where C is determined by (6.42). Since $p(v)$ is unimodal (Problem 6.76), these smallest confidence sets are intervals, $aS^2 < \sigma^2 < bS^2$. Values of a and b are tabled by Tate and Klett (1959), who also table the corresponding (different) values a', b' for the uniformly most accurate unbiased confidence intervals $a'S^2 < \sigma^2 < b'S^2$ (given in Example 5.5.1).

Instead of minimizing the Lebesgue measure $\int_A dv$ of the confidence sets A, one may prefer to minimize the scale-invariant measure

$$\int_A \frac{1}{v}\, dv. \tag{6.44}$$

To an interval (a, b), (6.44) assigns, in place of its length $b - a$, its logarithmic length $\log b - \log a = \log(b/a)$. The optimum solution A^{**} with respect to this new measure is again obtained by applying the Neyman–Pearson Lemma, and is given by

$$A^{**} = \{v : vp(v) > C\}, \tag{6.45}$$

which coincides with the uniformly most accurate unbiased confidence sets (Problem 6.77(i)).

One advantage of minimizing (6.44) instead of Lebesgue measure is that it then does not matter whether one estimates σ or σ^2 (or σ^r for some other power of r), since under (6.44), if (a, b) is the best interval for σ, then (a^r, b^r) is the best interval for σ^r (Problem 6.77(ii)). ■

Example 6.12.2 Let X_i $(i = 1, \ldots, r)$ be independently normally distributed as $N(\xi, 1)$. A slight generalization of Example 6.11.2 shows that uniformly most accurate equivariant confidence sets for $(\xi_1, \ldots, \xi_r)$ exist with respect to the group G of all rigid transformations and are given by

$$\sum (X_i - \xi_i)^2 \le C. \tag{6.46}$$

Suppose that the context of the problem does not possess the symmetry which would justify invoking invariance with respect to G, but does allow the weaker assumption of invariance under the group G_0 of translations $X'_i = X_i + a_i$. The totality of equivariant confidence sets with respect to G_0 is given by

$$(X_1 - \xi_1, \ldots, X_r - \xi_r) \in A, \tag{6.47}$$

where A is any fixed set in r-space satisfying

$$P_{\xi_1=\cdots=\xi_r=0}((X_1, \ldots, X_r) \in A) = 1 - \alpha. \tag{6.48}$$

Since uniformly most accurate equivariant confidence sets do not exist (Problem 6.75), let us consider instead the problem of determining the confidence sets of smallest Lebesgue measure. (This measure is invariant under G_0.) This is given by (6.43) with $v = (v_1, \ldots, v_r)$ and $p(v)$ the density of $(X_1, \ldots, X_r)$ when $\xi_1 = \cdots = \xi_r = 0$, and hence coincides with (6.46).

Quite surprisingly, the confidence sets (6.46) are inadmissible if and only if $r \geq 3$. A further discussion of this fact and references are deferred to Example 8.5.4. ■

Example 6.12.3 In the preceding example, suppose that the X_i are distributed as $N(\xi_i, \sigma^2)$ with σ^2 unknown, and that a variable S^2 is available for estimating σ^2. Assume S^2 is independent of the X's and that S^2/σ^2 has a χ^2 -distribution with f degrees of freedom.

The estimation of $(\xi_1, \ldots, \xi_r)$ by confidence sets on the basis of X's and S^2 remains invariant under the group G_0 of transformations

$$X'_i = bX_i + a_i, \qquad S' = bS, \qquad \xi'_i = b\xi_i + a_i, \quad \sigma' = b\sigma,$$

and the most general equivariant confidence set is of the form

$$\left(\frac{X_1 - \xi_1}{S}, \ldots, \frac{X_r - \xi_r}{S}\right) \in A, \tag{6.49}$$

where A is any fixed set in r-space satisfying

$$P_{\xi_1=\cdots=\xi_r=0}\left[\left(\frac{X_1}{S}, \ldots, \frac{X_r}{S}\right) \in A\right] = 1 - \alpha. \tag{6.50}$$

The confidence sets (6.49) can be written as

$$(\xi_1, \ldots, \xi_r) \in (X_1, \ldots, X_r) - SA, \tag{6.51}$$

where $-SA$ is the set obtained by multiplying each point of A by the scalar $-S$.

To see (6.51), suppose that $C(X_1, \ldots, X_r; S)$ is an equivariant confidence set for $(\xi_1, \ldots, \xi_r)$. Then the r-dimensional set C must satisfy

$$C(bX_1 + a_1, \ldots, bX_r + a_r; bS) = b[C(X_1, \ldots, X_r; S)] + (a_1, \ldots, a_r)$$

for all $a_1, \dots, a_r$ and all $b > 0$. It follows that $(\xi_1, \dots, \xi_r) \in C$ if and only if

$$\left(\frac{X_1 - \xi_1}{S}, \dots, \frac{X_r - \xi_r}{S}\right) \in \frac{(X_1, \dots, X_r) - C(X_1, \dots, X_r; S)}{S} = C(0, \dots, 0; 1) = A.$$

The equivariant confidence sets of smallest volume are obtained by choosing for A the set A^* given by (6.43) with $v = (v_1, \dots, v_r)$ and $p(v)$ the joint density of $(X_1/S, \dots, X_r/S)$ when $\xi_1 = \cdots = \xi_r = 0$. This density is a decreasing function of $\sum v_i^2$ (Problem 6.78), and the smallest equivariant confidence sets are therefore given by

$$\sum (X_i - \xi_i)^2 \le CS^2. \tag{6.52}$$

[Under the larger group G generated by all rigid transformations of $(X_1, \dots, X_r)$ together with the scale changes $X_i' = bX_i$, $S' = bS$, the same sets have the stronger property of being uniformly most accurate equivariant; see Problem 6.79.] ■

Examples 6.12.1–6.12.3 have the common feature that the equivariant confidence sets $S(X)$ for $\theta = (\theta_1, \dots, \theta_r)$ are characterized by an r-valued *pivotal quantity*, that is, a function $h(X, \theta) = (h_1(X, \theta), \dots, h_r(X, \theta))$ of the observations X and parameters θ being estimated that has a fixed distribution, and such that the most general equivariant confidence sets are of the form

$$h(X, \theta) \in A \tag{6.53}$$

for some fixed set A.[6] When the functions h_i are linear in θ, the confidence sets $C(X)$ obtained by solving (6.53) for θ are linear transforms of A (with random coefficients), so that the volume or invariant measure of $C(X)$ is minimized by minimizing

$$\int_A \rho(v_1, \dots, v_r)\, dv_1 \dots dv_r \tag{6.54}$$

for the appropriate ρ. The problem thus reduces to that of minimizing (6.54) subject to

$$P_{\theta_0}\{h(X, \theta_0) \in A\} = \int_A p(v_1, \dots, v_r)\, dv_1 \dots dv_r = 1 - \alpha, \tag{6.55}$$

where $p(v_1, \dots, v_r)$ is the density of the pivotal quantity $h(X, \theta)$. The minimizing A is given by

$$A^* = \left\{ v : \frac{p(v_1, \dots, v_r)}{\rho(v_1, \dots, v_r)} > C \right\}, \tag{6.56}$$

with C determined by (6.55).

[6] More general results concerning the relationship of equivariant confidence sets and pivotal quantities are given in Problems 6.71–6.74.

The following is one more illustration of this approach.

Example 6.12.4 Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be samples from $N(\xi, \sigma^2)$ and $N(\eta, \tau^2)$ respectively, and consider the problem of estimating $\Delta = \tau^2/\sigma^2$. Sufficiency and invariance under translations $X_i' = X_i + a_1$, $Y_j' = Y_j + a_2$ reduce the data to $S_X^2 = \sum(X_i, -\bar{X})^2$ and $S_Y^2 = \sum(Y_j - \bar{Y})^2$. The problem of estimating Δ also remains invariant under the scale changes

$$X_i' = b_1 X_i, \quad Y_j' = b_2 Y_j, \qquad 0 < b_1, b_2 < \infty,$$

which induce the transformations

$$S_X' = b_1 S_X, \qquad S_Y' = b_2 S_Y, \qquad \sigma' = b_1 \sigma, \qquad \tau' = b_2 \tau. \tag{6.57}$$

The totality of equivariant confidence sets for Δ is given by $\Delta/V \in A$, where $V = S_Y^2/S_X^2$ and A is any fixed set on the line satisfying

$$P_{\Delta=1}\left(\frac{1}{V} \in A\right) = 1 - \alpha. \tag{6.58}$$

To see this, suppose that $C(S_X, S_Y)$ are any equivariant confidence sets for Δ. Then C must satisfy

$$C(b_1 S_X, b_2 S_Y) = \frac{b_2^2}{b_1^2} C(S_X, S_Y), \tag{6.59}$$

and hence $\Delta \in C(S_X, S_Y)$ if and only if the pivotal quantity V/Δ satisfies

$$\frac{\Delta}{V} = \frac{S_X^2 \Delta}{S_Y^2} \in \frac{S_X^2}{S_Y^2} C(S_X, S_Y) = C(1, 1) = A.$$

As in Example 6.12.1, one may now wish to choose A so as to minimize either its Lebesgue measure $\int_A dv$ or the invariant measure $\int_A (1/v)\, dv$. The resulting confidence sets are of the form

$$p(v) > C \quad \text{and} \quad vp(v) > C, \tag{6.60}$$

respectively. In both cases, they are intervals $V/b < \Delta < V/a$ (Problem 6.80(i)). The values of a and b minimizing Lebesgue measure are tabled by Levy and Narula (1974); those for the invariant measure coincide with the uniformly most accurate unbiased intervals (Problem 6.80(ii)). ■

6.13 Confidence Bands for a Distribution Function

Suppose that $X = (X_1, \dots, X_n)$ is a sample from an unknown continuous cumulative distribution function F, and that lower and upper bounds L_X and M_X are to be determined such that with preassigned probability $1 - \alpha$ the inequalities

$$L_X(u) \le F(u) \le M_X(u) \quad \text{for all } u$$

hold for all continuous cumulative distribution functions F. This problem is invariant under the group G of transformations

$$X_i' = g(X_i), \quad i = 1, \dots, n,$$

where g is any continuous strictly increasing function. The induced transformation in the parameter space is $\bar{g}F = F(g^{-1})$.

If $S(x)$ is the set of continuous cumulative distribution functions

$$S(x) = \{F : L_x(u) \le F(u) \le M_x(u) \text{ for all } u\},$$

then

$$\begin{aligned} g^* S(x) &= \{\bar{g}F : L_x(u) \le F(u) \le M_x(u) \text{ for all } u\} \\ &= \{F : L_x[g^{-1}(u)] \le F(u) \le M_x[g^{-1}(u)] \text{ for all } u\}. \end{aligned}$$

For an equivariant procedure, this must coincide with the set

$$S(gx) = \left\{F : L_{g(x_1),\dots,g(x_n)}(u) \le F(u) \le M_{g(x_1),\dots,g(x_n)}(u) \text{ for all } u\right\}.$$

The condition of equivariance is therefore

$$\begin{aligned} L_{g(x_1),\dots,g(x_n)}[g(u)] &= L_x(u), \\ M_{g(x_1),\dots,g(x_n)}[g(u)] &= M_x(u) \quad \text{for all } x \text{ and } u. \end{aligned}$$

To characterize the totality of equivariant procedures, consider the *empirical distribution function* (EDF) T_x given by

$$T_x(u) = \frac{i}{n} \quad \text{for} \quad x_{(i)} \le u < x_{(i+1)}, \quad i = 0, \dots, n,$$

where $x_{(1)} < \cdots < x_{(n)}$ is the ordered sample and where $x_{(0)} = -\infty$, $x_{(n+1)} = \infty$. Then a necessary and sufficient condition for L and M to satisfy the above equivariance condition is the existence of numbers $a_0, \dots, a_n$; $a_0', \dots, a_n'$ such that

$$L_x(u) = a_i, \quad M_x(u) = a_i' \quad \text{for} \quad x_{(i)} < u < x_{(i+1)}.$$

That this condition is sufficient is immediate. To see that it is also necessary, let u, u' be any two points satisfying $x_{(i)} < u < u' < x_{(i+1)}$. Given any $y_1, \ldots, y_n$ and v with $y_{(i)} < v < y_{(i+1)}$, there exist $g, g' \in G$ such that

$$g(y_{(i)}) = g'(y_{(i)}) = x_{(i)}, \quad g(v) = u, \quad g'(v) = u'.$$

If L_x, M_x are equivariant, it then follows that $L_x(u') = L_y(v)$ and $L_x(u) = L_y(v)$, and hence that $L_x(u') = L_x(u)$ and similarly $M_x(u') = M_x(u)$, as was to be proved. This characterization shows L_x and M_x to be step functions whose discontinuity points are restricted to those of T_x.

Since any two continuous strictly increasing cumulative distribution functions can be transformed into one another through a transformation $\bar{g}$, it follows that all these distributions have the same probability of being covered by an equivariant confidence band. (See Problem 6.86.) Suppose now that F is continuous but no longer strictly increasing. If I is any interval of constancy of F, there are no observations in I, so that I is also an interval of constancy of the sample cumulative distribution function. It follows that the probability of the confidence band covering F is not affected by the presence of I and hence is the same for all continuous cumulative distribution functions F.

For any numbers a_i, a'_i let Δ_i, Δ'_i be determined by

$$a_i = \frac{i}{n} - \Delta_i, \qquad a'_i = \frac{i}{n} - \Delta'_i.$$

Then it was seen above that any numbers $\Delta_0, \ldots, \Delta_n$; $\Delta'_0, \ldots, \Delta'_n$ define a confidence band for F, which is equivariant and hence has constant probability of covering the true F. From these confidence bands a test can be obtained of the hypothesis of *goodness of fit* $F = F_0$ that the unknown F equals a hypothetical distribution F_0. The hypothesis is accepted if F_0 ties entirely within the band, that is, if

$$-\Delta_i < F_0(u) - T_x(u) < \Delta'_i$$
$$\text{for all} \quad x_{(i)} < u < x_{(i+1)} \quad \text{and all} \quad i = 1, \ldots, n.$$

Within this class of tests there exists no UMP member, and the most common choice of the Δ's is $\Delta_i = \Delta'_i = \Delta$ for all i. The acceptance region of the resulting *Kolmogorov–Smirnov test* can be written as

$$\sup_{-\infty<u<\infty} |F_0(u) - T_x(u)| < \Delta. \tag{6.61}$$

Tables of the null distribution of the Kolmogorov–Smirnov statistic are given by Birnbaum (1952). For large n, approximate critical values can be obtained from the limit distribution K of $\sqrt{n} \sup |F_0(u) - T_x(u)|$, due to Kolmogorov and tabled by Smirnov (1948). Derivations of K can be found, for example, in Feller (1948), Billingsley (1968), and Hájek et al. (1999). The large-sample properties of this test will be studied in Example 11.4.2 and Section 16.2. The more general problem of testing goodness of fit will be presented in Chapter 16.

6.14 Problems

Section 6.1

Problem 6.1 Let G be a group of measurable transformations of $(\mathcal{X}, \mathcal{A})$ leaving $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ invariant, and let $T(x)$ be a measurable transformation to $(\mathcal{T}, \mathcal{B})$. Suppose that $T(x_1) = T(x_2)$ implies $T(gx_1) = T(gx_2)$ for all $g \in G$, so that G induces a group G^* on $\mathcal{T}$ through $g^*T(x) = T(gx)$, and suppose further that the induced transformations g^* are measurable $\mathcal{B}$. Then G^* leaves the family $\mathcal{P}^T = \{P_\theta^T, \theta \in \Omega\}$ of distributions of T invariant.

Section 6.2

Problem 6.2 (i) Let $\mathcal{X}$ be the totality of points $x = (x_1, \ldots, x_n)$ for which all coordinates are different from zero, and let G be the group of transformations $x_i' = cx_i$, $c > 0$. Then a maximal invariant under G is $(\text{sgn } x_n, x_1/x_n, \ldots, x_{n-1}/x_n)$ where sgn x is 1 or -1 as x is positive or negative.

(ii) Let $\mathcal{X}$ be the space of points $x = (x_1, \ldots, x_n)$ for which all coordinates are distinct, and let G be the group of all transformations $x_i' = f(x_i)$, $i = 1, \ldots, n$, such that f is a $1:1$ transformation of the real line onto itself with at most a finite number of discontinuities. Then G is transitive over $\mathcal{X}$.

[(ii): Let $x = (x_1, \ldots, x_n)$ and $x' = (x_1', \ldots, x_n')$ be any two points of $\mathcal{X}$. Let $I_1, \ldots, I_n$ be a set of mutually exclusive open intervals which (together with their end points) cover the real line and such that $x_j \in I_j$. Let $I_1', \ldots, I_n'$ be a corresponding set of intervals for $x_1', \ldots, x_n'$. Then there exists a transformation f which maps each I_j continuously onto I_j', maps x_j into x_j', and maps the set of $n-1$ end points of $I_1, \ldots, I_n$ onto the set of end points of $I_1', \ldots, I_n'$.]

Problem 6.3 Suppose M is any $m \times p$ matrix. Show that $M^\top M$ is positive semidefinite. Also, show the rank of $M^\top M$ equals the rank of M, so that in particular $M^\top M$ is nonsingular if and only if $m \geq p$ and M is of rank p.

Problem 6.4 (i) A sufficient condition for (6.9) to hold is that D is a normal subgroup of G.

(ii) If G is the group of transformations $x' = ax + b$, $a \neq 0$, $-\infty < b < \infty$, then the subgroup of translations $x' = x + b$ is normal but the subgroup $x' = ax$ is not.

[The defining property of a normal subgroup is that given $d \in D$, $g \in G$, there exists $d' \in D$ such that $gd = d'g$. The equality $s(x_1) = s(x_2)$ implies $x_2 = dx_1$ for some $d \in D$, and hence $ex_2 = edx_1 = d'ex_1$. The result (i) now follows, since s is invariant under D.]

Section 6.3

Problem 6.5 Prove statements (i)-(iii) of Example 6.3.1.

Problem 6.6 Prove Theorem 6.3.1
(i) by analogy with Example 6.3.1, and
(ii) by the method of Example 6.3.2. [*Hint:* A maximal invariant under G is the set $\{g_1 x, \dots, g_N x\}$.]

Problem 6.7 Consider the situation of Example 6.3.1 with $n = 1$, and suppose that f is strictly increasing on (0, 1).
(i) The likelihood ratio test rejects if $X < \alpha/2$ or $X > 1 - \alpha/2$.
(ii) The MP invariant test agrees with the likelihood ratio test when f is convex.
(iii) When f is concave, the MP invariant test rejects when

$$\frac{1}{2} - \frac{\alpha}{2} < X < \frac{1}{2} + \frac{\alpha}{2}\,,$$

and the likelihood ratio test is the least powerful invariant test against both alternatives and has power $\leq \alpha$. When does the power $= \alpha$?

Problem 6.8 In Example 6.1.1, find a maximal invariant and the UMPI level α test.

Problem 6.9 Let X, Y have the joint probability density $f(x, y)$. Then the integral $h(z) = \int_{-\infty}^{\infty} f(y - z, y)dy$ is finite for almost all z, and is the probability density of $Z = Y - X$.
[Since $P\{Z \leq b\} = \int_{-\infty}^{b} h(z)dz$, it is finite and hence h is finite almost everywhere.]

Problem 6.10 (i) Let $X = (X_1, \dots, X_n)$ have probability density $(1/\theta^n) f[(x_1 - \xi)/\theta, \dots, (x_n - \xi)/\theta]$, where $-\infty < \xi < \infty$, $0 < \theta$ are unknown, and where f is even. The problem of testing $f = f_0$ against $f = f_1$ remains invariant under the transformations $x_i' = ax_i + b$ $(i = 1, \dots, n)$, $a \neq 0$, $-\infty < b < \infty$ and the most powerful invariant test is given by the rejection region

$$\int_{-\infty}^{\infty} \int_0^{\infty} v^{n-2} f_1(vx_1 + u, \dots, vx_n + u)\, dv\, du$$
$$> C \int_{-\infty}^{\infty} \int_0^{\infty} v^{n-2} f_0(vx_1 + u, \dots, vx_n + u)\, dv\, du.$$

(ii) Let $X = (X_1, \dots, X_n)$ have probability density $f(x_1 - \sum_{j=1}^k w_{1j}\beta_j, \dots, x_n - \sum_{j=1}^k w_{nj}\beta_j)$ where $k < n$, the w's are given constants, the matrix (w_{ij}) is of rank k, the β's are unknown, and we wish to test $f = f_0$ against $f = f_1$. The problem remains invariant under the transformations $x_i' = x_i + \Sigma_{j=1}^k w_{ij}\gamma_j, -\infty < \gamma_1, \dots, \gamma_k < \infty$, and the most powerful invariant test is given by the rejection region

$$\frac{\int \cdots \int f_1(x_1 - \sum w_{1j}\beta_j, \dots, x_n - \sum w_{nj}\beta_j) d\beta_1, \dots, d\beta_k}{\int \cdots \int f_0(x_1 - \sum w_{1j}\beta_j, \dots, x_n - \sum w_{nj}\beta_j) d\beta_1, \dots, d\beta_k} > C.$$

[A maximal invariant is given by $y =$

$$\left(x_1 - \sum_{r=n-k+1}^{n} a_{1_r} x_r,\ x_2 - \sum_{r=n-k+1}^{n} a_{2_r} x_r, \dots, x_{n-k} - \sum_{r=n-k+1}^{n} a_{n-k,r} x_r \right)$$

for suitably chosen constants a_{ir}.]

Problem 6.11 Let $X_1, \dots, X_m; Y_1, \dots, Y_n$ be samples from exponential distributions with densities for $\sigma^{-1}e^{-(x-\xi)/\sigma}$, for $x \ge \xi$, and $\tau^{-1}e^{-(y-\eta)/\tau}$ for $y \ge \eta$.

(i) For testing $\tau/\sigma \le \Delta$ against $\tau/\sigma > \Delta$, there exists a UMP invariant test with respect to the group $G: X_i' = aX_i + b, Y_j' = aY_j + c, a > 0, -\infty < b, c < \infty$, and its rejection region is

$$\frac{\sum[y_j - \min(y_1, \dots, y_n)]}{\sum[x_i - \min(x_1, \dots, x_m)]} > C.$$

(ii) This test is also UMP unbiased.

(iii) Extend these results to the case that only the r smallest X's and the s smallest Y's are observed.

[(ii): See Problem 5.15.]

Problem 6.12 If $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ are samples from $N(\xi, \sigma^2)$ and $N(\eta, \tau^2)$, respectively, the problem of testing $\tau^2 = \sigma^2$ against the two-sided alternatives $\tau^2 \ne \sigma^2$ remains invariant under the group G generated by the transformations $X_i' = aX_i + b, Y_i' = aY_i + c, \ (a \ne 0)$, and $X_i' = Y_i, Y_i' = X_i$. There exists a UMP invariant test under G with rejection region

$$W = \max\left\{ \frac{\sum(Y_i - \bar{Y})^2}{\sum(X_i - \bar{X})}, \frac{\sum(X_i - \bar{X})^2}{\sum(Y_i - \bar{Y})^2} \right\} \ge k.$$

[The ratio of the probability densities of W for $\tau^2/\sigma^2 = \Delta$ and $\tau^2/\sigma^2 = 1$ is proportional to $[(1 + w)/(\Delta + w)]^{n-1} + [(1 + w)/(1 + \Delta w)]^{n-1}$ for $w \ge 1$. The derivative of this expression is ≥ 0 for all Δ.]

Problem 6.13 Let $X_1, \ldots, X_n$ be a sample from a distribution with density

$$\frac{1}{\tau^n} f\left(\frac{x_1}{\tau}\right) \ldots f\left(\frac{x_n}{\tau}\right),$$

where $f(x)$ is either zero for $x < 0$ or symmetric about zero. The most powerful scale-invariant test for testing $H : f = f_0$ against $K : f = f_1$ rejects when

$$\frac{\int_0^\infty v^{n-1} f_1(vx_1) \ldots f_1(vx_n)\, dv}{\int_0^\infty v^{n-1} f_0(vx_1) \ldots f_0(vx_n)\, dv} > C.$$

Problem 6.14 *Normal versus double exponential*. For $f_0(x) = e^{-x^2/2}/\sqrt{2\pi}$, $f_1(x) = e^{-|x|}/2$, the test of the preceding problem reduces to rejecting when $\sqrt{\sum x_i^2}/\sum |x_i| < C$.

(Hogg, 1972.)

Note. The corresponding test when both location and scale are unknown is obtained in Uthoff (1973). Testing normality against Cauchy alternatives is discussed by Franck (1981).

Problem 6.15 *Uniform versus triangular*.

(i) For $f_0(x) = 1$ $(0 < x < 1)$, $f_1(x) = 2x$ $(0 < x < 1)$, the test of Problem 6.13 reduces to rejecting when $T = x_{(n)}/\bar{x} < C$.
(ii) Under f_0, the statistic $2n \log T$ is distributed as χ^2_{2n}.

(Quesenberry and Starbuck, 1976.)

Problem 6.16 Show that the test of Problem 6.10(i) reduces to

(i) $[x_{(n)} - x_{(1)}]/S < c$ for normal versus uniform;
(ii) $[\bar{x} - x_{(1)}]/S < c$ for normal versus exponential;
(iii) $[\bar{x} - x_{(1)}]/[x_{(n)} - x_{(1)}] < c$ for uniform versus exponential.

(Uthoff, 1970.)

Note. When testing for normality, one is typically not interested in distinguishing the normal from some other given shape but would like to know more generally whether the data are or are not consonant with a normal distribution. This is a special case of the problem of testing for goodness of fit, which is briefly discussed at the end of Section 6.13 and forms the topic of Chapter 16; also, see the many references in the notes to Chapter 16.

Problem 6.17 Let $X_1, \ldots, X_n$ be independent and normally distributed. Suppose X_i has mean μ_i and variance σ^2 (which is the same for all i). Consider testing the null hypothesis that $\mu_i = 0$ for all i. Using invariance considerations, find a UMP invariant test with respect to a suitable group of transformations in each of the following cases:

(i) σ^2 is known and equal to one.
(ii) σ^2 is unknown.

Section 6.4

Problem 6.18 (i) When testing $H : p \leq p_0$ against $K : p > p_0$ by means of the test corresponding to (6.15), determine the sample size required to obtain power β against $p = p_1$, $\alpha = 0.05$, $\beta = 0.9$ for the cases $p_0 = 0.1$, $p_1 = 0.15, 0.20, 0.25$; $p_0 = 0.05$, $p_1 = 0.10, 0.15, 0.20, 0.25$; $p_0 = 0.01$, $p_1 = 0.02, 0.05, 0.10, 0.15, 0.20$.

(ii) Compare this with the sample size required if the inspection is by attributes and the test is based on the total number of defectives.

Problem 6.19 *Two-sided t-test.*

(i) Let $X_1, \ldots, X_n$ be a sample from $N(\xi, \sigma^2)$. For testing $\xi = 0$ against $\xi \neq 0$, there exists a UMP invariant test with respect to the group $X_i' = cX_i$, $c \neq 0$, given by the two-sided t-test (5.17).

(ii) Let $X_1, \ldots, X_m$, and $Y_1, \ldots, Y_n$ be samples from $N(\xi, \sigma^2)$ and $N(\eta, \sigma^2)$ respectively. For testing $\eta = \xi$ against $\eta \neq \xi$ there exists a UMP invariant test with respect to the group $X_i' = aX_i + b$, $Y_j' = aY_j + b$, $a \neq 0$, given by the two-sided t-test (5.30).

[(i): Sufficiency and invariance reduce the problem to $|t|$, which in the notation of Section 4 has the probability density $p\delta(t) + p_\delta(-t)$ for $t > 0$. The ratio of this density for $\delta = \delta_1$ to its value for $\delta = 0$ is proportional to $\int_0^\infty (e^{\delta_1 v} + e^{-\delta_1 v}) g_{t^2}(v)\, dv$, which is an increasing function of t^2 and hence of $|t|$.]

Problem 6.20 *Testing a correlation coefficient.* Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a sample from a bivariate normal distribution.

(i) For testing $\rho \leq \rho_0$ against $\rho > \rho_0$ there exists a UMP invariant test with respect to the group of all transformations $X_i' = aX_i + b$, $Y_i' = cY_1 + d$ for which $a, c > 0$. This test rejects when the sample correlation coefficient R is too large.

(i) The problem of testing $\rho = 0$ against $\rho \neq 0$ remains invariant in addition under the transformation $Y_i' = -Y_i$, $X_i' = X_i$. With respect to the group generated by this transformation and those of (i) there exists a UMP invariant test, with rejection region $|R| \geq C$.

[(i): To show that the probability density $p_\rho(r)$ of R has monotone likelihood ratio, apply the condition of Problem 3.28(i), to the expression given for this density in Problem 5.67. Putting $t = \rho r + 1$, the second derivative $\partial^2 \log p_\rho(r)/\partial\rho\partial r$ up to a positive factor is

$$\frac{\sum_{i,j=0}^{\infty} c_i c_j t^{i+j-2} \left[(j-i)^2(t-1) + (i+j)\right]}{2\left[\sum_{i=0}^{\infty} c_i t^i\right]^2}.$$

To see that the numerator is positive for all $t > 0$, note that it is greater than

$$2\sum_{i=0}^{\infty} c_i t^{i-2} \sum_{j=i+1}^{\infty} c_j t^j \left[(j-i)^2(t-1) + (i+j)\right].$$

Holding i fixed and using the inequality $c_{j+1} < \frac{1}{2}c_j$, the coefficient of t^j in the interior sum is ≥ 0.]

Problem 6.21 Let (X_i, Y_i) be independent $N(\mu_i, \sigma^2)$ for $i = 1, \ldots, n$. The parameters $\mu_1, \ldots, \mu_n$ and σ^2 are all unknown. For testing $\sigma = 1$ against $\sigma > 1$, determine the UMPI level α test. Is the test also UMPU?

Problem 6.22 For testing the hypothesis that the correlation coefficient ρ of a bivariate normal distribution is $\leq \rho_0$, determine the power against the alternative $\rho = \rho_1$, when the level of significance α is .05, $\rho_0 = .3$, $\rho_1 = .5$, and the sample size n is 50, 100, 200.

Section 6.5

Problem 6.23 Almost invariance of a test ϕ with respect to the group G of either Problem 6.11(i) or Example 6.3.5 implies that ϕ is equivalent to an invariant test.

Problem 6.24 The totality of permutations of K distinct numbers $a_1, \ldots, a_K$, for varying $a_1, \ldots, a_K$ can be represented as a subset C_K of Euclidean K-space R_K, and the group G of Example 6.5.1 as the union of $C_2, C_3, \ldots$. Let ν be the measure over G which assigns to a subset B of G the value $\sum_{k=2}^{\infty} \mu_K(B \cap C_K)$, where μ_K denotes Lebesgue measure in E_K. Give an example of a set $B \subset G$ and an element $g \in G$ such that $\nu(B) > 0$ but $\nu(Bg) = 0$.

[If a, b, c, d are distinct numbers, the permutations g, g' taking (a, b) into (b, a) and (c, d) into (d, c) respectively are points in C_2, but gg' is a point in C_4.]

Section 6.6

Problem 6.25 Show that

(i) G_1 of Example 6.6.2 is a group;
(ii) the test which rejects when $X_{21}^2/X_{11}^2 > C$ is UMP invariant under G_1;
(iii) the smallest group containing G_1 and G_2 is the group G of Example 6.6.2.

Problem 6.26 Consider a testing problem which is invariant under a group G of transformations of the sample space, and let $\mathcal{C}$ be a class of tests which is closed under G, so that $\phi \in \mathcal{C}$ implies $\phi g \in \mathcal{C}$, where ϕg is the test defined by $\phi g(x) = \phi(gx)$. If there exists an a.e. unique UMP member ϕ_0 of $\mathcal{C}$, then ϕ_0 is almost invariant.

Problem 6.27 *Envelope power function*. Let $S(\alpha)$ be the class of all level-α tests of a hypothesis H, and let $\beta^*_\alpha(\theta)$ be the *envelope power function*, defined by

$$\beta^*_\alpha(\theta) = \sup_{\phi \in S(\alpha)} \beta_\phi(\theta),$$

where β_ϕ denotes the power function of ϕ. If the problem of testing H is invariant under a group G, then $\beta^*_\alpha(\theta)$ is invariant under the induced group $\bar{G}$.

Problem 6.28 (i) A generalization of equation (6.2) is

$$\int_A f(x)\, dP_\theta(x) = \int_{gA} f(g^{-1}x)\, dP_{\bar{g}\theta}(x).$$

(ii) If P_{θ_1} is absolutely continuous with respect to P_{θ_0}, then $P_{\bar{g}\theta_1}$ is absolutely continuous with respect to $P_{\bar{g}\theta_0}$ and

$$\frac{dP_{\theta_1}}{dP_{\theta_0}}(x) = \frac{dP_{\bar{g}\theta_1}}{dP_{\bar{g}\theta_0}}(gx) \qquad \left(\text{a.e. } P_{\theta_0}\right).$$

(iii) The distribution of $dP_{\theta_1}/dP_{\theta_0}(X)$ when X is distributed as P_{θ_0} is the same as that of $dP_{\bar{g}\theta_1}/dP_{\bar{g}\theta_0}(X')$ when X' is distributed as $P_{\bar{g}\theta_0}$.

Problem 6.29 *Invariance of likelihood ratio*. Let the family of distributions $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ be dominated by μ, let $p_\theta = dP_\theta/d\mu$, let μg^{-1} be the measure defined by $\mu g^{-1}(A) = \mu[g^{-1}(A)]$, and suppose that μ is absolutely continuous with respect to μg^{-1} for all $g \in G$.

(i) Then

$$p_\theta(x) = p_{\bar{g}\theta}(gx) \frac{d\mu}{d\mu g^{-1}}(gx) \qquad (\text{a.e. } \mu).$$

(ii) Let Ω and ω be invariant under $\bar{G}$, and countable. Then the likelihood ratio $\sup_\Omega p_\theta(x)/\sup_\omega p_\theta(x)$ is almost invariant under G.

(iii) Suppose that $p_\theta(x)$ is continuous in θ for all x, that Ω is a separable pseudometric space, and that Ω and ω are invariant. Then the likelihood ratio is almost invariant under G.

Problem 6.30 *Inadmissible likelihood ratio test*. In many applications in which a UMP invariant test exists, it coincides with the likelihood ratio test. That this is, however, not always the case is seen from the following example. Let $P_1, \ldots, P_n$ be n equidistant points on the circle $x^2 + y^2 = 4$, and $Q_1, \ldots, Q_n$ on the circle $x^2 + y^2 = 1$. Denote the origin in the (x, y) plane by O, let $0 < \alpha \le \frac{1}{2}$ be fixed, and let (X, Y) be distributed over the $2n + 1$ points $P_1, \ldots, P_n, Q_1, \ldots, Q_n, O$ with probabilities given by the following table:

	P_i	Q_i	O
H	α/n	$(1-2\alpha)/n$	α
K	p_i/n	0	$(n-1)/n$,

where $\sum p_i = 1$. The problem remains invariant under rotations of the plane by the angles $2k\pi/n$ $(k = 0, 1, \ldots, n-1)$. The rejection region of the likelihood ratio test consists of the points $P_1, \ldots, P_n$, and its power is $1/n$. On the other hand, the UMP invariant test rejects when $X = Y = 0$ and has power $(n-1)/n$.

Problem 6.31 Let G be a group of transformations of $\mathcal{X}$, and let $\mathcal{A}$ be a σ-field of subsets of $\mathcal{X}$, and μ a measure over $(\mathcal{X}, \mathcal{A})$. Then a set $A \in \mathcal{A}$ is said to be almost invariant if its indicator function is almost invariant.

(i) The totality of almost invariant sets forms a σ-field $\mathcal{A}_0$, and a critical function is almost invariant if and only if it is $\mathcal{A}_0$-measurable.
(ii) Let $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ be a dominated family of probability distributions over $(\mathcal{X}, \mathcal{A})$, and suppose that $\bar{g}\theta = \theta$ for all $\bar{g} \in \bar{G}, \theta \in \Omega$. Then the σ-field $\mathcal{A}_0$ of almost invariant sets is sufficient for $\mathcal{P}$.

[Let $\lambda = \sum c_i P_{\theta_i}$, be equivalent to $\mathcal{P}$. Then

$$\frac{dP_\theta}{d\lambda}(gx) = \frac{dP_{g^{-1}\theta}}{\sum c_i\, dP_{g^{-1}\theta_i}}(x) = \frac{dP_\theta}{d\lambda}(x) \qquad \text{(a.e. } \lambda\text{)},$$

so that $dP_\theta/d\lambda$ is almost invariant and hence $\mathcal{A}_0$-measurable.]

Problem 6.32 The UMP invariant test of Problem 6.14 is also UMP similar.

[Consider the problem of testing $\alpha = 0$ versus $\alpha > 0$ in the two-parameter exponential family with density

$$C(\alpha, \tau) \exp\left(-\frac{\alpha}{2\tau^2} \sum x_i^2 - \frac{1-\alpha}{\tau} \sum |x_i|\right), \qquad 0 \le \alpha < 1.]$$

Note. For the analogous result for the tests of Problem 6.15, 6.16, see Quesenberry and Starbuck (1976).

Problem 6.33 The following UMP unbiased tests of Chapter 5 are also UMP invariant under change in scale:

(i) The test of $g \le g_0$ in a gamma distribution (Problem 5.30).
(ii) The test of $b_1 \le b_2$ in Problem 5.18(i).

Section 6.7

Problem 6.34 The definition of d-admissibility of a test coincides with the admissibility definition given in Section 1.8 when applied to a two-decision procedure with loss 0 or 1 as the decision taken is correct or false.

Problem 6.35 (i) The following example shows that α-admissibility does not always imply d-admissibility. Let X be distributed as $U(0, \theta)$, and consider the tests φ_1 and φ_2 which reject when, respectively, $X < 1$ and $X < \frac{3}{2}$ for testing $H : \theta = 2$ against $K : \theta = 1$. Then for $\alpha = \frac{3}{4}$, φ_1 and φ_2 are both α-admissible but φ_2 is not d-admissible.

(ii) Verify the existence of the test φ_0 of Example 6.7.2.

Problem 6.36 (i) The acceptance region $T_1/\sqrt{T_2} \leq C$ of Example 6.7.3 is a convex set in the (T_1, T_2) plane.

(ii) In Example 6.7.3, the conditions of Theorem 6.7.1 are not satisfied for the sets $A : T_1/\sqrt{T_2} \leq C$ and $\Omega' : \xi > k$.

Problem 6.37 (i) In Example 6.7.4 show that there exist C_0, C_1 such that $\lambda_0(\eta)$ and $\lambda_1(\eta)$ are probability densities (with respect to Lebesgue measure).

(ii) Verify the densities h_0 and h_1.

Problem 6.38 Verify

(i) the admissibility of the rejection region (6.27);

(ii) the expression for $I(z)$ given in the proof of Lemma 6.7.1.

Problem 6.39 Let $X_1, \ldots, X_m; Y_1, \ldots, Y_n$ be independent $N(\xi, \sigma^2)$ and $N(\eta, \sigma^2)$ respectively. The one-sided t-test of $H : \delta = \xi/\sigma \leq 0$ is admissible against the alternatives (i) $0 < \delta < \delta_1$ for any $\delta_1 > 0$; (ii) $\delta > \delta_2$ for any $\delta_2 > 0$.

Problem 6.40 For the model of the preceding problem, generalize Example 6.7.3 (continued) to show that the two-sided t-test is a Bayes solution for an appropriate prior distribution.

Problem 6.41 Suppose $X = (X_1, \ldots, X_k)^\top$ is multivariate normal with unknown mean vector $(\theta_1, \ldots, \theta_k)^\top$ and known nonsingular covariance matrix Σ. Consider testing the null hypothesis $\theta_i = 0$ for all i against $\theta_i \neq 0$ for some i. Let C be any closed convex subset of k-dimensional Euclidean space, and let ϕ be the test that accepts the null hypothesis if X falls in C. Show that ϕ is admissible. *Hint*: First assume Σ is the identity and use Theorem 6.7.1. [An alternative proof is provided by Strasser (1985, Theorem 30.4).]

Section 6.9

Problem 6.42 *Wilcoxon two-sample test*. Let $U_{ij} = 1$ or 0 as $X_i < Y_j$ or $X_i > Y_j$, and let $U = \sum\sum U_{ij}$ be the number of pairs X_i, Y_j with $X_i < Y_j$.

(i) Then $U = \sum S_i - \frac{1}{2}n(n+1)$, where $S_1 < \cdots < S_n$ are the ranks of the Y's so that the test with rejection region $U > C$ is equivalent to the Wilcoxon test.
(ii) Any given arrangement of x's and y's can be transformed into the arrangement $x \ldots xy \ldots y$ through a number of interchanges of neighboring elements. The smallest number of steps in which this can be done for the observed arrangement is $mn - U$.

Problem 6.43 *Expectation and variance of Wilcoxon statistic.* If the X's and Y's are samples from continuous distributions F and G, respectively, the expectation and variance of the Wilcoxon statistic U defined in the preceding problem are given by

$$E\left(\frac{U}{mn}\right) = P\{X < Y\} = \int F\,dG \tag{6.62}$$

and

$$\begin{aligned} mnVar\left(\frac{U}{mn}\right) = & \int F\,dG + (n-1)\int (1-G)^2\,dF \\ & +(m-1)\int F^2\,dG - (m+n-1)\left(\int F\,dG\right)^2 . \end{aligned} \tag{6.63}$$

Under the hypothesis $G = F$, these reduce to

$$E\left(\frac{U}{mn}\right) = \frac{1}{2}, \qquad Var\left(\frac{U}{mn}\right) = \frac{m+n+1}{12mn}. \tag{6.64}$$

Problem 6.44 (i) Let $Z_1, \ldots, Z_N$ be independently distributed with densities $f_1, \ldots, f_N$, and let the rank of Z_i be denoted by T_i. If f is any probability density which is positive whenever at least one of the f_i is positive, then

$$P\{T_1 = t_1, \ldots, T_N = t_n\} = \frac{1}{N!} E\left[\frac{f_1\left(V_{(t_1)}\right)}{f\left(V_{(t_1)}\right)} \cdots \frac{f_N\left(V_{(t_N)}\right)}{f\left(V_{(t_N)}\right)}\right], \tag{6.65}$$

where $V_{(1)} < \cdots < V_{(N)}$ is an ordered sample from a distribution with density f.
(ii) If $N = m + n$, $f_1 = \cdots = f_m = f$, $f_{m+1} = \cdots = f_{m+n} = g$, and $S_1 < \cdots < S_n$ denote the ordered ranks of $Z_{m+1}, \ldots, Z_{m+n}$ among all the Z's, the probability distribution of $S_1, \ldots, S_n$ is given by (6.30).

[(i): The probability in question is $\int \ldots \int f_1(z_1) \ldots f_N(z_N)\,dz_1 \cdots dz_N$ integrated over the set in which z_i is the t_ith smallest of the z's for $i = 1, \ldots, N$. Under the transformation $w_{t_i} = z_i$ the integral becomes $\int \ldots \int f_1(w_{t_1}) \ldots f_N(w_{t_N})\,dw_1 \cdots dw_N$ integrated over the set $w_1 < \cdots < w_N$. The desired result now follows from the fact that the probability density of the order statistics $V_{(1)} < \cdots < V_{(N)}$ is $N!f(w_1) \cdots f(w_N)$ for $w_1 < \ldots < w_N$.]

Problem 6.45 (i) For any continuous cumulative distribution function F, define $F^{-1}(0) = -\infty$, $F^{-1}(y) = \inf\{x : F(x) = y\}$ for $0 < y < 1$, $F^{-1}(1) = \infty$ if $F(x) < 1$ for all finite x, and otherwise $\inf\{x : F(x) = 1\}$. Then $F[F^{-1}(y)] = y$ for all $0 \le y \le 1$, but $F^{-1}[F(y)]$ may be $< y$.

(ii) Let Z have a cumulative distribution function $G(z) = h[F(z)]$, where F and h are continuous cumulative distribution functions, the latter defined over (0,1). If $Y = F(Z)$, then $P\{Y < y\} = h(y)$ for all $0 \le y \le 1$.

(iii) If Z has the continuous cumulative distribution function F, then $F(Z)$ is uniformly distributed over (0, 1).

[(ii): $P\{F(Z) < y\} = P\{Z < F^{-1}(y)\} = F[F^{-1}(y)] = y$.]

Problem 6.46 Let Z_i have a continuous cumulative distribution function F_i ($i = 1, \ldots, N$), and let G be the group of all transformations $Z'_i = f(Z_i)$ such that f is continuous and strictly increasing.

(i) The transformation induced by f in the space of distributions is $F'_i = F_i(f^{-1})$.

(ii) Two N-tuples of distributions $(F_1, \ldots, F_N)$ and $(F'_1, \ldots, F'_N)$ belong to the same orbit with respect to $\bar{G}$ if and only if there exist continuous distribution functions $h_1, \ldots, h_N$ defined on (0,1) and strictly increasing continuous distribution functions F and F' such that $F_i = h_i(F)$ and $F'_i = h_i(F')$.

[(i): $P\{f(Z_i) \le y\} = P\{Z_i \le f^{-1}(y)\} = F_i[f^{-1}(y)]$.
(ii): If $F_i = h_i(F)$ and the F'_i are on the same orbit, so that $F'_i = F_i(f^{-1})$, then $F'_i = h_i(F')$ with $F' = F(f^{-1})$. Conversely, if $F_i = h_i(F)$, $F'_i = h_i(F')$, then $F'_i = F_i(f^{-1})$ with $f = F'^{-1}(F)$.]

Problem 6.47 Under the assumptions of the preceding problem, if $F_i = h_i(F)$, the distribution of the ranks $T_1, \ldots, T_N$ of $Z_1, \ldots, Z_N$ depends only on the h_i, not on F. If the h_i are differentiable, the distribution of the T_i is given by

$$P\{T_1 = t_1, \ldots, T_N = t_n\} = \frac{E\left[h'_1\left(U_{(t_1)}\right) \ldots h'_N\left(U_{(t_N)}\right)\right]}{N!}, \tag{6.66}$$

where $U_{(1)} < \cdots < U_{(N)}$ is an ordered sample of size N from the uniform distribution $U(0, 1)$. [The left-hand side of (6.66) is the probability that of the quantities $F(Z_1), \ldots, F(Z_N)$, the ith one is the t_ith smallest for $i = 1, \ldots, N$. This is given by $\int \ldots \int h'_1(y_1) \ldots h'_N(y_N)\, dy$ integrated over the region in which y_i is the t_ith smallest of the y's for $i = 1, \ldots, N$. The proof is completed as in Problem 6.44.]

Problem 6.48 *Distribution of order statistics.*

(i) If $Z_1, \ldots, Z_N$ is a sample from a cumulative distribution function F with density f, the joint density of $Y_i = Z_{(s_i)}$, $i = 1, \ldots, n$, is

$$\frac{N! f(y_1) \ldots f(y_n)}{(s_1 - 1)!(s_2 - s_1 - 1)! \ldots (N - s_n)!} \tag{6.67}$$
$$\times [F(y_1)]^{s_1 - 1} [F(y_2) - F(y_1)]^{s_2 - s_1 - 1} \ldots [1 - F(y_n)]^{N - s_n}$$

for $y_1 < \cdots < y_n$.

(ii) For the particular case that the Z's are a sample from the uniform distribution on (0,1), this reduces to

$$\frac{N!}{(s_1-1)!(s_2-s_1-1)!\dots(N-s_n)!} \tag{6.68}$$
$$y_1^{s_1-1}(y_2-y_1)^{s_2-s_1-1}\dots(1-y_n)^{N-s_n}.$$

For $n=1$, (6.68) is the density of the beta distribution $B_{s,N-s+1}$, which therefore is the distribution of the single order statistic $Z_{(s)}$ from $U(0,1)$.

(iii) Let the distribution of $Y_1,\dots,Y_n$ be given by (6.68), and let V_i be defined by $Y_i = V_iV_{i+1}\dots V_n$ for $i=1,\dots,n$. Then the joint distribution of the V_i is

$$\frac{N!}{(s_1-1)!\dots(N-s_n)!}\prod_{i=1}^{n} v_i^{s_i-1}(1-v_i)^{s_{i+1}-s_i-1} \qquad (s_{n+1}=N+1),$$

so that the V_i are independently distributed according to the beta distribution $B_{s_i-i+1,s_{i+1}-s_i}$.

[(i): If $Y_1 = Z_{(s_1)},\dots,Y_n = Z_{(s_n)}$ and $Y_{n+1},\dots,Y_N$ are the remaining Z's in the original order of their subscripts, the joint density of $Y_1,\dots,Y_n$ is $N(N-1)\dots(N-n+1)\int\dots\int f(y_{n+1})\dots f(y_N)\,dy_{n+1}\dots dy_N$ integrated over the region in which s_1-1 of the y's are $< y_1$, s_2-s_1-1 between y_1 and y_2, ..., and $N-s_n > y_n$. Consider any set where a particular s_1-1 of the y's is $< y_1$, a particular s_2-s_1-1 of them is between y_1 and y_2, and so on, There are $N!/(s_1-1)!\dots(N-s_n)!$ of these regions, and the integral has the same value over each of them, namely $[F(y_1)]^{s_1-1}[F(y_2)-F(y_1)]^{s_2-s_1-1}\dots[1-F(y_n)]^{N-s_n}$.]

Problem 6.49 (i) If $X_1,\dots,X_m$ and $Y_1,\dots,Y_n$ are samples with continuous cumulative distribution functions F and $G=h(F)$ respectively, and if h is differentiable, the distribution of the ranks $S_1 < \dots < S_n$ of the Y's is given by

$$P\{S_1=s_1,\dots,S_n=s_n\} = \frac{E\left[h'\left(U_{(s_1)}\right)\dots h'\left(U_{(s_n)}\right)\right]}{\binom{m+n}{m}}, \tag{6.69}$$

where $U_{(1)} < \cdots < U_{(m+n)}$ is an ordered sample from the uniform distribution $U(0,1)$.

(ii) If in particular $G=F^k$, where k is a positive integer, (6.69) reduces to

$$P\{S_1=s_1,\dots,S_n=s_n\} \tag{6.70}$$
$$= \frac{k^n}{\binom{m+n}{m}}\prod_{j=1}^{n}\frac{\Gamma\left(s_j+jk-j\right)}{\Gamma\left(s_j\right)}\cdot\frac{\Gamma\left(s_{j+1}\right)}{\Gamma\left(s_{j+1}+jk-j\right)}.$$

Problem 6.50 For sufficiently small $\theta > 0$, the Wilcoxon test at level

$$\alpha = k \Big/ \binom{N}{n}, \qquad k \text{ a positive integer,}$$

maximizes the power (among rank tests) against the alternatives (F, G) with $G = (1-\theta)F + \theta F^2$.

Problem 6.51 An alternative proof of the optimum property of the Wilcoxon test for detecting a shift in the logistic distribution is obtained from the preceding problem by equating $F(x-\theta)$ with $(1-\theta)F(x) + \theta F^2(x)$, neglecting powers of θ higher than the first. This leads to the differential equation $F - \theta F' = (1-\theta)F + \theta F^2$, the solution of which is the logistic distribution.

Problem 6.52 Let $\mathcal{F}_0$ be a family of probability measures over $(\mathcal{X}, \mathcal{A})$, and let $\mathcal{C}$ be a class of transformations of the space $\mathcal{X}$. Define a class $\mathcal{F}_1$ of distributions by $F_1 \in \mathcal{F}_1$ if there exists $F_0 \in \mathcal{F}_0$ and $f \in \mathcal{C}$ such that the distribution of $f(X)$ is F_1 when that of X is F_0. If ϕ is any test satisfying (a) $E_{F_0}\phi(X) = \alpha$ for all $F_0 \in \mathcal{F}_0$, and (b) $\phi(x) \le \phi[f(x)]$ for all x and all $f \in \mathcal{C}$, then ϕ is unbiased for testing $\mathcal{F}_0$ against $\mathcal{F}_1$

Problem 6.53 Let $X_1, \dots, X_m; Y_1, \dots, Y_n$ be samples from a common continuous distribution F. Then the Wilcoxon statistic U defined in Problem 6.42 is distributed symmetrically about $\frac{1}{2}mn$ even when $m \neq n$.

Problem 6.54 (i) If $X_1, \dots, X_m$ and $Y_1, \dots, Y_n$ are samples from $F(x)$ and $G(y) = F(y - \Delta)$, respectively, (F continuous), and $D_{(1)} < \cdots < D_{(mn)}$ denote the ordered differences $Y_j - X_i$, then

$$P\left[D_{(k)} < \Delta < D_{(mn+1-k)}\right] = P_0[k \le U \le mn - k],$$

where U is the statistic defined in Problem 6.42 and the probability on the right side is calculated for $\Delta = 0$.

(ii) Determine the above confidence interval for Δ when $m = n = 6$, the confidence coefficient is $\frac{20}{21}$, and the observations are x : 0.113, 0.212, 0.249, 0.522, 0.709, 0.788, and y : 0.221, 0.433, 0.724, 0.913, 0.917, 1.58.

(iii) For the data of (ii) determine the confidence intervals based on Student's t for the case that F is normal.

Hint: $D_{(i)} \le \Delta < D_{(i+1)}$ if and only if $U_\Delta = mn - i$, where U_Δ is the statistic U of Problem 6.42 calculated for the observations

$$X_1, \dots, X_m; Y_1 - \Delta, \dots, Y_n - \Delta.$$

[An alternative measure of the amount by which G exceeds F (without assuming a location model) is $p = P\{X < Y\}$. The literature on confidence intervals for p is reviewed in Mee (1990).]

Problem 6.55 (i) Let X, X' and Y, Y' be independent samples of size 2 from continuous distributions F and G, respectively. Then

$$\begin{aligned} p &= P\{\max(X, X') < \min(Y, Y')\} + P\{\max(Y, Y') < \min(X, X')\} \\ &= \tfrac{1}{3} + 2\Delta, \end{aligned}$$

where $\Delta = \int (F - G)^2 \, d[(F + G)/2]$.
(ii) $\Delta = 0$ if and only if $F = G$.
[(i): $p = \int (1 - F)^2 \, dG^2 + \int (1 - G)^2 \, dF^2$ which after some computation reduces to the stated form.
(ii): $\Delta = 0$ implies $F(x) = G(x)$ except on a set N which has measure zero both under F and G. Suppose that $G(x_1) - F(x_1) = \eta > 0$. Then there exists x_0 such that $G(x_0) = F(x_0) + \frac{1}{2}\eta$ and $F(x) < G(x)$ for $x_0 \le x \le x_1$. Since $G(x_1) - G(x_0) > 0$, it follows that $\Delta > 0$.]

Problem 6.56 *Continuation.*

(i) There exists at every significance level α a test of $H : G = F$ which has power $> \alpha$ against all continuous alternatives (F, G) with $F \neq G$.
(ii) There does not exist a nonrandomized unbiased rank test of H against all $G \neq F$ at level

$$\alpha = 1 \Big/ \binom{m+n}{n}.$$

[(i): let $X_i, X_i'; Y_i, Y_i'$ $(i = 1, \dots, n)$ be independently distributed, the X's with distribution F, the Y's with distribution G, and let $V_i = 1$ if $\max(X_i, X_1') < \min(Y_i, Y_i')$ or $\max(Y_i, Y_i') < \min(X_i, X_i')$, and $V_i = 0$ otherwise. Then $\sum V_i$ has a binomial distribution with the probability p defined in Problem 6.55, and the problem reduces to that of testing $p = \frac{1}{3}$ against $p > \frac{1}{3}$.
(ii): Consider the particular alternatives for which $P\{X < Y\}$ is either 1 or 0.]

Problem 6.57 (i) Let $X_1, \dots, X_m; Y_1, \dots, Y_n$ be i.i.d. according to a continuous distribution F, let the ranks of the Y's be $S_1 < \cdots < S_n$, and let $T = h(S_1) + \cdots + h(S_n)$. Then if either $m = n$ or $h(s) + h(N + 1 - s)$ is independent of s, the distribution of T is symmetric about $n \sum_{i=1}^N h(i)/N$.
(ii) Show that the two-sample Wilcoxon and normal scores statistics are symmetrically distributed under H, and determine their centers of symmetry.
[(i): Let $S_i' = N + 1 - S_i$, and use the fact that $T' = \sum h(S_j')$ has the same distribution under H as T.]

Section 6.10

Problem 6.58 (i) Let m and n be the numbers of negative and positive observations among $Z_1, \dots, Z_N$, and let $S_1 < \cdots < S_n$ denote the ranks of the positive Z's among $|Z_1|, \dots |Z_N|$. Consider the $N + \frac{1}{2}N(N-1)$ distinct sums $Z_i + Z_j$ with $i = j$ as well as $i \neq j$. The Wilcoxon signed-rank statistic $\sum S_j$, is equal to the number of these sums that are positive.

(ii) If the common distribution of the Z's is D, then

$$E\left(\sum S_j\right) = \tfrac{1}{2}N(N+1) - ND(0) - \tfrac{1}{2}N(N-1)\int D(-z)\,dD(z).$$

[(i) Let K be the required number of positive sums. Since $Z_i + Z_j$ is positive if and only if the Z corresponding to the larger of $|Z_i|$ and $|Z_j|$ is positive, $K = \sum_{i=1}^N \sum_{j=1}^N U_{ij}$ where $U_{ij} = 1$ if $Z_j > 0$ and $|Z_i| \leq Z_j$ and $U_{ij} = 0$ otherwise.]

Problem 6.59 Let $Z_1, \dots, Z_N$ be a sample from a distribution with density $f(z - \theta)$, where $f(z)$ is positive for all z and f is symmetric about 0, and let m, n, and the S_j be defined as in the preceding problem.

(i) The distribution of n and the S_j is given by

$$P\{\text{the number of positive } Z\text{'s is } n \text{ and } S_1 = s_1, \dots, S_n = s_n\} \tag{6.71}$$
$$= \frac{1}{2^N} E\left[\frac{f\left(V_{(r_1)} + \theta\right)\dots f\left(V_{(r_m)} + \theta\right) f\left(V_{(s_1)} - \theta\right)\dots f\left(V_{(s_n)} - \theta\right)}{f\left(V_{(1)}\right)\dots f\left(V_{(N)}\right)}\right],$$

where $V_{(1)} < \dots < V_{(N)}$, is an ordered sample from a distribution with density $2f(v)$ for $v > 0$, and 0 otherwise.

(ii) The rank test of the hypothesis of symmetry with respect to the origin, which maximizes the derivative of the power function at $\theta = 0$ and hence maximizes the power for sufficiently small $\theta > 0$, rejects, under suitable regularity conditions, when

$$-E\left[\sum_{j=1}^n \frac{f'(V_{(s_j)})}{f(V_{(s_j)})}\right] > C.$$

(iii) In the particular case that $f(z)$ is a normal density with zero mean, the rejection region of (ii) reduces to $\sum E(V(s_j) > C$, where $V_{(1)} < \dots < V_{(N)}$ is an ordered sample from a χ-distribution with 1 degree of freedom.

(iv) Determine a density f such that the one-sample Wilcoxon test is most powerful against the alternatives $f(z - \theta)$ for sufficiently small positive θ.

[(i): Apply Problem 6.44(i) to find an expression for $P\{S_1 = s_1, \dots, S_n = s_n$ given that the number of positive Z's is $n\}$.]

Problem 6.60 An alternative expression for (6.71) is obtained if the distribution of Z is characterized by (ρ, F, G). If then $G = h(F)$ and h is differentiable, the distribution of n and the S_j is given by

$$\rho^m(1-\rho)^n E\left[h'(U_{(s_1)})\cdots h'(U_{(s_n)})\right], \tag{6.72}$$

where $U_{(1)}, < \cdots < U_{(N)}$ is an ordered sample from $U(0, 1)$.

Problem 6.61 *Unbiased tests of symmetry.* Let $Z_1, \ldots, Z_N$, be a sample, and let ϕ be any rank test of the hypothesis of symmetry with respect to the origin such that $z_i \le z_i'$ for all i implies $\phi(z_1, \ldots, z_N) \le \phi(z_1', \ldots, z'N)$. Then ϕ is unbiased against the one-sided alternatives that the Z's are stochastically larger than some random variable that has a symmetric distribution with respect to the origin.

Problem 6.62 *The hypothesis of randomness.*[7] Let $Z_1, \ldots, Z_N$ be independently distributed with distributions $F_1, \ldots, F_N$, and let T_i denote the rank of Z_i among the Z's. For testing the *hypothesis of randomness* $F_1 = \cdots = F_N$ against the alternatives K of an *upward trend*, namely, that Z_i is stochastically increasing with i, consider the rejection regions

$$\sum it_i > C \tag{6.73}$$

and

$$\sum iE(V_{(t_i)}) > C, \tag{6.74}$$

where $V_{(1)} < \cdots < V_{(N)}$ is an ordered sample from a standard normal distribution and where t_i is the value taken on by T_i.

(i) The second of these tests is most powerful among rank tests against the normal alternatives $F = N(\gamma + i\delta, \sigma^2)$ for sufficiently small δ.
(ii) Determine alternatives against which the first test is a most powerful rank test.
(iii) Both tests are unbiased against the alternatives of an upward trend; so is any rank test ϕ satisfying $\phi(z_1, \ldots, z_N) \le \phi(z_1', \ldots, z_N')$ for any two points for which $i < j$, $z_i < z_j$ implies $z_i' < z_j'$ for all i and j.

[(iii): Apply Problem 6.52 with $\mathcal{C}$ the class of transformations $z_1' = z_1$, $z_i' = f_i(z_i)$ for $i > 1$, where $z < f_2(z) < \cdots < f_N(z)$ and each f_i is nondecreasing. If $\mathcal{F}_0$ is the class of N-tuples $(F_1, \ldots, F_N)$ with $F_1 = \cdots = F_N$, then $\mathcal{F}_1$ coincides with the class K of alternatives.]

Problem 6.63 In the preceding problem let $U_{ij} = 1$ if $(j-i)(Z_j - Z_i) > 0$, and $= 0$ otherwise.

(i) The test statistic $\sum iT_i$, can be expressed in terms of the U's through the relation

$$\sum_{i=1}^N iT_i = \sum_{i<j}(j-i)U_{ij} + \frac{N(N+1)(N+2)}{6}.$$

(ii) The smallest number of steps [in the sense of Problem 6.42(ii)] by which $(Z_1, \ldots, Z_N)$ can be transformed into the ordered sample $(Z_{(1)}, \ldots, Z_{(N)})$ is

[7] Some tests of randomness are treated in Diaconis (1988).

$[N(N-1)/2] - U$, where $U = \sum_{i<j} U_{ij}$. This suggests $U > C$ as another rejection region for the preceding problem.

[(i): Let $V_{ij} = 1$ or 0 as $Z_i \le Z_i$ or $Z_i > Z_j$. Then $T_j = \sum_{i=1}^N V_{ij}$, and $V_{ij} = U_{ij}$ or $1 - U_{ij}$ as $i < j$ or $i \ge j$. Expressing $\sum_{j=1}^N jT_j = \sum_{j=1}^N j \sum_{i=1}^N V_{ij}$ in terms of the U's and using the fact that $U_{ij} = U_{ji}$, the result follows by a simple calculation.]

Problem 6.64 *The hypothesis of independence*. Let $(X_1, Y_1), \ldots, (X_N, Y_N)$ be a sample from a bivariate distribution, and $(X_{(1)}, Z_1), \ldots, (X_{(N)}, Z_N)$ be the same sample arranged according to increasing values of the X's so that the Z's are a permutation of the Y's. Let R_i be the rank of X_i among the X's, S_i the rank of Y_i among the Y's, and T_i the rank of Z_i among the Z's, and consider the hypothesis of independence of X and Y against the alternatives of positive regression dependence.

(i) Conditionally, given $(X_{(1)}, \ldots, X_{(N)})$, this problem is equivalent to testing the hypothesis of randomness of the Z's against the alternatives of an upward trend.

(ii) The test (6.73) is equivalent to rejecting when the *rank correlation coefficient*

$$\frac{\sum(R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum(R_i - \bar{R}^2) \sum(S_i - \bar{S})^2}} = \frac{12}{N^3 - N} \sum \left(R_i - \frac{N+1}{2}\right)\left(S_i - \frac{N+1}{2}\right)$$

is too large.

(iii) An alternative expression for the rank correlation coefficient[8] is

$$1 - \frac{6}{N^3 - N} \sum (S_i - R_i)^2 = 1 - \frac{6}{N^3 - N} \sum (T_i - i)^2.$$

(iv) The test $U > C$ of Problem 6.63(ii) is equivalent to rejecting when Kendall's t-statistic $\sum_{i<j} V_{ij}/N(N-1)$ is too large where V_{ij} is $+1$ or -1 as $(Y_j - Y_i)(X_j - X_i)$ is positive or negative.

(v) The tests (ii) and (iv) are unbiased against the alternatives of positive regression dependence.

Section 6.11

Problem 6.65 In Example 6.11.1, a family of sets $S(x, y)$ is a class of equivariant confidence sets if and only if there exists a set $\mathcal{R}$ of real numbers such that

$$S(x, y) = \bigcup_{r \in \mathcal{R}} \{(\xi, \eta) : (x - \xi)^2 + (y - \eta)^2 = r^2\}.$$

[8] For further material on these and other tests of independence, see Kendall (1970), Aiyar, Guillier, and Albers (1979), Kallenberg and Ledwina (1999).

Problem 6.66 Let $X_1, \dots, X_n; Y_1, \dots, Y_n$ be samples from $N(\xi, \sigma^2)$ and $N(\eta, \tau^2)$, respectively. Then the confidence intervals (5.42) for τ^2/σ^2, which can be written as

$$\frac{\sum(Y_j - \bar{Y})^2}{k\sum(X_i - \bar{X})^2} \le \frac{\tau^2}{\sigma^2} \le \frac{k\sum(Y_j - \bar{Y})^2}{\sum(X_i - \bar{X})^2},$$

are uniformly most accurate equivariant with respect to the smallest group G containing the transformations $X_i' = aX + b$, $Y_i' = aY + c$ for all $a \neq 0$, b, c and the transformation $X_i' = dY_i$, $Y_i' = X_i/d$ for all $d \neq 0$.
[Cf. Problem 6.12.]

Problem 6.67 (i) *One-sided equivariant confidence limits*. Let θ be real-valued, and suppose that, for each θ_0, the problem of testing $\theta \le \theta_0$ against $\theta > \theta_0$ (in the presence of nuisance parameters ϑ) remains invariant under a group G_{θ_0} and that $A(\theta_0)$ is a UMP invariant acceptance region for this hypothesis at level α. Let the associated confidence sets $S(x) = \{\theta : x \in A(\theta)\}$ be one-sided intervals $S(x) = \{\theta : \underline{\theta}(x) \le \theta\}$, and suppose they are equivariant under all G_θ and hence under the group G generated by these. Then the lower confidence limits $\underline{\theta}(X)$ are uniformly most accurate equivariant at confidence level $1 - \alpha$ in the sense of minimizing $P_{\theta,\vartheta}\{\underline{\theta}(X) \le \theta'\}$ for all $\theta' < \theta$.

(ii) Let $X_1, \dots, X_n$ be independently distributed as $N(\xi, \sigma^2)$. The upper confidence limits $\sigma^2 \le \sum(X_i - \bar{X})^2/C_0$ of Example 5.5.1 are uniformly most accurate equivariant under the group $X_i' = X_i + c$, $-\infty < c < \infty$. They are also equivariant (and hence uniformly most accurate equivariant) under the larger group $X_i' = aX_i + c$, $-\infty < a, c < \infty$.

Problem 6.68 *Counterexample*. The following example shows that the equivariance of $S(x)$ assumed in the paragraph following Lemma 6.11.1 does not follow from the other assumptions of this lemma. In Example 6.5.1, let $n = 1$, let $G^{(1)}$ be the group G of Example 6.5.1, and let $G^{(2)}$ be the corresponding group when the roles of Z and $Y = Y_1$ are reversed. For testing $H(\theta_0) : \theta = \theta_0$ against $\theta \neq \theta_0$ let G_{θ_0} be equal to $G^{(1)}$ augmented by the transformation $Y' = \theta_0 - (Y_1 - \theta_0)$ when $\theta \le 0$, and let G_{θ_0} be equal to $G^{(2)}$ augmented by the transformation $Z' = \theta_0 - (Z - \theta_0)$ when $\theta > 0$. Then there exists a UMP invariant test of $H(\theta_0)$ under G_{θ_0} for each θ_0, but the associated confidence sets $S(x)$ are not equivariant under $G = \{G_\theta, -\infty < \theta < \infty\}$.

Problem 6.69 (i) Let $X_1, \dots, X_n$ be independently distributed as $N(\xi, \sigma^2)$, and let $\theta = \xi/\sigma$. The lower confidence bounds $\underline{\theta}$ for θ, which at confidence level $1 - \alpha$ are uniformly most accurate invariant under the transformations $X_i' = aX_i$, are

$$\underline{\theta} = C^{-1}\left(\frac{\sqrt{n}\bar{X}}{\sqrt{\sum(X_i - \bar{X})^2/(n-1)}}\right),$$

where the function $C(\theta)$ is determined from a table of noncentral t so that

$$P_\theta \left\{ \frac{\sqrt{n}\bar{X}}{\sqrt{\sum(X_i - \bar{X})^2/(n-1)}} \le C(\theta) \right\} = 1 - \alpha.$$

(ii) Determine $\underline{\theta}$ when the x's are 7.6, 21.2, 15.1, 32.0, 19.7, 25.3, 29.1, 18.4 and the confidence level is $1 - \alpha = .95$.

Problem 6.70 (i) Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a sample from a bivariate normal distribution, and let

$$\underline{\rho} = C^{-1} \left(\frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}} \right),$$

where $C(\rho)$ is determined such that

$$P_\theta \left\{ \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}} \le C(\rho) \right\} = 1 - \alpha.$$

Then $\underline{\rho}$ is a lower confidence limit for the population correlation coefficient ρ at confidence level $1 - \alpha$; it is uniformly most accurate invariant with respect to the group of transformations $X_i' = aX_i + b$, $Y_i' = cY_i + d$, with $ac > 0$, $-\infty < b$, $d < \infty$.

(ii) Determine $\underline{\rho}$ at level $1 - \alpha = .95$ when the observations are (12.9,.56), (9.8,.92), (13.1,.42), (12.5,1.01), (8.7,.63), (10.7,.58), (9.3,.72), (11.4,.64).

Note. The following problems explore the relationship between pivotal quantities and equivariant confidence sets. For more details see Arnold (1984).

Let X be distributed according $P_{\theta,\vartheta}$, and consider confidence sets for θ that are equivariant under a group G^*, as in Section 6.11. If w is the set of possible θ-values, define a group $\tilde{G}$ on $\mathcal{X} \times w$ by $\tilde{g}(\theta, x) = (gx, \bar{g}\theta)$.

Problem 6.71 Let $V(X, \theta)$ be any pivotal quantity [i.e., have a fixed probability distribution independent of (θ, ϑ)], and let B be any set in the range space of V with probability $P(V \in B) = 1 - \alpha$. Then the sets $S(x)$ defined by

$$\theta \in S(x) \quad \text{if and only if} \quad V(\theta, x) \in B \tag{6.75}$$

are confidence sets for θ with confidence coefficient $1 - \alpha$.

Problem 6.72 (i) If $\tilde{G}$ is transitive over $\mathcal{X} \times w$ and $V(X, \theta)$ is maximal invariant under $\tilde{G}$, then $V(X, \theta)$ is pivotal.

(ii) By (i), any quantity $W(X, \theta)$ which is invariant under $\tilde{G}$ is pivotal; give an example showing that the converse need not be true.

Problem 6.73 Under the assumptions of the preceding problem, the confidence set $S(x)$ is equivariant under G^*.

Problem 6.74 Under the assumptions of Problem 6.72, suppose that a family of confidence sets $S(x)$ is equivariant under G^*. Then there exists a set B in the range space of the pivotal V such that (6.75) holds. In this sense, all equivariant confidence sets can be obtained from pivotals.

[Let A be the subset of $\mathcal{X} \times w$ given by $A = \{(x, \theta) : \theta \in S(x)\}$. Show that $\tilde{g}A = A$, so that any orbit of $\tilde{G}$ is either in A or in the complement of A. Let the maximal invariant $V(x, \theta)$ be represented as in Section 6.2 by a uniquely defined point on each orbit, and let B be the set of these points whose orbits are in A. Then $V(x, \theta) \in B$ if and only if $(x, \theta) \in A$.] *Note*. Problem 6.74 provides a simple check of the equivariance of confidence sets. In Example 6.12.2, for instance, the confidence sets (6.46) are based on the pivotal vector $(X_1 - \xi_1, \dots, X_r - \xi_r)$, and hence are equivariant.

Section 6.12

Problem 6.75 In Examples 6.12.1 and 6.12.2 there do not exist equivariant sets that uniformly minimize the probability of covering false values.

Problem 6.76 In Example 6.12.1, the density $p(v)$ of $V = 1/S^2$ is unimodal.

Problem 6.77 Show that in Example 6.12.1,

(i) the confidence sets $\sigma^2/S^2 \in A^{**}$ with A^{**} given by (6.45) coincide with the uniformly most accurate unbiased confidence sets for σ^2;
(ii) if (a, b) is best with respect to (6.44) for σ, then (a^r, b^r) is best for σ^r $(r > 0)$.

Problem 6.78 Let $X_1, \dots, X_r$ be i.i.d. $N(0, 1)$, and let S^2 be independent of the X's and distributed as χ^2_ν. Then the distribution of $(X_1/S\sqrt{\nu}, \dots, X_r/S\sqrt{\nu})$ is a central multivariate t-distribution, and its density is

$$p(v_1, \dots, v_r) = \frac{\Gamma(\frac{1}{2}(\nu + r))}{(\pi\nu)^{r/2}\Gamma(\nu/2)} \left(1 + \frac{1}{\nu} \sum v_i^2\right)^{-\frac{1}{2}(\nu+r)} .$$

Problem 6.79 The confidence sets (6.52) are uniformly most accurate equivariant under the group G defined at the end of Example 6.12.3.

Problem 6.80 In Example 6.12.4, show that

(i) both sets (6.60) are intervals;
(ii) the sets given by $vp(v) > C$ coincide with the intervals (5.41).

Problem 6.81 Let $X_1, \dots, X_m$; $Y_1, \dots, Y_n$ be independently normally distributed as $N(\xi, \sigma^2)$ and $N(\eta, \sigma^2)$ respectively. Determine the equivariant confidence sets for $\eta - \xi$ that have smallest Lebesgue measure when

(i) σ is known;
(ii) σ is unknown.

Problem 6.82 Generalize the confidence sets of Example 6.11.3 to the case that the X_i are $N(\xi_i, d_i\sigma^2)$ where the d's are known constants.

Problem 6.83 Solve the problem corresponding to Example 6.12.1 when

(i) $X_1, \ldots, X_n$ is a sample from the exponential density $E(\xi, \sigma)$, and the parameter being estimated is σ;
(ii) $X_1, \ldots, X_n$ is a sample from the uniform density $U(\xi, \xi + \tau)$, and the parameter being estimated is τ.

Problem 6.84 Let $X_1, \ldots, X_n$ be a sample from the exponential distribution $E(\xi, \sigma)$. With respect to the transformations $X_i' = bX_i + a$ determine the smallest equivariant confidence sets

(i) for σ, both when size is defined by Lebesgue measure and by the equivariant measure (6.44);
(ii) for ξ.

Problem 6.85 Let X_{ij} ($j = 1, \ldots, n_i;\ i = 1, \ldots, s$) be samples from the exponential distribution $E(\xi_i, \sigma)$. Determine the smallest equivariant confidence sets for $(\xi_1, \ldots, \xi_r)$ with respect to the group $X_{ij}' = bX_{ij} + a_i$.

Section 6.13

Problem 6.86 If the confidence sets $S(x)$ are equivariant under the group G, then the probability $P_\theta\{\theta \in S(X)\}$ of their covering the true value is invariant under the induced group $\bar{G}$.

Problem 6.87 Consider the problem of obtaining a (two-sided) confidence band for an unknown continuous cumulative distribution function F.

(i) Show that this problem is invariant both under strictly increasing and strictly decreasing continuous transformations $X_i' = f(X_i)$, $i = 1, \ldots, n$, and determine a maximal invariant with respect to this group.
(ii) Show that the problem is not invariant under the transformation

$$X_i' = \begin{cases} X_i & \text{if } |X_i| \geq 1, \\ X_i - 1 & \text{if } 0 < X_i < 1, \\ X_i + 1 & \text{if } -1 < X_i < 0. \end{cases}$$

[(ii): For this transformation g, the set $g^*S(x)$ is no longer a band.]

6.15 Notes

Invariance considerations were introduced for particular classes of problems by Hotelling (1936) and Pitman (1939b). The general theory of invariant and almost invariant tests, together with its principal parametric applications, was developed by Hunt and Stein (1946) in an unpublished paper. In their paper, invariance was not proposed as a desirable property in itself but as a tool for deriving most stringent tests (cf. Chapter 8). Apart from this difference in point of view, the present account is based on the ideas of Hunt and Stein, about which E. L. Lehmann learned through conversations with Charles Stein during the years 1947–1950.

Of the admissibility results of Section 6.7, Theorem 6.7.1 is due to Birnbaum (1955) and Stein (1956a); Example 6.7.3 (continued) and Lemma 6.7.1, to Kiefer and Schwartz (1965).

The problem of minimizing the volume or diameter of confidence sets is treated in DasGupta (1991).

Deuchler (1914) appears to contain the first proposal of the two-sample procedure known as the Wilcoxon test, which was later discovered independently by many different authors. A history of this test is given by Kruskal (1957). Hoeffding (1951) derives a basic rank distribution of which (6.22) is a special case, and from it obtains locally optimum tests of the type (6.23).