ADMISSIBILITY OF PROCEDURES IN TWO-DIMENSIONAL LOCATION PARAMETER PROBLEMS

By Lawrence D. Brown[1] and Martin Fox[2]

Cornell University and Michigan State University

Conditions are given for admissibility of procedures invariant under two-dimensional translation. These conditions may be applied to obtain admissibility of (i) the best invariant procedure for estimating a two-dimensional location parameter and (ii) procedures which are Bayes in the class of procedures invariant under two-dimensional translation, the prior distribution being over values of some other parameter.

1. Introduction. Brown (1966) gave general conditions for the admissibility of best invariant estimates of one-dimensional location parameters. Using his methods, Fox (1971) did the same in the case of multiple decision problems invariant under a one-dimensional location parameter. In the present paper we present a theorem combining the extension of the above results to the two-dimensional case. In the three-dimensional case Brown (1966) gave general conditions for inadmissibility of best invariant estimates while Portnoy and Stein (1971) gave an example of an inadmissible best invariant test.

Notation and assumptions are given in Section 2 while the main theorem and its proof appear in Section 3. Section 4 contains lemmas concerning verification of the conditions in the case of estimation. The application to multiple decision problems is given in Section 5 and is extended in Section 7. An indication of how the one-dimensional results of Brown (1966) may be extended along the lines of this paper is found in Brown and Fox (1972).

In the one-dimensional case one of the conditions involves finiteness of a moment of order one higher than needed for finite risk. Counterexamples when this condition is violated have been given by Perng (1970) and by Fox and Perng (1969). In the present paper we require finiteness of a moment of order two higher than needed for finite risk. Section 6 contains examples patterned after those for the one-dimensional case. Unfortunately we have been unable to verify one other condition and one example violates a condition other than the moment condition. The conditions causing these difficulties have no parallels in Brown (1966) or in Fox (1971).

2. Notation and assumptions. Let X, Y be random variables taking values in $\mathscr{R}^2 \times \mathscr{Y}$, $\mathscr{B}_{\mathscr{R}^2} \times \mathscr{B}_{\mathscr{Y}}$ where $\mathscr{R}^2$, $\mathscr{B}_{\mathscr{R}^2}$ denotes two-dimensional Euclidean space

248

with the usual Borel σ-field. Suppose the distribution of X, Y is given by

(2.1) $\Pr_\theta \{X, Y \in S\} = \iint_S p(x - \theta, y) \, dx \mu(dy)$

where $\theta \in \mathscr{R}^2$, dx denotes Lebesgue measure on $\mathscr{R}^2$, $\mathscr{B}_{\mathscr{R}^2}$, and

$$\int p(x - \theta, y) \, dx = 1 \, .$$

The parameter space is therefore $\mathscr{R}^2 = \{\theta\}$. The family of distributions is "invariant" (or, "equivariant") under translation. Y is the "maximal invariant" statistic.

We suppose the decision space $\mathscr{D}$ has a group of transformations defined on it which is a homeomorphic image of the translation group in $\mathscr{R}^2$. Thus, for each $z \in \mathscr{R}^2$ there is a transformation $g_z : \mathscr{D} \to \mathscr{D}$, and $g_{z_1 + z_2} = g_{z_1} \cdot g_{z_2}$. For example, in an estimation problem $\mathscr{D}$ might be $\mathscr{R}^2$ itself and $g_z d = d + z$; or in a multiple decision problem (see Section 5) g_z is the identity map $g_z d = d$ for all $z \in \mathscr{R}^2$.

Suppose further that the loss function is invariant; i.e. that the loss function $L : \mathscr{D} \times \mathscr{R}^2 \times \mathscr{R}^2 \times \mathscr{Y} \to [0, \infty)$ satisfies

(2.2) $L(d, \theta, x, y) = W(g_{-\theta} d, x - \theta, y)$

where, of course $W : \mathscr{D} \times \mathscr{R}^2 \times \mathscr{Y} \to [0, \infty)$.

It is also convenient to assume that $\mathscr{D}$ is a measurable space with σ-field, $\mathscr{B}_{\mathscr{D}}$, W is a measurable map with respect to $\mathscr{B}_{\mathscr{D}} \times \mathscr{B}_{\mathscr{R}^2} \times \mathscr{B}_{\mathscr{Y}}$ and that the map $(d, \theta, z) \to (g_z d, \theta + z)$ is measurable from $\mathscr{B}_{\mathscr{D}} \times \mathscr{B}_{\mathscr{R}^2} \times \mathscr{B}_{\mathscr{R}^2}$ to $\mathscr{B}_{\mathscr{D}} \times \mathscr{B}_{\mathscr{R}^2}$. This last fact implies in particular that the maps $d \to g_z d$ and $(d, \theta) \to g_{-\theta} d$ are measurable.

Certain aspects involving invariant statistical problems with a parameter space larger than $\mathscr{R}^2$ can also be treated using the results of Section 3. This will be done in Section 5 for the specific example of an invariant multiple decision procedure. The general situation will be discussed in Section 7. (But no theorem will be proved there!)

A decision procedure is a measurable map

$$\delta : \mathscr{R}^2 \times \mathscr{Y} \to \mathscr{D} \, .$$

Its risk function is

$$R(\theta, \delta) = E_\theta(W(g_{-\theta} \delta(X, Y), X - \theta, Y) \, .$$

Formally, then, all decision procedures are non-randomized. However, if $\mathscr{D}$ is a locally compact, σ-compact topological space and $\mathscr{B}_{\mathscr{D}}$ is the Borel field on $\mathscr{D}$ then we have not lost any statistical generality in using this more convenient formulation rather than a formulation allowing the use of randomized estimators. This follows from Wald and Wolfowitz (1950).

A procedure is invariant if

$$\delta(x + z, y) = g_z \delta(x, y) \, .$$

Since W is invariant (2.2) the risk of any invariant procedure is a constant. Let

250   LAWRENCE D. BROWN AND MARTIN FOX

$\mathscr{I}$ denote the class of invariant procedures. Note the standard fact that if $\delta \in \mathscr{I}$ then $\delta(x, y) = g_x \delta(0, y)$. Hence the function

$$\gamma(y) = \delta(0, y)$$

determines δ with the aid of the maps g_x.

We now formulate the assumptions which will be required for the admissibility theorem of the next section. These assumptions will be only briefly discussed here. More detailed discussion of the latter assumptions will be left for Sections 4 and 5 which deal with specific types of location problems. The first three assumptions to follow are closely related to Assumptions 2.1.1, 2.1.3, and 2.1.2 respectively of Brown (1966).

ASSUMPTION 1. There exists a procedure $\delta_0 \in \mathscr{I}$ such that

$$R_0 = R(\theta, \delta_0) = \inf_{\delta \in \mathscr{I}} R(\theta, \delta) < \infty .$$

δ_0 is called the *best invariant procedure*.

The existence of δ_0 follows from standard statistical results involving very mild regularity conditions. That essential uniqueness, implied by a later regularity condition (see Lemma 2.1), is needed for any general admissibility result can be inferred from analogous examples involving one-dimensional location problems; see Farrell (1964) and Perng (1970).

ASSUMPTION 2.

(2.3) $\int \mu(dy) \int ||x||^2 W(\delta_0(x, y), x, y) p(x, y)\, dx < \infty .$

Again, by analogy with one-dimensional problems one would expect a moment condition to be needed for any general admissibility result. Further, we conjecture that the exponent 2 in (2.3) is the smallest exponent which yields a general admissibility theorem like Theorem 3.1. The examples of Section 6 give partial support to this conjecture.

ASSUMPTION 3. If δ is any procedure such that $R(\theta, \delta) \leq R_0$ for all $\theta \in \mathscr{A}^{2}$, then there exist procedures δ_L $(L \to \infty)$ with the following properties:

$$R(\theta, \delta_L) \leq R(\theta, \delta) \qquad \text{for} \quad ||\theta|| \leq L ,$$

and for $L < M$ and $||x|| < 3L/2$ we have

(2.4) $\int W(g_{-(x-z)} \delta_L(x, y), z, y) p(z, y)\, dz \leq \int W(g_{-(x-z)} \delta_M(x, y), z, y) p(z, y)\, dz ,$

and

(2.5) $\int \mu(dy) \int_{||x||<L/2} dx \int_{||\theta|| \geq L} W(g_{-\theta} \delta_L(x, y), x - \theta, y) p(x - \theta, y)\, d\theta \to 0$

as $L \to \infty$.

The hypothesis of this assumption is peculiar. It begins, in essence, with the negation of the conclusion of Theorem 3.1. Nevertheless, Assumption 3 need not be difficult to verify. It is a condition on the tails of p. For, if for some t' we have $\int \mu(dy) \int_{||x||>t'} p(x, y)\, dx = 0$, then we may choose $\delta_L = \delta$ and then for

L sufficiently large ($L/2 \geq t'$) the left-hand side of (2.5) is zero, and so Assumption 3 is satisfied.

Also, if $W \leq B$ and the moment condition

$$(2.6) \qquad \int \mu(dy) \int ||x||^2 p(x, y)\, dx < \infty$$

is satisfied, then letting $\delta_L = \delta$ we have

$$
\begin{aligned}
(2.7) \qquad & \int_{||\theta|| \geq L} d\theta \int \mu(dy) \int_{||x|| < L/2} W(g_{-\theta}\delta(x, y), x - \theta, y)p(x - \theta, y)\, dx \\
& \leq B \int_{||\theta|| \geq L} d\theta \int \mu(dy) \int_{||z|| > ||\theta|| - L/2} p(z, y)\, dz \\
& \leq B \int \mu(dy) \int_{||z|| \geq L/2} \left(\int_{||\theta|| < L/2 + ||z||} d\theta \right) p(z, y)\, dz \\
& \leq 4\pi B \int \mu(dy) \int_{||z|| \geq L/2} ||z||^2 p(z, y)\, dz \\
& = o(1) \qquad\qquad\qquad\qquad\qquad\qquad \text{as } L \to \infty.
\end{aligned}
$$

Thus Assumption 3 is satisfied in this case also.

The last assumption has no analog in Brown (1966) and is qualitatively different from the previous conditions. It is not mainly a condition on the tails of p and/or the growth of W. Rather it is a kind of condition on the local behavior of $W(\cdot, x, y)$ near $\delta_0(x, y)$.

ASSUMPTION 4. There is a non-increasing function $k : (0, \infty) \to (0, \infty)$ such that

$$(2.8) \qquad \int_0^\infty k(v)\, dv = k_5 < \infty,$$

and such that for any $\delta \in \mathscr{S}$ inequalities (2.9) and (2.10) below are satisfied:

$$
\begin{aligned}
(2.9) \qquad & \int \mu(dy) \int_{||x|| < v} (W(\delta_0(x, y), x, y) - W(\delta(x, y), x, y))p(x, y)\, dx \\
& \leq k(v)(\int \mu(dy) \int (W(\delta(x, y), x, y) - W(\delta_0(x, y), x, y))p(x, y)\, dx)^{\frac{1}{2}}
\end{aligned}
$$

and

$$
\begin{aligned}
(2.10) \qquad & \int \mu(dy) \int_{||x|| > v} [W(\delta_0(x, y), x, y) - W(\delta(x, y), x, y)]^+ p(x, y)\, dx \\
& \leq k(v)(\int \mu(dy) \int (W(\delta(x, y), x, y) - W(\delta_0(x, y), x, y))p(x, y)\, dx)^{\frac{1}{2}}.
\end{aligned}
$$

The following trivial lemma implies a form of essential uniqueness of δ_0. For estimation, essential uniqueness is implied in the conclusion of this lemma if, as in Farrell's (1964) case, $W(d, x, y)$ is strictly increasing as d moves away from 0 in any direction.

LEMMA 2.1. *Let δ_0 be any best invariant procedure satisfying (2.10). If δ is also best invariant, then*

$$(2.11) \qquad W(\delta_0(x, y), x, y)p(x, y) = W(\delta(x, y), x, y)p(x, y) \quad \text{a.e.} \quad (dx\mu(dy)).$$

Assumption 4 is usually somewhat more difficult to verify than are the preceding assumptions. Merely to indicate to the reader that it may be satisfied we consider below the case where $\mathscr{D} = \mathscr{R}^2$, $W(d, x, y) = ||d||^2$ (i.e., $L(d, \theta, x, y) = ||d - \theta||^2$), and (for simplicity) $E_0(x) = 0$ so that $\delta_0(x) = x$.

We will need a slightly more stringent condition than Assumption 2; namely,

for some $\varepsilon > 0$

(2.12) $\int \mu(dy) \int \|x\|^2 (\ln^{2+\varepsilon} \|x\|) W(x, x, y) p(x, y) \, dx < \infty$.

With this assumption and W as above

(2.13) $k(v) = 2(\int \mu(dy) \int_{\|x\| > v} \|x\|^2 p(x, y) \, dx)^{\frac{1}{2}}$

satisfies (2.8). Clearly, $\int_0^1 k(v) \, dv < \infty$ and, for $v \geqq 1$, we have $k(v) \leqq$ a $v^{-1} \ln^{-(1+\varepsilon/2)} v$ so that $\int_1^\infty k(v) \, dv < \infty$. Under these assumptions, the left-hand side of (2.9) satisfies

$$\int \mu(dy) \int_{\|x\| < v} (\|x\|^2 - \|x + \gamma(y)\|^2) p(x, y) \, dx$$

$$\leqq \int \mu(dy) \int_{\|x\| < v} -2x \cdot \gamma(y) p(x, y) \, dx$$

(2.14) $= \int \mu(dy) \int_{\|x\| > v} 2x \cdot \gamma(y) p(x, y) \, dx$

$$\leqq 2(\int \mu(dy) \int_{\|x\| > v} \|x\|^2 p(x, y) \, dx)^{\frac{1}{2}} (\int \mu(dy) \int \|\gamma(y)\|^2 p(x, y) \, dx)^{\frac{1}{2}}$$

$$= k(v)(\int \mu(dy) \int (W(\delta(x, y), x, y) - W(\delta_0(x, y), x, y)) p(x, y) \, dx)^{\frac{1}{2}} .$$

This verifies (2.9). Verification of (2.10) is similar since the left-hand side of (2.10) is dominated by

$$\int \mu(dy) \int_{\|x\| > v} |2x \cdot \gamma(y)| p(x, y) \, dx .$$

Note that the above argument strongly uses the special form $W(d, x, y) = \|d\|^2$. Thus it is not to be expected—and is not the case—that the moment condition (2.12) or any other moment condition suffices by itself to check Assumption 4 in more general cases. A similar slight strengthening of Assumption 2 is required in Stein (1959) and James and Stein (1961).

3. The admissibility theorem.

THEOREM 3.1. *In the problem of Section 2 suppose Assumptions 1—4 are satisfied. Then the best invariant estimator δ_0 is admissible.*

PROOF. Suppose δ_0 is inadmissible. Then there is a procedure, δ, which is better than δ_0. $R(\theta, \delta) \leqq R_0$. As in (2.1.5)—(2.1.7) of Brown (1966) we may interchange orders of integration and substitute $z = x - \theta$ for θ to yield

$$0 \leqq \int_{\|\theta\| < L} [R_0 - R(\theta, \delta_L)] \, d\theta$$

$$= \int \mu(dy) \int_{\|x\| < L/2} dx \int_{\|z - x\| < L} \bar{w}_L(z, x, y) p(z, y) \, dz$$

(3.1) $+ \int \mu(dy) \int_{L/2 \leqq \|x\| < 3L/2} dx \int_{\|z - x\| < L} \bar{w}_L(z, x, y) p(z, y) \, dz$

$$+ \int \mu(dy) \int_{\|x\| \geqq 3L/2} dx \int_{\|z - x\| < L} \bar{w}_L(z, x, y) p(z, y) \, dz$$

$$= I_1(L) + I_2(L) + I_3(L)$$

where in view of the invariance of W and using (2.2) we set

$$\bar{w}_L(z, x, y) = W(\delta_0(z, y), z, y) - W(g_{-(x-z)} \delta_L(x, y), z, y)$$

and $I_i(L)$ are defined in the above expression.

In view of Assumption 2, the quantity $I_3(L)$ satisfies

$$
\begin{aligned}
(3.2) \qquad I_3(L) &\leq \int \mu(dy) \int_{||z||>L/2} \int_{||x-z||\leq L} dx\, W(\delta_0(z, y), z, y)p(z, y)\, dz \\
&\leq \pi \int \mu(dy) \int_{||z||>L/2} 4||z||^2 W(\delta_0(z, y), z, y)p(z, y)\, dz \\
&= o(1) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{as } L \to \infty .
\end{aligned}
$$

From Assumptions 4 and 2 it follows that

$$
\begin{aligned}
(3.3) \qquad h(L) &= \int \mu(dy) \int_{||x||<L/2} dx \int_{||z-x||\geq L} \bar{w}_L(z, x, y)p(z, y)\, dz \\
&= o(1) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{as } L \to \infty .
\end{aligned}
$$

Furthermore, I_2 satisfies

$$
(3.4) \qquad I_2(L) \leq \int_{L/2\leq||x||\leq 3L/2} R_0\, dx = O(L^2) .
$$

Subtract $h(L)$ from both sides of formula (3.1). Transpose the first integral to the left side of the inequality and apply the above estimates to the terms which remain on the right. The result is

$$
\begin{aligned}
(3.5) \qquad I(L) &= \int \mu(dy) \int_{||x||<L/2} dx \int -\bar{w}_L(z, x, y)p(z, y)\, dz \\
&= \int \mu(dy) \int_{||x||<L/2} dx \int_{||z-x||<L} -\bar{w}_L(z, x, y)p(z, y)\, dz - h(L) \\
&\leq I_2(L) - h(L) + o(1) = O(L^2) \qquad\qquad\qquad \text{as } L \to \infty .
\end{aligned}
$$

For x fixed the procedure $g_{-(x-z)}\delta_L(x, y)$ is invariant. Thus, since δ_0 is the best invariant procedure, $I(L)$ is nonnegative. It is non-decreasing in L because δ_0 is best invariant and because of (2.4) of Assumption 3.

We now use the crucial Assumption 4, and then the Cauchy–Schwarz inequality to obtain

$$
\begin{aligned}
A^{-1} \int_A^{2A} I_2(L)\, dL &\leq A^{-1} \int_A^{2A} dL \int_{L/2\leq||x||\leq 3L/2} dx \\
&\quad \times \{(\int \mu(dy) \int_{||z||<|L-||x|||} \bar{w}_L(z, x, y)p(z, y)\, dz)^+ \\
&\quad + (\int \mu(dy) \int_{||z||>|L-||x|||} \bar{w}_L^+(z, x, y)p(z, y)\, dz)\} \\
&\leq 2A^{-1} \int_A^{2A} dL \int_{L/2\leq||x||\leq 3L/2} dx\, k(|L-||x|||) \\
&\quad \times (\int \mu(dy) \int -\bar{w}_L(z, x, y)p(z, y)\, dz)^{\frac{1}{2}} \\
&\leq 2A^{-1} \int_{A/2\leq||x||\leq 3A} dx\, (\int \mu(dy) \int -\bar{w}_{6A}(z, x, y)p(z, y)\, dz)^{\frac{1}{2}} \\
&\quad \times \int_{\max\{A,2||x||/3\}}^{\min\{2A,2||x||\}} k(|L-||x|||)\, dL \\
&\leq k_6(\int_{A/2\leq||x||\leq 3A} dx \int \mu(dy) \int -\bar{w}_{6A}(z, x, y)p(z, y)\, dz)^{\frac{1}{2}} \\
&\leq k_6(I(6A) - I(A))^{\frac{1}{2}}
\end{aligned}
$$

where

$$
k_6 = 2A^{-1}\{\int_{A/2\leq||x||\leq 3A} dx\, [\int_{\max\{A,2||x||/3\}}^{\min\{2A,2||x||\}} k(|L-||x|||)\, dL]^2\}^{\frac{1}{2}}
$$

which is finite by Assumption 4.

(Note that the next to last step of the above argument fails in $\mathscr{R}^3$ and higher dimensions.)

Thus, for any A there is an L with $A \leq L \leq 2A$ such that

$$
(3.7) \qquad I(6A) - I(A) \geq (I_2(L))^2/k_6^2 .
$$

It follows from (3.5), (3.7) and the monotonicity of $I(\cdot)$ that

$$(3.8) \qquad\qquad I(6L) - I(L/2) \geqq I^2(L)/k_6{}^2 + o(1)I(L) \, .$$

Suppose I is unbounded. By (3.8), for L_0 sufficiently large and $L \geqq L_0$, $I(6L)/I(L) \geqq 6^3$. By induction, for $n = 1, 2, \cdots$ we obtain $I(6^n L_0) \geqq 6^{3n} I(L_0)$. Since $I(L) = O(L^2)$, we obtain a contradiction by dividing both sides by $(6^n L_0)^2$ and letting $n \to \infty$. Thus, I is bounded.

Boundedness and monotonicity of I imply that the left-hand side of (3.8) converges to zero. Thus, $I(L) \equiv 0$ which implies, by Assumption 1, that $\delta = \delta_0$ a.e. $(dx\mu(dy))$ and, hence, that $R(\theta, \delta) \equiv R_0$. Thus, δ_0 is admissible and the theorem is proved.

REMARK. Even if the distributions of X given Y do not possess densities with respect to Lebesgue measure, the above argument can still be applied. Suppose the distribution of X given Y is

$$P_\theta(A \mid y) = P_0(A - \theta \mid y)$$

where $A - \theta = \{x : x + \theta \in A\}$ and P_0 is a measurable conditional probability measure (i.e. $P(A \mid \cdot)$ is measurable). Suppose the obvious analogs of Assumptions 1—4 are valid. Then the above proof, with the obvious notational changes, remains valid to show that the best invariant procedure will be almost admissible $(d\theta)$ but may not be admissible.

REMARK. In a similar manner a theorem can be proved if only an almost everywhere version of Assumption 4 is valid. To be precise, suppose (2.9) reads: For any procedure δ, for almost all x (dx) we have

$$\int \mu(dy) \int_{\|z\| < v} [W(\delta_0(z, y), z, y) - W(g_{-(x-z)}\delta(x, y), z, y)]p(z, y) \, dz$$

$$\leqq k(v)(\int \mu(dy) \int [W(g_{-(x-z)}\delta(x, y), z, y) - W(\delta_0(z, y), z, y)]p(z, y) \, dz)^{\frac{1}{2}} \, .$$

A similar expression replaces (2.10). Then the argument in the proof of Theorem 3.1 yields the conclusion that δ_0 is almost admissible. We note that the assumptions of this remark are applicable to certain problems of "estimation" by (possibly randomized) confidence sets. The conclusion in these cases is that the best invariant confidence procedure is almost admissible. Joshi (1970) has termed this "weak admissibility."

4. Estimation. In the formulation of Section 2 suppose $\mathscr{D} = \mathscr{R}^2$, $g_z d = d + z$, and the loss function is simply a function of the argument $d - \theta$ only. For simplicity we write

$$W(d - \theta) = L(d, \theta, x, y) = W(g_{-\theta}d, x - \theta, y)$$

which defines the function $W(\cdot)$ of one argument. Assume $\sup_{\|t\| < K} W(t) < \infty$ for all $K < \infty$. There is no loss of generality in assuming the variables have been modified so that a best invariant estimator (assumed to exist) is given by

$$(4.1) \qquad\qquad \delta_0(x, y) = x \, .$$

In this section we describe how Theorem 3.1 applies to such an estimation problem. Primarily, we give more easily interpreted conditions which imply Assumptions 3 and 4. In that sense the results of this section are a straightforward analytic exercise; the results of the next section are much more interesting statistically.

In the estimation context of this section the validity of Assumption 3 is also often not hard to check. We have already given conditions which imply Assumption 3 if either W is bounded or if p has compact support. The following general statement and example may help in other situations.

Condition C. For each $y \in \mathscr{Y}$, there exists an increasing function $B_y : R \to R$ with $B_y(L) > L$ such that

(4.2) $\|t_1\| \leq L$, $\|t_2\| \geq B_y(L) - L$ implies $W(t_1) \leq W(t_2)$,

and

(4.3) $\|t_1\| \leq 3L/2$, $\|t_2\| \geq B_y(L) - 3L/2$

implies $\int W(z + t_1)p(x, y)\, dz \leq \int W(z + t_2)p(z, y)\, dz$

for all $y \in \mathscr{Y}$, and such that

(4.4) $\lim_{L \to \infty} L^2 \int \mu(dy) \int_{\|z\| > L/2} \sup\{W(z+t) : \|t\| < 3L/2 + B_y(L)\}p(z, y)\, dz = 0$.

Before showing that Condition C implies Assumption 3 we will verify that it is satisfied in a standard class of examples. Suppose $W(t) = \|t\|^k$, $0 < k < \infty$, and Assumption 2 is satisfied. Condition C is also satisfied. If $k \geq 1$, choose $B_y(L) = 4R_0^{1/k}(y) + 9L$ where $R_0(y) = \int W(\delta_0(x, y))p(x, y)\, dx$. Since $B(L) > L$ and $W(t)$ is an increasing function of $\|t\|$, we see that (4.2) is trivial. Also (4.4) is easy since $W(z + (3L/2 + B(L))^k) < (22 + 4R_0^{1/k})^k W(z)$ for $\|z\| > L/2$ and $L > 2$. Assumption 2 then implies (4.4). For (4.3) when $k \geq 1$ note that $\|z + t_1\|^k \leq 2^{k-1}\|z\|^k + 2^{k-1}\|t_1\|^k$ and, by a Chebyshev type inequality,

$\mathrm{Pr}_0\{\|X\| < (B_y(L) - 3L/2)/2 \,|\, Y = y\} \geq 1 - R_0(y)/(B_y(L)/2 - 3L/4)^k$.

Thus, for $\|t_1\| \leq 3L/2$ and $\|t_2\| > B_y(L) - 3L/2$ we obtain

$$\int W(z + t_1)p(z, y)\, dy \leq 2^{k-1}R_0(y) + 2^{k-1}(3L/2)^k$$
$$< \tfrac{1}{2}((B_y(L) - 3L/2)/2)^k$$
$$\leq \int_{\|z\| < (B_y(L)-3L/2)/2} W(z + t_2)p(z, y)\, dz.$$

This implies that (4.3) is satisfied. For $0 < k < 1$ a slightly different choice of B_y is needed, but the argument is analogous to the above.

LEMMA 4.1.   *Condition* C *implies Assumption* 3.

PROOF.   Given $\bar{\delta}$ define

(4.5) $\delta_L(x, y) = \delta(x, y)$ if $\|\delta(x, y)\| \leq B_y(L)$

$= 0$ if $\|\delta(x, y)\| > B_y(L)$.

Now (4.2) of Condition C immediately implies that $W(\delta_L(x, y) - \theta) \leq W(\delta(x, y) - \theta)$ for $||\theta|| \leq L$. Hence $R(\theta, \delta_L) \leq R(\theta, \delta)$ for $||\theta|| \leq L$. Also (4.3) is precisely (2.4). For (2.5) note that $||x|| \leq L/2$ and $||\theta|| \geq L$ implies that $||z|| \geq L/2$ where $z = x - \theta$ and also that $W(\delta_L(x, y) - \theta) = W(z + \gamma_L(x, y))$ where $\gamma_L(x, y) = \delta_L(x, y) - x$. From (4.5) it follows that $||\gamma_L(x, y)|| \leq 3L/2 + B_y(L)$. It is then easy to check that (4.4) implies (2.5). This completes the proof.

It seems a little harder to give a precise set of easily usable conditions which imply Assumption 4. In Section 2 we gave a condition which implies Assumption 4 when $W(t) = ||t||^2$ (squared error loss). We describe below a fairly general method of checking Assumption 4 which generalizes the considerations in Section 2. This will be followed by some more specific examples.

For $m \in (0, \infty)$, $v \in (0, \infty)$ let

$$m^2 C^2(m, y) \leq \inf \{\int (W(x + t) - W(x))p(x, y)\, dx : ||t|| = m\},$$

$$mD_v(m, y) \geq \sup \{\int_{||x|| < w} (W(x) - W(x + t))p(x, y)\, dx : w \geq v, ||t|| = m\},$$

$$mE_v(m, y) \geq \sup \{\int_{||x|| > v} (W(x) - W(x + t))^+ p(x, y)\, dx : ||t|| = m\},$$

$$k(v, y) = \sup \left\{\max \left(\frac{D_v(m, y)}{C(m, y)}, \frac{E_v(m, y)}{C(m, y)}\right) : 0 < m < \infty\right\},$$

$$k(v) = (\int k^2(v, y)\mu(dy))^{\frac{1}{2}}.$$

LEMMA 4.2. *If $\int_0^\infty k(v)\, dv < \infty$ then Assumption 4 is satisfied.*

PROOF. The hypothesis of the lemma guarantees that k satisfies condition (2.8) of Assumption 4. The above definition of k implies that k is non-increasing. It remains to verify (2.9) and (2.10). Consider (2.9). Since $\delta \in \mathscr{S}$, we have $\delta(x, y) = x + \gamma(y)$. By assumption $\delta_0(x, y) = x$ so

$$k(v)(\int \mu(dy) \int (W(x + \gamma(y)) - W(x))p(x, y)\, dx)^{\frac{1}{2}}$$

$$\geq k(v)(\int ||\gamma(y)||^2 C^2(||\gamma(y)||, y)\mu(dy))^{\frac{1}{2}}$$

(4.6)
$$\geq k(v) \left(\int \frac{D_v^2(||\gamma(y)||, y)}{C^2(||\gamma(y)||, y)} \mu(dy)\right)^{-\frac{1}{2}} \int ||\gamma(y)||D_v(||\gamma(y)||, y)\mu(dy)$$

$$\geq \int ||\gamma(y)||D_v(||\gamma(y)||, y)\mu(dy)$$

$$\geq \int \mu(dy) \int_{||x|| < v} (W(x) - W(x + \gamma(y)))p(x, y)\, dx.$$

(In the third line of (4.6) we have used the Cauchy–Schwarz inequality and in the fourth line the definition of k.) This verifies (2.9) while (2.10) is verified by a similar sequence of steps. The proof is complete.

The following is a moderately general and reasonably convenient application of Lemma 4.2 to the case of convex loss functions. If M is a symmetric 2×2 matrix, let $>M<$ denote the smallest eigenvalue of M. For a real valued function M and $\mathscr{R}^2$, let $M'(t) = ((\partial/\partial t_1)M(t), (\partial/\partial t_2)M(t))^\tau$ and

$$M''(t) = \begin{bmatrix} \dfrac{\partial^2}{\partial t_1^2} M(t) & \dfrac{\partial^2}{\partial t_1 \partial t_2} M(t) \\[4mm] \dfrac{\partial^2}{\partial t_2 \partial t_1} M(t) & \dfrac{\partial^2}{\partial t^2} M(t) \end{bmatrix}.$$

EXAMPLE 4.1. Suppose W is convex, and possesses continuous second partial derivatives, and

$$(\int \mu(dy) \int w(x + t)p(x, y) \, dx)'|_{t=0} = \int \mu(dy) \int W'(x)p(x, y) \, dx = 0 \ .$$

Let

$$\lambda(m) = \inf \{> W''(x)< : \, ||x|| \geqq m\} \ .$$

Note that $\lambda \geqq 0$. Suppose there is an $\alpha > 0$ such that for every $t \in \mathscr{P}^2$ and $y \in \mathscr{Y}$ we have

(4.7) $$\int_{t^\tau \cdot x \geqq 0} \lambda(||x||)p(x, y) \, dx \geqq \alpha > 0 \ .$$

(Note that (4.7) is satisfied if $\lambda(m) \geqq \alpha$ for all m as is the case when $W(x) = ||x||^2$. Also, if $W(x) = ||x||^k \ (k > 2)$, then $\lambda(m)$ is an increasing nonnegative function of m. Hence (4.7) is satisfied if

$$\int_{t^\tau \cdot x \geqq 0; ||x|| > \alpha_1 > 0} p(x, y) \, dx > \alpha_2 > 0$$

for $t \in \mathscr{P}^2$, $y \in \mathscr{Y}$, which is the case in a wide variety of examples.) Then Assumption 4 is satisfied if

(4.8) $$\int_0^\infty (\int (\int_{||x|| > v} ||W'(x)||p(x, y) \, dx)^2 \mu(dy))^{\frac{1}{2}} \, dv < \infty \ .$$

(If $W(x) = ||x||^k$, $k \geqq 2$ (4.8) is satisfied if

$$\int \mu(dy) \int ||x||^{2k}(ln^{2+\epsilon}||x||)p(x, y) \, dx < \infty \ .$$

Note that if $k > 2$ this condition is more stringent than (2.12).)

Since W is convex $W(x) - W(x + \gamma) \leqq -\gamma^\tau W'(x)$. Define

$$||\gamma||D_v(||\gamma||, y) = ||\gamma||(\int_{||x|| > v} ||W'(x)||p(x, y) \, dx)$$
$$\geqq \gamma^\tau \int_{||x|| > v} W'(x)p(x, y) \, dx$$
$$\geqq \int_{||x|| < v} (W(x) - W(x + \gamma))p(x, y) \, dx \ .$$

By a fairly similar line of reasoning it suffices to choose $E_v = D_v$.

A Taylor expansion of W yields $W(x + \gamma) - W(x) = \gamma^\tau W'(x) + \gamma^\tau W''(\bar{x})\gamma$ where $\bar{x}$ is on the line segment joining x and $x + \gamma$. Let $\hat{x}$ denote the point on this segment which is closest to 0. Then $W(x + \gamma) - W(x) \geqq \gamma^\tau W'(x) + ||\gamma||^2 \lambda(||\hat{x}||)$. If $\gamma^\tau x \geqq 0$, then $\hat{x} = x$. We have

$$\int (W(x + \gamma) - W(x))p(x, y) \, dx \geqq \gamma^\tau \int W'(x)p(x, y) \, dx + ||\gamma||^2 \int \lambda(||\hat{x}||) \, dx$$
$$\geqq ||\gamma||^2 \int_{\gamma^\tau x \geqq 0} \lambda(||x||) \, dx$$
$$\geqq ||\gamma||^2 \alpha \ .$$

Hence we may choose $C^2(\gamma, y) = \alpha$. Using the above definitions for C, D, E in Lemma 4.2 verifies that the condition (4.8) is indeed sufficient to imply that Assumption 4 is satisfied.

The methods of the preceding example may also be applied in other situations when W is convex, but we will not pursue the matter. In the following example we take p to be a normal density and W spherically symmetric, but not necessarily convex. The choice of normality is mainly for convenience, to guarantee

existence and ease of computability of certain derivatives. Many other smooth spherically symmetric densities could be used in place of the normal density, but again we will not pursue the matter further here.

EXAMPLE 4.2. Suppose the distribution of X is normal with mean θ and co-variance matrix I and suppose $W(t) = V(||t||)$ where V is non-decreasing and not identically constant. (The variable Y is superfluous and we omit it from the notation.) Then Assumption 4 is satisfied if $\int ||x||^4 W(x + t)p(x) dx < \infty$ for all $t \in \mathscr{N}^{2}$. This is certainly true if $V(r) = O(e^{kr})$ as $r \to \infty$, for some $k < \infty$.

Let $t = (a, 0)$, $a \geq 0$, and $h(a) = \int (W(x + t) - W(x))p(x) dx$. Computation yields

$$h'(a) = \int x_1 W(x + t)p(x) dx$$

and

$$h''(a) = \int (x_1^2 - 1)W(x + t)p(x) dx$$

for all a. Note that h' and h'' are continuous.

By symmetry $h'(0) = 0$. Also $h'(a) > 0$ for $a > 0$. Since $(x_1^2 - 1)$ and $W(x_1)$ are non-decreasing functions of x_1^2 and since $\int (x_1^2 - 1)p(x) dx = 0$ it follows that $h''(0) > 0$. Hence $h''(a) \geq \varepsilon_1$ for $a \leq \varepsilon_2$ for some $\varepsilon_1 > 0$, $\varepsilon_2 > 0$. Note that there exists an $\varepsilon_3 > 0$ such that $\int x_1 W(x + t)p(x) dx \geq \varepsilon_3 > 0$ for $a > \varepsilon_2$. Because of spherical symmetry we may set $a^2 C^2(a) = h(a)$, and we have

(4.9) $a^2 C^2(a) \geq \varepsilon_1 a^2$ if $0 < a \leq \varepsilon_2$

$= \varepsilon_1 \varepsilon_2^2 + \varepsilon_3 a$ if $a > \varepsilon_2$.

Another computation yields

$$\frac{d}{da} \int_{||x||>v} (W(x) - W(x + t))^+ p(x) dx$$

$$= \frac{d}{da} \int_{||x||>v : x_1 < -a/2} (W(x) - W(x + t))p(x) dx$$

(4.10) $= - \int_{||x||>v : x_1 < -a/2} x_1 W(x + t)p(x) dx$

$\leq - \int_{||x||>v : x_1 < -a/2} x_1 W(x)p(x) dx$

$\leq \min (N_1, N_4/v^3, 8N_4/a^3)$

where

$$N_k = \int ||x||^k W(x)p(x) dx < \infty \qquad \text{for } k = 1, 4 .$$

Integrating (4.10), we may take

$$aE_v(a) = \min (N_1 a, N_3 a/v^3) \qquad \text{if } a^3 \leq \max \left(\frac{8N_4}{N_1}, 8v^3\right)$$

(4.11) $= 3N_1^{\frac{2}{3}} N_4^{\frac{1}{3}} - 4N_4/a^2$ if $a^3 > \dfrac{8N_4}{N_1}$ and $N_1 \leq N_4/v^3$

$= 3N_4/v^2 - 4N_4/a^2$ if $a > 2v$ and $N_1 > N_4/v^3$.

From the spherical symmetry and non-decreasing property of V it follows that

$$\int_{||x||<v} (W(x) - W(x + t))p(x) dx \leq 0 .$$

Hence we may choose $D_v \equiv 0$. Computing from (4.8) and (4.11) then yields that $k(v) = \sup_{0<a<\infty} (E_v(a)/C(a))$ is bounded and satisfies $k(v) = O(1/v^2)$ as $v \to \infty$.

5. Multiple decisions. It is clear (see Section 2) how Theorem 3.1 may be applied to the problem of estimating θ. We now consider the application to multiple decision problems in the presence of a nuisance location parameter.

In place of (2.1) we assume the distribution of X, Y is given by

(5.1) $$\Pr_{\theta, j}\{X, Y \in S\} = \int_S \int f_j(x - \theta, y)\, dx \nu_j(dy)$$

where now the unknown parameter is (θ, j) with θ as in Section 2, $j = 1, \cdots, r$, each $\int f_j(x - \theta, y)\, dx = 1$ and each ν_j is a probability measure. In the invariant multiple decision problem, we have $\mathscr{D} = \{1, \cdots, m\}$ and loss of the form $V(j, d, y)$ when the true parameter value is (θ, j) and decision $d \in \mathscr{D}$ is made.

Contrary to the case of estimation, invariant procedures now depend only on Y.

Let $\xi = (\xi_1, \cdots, \xi_r)$ be a prior distribution on the index j. Without loss of generality assume each $\xi_j > 0$. Set

(5.2) $$L(d, \theta, x, y) = \frac{\sum_j \xi_j V(j, d, y) f_j(x - \theta, y) \dfrac{d\nu_j}{d\mu}(y)}{\sum_j \xi_j f_j(x - \theta, y) \dfrac{d\nu_j}{d\mu}(y)}$$

where $\mu = \sum_j \xi_j \nu_j$ and

(5.3) $$p(x - \theta, y) = \sum_j \xi_j f_j(x - \theta, y) \frac{d\nu_j}{d\mu}(y).$$

Then, using (5.1), (5.2) and (5.3) yields

$$\Pr_\theta\{X, Y \in S\} = \int_S \int p(x - \theta, y)\, dx \mu(dy)$$

as in (2.1). Furthermore, $\int p(x - \theta, y)\, dx = 1$ and

(5.4) $$R(\theta, \delta) = \sum_j \xi_j \int \nu_j(dy) \int V(j, \delta(x, y), y) f_j(x - \theta, y)\, dx$$

which is the usual Bayes risk for fixed θ.

The procedure δ_0 of Assumption 1 is the procedure which is Bayes in the class of invariant procedures. In the case of hypothesis testing ($r = m = 2$ and $V(j, d, y) = 0$ when $j = d$) if V is independent of y, then δ_0 is the best invariant test of some size.

We now modify Assumptions 1 to 4 to fit the multiple decision problem.

ASSUMPTION 1'. The procedure which is Bayes in the class of invariant procedures exists.

ASSUMPTION 2'. For all $j = 1, \cdots, r$, we have

$$\int \nu_j(dy) \int ||x||^2 V(j, \delta_0(x, y), y) f_j(x, y)\, dx < \infty.$$

When V does not depend on y this assumption is equivalent to $E_{0j}||X||^2 < \infty$. As in (5.4) Assumptions 2 and 2' are equivalent.

When V is independent of y, it is clear that W is bounded. In this case, $E_{0j}||X||^2 < \infty$ implies Assumption 3 as shown in (2.7).

ASSUMPTION 3'. If δ is any procedure for which $R(\theta, \delta) \leqq R_0$ for all $\theta \in \mathscr{R}^2$, then there exist procedures δ_L, $L \to \infty$ with the following properties:

(i) $R(\theta, \delta_L) \leqq R(\theta, \delta)$ for $\|\theta\| \leqq L$;
(ii) for $L < M$ and $\|x\| < 3L/2$ we have

$$(5.5) \qquad V(j, \delta_L(x, y), y) \int f_j(z, y) \, dz \leqq V(j, \delta_M(x, y), y) \int f_j(z, y) \, dz$$

for all $j = 1, \cdots, r$; and
(iii)

$$(5.6) \qquad \int \nu_j(dy) \int_{\|x\| < L/2} dx \int_{\|\theta\| \geqq L} V(j, \delta_L(x, y), y) f_j(x - \theta, y) \, d\theta \to 0$$

$$\text{as } L \to \infty .$$

Now $gd = d$ so that (5.5) implies (2.4). That (5.6) implies (2.5) follows by straightforward calculation.

ASSUMPTION 4'. There is a non-increasing function $k : (0, \infty) \to (0, \infty)$ satisfying (2.8) and such that for any $\delta \in \mathscr{I}$ and all $j = 1, \cdots, r$ we have

$$(5.7) \qquad \int \nu_j(dy) \int_{\|x\| < v} [V(j, \delta_0(x, y), y) - V(j, \delta(x, y), y)] f_j(x, y) \, dx$$

$$\leqq k(v)\{\int \mu(dy) \int [W(\delta(x, y), x, y) - W(\delta_0(x, y), x, y)] p(x, y) \, dx\}^{\frac{1}{2}}$$

and

$$(5.8) \qquad \int \nu_j(dy) \int_{\|x\| > v} [V(j, \delta_0(x, y), y) - V(j, \delta(x, y), y)]^+ f_j(x, y) \, dx$$

$$\leqq k(v)\{\int \mu(dy) \int [W(\delta(x, y), x, y) - W(\delta_0(x, y), x, y)] p(x, y) \, dx\}^{\frac{1}{2}} .$$

By straightforward calculations, (5.7) implies (2.9) and (5.8) implies (2.10).

Admissibility as used in the proof of Theorem 3.1 (namely, $R(\theta, \delta) \leqq R_0$ for all θ implies equality) is not the usual definition for the multiple decision problem. Let

$$R^{(j)}(\theta, \delta) = E_{\theta j} V(j, \delta(X, Y), Y)$$

and $R_0^{(j)} = R^{(j)}(\theta, \delta_0)$, the latter being clearly independent of θ. Then, admissibility in the usual sense of δ_0 requires that $R^{(j)}(\theta, \delta) \leqq R_0^{(j)}$ for all θ, j implies equality. However, $R(\theta, \delta) = \sum_j \xi_j R^{(j)}(\theta, \delta)$ and $R_0 = \sum_j \xi_j R_0^{(j)}$. Since each $\xi_j > 0$, we obtain

COROLLARY 5.1. *In the multiple decision problem of this section suppose Assumptions 1'—4' are satisfied. Then, the procedure δ_0 which is Bayes in the class of invariant procedures is admissible in the usual sense.*

Note that admissibility in the usual sense of an invariant procedure implies it is Bayes in the class of invariant procedures with respect to some prior ξ on $\{1, \cdots, r\}$ (possibly with some $\xi_j = 0$).

We now proceed, as in Section 4, to obtain conditions implying those of Theorem 3.1. The theorem has been chosen rather than Corollary 5.1 since its conditions are slightly weaker.

First assume $V(j, d, y)$ is independent of y. This implies that V, and hence W,

are bounded functions. It has already been noted in Section 2 that in this case Assumption 2 implies Assumption 3.

We now consider hypothesis testing with $V(j, d, y) = 1 - \delta_{jd}$ $(j, d = 1, 2)$. Let ξ be the prior probability that $j = 1$. An extension to more general multiple decision problems of the lemma which follows (analogous to Lemma 4.2) can be obtained but is notationally difficult. The pattern is fully illustrated by the case stated above.

For $\delta \in \mathscr{S}$, let

$$A = \{y : \delta_0(x, y) = 1\}$$
$$B = \{y : \delta_0(x, y) = 2\}$$

$$C^2(y) \leq \left[\xi \frac{d\nu_1}{d\mu}(y) - (1 - \xi)\frac{d\nu_2}{d\mu}(y) \right][I_A(y) - I_B(y)]$$

$$D_v(y) \geq \sup\left\{ \int_{||x||<w} \left[-\xi f_1(x, y)\frac{d\nu_1}{d\mu}(y) + (1 - \xi)f_2(x, y)\frac{d\nu_2}{d\mu}(y) \right] dx : w \geq v \right\}$$

$$\text{if } y \in A$$

$$\geq \sup\left\{ \int_{||x||<w} \left[\xi f_1(x, y)\frac{d\nu_1}{d\mu}(y) - (1 - \xi)f_2(x, y)\frac{d\nu_2}{d\mu}(y) \right] dx : w \geq v \right\}$$

$$\text{if } y \in B$$

but, in any case, let $D_v(y) \geq 0$. Further, let

$$E_v(y) \geq \int_{||x||>v} \left[\left\{ \xi f_1(x, y)\frac{d\nu_1}{d\mu}(y) - (1 - \xi)f_2(x, y)\frac{d\nu_2}{d\mu}(y) \right\} \{I_B(y) - I_A(y)\} \right]^+ dx$$

$$k(v, y) = \max\left(\frac{D_v(y)}{C(y)}, \frac{E_v(y)}{C(y)} \right)$$

$$k(v) = [\int k^2(v, y)\mu(dy)]^{\frac{1}{2}}.$$

LEMMA 5.1. *If $\int_0^\infty k(v)\, dv < \infty$, then Assumption 4' is satisfied.*

The proof is so similar to that of Lemma 4.2 that we omit it.

The following example is an application of Corollary 5.1. Let $Z^{(1)}, \cdots, Z^{(n)}$ $(Z^{(i)} = (Z_1^{(i)}, Z_2^{(i)}))$ be independent, normally distributed on $\mathscr{R}^2$ with $EZ^{(i)} = \theta$ $(i = 1, \cdots, n)$ and each with covariance matrix $\left(\begin{smallmatrix}1 & \rho\\ \rho & 1\end{smallmatrix}\right)$. We wish to test $H_0 : \rho = 0$ versus $H_1 : \rho = \pm\rho_1$.

Define the relationship between ρ and j by:

$$\rho \quad 0 \quad \rho_1 \quad -\rho_1$$
$$j \quad 1 \quad 2 \quad 3 \ .$$

Set $X = \bar{Z}$ and $Y = (Y_1, Y_2, Y_3)^\tau$ where $Y_1 = \sum_1^n (Z_1^{(i)} - \bar{Z}_1)^2$, $Y_2 = \sum_1^n (Z_2^{(i)} - \bar{Z}_2)^2$, and $Y_3 = \sum_1^n (Z_1^{(i)} - \bar{Z}_1)(Z_2^{(i)} - \bar{Z}_2)$. Finally, let $\xi_1 > 0$, $\xi_2 = \xi_3 > 0$ and the loss function be

$$V(j, d, y) = 0 \quad \text{if } j = 1, d = 1 \text{ or } j = 2, 3, d = 2$$
$$= 1 \quad \text{if } j = 2, 3, d = 1$$
$$= \exp\left[-\frac{\rho_1^2}{1 - \rho_1^2}(y_1 + y_2) \right] \quad \text{if } j = 1, d = 2.$$

The unique test Bayes in the class of invariant tests rejects H_0 when $|Y_3|$ is large. Furthermore, any such test is Bayes in this class for some choice of $\xi_1 > 0$ and $\xi_2 = \xi_3 > 0$.

Clearly Assumption 2′ is satisfied.

With $\delta_L = \delta$, the inner integral in (5.6) decreases exponentially as $L \to \infty$. Hence, Assumption 3′ is satisfied.

Extending the application of Lemma 5.1 in the obvious way to this case we find we are able to set $C^2(y) = \alpha < 1$, and $D_v(y) = E_v(y) = e^{-v}h(y)$ where h^2 is μ-integrable. Then, $k(v) = \lambda e^{-v}$, integrable. This verifies Assumption 4′.

Thus we see that, for any $C > 0$, the test which rejects H_0 when $|Y_3| > C$ is admissible in our problem.

The admissibility of this test with this loss function is not as interesting as is admissibility under the loss function

$$V(j, d, y) = 0 \qquad \text{if} \quad j = 1, d = 1 \quad \text{or} \quad j = 2, 3, d = 2$$
$$= 1 \qquad \text{otherwise} .$$

Admissibility in this case may be obtained as above by observing that the calculations from (5.2) to (5.4) are valid even if ξ depends on y. We will write $\xi(y)$. Then set

$$\xi_1(y) = 1 \bigg/ \bigg[1 + \alpha \exp\bigg(-\frac{\rho_1^2}{1 - \rho_1^2} (y_1 + y_2) \bigg) \bigg]$$

for $\alpha > 0$ and $\xi_2(y) = \xi_3(y) = [1 - \xi_1(y)]/2$.

This result for the 0–1 loss function clearly implies admissibility of this test for any alternative which includes symmetrically placed values of ρ and, in particular, the known result for the alternative $\rho \neq 0$. In these cases we may even let the loss function be any symmetric function of ρ for the case of accepting H_0 when false.

We have been unable to extend these methods to show that, if the variances are unknown, then the test which rejects when $|Y_3|/(Y_1 Y_2)^{\frac{1}{2}} > C$ is admissible. However, an extension is possible to show admissibility of the previous test when the variances are unknown but their product is known.

6. Examples. The examples of this section fail to satisfy the moment condition (Assumption 2) and the conclusion of the theorem. Unfortunately, we have been unable to verify all the remaining assumptions. No proofs of assertions made will be given.

EXAMPLE 6.1. Estimation. This example is patterned after Perng's (1970).

Let $\mathscr{D} = \mathscr{R}^2$ with $L(d, \theta, x, y) = ||d - \theta||^2$. Let $\mathscr{Y} = \mathscr{R}$ and $\mathscr{B}_y$ be the usual Borel σ-field. Finally, for $\eta > 0$, let

$$\mu(dy) = Ky^{-(5-\eta)} dy \qquad \text{if} \quad y > 1$$
$$= 0 \qquad \text{if} \quad y \leq 1 .$$

Set $p(x, y) = y^{-2}h(x/y)$ where

$$h(z) = 1/\pi \qquad \text{if} \quad ||z|| < 1$$
$$= 0 \qquad \text{if} \quad ||z|| \geq 1 \, .$$

The unique a.e. (μ) best invariant procedure is $\delta_0(x, y) = x$. Furthermore, $\int \mu(dy) \int ||x||^\alpha W(\delta_0(x, y), x, y)p(x, y) \, dx < \infty$ if, and only if, $\alpha < 2 - \eta$.

It is easy to verify (2.9) of Assumption 4. However, (2.10) does not hold. If $\delta(x, y) = x + \gamma(y)$ with $||\gamma(y)||^2$, then the integral on the right side of (2.14) is finite provided $\alpha < 4 - \eta$ while the left side behaves as $v^{-(4-\eta-\alpha)}$, not integrable when $\alpha \geq 3 - \eta$. We have been unable to verify Assumption 3.

A better estimate is $\delta(x, y) = x + y\lambda(x/y)$ with

$$\lambda(t) = -\varepsilon\alpha(1 - \varepsilon^2||t||^2)t \qquad \text{if} \quad ||t|| < 1/\varepsilon$$
$$= 0 \qquad \text{if} \quad ||t|| \geq 1/\varepsilon$$

where $\alpha, \varepsilon > 0$ are small, α small relative to ε.

EXAMPLE 6.2. Testing. This example is patterned after that in Fox and Perng (1969).

Let $\mathcal{D} = \{1, 2\}$, $V(j, d, y) = 1 - \delta_{j,d}$, $\mathcal{Y} = \mathcal{R}^3$ and $\mathcal{B}_{\mathcal{Y}}$ be the usual Borel σ-field on $\mathcal{R}^3$. Consider Y in the form (Y_1, Y_2) with Y_1 taking values in $\mathcal{R}$ and Y_2 taking values in $\mathcal{R}^2$. Let Y_1 and Y_2 be independent. Let (y_1, y_2) be the corresponding decomposition of $y \in \mathcal{R}^3$. Let Y_1 have density with respect to Lebesgue measure g_1 when $j = 1$ and g_2 when $j = 2$ given by

$$g_1(u) = g_2(-u) = C_1 u^{-3} \qquad \text{if} \quad u > 1$$
$$= C_2|u|^{-3} \qquad \text{if} \quad u < -1$$
$$= 0 \qquad \text{if} \quad |u| \leq 1$$

where $C_1/C_2 < \xi_1/\xi_2 < C_2/C_1$. For $j = 1, 2$ let Y_2 be distributed uniformly in the first quadrant on the perimeter of the unit circle. Then,

$$\delta_0(x, y) = 1 \qquad \text{if} \quad y_1 \leq 0$$
$$= 2 \qquad \text{if} \quad y_1 > 0$$

is the a.e. (μ) unique best invariant procedure.

Let Z be distributed uniformly in the first quadrant in the interior of the circle of radius η. When $j = 1$ let

$$X = Y_1^{1+\varepsilon}Y_2 + Z + \theta \qquad \text{if} \quad Y_1 > 0$$
$$= -Z + \theta \qquad \text{if} \quad Y_1 \leq 0$$

and when $j = 2$ let

$$X = -Z + \theta \qquad \text{if} \quad Y_1 > 0$$
$$= Z + \theta \qquad \text{if} \quad Y_1 \leq 0 \, .$$

Clearly, Assumption 2 fails to hold. In fact, $E_{20}||X||^\alpha < \infty$ for all $\alpha > 0$ while

$E_{10}||X||^\alpha < \infty$ if, and only if, $\alpha < 2/(1 + \varepsilon)$. Unfortunately we shall have to take $0 < \varepsilon < 1$ with ε sufficiently large in order to obtain inadmissibility.

We have been unable to verify Assumption 3. However, both parts of Assumption 4 are readily verified either directly or using Lemma 5.1. See the Remark below.

For $\eta > 0$, sufficiently small, $\varepsilon < 1$, sufficiently close to 1, $C_1 < C_2$, sufficiently close to one another, and $a > 2$, sufficiently close to 2, a better procedure is

$$\delta(x, y) = 1 \qquad \text{if} \quad y_1 > 1; y_1^{2(1+\varepsilon)} \le ||x||^2 \le a y_1^{1+\varepsilon} y_2^{-}x; x_1, x_2 > a - 1$$
$$= 2 \qquad \text{if} \quad y_1 < -1; ||x|| \le |y_1|^{1+\varepsilon}; x_1, x_2 > a - 1$$
$$= \delta_0(x, y) \qquad \text{otherwise .}$$

REMARK. The verification that this example satisfies Assumption 4 could have used Lemma 5.1.

7. **A further application.** In this section we give a generalized version of the application given in Section 5. A special case of this generalization would be estimation of a parameter η in the presence of a nuisance location parameter θ.

We generalize (5.1) to

$$P_{\theta, \eta}\{X, Y \in S\} = \int_S \int f_\eta(x - \theta, y) \, dx \nu_\eta(dy)$$

where $\eta \in H$. Let $\mathscr{A}$ be a σ-field of subsets of H and ξ be a probability measure (prior distribution of η) on $\mathscr{A}$. Assume

(i) $f_\eta(x, y)$ is jointly measurable in x, y and η;
(ii) $f_\eta(x, y)$ is a probability density function on $\mathscr{A}^2$ for each fixed η and y;
(iii) $\nu_\eta(B)$ is measurable in η for each fixed B;
(iv) there exists a probability measure μ such that $\nu_\eta \ll \mu$ for all $\eta \in H$; and
(v) $\int (d\nu_\eta/d\mu)(y)\xi(d\eta) = 1$.

We wish to make a decision $d \in \mathscr{D}$ subject to loss $V(\eta, d, \theta, x, y)$. Proceeding as before, set

$$L(d, \theta, x, y) = \frac{\int \xi(d\eta) V(\eta, d, \theta, x, y) f_\eta(x - \theta, y) \dfrac{d\nu_\eta}{d\mu}(y)}{\int \xi(d\eta) f_\eta(x - \theta, y) \dfrac{d\nu_\eta}{d\mu}(y)} .$$

For (2.2) to be valid we require a similar statement for V, that is,

$$V(\eta, d, \theta, x, y) = V^*(\eta, g_{-\theta}d, x - \theta, y) .$$

Also set

$$p(x - \theta, y) = \int \xi(d\eta) f_\eta(x - \theta) \frac{d\nu_\eta}{d\mu}(y) .$$

Then,

$$P_\theta\{X, Y \in S\} = \int_S \int p(x - \theta, y) \, dx \mu(dy)$$

and

$$\int p(x - \theta, y) \, dx = 1 .$$

Furthermore, assuming Fubini's theorem applies (which only requires that $\bar{\delta}$ have finite risk),

$$R(\theta, \bar{\delta}) = \int \xi(d\eta) \int V(d, \theta, \eta, x, y) f_\gamma(x - \theta, y) \nu_\gamma(dy),$$

the usual Bayes risk for fixed θ.

Assumptions 1–4 in this case are implied by the following.

ASSUMPTION 1″. The Bayes invariant procedure exists.

ASSUMPTION 2″. For all $\eta \in H$ we have

$$\int \nu_\gamma(dy) \int ||x||^2 V(\eta, d, \theta, x, y) f_\gamma(x, y) \, dx \leqq M < \infty.$$

ASSUMPTION 3″. If $\bar{\delta}$ is any procedure for which $R(\theta, \bar{\delta}) \leq R_0$ for all $\theta \in \mathscr{R}^2$, then there exist procedures $\bar{\delta}_L$, $L \to \infty$ with the following properties

  (i) $R(\theta, \bar{\delta}_L) \leq R(\theta, \bar{\delta})$ for $||\theta|| \leq L$;
  (ii) for $\eta \in H$, $L < M$, $||x|| < 3L/2$ we have

$$\int V(\eta, g_{-(x-z)}\bar{\delta}_L(x, y), x - z, z, y) f_\gamma(z, y) \, dz$$
$$\leqq \int V(\eta, g_{-(x-z)}\bar{\delta}_M(x, z), x - z, z, y) f_\gamma(z, y) \, dz;$$

and

  (iii) $\int \nu_\gamma(dy) \int_{||x||<L/2} dx \int_{||\theta|| \geqq L} V(\eta, g_{-\theta}\bar{\delta}_L(x, y), \theta, x - \theta, y) f_\gamma(x - \theta, y) \, d\theta \to 0$ uniformly in $\eta \in H$ as $L \to \infty$.

ASSUMPTION 4″. There exists a non-increasing function $k: (0, \infty) \to (0, \infty)$ satisfying (2.11) such that for all $\eta \in H$ and $\bar{\delta} \in \mathscr{I}$ we have

$$\int \nu_\gamma(dy) \int_{||x||<v} [V(\eta, \delta_0(x, y), 0, x, y) - V(\eta, \delta(x, y), 0, x, y)] f_\gamma(x, y) \, dx$$
$$\leqq k(v)\{\int \mu(dy) \int [W(\delta(x, y), x, y) - W(\delta_0(x, y), x, y)] p(x, y) \, dx\}^{\frac{1}{2}}$$

and

$$\int \nu_\gamma(dy) \int_{||x||>v} [V(\eta, \delta_0(x, y), 0, x, y) - V(\eta, \delta(x, y), 0, x, y)]^+ f_\gamma(x, y) \, dx$$
$$\leqq k(v)\{\int \mu(dy) \int [W(\delta(x, y), x, y) - W(\delta_0(x, y), x, y)] p(x, y) \, dx\}^{\frac{1}{2}}.$$

We then obtain

COROLLARY 7.1. *In the problem of this section suppose Assumptions* 1″—4″ *are satisfied. Then the procedure* δ_0 *which is Bayes in the class of invariant procedures is almost* (ξ) *admissible.*

REFERENCES

BROWN, L. D. (1966). On the admissibility of invariant estimators of one or more location parameters. *Ann. Math. Statist.* 37 1087–1137.
BROWN, L. D. and FOX, M. (1971). Admissibility in statistical problems involving a location or scale parameter. To appear in *Ann. Statist.* 2.
FARRELL, R. (1964). Estimators of a location parameter in the absolutely continuous case. *Ann. Math. Statist.* 35 949–959.
FOX, M. (1971). Admissibility of multiple decision procedures Bayes in the class invariant under a real location parameter. *Ann. Math. Statist.* 42 1553–1561.

Fox, M. and Perng, S. K. (1969). Inadmissibility of the best invariant test when the moment is infinite under one of the hypotheses. *Ann. Math. Statist.* **40** 1483–1485.

James, W. and Stein, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 361–379 Univ. of California Press.

Joshi, V. M. (1970). Admissibility of invariant confidence procedures for estimating a location parameter. *Ann. Math. Statist.* **41** 1568–1581.

Perng, S. K. (1970). Inadmissibility of various "good" statistical procedures which are translation invariant. *Ann. Math. Statist.* **41** 1311–1321.

Portnoy, S. and Stein, C. (1971). Inadmissibility of the best invariant test in three or more dimensions. *Ann. Math. Statist.* **42** 799–801.

Stein, C. (1959). The admissibility of Pitman's estimator for two location parameters. Stanford Univ. Technical Report No. 25.

Wald, A. and Wolfowitz, J. (1950). Bayes solutions of sequential decision problems. *Ann. Math. Statist.* **21** 82–99.

DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NEW YORK 14850

DEPARTMENT OF STATISTICS AND PROBABILITY
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48823