ON UNBIASED ESTIMATION[1]

By L. Schmetterer[2]

University of California, Berkeley

The theory of unbiased estimation has been mainly developed for quadratic loss-functions. The purpose of the present paper is to generalize this theory to convex loss-functions, and especially to loss-functions which are pth powers ($p \geqq 1$). The treatment of these cases needs in part quite different tools than in the quadratic case. Theorems of Stein and Bahadur are generalized. The contents of the paper have, however, some relations to results previously obtained by Barankin.

Let (R, S) be a measurable space and let $\mathfrak{P}$ be a nonempty class of probability measures P on S. Let g be any real valued function from $\mathfrak{P}$ into euclidean R_1. A real-valued measurable function on R for which $\int_R hdP$ exists for all $P \varepsilon \mathfrak{P}$ is called an unbiased estimator for g if

$$E(h; P) = \int_R hdP = g(P), \tag{1}$$

for all $P \varepsilon \mathfrak{P}$.

The set of all h's which satisfy (1) will be designated by H_g. Let $\omega(z)$ be any nonnegative Borel-measurable function defined on $-\infty < z < \infty$. Denote by $H_g(\omega; P)$ the set of all $h \varepsilon H_g$ for which $E(\omega(h - g(P)); P)$ with $P \varepsilon \mathfrak{P}$ exists.

Definition 1: $h_0 \varepsilon H_g(\omega; P_0)$ is called locally ω-minimal a $P_0 \varepsilon \mathfrak{P}$ if

$$E(\omega(h_0 - g(P_0));)P_0) \leqq E(\omega(h - g(P_0)); P_0)$$

for all $h \varepsilon H_g(\omega; P_0)$.

Definition 2: $h_0 \varepsilon \bigcap_{P \varepsilon \mathfrak{P}} H_g(\omega; P)$ is called uniformly ω-minimal if

$$E(\omega(h_0 - g(P)); P) \leqq E(\omega(h - g(P)); P)$$

for all $h \varepsilon \bigcap_{P \varepsilon \mathfrak{P}} H_g(\omega; P)$ and every $P \varepsilon \mathfrak{P}$.

If $\omega(z)$ is of the form $| z |^p$, $p \geqq 1$, then we shall also use the phrase p-minimal instead of ω-minimal. The significance of $H_g(p; P)$ is obvious.

The case $\omega(z) = z^2$ is frequently treated in the literature. Only a few papers exist which are occupied with more general loss functions $\omega(z)$. I refer in this connection to investigations by Barankin [1].

We now give

Definition 3. Let $V_p (p \geqq 1)$ be the class of all unbiased estimators v for $g \equiv 0$ such that $E(|v|^p; P)$ exists for all $P \varepsilon \mathfrak{P}$, and let $V_p^{P_0}$ be the class of all un-

Received April 17, 1959; revised June 20, 1960.

[1] Research done at the Adolf C. and Mary Sprague Miller Institute for Basic Research in Science, Berkeley, California.

[2] Now at the University of Hamburg.

biased estimators v for $g \equiv 0$ such that $E(|v|^p; P_0)$, $P_0 \varepsilon \mathfrak{P}$, exists. The class of all measurable functions h which satisfy $E(|h|^p; P) < \infty$ for all $P \varepsilon \mathfrak{P}$ will be denoted by E_p.

For any $p \geqq 1$ and any measurable function h on R we will write $\| h \|_{p,P}$ for $(\int_R |h|^p dP)^{1/p}$. The Banach space of all functions h with finite norm $\| h \|_{p,P}$ will be denoted by L_p^P.

In [2], the following theorem was proved by the author.

THEOREM 1. *$h_0 \varepsilon \bigcap_{P \varepsilon \mathfrak{P}} H_g(p; P)$ is uniformly p-minimal ($p > 1$) if and only if the Fréchet-differential, $dL(h_0 - g(P); v)$, of the norm $\| h_0 - g(P) \|_{p,P}$ vanishes for all $v \varepsilon V_p$ and each $P \varepsilon \mathfrak{P}$.*

Clearly, a similar theorem is valid for unbiased estimators which are locally p-minimal at P_0 replacing V_p by $V_p^{P_0}$.

Moreover, I will make use of the following theorem [2], [3, p. 63].

THEOREM 2. *If $\omega(z)$ is strictly convex, then there exists at most one unbiased estimator which is locally or uniformly ω-minimal.*

REMARK. Clearly, the exact meaning of Theorem 2 is the following: If $h_0 \varepsilon H_g(\omega; P_0)$ is locally ω-minimal in P_0, then, for any other locally ω-minimal $h \varepsilon H_g(\omega; P_0)$, we have $P_0(\{h \neq h_0\}) = 0$, and, if $h_0 \varepsilon \bigcap_{P \varepsilon \mathfrak{P}} H_g(\omega; P)$ is uniformly ω-minimal, then, for all $P \varepsilon \mathfrak{P}$, we have $P(\{h \neq h_0\}) = 0$ for any other uniformly ω-minimal $h \varepsilon \bigcap_{P \varepsilon \mathfrak{P}} H_g(\omega; P)$. Similar remarks apply to analogous cases. We shall now prove

THEOREM 3. *Let $\mathfrak{P}$ be dominated by a probability measure μ with $\mu \varepsilon \mathfrak{P}$. The generalized density $dP/d\mu$ of $P \varepsilon \mathfrak{P}$ will be denoted by f_P. Suppose that $f_P \varepsilon L_q^\mu (q > 1)$ for all $P \varepsilon \mathfrak{P}$. Let G be the set of all real-valued functions g_k on $\mathfrak{P}$ of the form $P \rightarrow E(k; P)$ with $k \varepsilon L_p^\mu$ and $1/p + 1/q = 1$. $h \varepsilon H_g(p; \mu)$ with $g \varepsilon G$ is locally p-minimal at μ if and only if there exists a mapping T defined on G into the real numbers such that*

$$T(g_k) = \int_R k \, | h - g(\mu) |^{p/q} \operatorname{sgn} (h - g(\mu)) \, d\mu \tag{2}$$

for all $k \varepsilon L_q^\mu$. The value of the minimum is given by $T(g - g(\mu))$.

The proof is based on two lemmas.

LEMMA 1. *Let B be a Banach space. Denote its norm by $\| \cdot \|$. Let B^* be the conjugate space of B and let $M^0 \subset B^*$ be the annihilator of M, where M is a closed linear manifold of B. Let $Q = B/M$ be the quotient space of B and M and let φ be the canonical mapping of B onto Q. Introducing the norm*

$$\| y \| = \inf_{\varphi(x) = y} \| x \|,$$

Q also becomes a Banach space. Let Q^* be the conjugate space of Q. The mapping φ^*, the transformation adjoint to φ, is a one-to-one linear and isometric mapping of Q^* onto M^0 [4, p. 115].

LEMMA 2. *V_p^μ is a closed linear manifold of L_p^μ.*

PROOF. It is clear that V_p^μ is a linear manifold. Moreover V_p^μ is closed in L_p^μ because strong convergence in L_p^μ implies weak convergence.

Proof of Theorem 3. First, let T be a mapping of G into the set of real numbers, which satisfies (2) for some $h \varepsilon H_g(p; \mu)$. Choose $B = L_p^\mu$ and $M = V_p^\mu$. There is a one-to-one correspondence between G and the set of all classes $H_{g_k}(p;\mu)$ (with $k \varepsilon L_p^\mu$). Thus, there is a one-to-one correspondence between G and $Q = L_p^\mu / V_p^\mu$. Let us now consider T as a functional on Q. Clearly, T must be linear and bounded. Now an application of Lemma 1 shows that necessarily

$$\int_R v \mid h - g(\mu) \mid^{p/q} \operatorname{sign}\,(h - g(\mu))\, d\mu = 0 \tag{3}$$

for all $v \varepsilon V_p^\mu$. But Theorem 1 implies that h is locally p-minimal in μ. On the other hand, if $h \varepsilon H_g(p; \mu)$ is locally p-minimal we again have (3) for all $v \varepsilon V_p^\mu$ according to Theorem 1. Denote the linear functional defined by

$$\int_R k \mid h - g(\mu) \mid^{p/q} \operatorname{sign}\,(h - g(\mu))\, d\mu$$

for all $k \varepsilon L_p^\mu$ by L. We can define T by $\varphi^{*-1}(L)$. Clearly,

$$T(g - g(\mu)) = \int_R \mid h - g(\mu) \mid^p d\mu.$$

For the case $p = 2$, Theorem 3 has been proved by Stein [5] by a different method.

Next we give

Definition 5. Let $p > 1$ and $1/p + 1/q = 1$. We define a transformation N of L_q^μ to L_p^μ by $f \rightarrow \mid f \mid^{p/q} \operatorname{sgn} f$ for all $f \varepsilon L_p^\mu$. If f runs through a subset $C \subset L_p^\mu$ we write for the set of all Nf with $f \varepsilon C$ simply NC. Clearly, for all $k \varepsilon L_q^\mu$, $N^{-1}k$ exists and is given by $\mid k \mid^{q/p} \operatorname{sgn} k$.

It is not difficult to find applications of Theorem 3 which are generalizations of corresponding applications by Stein. This leads, e.g., to

Theorem 3′. *Let there be given a σ-algebra $\mathfrak{S}$ of subsets of $\mathfrak{P}$, let there be given a σ-finite totally additive (in general) signed measure m over $(\mathfrak{P}, \mathfrak{S})$, and suppose that f_P satisfies the conditions of Theorem 3. Suppose further that f_P, considered as a function on $R \times \mathfrak{P}$, is measurable. If $\int_R \mid k \mid \int_{\mathfrak{P}} f_P^q\, d \mid m \mid d\mu$ exists for all $k \varepsilon L_p^u$ and if $E(N^{-1} \int_{\mathfrak{P}} f_P\, dm; \mu) = 0$, then $N^{-1} \int_{\mathfrak{P}} f_P\, dm$ is locally p-minimal at μ.*

Proof. Denote the mapping $P \rightarrow E(k; P)$ with $k \varepsilon L_p^\mu$ by g_k. It is enough to observe that

$$T(g_k) = \int_{\mathfrak{P}} \int_R k f_P\, dm\, d\mu$$

for all $k \varepsilon L_p^\mu$ exists and satisfies the conditions of Theorem 3.

We will illustrate this theorem for the case $p = 3$ by a simple example which however is general enough to serve as a pattern for the general finite dimensional case.

Example 1: Suppose that $R = \{x_1, x_2, x_3, x_4\}$ is a finite set and S the set of all subsets of R. Define

$$P_1(x_i) = a_i, \quad a_i \geqq 0, \quad 1 \leqq i \leqq 4, \quad \sum_{l=1}^{4} a_i = 1.$$

$$P_2(x_i) = \alpha > 0, \quad i = 1, 3; \quad P_2(x_i) = \beta > 0, \quad i = 2, 4, \alpha + \beta = \tfrac{1}{2}$$

$$\mu(x_i) = \beta, \quad i = 1, 3; \quad \mu(x_i) = \alpha, \quad i = 2, 4$$

Let $\mathfrak{S}$ be the set of all subsets of $\mathfrak{P} = \{P_1, P_2, \mu\}$ and define the measure m by: $m(P_1) = 0$, $m(P_2) = \lambda_2$, $m(\mu) = \lambda_3$, where λ_2 and λ_3 are any real numbers.

Obviously P_1 and P_2 are dominated by μ and we have

$$f_{P_1}(x_i) = a_i/\beta, \quad i = 1, 3; \quad f_{P_1}(x_i) = a_i/\alpha, \quad i = 2, 4$$

$$f_{P_2}(x_i) = \alpha/\beta, \quad i = 1, 3; \quad f_{P_2}(x_i) = \beta/\alpha, \quad i = 2, 4$$

We will now determine unbiased estimators which are locally 3-minimal at μ.

We have: $\int_{\mathfrak{P}} f_P(x_i)\, dm = (\alpha/\beta)\lambda_2 + \lambda_3$, $i = 1, 3$, $\int_{\mathfrak{P}} f_P(x_i) dm = (\beta/\alpha)\lambda_2 + \lambda_3$, $i = 2, 4$.

According to Theorem 3′ we have to determine λ_2 and λ_3 in such a manner that

$$\beta \mid (\alpha/\beta)\, \lambda_2 + \lambda_3 \mid^{\frac{1}{2}} \operatorname{sgn} (\alpha/\beta)\lambda_2 + \lambda_3) \\ + \alpha \mid (\beta/\alpha)\lambda_2 + \lambda_3 \mid^{\frac{1}{2}} \operatorname{sgn} ((\beta/\alpha)\lambda_2 + \lambda_3) = 0$$

It follows by a simple calculation that, if y is any real number and if g is a function over $\mathfrak{P}$ defined by

$$g(P_1) = (\mid y \mid \mid \alpha^2 - \beta^2 \mid)^{\frac{1}{2}}(\beta^2 + \alpha^2)^{-1}((a_1 + a_3)(\alpha/\beta) \operatorname{sgn} (y(\alpha^2 - \beta^2) \\ + (a_2 + a_4)(\beta/\alpha) \operatorname{sgn} (y(\beta^2 - \alpha^2))$$

$$g(P_2) = (\mid y \mid \mid \alpha^2 - \beta^2 \mid)^{\frac{1}{2}}(\beta^2 + \alpha^2)^{-1}((\alpha^2/\beta) \operatorname{sgn} (y(\alpha^2 - \beta^2)) \\ + (\beta^2/\alpha) \operatorname{sgn} (y(\beta^2 - \alpha^2))$$

$$g(\mu) = 0,$$

then

$$h(x_i) = (\mid y \mid \mid \alpha^2 - \beta^2 \mid)^{\frac{1}{2}}\alpha/\beta(\beta^2 + \alpha^2) \operatorname{sgn} (y(\alpha^2 - \beta^2)), \qquad i = 1, 3,$$

$$h(x_i) = (\mid y \mid \mid \beta^2 - \alpha^2 \mid)^{\frac{1}{2}}\beta/\alpha(\beta^2 + \alpha^2) \operatorname{sgn} (y(\beta^2 - \alpha^2)), \qquad i = 2, 4$$

is the unbiased estimator for the function g which is locally 3-minimal at μ.

Clearly, if we had taken $m(P_1) = \lambda_1 \neq 0$, then we would have obtained a two-parametric class of unbiased estimators which are locally 3-minimal at μ for a corresponding two-parametric class of functions g which vanish at μ. Hence it is possible to determine the locally 3-minimal unbiased estimator for every function g on $\mathfrak{P}$ that vanishes at μ by solving an algebraic equation for $\lambda_1, \lambda_2, \lambda_3$, which is at most of the second degree (Cf. also example 2).

Let G have the same significance as in Theorem 3 and let G_0 be the subset of all functions $g \,\varepsilon\, G$ with $g(\mu) = 0$. We denote the set of all unbiased estimators for $g \varepsilon G_0$ which are locally p-minimal at μ by $T^{\mu}_{p,0}$ and the corresponding set for all $g \varepsilon G$ by T^{μ}_{p}.

We now prove

THEOREM 4. *Suppose that* $\mathfrak{P}$ *satisfies the conditions of Theorem* 3. $T^{\mu}_{p,0}$ *can be mapped by a one-to-one transformation onto a subset* W *of a closed linear manifold* $U \subset L^{\mu}_{q}$ *where* U *is the closed linear manifold spanned by all* f_P *and* W *is the set of all* $k \varepsilon U$ *with* $E(N^{-1}k; \mu) = 0$. $U \cap NH_g(p; \mu)$ *contains for each* $g \varepsilon G_0$ *exactly one element* k_g *and* $N^{-1}k_g$ *is an unbiased estimator for* g *and locally* p*-minimal at* μ.

Of course, this theorem is strongly related to Theorem 3. First we formulate a theorem of Barankin [1] as

LEMMA 3. *Suppose that* $\mathfrak{P}$ *is dominated by* μ *with* $\mu \varepsilon \mathfrak{P}$. *Suppose further that* $f_P \varepsilon L^{\mu}_{p}(1 \leqq q < \infty)$ *for all* $P \varepsilon \mathfrak{P}$. *Then there exists for each nonempty class* $H_g(p; \mu)$ *at least one unbiased estimator which is locally* p*-minimal at* μ, *where* $1/p + 1/q = 1$.

PROOF OF THE THEOREM. Let $k \varepsilon U$ and so

$$E(N^{-1}k; \mu) = 0. \tag{3'}$$

According to Definition 3 we have for each $v \varepsilon V^{\mu}_{p}$ and all f_P

$$\int_R vf_P \, d\mu = 0. \tag{4}$$

If

$$k = \sum_{i=1}^{n} \alpha_i f_{P_i} \tag{5}$$

for any natural n, any real number α_i and $P_i \varepsilon \mathfrak{P}$, $i = 1, \cdots, n$, then (4) implies

$$\int_R vk \, d\mu = 0 \tag{6}$$

for all $v \varepsilon V^{\mu}_{p}$. If k satisfies condition (3′), then an application of Theorem 1 shows that $N^{-1}k$ is locally p-minimal at μ.

If $\| k_n - k \|_{q,\mu} \to 0$, where the k_n are of the form (5), and if k fulfills (3′), then k also satisfies (6) for all $v \varepsilon V^{\mu}_{p}$. This implies that $N^{-1}k$ is locally p-minimal at μ.

Now we have to show that $U \cap NH_g(p; \mu)$ is not empty for every $g \varepsilon G_0$. An application of Theorem 2 entails that this intersection contains at most one element. There exists, according to Lemma 3, an element $h \varepsilon H_g(p; \mu)$ which is locally p-minimal at μ. Moreover, Barankin has proved the existence of a sequence $k_n \varepsilon U$, such that

$$\int_R k_n h \, d\mu \to \| h \|^{p}_{p,\mu} \quad \text{and} \quad \| k_n \|_{q,\mu} \to \| Nh \|_{q,\mu}.$$

It follows, using a theorem of Radon [6], that $\| k_n - Nh \|_{q,\mu} \to 0$.

COROLLARY. *For* $p = 2$, U *and* T^{μ}_{2} *are identical* [7].

This follows from the fact that N is the identity for $p = 2$.

Let us denote by $\tilde{T}_p^P$ the set of all estimators which are p-minimal at $P \varepsilon \mathfrak{P}$ for some real-valued mapping on $\mathfrak{P}$.

THEOREM 5. *Let $\mathfrak{P}$ be any (not necessarily dominated) set of probablity measures defined on S and suppose $P \varepsilon \mathfrak{P}$. If $h \varepsilon \tilde{T}_p^P$, then, for any constant λ, $h + \lambda$ and λh are also in $\tilde{T}_p^P$. $\tilde{T}_p^P$ is in general not linear.*

PROOF. The first positive part of the theorem is a trivial application of Theorem 1. Further, it is almost obvious that $\tilde{T}_p^P$ for $p \neq 2$ is not linear. We consider a simple example.

EXAMPLE 2: Let $t_1, t_2, t_3 (0 < t_i < 1)$ be a set of three real numbers including $\frac{1}{2}$. Let $a_1, \cdots, a_4$ be any different real numbers. Let $\mathfrak{P} = (P_{t_1}, P_{t_2}, P_{t_3})$ be given by $P_{t_i}(a_1) = (1 - t_i)^2$, $P_{t_i}(a_2) = t_i - t_i^2/2$, $P_{t_i}(a_3) = t_i - t_i^2$, $P_{t_i}(a_4) = t_i^2/2$ and $P_{t_i}(M) = 0$ for each set M of real numbers which does not contain at least one of the numbers $a_1, \cdots, a_4$.

Consider the two functionals on $\mathfrak{P}$, $g_1(t_i) = t_i$, $g_2(t_i) = t_1^2$, $i = 1, 2, 3$. It is easy to see that the set H_{g_1} consists of the following functions:

$$h^{(1)}(a_1) = 0, h^{(1)}(a_2) = 1 - x, h^{(1)}(a_3) = x, h^{(1)}(a_4) = 1 + x, \quad -\infty < x < \infty.$$

For the determination of the unbiased estimator $h_0^{(1)}$ which is locally 3-minimal at $P_{\frac{1}{2}}$ one obtains the equation

$$\tfrac{3}{8}(\tfrac{1}{2} - x)^2 \operatorname{sgn}(x - \tfrac{1}{2}) + \tfrac{1}{4}(x - \tfrac{1}{2})^2 \operatorname{sgn}(x - \tfrac{1}{2}) + \tfrac{1}{8}(\tfrac{1}{2} + x)^2 \operatorname{sgn}(\tfrac{1}{2} + x) = 0$$

and $h_0^{(1)}$ is determined by the solution $x_0^{(1)} = (3 - (5)^{\frac{1}{2}})/4$. The set H_{g_2} consists of the following functions

$$h^{(2)}(a_1) = 0, h^{(2)}(a_2) = -x, h^{(2)}(a_3) = x, h^{(2)}(a_4) = 2 + x, \quad -\infty < x < \infty.$$

For the determination of $h_0^{(2)}$ we must consider the equation

$$\tfrac{2}{8}(x + \tfrac{1}{4})^2 \operatorname{sgn}(x + \tfrac{1}{4}) + \tfrac{1}{4}(x - \tfrac{1}{4})^2 \operatorname{sgn}(x - \tfrac{1}{4}) + \tfrac{1}{8}(\tfrac{7}{4} + 4)^2 \operatorname{sgn}(\tfrac{7}{4} + x) = 0$$

and the relevant solution of this equation is given by $x_0^{(2)} = (3 - (53)^{\frac{1}{2}})/8$.

Finally, let $g_3(t_i) = g_1(t_i) + g_2(t_i)$, $i = 1, 2, 3$. The set H_{g_3} consists of the functions

$$h^{(3)}(a_1) = 0, h^{(3)}(a_2) = 1 - y, h^{(3)}(a_3) = y, h^{(3)}(a_4) = 3 + y, \quad -\infty < y < \infty.$$

The solution $y_0 = (9 - (141)^{\frac{1}{2}})/8$ of the equation

$$\tfrac{3}{8}(\tfrac{1}{4} - y)^2 \operatorname{sgn}(y - \tfrac{1}{4}) + \tfrac{1}{4}(y - \tfrac{3}{4})^2 \operatorname{sgn}(y - \tfrac{3}{4}) + \tfrac{1}{8}(\tfrac{9}{4} + y)^2 \operatorname{sgn}(\tfrac{9}{4} + y) = 0$$

determines $h_0^{(3)}$. Clearly, $h_0^{(1)} + h_0^{(2)} \neq h_0^{(3)}$.

THEOREM 6. *Suppose that $\mathfrak{P}$ satisfies the conditions of Theorem 3. Then T_p^μ is closed in L_p^μ.*

We need the following

LEMMA 4. *Let f_1 and f_2 be in L_p^P. We have the inequality*

$$(7) \qquad \int_R | Nf_1 - Nf_2 |^q \, dP \leqq C(p, \| f_1 \|_{p,P}, \| f_2 \|_{p,P}) \| f_1 - f_2 \|_{p,P}$$

where $p > 1$ and $1/p + 1/q = 1$ and $C(p, \| f_1 \|_{p,P}, \| f_2 \|_{p,P}) = 2^{q+1}p(\| f_1 \|_{p,P} + \| f_2 \|_{p,P})^{p/q}$.

PROOF. For $r \geqq 1$ and any real numbers y, z the following inequalities are valid:

$$(8) \qquad | y - z |^r \leqq 2^r \, | \, | y |^r \operatorname{sgn} y - | z |^r \operatorname{sgn} z \, |$$

and

$$(9) \qquad | \, | y |^r \operatorname{sgn} y - | z |^r \operatorname{sgn} z \, | \leqq 2r \, | y - z | \, (\, | y | + | z | \,)^{r-1}.$$

(For a proof, compare ([8], p. 221)).

We use first (8) for $y = | f_1 |^{p/q} \operatorname{sgn} f_1$, $z = | f_2 |^{p/q} \operatorname{sgn} f_2$ and $r = q$ and then (9) for $y = f_1$, $z = f_2$ and $r = p$ and so obtain

$$\| f_1 |^{p/q} \operatorname{sgn} f_1 - | f_2 |^{p/q} \operatorname{sgn} f_2 |^q \leqq 2^{q+1} \, p \, | f_1 - f_2 | \, (| f_1 | + | f_2 |)^{p-1}$$

(up to sets of P-measure 0 of course). Integrating, and applying Hölder's and Minkowski's inequalities, gives (7).

PROOF OF THE THEOREM. Let $h_n \, \varepsilon \, T_p^\mu$ and $\| h_n - h \|_{p,\mu} \to 0$ for some $h \, \varepsilon \, L_p^\mu$. We have $g_n(P) = E(h_n \, ; P) \to E(h \, ; P) = g(P)$ for all $P \, \varepsilon \, \mathfrak{P}$ because $f_P \, \varepsilon \, L_q^\mu$. It follows that

$$\| h_n - g_n(\mu) - (h - g(\mu)) \|_{p,\mu} \to 0.$$

The inequality (7) of Lemma 4 implies

$$(10) \qquad \| N(h_n - g_n(\mu)) - N(h - g(\mu)) \|_{q,\mu} \to 0.$$

Now

$$\int_R v \, | h_n - g_n(\mu) |^{p/q} \operatorname{sgn} (h_n - g_n(\mu)) \, d\mu = 0$$

for all $v \, \varepsilon \, V_p^\mu$ and $n = 1, 2, \cdots$. Therefore, (10) implies

$$\int_R v \, | h - g(\mu) |^{p/q} \operatorname{sgn} (h - g(\mu)) \, d\mu = 0$$

for all $v \, \varepsilon \, V_p^\mu$.

It is well known that there is a strong relation between the concepts of sufficiency and of uniform ω-minimality. In this connection the following definition [9] is important.

DEFINITION 6. A subalgebra S_0 of S is called p-complete if zero is the unique S_0-measurable element of V_p.

There is the following important result [10], [11], [9] which we formulate as

LEMMA 5. *If there exists a sufficient and p-complete subalgebra S_0 of S for $\mathfrak{P}$, then* *an S_0-measurable uniformly p-minimal estimator exists for each g if* $\bigcap_{P \varepsilon \mathfrak{P}} H_g(\omega; P)$ *is non empty.*

For the case $p = 2$ Bahadur [7] gave an interesting inverse theorem. It seems that such a theorem does not exist in the more general cases. But by modifying Bahadur's ideas it is possible to give the following

THEOREM 7. *Let $\mathfrak{P}$ be any class of probability measures. Consider the set C_p of all characteristic functions of sets in S, which are uniformly p-minimal ($p > 1$). Denote by S_0 the smallest subalgebra of S, such that all functions of C_p are S_0-measurable. Then S_0 is a p-complete subfield and all S_0-measurable functions E_p (Definition 3) are uniformly p-minimal estimators.*

PROOF. Let $A \subset R$ be any set. We denote by c_A the characteristic function of this set. Consider now a set $A \varepsilon S$ and suppose that $c_A \varepsilon C_p$. Then we have for all $v \varepsilon V_p$ and each P $\int_R vN(c_A - P(A))\, dP = 0$. Suppose $0 < P(A) < 1$. It follows that

$$\int_A v[(1 - P(A))^{p/q} + (P(A))^{p/q}]\, dP = 0$$

for all $v \varepsilon V_p$ and each $P \varepsilon \mathfrak{P}$. This means $\int_A v\, dP = 0$ for all $v \varepsilon V_p$ and each $P \varepsilon \mathfrak{P}$, or

$$\int_R vc_A\, dP = 0. \tag{11}$$

Obviously, (11) holds also for the cases $P(A) = 0$ and $P(A) = 1$. We have $0 \leqq c_A \leqq 1$ and so by (11) $vc_A \varepsilon V_p$ for all $v \varepsilon V_p$. Now, if $B \varepsilon S$ is a different set with $c_B \varepsilon C_p$, we have instead of (11) $\int_R v\, c_B\, dP = 0$ for all $v \varepsilon V_p$ and each $P \varepsilon \mathfrak{P}$. It follows that

$$\int_R vc_A\, c_B\, dP = \int_R vc_{A \cap B}\, dP = 0 \tag{12}$$

for all $v \varepsilon V_p$ and each $P \varepsilon \mathfrak{P}$. Suppose $0 < P(A \cap B) < 1$. Consider

$$\int_R vN(c_{A\cap B} - P(A \cap B))\, dp$$

for a $v \varepsilon V_p$ and a $P \varepsilon \mathfrak{P}$. (12) implies that this integral vanishes and so $c_{A\cap B} \varepsilon C_p$.

Moreover, if $c_A \varepsilon C_p$ it follows that $1 - c_A = c_{R-A} \varepsilon C_p$ by using Theorem 5.

Finally, consider a denumerable class of sets $A_i \varepsilon S$, which are pairwise disjoint and so that $c_{A_i} \varepsilon C_p$. Denote $\bigcup_{i=1}^{\infty} A_i$ by A. We have $c_A = \sum_{i=1}^{\infty} c_{A_i}$ and so for all $P \varepsilon \mathfrak{P}$ $\| \sum_{i=1}^{n} c_{A_i} - c_A \|_{p,P} \to 0$. Theorem 6 gives $c_A \varepsilon C_p$. Thus, we have proved that the class of all sets $A \varepsilon S$ for which c_A belongs to C_p, forms a σ-algebra and obviously this must be S_0. It is easy to show that $\alpha_1 c_{A_1} + \alpha_2 c_{A_2}$ for any real numbers α_i and $c_{A_i} \varepsilon C_p$ is a uniformly p-minimal estimator. Let $h \varepsilon E_p$ and let h be S_0-measurable. Then there exists always a sequence of functions of the form $\sum_{i=1}^{k_n} \alpha_i c_{A_i}$, α_i real numbers, $c_{A_i} \varepsilon C_p$, which converge to h, and such that $\| \sum_{i=1}^{k_n} \alpha_i c_{A_i} - h \|_{p,P} \to 0$ for very $P \varepsilon \mathfrak{P}$. It follows that h is uniformly p-minimal.

If $h \varepsilon E_p$ is S_0-measurable and an unbiased estimator for $g \equiv 0$, h must be uniformly p-minimal and so equal to zero. Moreover it is easy to show that S_0 is necessary. (Cf. [7].)

Concerning sufficiency it is possible to show

THEOREM 8. *Let $\mathfrak{P}$ be a dominated class of probability measures, μ a measure equivalent to $\mathfrak{P}$, and $\mu \varepsilon \mathfrak{P}$. Suppose that $\mathfrak{P}$ is a convex set. Consider the set $T_{p,b}$ of all bounded uniformly p-minimal estimators, and denote by S^0 the smallest sub-algebra of S such that all elements of $T_{p,b}$ are S^0-measurable. We assume further: If a real-valued function g on $\mathfrak{P}$ has a bounded unbiased estimator then it has also a uniformly p-minimal unbiased estimator. Then S^0 is sufficient for $\mathfrak{P}$.*

PROOF. We remark that the existence of a measure μ which is equivalent to $\mathfrak{P}$ can be proved in the dominated case [12]. Let P_1, $P_2 \varepsilon \mathfrak{P}$ and $P_1 \neq P_2$. The measure

$$\lambda = \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 \mu, \quad \alpha_i > 0, \quad \alpha_1 + \alpha_2 + \alpha_3 = 1$$

is equivalent to μ and so to $\mathfrak{P}$ and moreover, $\lambda \varepsilon \mathfrak{P}$. We have

$$1 = \alpha_1(dP_1/d\lambda) + \alpha_2(dP_2/d\lambda) + \alpha_3(d\mu/d\lambda).$$

It follows that $dP_i/d\lambda = f_{P_i}$, $i = 1, 2$ is bounded. Consider $E(f_{P_i}; P) = g_i(P)$ for all $P \varepsilon \mathfrak{P}$. By the boundedness of f_{P_i} there exist uniformly p-minimal unbiased estimators h_i for g_i.

Let V be the class of all unbiased estimators v for the zero-functional on $\mathfrak{P}$. We have

$$\int_R v f_{P_i} \, d\lambda = 0 \tag{13}$$

for all $v \varepsilon V$ and so for all $v \varepsilon V_p^\lambda$.

If $E(N^{-1} f_{P_i}; \lambda) = 0$ then according to Theorem 4, $N^{-1} f_{P_i}$ must be locally p-minimal at λ for g_i. But λ is equivalent to $\mathfrak{P}$ and moreover $N^{-1} f_{P_i} \varepsilon E_p$. Thus, we must have $N^{-1} f_{P_i} = h_i$ according to Theorem 2. Since $N^{-1} f_{P_i}$ is bounded, we have $h_i \varepsilon T_{p,b}$. Therefore, f_{P_i} is S^0-measurable. However, in general $E(N^{-1} f_{P_i}; \lambda) \neq 0$.

Let γ be any real number. By (13) we also have

$$\int_R v(f_{P_i} + \gamma) \, d\lambda = 0 \quad \text{for all } v \varepsilon V.$$

Consider

$$\int_R | f_{P_i} + \gamma |^{q/p} \operatorname{sgn} (f_{P_i} + \gamma) \, d\lambda.$$

It is easy to show that this integral is a continuous function η of γ for $-\infty <$

$\gamma < \infty$ by using Lemma 4. If $\gamma > 0$ is large enough, $\eta(\gamma)$ must be >0, because f_{P_i} is bounded. If $\gamma < 0$ and $|\gamma|$ is large enough, $\eta(\gamma)$ is < 0. Hence, there is at least one $\gamma = \gamma_0$ with $\eta(\gamma_0) = 0$.

We have to repeat the previous argument with f_{P_i} replaced by $(f_{P_i} + \gamma_0)$. We obtain again the result that f_{P_i} is S^0-measurable. Thus we have proved that S^0 is pairwise sufficient for $\mathfrak{P}$. This involves sufficiency for the dominated case [12].

REFERENCES

[1] E. W. Barankin, "Locally best unbiased estimates," *Ann. Math. Stat.*, Vol. 20 (1949), pp. 477–502.

[2] L. Schmetterer, "Bemerkungen zur Theorie der erwartungstreuen Schätzungen," *Mitteil.-Bl. math. Statistik*, Vol. 9 (1957), pp. 147–152.

[3] D. A. S. Fraser, *Nonparametric Methods in Statistics*, John Wiley and Sons, New York, 1957.

[4] N. Bourbaki, Livre V, *Espaces Vectoriels Topologiques*, Chapitre III-V, Hermann et Cie, Paris, 1955.

[5] Charles Stein, "Unbiased estimates with minimum variances," *Ann. Math. Stat.*, Vol. 21 (1950), pp. 406–415.

[6] J. Radon, "Theorie und Anwendungen der absolut additiven Mengenfunktionen," *Sitzungsberichte der math. naturwiss. Klasse der Akad. der Wiss*, Wien CXII, Abt. IIa (1913), pp. 1295–1438.

[7] R. R. Bahadur, "On unbiased estimates of uniformly minimum variance," *Sankhyā*, Vol. 18 (1957), pp. 211–224.

[8] N. Bourbaki, Livre VI, *Integration*, Chapitre I-IV (1952), Hermann et Cie, Paris.

[9] E. L. Lehmann and H. Scheffé, "Completeness, similar regions and unbiased estimation," *Sankhyā*, Vol. 10 (1950), pp. 305–339.

[10) C. Blackwell, "Conditional expectation and unbiased sequential estimation," *Ann. Math. Stat.*, Vol. 18 (1947), pp. 105–110.

[11] E. W. Barankin, "Extension of a theorem of Blackwell," *Ann. Math. Stat.*, Vol. 21 (1950), pp. 280–284.

[12] R. P. Halmos and L. J. Savage, "Application of the Radon-Nikodym theorem to the theory of sufficient statistics," *Ann. Math. Stat.*, Vol. 21 (1949), pp. 225–241.