## LOCALLY BEST UNBIASED ESTIMATES[1]

By E. W. Barankin

*University of California, Berkeley*

**Summary.** The problem of unbiased estimation, restricted only by the postulate of section 2, is considered here. For a chosen number $s > 1$, an unbiased estimate of a function $g$ on the parameter space, is said to be best at the parameter point $\theta_0$ if its $s$th absolute central moment at $\theta_0$ is finite and not greater than that for any other unbiased estimate. A necessary and sufficient condition is obtained for the existence of an unbiased estimate of $g$. When one exists, the best one is unique. A necessary and sufficient condition is given for the existence of only one unbiased estimate with finite $s$th absolute central moment. The $s$th absolute central moment at $\theta_0$ of the best unbiased estimate (if it exists) is given explicitly in terms of only the function $g$ and the probability densities. It is, to be more precise, specified as the l.u.b. of certain set $\mathfrak{A}$ of numbers. The best estimate is then constructed (as a limit of a sequence of functions) with the use of only the data (relating to $g$ and the densities) associated with any particular sequence in $\mathfrak{A}$ which converges to the l.u.b. of $\mathfrak{A}$.

The case $s = \infty$ is considered apart. The case $s = 2$ is studied in greater detail. Previous results of several authors are discussed in the light of the present theory. Generalizations of some of these results are deduced. Some examples are given to illustrate the applications of the theory.

**1. Introduction.** Let $\Omega$ be a space of points $x$, and $\mu$ be a totally additive measure defined on a $\sigma$-field $\mathcal{F}$ of subsets of $\Omega$. Let $\mathfrak{P} = \{p_\theta, \theta \in \Theta\}$ be a family of probability densities in $\Omega$ with respect to the measure $\mu$. $\Theta$ is any index set; we lay down no conditions on its structure. We are concerned here with the existence and characterization of unbiased estimates of a real-valued function $g$ on $\Theta$, which are in some suitable sense "best" for a prescribed parameter point $\theta_0$. That is, a real-valued, measurable ($\mu$) function $f_0$ on $\Omega$ such that

$$(1) \qquad \int_\Omega f_0 p_\theta \, d\mu = g(\theta), \qquad \theta \in \Theta,$$

and which satisfies a specified criterion of bestness for $\theta = \theta_0$. This criterion is usually taken to be

$$(2) \qquad \int_\Omega (f_0 - g(\theta_0))^2 p_{\theta_0} \, d\mu \leqq \int_\Omega (f - g(\theta_0))^2 p_{\theta_0} \, d\mu, \qquad f \in \overline{\mathfrak{M}},$$

where $\overline{\mathfrak{M}}$ denotes the class of all unbiased estimates of $g$; i.e., the class of all $f$ satisfying (1). The obvious advantage in the definition (2) is the algebraic

[1] This article was prepared while the author was under contract with the Office of Naval Research.

pliability. The obvious disadvantage is that $\overline{\mathfrak{M}}$ may contain no estimate with finite variance (cf. section 9).

For the investigation of the fundamental questions, posed above, relating to unbiased estimates, we shall not restrict ourselves to (2). We consider chosen and fixed, a number $s > 1$, and lay down the

DEFINITION. *$f_0 \in \overline{\mathfrak{M}}$ is best at $\theta_0$ if*

$$\infty > \int_\Omega |f_0 - g(\theta_0)|^s p_{\theta_0}\, d\mu \leqq \int_\Omega |f - g(\theta_0)|^s p_{\theta_0}\, d\mu, \qquad f \in \overline{\mathfrak{M}}.$$

With this, and under the condition of a rather natural postulate on $\mathfrak{P}$ (cf. section 2), we exhibit a necessary and sufficient condition for the existence of an unbiased estimate of $g$ having a finite $s$th absolute central moment at $\theta_0$.[2]

Except for the discussion, in section 3, of the case in which $g$ is constant on $\Theta$, we do not consider directly the estimation of $g$, but rather that of $h = g - g(\theta_0)$. Lemma 1, of section 2, gives the solution of the problem for $g$ when that for $h$ is known. After section 3, it is assumed exclusively that $h$ is not $\equiv 0$, except where the contrary is explicitly stated.

In case $s$ is finite, the existence theorem section 4, Theorem 2, asserts also the uniqueness of the best unbiased estimate of $h$. It is interesting to observe the similarity between the proof of this uniqueness and Fisher's proof of the (what might be called) asymptotic uniqueness of an efficient estimator [2 pp. 704, 705]. The case $s = \infty$[3] is discussed in section 5; in this case we find that, in general, the best estimate is not unique. However, for $s$ both finite and infinite, and as well when $g$ is constant ($\therefore h \equiv 0$), we give a necessary and sufficient condition that there be a unique unbiased estimate with finite s.a.c.m.[4] (cf. section 4, Corollary 2-1, and section 5, Theorem 3 (iii)).

Theorem 2 determines the s.a.c.m. of the best estimate as the l.u.b. of a set of numbers given explicitly; and thereby, in particular, throws open the class of all lower bounds of the minimum s.a.c.m. Investigations after such lower bounds, in the classical case $s = 2$, have led to the well-known results of Cramér-Rao [3 p. 480, (32.3.3)], and Bhattacharyya [4, p. 3, (1.10)]. In section 6, which is devoted to obtaining various special lower bounds, we show how those particular bounds fall out. It should be remarked, however, that our conditions on $\mathfrak{P}$ are in general different from those of the above authors.

[2] For the case $s = 2$ an alternative existence condition, antedating these results, but not yet published, has been obtained by C. Stein.

[3] If we use, in the above definition, the $s$th root of the $s$th absolute central moment, instead of the latter itself, then the bestness criterion for $s = \infty$ is the limiting criterion for $s \to \infty$; viz.,

$$\infty > \operatorname*{ess.\,sup.}_{x \in \Omega} |f_0 - g(\theta_0)| \leqq \operatorname*{ess.\,sup}_{x \in \Omega} |f - g(\theta_0)|, \qquad f \in \overline{\mathfrak{M}},$$

where ess. sup. refers to the measure $\nu(A) = \int_A p_{\theta_0}\, d\mu$.

[4] The abbreviation s.a.c.m. will henceforth be used to indicate $s$th absolute central moment at $\theta_0$.

In section 7 we give, in Theorem 7 and its corollary, a construction of the best estimate, depending only on the knowledge of the minimum s.a.c.m. The latter, as indicated in the preceding paragraph, is always known independently of any knowledge of the best estimate. We use these results to obtain explicitly (Theorems 8 and 9) the best estimates, for arbitrary $s$, in two cases where we assume the minimum s.a.c.m. known. These cases, when $s = 2$, give the minimum variance as determined by the equality sign in the Cramér-Rao and Bhattacharyya inequalities, respectively.

Section 8 is given to a brief discussion of the special case $s = 2$. Finally, in section 9, we present a detailed study of an example.

At the suggestion of the referee we have added an appendix in which is given a brief running description of the fundamental ideas of Banach spaces that come into use here. The italicized phrases are those mentioned explicitly in the course of the paper.

We shall merely mention here certain points which will be elaborated further in future communications. (1) The general theory developed here pertains as well to sequential as to nonsequential estimation; one has only to make the proper identification of $\Omega$, $\mathcal{F}$, $\mu$, and $\mathfrak{P}$. Moreover, as applied to sequential estimation, the theory will determine the optimum stopping regions. (2) The discussion of section 5 below can be carried through with "ess. sup." referring to the measure $\mu$, and $\mathfrak{L}_1$ being the space of functions on $\Omega$ which are integrable ($\mu$); and for this, no restrictions whatsoever on the densities $p_\theta$ are required (cf. the postulate of section 2), since the $p_\theta$ are elements of this $\mathfrak{L}_1$ solely by virtue of their properties as probability densities. This development would, for example, be sufficient to yield the estimate of Girshick, Mosteller, and Savage [5] in the case of sequential binomial estimation. Also, this unrestricted analysis is fundamental for the problem of similar regions (a case of the bounded unbiased estimation of a constant function). (3) For any $s > 1$ it may be observed in the result of Theorem 7 below, that the best (at $\theta_0$) estimate depends only on a sufficient statistic; this is clear from Neyman's theorem on sufficient statistics [6], since the best estimate depends only on ratios of the density functions $p_\theta$. But more than this, Blackwell's method [7] of deriving a uniformly (over the parameter set) better unbiased estimate from a given unbiased estimate can be proved to remain valid also when the measure of dispersion is the $s$th absolute central moment, $s > 1$. And for this, the postulate of section 2 is not required. (4) Finally, we point out that, with the proper specializations of $\Theta$, Cramér's theorem on the ellipsoid of concentration [8], Bhattacharyya's multidimensional inequality [9], and the extensions of the Rao, Cramér, and Bhattacharyya bounds to sequential estimation—as, for example, by Blackwell and Girshick [1], Wolfowitz [10], and Seth [20]—can be drawn from Theorem 4 below.

The inspiration for the mode of analysis in the following pages, and the major part of its substance, come from F. Riesz: his book [11 Ch. III] and the article [12] (in particular sections 8–11 thereof). In strictly mathematical terminology, Theorems 2 and 3 are given in [11] for the sequence-spaces $\ell_r$; and

Theorem 2 in [12] for the spaces $\mathfrak{L}_r$ of functions on the real interval [0, 1] with Lebesgue integrable $r$th powers. The proofs are given there for the case of a denumerable set $\Theta$; in [12] an indication is given of the extension to a non-denumerable $\Theta$. Our proof of Theorems 2 and 3, however, follows that given by Banach [13, p. 74] for the case of denumerable $\Theta$. It is based on two results, a theorem of Hahn-Banach [13, p. 55, Theorem 4], and the representation theorem (suitable for the general type of $\mathfrak{L}_r$ that we consider) for bounded linear functionals on $\mathfrak{L}_r$ [14, p. 338, Theorem 46]. The first of these, and the representation theorem for any $r > 1$, spring in fact from the same article [12, p. 475] of Riesz. In the case $r = 1$, the representation theorem is due originally to Steinhaus [15]; in the case $r = 2$, it was developed simultaneously in 1907 by Riesz [16] and Fréchet [17].

Riesz' proofs of the sufficiency of the condition in Theorem 2 proceed by constructing an explicit sequence of functions on $\Omega$ which converge strongly in $\mathfrak{L}_s$ to the (in the present statistical terminology) best estimate. Precisely, if in Theorem 7 below, we take, for each $n = 1, 2, \cdots$, the numbers $\alpha_1^n, \alpha_2^n, \cdots, \alpha_{k_n}^n$ so that the expression

$$\frac{\left| \sum_{i=1}^{k_n} \alpha_i^n h(\theta_i^n) \right|}{\left\| \sum_{i=1}^{k_n} \alpha_i^n \pi_{\theta_n^i} \right\|_r}$$

is maximum, then the assertion of this theorem is that of Riesz. However, Theorem 7 is established here without this strict requirement on the $\alpha_i^n$. The dropping of this restriction was essential for the proofs of Theorems 8 and 9. The latter two theorems are, in fact, proved with the use of Corollary 7–1, which is an even stronger result than Theorem 7. This corollary falls out of the proof of Theorem 7 immediately, in consequence of our use of Lemma 2 for that proof. The lemma, moreover, eases the proof of Theorem 7 markedly, in doing away with the need for any differentiation.

**2. Preliminary considerations.** We begin then by introducing the absolutely continuous (with respect to $\mu$) measure, defined on $\mathcal{F}$,

$$\nu(A) = \int_A p_{\theta_0}\, d\mu, \qquad A \in \mathcal{F}.$$

A function $\phi$ is summable ($\nu$) over $\Omega$ if and only if $\phi \cdot p_{\theta_0}$ is summable ($\mu$) over $\Omega$; and we have

$$\int_\Omega \phi\, d\nu = \int_\Omega \phi \cdot p_{\theta_0}\, d\mu,$$

(cf. [18, pp. 36–38]). Assuming that each of the ratios

$$\pi_\theta(x) = \frac{p_\theta(x)}{p_{\theta_0}(x)}, \qquad \theta \in \Theta$$

is defined almost everywhere ($\mu$) throughout $\Omega$, it follows that $f$ is an unbiased estimate of $g$ if and only if

$$\int_\Omega f\pi_\theta \, d\nu = g(\theta), \qquad \theta \in \Theta. \tag{3}$$

We define

$$h(\theta) = g(\theta) - g(\theta_0).$$

Since

$$\int_\Omega \pi_\theta \, d\nu = 1, \qquad \theta \in \Theta,$$

it is clear from (3) that $f$ is an unbiased estimate of $g$ if and only if $f - g(\theta_0)$ is an unbiased estimate of $h$. Moreover, $f$ is best, for $g$, at $\theta_0$ if and only if $f - g(\theta_0)$ is best, for $h$, at $\theta_0$.

Define

$$r = \frac{s}{s-1},$$

and let $\mathfrak{L}_r$ and $\mathfrak{L}_s$ be the spaces, normed in the usual way, of real-valued functions on $\Omega$, with summable ($\nu$) absolute $r$th and $s$th powers, respectively. We denote the respective norms by $\| \ \|_r$ and $\| \ \|_s$; that is, if $\xi \in \mathfrak{L}_r$ and $\eta \in \mathfrak{L}_s$,

$$\| \xi \|_r = \left( \int_\Omega | \xi |^r \, d\nu \right)^{1/r},$$

and

$$\| \eta \|_s = \left( \int_\Omega | \eta |^s \, d\nu \right)^{1/s}.$$

We note that these spaces, for $s < \infty$, are weakly compact (cf. [21]). This property will be used in the proof of Theorem 7. Also, we shall make explicit use of the representation theorem for linear functionals on $\mathfrak{L}_r$ [14, p. 338, Theorem 46].

The assumptions on $\mathfrak{P}$, or on $\mathfrak{P}_0 = \{\pi_\theta, \theta \in \Theta\}$, will now be the following.

Postulate: *The functions $\pi_\theta$ are defined almost everywhere ($\mu$) in $\Omega$, and are elements of $\mathfrak{L}_r$.*

The foregoing considerations combine to give the following equivalence.

Lemma 1. *$\phi_0 + g(\theta_0)$ is an unbiased estimate of $g$, which is best at $\theta_0$, if and only if (i) $\phi_0$ satisfies the equations*

$$\int_\Omega \phi \cdot \pi_\theta \, d\nu = h(\theta), \qquad \theta \in \Theta, \tag{4}$$

*and (ii) when $\phi$ is any other function satisfying (4), we have*

$$\| \phi \|_s \geqq \| \phi_0 \|_s ;$$

*that is, if and only if* $\phi_0$ *is an unbiased estimate of h with minimum (finite) norm in* $\mathfrak{L}_s$. *The s.a.c.m. of* $\phi_0 + g(\theta_0)$ *is precisely* $\| \phi_0 \|_s$.

Starting with section 4, we shall deal directly with the estimation of $h$.

**3. The case of constant $g$.** Throughout the remainder (section 4 et seq.) of this article, the function $h$ is assumed, unless the contrary is explicitly stated, to be non-constant; that is, since $h(\theta_0) = 0$, not $\equiv 0$. We can, and shall in this section, obtain the results of the desired kind for the case of a constant function $g$, by a brief, direct attack.

Let $g(\theta) \equiv g_0$, a constant. Then of course $h(\theta) \equiv 0$. One unbiased estimate of $g$ is immediately obvious, viz., $f_1(x) \equiv g_0$. The s.a.c.m. of $f_1$ is 0.

There will exist other[5] unbiased estimates of $g$ with finite s.a.c.m. if and only if there exist non-null unbiased estimates, in $\mathfrak{L}_s$, of $0 \equiv h$. That is, by virtue of the isomorphism between $\mathfrak{L}_s$ and the space of linear functionals on $\mathfrak{L}_r$, there will exist an unbiased estimate of $g$ with finite s.a.c.m., distinct from $f_1$, if and only if there exists a non-null functional on $\mathfrak{L}_r$ which vanishes on the elements of $\mathfrak{P}_0 = \{\pi_\theta, \theta \in \Theta\}$. And a necessary and sufficient condition that such a functional exist is that $\mathfrak{P}_0$ be not a fundamental set in $\mathfrak{L}_r$ [13, p. 58, Theorem 7].

Observe finally that, in any case, $f_1$ is the unique unbiased estimate of $g$ with vanishing s.a.c.m.

We collect these results in the following statement.

THEOREM 1. *If* $g(\theta) \equiv g_0$, *a constant, then there is a unique best unbiased estimate of* $g$; *viz.,* $f_1(x) \equiv g_0$. *And the s.a.c.m. of* $f_1$ *is* 0.

*A necessary and sufficient condition that there exist no other unbiased estimates of g having finite s.a.c.m. is that the set* $\mathfrak{P}_0$ *be fundamental in* $\mathfrak{L}_r$.

As an illustration of the ideas of this section, consider the following example: $\Omega$ is the real interval [0, 1]; $\mu$ is Lebesgue measure; $\Theta$ is the set of non-negative integers; and

$$p_\theta(x) = (\theta + 1)x^\theta.$$

And take $\theta_0 = 0$. Then, $\nu$ is again Lebesgue measure, and $\pi_\theta = p_\theta$ for each $\theta$. For definiteness, take $r = 2$ (the results in this case are the same for any $r \geqq 1$). It is well-known that the non-negative integer powers of $x$ form a fundamental set in $\mathfrak{L}_2$ on a finite real interval. That is, if $\xi$ is a function on [0, 1], such that $\int_0^1 \xi^2 \, dx < \infty$, and if $\epsilon > 0$, then there exist an integer $n$ and coefficients $b_0$,

[5] That is, distinct from $f_1$ in the sense of $\mathfrak{L}_s$; or, equivalently, differing from $f_1$ on a set of positive ($\nu$) measure. Whenever, in the sequel, an equation $\xi_1 = \xi_2$ appears, for two functions $\xi_1$ and $\xi_2$ in $\mathfrak{L}_r$ or $\mathfrak{L}_s$, equality almost everywhere ($\nu$) in $\Omega$ will be understood. It is a consequence of our postulate that if two functions on $\Omega$ are equal almost everywhere ($\nu$), they are equal almost everywhere ($\nu'$), where $\nu'$ is anyone of the measures $\nu'(A) = \int_A p_{\theta'} \, d\mu$, $\theta' \in \Theta$.

$b_1, \cdots, b_n$ such that

$$\int_0^1 \left(\xi - \sum_{i=0}^{n} b_i x^i\right)^2 dx < \epsilon.$$

Hence, in this case an unbiased estimate with finite variance at $\theta = 0$ is unique (as well for a non-constant function $g$ as for one which is constant over $\Theta$; cf. section 4, Corollary 2-1).

**4. The main theorem for non-constant $h$.** We shall denote by $\mathfrak{M}_s$ the class (or the set in $\mathfrak{L}_s$) of all unbiased estimates of $h$ that belong to $\mathfrak{L}_s$.

THEOREM 2. (i) *A necessary and sufficient condition that $\mathfrak{M}_s$ be non-empty is that there exist a constant $C$ such that for every set of $n$ functions $\pi_{\theta_1}, \pi_{\theta_2}, \cdots, \pi_{\theta_n}$, in $\mathfrak{P}_0$, and every set of $n$ real numbers $a_1, a_2, \cdots, a_n$, we have, for every $n = 1, 2, \cdots$,*

$$\left| \sum_{i=1}^{n} a_i h(\theta_i) \right| \leqq C \left\| \sum_{i=1}^{n} a_i \pi_{\theta_i} \right\|_r. \tag{5}$$

(ii) *For every $\phi \in \mathfrak{M}_s$ we have $\| \phi \|_s \geqq C_0$, where $C_0$ is the g.l.b. of the set of admissible constants $C$ in* (5).

(iii) *If $\mathfrak{M}_s$ is non-empty there is a unique $\phi_0 \in \mathfrak{M}_s$ with $\| \phi_0 \|_s = C_0$. Thus, $\phi_0$ is the unique unbiased estimate of $h$ which is best at $\theta_0$.*

The non-constancy of $h$ clearly implies $C_0 > 0$.

The necessity of condition (5) is immediate. Suppose $\phi \in \mathfrak{M}_s$, so that $\phi$ satisfies equations (4); then, for any $\theta_1, \theta_2, \cdots, \theta_n$, and any real numbers $a_1, a_2, \cdots, a_n$,

$$\sum_{i=1}^{n} a_i h(\theta_i) = \int_\Omega \phi \cdot \sum_{i=1}^{n} a_i \pi_{\theta_i} \cdot d\nu.$$

By the Hölder inequality it follows that

$$\left| \sum_{i=1}^{n} a_i h(\theta_i) \right| \leqq \|\phi\|_s \cdot \left\| \sum_{i=1}^{n} a_i \pi_{\theta_i} \right\|_r.$$

Hence (5) is satisfied with $C = \| \phi \|_s$.

Part (ii) of the theorem is hereby proved as well.

Suppose $\mathfrak{M}_s$ non-empty, and $\phi_0, \phi_1$ in $\mathfrak{M}_s$, such that $\| \phi_0 \|_s = \| \phi_1 \|_s = C_0$. Then $1/2\ (\phi_0 + \phi_1) \in \mathfrak{M}_s$ and therefore

$$1/2\ \| \phi_0 + \phi_1 \|_s \geqq C_0.$$

But, by the Minkowski inequality,

$$1/2\ \| \phi_0 + \phi_1 \|_s \leqq 1/2\ (\| \phi_0 \|_s + \| \phi_1 \|_s) = C_0,$$

Hence

$$\| \phi_0 + \phi_1 \|_s = \| \phi_0 \|_s + \| \phi_1 \|_s.$$

This equality implies $\phi_1 = \alpha\, \phi_0$ for some positive $\alpha$. But since the norms of $\phi_0$ and $\phi_1$ are equal (and $\neq 0$) $\alpha$ must be unity. Thus the uniqueness of $\phi_0$ is proved.

It remains now to prove, assuming (5) satisfied, the existence of $\phi_0$. Consider the functional $F$ on $\mathfrak{P}_0$ defined by

$$F(\pi_\theta) = h(\theta).$$

The Hahn-Banach theorem alluded to in section 1 (viz., [13, p. 55, Theorem 4]) has precisely (5) as a necessary and sufficient condition for the existence of a linear functional $G$ on $\mathfrak{L}_r$ satisfying

$$\text{(a) } G(\pi_\theta) = h(\theta),\ \theta \in \Theta;$$
$$\text{(b) } \| G \| \leqq C;$$

where $\| G \|$ is the norm of $G$, i.e.,

$$\| G \| = \underset{\xi \in \mathfrak{L}_r}{\text{l.u.b.}} \frac{| G(\xi) |}{\| \xi \|_r}.$$

In particular, taking $C = C_0$, there is a linear functional $G_0$ on $\mathfrak{L}_r$ with

$$\text{(a}'\text{) } G_0(\pi_\theta) = h(\theta),\ \theta \in \Theta$$
$$\text{(b}'\text{) } \| G_0 \| \leqq C_0 .$$

But, for an element $\sum_{i=1}^{n} a_i \pi_{\theta i}$ in the linear manifold $[\mathfrak{P}_0]$ spanned by the $\pi_\theta$,

$$G_0 \left(\sum_i a_i \pi_{\theta_i}\right) = \sum_i a_i h(\theta_i),$$

so that

$$\| G_0 \| \geqq \underset{\xi \in [\mathfrak{P}_0]}{\text{l.u.b.}} \frac{| G_0(\xi) |}{\| \xi \|_r} = C_0 .$$

Hence (b′) is replaced by the precise statement

$$\text{(b}''\text{) } \| G_0 \| = C_0 .$$

Now the representation theorem for linear functionals on $\mathfrak{L}_r$ asserts the existence of $\phi_0 \in \mathfrak{L}_s$, such that

$$G_0(\xi) = \int_\Omega \phi_0 \cdot \xi \, d\nu,$$

and

$$\| \phi_0 \|_s = \| G_0 \| = C_0 .$$

This taken with (a′) establishes the existence of $\phi_0 \in \mathfrak{L}_s$ satisfying

$$\begin{cases} \displaystyle\int_\Omega \theta_0 \pi_\theta \, d\nu = h(\theta), & \theta \in \Theta \\ \| \phi_0 \|_s = C_0 . & \end{cases}$$

and this completes the proof of the theorem.

It is readily seen that $\mathfrak{M}_s$ will consist of more than just $\phi_0$ if and only if there exists a non-null functional on $\mathfrak{L}_r$ which vanishes on $\mathfrak{P}_0$. Our discussion in section 3 therefore enables us to assert the following.

COROLLARY 2-1. *$\mathfrak{M}_s$, when it is non-empty, consists of $\phi_0$ alone if and only if $\mathfrak{P}_0$ is fundamental in $\mathfrak{L}_r$.*

A word is in order concerning the following two consequences of the boundedness of the measure $\nu$: (i) if $\mathfrak{P}_0 \subset \mathfrak{L}_r$, then also $\mathfrak{P}_0 \subset \mathfrak{L}_{r'}$ for every $r' < r$; (ii) if $\phi \in \mathfrak{L}_s$ then also $\phi \in \mathfrak{L}_{s'}$ for every $s' < s$. Otherwise stated: (i′) if $\mathfrak{P}_0$ satisfies the postulate of section 2 for the number $r$, it likewise satisfies this postulate for every (admissible) $r' < r$; (ii′) if $\mathfrak{M}_s$ is non-empty, then $\mathfrak{M}_{s'}$ is non-empty for every $s' < s$. Regarding (i′) we shall make only the obvious remark that although $\mathfrak{P}_0$ satisfies the postulate for every $r' < r$, there may be values of $r' < r$ such that no $C$ for (5) exists; this will be exemplified in section 9. Where (ii′) is concerned, it is clear that the non-emptiness of $\mathfrak{M}_s$ will not necessarily imply that $\mathfrak{P}_0 \subset \mathfrak{L}_{s'/s'-1}$ for every $s' < s$, even though for every such $s'$ $\mathfrak{M}_{s'}$ is non-empty. If for every $\phi \in \Theta$ other than $\theta_0$ we have $\pi_\theta \notin \mathfrak{L}_{s'/s'-1}$, for some particular $s' < s$, then we may have the situation in which there are elements in $\mathfrak{M}_{s'}$ with norms arbitrarily close to 0. However, this cannot be the case if (a) for some $\theta$ other than $\theta_0$, $\pi_\theta \in \mathfrak{L}_{s'/s'-1}$, and (b) $h$ does not vanish identically on $\Theta'$, the set of those $\theta$ for which $\pi_\theta \in \mathfrak{L}_{s'/s'-1}$. For, when these two conditions are satisfied, Theorem 2 applies to $h$ as defined on $\Theta'$; consequently there is a positive lower bound for the $s'$—norms of the unbiased estimates of $h$ over $\Theta'$. And since every element of $\mathfrak{M}_s'$ is, in particular, an unbiased estimate of $h$ over $\Theta'$, it follows that the norms of those elements are bounded below by a positive number.

**5. The case $s = \infty$ $(r = 1)$.** Let $\mathfrak{M}_\infty$ denote the class of essentially bounded $(\nu)$ unbiased estimates of $h$; and let bestness at $\theta_0$ be defined with respect to the essential absolute suprema of the elements of this class. That is, the unbiased estimate $\phi_0$, of $h$, is best at $\theta_0$ if

$$\operatorname*{ess.\,sup.}_{x \in \Omega} |\phi_0(x)| < \infty,$$

and if, when $\phi$ is another unbiased estimate of $h$, we have

$$\operatorname*{ess.\,sup.}_{X \in \Omega} |\phi_0(x)| \leqq \operatorname*{ess.\,sup.}_{X \in \Omega} |\phi(x)|.$$

The fundamental postulate for the functions $\pi_\theta$ is, in this case, that $\mathfrak{P}_0 \subset \mathfrak{L}_1$.

Now, $\mathfrak{L}_\infty$, the space of essentially bounded, measurable $(\nu)$ functions on $\Omega$, normed by ess. sup., is the space of linear functionals on $\mathfrak{L}_1$ [14, p. 338]. Examination of the proof of Theorem 2 will show that that proof goes through also in the present case in all but one detail: we cannot here in general prove the uniqueness of the best estimate. The proof of uniqueness breaks down since the equality

$$\text{ess. sup. } |\phi_0(x) + \phi_1(x)| = \text{ess. sup. } |\phi_0(x)| + \text{ess. sup. } |\phi_1(x)|$$

does not imply that $\phi_1$ is a constant multiple of $\phi_0$. Of course, if $\mathfrak{P}_0$ is fundamental in $\mathfrak{L}_1$, we have a fortiori the uniqueness of the best estimate.

The results for the case $s = \infty$ are then the following.

THEOREM 3. (i) *A necessary and sufficient condition that $\mathfrak{M}_\infty$ be non-empty is that there exist a constant $C$ such that for every set of $n$ functions $\pi_{\theta_1}, \pi_{\theta_2}, \cdots, \pi_{\theta_n}$, in $\mathfrak{P}_0$, and every set of $n$ real numbers $a_1, a_2, \cdots, a_n$, we have, for every $n = 1, 2, \cdots$,*

$$\left| \sum_{i=1}^{n} a_i h(\theta_i) \right| \leqq C \left\| \sum_{i=1}^{n} a_i \pi_{\theta_i} \right\|_1 .$$

(ii) *For every $\phi \in \mathfrak{M}_\infty$ we have $\| \phi \|_\infty \geqq C_0$, where $C_0$ is the g.l.b. of the set of admissible constants $C$ above.*

(iii) *When $\mathfrak{M}_\infty$ is non-empty, it contains elements with norm equal to $C_0$. These are the best* (at $\theta_0$) *unbiased estimates of $h$. When $\mathfrak{P}_0$ is not fundamental in $\mathfrak{L}_1$, there need not exist a unique best estimate.*

We close this section with the remark that Theorem 1 remains valid, as it stands, in the case $s = \infty$.

**6. Particular lower bounds for the minimum s.a.c.m.** In order to stress their significance in the statistical context, we shall give the statements of this section with the help of the symbol $\sigma_s(\phi)$ for the $s$th root of the s.a.c.m. of the unbiased estimate $\phi$, of $h$. We have of course, the relation

$$\sigma_s(\phi) = \| \phi \|_s .$$

Now, one of the most important aspects of Theorem 2 is that it presents us immediately with an explicit evaluation of the minimum $\sigma_s(\phi)$ for all $\phi \in \mathfrak{M}_s$. We state the formula in the form of a theorem.

THEOREM 4. *Let $\mathfrak{R}$ denote the set of all real numbers. Then,*

$$\underset{\phi \in \mathfrak{M}_s}{\text{g.l.b.}}\, \sigma_s(\phi) = \underset{\substack{\theta_1, \theta_2, \cdots, \theta_n \in \Theta \\ a_1, a_2, \cdots, a_n \in \mathfrak{R} \\ n = 1, 2, \cdots}}{\text{l.u.b.}} \frac{\left| \sum_{i=1}^{n} a_i h(\theta_i) \right|}{\left\| \sum_{i=1}^{n} a_i \pi_{\theta_i} \right\|_r}$$

For brevity, let us set

$$\underset{\phi \in \mathfrak{M}_s}{\text{g.l.b.}}\, \sigma_s(\phi) = \sigma_s^{\min}.$$

Since this theorem expresses $\sigma_s^{\min}$ as the l.u.b. of an explicit set of numbers, it is clear that *the class of all lower bounds of $\sigma_s^{\min}$ is thereby thrown open to us.* It follows that, when $s = r = 2$ and our hypotheses on $\mathfrak{P}$ are fulfilled, the classical lower bounds of Cramér-Rao [3, p. 480] and Bhattacharyya [4, p. 3] are particularized consequences of Theorem 4. In the results that follow here we shall indicate the deduction of those classical bounds. We need not, however, restrict $s$.

For a moment, let us denote by $\pi(x)$ the function on $\Theta$ which assigns the value $\pi_\rho(x)$ to the point $\rho \in \Theta$, and let $\Theta$ be an interval on the real axis. Then we shall, below, write $\pi_\theta'$ for the function (when it exists) on $\Omega$ which assigns the value $(d\pi(x)/d\rho)_{\rho=\theta}$ to $x \epsilon \Omega$. Similarly, $\pi''_\theta$ for the function assigning the value $(d^2\pi(x)/d\rho^2)_{\rho=\theta}$ to $x$; and so on.

THEOREM 5. *Suppose the following conditions fulfilled:*

(i) $\Theta = \mathcal{I}$, *an interval on the real axis;*

(ii) $h$ *is differentiable on* $\Theta' \subseteqq \mathcal{I}$;

(iii) *for each* $\theta \epsilon \Theta'$, $\pi'_\theta$ *is defined almost everywhere* ($\nu$), *and is an element of* $\mathfrak{L}_r$;

(iv) *for each* $\theta \epsilon \Theta'$,

$$\lim_{\rho \to \theta} \left\| \frac{\pi_\rho - \pi_\theta}{\rho - \theta} - \pi'_\theta \right\|_r = 0.$$

*Then, for any* $m + n$ ($m, n = 1, 2, \cdots$) *points* $\theta_1, \theta_2, \cdots, \theta_m$ *in* $\mathcal{I}$, *and* $\theta'_1, \theta'_2, \cdots, \theta'_n$ *in* $\Theta'$, *and any* $m + n$ *real numbers* $a_1, a_2, \cdots, a_m, b_1, b_2, \cdots, b_n$ *such that*

$$\left\| \sum_{i=1}^m a_i \pi_{\theta_i} + \sum_{i=1}^n b_i \pi'_{\theta'_i} \right\|_r \neq 0,$$

*we have*

$$\sigma_s^{\min} \geqq \frac{\left| \sum_{i=1}^m a_i h(\theta_i) + \sum_{i=1}^n b_i h'(\theta'_i) \right|}{\left\| \sum_{i=1}^m a_i \pi_{\theta_i} + \sum_{i=1}^n b_i \pi'_{\theta'_i} \right\|_r}. \tag{6}$$

The prime on the $h$ in (6) denotes the derivative of $h$.

To prove this theorem, observe first that by virtue of Theorem 4, we may write

$$\sigma_s^{\min} \geqq \frac{\left| \sum_{i=1}^m a_i h(\theta_i) + \sum_{i=1}^n b_i \frac{h(\rho_i) - h(\theta'_i)}{\rho_i - \theta'_i} \right|}{\left\| \sum_{i=1}^m a_i \pi_{\theta_i} + \sum_{i=1}^n b_i \frac{\pi_{\rho_i} - \pi_{\theta'_i}}{\rho_i - \theta'_i} \right\|_r}$$

for every set of points $\rho_1, \rho_2, \cdots, \rho_n$ in $\mathcal{I}$ such that the denominator of the right-hand side is defined and $\neq 0$. Therefore, also

$$\sigma_s^{\min} \geqq \lim_{\substack{\rho_i \to \theta_i \\ i=1,2,\cdots,n}} \frac{\left| \sum_{i=1}^m a_i h(\theta_i) + \sum_{i=1}^n b_i \frac{h(\rho_i) - h(\theta'_i)}{\rho_i - \theta'_i} \right|}{\left\| \sum_{i=1}^m a_i \pi_{\theta_i} + \sum_{i=1}^n b_i \frac{\pi_{\rho_i} - \pi_{\theta'_i}}{\rho_i - \theta'_i} \right\|_r}. \tag{7}$$

Now, by condition (iv), the element

$$\sum_{i=1}^m a_i \pi_{\theta_i} + \sum_{i=1}^n b_i \frac{\pi_{\rho_i} - \pi_{\theta'_i}}{\rho_i - \theta'_i}$$

of $\mathfrak{L}_r$ converges, in the strong sense in $\mathfrak{L}_r$, to

$$\sum_{i=1}^m a_i \pi_{\theta_i} + \sum_{i=1}^n b_i \pi'_{\theta'_i},$$

as $\rho_i \to \theta_i'$, $i = 1, 2, \cdots, n$. Consequently we have convergence of the norm; that is, the denominator of the right-hand side of (7) converges to the denominator of (6). (The latter is $\neq 0$, so that for all $\rho_i$ sufficiently close to $\theta_i'$, $i = 1, 2, \cdots, n$, the ratios in (7) are defined.) There is no difficulty about the convergence of the numerator of (7) to that of (6). The theorem is thus proved.

COROLLARY 5-1. *Under the hypothesis of Theorem* 5, *we have, in particular, when* $\theta_0 \in \Theta'$ *and* $\| \pi'_{\theta_0} \|_r \neq 0$,

$$\sigma_s^{\min} \geqq \frac{|h'(\theta_0)|}{\| \pi'_{\theta_0} \|_r}. \tag{8}$$

If we denote by $p$ the function on $\Omega \times \Theta$ which assigns the value $p_\theta(x)$ to the point $(x, \theta)$, and write (8) in the form

$$(\sigma_s^{\min})^r \geqq \frac{|h'(\theta_0)|^r}{\int_\Omega \left| \frac{\partial \log p}{\partial \theta} \right|^r_{\theta = \theta_0} p_{\theta_0} \, d\mu}, \tag{8'}$$

the generalization of the Cramér-Rao inequality afforded by (8) becomes evident.

Using the result and method of Theorem 5, we can establish the next in a hierarchy of theorems.

THEOREM 6. *Suppose the hypothesis of Theorem* 5 *satisfied, and the following condition fulfilled: for each* $\theta$ *in a non-empty subset* $\Theta''$ *of* $\Theta'$, (i) $h''(\theta)$ (*the second derivative*) *exists and* (ii) $\pi''_\theta$ *is defined almost everywhere* ($\nu$), *is an element of* $\mathfrak{L}_r$, *and satisfies*

$$\lim_{\rho \to \theta} \left\| \frac{\pi'_\rho - \pi'_\theta}{\rho - \theta} - \pi''_\theta \right\|_r = 0.$$

*Then, for any* $m + n + q (m, n, q = 1, 2, \cdots)$ *points* $\theta_1, \theta_2, \cdots, \theta_m$ *in* $\mathcal{I}$, $\theta_1', \theta_2', \cdots, \theta_n'$ *in* $\Theta'$, *and* $\theta_1'', \theta_2'', \cdots, \theta_q''$ *in* $\Theta''$, *and any* $m + n + q$ *real numbers* $a_1, a_2, \cdots, a_m, b_1, b_2, \cdots, b_n, c_1, c_2, \cdots, c_q$ *such that*

$$\left\| \sum_{i=1}^m a_i \pi_{\theta_i} + \sum_{i=1}^n b_i \pi'_{\theta_i'} + \sum_{i=1}^q c_i \pi''_{\theta_i''} \right\|_r \neq 0,$$

*we have*

$$\sigma_s^{\min} \geqq \frac{\left| \sum_{i=1}^m a_i h(\theta_i) + \sum_{i=1}^n b_i h'(\theta_i') + \sum_{i=1}^q c_i h''(\theta_i'') \right|}{\left\| \sum_{i=1}^m a_i \pi_{\theta_i} + \sum_{i=1}^n b_i \pi'_{\theta_i'} + \sum_{i=1}^q c_i \pi''_{\theta_i''} \right\|_r}.$$

Just as in the case of the previous theorem, we have here an immediate corollary.

COROLLARY 6-1. *Under the hypothesis of Theorem* 6, *we have in particular, when* $\theta_0 \in \Theta' \cdot \Theta''$,

$$\sigma_s^{\min} \geqq \frac{|bh'(\theta_0) + ch''(\theta_0)|}{\| b\pi'_{\theta_0} + c\pi''_{\theta_0} \|_r}, \tag{9}$$

*for any two real numbers, b and c, such that the denominator of the right-hand side does not vanish.*

Consider (9) in the particular case $s = r = 2$. In this case, (9) may be written, explicitly,

$$(10)\qquad (\sigma_2^{\min})^2 \geqq \frac{|bh'(\theta_0) + ch''(\theta_0)|^2}{\int_\Omega \frac{1}{p_{\theta_0}}\left(b\frac{\partial p}{\partial\theta} + c\frac{\partial^2 p}{\partial\theta^2}\right)^2_{\theta=\theta_0} d\mu}.$$

In particular, (10) holds for values of $b$ and $c$ which maximize the right-hand side. And that maximum value is found, in the usual way, to be

$$J^{11}[h'(\theta_0)]^2 + 2J^{12}h'(\theta_0)h''(\theta_0) + J^{22}[h''(\theta_0)]^2,$$

where the matrix

$$\begin{pmatrix} J^{11} & J^{12} \\ J^{12} & J^{22} \end{pmatrix}$$

is the inverse of the matrix

$$\begin{pmatrix} \int_\Omega \frac{1}{p_\theta}\left(\frac{\partial p}{\partial\theta}\right)^2 d\mu & \int_\Omega \frac{1}{p_\theta}\frac{\partial p}{\partial\theta}\frac{\partial^2 p}{\partial\theta^2} d\mu \\ \int_\Omega \frac{1}{p_\theta}\frac{\partial p}{\partial\theta}\frac{\partial^2 p}{\partial\theta^2} d\mu & \int_\Omega \frac{1}{p_\theta}\left(\frac{\partial^2 p}{\partial\theta^2}\right)^2 d\mu \end{pmatrix}.$$

Thus, we have

$$(11)\qquad (\sigma_2^{\min})^2 \geqq J^{11}[h'(\theta_0)]^2 + 2J^{12}h'(\theta_0)h''(\theta_0) + J^{22}[h''(\theta_0)]^2.$$

This is seen to be Bhattacharyya's result for the case of derivatives up to second order.

It is obvious how we extend Theorem 6 to obtain a similar result involving the functions $\pi_\theta, \pi_\theta', \pi_\theta'', \cdots, \pi_\theta^{(n)}$, for any assigned $n$. And it is thereafter clear how, in the case $s = r = 2$, Bhattacharyya's general inequality may be deduced.

It is clear that we can proceed from Theorem 4, under suitable conditions, to lower bounds for $\sigma_s^{\min}$ which involve integrals of the functions $\pi(x)$ (and the corresponding integrals of $h$) as well as the derivatives of these functions.

In closing this section we note that all the above considerations apply equally to the case $s = \infty$.

**7. Determination of the best estimate.** We shall now prove the following theorem, which provides an explicit construction of the best (at $\theta_0$) estimate of $h$. We repeat that $s$ is now taken to be finite.

THEOREM 7. *Let $\mathfrak{M}_s$ be non-empty, and $\phi_0$ be the best (at $\theta_0$) unbiased estimate of h. Let $\{\theta_i^n, i = 1, 2, \cdots, k_n\}, n = 1, 2, \cdots$, be a sequence of (finite) sets of points of $\Theta$, and $\{\alpha_i^n, i = 1, 2, \cdots, k_n\}, n = 1, 2, \cdots$, a sequence of sets of real numbers, such that*

$$\lim_{n\to\infty} \frac{\left|\sum_{i=1}^{k_n} \alpha_i^n h(\theta_i^n)\right|}{\left\|\sum_{i=1}^{k_n} \alpha_i^n \pi_{\theta_i^n}\right\|_r} = C_0 = \|\phi_0\|_s = \sigma_s^{\min}.$$

*Then the functions* $\zeta_n$ :

$$\zeta_n(x) = \frac{\sum_{i=1}^{k_n} \alpha_i^n h(\theta_i^n)}{\left\| \sum_{i=1}^{k_n} \alpha_i^n \pi_{\theta_i^n} \right\|_r^r} \cdot \left| \sum_{i=1}^{k_n} \alpha_i^n \pi_{\theta_i^n}(x) \right|^{r/s} \operatorname{sgn}\left( \sum_{i=1}^{k_n} \alpha_i^n \pi_{\theta_i^n}(x) \right)$$

*(are elements of* $\mathfrak{L}_s$ *and) converge strongly in* $\mathfrak{L}_s$ *to* $\phi_0$ .
The strong convergence here means precisely that

$$\lim_{n\to\infty} \int_\Omega | \zeta_n - \phi_0 |^s \, d\nu = 0.$$

Clearly, we may, with no loss in generality, assume the numbers $\alpha_i^n$ to be such that

$$\left\| \sum_{i=1}^{k_n} \alpha_i^n \pi_{\theta_i^n} \right\|_r = 1, n = 1, 2, \cdots . \tag{12}$$

*We shall suppose this to be the case throughout the proof.* Then the essential property of the $\theta_i^n$ and the $\alpha_i^n$ is that

$$\lim_{n\to\infty} \left| \sum_{i=1}^{k_n} \alpha_i^n h(\theta_i^n) \right| = C_0 . \tag{13}$$

And in this normalized situation, the functions $\zeta_n$ will be given by

$$\zeta_n(x) = \sum_{i=1}^{k_n} \alpha_i^n h(\theta_i^n) \cdot \left| \sum_{i=1}^{k_n} \alpha_i^n \pi_{\theta_i^n}(x) \right|^{r/s} \operatorname{sgn}\left( \sum_{i=1}^{k_n} \alpha_i^n \pi_{\theta_i^n}(x) \right). \tag{14}$$

That these functions are elements of $\mathfrak{L}_s$ is easily seen; in fact,

$$\| \zeta_n \|_s = \left| \sum_{i=1}^{k_n} \alpha_i^n h(\theta_i^n) \right| .$$

The proof of this theorem will consist mainly in the application of the following two lemmas.

LEMMA 2. *Let* $0 \neq \eta \in \mathfrak{L}_s$ , *and* $\{\xi_n , n = 1, 2, \cdots\}$ *be a sequence of functions in* $\mathfrak{L}_r$ *such that*

$$\text{(i)} \quad \| \xi_n \|_r = 1, \qquad n = 1, 2, \cdots$$

$$\text{(ii)} \quad \lim_{n\to\infty} \int_\Omega \xi_n \eta \, d\nu = \| \eta \|_s .$$

*Then* $\xi_n$ *converges strongly in* $\mathfrak{L}_r$ *to the function*

$$\xi_0 = \frac{1}{\| \eta \|_s^{s/r}} | \eta |^{s/r} \operatorname{sgn} \eta.$$

Let us observe first that

$$\int_\Omega \xi_0 \eta \, d\nu = \| \eta \|_s \tag{15}$$

and

$$\| \xi_0 \| = 1.$$

Furthermore, $\xi_0$ is the unique element with norm $\leqq 1$ in $\mathfrak{L}_r$ having the property (15). For, if also,

$$\int_\Omega \xi_0' \eta \, d\nu = \| \eta \|_s, \qquad \| \xi_0' \|_r \leqq 1,$$

we then have

$$\int_\Omega \tfrac{1}{2}(\xi_0 + \xi_0') \cdot \eta \, d\nu = \| \eta \|_s ;$$

and from this,

$$\tfrac{1}{2} \| \xi_0 + \xi_0' \|_r \| \eta \|_s \geqq \| \eta \|_s .$$

That is,

$$\| \xi_0 + \xi_0' \|_r \geqq 2 \geqq \| \xi_0 \|_r + \| \xi_0' \|_r .$$

From this, and (Minkowski)

$$\| \xi_0 + \xi_0' \|_r \leqq \| \xi_0 \|_r + \| \xi_0' \|_r ,$$

we have

$$\| \xi_0 + \xi_0' \|_r = \| \xi_0 \|_r + \| \xi_0' \|_r .$$

Therefore, for some $a > 0$, $\xi_0' = a\xi_0$. But we must have $a = 1$ if $\xi_0$ and $\xi_0'$ are both to satisfy (15), as assumed. Hence $\xi_0' = \xi_0$.

Now consider the sequence $\{\xi_n\}$. Choose a sub-sequence $\{\xi_{n_i}\}$ that converges weakly to, say, $\xi'$. Then $\| \xi' \|_r \leqq 1$. We have

$$\int_\Omega \xi' \eta \, d\nu = \lim_{i \to \infty} \int_\Omega \xi_{n_i} \eta \, d\nu = \| \eta \|_s .$$

Hence, $\xi' = \xi_0$. And since $1 = \| \xi_{n_i} \|_r \to 1 = \| \xi_0 \|_r$, it follows that $\xi_{n_i}$ converges strongly to $\xi_0$ (cf. [13, p. 139, section 3]).

Suppose there is a subsequence $\{\xi_{n_i}'\}$ of $\{\xi_n\}$ such that

$$\| \xi_{n_i}' - \xi_0 \| > \delta > 0, \qquad i = 1, 2, \cdots .$$

We have, nonetheless, for this subsequence, the hypotheses of our lemma satisfied. We can therefore apply the argument of the previous paragraph to extract a subsequence of $\{\xi_{n_i}'\}$, which converges strongly to $\xi_0$. This is in obvious contradiction to the above $\delta$-assumption, and the lemma is hereby proved.

LEMMA 3. *Lemma 2 remains true with the roles of $\mathfrak{L}_r$ and $\mathfrak{L}_s$ interchanged.*

This is obvious.

Returning now to the proof of Theorem 7, let us first, for the sake of brevity, introduce the notation:

$$c_n = \sum_{i=1}^{k_n} \alpha_i^n \, h(\theta_i^n),$$

$$\gamma_n = \operatorname{sgn}\left(\sum_{i=1}^{k_n} \alpha_i^n \, h(\theta_i^n)\right),$$

$$\psi_n = \sum_{i=1}^{k_n} \alpha_i^n \, \pi_{\theta_i^n}.$$

From

$$\int_\Omega \phi_0 \, \pi_\theta \, d\nu = h(\theta), \qquad \theta \in \Theta,$$

we easily obtain

$$\int_\Omega \phi_0 \, \psi_n \, d\nu = c_n, \qquad n = 1, 2, \cdots,$$

which we may write

$$\int_\Omega \phi_0 \cdot \gamma_n \, \psi_n \, d\nu = | \, c_n \, |, \qquad n = 1, 2, \cdots.$$

Since $| \, c_n \, | \to || \, \phi_0 \, ||_s$ (cf. (13)) and $|| \, \gamma_n \psi_n \, ||_r = 1$, $n = 1, 2, \cdots$, (cf. (12)), we have, by Lemma 2, that $\gamma_n \psi_n$ converges strongly to

$$\psi_0 = \frac{1}{C_0^{s/r}} \, | \, \phi_0 \, |^{s/r} \operatorname{sgn} \phi_0. \tag{16}$$

The functions (cf. (14))

$$\zeta_n = c_n \, | \, \psi_n \, |^{r/s} \operatorname{sgn} \psi_n$$

obviously satisfy

$$\int_\Omega \zeta_n \cdot \gamma_n \, \psi_n \, d\nu = | \, c_n \, |, \qquad n = 1, 2, \cdots$$

And from this we conclude that

$$\lim_{n \to \infty} \int_\Omega \zeta_n \, \psi_0 \, d\nu = C_0,$$

or

$$\lim_{n \to \infty} \int_\Omega \frac{\zeta_n}{| \, c_n \, |} \, \psi_0 \, d\nu = 1 = || \, \psi_0 \, ||_r.$$

We may apply Lemma 3 to this result, since $|| \, \zeta_n / | \, c_n \, | \, ||_r = 1$, $n = 1, 2, \cdots$. And we thereby conclude that $\zeta_n / | \, c_n \, |$ converges strongly to

$$| \, \psi_0 \, |^{r/s} \operatorname{sgn} \psi_0,$$

which, on substituting from the definition (16) of $\psi_0$, we find to be just

$$\frac{\phi_0}{C_0}.$$

Since $| c_n | \to C_0$, it follows immediately that $\zeta_n$ converges strongly to $\phi_0$, and the theorem is proved.

The following corollary is actually of greater use in applications than Theorem 7 itself, for the reason that it leaves no doubt about the form of $\lim \zeta_n$ (i.e., $\phi_0$) when we know explicitly the form of $\lim \gamma_n \psi_n$.

COROLLARY 7–1. *Assume the hypothesis of Theorem 7. Then the functions*

$$\frac{\operatorname{sgn}\left(\sum_{i=1}^{k_n} \alpha_i^n h(\theta_i^n)\right)}{\left\| \sum_{i=1}^{k_n} \alpha_i^n \pi_{\theta_i^n} \right\|_r} \cdot \sum_{i=1}^{k_n} \alpha_i^n \pi_{\theta_i^n}$$

*converge strongly, in $\mathfrak{L}_r$, to a function $\psi_0$, and*

$$\phi_0 = C_0 | \psi_0 |^{r/s} \operatorname{sgn} \psi_0 .$$

This is clear from the proof of the theorem.

By way of illustrating the application of these results, we shall prove the following theorem.

THEOREM 8. *Assume the hypothesis of Theorem 5. And, further, let the equality sign hold in* (8). *Then,*

$$\phi_0(x) = \frac{h'(\theta_0)}{\| \pi'_{\theta_0} \|_r^r} \cdot | \pi'_{\theta_0}(x)|^{r/s} \operatorname{sgn} \pi'_{\theta_0}(x).$$

Since (8) is an equality, we may under the hypothesis of Theorem 5, consider that we have

$$C_0 = \lim_{n\to\infty} \frac{\left| \frac{1}{\rho_n - \theta_0} h(\rho_n) - \frac{1}{\rho_n - \theta_0} h(\theta_0) \right|}{\left\| \frac{1}{\rho_n - \theta_0} \pi_{\rho_n} - \frac{1}{\rho_n - \theta_0} \pi_{\theta_0} \right\|_r}. \tag{17}$$

where $\{\rho_n\}$ is a sequence in $\mathcal{J}$ converging to $\theta_0$. The numerator of the right-hand side of (17), sans the vertical bars, converges to $h'(\theta_0)$ (which is $\neq 0$, since $C_0 \neq 0$); hence, for all sufficiently large $n$, that expression has the signum of $h'(\theta_0)$. The functions whose norms appear in the denominator of (17) we know to converge strongly in $\mathfrak{L}_r$ to $\pi'_{\theta_0}$ (by the hypothesis of Theorem 5). Hence, for this case, the function $\psi_0$ of Corollary 7-1 is

$$\psi_0 = \frac{\operatorname{sgn} h'(\theta_0)}{\| \pi'_{\theta_0} \|_r} \pi'_{\theta_0}.$$

Therefore, by the same corollary,

$$\phi_0(x) = \frac{| h'(\theta_0) |}{\| \pi'_{\theta_0} \|_r} \cdot \left| \frac{\operatorname{sgn} h'(\theta_0)}{\| \pi'_{\theta_0} \|_r} \pi'_{\theta_0} \right|^{r/s}$$

$$\cdot \operatorname{sgn} h'(\theta_0) \cdot \operatorname{sgn} \pi_{\theta_0}(x)$$
$$= \frac{h'(\theta_0)}{\| \pi'_{\theta_0} \|_r^r} \, | \pi'_{\theta_0}(x) |^{r/s} \operatorname{sgn} \pi'_{\theta_0}(x).$$

And this is the result asserted in the theorem.

The reader will have no difficulty in establishing, in the exact pattern of the preceding proof, the following.

THEOREM 9. *Assume the hypothesis of Theorem* 6. *And, further, let the equality sign hold in* (9) *for* $b = b_0$, $c = c_0$ . [6] *Then,*

$$\phi_0(x) = \frac{b_0 h'(\theta_0) + c_0 h''(\theta_0)}{\| b_0 \pi'_{\theta_0} + c_0 \pi''_{\theta_0} \|_r^r} \cdot | b_0 \pi'_{\theta_0}(x) + c_0 \pi''_{\theta_0}(x) |^{r/s} \cdot \operatorname{sgn} (b_0 \pi'_{\theta_0}(x) + c_0 \pi''_{\theta_0}(x)).$$

It is evident that results of the type in these theorems may be built up as well with integrals over the parameter space.

A question of considerable practical importance is that of the rapidity of convergence of the $\zeta_n$ to $\phi_0$ . An answer to this question, on the level of generality we are maintaining in this study, consists in relating this convergence to that of the $| c_n |$ to $C_0$ . In the case $s = r = 2$, the answer is immediate and exact:

$$\begin{aligned} \| \zeta_n - \phi_0 \|_2^2 &= \int_\Omega (\zeta_n - \phi_0)^2 \, d\nu \\ &= \int_\Omega \zeta_n^2 \, d\nu - 2 \int_\Omega \phi_0 \zeta_n \, d\nu + \int_\Omega \phi_0^2 \, d\nu \\ &= | c_n |^2 - 2 | c_n |^2 + C_0^2 \\ &= C_0^2 - | c_n |^2. \end{aligned}$$

Thus, if one unbiased estimate is known, it provides, since its norm is $\geqq C_0$, an upper bound for $\| \zeta_n - \phi_0 \|_2$ . The same is true in the general case (any $s$) once we have established an upper bound, depending on $C_0$ and $| c_n |$, for $\| \zeta_n - \phi_0 \|_s$ . But in the general case, a good upper bound does not seem to be so close at hand. There are indications of the direction in which one must proceed, and we hope to draw some significant results out of these before long.

**8. The case** $s = r = 2$. The particular aspects of this case (where bestness of an estimate has reference to its *variance*), which arise out of the coincidence of $\mathfrak{L}_r$ and $\mathfrak{L}_s$ , merit some discussion. We shall denote the inner product, $\int_\Omega \xi \eta \, d\nu$, of two functions $\xi$ and $\eta$ in $\mathfrak{L}_2$ , as usual by $(\xi, \eta)$. Let $\{\mathfrak{P}_0\}$ denote the closed linear manifold in $\mathfrak{L}_2$ spanned by the $\pi_\theta$ .

THEOREM 10. *Let* $\mathfrak{M}_2$ *be non-empty. Then* $\phi_0$ *is the unique element of* $\mathfrak{M}_2$ *which lies in* $\{\mathfrak{P}_0\}$.

[6] In the case $s = 2$, $b_0$ and $c_0$ are the values which render (11) an equality.

To begin with it is clear that the functions $\zeta_n$ of Theorem 7, in the present case $s = r = 2$, are all elements of $[\mathfrak{P}_0]$, the linear manifold spanned by the $\pi_\theta$. Hence, since $\phi_0$ is the strong limit of these elements, $\phi_0 \in \{\mathfrak{P}_0\}$.

Now suppose also $\phi_1 \in \mathfrak{M}_2$, $\phi_1 \in \{\mathfrak{P}_0\}$. Then, from

$$(\phi_0, \pi_\theta) = h(\theta), \qquad \theta \in \Theta,$$

$$(\phi_1, \pi_\theta) = h(\theta), \qquad \theta \in \Theta,$$

we have

$$(\phi_1 - \phi_0, \pi_\theta) = 0, \qquad \theta \in \Theta,$$

and, by continuity of the inner product,

$$(\phi_1 - \phi_0, \xi) = 0, \qquad \xi \in \{\mathfrak{P}_0\};$$

that is, $\phi_1 - \phi_0 \in \{\mathfrak{P}_0\}^\perp$. But, from $\phi_0 \in \{\mathfrak{P}_0\}$ and $\phi_1 \in \{\mathfrak{P}_0\}$ it follows that $\phi_1 - \phi_0 \in \{\mathfrak{P}_0\}$. Hence $\phi_1 - \phi_0 = 0$, and this proves the exclusiveness of the property for $\phi_0$.

Another characterization of $\phi_0$ is given by the following corollary.

COROLLARY 10-1. *If $\mathfrak{M}_2$ is non-empty, then $\phi_0$ is the unique element of $\mathfrak{M}_2$ which satisfies the system of equations in $\xi$ : $(\phi, \xi) = \|\xi\|_2^2$, $\phi \in \mathfrak{M}_2$.*

To see that $\phi_0$ has the asserted property, let $\phi$ be any element of $\mathfrak{M}_2$, and set $\phi = \xi + \eta$, with $\xi \in \{\mathfrak{P}_0\}$ and $\eta \in \{\mathfrak{P}_0\}^\perp$. From

$$(\xi, \pi_\theta) = (\xi + \eta, \pi_\theta) = (\phi, \pi_\theta) = h(\theta),$$

it follows that $\xi \in \mathfrak{M}_2$. Hence $\xi = \phi_0$. And so,

$$(\phi, \phi_0) = (\phi_0 + \eta, \phi_0) = \|\phi_0\|_2^2.$$

If $\phi_1 \in \mathfrak{M}_2$ has this property also, then both

$$(\phi_1, \phi_0) = \|\phi_0\|_2^2$$

and

$$(\phi_0, \phi_1) = \|\phi_1\|_2^2;$$

and therefore

$$\|\phi_1\|_2 = \|\phi_0\|_2.$$

This proves $\phi_1 = \phi_0$, and so the corollary.

**9. An example.** Let $\Omega$ be Euclidean $n$-space, $x = (x_1, x_2, \cdots, x_n)$; $\mu$, Lebesgue measure; $\Theta$, the set of real numbers; and

$$p_\theta(x) = \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \theta)^2\right\}.$$

And finally, let $\theta_0 = 0$. Then

$$\pi_\theta(x) = \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(-2\theta x_i + \theta^2)\right\}.$$

If $0 < b < \frac{1}{2}$, and we define

$$\phi_1(x) = (1 - 2b)^{n/2} \exp\left\{b \sum_{i=1}^{n} x_i^2\right\} - 1,$$

we have, for each $\theta$,

$$\int_\Omega \phi_1(x) p_\theta(x)\, d\mu = \exp\left\{\frac{nb}{1 - 2b}\theta^2\right\} - 1.$$

Thus, $\phi_1$ is an unbiased estimate of the function $h$:

$$h(\theta) = \exp\left\{\frac{nb}{1 - 2b}\theta^2\right\} - 1.$$

If we examine

$$\| \phi_1 \|_s^s = \frac{1}{(2\pi)^{n/2}} \int_\Omega \left| (1 - 2b)^{n/2} \exp\left\{b \sum_{i=1}^{n} x_i^2\right\} - 1 \right|^s \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right\} d\mu;$$

we find that this integral converges only for $s < 1/2b$. Shifting the emphasis, we may state: *for the function* $h$, *defined by*

$$h(\theta) = e^{\alpha\theta^2} - 1, \qquad \alpha > 0,$$

*there exists an unbiased estimate with finite* $s$th *moment at* $\theta = 0$, *for each*

$$s < \frac{n + 2\alpha}{2\alpha}.$$

Next, observe that

$$\| \pi_\theta \|_r^r = \frac{1}{(2\pi)^{n/2}} \int_\Omega \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} (x_i^2 - 2r\theta x_i + r\theta^2)\right\} d\mu$$
$$= \exp\{\tfrac{1}{2} nr(r - 1)\theta^2\},$$

so that the $\pi_\theta$ are elements of $\mathfrak{L}_r$ for each $r > 1$. The ratio

$$\frac{| h(\theta) |}{\| \pi_\theta \|_r} = (e^{\alpha\theta^2} - 1) \exp\{-\tfrac{1}{2}n(r - 1)\theta^2\}$$

is seen to diverge as $\theta \to \infty$, if

$$\tfrac{1}{2} n(r - 1) < \alpha.$$

Hence, by Theorem 2, there exists no unbiased estimate of $h$ belonging to $\mathfrak{L}_s$ for a value of $s$ such that the number

$$r = \frac{s}{s - 1}$$

satisfies the inequality just above; that is, for a value of $s$ greater than

$$\frac{n + 2\alpha}{2\alpha}.$$

Otherwise stated: *there exists no unbiased estimate of h with finite* sth *moment at* $\theta = 0$, *for*

$$s > \frac{n + 2\alpha}{2\alpha}.$$

It is most likely true that this last statement holds, in general, with

$$s \geqq \frac{n + 2\alpha}{2\alpha}.$$

We shall consider here only the case

$$\frac{n + 2\alpha}{2\alpha} = 2;$$

and since the analysis is the same for every pair $n$, $\alpha$ satisfying this equality, we treat the particular case of

$$n = 1, \qquad \alpha = \tfrac{1}{2}.$$

Thus, we shall show: *for* $n = 1$, *there exists no unbiased estimate of* $h_2$,

$$h_2(\theta) = e^{\frac{1}{2}\theta^2} - 1,$$

*with finite variance at* $\theta = 0$.

We must show that the ratios

$$\frac{\left|\sum_{i=1}^{m} a_i(e^{\frac{1}{2}\theta_i^2} - 1)\right|}{\left\|\sum_{i=1}^{m} a_i \pi_{\theta_i}\right\|_2}$$

are not bounded for all choices of $m$ (distinct) $\theta_i$'s, and all sets of $m$ real numbers $a_i$, and all $m$. This is clearly equivalent to showing the same for the ratios

$$Q(m, a_i, \theta_i) = \frac{\left|\sum_{i=1}^{m} a_i(1 - e^{-\frac{1}{2}\theta_i^2})\right|}{\left\|\sum_{i=1}^{m} a_i e^{-\frac{1}{2}\theta_i^2} \pi_{\theta_i}\right\|_2}.$$

Now we find, by direct computation,

$$\left\|\sum_{i=1}^{m} a_i e^{-\frac{1}{2}\theta_i^2} \pi_{\theta_i}\right\|_2^2 = \sum_{i,j=1}^{m} e^{-\frac{1}{4}(\theta_i - \theta_j)^2} a_i a_j.$$

And the solution of the familiar extremum problem:

$$\sup_{(a_i)} \left|\sum_{i=1}^{m} a_i(1 - e^{-\frac{1}{2}\theta_i^2})\right| \quad \text{subject to} \quad \sum_{i,j=1}^{m} e^{-\frac{1}{4}(\theta_i - \theta_j)^2} a_i a_j = 1$$

yields

$$\sup_{(a_i)} Q^2(m, a_i, \theta_i) = \sum_{i,j=1}^{m} v_{ij}(1 - e^{-\frac{1}{2}\theta_i^2})(1 - e^{-\frac{1}{2}\theta_j^2}),$$

where the matrix

$$V = (v_{ij}), \qquad i, j = 1, 2, \cdots, m,$$

is the inverse of the matrix

$$U = (e^{-\frac{1}{2}(\theta_i - \theta_j)^2}), \qquad i, j = 1, 2, \cdots, m.$$

We now take, in particular,

$$\theta_i = it, \qquad i = 1, 2, \cdots, m,$$

where $t$ is a positive number. Clearly, there exists a number $t_0$ such that for $t > t_0$,

$$U(t) = (e^{-\frac{1}{2}(i-j)^2 t^2})$$

is non-singular. Also,

$$\lim_{t \to \infty} U(t) = I,$$

the identity matrix. Then, for $t > t_0$, $V = U^{-1}$ is a continuous function of $U$, so that

$$\lim_{t \to \infty} V(t) = (\lim_{t \to \infty} U(t))^{-1} = I.$$

Hence,

$$\lim_{t \to \infty} v_{ij}(t) = \delta_{ij}.$$

It follows that

$$\lim_{t \to \infty} \sup_{(a_i)} Q^2(m, a_i, it) = m,$$

and therefore,

$$\sup_{(a_i, \theta_i)} Q^2(m, a_i, \theta_i) \geqq m.$$

(A simple argument on the characteristic values of $U$ shows that there is actually equality here.) This result gives the unboundedness of the ratios $Q_j$ and our proposition is proved, by virtue of Theorem 2.

## APPENDIX

The spaces $\mathfrak{L}_r$ and $\mathfrak{L}_s$ are instances of a Banach space over the reals; that is, a complete, normed, linear vector space, closed under multiplication by real numbers. That the space, say $\mathfrak{B}$, is normed is to say that there is a non-negative, real-valued function, $\| \ \|$, defined on $\mathfrak{B}$, with the properties:

$$\| \xi \| = 0 \quad \text{if and only if } \xi \text{ is the null vector,}$$

$$\| a\xi \| = | a | \cdot \| \xi \|,$$

$$\| \xi + \eta \| \leqq \| \xi \| + \| \eta \|;$$

where $\xi, \eta \in \mathfrak{B}$ and $a$ is real. The number $\| \xi \|$ is called the *norm* of $\xi$.

The function $\|\xi - \eta\|$ on pairs $\xi$, $\eta$ of vectors is a distance function in the usual sense. With it, *strong convergence* (or simply *convergence*) is defined in $\mathfrak{B}$: $\xi_n$ *converges strongly* to $\xi$ when $\lim_{n\to\infty} \|\xi_n - \xi\| = 0$. In symbols: $\xi_n \to \xi$ or $\lim \xi_n = \xi$. The usual set-theoretic notions are now defined in the obvious way; e.g., limit point of a set, closed set, etc. That the space $\mathfrak{B}$ is complete means that every sequence $\{\xi_n\}$ satisfying $\lim_{m,n\to\infty} \|\xi_m - \xi_n\| = 0$ converges to a (unique) element $\xi \in \mathfrak{B}$.

A *linear manifold* $\mathfrak{M}$ in $\mathfrak{B}$ is a subset of $\mathfrak{B}$ with the property that for any two elements $\xi$, $\eta \in \mathfrak{M}$ and any two real numbers $a$, $b$, we have also $a\xi + b\eta \in \mathfrak{M}$. A *closed linear manifold* is a linear manifold that is closed in the set-theoretic sense. If $S$ is any subset of $\mathfrak{B}$, then the set, $[S]$, of all finite linear combinations of elements of $S$ is a linear manifold; it is the *linear manifold spanned by* $S$. The closure of $[S]$, denoted by $\{S\}$, is called the *closed linear manifold spanned by* $S$. In general, $[S]$ is a proper subset of $\{S\}$. A set $S \subseteq \mathfrak{B}$ is called *fundamental* when $\{S\} = \mathfrak{B}$.

A *linear functional*, $G$, on $\mathfrak{B}$ is a real-valued function with the property that for any two elements $\xi$, $\eta \in \mathfrak{B}$ and any two real numbers $a$, $b$, we have $G(a\xi + b\eta) = aG(\xi) + bG(\eta)$. The linear functional $G$ is said to be *bounded* when the number

$$\|G\| = \underset{\|\xi\|\neq 0}{\text{l.u.b.}} \frac{|G(\xi)|}{\|\xi\|}$$

is finite. $\|G\|$ is called the *norm* of $G$. (Throughout the text of the paper, the qualification "bounded" has been understood in all references to linear functionals). If we define the sum of two linear functionals $F$ and $G$ by $(F + G)(\xi) = F(\xi) + G(\xi)$, and make the other requisite definitions in the obvious way, we find that the bounded linear functionals on $\mathfrak{B}$ form a linear vector space over the reals. The function $\|\ \|$ on the bounded linear functionals, which we have already called a norm, is in fact a norm in the Banach space sense. This vector space, so normed, is readily shown to be complete. Hence it is a Banach space—usually called the conjugate space to $\mathfrak{B}$. It is this space we have referred to in the text as the *space of linear functionals on* $\mathfrak{B}$.

If a sequence $\{\xi_n\}$ of elements of $\mathfrak{B}$ has the property that $\lim_{n\to\infty} G(\xi_n) = G(\xi)$ for every bounded linear functional $G$, then $\xi_n$ is said to *converge weakly to* $\xi$. If, of the sequence $\{\xi_n\}$, we know only that $\lim_{n\to\infty} G(\xi_n)$ exists for every bounded linear functional, we say simply that the sequence is weakly convergent. The space $\mathfrak{B}$ is called weakly complete if every weakly convergent sequence converges weakly to a limit. The spaces $\mathfrak{L}_r$, $r \geqq 1$ are weakly complete. $\mathfrak{B}$ is said to be *weakly compact* if every bounded set $S \subset \mathfrak{B}$ contains a weakly convergent sequence. That $S$ is "bounded" means $\underset{\xi\in S}{\text{l.u.b.}} \|\xi\| < \infty$.

A real Hilbert space $\mathfrak{H}$ is a real Banach space on which there is defined an inner product; that is, a function $(\xi, \eta)$ on pairs of elements $\xi$, $\eta$, with the properties

$$(\xi, \eta) = (\eta, \xi),$$

$$(a\xi, \eta) = a(\xi, \eta),$$

$$(\xi + \zeta, \eta) = (\xi, \eta) + (\zeta, \eta),$$

$$\| \xi \|^2 = (\xi, \xi).$$

The inner product is a *continuous* function of both its arguments; i.e., $\lim \xi_m = \xi$ and $\lim \eta_n = \eta$ imply $\lim (\xi_m, \eta_n) = (\xi, \eta)$. The space $\mathfrak{L}_2$ in the text is a Hilbert space when we take $(\xi, \eta) = \int_\Omega \xi\eta \, d\nu$. Two elements $\xi$, $\eta$ which are such that $(\xi, \eta) = 0$ are said to be orthogonal. If $S$ is any set in $\mathfrak{H}$, then the set of elements of $\mathfrak{H}$ each of which is orthogonal to every element of $S$ is called the *orthocomplement* of $S$, and is denoted by $S^\perp$.

For further elaboration the reader is referred to [13] and [19].

## REFERENCES


[1] D. Blackwell and M. A. Girshick, "A lower bound for the variance of some unbiased sequential estimates," *Annals of Math. Stat.*, Vol. 18 (1947), pp. 277–280.

[2] R. A. Fisher, "Theory of statistical estimation," *Camb. Phil. Soc. Proc.*, Vol. 22, (1925), pp. 700–725.

[3] H. Cramér, *Mathematical Methods of Statistics*, Princeton Press, Princeton, 1946.

[4] A. Bhattacharyya, "On some analogues of the amount of information and their use in statistical estimation," *Sankhyā*, Vol. 8 (1946), pp. 1–14.

[5] M. A. Girshick, F. Mosteller, and L. J. Savage, "Unbiased estimates for certain binomial sampling problems," *Annals of Math. Stat.*, Vol. 17 (1946), pp. 13–23.

[6] J. Neyman, "Su un teorema concernente le cosidette statistiche sufficienti," *Giornale dell'Istituto Italiano degli Attuari*, Vol. 6 (1935), pp. 320–334.

[7] D. Blackwell, "Conditional expectation and unbiased sequential estimation," *Annals of Math. Stat.*, Vol. 18 (1947), pp. 105–110.

[8] H. Cramér, "A contribution to the theory of statistical estimation," *Skand. Aktuar. tids.*, Vol. 29 (1946), pp. 85–94.

[9] A. Bhattacharyya, "On some analogues of the amount of information and their use in statistical estimation (cont'd)," *Sankhyā*, Vol. 8 (1947), pp. 201–218.

[10] J. Wolfowitz, "The efficiency of sequential estimates and Wald's equation for sequential processes," *Annals of Math. Stat.*, Vol. 18 (1947), pp. 215–230.

[11] F. Riesz, *Les Systèmes d'Équations Linéaires a une Infinité d'Inconnues*, Gauthier-Villars, Paris, 1913.

[12] F. Riesz, "Untersuchungen über Systeme integrierbare Funktionen," *Math. Annalen*, Vol. 69 (1910), pp. 449–497.

[13] S. Banach, *Théorie des Opérations Linéaires*, Garasinski, Warsaw, 1932.

[14] N. Dunford, "Uniformity in linear spaces," *Am. Math. Soc. Trans.*, Vol. 44 (1938), pp. 305–356.

[15] M. H. Steinhaus, "Additive und stetige funktionaloperationen," *Math. Zeits.*, Vol. 5 (1918), pp. 186–221.

[16] F. Riesz, "Sur une espèce de géométrie analytique des systèmes de fonctions sommables," *Comptes Rendus*, Vol. 144 (1907), pp. 1409–1411.

[17] M. Fréchet, "Sur les ensembles de fonctions et les opérations linéaires," *Comptes Rendus*, Vol. 144 (1907), pp. 1414–1416.

[18] S. Saks, *Theory of the Integral*, Stechert, New York, 1937.

[19] J. von Neumann, *Functional Operators*, (Mimeographed notes) Princeton, 1935.

[20] G. R. Seth, "On the variance of estimates," *Annals of Math. Stat.*, Vol. 20 (1949), pp. 1–27.

[21] B. J. Pettis, "A note on regular Banach spaces," *Am. Math. Soc. Bull.*, Vol. 44 (1938), pp. 420–428.