The Annals of Mathematical Statistics
1971, Vol. 42, No. 5, 1553–1561

ADMISSIBILITY OF CERTAIN LOCATION INVARIANT MULTIPLE DECISION PROCEDURES[1]

BY MARTIN FOX

Michigan State University

Random variables $X, Y_1, Y_2, \cdots$ are available for observation with X real valued and $Y_1, Y_2, \cdots$ taking values in arbitrary spaces. The distribution of $Y = (Y_1, Y_2, \cdots)$ is given by μ_j $(j = 1, \cdots, r)$ and the conditional density with respect to Lebesgue measure given $Y_i = y_i$ $(i = 1, \cdots, n-1)$ is $p_{jn}(x-\theta, y)$ where $y = (y_1, y_2, \cdots)$. The parameters j and θ are unknown. A decision $k \in \{1, \cdots, m\}$ is to be made with loss $W(j, k, n, y)$ when n observations are taken. Following Brown's (1966) methods admissibility is proved for the decision procedure which is Bayes in the class of invariant procedures. The result contains that of Lehmann and Stein (1953).

1. Introduction. The admissibility of best invariant sequential estimates of a real location parameter was studied by Brown (1966). A related testing problem, namely of testing that observations come from one location family versus the alternative that they come from another (real location parameter in both cases), was treated by Lehmann and Stein (1953) for the fixed sample size case. In the present paper Brown's methods are adapted to invariant sequential multiple decision problems when the sample comes from one of several location families. As in Brown's paper the results apply to fixed sample size problems as a special case.

Let $H = \times_1^\infty E_i$ where E_1 is the real line and $E_2, E_3, \cdots$ are arbitrary. Let $\mathscr{B}_1$ be the Borel σ-field on E_1 and $\mathscr{B}_i$ be an arbitrary σ-field on E_i $(i = 2, 3, \cdots)$. Let $\mathscr{A}_i$ be the σ-field of cylinder sets with measurable bases in the ith coordinate and $\mathscr{A}^{(n)} = \mathscr{F}(\bigcup_1^n \mathscr{A}_i)$ where $\mathscr{F}(\mathscr{C})$ is the smallest σ-field containing $\mathscr{C}$. In Brown (1966) the E_i are all taken to be the real line. His results in Section 2 except for those in Subsection 2.2 do not require his more restrictive conditions. It will be seen that in the present setting the results in Section 3 (which parallel Brown's Subsection 2.2) do not require his restrictive assumptions.

Let $(X, Y_1, Y_2, \cdots)$ be a random variable on $(H, \mathscr{A}^{(\infty)})$ such that

(i) X takes values in E_1 and is $\mathscr{A}_1$-measurable and

(ii) Y_i takes values in E_{i+1} and is $\mathscr{A}_{i+1}$-measurable $(i = 1, 2, \cdots)$.

Let $Y = (Y_1, Y_2, \cdots)$. For all $\theta \in E$ and $j = 1, \cdots, r$ if $S \in \mathscr{A}^{(n)}$, then

$$P_{\theta j}((X, Y) \in S) = \iint_S p_{jn}(x-\theta, y)\, dx\, \mu_j(dy)$$

and set $p_j = p_{j\infty}$. Here $\int_E p_{jn}(x, \cdot)dx = 1$ a.e. (μ_j) and each μ_j is a probability measure on $\mathscr{F}(\bigcup_2^\infty \mathscr{A}_i)$. Thus, $p_{jn}(x-\theta, y)$ is the conditional probability density

Received July 1, 1970.

[1] Research supported by NSF Grant GP-13484.

with respect to Lebesgue measure, henceforth denoted by λ, of X given $Y_i = y_i (i = 1, \cdots, n-1)$ and is $\mathscr{A}^{(n)}$-measurable.

This model is motivated by considering $Z_1, Z_2, \cdots$ which are independent, identically distributed random variables whose common probability density with respect to λ is $f_j(x-\theta)$. We then set $X = Z_1$ and $Y_i = Z_{i+1} - Z_1$ $(i = 1, 2, \cdots)$ so Y is the maximal invariant under the translation group. Section 5 contains an example of this type for fixed sample size.

We wish to make a decision $k \in \{1, \cdots, m\}$ without knowledge of the true values of j and of θ. Observations are taken sequentially on $X, Y_1, Y_2, \cdots$. A decision procedure δ consists of a stopping rule σ and a behaviorial decision rule ε where $\sigma : H \to I$ and $\varepsilon : H \to \Lambda_m$. Here I is the set of natural numbers and Λ_v is the probability simplex in v-dimensional Euclidean space. The interpretation of $\sigma = n$ is that sampling stops after observing $X, Y_1, \cdots, Y_{n-1}$ so that we require, for all Lebesgue measurable $A \subset \Lambda_m$, that

$$\{(x, y) : \sigma(x, y) = n, \varepsilon(x, y) \in A\} \in \mathscr{A}^{(n)}.$$

If $\sigma(x, y) = 0$ for some (x, y), then this is true for all (x, y) and ε is a constant.

The loss may depend on

(i) the index j of the true family of distributions;
(ii) the decision k;
(iii) the number n of observations; and
(iv) the value y of Y.

Hence, the loss function will be denoted by $W(j, k, n, y)$. By letting $W(j, k, n, y) = \infty$ when $n \neq n_0$ we obtain a fixed sample size problem. In this case we abuse notation by deleting n from the arguments of W. The risk of the decision procedure $\delta = (\sigma, \varepsilon)$ is

$$R(\theta, j, \delta) = \sum_{k=1}^{m} \int\int \varepsilon_k(x, y) W(j, k, \sigma(x, y), y) p_j(x-\theta, y)\, dx\, \mu_j(dy)$$

where ε_k is the kth coordinate function of ε. Then, a decision procedure δ is (translation) invariant if it does not depend on x or, more formally, if δ is $\mathscr{F}(\bigcup_2^\infty \mathscr{A}_i)$ measurable.

Let $\mathscr{I}$ denote the set of all $\mathscr{F}(\bigcup_2^\infty \mathscr{A}_i)$-measurable functions (i.e., invariant measurable functions) on H. For $\delta = (\sigma, \varepsilon) \in \mathscr{I}$ the risk is independent of θ so we suppress this argument from the function R and observe that in this case

$$R(j, \delta) = \sum_{k=1}^{m} \int \varepsilon_k(y) W(j, k, \sigma(y), y) \mu_j(dy).$$

Let $\xi \in \Lambda_r$, that is, ξ is an a priori distribution on the index j of the family of distributions. We then extend the domain of definition of R by setting

$$R(\theta, \xi, \delta) = \sum_{j=1}^{r} \xi_j R(\theta, j, \delta).$$

Now let δ_0 be Bayes in $\mathscr{I}$ with respect to ξ, that is, $\delta_0 \in \mathscr{I}$ and

$$R(\xi, \delta_0) = \inf_{\delta \in \mathscr{I}} R(\xi, \delta).$$

Set $R_0(j) = R(j, \delta_0)$. Without loss of generality assume $\xi_j > 0$ $(j = 1, \cdots, r)$.

The purpose of this paper is to present the

THEOREM. *Let* $\delta_0 = (\sigma_0, \varepsilon_0)$ *be Bayes in* $\mathscr{I}$ *with respect to* ξ. *Assume*

(i) *if* $\delta_i \in \mathscr{I}$ *and* $R(j, \delta_i) \to R_0(j)$ $(j = 1, \cdots, r)$, *then* $\delta_i \to \delta_0$ *in measure* (μ_j);

(ii) *for* $j = 1, \cdots, r$; $k = 1, \cdots, m$ *we have*

$$\int \mu_j(dy)\{W(j, k, \sigma_0(y), y) \int |x| p_j(x, y)\, dx\} < \infty; \tag{1.1}$$

and

(iii) *for* $j = 1, \cdots, r$; $k = 1, \cdots, m$ *we have*

$$\int_0^\infty dz\, \{\sup_{\delta \in \mathscr{I}} \int \mu_j(dy)\{[\varepsilon_{0k}(y)W(j, k, \sigma_0(y), y) - \varepsilon_k(y)W(j, k, \sigma(y), y)] \int_{-z}^{z} p_j(x, y)\, dx\}\} < \infty. \tag{1.2}$$

Then, δ_0 *is admissible.*

This theorem is analogous to Brown's Theorem 2.1.1. In this case, contrary to Brown's, all procedures for which no observations are taken are invariant. Note that (1.1) implies $R_0 < \infty$.

Setting $r = m = 2$, consider a fixed sample size problem in which $W(j, k, y) = 0$ or 1 according as $j = k$ or $j \neq k$. This is the hypothesis testing problem considered by Lehmann and Stein. Any test which is best invariant of its size and which requires randomization with probability zero (as in Theorem 1 of Lehmann and Stein (1953)) is Bayes in $\mathscr{I}$ with respect to some $\xi \in \Lambda_2$. Furthermore, (1.1) is the moment condition in Lehmann and Stein (1953). An example has been given (Fox and Perng (1969)) showing inadmissibility may result when this moment condition is violated.

The proof of the theorem parallels that in Brown (1966) Section 2.1. A brief discussion of changes needed is given in Section 2. Sections 3 and 4, respectively, contain lemmas concerning Conditions (i) and (iii) of the theorem. These lemmas and their proofs parallel corresponding lemmas in Brown (1966) Sections 2.2, 2.3. Complete proofs are not given in Section 3. Those in Section 4 are sufficiently short so that they are included.

Lemmas 3.2 and 4.3 show that the theorem contains the corollary given by Lehmann and Stein (1953). Complete proofs are found in Fox (1970).

2. Outline of the proof of the theorem. Let $R(\theta, j, \delta) \leqq R_0(j)$ for all $\theta \in E_1$ and $j = 1, \cdots, r$. Let

$$\omega(j, z, n, z', n', y) = \sum_{k=1}^{m} [z_k W(j, k, n, y) - z_k' W(j, k, n', y)] \tag{2.1}$$

and $\bar{\omega}(j, x, y) = \omega(j, \varepsilon_0(y), \sigma_0(y), \varepsilon(x, y), \sigma(x, y))$.

The assumption of the previous paragraph implies

$$\int_{-L}^{L} d\theta \int \mu_j(dy) \int \bar{\omega}(j, x, y) p_j(x - \theta, y)\, dx \geqq 0$$

for all $j = 1, \cdots, r$ and $L > 0$. Note that the statement of the theorem does not restrict attention to procedures with $\sigma \geqq 1$. In the present case, contrary to Brown's, procedures with $\sigma = 0$ are invariant.

For each $j = 1, \cdots, r$ the steps of the proof proceed as in Brown (1966) up to (2.1.18). A lemma parallel to Lemma 2.1.1 is not needed. Instead, the corresponding step results directly as follows.

Let

$$\begin{aligned} H_j &= \limsup_{L\to\infty} \int \mu_j(dy) \int_{-3L/2}^{-L/2} dx\,[\bar\omega(j,x,y) \int_{-L/2}^{L+x} p_j(z,y)\,dz] \\ &= \limsup_{L\to\infty} \int \mu_j(dy)\{\int_{-L-A}^{-L+A} dx \int_{-L/2}^{L+x} dz + \int_{-3L/2}^{-L-A} dx \int_{-L/2}^{L+x} dz \\ &\qquad + \int_{-L+A}^{-L/2} dx \int_{-L-x}^{L+x} dz + \int_{-L+A}^{-L/2} dx \int_{-L/2}^{-L-x} dz\} \\ &\qquad \cdot \{\bar\omega(j,x,y)p_j(z,y)\} \qquad (2.2) \\ &\leqq \limsup_{L\to\infty} \int_{-L-A}^{-L+A} dx\,[\bar\omega(j,x,y) \int \mu_j(dy) \int_{-L/2}^{L+x} p_j(z,y)\,dz] \\ &\quad + \limsup_{L\to\infty} \int \mu_j(dy)\{\int_{-L/2}^{-A} dz \int_{z-L}^{-A-L} dx + \int_{-L+A}^{-L/2} dx \int_{-L-x}^{L+x} dz \\ &\qquad + \int_{-L/2}^{-A} dz \int_{-L+A}^{-z-L} dx\} \\ &\qquad \cdot \{\bar\omega(j,x,y)p_j(z,y)\}. \end{aligned}$$

Consider the first term in the last expression in (2.2). The outer integration yields a term which is an integral over $S(L)$ (and hence converges to zero by a step parallel to (2.1.18)) plus an integral over the remainder of the region. Since $-L-A < x < -L+A$ over the entire region of integration, on this remaining part we have $\sigma(x, y) = \sigma_0(y)$ and $|\varepsilon_k(x, y)-\varepsilon_0(y)| \leqq \alpha$ ($k = 1, \cdots, n$). Thus, this remainder is at most

$$\begin{aligned} &\alpha \int_{-L-A}^{-L+A} dx \int \mu_j(dy)[\sum_{k=1}^m W(j,k,\sigma_0(y),y) \int_{-L/2}^{L+x} p_j(z,y)\,dz] \\ &\quad = \alpha \sum_{k=1}^m \int \mu_j(dy)[W(j,k,\sigma_0(y),y) \int_{-L/2}^{A} (A-z)p_j(z,y)\,dz] \\ &\quad \leqq \alpha M \end{aligned}$$

for some finite M as a consequence of (1.1). But $\alpha > 0$ is arbitrary so even the remainder converges to 0 as $L \to \infty$.

Once again the proof parallels that in Brown (1966), this time through (2.1.24). Since inequalities parallel to (2.1.13) and (2.1.24) hold for all $j = 1, \cdots, r$ we have

$$\begin{aligned} 0 &\leqq \sum_{j=1}^r \xi_j \liminf_{L\to\infty} \int \mu_j(dy) \int_{-L/2}^{L/2} dx\,[\bar\omega(j,x,y) \int_{-L/2}^{L/2} p_j(z,y)\,dz] \\ &\leqq \liminf_{L\to\infty} \sum_{j=1}^r \xi_j \int \mu_j(dy) \int_{-L/2}^{L/2} dx\,[\bar\omega(j,x,y) \int_{-L/2}^{L/2} p_j(z,y)\,dz] \qquad (2.3) \\ &\leqq \sum_{j=1}^r \xi_j \int dx\,[\int \mu_j(dy)\{\bar\omega(j,x,y) \int p_j(z,y)\,dz\}] \\ &= \int dx\,[\sum_{j=1}^r \xi_j \int \mu_j(dy)\{\bar\omega(j,x,y) \int p_j(z,y)\,dz\}]. \end{aligned}$$

Since $\delta_x \in \mathscr{I}$ and δ_0 is Bayes in $\mathscr{I}$ with respect to ξ, the last quantity in brackets in (2.3) is nonpositive and, hence, is 0 for almost all $x(\lambda)$. But Assumption (i) implies for each $j = 1, \cdots, r$ that δ_0 is the unique a.e. $(\sum_1^r \mu_j)$ decision rule which is Bayes in $\mathscr{I}$ with respect to ξ. As a consequence, $\delta(x, y) = \delta_x(y) = \delta_0(y)$ a.e. $(\lambda \times \sum_1^r \mu_j)$ which completes the proof.

3. Discussion of Condition (i). We have seen that Condition (i) implies that δ_0 is the unique Bayes procedure in $\mathscr{I}$ with respect to ξ. In Lehmann and Stein (1953), with fixed sample size, unicity of the best invariant test of a given size of $H_1 : j = 1$ versus $H_2 : j = 2$ is guaranteed by the assumption that the μ_1-probability of randomization is 0. Perng (1967) showed that without unicity the best invariant test may be inadmissible. We study in this section the relationship between unicity and Condition (i) and will see that the latter is not much stronger than the former. In fact, in the fixed sample size case, Lemma 3.2 states the two are equivalent.

For $\delta \in \mathscr{I}$, let $\xi_{j'}^*(y)$ be the posterior probability that $j = j'$ given $Y = y$ and

$$R(\xi, \delta \mid y) = \textstyle\sum_{j=1}^r \xi_j^*(y) \sum_{k=1}^m \varepsilon_k(y) W(j, k, \sigma(y), y),$$

the conditional Bayes risk of δ given $Y = y$. Thus, $R(\xi, \delta) = \sum_1^r \xi_j \int R(\xi, \delta \mid y) \times \mu_j(dy)$.

Let σ' be a fixed invariant stopping rule. Assume, for the remainder of this section, that there exists $\delta = (\varepsilon, \sigma') \in \mathscr{I}$ such that $R(\xi, \delta) < \infty$. Let

$$B_0(y) = \{v : v \in \Lambda_m, R(\xi, (v, \sigma') \mid y) = \inf_{t \in \Lambda_m} R(\xi, (t, \sigma') \mid y)\}; \tag{3.1}$$

$$\varepsilon_{0k}'(y) = \inf \{v_k : v \in B_{k-1}(y)\} \qquad (k = 1, \cdots, m) \tag{3.2}$$

and

$$B_k(y) = \{v : v \in B_{k-1}(y), v_k = \varepsilon_{0k}'(y)\} \qquad (k = 1, \cdots, m-1). \tag{3.3}$$

The first lemma parallels Lemma 2.2.1 in Brown (1966).

LEMMA 3.1. *Let σ', $\delta = (\varepsilon, \sigma')$ and ε_0' be as in the preceding paragraph. Then, ε_0' is finitely measurable, $\delta_0' = (\varepsilon_0', \sigma') \in \mathscr{I}$ and $R(\xi, \delta_0') = \inf_{\delta = (\varepsilon, \sigma') \in \mathscr{I}} R(\xi, \delta)$. The decision procedure δ_0' is the unique* a.e. $(\sum_1^r \mu_j)$ *procedure of the form* $(\varepsilon, \sigma') \in \mathscr{I}$ *having Bayes risk $R(\xi, \delta_0')$ if, and only if, for almost all y $(\sum_1^r \mu_j)$ the sets defined in* (3.2) *are singletons for each $k = 1, \cdots, m$. In this case, $R(\xi, \delta_v') \to R(\xi, \delta_0')$ for a sequence $\delta_v' = (\varepsilon_v', \sigma') \in \mathscr{I}$ implies $\varepsilon_v' \to \varepsilon_0$ in measure $(\sum_1^r \mu_j)$.*

PROOF. Note that p_{jn} and $W(j, k, n, \cdot)$ must be $\mathscr{A}^{(n)}$-measurable. On $\Sigma_n = \{y : \sigma'(y) = n\}$ let

$$\begin{aligned} f_n(v, y) &= \textstyle\sum_{j=1}^r \xi_j \sum_{k=1}^m v_k W(j, k, n, y) \\ &= R((v, \sigma') \mid y). \end{aligned} \tag{3.4}$$

In (3.4) we can take $y = (y_1, \cdots, y_{n-1})$ so that f_n is Borel measurable on $\Lambda_m \times \Sigma_n$ and, clearly, is continuous in v for each $y \in \Sigma_n$. Hence, $B_0(y)$ is closed and, furthermore, $g_n = \inf_{v \in \Lambda_m} f_n(v, \cdot)$ is measurable. Hence, for any fixed $a \in \Lambda_m$ the measurability of $\{y : \varepsilon_{0k}'(y) > a_k\} = \{y : f_n(a, y) - g_n(y) > 0\}$ follows. Thus, ε_0' is Borel measurable on Σ_n for each n. Since $\sigma' < \infty$ a.e. $(\sum_1^r \mu_j)$, we see that ε_0' is finitely measurable. Also $\delta_0' \in \mathscr{I}$ and is a decision rule since $\{(x, y) : \sigma'(y) = n, \varepsilon_0' \in A\} \in \mathscr{A}^{(n)}$ for every Lebesgue measurable $A \subset \Lambda_m$.

The remainder of the proof is parallel to that of Brown (1966) Lemma 2.2.1. The second lemma parallels his Lemma 2.2.2.

LEMMA 3.2. *In the fixed sample size case a procedure δ_0 which is Bayes in $\mathscr{I}$ exists. This procedure is unique* a.e. $(\sum_1^r \mu_j)$ *if, and only if, Assumption* (i) *of the theorem is valid.*

PROOF. The first statement in Lemma 3.1 guarantees existence of δ_0. The last statement proves that unicity implies Assumption (i). The reverse implication has already been noted.

The last lemma parallels Lemma 2.2.3 in Brown (1966).

LEMMA 3.3. *Assume* $W(j, k, \infty, y) \equiv \infty$ *and*

$$\liminf_{n\to\infty} W(j, k, n, y) = \infty.$$

Then, there exists at least one procedure Bayes in $\mathscr{I}$. Furthermore, Assumption (i) *of the theorem is satisfied if, and only if, the procedure Bayes in $\mathscr{I}$ is unique* a.e. $(\sum_1^r \mu_j)$.

PROOF. The proof parallels that of Brown (1966) Lemma 2.2.3 through the assertion following (2.2.12) that the sequence of stopping rules converges in measure (in this case with respect to $\sum_1^r \mu_j$). Convergence of the sequence of behavioral decision rules follows in a similar fashion since if $\delta^* = (\sigma^*, \varepsilon^*)$ is the unique procedure Bayes in $\mathscr{I}$, then a.e. $(\sum_1^r \mu_j)$ only one component of ε^* is positive (and, hence, equals one).

4. Discussion of Condition (iii). We consider lemmas concerning Condition (iii) of the theorem. Recall the definition of ω in (2.1) and let

$$I_j = \int_0^\infty dz \,\{\sup_{\delta\in\mathscr{I}} \int \omega(j, \varepsilon_0(y), \sigma_0(y), \varepsilon(y), \sigma(y), y)\cdot \mu_j(dy) \int_{-z}^{z} p_j(x, y)\, dx\}.$$

Condition (iii) is equivalent to $I_j < \infty$ $(j = 1, \cdots, r)$.

The first lemma parallels Lemma 2.3.1 of Brown (1966).

LEMMA 4.1. *If there exists* $z_0 < \infty$ *such that*

$$\int \mu_j(dy) \int_{-z_0}^{z_0} p_j(z, y)\, dx = 1 \tag{4.1}$$

for each $j = 1, \cdots, r$, *then Condition* (iii) *is satisfied.*

PROOF. By (4.1) for $z > z_0$ we have

$$\sup_{\delta\in\mathscr{I}} \int \omega(j, \varepsilon_0(y), \sigma_0(y), \varepsilon(y), \sigma(y), y)\mu_j(dy) \int_{-z}^{z} p_j(x, y)\, dx = 0.$$

Hence,

$$\begin{aligned} I_j &\leqq \int_0^{z_0} dz \sum_{k=1}^m \int \varepsilon_{0k}(y) W(j, k, \sigma_0(y), y)\mu_j(dy) \cdot \int p_j(x, y)\, dx \\ &\leqq z_0 R_0(j) < \infty. \end{aligned}$$

The next lemma covers the case in which procedures having finite risk have sample size bounded by $N < \infty$. It parallels Lemma 2.3.2 of Brown (1966).

LEMMA 4.2. *Suppose there exists* $N < \infty$ *such that* $W(j, k, n, y) = \infty$ *for* $n > N$ *and that*

$$(4.2)\qquad \int W(j,k,n,y)\mu_j(dy)\int |x|p_j(x,y)\,dx < \infty$$

for $j = 1, \cdots, r$; $k = 1, \cdots, m$; $n \leqq N$. *Then Assumption* (iii) *is satisfied.*

PROOF. Since δ_0 is Bayes in $\mathscr{I}$ we have

$$\sum_{j=1}^r \xi_j \int \omega(j, \varepsilon_0(y), \sigma_0(y), \varepsilon(y), \sigma(y), y)\mu_j(dy)\int p_j(x,y)\,dx \leqq 0.$$

Hence,

$$(4.3)\qquad \sum_{j=1}^r \xi_j \int \omega(j, \varepsilon_0(y), \sigma_0(y), \varepsilon(y), \sigma(y), y)\mu_j(dy)\int_{-z}^{z} p_j(x,y)\,dx$$
$$\leqq -\sum_{j=1}^r \xi_j \int \omega(j, \varepsilon_0(y), \sigma_0(y), \varepsilon(y), \sigma(y), y)\mu_j(dy)\int_{|x|>z} p_j(x,y)\,dx.$$

From (4.2) and (4.3) we obtain

$$\sum_1^r \xi_j I_j \leqq \sum_{j=1}^r \xi_j \int_0^\infty dz\,\{\sup_{\delta\in\mathscr{I}}\int \sum_{k=1}^m \varepsilon_k(y) W(j,k,\sigma(y),y)\mu_j(dy)$$
$$\cdot\int_{|x|>z} p_j(x,y)\,dx\}$$
$$(4.4)\qquad \leqq \sum_{k=1}^m \sum_{n=1}^N \sum_{j=1}^r \xi_j \int_0^\infty dz \int W(j,k,n,y)\int_{|x|>z} p_j(x,y)\,dx$$
$$= \sum_{k=1}^m \sum_{n=1}^N \sum_{j=1}^r \xi_j \int W(j,k,n,y)\mu_j(dy)\int |x|p_j(x,y)\,dx$$
$$< \infty.$$

But each $\xi_j > 0$ so $I_j < \infty$ $(j = 1, \cdots, r)$.

The next lemma parallels Lemma 2.3.3 of Brown (1966). It shows that the moment condition of Lehmann and Stein (1953) implies Assumption (iii).

LEMMA 4.3. *In the fixed sample size case if*

$$\int W(j,k,y)\mu_j(dy)\int |x|p_j(x,y)\,dx < \infty$$

for $j = 1, \cdots, r$; $k = 1, \cdots, m$, *then Assumption* (iii) *is satisfied.*

PROOF. The proof of Lemma 4.2 will suffice with the modification that in (4.4) the sum over n contains only the term corresponding to the fixed sample size.

We now consider the interesting case in which

$$(4.5)\qquad W(j,k,n,y) = W_1(j,k) + W_2(n).$$

Let $\mathscr{I}_n$ be the set of invariant behavioral decision rules depending only on $(X, Y_1, \cdots, Y_{n-1})$. The final lemma of this section parallels Lemma 2.3.4 of Brown (1966).

LEMMA 4.4. *Let W satisfy* (4.5). *Suppose for some $c > 0$ that $W_2(n) - W_2(n-1) \geqq c$ for $n = 1, 2, \cdots$ and that $W_2(n) = O(n^\beta)$ as $n \to \infty$ for some $\beta \geqq 1$. Assume, for some $\alpha > 0$, that*

$$\int \mu_j(dy) \int |x|^{1+\alpha} p_j(x, y)\, dx < \infty \tag{4.6}$$

for $j = 1, \cdots, r$ and that, for some $\gamma > 0$, there exists a sequence $\varepsilon_n \in \mathscr{I}_n (n = 1, 2, \cdots)$ such that

$$\mu_j\{y: \textstyle\sum_{j'=1}^r \xi_{j'} \sum_{k=1}^m \varepsilon_{nk}(y) W_1(j, k) \int p_{j'}(x, y)\, dx > c/2\} = O(n^{-\beta(1+2/\alpha)-\gamma}) \tag{4.7}$$

as $n \to \infty$ for $j = 1, \cdots, r$. Then, Assumption (iii) *is satisfied.*

PROOF. Let $\delta^* = (\varepsilon^*, \sigma^*) \in \mathscr{I}$ where

$$\sigma^*(y) = \min\{n: \textstyle\sum_{j=1}^r \xi_j \sum_{k=1}^m \varepsilon_{nk}(y) W_1(j, k) \int p_{nj}(x, y)\, dx \leqq c/2\}$$

and $\varepsilon^*(y) = \varepsilon_n(y)$ for all y such that $\sigma^*(y) = n$. From (4.7) we obtain $\mu_j\{y: \sigma^*(y) \geqq n\} = O(n^{-\beta(1+2/\alpha)-\gamma})$ so that $\sigma^* < \infty$ a.e. $(\sum_1^r \mu_j)$ and

$$\int [W_2(\sigma^*(y))]^{1+2/\alpha} \mu_j(dy) < \infty. \tag{4.8}$$

If $(y_1, \cdots, y_{n-1})$ has been observed with $\sigma^*(y_1, \cdots, y_{n-1}) = n$, then the expected loss due to making some decision using ε_n is at most $c/2$ which is less than the cost of another observation. Hence, $\sigma_0 \leqq \sigma^*$ a.e. $(\sum_1^r \mu_j)$.

From Hölder's inequality, (4.6), (4.8) and the fact that $\sigma_0 \leqq \sigma^*$ a.e. (μ_j) there exists $b_1 < \infty$ such that

$$\begin{aligned} &\int W_2(\sigma_0(y)) \mu_j(dy) \int_{|x|>z} p_j(x, y)\, dx \\ &\quad \leqq \{\int [W_2(\sigma_0(y))]^{1+2/\alpha} \mu_j(dy)\}^{\alpha/(2+\alpha)} \cdot [\int \mu_j(dy) \int_{|x|>z} p_j(x, y)\, dx]^{2/(2+\alpha)} \\ &\quad \leqq b_1 z^{-2(1+\alpha)/(2+\alpha)}. \end{aligned} \tag{4.9}$$

Letting $W^* = \max_{j,k} W_1(j, k)$ and using (4.6) we see that there exists $b_2 < \infty$ such that

$$\begin{aligned} &\textstyle\sum_{k=1}^m W(j, k) \int \varepsilon_{0k}(y) \mu_j(dy) \int_{|x|>z} p_j(x, y)\, dx \\ &\quad \leqq W^* z^{-(1+\alpha)} \int \mu_j(dy) \int |x|^{1+\alpha} p_j(x, y)\, dx \\ &\quad \leqq b_2 z^{-(1+\alpha)}. \end{aligned} \tag{4.10}$$

Then, (4.5), (4.9) and (4.10) yield

$$\begin{aligned} I_j &\leqq \textstyle\int_0^\infty dz\, \{\sum_{k=1}^m W_1(j, k) \int \varepsilon_{0k}(y) \mu_j(dy) \int_{|x|>z} p_j(x, y)\, dx \\ &\quad + \int W_2(\sigma_0(y)) \mu_j(dy) \int_{|x|>z} p_j(x, y)\, dx\} \\ &\leqq \textstyle\int_1^\infty [b_2 z^{-(1+\alpha)} + b_1 z^{-2(1+\alpha)/(2+\alpha)}]\, dz + R_0(j) \\ &< \infty. \end{aligned}$$

The author has been unable to prove a lemma paralleling Lemma 2.3.5 of Brown (1966).

5. An example. Let Z_1, Z_2 be independent random variables each distributed normally with mean 0 and variance $\sigma_j^2/2$. Set $X = Z_1$ and $Y = Z_1 - Z_2$. Assume $\sigma_j = j$ for some $j = 1, 2, 3$. Assume the loss does not depend on Y and is given by the table below when action k is taken.

		k	
j	1	2	3
1	0	1	1
2	1	0	1
3	1	1	0

Let $\xi_j = \frac{1}{3}$ $(j = 1, 2, 3)$. The procedure which is Bayes in the class of invariant procedures (and, hence, admissible) makes decision 1 if $Y^2 \leqq \frac{8}{3} \log 2$, decision 2 if $\frac{8}{3} \log 2 < Y^2 \leqq 72/5 \log (\frac{3}{2})$, and decision 3 if $Y^2 > 72/5 \log (\frac{3}{2})$.

In this example only two observations have been taken for clarity of exposition. If more observations are taken, we set $Y_i = Z_1 - Z_{i+1} (i = 1, \cdots, n-1)$ and then our decision rule is of a similar form using $\sum_{i=1}^{n} Y_i^2$.

Setting $r = m$ is also not needed. We can, for example, have a fourth possible decision with constant loss a. If a is sufficiently small (less than the maximum posterior probability given Y of an "error" using our procedure) then for some values of Y this decision will be made. For a smaller than the minimum posterior probability given Y of an "error," the fourth decision will always be made.

REFERENCES

BROWN, LAWRENCE DAVID (1966). On the admissibility of invariant estimators of one or more location parameters. *Ann. Math. Statist.* **37** 1087–1136.

FOX, MARTIN (1970). Admissibility of multiple decision procedures Bayes in the class invariant under a real location parameter. Michigan State University RM-251.

FOX, MARTIN 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.

LEHMANN, E. L. and STEIN, C. M. (1953). The admissibility of certain invariant statistical tests involving a translation parameter. *Ann. Math. Statist.* **24** 473–479.

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