This dissertation has been
microfilmed exactly as received 69-4206

ABBEY, Joseph Leo, 1940-
SERIES REPRESENTATION OF MULTIVARIATE AND SEQUENTIAL UNIFORMLY MINIMUM VARIANCE UNBIASED ESTIMATES FOR EXPONENTIAL FAMILIES.

Iowa State University, Ph.D., 1968
Statistics

University Microfilms, Inc., Ann Arbor, Michigan

SERIES REPRESENTATION OF MULTIVARIATE AND SEQUENTIAL UNIFORMLY MINIMUM VARIANCE UNBIASED ESTIMATES FOR EXPONENTIAL FAMILIES

by

Joseph Leo Abbey

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate Faculty

Iowa State University
Of Science and Technology
Ames, Iowa

1968

TABLE OF CONTENTS

I. INTRODUCTION AND REVIEW

This dissertation deals with the problem of statistical estimation involving the Koopman-Darmois class of exponential densities. Let x be a sample from a sample space $(X, \mathcal{A})$, where X is a Borel subset of n dimensional Euclidean space, $1 \leq n \leq N$ (N finite) and $\mathcal{A}$ is the Borel field of subsets of X. Let $\mathcal{P} = \{p_\theta(x) \mid \theta \in \Omega\}$ be a set of probability measures on $\mathcal{A}$ such that $p_\theta(x) \in \mathcal{P}$ admits an exponential density function $f(x, \theta)$ with respect to a fixed σ-finite measure $\mu(x)$ which may be Lebesgue or counting measure. The density function, $f(x, \theta)$, will be assumed to be a member of the Koopman-Darmois class of exponential densities and Ω will be taken to be an open set in p-dimensional Euclidean space called the natural parameter space. Let $g(\theta)$ be an estimable (has at least one square integrable $\mathcal{P}$ unbiased estimator) parametric function. Our problem is to consider the existence of a uniformly minimum variance unbiased estimator U.M.V.U.E. for $g(\theta)$ and in the case of existence, to obtain a series representation for it. (This formulation enables us to deal with both discrete and continuous random variables.)

Abbey (1) considered the special case where $n = N$ is fixed and $p = 1$ (θ is a scalar). In this case it is well known that a complete sufficient statistic exists and that every estimable parametric function $g(\theta)$ possesses a unique U.M.V.U.E. which is obtainable by the method of Rao and Blackwell provided an unbiased estimator of $g(\theta)$ is available. A method equivalent to Rao-Blackwellisation with the advantage of not requiring explicit knowledge of any unbiased estimator was developed and a series representation of the U.M.V.U.E. together with an expression for its variance was given. The key requirement for the approach adopted was the equality of two subspaces U_θ and V_θ of the Hilbert space, H_θ, of $\mathcal{Q}$-measurable, square integrable $p_\theta(x)$ (θ arbitrary but fixed) functions of x. We will now briefly review this approach and show that while it generalizes to the case $p > 1$ considered in Chapter III for which n is fixed, even when the assumptions for its validity are satisfied, it breaks down for the sequential generalization dealt with in Chapter II. Further the assumptions needed are not valid in general, however, in the special case considered in Abbey (1) and its generalization in Chapter III of the present investigation, we will show that this deficiency is readily taken care of.

In the formal exposition of the rest of this investigation the sample space will always be denoted by $(X, \mathcal{Q})$ and the above assumptions about it will be made throughout. The set of measures on $\mathcal{Q}$ will also be denoted by $\mathcal{P}$ and the assumptions made here will again be kept throughout. It is convenient and of some theoretical advantage to discuss subfields of $\mathcal{Q}$ rather than the statistics which induce them. In this regard the subfield $\mathcal{Q}_o$ induced by the sufficient statistic will be referred to as the sufficient subfield and in general subfields will be named by properties of the inducing statistic. The complete (completeness in the sense of Definition 2.1) subfield will be denoted by $\mathcal{Q}_c$. It follows from Bahadur (5) that under the assumptions made about the sample space $(X, \mathcal{Q})$ and the set of measures $\mathcal{P}$, it may be assumed that any subfield corresponds to a statistic and an estimator measurable with respect to a subfield is an estimator which depends on x only through the corresponding statistic.

It is also convenient to take the empty set as the only $\mathcal{A}$-measurable $\mathcal{P}$-null set. This is to eliminate the need for a null set qualification to many of the definitions, arguments and conclusions to follow. Thus the assertion $h_c(x)$ is the unique U.M.V.U.E. for $g(\theta)$ strictly speaking means if $h_o(x)$ is also a U.M.V.U.E. for $g(\theta)$ then $p_\theta(h_c(x) = h_o(x)) = 1$ for each $p_\theta \in \mathcal{P}$. Specifically the following convention is followed throughout this investigation. If A and B are sets in $\mathcal{A}$, $A = B$ means $(A \cap B') \cup (B \cap A')$ is $\mathcal{P}$ null; if $h_1(x)$, $h_2(x)$ are estimators, $h_1(x) = h_2(x)$ means $\{x \in X | h_1(x) \neq h_2(x)\}$ is $\mathcal{P}$-null. The relations of inclusion and equality between classes of sets of X, and between classes of estimators, are to be interpreted in terms of this convention.

Let θ_o be an arbitrary but fixed point in Ω. The approach adopted in Abbey (1) is to obtain a representation of the locally best (minimum variance at θ_o) unbiased estimator of $g(\theta)$ and then conclude that this is indeed the U.M.V.U.E.

Let H_o denote the real Hilbert space defined by

$$H_o = \mathcal{L}_2(X, \mathcal{A}, \lambda) \tag{1.1}$$

where

$$\lambda = p_{\theta_o}(x) \quad . \tag{1.2}$$

For any $h_1, h_2 \in H_o$ let

$$(h_1, h_2) = \int_X h_1 \cdot h_2 \, d\lambda \tag{1.3}$$

and

$$\|h_1\|^2 = (h_1, h_1) \quad . \tag{1.4}$$

Let

$$\ell_\theta(x) = f(x, \theta)/f(x, \theta_o) \qquad \theta \in \Omega \tag{1.5}$$

and

$$L = \{\ell_\theta \mid \theta \in \Omega\} \quad . \tag{1.6}$$

For any positive integer n, let the n^{th} partial derivative of ℓ_θ with respect to θ evaluated at θ_o be denoted by $\psi_n(x)$; then

$$\psi_n(x) = f^{(n)}(x, \theta_o)/f(x, \theta_o) \tag{1.7}$$

and

$$\psi_o(x) = 1 \quad . \tag{1.8}$$

Let

$$S = \{\psi_n(x) \mid n = 0, 1, 2, \ldots\} \quad . \tag{1.9}$$

It is easy to verify that the functions $\ell_\theta(x)$ and $\psi_n(x)$ are well defined under the null set qualifications above. Further, it can be shown that

$$S \subset H_o \quad . \tag{1.10}$$

Let

$$V_o = \overline{\text{span}}\ \{S\} \quad . \tag{1.11}$$

We will assume that

$$L \subset H_o \quad . \tag{1.12}$$

This last assumption is not wholly valid in general in the sense that only a proper subset of L may be in H_o. However this does not invalidate the arguments since it can be shown from the properties of the density function that if $L_o = \{\ell_\theta \in L \mid \ell_\theta \in H_o\}$ then $\Omega_o = \{\theta \in \Omega \mid \ell_\theta \in L_o\}$ contains an open interval and from the analysis, the locally best estimator of $g(\theta)$ which is also $\mathcal{A}_o$-measurable will then be unbiased only for $\theta \in \Omega_o$. However from the completeness of the sufficient statistic, it is easily shown that this locally best estimator is in fact unbiased for $g(\theta)$ over the whole parameter space Ω. In the general case where $\mathcal{A}_o$ is not complete this argument breaks down.

Now let U_o be the subspace of H_o spanned by L. That is

$$U_o = \overline{\text{span}}\ \{L\} \quad . \tag{1.12}$$

It follows from Bahadur (5) that a necessary and sufficient condition for any element in V_o to be the locally best unbiased estimator of its expectation is that

$$U_o \subseteq V_o \quad . \tag{1.13}$$

Abbey (1) showed using certain analytic properties of estimable parametric functions that for the fixed sample, scalar parameter case we have

$$V_o = W_o \tag{1.14}$$

where

$$W_o = \mathcal{L}_2(X, \mathcal{a}_o, \lambda) \quad . \tag{1.15}$$

However each element of L is $\mathcal{a}_o$-measurable. In fact from Theorem 6.2 of Bahadur (6) $\mathcal{a}_o$ is the smallest subfield such that each $\ell_\theta(x)$ is $\mathcal{a}_o$-measurable. Hence under the assumption $L \subset H_o$ we have from Equations 1.14 and 1.15

$$U_o \subseteq V_o \tag{1.16}$$

which verifies Equation 1.13.

In Chapter II we will use certain results in the theory of moments to show that under certain assumptions, which in particular hold for the fixed sample case,

$$V_o = \overline{\text{span}}\,\{S\} = \mathcal{L}_2(X, \mathcal{a}_c, \lambda) \tag{1.17}$$

where $\mathcal{a}_c$ is the complete subfield. It will then follow from Bahadur (5) that under these conditions

$$V_o \subseteq W_o \tag{1.18}$$

with proper inclusion if $\mathcal{G}_c \neq \mathcal{G}_o$. It must be pointed out however that in this latter case we also have $U_o \subseteq W_o$ and while it is therefore possible that Equation 1.13 may still hold, it has not been possible to verify this.

When $\mathcal{G}_c'$ is not also sufficient, however, the orthogonal projection of a given $h(x)$ in H_o onto V_o may depend on θ_o. Further since θ_o is arbitrary this means that we may get a different statistic, by projection, at each θ value. In this case we obtain an unachievable bound for estimators of the expectation of $h(x)$ from the variances of these projections. On the other hand if the projection is independent of θ, then we obtain a U.M.V.U.E. for the expectation of $h(x)$. In this latter case we will say that the expectation of $h(x)$ is $\mathcal{G}_c$-estimable.

We will also study V_o directly and show that for any $h(x) \in H_o$ which is unbiased for a given $g(\theta)$, the orthogonal projection of $h(x)$ onto V_o is locally best unbiased for $g(\theta)$ provided it is square integrable $\mathcal{P}$. Further if $g(\theta)$ possesses a U.M.V.U.E. then the above locally best unbiased estimator for $g(\theta)$ is in fact the U.M.V.U.E. since the densities have the same support. It may be recalled, Bahadur (5), DeGroot (10), that in the sequential case considered in Chapter II there may exist estimable parametric functions which do not possess U.M.V.U.E.'s.

II. THE SEQUENTIAL CASE

A. Introduction

Consider the following fixed sample situation. Let $x = (x_1, x_2, \ldots, x_n)$ be a random sample from a sample space $(E^n, \mathcal{B}^n)$ where E^n is n-dimensional Euclidean space and $\mathcal{B}^n$ is the Borel field of subsets of E^n, and suppose that x is distributed in $(E^n, \mathcal{B}^n)$ according to p_θ, one of a certain set $\mathcal{P} = \{p_\theta | \theta \in \Omega\}$ of probability measures on $\mathcal{B}^n$. We will assume that each $p_\theta \in \mathcal{P}$ admits an exponential density function $f(x; \theta)$ with respect to a fixed σ-finite measure $\mu(x)$ which may be Lebesgue or counting measure and that Ω is an open interval of E^1. Following Lehmann (17) we will write

$$f(x; \theta) = \alpha(x) \beta(\theta) \exp \theta T(x) \qquad \theta \in \Omega \quad . \tag{2.1}$$

In this case n is fixed and $T(x)$ a real valued, Borel-measurable function on X is the complete sufficient statistic.

We now consider the following sequential situation. We are given a sampling rule for taking successive observations which is such that at the m^{th} stage, the decision of whether or not to make the $(m+1)^{th}$ observation depends only on the value of $T_m(x)$. It then follows from Blackwell (7), that if the total number of observations is n (a random variable taking values $1, 2, 3, \ldots, N$ with N a fixed finite positive integer), then (n, T_n) is a sufficient statistic for θ. We will take for our sample space the union $X = \bigcup_{n=1}^{N} X_n$ of Borel subsets X_n of n dimensional Euclidean space E^n for which exactly n observations are taken.

For each n, let $\mathcal{A}^n$ be the Borel field of subsets of X_n and $\mathcal{A}_o^n$ the subfield of $\mathcal{A}^n$ induced on X_n by $T_n(x)$. The sample space for our problem may then be defined as the probability measure space $(X, \mathcal{A}, \mathcal{P})$ where

$$p_\theta(X) = \sum_{n=1}^{N} p_\theta(X_i) = 1 \quad \text{for all } \theta \in \Omega \tag{2.2}$$

and

$$\mathcal{A} = \{ \bigcup_{n=1}^{N} A_n | A_n \in \mathcal{A}^n \} \quad . \tag{2.3}$$

Consider the subfield $\mathcal{A}_o$ of $\mathcal{A}$ induced on X by the sufficient statistic $(n, T_n(x))$. Then $\mathcal{A}_o$ is defined by

$$\mathcal{A}_o = \{ \bigcup_{n=1}^{N} A_n | A_n \in \mathcal{A}_o^n \} \quad . \tag{2.4}$$

Since for each n, the null set $\emptyset \in \mathcal{A}^n$ it follows that $\mathcal{A}_o^n \subset \mathcal{A}_o$ for all n. Indeed $\mathcal{A}_o$ may be characterized as the smallest Borel field containing $\mathcal{A}_o^n$ $n = 1, 2, \ldots, N$. $\mathcal{A}_o$ is sufficient for the measures $\mathcal{P}$ on $\mathcal{A}$. In general, however, it is not complete. We will examine conditions under which $\mathcal{A}_o$ is both sufficient and complete by defining the complete subfield for the measures $\mathcal{P}$.

Consider the real valued, $\mathcal{A}$-measurable function $T(x)$ on X given by

$$T(x) = T_n(x) \quad \text{if} \quad x \in X_n \quad n = 1, 2, \ldots, N \quad . \tag{2.5}$$

Let $\mathcal{A}_c$ denote the subfield of $\mathcal{A}$ induced on X by $T(x)$. Then we have the following result.

Lemma 2.1. The subfield $\mathcal{A}_c$ of $\mathcal{A}$ induced by the statistic $T(x)$ defined in Equation 2.5 is a subfield of the sufficient subfield $\mathcal{A}_o$.

Proof. Follows from the fact that we can define $T(x)$ as a $\mathcal{A}_o$-measurable function of $(n, T_n(x))$ by

$$T(x) = \sum_{n=1}^{N} T_n(x) \qquad x \in X .$$

Alternatively let Y be the range of $T(x)$. Clearly Y is a Borel subset of E^1. Let $\mathcal{B}$ be the Borel field on Y and $B \in \mathcal{B}$ be an arbitrary set in Y. Then

$$\begin{aligned} T^{-1}(B) &= \{x \in X \mid T(x) \in B\} \\ &= \bigcup_{n=1}^{N} \{x \in X \mid T_n(x) \in B\} \end{aligned} .$$

It then follows from the definition of $\mathcal{A}_o$ in Equation 2.3 that $T^{-1}(B) \in \mathcal{A}_o$ and since B is arbitrary

$$\mathcal{A}_c = T^{-1}(\mathcal{B}) \subseteq \mathcal{A}_o .$$

In general we have proper inclusion of $\mathcal{A}_c$ in $\mathcal{A}_o$. The following lemma gives a necessary and sufficient condition for equality of $\mathcal{A}_c$ and $\mathcal{A}_o$.

Lemma 2.2. A necessary and sufficient condition for the equality of $\mathcal{A}_c$ and $\mathcal{A}_o$ is that the ranges of $T_n(x)$ $n = 1, 2, \ldots, N$ be non-overlapping.

<u>Proof.</u> Necessity. Suppose $\mathcal{A}_o = \mathcal{A}_c$. Then for each n, $T_n(x)$ is $\mathcal{A}_c$ measurable.

Let

$$\emptyset \neq B \subseteq T_{n_1}(X_{n_1}) \cap T_{n_2}(X_{n_2}) \quad n_1, n_2 \text{ fixed}$$

and

$$A_{n_i} = T_{n_i}^{-1}(B) \qquad i = 1, 2 \quad .$$

However

$$T^{-1}(B) = A_{n_1} \cup A_{n_2}$$

and hence $A_{n_1} \notin \mathcal{A}_c$, unless $A_{n_2} = \emptyset$ in which case $B = T_{n_2}(\emptyset)$ and hence $B = \emptyset$. Therefore $A_{n_2} \neq \emptyset$ and $T_{n_1}(x)$ is not $\mathcal{A}_c$ measurable which contradicts $\mathcal{A}_o = \mathcal{A}_c$ and hence $B = \emptyset$. The result now follows from the fact that B, n_1, n_2 are arbitrary.

Sufficiency. Suppose for any n_1, n_2,

$$T_{n_1}(X_{n_1}) \cap T(X_{n_2}) = \emptyset \quad .$$

Then each $T_n(x)$ is $\mathcal{A}_c$-measurable and hence each $\mathcal{A}_o^n$ is contained in $\mathcal{A}_c$. However, since $\mathcal{A}_o$ is the smallest Borel field containing $\mathcal{A}_o^n$

$n = 1, 2, \ldots, N$, we have

$$\mathcal{G}_o \subseteq \mathcal{G}_c$$

and hence using Lemma 2.1

$$\mathcal{G}_o = \mathcal{G}_c \quad .$$

We will use the following definition of completeness due to Bahadur (5).

Definition 2.1. (Bahadur) We will call a given σ-algebra, S, on X complete if there are no non-trivial unbiased estimators of zero which are S-measurable and have finite variance for all $\theta \in \Omega$.

Let $u(m)$ $m = 1, 2, \ldots$ denote the m^{th} moment of $T(x)$ at θ, an arbitrary but fixed point in Ω. That is

$$u(m) = \int_X T(x)^m \, dp_\theta(x) \tag{2.6}$$

$$= \sum_{n=1}^{N} \int_{X_n} T_n(x)^m \, dp_\theta(x) \quad . \tag{2.7}$$

It follows from a theorem of Lehmann (17) that $u(m)$ exists and

$$u(m) = \sum_{n=1}^{N} \gamma_n^{(m)}(\theta) / \gamma_n(\theta) \qquad m = 1, 2, \ldots \tag{2.8}$$

where

$$\gamma_n(\theta)^{-1} = \int_{X_n} \alpha_n(x) \exp \theta \, T_n(x) \, d\mu(x) \quad .$$

Lemma 2.3. The set of moments $\{u(m)\}$ determines the distribution function of $T(x)$ at each θ in Ω.

Proof. From Feller (11) it is sufficient to show that the series

$$\sum_{k=0}^{\infty} \frac{(-1)^k u(k)}{k!} (\theta - \theta_o)^k$$

converges for all $|\theta - \theta_o| < \delta_o$, for some $\delta_o > 0$.

Consider the series

$$\sum_{k=0}^{\infty} \frac{u(k)}{k!} (\theta - \theta_o)^k \quad .$$

Then for all $|\theta - \theta_o| < \delta_1$ the series converges absolutely where

$$\delta_1^{-1} = \limsup_{n \to \infty} \left|\frac{u_k}{k!}\right|^{\frac{1}{k}} \quad .$$

Now from Equation 2.8

$$u_k = \sum_{n=1}^{N} \gamma_n^{(k)}(\theta_o)/\gamma_n(\theta_o) \quad .$$

Further

$$|u_k|^{\frac{1}{k}} \leq \sum_{n=1}^{N} |\gamma_n^{(k)}(\theta_o)/\gamma_n(\theta_o)|^{\frac{1}{k}} \quad .$$

And

$$\limsup_{k \to \infty} \sum_{n=1}^{N} |\gamma_n^{(k)}(\theta_o)/k!\, \gamma_n(\theta_o)|^{\frac{1}{k}} \leq \sum_{n=1}^{N} \limsup_{k \to \infty} |\gamma_n^{(k)}(\theta_o)/k!\, \gamma_n(\theta_o)|^{\frac{1}{k}} \quad .$$

However from the analytic properties of $\gamma_n(\theta)$ $n = 1, 2, \ldots, N$

$$\limsup_{k \to \infty} |\gamma_n^k(\theta_o)/k!\, \gamma_n(\theta_o)|^{\frac{1}{k}} < \infty \quad n = 1, 2, \ldots, N \quad .$$

Hence

$$\delta_1^{-1} = \limsup_{n \to \infty} | u_k/k! |^{\frac{1}{k}} < \infty$$

which implies $\delta_1 > 0$. The result follows by choosing $\delta_o \leq \delta_1$.

Using Stirling's approximation

$$\delta_1^{-1} = \limsup | u_k |^{\frac{1}{k}}/k \tag{2.10}$$

which is the sufficient condition given by Kendall (15), (16). The next lemma is given in Akhiezer (3) and we state it here without proof.

Lemma 2.4. (Akhiezer) If the moments of T determine the distribution function of T, then the polynomials in T are dense in the Hilbert space of square integrable functions of T.

It therefore follows from Halmos (14) that these polynomials regarded as $\mathcal{a}_c$-measurable functions of x are dense in the Hilbert space V_θ given by

$$V_\theta = \mathcal{L}_2(X, \mathcal{a}_c, p_\theta(x)) \qquad \theta \in \Omega \quad . \tag{2.11}$$

Consider the density function $f(x, \theta)$, with respect to μ, of $p_\theta(x)$ on $\mathcal{a}$. Under the null set qualifications stated in Chapter 1 we have $f(x, \theta) > 0$ $x \in X$. Let

$$\psi_k(x) = f^{(k)}(x, \theta)/f(x, \theta) \qquad k = 1, 2, \ldots \tag{2.12}$$

$$\psi_o(x) = 1 \qquad k = 0, 1, 2, 3, \ldots \tag{2.13}$$

Assumption 2.1. The set of functions $\{\psi_k\}$ is dense in the Hilbert space V_θ $(\theta \in \Omega)$ defined by Equation 2.11.

It is easy to verify that in the fixed sample case to be considered in Section C of this chapter, ψ_k is a polynomial in $T(x)$ of exact degree k, and hence the elements of $\{\psi_k | k = 0, 1, 2, \ldots\}$ are linearly independent. Further any arbitrary polynomial in $T(x)$ of degree $\leq k$ is uniquely expressible as a linear combination of ψ_o, ψ_1, $\psi_2, \ldots, \psi_k$. It therefore follows from Cheney (9) that the set $\{\psi_k\}$ generates the set of all polynomials in $T(x)$ and is therefore dense in V_θ using Lemma 2.5.

In the sequential case (truncated or non-truncated) Lemma 2.5 is not directly applicable since in this case it may be verified for example by the use of Equation 2.16 that the elements of $\{\psi_k\}$ are homogeneous polynomials in $T(x)$ and $\nu(x)$ (where $\nu(x) = n$ if $x \in X_n$). Hence apart from some special cases these homogeneous polynomials will at best be dense in the Hilbert space of square integrable functions of both $T(x)$ and $\nu(x)$ and not just V_θ. It appears from this that in general the space spanned by $\{\psi_k\}$ will contain V_θ as a proper subset. This conjecture has however not been verified.

The negative binomial application is interesting in this regard since it calls for cessation in taking observations when $T_n(x) = c$ (a constant). It is easily seen again by use of Equation 2.16 that in this case ψ_k is a polynomial in $\nu(x)$. Hence the conditions for completeness to be given here does not apply to this application since it is based on the assumption that $T_n(x)$ assumes a range of values for each n.

If the above conjecture is correct, then Assumption 2.1 is not needed to prove the completeness of the subfield induced by $T(x)$. For in that case any element in V_θ is also an element of the $\overline{\text{span}}\ \{\psi_k\}$ and hence if it is unbiased for zero then it is orthogonal to each element of $\{\psi_k\}$ and hence belongs to the orthogonal complement in H_θ of $\overline{\text{span}}\ \{\psi_k\}$. This implies that the function in question must be the zero function.

Definition 2.2. A parametric function $g(\theta)$ will be said to be estimable if there exists at least one $\mathcal{C}$-measurable, square integrable $(\mathcal{P})$ real valued function $h(x)$ such that

$$\int_X h(x)\, dp_\theta(x) = g(\theta) \quad \text{for all} \quad \theta \in \Omega \ . \tag{2.14}$$

We will call such an $h(x)$ an unbiased estimator of $g(\theta)$.

Lemma 2.5. Let $h(x)$ be an unbiased estimator of $g(\theta)$. Then

(i) the parametric function

$$g(\theta) = \int_X h(x)\, dp_\theta(x)$$

$$= \sum_{n=1}^{N} \int_{X_n} h(x)\, \alpha_n(x)\, \beta_n(\theta) \exp \theta\, T_n(x)\, d\mu(x)$$

considered as a function of the complex variable $\theta = \xi + i\nu$ is an analytic function in the region R of parameter points for which ξ is an interior point of Ω.

(ii) the derivatives of all orders with respect to θ of $g(\theta)$ can be computed under the integral sign.

Proof. It is sufficient to prove the theorem for

$$g_n(\theta) = \int_{X_n} h(x)\,\alpha_n(x) \exp \theta\, T_n(x)\, d\mu(x)$$

since for each n $\beta_n(\theta)$ already has the properties stated in the theorem and if the theorem holds for all n then it also holds for the sum.

Now for θ_o fixed but arbitrary interior point of Ω

$$\frac{g_n(\theta) - g_n(\theta_o)}{\theta - \theta_o} = \int h(x)\left[\frac{\exp[(\theta-\theta_o)T_n(x)]-1}{\theta - \theta_o}\right]\alpha_n(x) \exp \theta_o T_n(x) d\mu(x) \quad .$$

Now for $|\theta - \theta_o| < \delta$

$$\left| h(x)\left[\frac{\exp[(\theta-\theta_o)T_n(x)] - 1}{\theta - \theta_o}\right]\right|$$

$$\leq |h(x)| \frac{\exp|\delta|\, T_n(x)}{\delta}$$

$$\leq |h(x)| \left[\frac{\exp \delta\, T_n(x) + \exp - \delta\, T_n(x)}{\delta}\right] \quad .$$

Since the right hand side is integrable $p_{\theta_o}(x)$ it follows from the Lebesgue Dominated Convergence Theorem that for any sequence of points θ_n tending to θ_o, the difference quotient of $g_n(\theta)$ tends to

$$\int h(x)\, T_n(x)\, \alpha_n(x) \exp \theta_o\, T_n(x)\, d\mu(x)$$

which completes the proof since θ_o is arbitrary. It can be verified by induction that the higher derivatives of $g(\theta)$ can be obtained by differentiation under the integral sign.

Theorem 2.1. The subfield $\mathfrak{A}_c$, induced on X by the statistic $T(x) = T_n(x)$ if $x \in X_n$, is complete in the sense of Definition 2.1.

Proof. We will show that if $h(x)$ is an unbiased estimator (in the sense of Definition 2.2) then $h(x) = 0$ provided $h(x)$ is $\mathfrak{A}_c$-measurable.

If $h(x)$ is $\mathfrak{A}_c$-measurable, square integrable $(\mathcal{P})$ unbiased estimator of zero, then

$$h(x) \in \mathcal{L}_2(X, \mathfrak{A}_c, p_\theta(x)) = V_\theta \quad \text{for all} \quad \theta \in \Omega \quad .$$

However, from Lemma 2.5, $h(x)$ belongs to the orthogonal complement of V_θ in $H_\theta = \mathcal{L}_2(X, \mathfrak{A}, p_\theta(x))$ since the inner product

$$(h(x), \psi_k(x)) = 0 \quad k = 0, 1, 2, \ldots$$

and $\{\psi_k\}$ is dense in V_θ. Hence $h(x) = 0$ which completes the proof.

We now establish necessary and sufficient conditions for the existence of a complete sufficient subfield. It will be clear that in this case the sufficient statistic is also complete.

Theorem 2.2. A necessary and sufficient condition for the existence of a complete sufficient statistic is that the range spaces of $T_n(x)$ should be non-overlapping. In this case the statistic $(n, T_n(x))$ is both complete and sufficient.

Proof. Follows directly from Lemma 2.2 and Theorem 2.1.

It is evident that the requirement in the above theorem for the existence of a complete sufficient statistic will, in general, not be obtainable especially in cases where the random variable x takes both positive and negative values. In the case of the normal family, this result has also been given by Lehmann and Stein (18). However, it follows from Bahadur (5) that whenever there is no complete sufficient statistic then there exist estimable parametric functions that do not admit unbiased estimators of uniformly minimum variance and further than the class of uniformly minimum variance unbiased estimators cannot be characterized as the class of all square integrable $(\mathcal{P})$ estimators which are measurable with respect to some subfield unless X is finite or $\mathcal{P}$ is a finite set of measures. [It may be recalled that if a complete sufficient statistic exists then the class of uniformly minimum variance unbiased estimators is the class of all square integrable $(\mathcal{P})$ functions which are measurable with respect to the complete sufficient subfield.]

In the next section we will consider the orthogonal projection of an arbitrary estimator, of a given estimable $g(\theta)$, unto the space spanned by $\{\psi_n\}$ both under Assumption 2.1 and without it. It must be pointed out that the series representation of any element of $\overline{\text{span}}\,\{\psi_k\}$ does not depend in any way on the validity of Assumption 2.1. We will also discuss the achievability of certain lower bounds obtained by Blackwell and Girshick (8) and Wolfowitz (23).

B. Existence and Construction of the U.M.V.U.E.

We noted in the previous section that in general the class of U.M.V.U.E.'s cannot be characterized as the class of $\mathfrak{A}$-measurable, square integrable $(\mathcal{P})$ functions (with $\mathfrak{A}$, some σ-algebra). We will therefore restrict attention to the subclass of U.M.V.U.E.'s which are $\mathfrak{A}_c$-measurable.

Definition 2.3. A parametric function, $g(\theta)$, will be said to be $\mathfrak{A}_c$-estimable if there exists at least one $\mathfrak{A}_c$-measurable, square integrable $(\mathcal{P})$ real valued function $h(x)$ such that

$$\int h(x)\, dp_\theta(x) = g(\theta) \quad \text{for all} \quad \theta \in \Omega \ . \tag{2.14}$$

It may be pointed out that the class of $\mathfrak{A}_c$-measurable estimators is not empty since it includes all bounded estimators. In fact from Bahadur (5) the class of all bounded U.M.V.U.E.'s is the class of all bounded $\mathfrak{A}_c$-measurable functions and in the case where X is finite, (for example the binomial case) the class of U.M.V.U.E.'s is the class of $\mathfrak{A}_c$-measurable estimators.

Let $g(\theta)$ be an estimable parametric function and let $h(x)$ be an unbiased estimator of $g(\theta)$. Since $\mathfrak{A}_c \subseteq \mathfrak{A}_o$ (Lemma 2.1) it follows from Bahadur (6) that

$$E_\theta(h(x)\,|\,\mathfrak{A}_c) = E_\theta(E_\theta(h(x)\,|\,\mathfrak{A}_o)\,|\,\mathfrak{A}_c) \tag{2.15}$$

and since the conditional expectation operator given $\mathfrak{A}_c$ is also the orthogonal projection operator onto $\mathcal{L}_2(X, \mathfrak{A}_c, p_\theta(x))$ we may restrict ourselves to the subspaces $W_\theta = \mathcal{L}_2(X, \mathcal{C}_o, p_\theta(x)$ and $V_\theta = \mathcal{L}_2(X, \mathcal{C}_c, p_\theta(x))$ for each $\theta \in \Omega$.

Since $\mathcal{C}_o$ is sufficient $E_\theta(h(x)|\mathcal{C}_o)$ is independent of θ. In general however $E_\theta(h(x)|\mathcal{C}_c)$ will be a function of θ and will not therefore be an estimator unless the parametric function $g(\theta)$, estimated by $h(x)$, is $\mathcal{C}_c$-estimable. In the former case we will obtain an unachievable variance bound from the variance of $E_\theta(h(x)|\mathcal{C}_c)$ but in the latter case we obtain the U.M.V.U.E. of $g(\theta)$. It may be noted from Equation 2.15 that the $\mathcal{C}_c$-measurable estimator of $g(\theta)$ may not be improved upon by Rao-Blackwellization since it is already $\mathcal{C}_o$-measurable. Further from the completeness of $\mathcal{C}_c$, the estimator is unique and hence it is the U.M.V.U.E. for $g(\theta)$. We state this as a lemma.

Lemma 2.6. If a parametric-function $g(\theta)$ is $\mathcal{C}_c$-estimable then it possesses a unique U.M.V.U.E. which may be characterized as the $\mathcal{C}_c$-measurable unbiased estimator of the parametric function.

The method for constructing the U.M.V.U.E. of an $\mathcal{C}_c$-estimable parametric function is to obtain a complete orthonormal set for V_θ from the polynomials $\psi_k(x)$ given in Equations 2.12 and 2.13. The choice of this particular set of polynomials is based on the fact that expressions for the U.M.V.U.E. may be obtained without knowledge of any unbiased estimator of $g(\theta)$ since the inner products required for the series are obtainable as linear combinations of the derivatives of $g(\theta)$. We will illustrate this for a subclass of the exponential family.

Consider the subclass of the exponential family with density $f(x, \theta)$ satisfying the equation

$$T_n(x) - n\theta = K(\theta)\psi_1(x, n) \quad n = 1, 2, \ldots, N \tag{2.16}$$

where $K(\theta)$ is a function of θ only, and possesses derivatives up to the third order with

$$\frac{d^3K}{d\theta^3} = 0 \quad \text{for all } \theta \in \Omega \ . \tag{2.17}$$

That is

$$K(\theta) = a\theta^2 + b\theta + c \ . \tag{2.18}$$

For this subclass it can be shown that the polynomials $\psi_m(x)$ in $T(x)$ satisfy the recursive relation

$$\psi_{m+1} = (\psi_1 - A_m)\psi_m - B_m\psi_{m-1} \quad m = 1, 2, \ldots \tag{2.19}$$

where

$$A_m = mK'(\theta)/K \tag{2.20}$$

and

$$B_m = m(m-1)K''(\theta)/2K + mn/K \ . \tag{2.21}$$

Using Equation 2.19 and the fact that

$$\frac{\partial\psi_m}{\partial\theta} = \psi_{m+1} - \psi_1\psi_m \tag{2.22}$$

it can be verified that

$$\psi_1^j \cdot \psi_i = \psi_{i+j} + \sum_{u=1}^{2j-1} d_u^j \psi_{i+j-u} + d_{2j}^j \psi_{i-j} \qquad j < i \tag{2.23}$$

where the d_u^j are $K(\theta)$, its derivatives and positive powers of n.

Since ψ_j is a polynomial in ψ_1 and $\nu(x)$ (where $\nu(x) = n$ if $x \in X_n$) of exact degree j it follows from Equation 2.23 that

$$(\psi_j, \psi_i)_\theta = E_\theta(\psi_j \psi_i) = \sum_{u=1}^{i+j} \sum_{\ell=1}^{i+j} e_{u\ell} \frac{d^u}{d\theta^u} E_\theta(n^\ell) \tag{2.24}$$

where the $e_{u\ell}$ are functions of $K(\theta)$ and its derivatives. And

$$\psi_1^i \psi_i = \psi_{2i} + \sum_{u=1}^{2i-1} d_u^i \psi_{2i-u} + d_{2i}^i \psi_o \tag{2.25}$$

where

$$d_{2i}^i = \prod_{\ell=1}^{i} B_\ell \quad . \tag{2.26}$$

Hence

$$\|\psi_i\|_\theta^2 = (\psi_i, \psi_i)_\theta = \sum_u^{2i} \sum_\ell^{2i} e_{u\ell} \frac{d^u}{d\theta^u} E_\theta(n^\ell) + E_\theta(\prod_{\ell=1}^{i} B_\ell) \quad . \tag{2.27}$$

In particular

$$(\psi_1, \psi_i)_\theta = \frac{i}{K} \frac{d^{i-1}}{d\theta^i} E_\theta(n) \qquad i > 1 \tag{2.28}$$

and

$$\|\psi_1\|_\theta^2 = E_\theta(n)/K \quad . \tag{2.29}$$

Now from Equation 2.16 it may be verified that

$$K^{-1} = E_\theta\left[\frac{\partial \log f_1(x, \theta)}{\partial \theta}\right]^2 \tag{2.30}$$

where $f_1(x, \theta)$ is the density of the one dimensional random variable x. Hence

$$\|\psi_1\|^2 = E_\theta(n) \cdot E_\theta\left[\frac{\partial \log f_1(x, \theta)}{\partial \theta}\right]^2 \quad . \tag{2.31}$$

We will now use the elements of $\{\psi_k\}$ to construct a complete orthonormal set $\{\varphi_k\}$ for the Hilbert space V_θ. Let

$$\lambda_{ij} = (\psi_i, \psi_j)_\theta \quad . \tag{2.32}$$

Then we may define φ_k by

$$\varphi_k = \begin{vmatrix} \psi_1 & \psi_2 & \cdots & \psi_k \\ \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1k} \\ \vdots & & & \\ \lambda_{(k-1)1} & \lambda_{(k-1)2} & \cdots & \lambda_{(k-1)k} \end{vmatrix} \Bigg/ (\Lambda_k \Lambda_{k-1})^{\frac{1}{2}} \tag{2.33}$$

and

$$\varphi_o = 1 \tag{2.34}$$

where

$$\Lambda_k = \begin{vmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1k} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2k} \\ \vdots & & & \\ \lambda_{k1} & \lambda_{k2} & \cdots & \lambda_{kk} \end{vmatrix} \quad . \tag{2.35}$$

It can be verified that $\{\varphi_k\}$ is an orthonormal set and since $\{\psi_k\}$ is everywhere dense in V_θ, that $\{\varphi_k\}$ is a complete orthonormal set for V_θ. Now given any estimable parametric function $g(\theta)$, let U_g be the set of unbiased estimators for $g(\theta)$. Then we have the following result.

Theorem 2.3. Suppose $g(\theta)$ is estimable. Then any unbiased estimator $h(x)$ of $g(\theta)$ has a unique orthogonal projection $h_c(x)$ onto V_θ which is such that

$$\|h_c(x)\|_\theta \leq \|h(x)\|_\theta \quad \text{for any } h(x) \in U_g$$

with strict inequality unless $h(x)$ is $\mathcal{a}_c$-measurable, in which case $h_c(x) = h(x)$. Further if $g(\theta)$ is $\mathcal{a}_c$-estimable then $h_c(x)$ is the U.M.V.U.E. for $g(\theta)$.

Proof. Let $h(x)$ be an arbitrary element of U_g. Then the orthogonal projection of $h(x)$ onto V_θ has a limit in the mean (ℓ.i.m.) representation

$$h_c(x) = \ell.\text{i.m.} \sum_{k=0}^{\infty} (h, \varphi_k)\varphi_k \tag{2.36}$$

with the square norm given by

$$\|h_c(x)\|_\theta^2 = \sum_{k=0}^{\infty} (h, \varphi_k)^2 \leq \|h(x)\|_\theta^2 \tag{2.37}$$

using the property that any orthogonal projection operator has norm ≤ 1. Further the representation of $h_c(x)$ given above is independent of $h(x)$ since

$$(h, \varphi_k) = \begin{bmatrix} g'(\theta) & g''(\theta) & \cdots & g^{(k)}(\theta) \\ \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1k} \\ \vdots & & & \\ \lambda_{(k-1)1} & \lambda_{(k-1)2} & \cdots & \lambda_{(k-1)k} \end{bmatrix} \Bigg/ (\Lambda_k \Lambda_{k-1})^{\frac{1}{2}} \qquad k = 1, 2, \ldots \tag{2.38}$$

and

$$(h, \varphi_o) = g(\theta) \quad . \tag{2.39}$$

Hence (h, φ_k) is independent of $h(x)$ and depends only on $g(\theta)$. In fact (h, φ_k) is a linear combination of $g(\theta)$ and its derivatives. Therefore

$$\|h_c(x)\|_\theta \leq \|h(x)\|_\theta \quad \text{for any } h(x) \in U_g \text{ and } \theta \in \Omega \tag{2.40}$$

with strict inequality unless $h(x) \in V_\theta$ in which case it is $\mathcal{G}_c$-measurable and from the uniqueness of $h_c(x)$ is in fact equal to $h_c(x)$.

Since $\mathcal{G}_c$ is not a sufficient subfield, in general $h_c(x)$, which is also the conditional expectation of any unbiased estimator, will not be independent of θ. Hence $h_c(x)$ may not be an unbiased estimator of $g(\theta)$. However if $g(\theta)$ is $\mathcal{G}_c$-estimable then $h_c(x)$ will be the unique $\mathcal{G}_c$-measurable unbiased estimator of $g(\theta)$ and from Equation 2.40 $h_c(x)$ is the U.M.V.U.E. of $g(\theta)$. This completes the proof.

It is clear from the above theorem that in the general case in which $g(\theta)$ is estimable but not necessarily $\mathcal{G}_c$-estimable, the variance of $h_c(x)$ given by

$$\mathrm{Var}(h_c(x))_\theta = \sum_{k=1}^{\infty} (h, \varphi_k)^2 \qquad (2.41)$$

provides an unachievable lower bound for the variances of unbiased estimators of $g(\theta)$.

In this connection it is interesting to note, for the special case $g(\theta) = \theta$, the relationship between the variance of $h_c(x)$ and certain other bounds obtained by Blackwell and Girshick (8) and Wolfowitz (23) under different sets of assumptions, both different from those made here. For this special case we have by direct substitution in Equations 2.38 and 2.41

$$\mathrm{Var}(h_c(x)) = 1/\|\psi_1\|_\theta^2 + \sum_{k=2}^{\infty} (h, \varphi_k)\varphi_k \quad . \qquad (2.42)$$

And from Equation 2.31

$$\|\psi_1\|_\theta^2 = E_\theta(n)\, E_\theta\left[\frac{\partial \log f_1}{\partial \theta}\right]^2 \quad . \qquad (2.43)$$

Substituting Equation 2.43 in Equation 2.42 the first expression on the right hand side is the Blackwell-Girshick-Wolfowitz bound which is thus a special case of the bound given here. Clearly then the Blackwell-Girshick-Wolfowitz bound is not achievable even when θ is $\mathfrak{a}_c$-estimable unless

$$\sum_{k=2}^{\infty} (h, \varphi_k)\varphi_k = 0 \quad . \qquad (2.44)$$

By direct substitution or from Seth (22), Equation 2.44 holds if and only if

$$\lambda_{1k} = 0 \quad \text{for all } k > 1 \quad . \qquad (2.45)$$

However, from Equation 2.28 this is possible only if all the derivatives of $E_\theta(n)$ vanish identically and since by Lemma 2.5 $E_\theta(n)$ is analytic, this implies $E_\theta(n)$ is a constant, N say, which corresponds to the fixed sample case. This result on the unachievability of the Blackwell-Girshick-Wolfowitz bound has also been obtained by Blackwell and Girshick (8) and Seth (22).

The results obtained here are particularly useful when sampling from binomial populations in which case X is finite and every estimable parametric function which possesses a U.M.V.U.E. is $\mathcal{A}_c$-estimable and hence $h_c(x)$ is the U.M.V.U.E. As was noted earlier, in this case, the class of U.M.V.U.E.'s is precisely the class of $\mathcal{A}_c$-measurable square integrable $(\mathcal{P})$ functions. It is interesting to note that for the negative binomial application for which Assumption 2.1 has been verified it is possible to reparameterize the density function so that the Blackwell-Girschick-Wolfowitz bound is achieved.

Suppose now that Assumption 2.1 does not hold. Consider the subspace M_θ of H_θ spanned by $\{\psi_k\}$. It is easily verified that each ψ_k is $\mathcal{A}_o$ measurable and hence

$$M_\theta \subseteq \mathcal{L}_2(X, \mathcal{A}_o, p_\theta(x)) \quad \theta \in \Omega \quad .$$

This last result implies that any unbiased estimator in M_θ cannot be improved upon by Rao-Blackwellisation.

The next two lemmas establish certain optimal properties for the elements of M_θ.

Lemma 2.7. Any $h(x) \epsilon M_\theta$ is locally best for its expectation.

Proof. From a well known theorem, a necessary and sufficient condition for an estimator $h(x)$ to be locally best at θ for its expectation is that

$$(h, z)_\theta = 0 \quad \text{for all} \quad z \epsilon U_z$$

where U_z is the set of all unbiased estimators of zero in H_θ. However for any $z \epsilon U_z$

$$(\psi_k, z)_\theta = 0 \quad k = 0, 1, 2, \ldots \ .$$

Hence for any $h \epsilon M_\theta$, $(h, z)_\theta = 0$ and hence result.

Lemma 2.8. Given an estimable parametric function which possesses a U.M.V.U.E., if $h(x) \epsilon M_\theta$ for all $\theta \epsilon \Omega$, and is unbiased for $g(\theta)$ then $h(x)$ is the U.M.V.U.E. for $g(\theta)$.

Proof. From Lemma 2.7, $h(x)$ is the locally best unbiased estimator of $g(\theta)$. Hence if $h_o(x)$ is the U.M.V.U.E. of $g(\theta)$ then we have

$$\|h(x)\|_\theta = \|h_o(x)\|_\theta = C_\theta \quad \text{say}$$

and since $\frac{1}{2}(h(x) + h_o(x))$ is unbiased for $g(\theta)$ we have

$$\|\tfrac{1}{2}(h + h_o)\|_\theta \geq C_\theta \ .$$

However by Minkowski's inequality

$$\left\|\frac{1}{2}(h + h_o)\right\|_\theta \leq \frac{1}{2}C_\theta + \frac{1}{2}C_\theta = C_\theta .$$

From the last two inequalities we have

$$\|h + h_o\| = \|h\| + \|h_o\|$$

which implies $h(x) = h_o(x)$ since under the null set qualifications referred to in Chapter I, the empty set is the only $\mathcal{P}$-null set. We therefore have the following theorem.

Theorem 2.4. Suppose $g(\theta)$ is estimable. Let $h_M(x)$ be the orthogonal projection of any unbiased estimator of $g(\theta)$ onto M_θ, where θ is arbitrary but fixed. If $h_M(x) \in M_\theta$ for all $\theta \in \Omega$ then it is the locally best and hence U.M.V.U.E. for $g(\theta)$ if the latter exists.

Proof. From Lemmas 2.7 and 2.8 it is sufficient to show that if $h_M(x) \in M_\theta$ for all $\theta \in \Omega$ then $h_M(x)$ is unbiased for $g(\theta)$. However, from Lemma 2.5, $G(\theta)$, the expectation of $h_M(x)$, is analytic in θ and further by differentiating under the integral sign,

$$G^{(k)}(\theta) = (h_M, \psi_k)_\theta = (h, \psi_k)_\theta = g^{(k)}(\theta)$$

for any $h \in U_g$ and any integral k. Hence $G(\theta) = g(\theta)$ for all $\theta \in \Omega$; which proves the result.

It is clear from the above discussion that the series representation given for elements in V_θ in Equation 2.36 as well as the expression for the variance in Equation 2.37 remain valid for functions in M_θ. This follows from the fact that these expressions depend on the spanning of the relevant space by $\{\psi_k\}$ and not on what the subspace itself is.

It must be noted that the restriction to $\mathfrak{a}_c$-estimable functions is no longer required and that provided the projection is square integrable ($\mathcal{P}$) then the expression given in Equation 2.37 is the variance of the locally best unbiased estimator and hence provides an achievable variance bound expression at the point θ. If however the projection fails to be square integrable ($\mathcal{P}$) then we obtain a local unachievable variance bound which is still sharper than the Blackwell-Girshick-Wolfowitz bound.

We have so far not considered the pointwise convergence of the series representation of $h_c(x)$ even for the cases in which $g(\theta)$ is $\mathfrak{a}_c$-estimable and hence $h_c(x)$ is a statistic. No criteria have been obtained for the general case. In the next section we will consider the fixed sample situation for which Assumption 2.1 is valid as a special case of the sequential and consider the problem of pointwise convergence then. It may be noted that the results there have been obtained directly using the method developed here by Abbey and David (2).

It is also possible to generalize the results here in two possible directions. First we could regard the parameter space as an open set in p-dimensional Euclidean space $(p > 1)$. In this case $T_n(x)$ $(n = 1, 2, \ldots, N)$ will also be a p-dimensional vector of functions and so will $T(x)$. We will consider this possibility, for fixed sample situations only, in Chapter III. It is also possible to consider the more general sequential schemes in which sampling does not necessarily cease after a predetermined number, N, of steps. It seems that with some further assumptions, see for example Seth (22), the method developed here may be applicable to this case although no definite results in this direction have been obtained.

C. Fixed Sample Case and Pointwise Convergence

In this section we will consider the fixed sample situation and give criteria for the pointwise convergence of the series given in Equation 2.36 for the U.M.V.U.E. $h_c(x)$ of the parametric function $g(\theta)$. We will consider the fixed sample situation as a special case, where the random variable n takes only one value N, of the sequential problem dealt with in the last two sections.

The sample $x = (x_1, x_2, \ldots, x_N)$ is, here, taken from the sample space $(X, \mathcal{A})$ where X is a Borel subset of N dimensional Euclidean space, E^N, and $\mathcal{A}$ is the Borel field of subsets of X. As before we have a set of probability measures $\mathcal{P} = \{p_\theta(x) \mid \theta \in \Omega\}$ defined on $\mathcal{A}$ and such that each $p_\theta(x)$ admits an exponential density function $f(x, \theta)$ with respect to a fixed σ-finite measure $\mu(x)$ and Ω is an open interval of E^1. From Equation 2.1 we have

$$f(x, \theta) = \alpha(x)\beta(\theta)\exp\theta T(x) \qquad \theta \in \Omega \quad . \tag{2.46}$$

It is easily seen that the sufficient statistic $(N, T_N(x))$ and the statistic

$T(x)$ defined by Equation 2.5 are the same and hence the subfields $\mathcal{A}_o$ induced by $(N, T_N(x))$ and $\mathcal{A}_c$ induced by $T(x)$ are equal. Clearly then $\mathcal{A}_o$ is the complete sufficient subfield of $\mathcal{A}$ as is well known and further, in this situation, every estimable parametric function, $g(\theta)$, possesses a unique U.M.V.U.E. which is its only unbiased estimator that is $\mathcal{A}_o$-measurable. In fact it is well known that in this case, the class of all U.M.V.U.E.'s may be characterized as the class of all $\mathcal{A}_o$-measurable square integrable $(\mathcal{P})$ real valued functions on X.

For any arbitrary $\theta \in \Omega$, let H_θ be the real Hilbert space defined by

$$H_\theta = \mathcal{L}_2(X, \mathcal{A}, p_\theta) \tag{2.47}$$

with the usual inner product. It follows from the above remarks that the subspaces W_θ and V_θ, defined respectively by

$$W_\theta = \mathcal{L}_2(X, \mathcal{A}_o, p_\theta) \tag{2.48}$$

and

$$V_\theta = \mathcal{L}_2(X, \mathcal{A}_c, p_\theta) \tag{2.49}$$

are identical. Further from Lemma 2.4, the polynomials in $T(x)$ are dense in W_θ which implies that the set of functions $\{\psi_n\}$ defined in Equations 2.12 and 2.13 is dense in W_θ. Hence if $\{\varphi_n\}$ is a complete orthonormal set obtained from $\{\psi_n\}$, then the U.M.V.U.E., $h_c(x)$ of a given estimable parametric function $g(\theta)$ is given by

$$h_c(x) = \ell.i.m. \sum_{n=0}^{\infty} (h, \varphi_n)\varphi_n \tag{2.50}$$

with variance

$$V(h_c) = \sum_{n=1}^{\infty} (h, \varphi_n)^2 \tag{2.51}$$

since

$$(h, \varphi_o) = g(\theta) \quad . \tag{2.52}$$

It follows from Lemma 2.5 that for any unbiased estimator $h(x)$ of $g(\theta)$ we have

$$(h, \psi_n) = (h_c, \psi_n) = g^{(n)}(\theta) \quad . \tag{2.53}$$

Hence since φ_n is a linear combination of $\psi_o, \psi_1, \ldots, \psi_{n-1}, \psi_n$ it is clear that (h, φ_n) is a linear combination of $g(\theta)$ and its first n derivatives. These results imply that the expressions given in Equations 2.50 and 2.51 are obtainable without explicit knowledge of any unbiased estimator of $g(\theta)$.

We will now restrict attention to the subclass of exponential densities satisfying Equations 2.16, 2.17 and 2.18. From Equations 2.19, 2.20 and 2.21 we have respectively

$$\psi_{m+1} = (\psi_1 - A_m)\psi_m - B_m \psi_{m-1} \quad m = 1, 2, \ldots \tag{2.54}$$

where

$$A_m = mK'/K \tag{2.55}$$

and

$$B_m = m(m-1)K''/2K + mN/K \quad . \qquad (2.56)$$

Finally from Equations 2.24 and 2.27 with n (= N) fixed we have

$$(\psi_j, \psi_i)_\theta = 0 \qquad \text{for all } i \neq j \qquad (2.57)$$

and

$$\|\psi_i\|_\theta^2 = \prod_{\ell=1}^{i} B_\ell \quad i = 1, 2, \ldots \quad . \qquad (2.58)$$

For this subclass of exponential densities, then, the elements of the set $\{\psi_n\}$ are mutually orthogonal and the complete orthonormal set $\{\varphi_n\}$ defined by Equations 2.33 and 2.34 assume the particularly simple form

$$\varphi_n = \psi_n/\|\psi_n\| \qquad (2.59)$$

$$\varphi_o = 1 \quad . \qquad (2.60)$$

The following is a corollary to Theorem 2.3.

Corollary 2.1. For the subclass of the exponential densities satisfying Equations 2.16, 2.17 and 2.18, any estimable parametric function $g(\theta)$ possesses a unique U.M.V.U.E., $h_c(x)$, given by

$$h_c(x) = \ell.\text{i.m.} \sum_{n=0}^{\infty} g^{(n)}(\theta)\, f^{(n)}(x, \theta)/\|\psi_n\|^2 f(x, \theta)$$

with variance

$$V(h_c) = \sum_{n=1}^{\infty} [g^{(n)}(\theta)]^2/\|\psi_n\|^2 < \infty \quad .$$

Proof. Follows from Theorem 2.3, since in this situation every estimable parametric function $g(\theta)$ is also $\mathcal{G}_o$-measurable. The rest follows from Equations 2.59 and 2.60 and the implications that for any unbiased estimator $h(x)$ of $g(\theta)$,

$$(h, \varphi_n) = g^{(n)}(\theta)/\|\psi_n\| \quad n = 1, 2, \ldots$$

and

$$(h, \varphi_o) = g(\theta) \quad .$$

From Equation 2.18 it is clear that

$$K''(\theta) = 2a \quad \text{for all } \theta \in \Omega \quad . \tag{2.61}$$

Hence if $a < 0$ then we have from Equation 2.56

$$B_n \leq 0 \quad \text{for all } n \geq \frac{a-1}{a}(a-N)/a \quad . \tag{2.62}$$

However from Equation 2.58 this would imply that

$$\|\psi_n\|^2 \leq 0 \quad \text{for all } n \geq \frac{a-1}{a}(a-N)/a \tag{2.63}$$

which implies that

$$\|\psi_n\|^2 = 0 \quad \text{for all } n \geq \frac{a-1}{a}(a-N)/a \tag{2.64}$$

which further implies that

$$\psi_n(x) = 0 \quad \text{for all } n \geq \frac{a-N}{a}(a-N)/a \quad . \tag{2.65}$$

In this case the series given in Corollary 2.1 both terminate after a finite number of terms and hence in particular the series representation of the U.M.V.U.E., $h_c(x)$ is trivially convergent pointwise to $h_c(x)$. Let N_o be the greatest integer less than $(a-N)/a$. It appears from the construction that in this case where $\psi_n(x)$ vanishes for all $n > N_o$, no $g(\theta)$ having non-vanishing derivatives of higher order than N_o is estimable. An example of this occurs with the binomial distribution where

$$K(\theta) = \theta(1 - \theta) \tag{2.66}$$

and

$$a = -1 \quad .$$

Here

$$\frac{a-N}{a} = N + 1 \quad .$$

Hence

$$N_o = N \quad . \tag{2.67}$$

And therefore only polynomials in θ of degrees $\leq N$ may be estimable. This result is also given in DeGroot (10).

We will now consider the convergence of the series in the nontrivial situations $a \geq 0$. It is clear that the problem here is in the area of convergence of generalized Fourier series and it is possible that in specific situations the convergence may be obtained by more powerful techniques than those that will be considered here.

Two criteria will be considered here. The first one utilizes the analytic properties of $g(\theta)$ established in Lemma 2.5 and is derived from the following lemma.

Lemma 2.9. If $h_n(x)$ is a sequence of integrable (p) functions such that

$$\sum_{n=0}^{\infty} \int |h_n(x)| \, dp < \infty \quad .$$

Then $\Sigma h_n(x)$ converges pointwise to an integrable (p) function $h_c(x)$ and

$$\int h_c(x) = \sum_{n=0}^{\infty} \int h_n(x) \, dp \quad .$$

<u>Proof.</u> Consider the sequence S_N of partial sums, of the series

$$\sum_{n=0}^{\infty} |h_n(x)| \quad ,$$

given by

$$S_N = \sum_{n=0}^{N} |h_n(x)| \quad .$$

Clearly S_N is a non-decreasing sequence of real valued non-negative functions and hence

$$\int \sum_{n=0}^{\infty} |h_n(x)| \, dp = \sum_{n=0}^{\infty} \int |h_n(x)| \, dp < \infty \quad . \tag{2.68}$$

Hence

$$\sum_{n=0}^{\infty} |h_n(x)| \quad \text{if finite for } x \in X$$

and

$$h_c(x) = \sum_{n=0}^{\infty} h_n(x) \text{ is finite for } x \in X$$

since absolute convergence implies convergence.

Again

$$\left| \sum_{n=0}^{\infty} h_n(x) \right| \leq \sum_{n=0}^{N} |h_n(x)| \leq \sum_{n=0}^{\infty} |h_n(x)| \tag{2.69}$$

and from Equation 2.68, the right hand side of Equation 2.69 is integrable. Hence

$$\int \sum_{n=0}^{\infty} h_n(x)\, dp = \int h_c(x)\, dp = \sum_{n=0}^{\infty} \int h_n(x)\, dp$$

using the Lesbegue Dominated Convergence Theorem.

It may be noted that we have used the same function $h_c(x)$ for both the (ℓ.i.m.) limit as well as for the pointwise limit. That this is the case follows from the next lemma given in Feller(11) which we state here without proof.

Lemma 2.10. Given a sequence $\{h_n(x)\}$ such that

$$\ell.\text{i.m. } h_n(x) = h_o(x)$$

and

$$\lim_{n \to \infty} h_n(x) = h_c(x) \qquad x \in X \quad .$$

Then

$$h_o(x) = h_c(x) \quad .$$

It follows from Lemmas 2.9 and 2.10 that in order to obtain the convergence of the series given in Corollary 2.1 for $h_c(x)$ it is sufficient to consider the convergence of

$$\sum_{n=0}^{\infty} \int |g^{(n)}(\theta)| \, |f^{(n)}(x, \theta)| / \|\psi_n\|^2 |f(x, \theta)| \quad .$$

Using the Cauchy-Schwarz inequality and the definition of ψ_n we have

$$\sum_{n=0}^{\infty} \int |g^{(n)}(\theta)| \, |f^{(n)}(x, \theta)| / \|\psi_n\|^2 |f(x, \theta)|$$

$$\leq \sum_{n=0}^{\infty} |g^{(n)}(\theta)| / \|\psi_n\| \quad .$$

Hence if under some conditions the above series converges then under the same set of conditions the series representation of $h_c(x)$ converges pointwise to $h_c(x)$. Consider now the following cases.

Case I. $\qquad 0 < a \leq N$

In this case

$$\|\psi_n\|^2 \geq (a/K)^n (n!)^2 \quad .$$

Hence

$$\sum_{n=0}^{\infty} |g^{(n)}(\theta)| / \|\psi_n\| \leq \sum_{n=0}^{\infty} |g^{(n)}(\theta)| \, d^n / n! \tag{2.7.0}$$

where

$$d = (K/a)^{\frac{1}{2}} \quad .$$

Without loss of generality let $\theta = 0 \in \Omega$ then from the analytic properties of $g(\theta)$ given in Lemma 2.5

$$g(\theta) = \sum_{n=0}^{\infty} g^{(n)}(\theta)\, \theta^n/n! \quad \text{for} \quad |\theta| < \delta \tag{2.71}$$

where $\delta > 0$ is the radius of convergence of the Taylor series of $g(\theta)$ about $\theta = 0$. Hence if $d < \delta$, then it follows from the absolute convergence of the Taylor series in Equation 2.71 that the right hand side of Equation 2.70 converges and hence that the series for $h_c(x)$ also converges pointwise.

Case II. $\quad a > N$

Here

$$\|\psi_n\|^2 \geq (N/K)^n (n!)^2 \quad .$$

Hence from the arguments used in Case I, it is clear that the series for $h_c(x)$ converges if $d(=(K/N)) < \delta$ and δ is as defined before. It may be noted that in this particular case there is a strong possibility of convergence at least for sufficiently large values of N.

The case $a = 0$ presents some new difficulties with the criterion given here and will be dealt with in conjunction with the others under a different criterion.

This other criterion is derived from the following result given in Loeve (19) and is stated here without proof.

Lemma 2.11. Given a sequence $\{h_n(x)\}$ in W_θ such that

$$\sum_{n=1}^{\infty} \|h_n(x)\|^2 < \infty \quad .$$

Then $\sum_{n=1}^{\infty} h_n(x)$ converges pointwise provided

$$\sum_{n=1}^{\infty} n \log n \, \|h_n(x)\|^2 < \infty \quad . \tag{2.72}$$

Clearly the convergence of $\sum_{n=1}^{\infty} h_n(x)$ is sufficient. Further

$$n \log n \, \|h_n(x)\|^2 \leq n(n-1) \|h_n(x)\|^2 \quad n = 1, 2, \ldots \, . \tag{2.73}$$

Hence if

$$\sum_{n=1}^{\infty} n(n-1) \|h_n(x)\|^2 < \infty \quad ,$$

then the conditions of the lemma are satisfied since in our case with

$$h_n(x) = g^{(n)}(\theta) \, f^{(n)}(x, \theta) / \|\psi_n\|^2 f(x, \theta)$$

we have from Corollary 2.1

$$\sum_{n=1}^{\infty} \|h_n(x)\|^2 < \infty \quad .$$

Consider now the case $a = 0$. Then

$$\|\psi_n\|^2 = n! \, (N/K)^n \quad .$$

Hence

$$\sum_{n=1}^{\infty} n(n-1) \|h_n(x)\|^2 = \sum_{n=2}^{\infty} [g^{(n)}(\theta)]^2 (K/N)^n / (n-2)!$$

However the power series

$$\sum_{n=1}^{\infty} [g^{(n)}(\theta)]^2 (K/N)^n / n!$$

converges for all $N > 1$ since it is the variance of $h_c(x)$ given in Corollary 2.1 and its convergence does not depend on N but only on the estimability of $g(\theta)$. Hence the power series

$$\sum_{n=1}^{\infty} [g^{(n)}(\theta)]^2 d^n/n!$$

converges for all $0 < d < K$. This, however, implies that the twice differentiated series also converges for all d such that $0 < d < K$. Hence at least for $N > 1$,

$$\sum_{n=2}^{\infty} [g^{(n)}(\theta)]^2 (K/N)^n/(n-2)!$$

converges.

It now follows from Lemma 2.11 that the series representation of $h_c(x)$ is pointwise convergent. In the next section we will give some applications and also discuss a few related results. Abbey (1) has also applied this technique to the estimation of θ^2 in the case of normally distributed variables with unknown mean θ and variance 1.

D. Some Applications and Related Results

It is clear from the discussion in Bahadur (5) on Bhattacharyya bounds that the variance of the U.M.V.U.E. given here, for the case, $g(\theta) = \theta$ equals the limiting Bhattacharyya bound. We will illustrate this in the case of the negative binomial where the limiting Bhattacharyya bound has been obtained by Murty (20).

From Bahadur (5) the relationship between this method and the Rao-Blackwell method is apparent. In the sequential case where $\mathcal{C}_c \subset \mathcal{C}_o$, the method of Rao and Blackwell may improve a given unbiased estimator but may not yield the U.M.V.U.E. even if it exists. On the other hand if the parametric function is $\mathcal{C}_c$-estimable the method developed here does yield the U.M.V.U.E. If however $g(\theta)$ is estimable but not $\mathcal{C}_c$-estimable, then an unachievable variance bound is obtained by the approach adopted here. In the fixed sample situation and also those special cases of the sequential situation where $\mathcal{C}_c = \mathcal{C}_o$ the method developed here is equivalent to that of Rao and Blackwell.

Consider now the following negative binomial situation

$$f(x, p) = pq^x \quad 0 < p < 1 \quad x = 0, 1, 2, \ldots \; . \tag{2.74}$$

Taking p as the parameter, Murty (20) considered the Bhattacharyya bounds for variances of unbiased estimators of $g(p) = p$. He obtained the k^{th} bound as

$$L_k = p^2 q(q^{k-1} + q^{k-2} + \ldots + 1) \quad . \tag{2.75}$$

Hence

$$\lim_{k \to \infty} L_k = p^2 q(1 - q)^{-1} = pq \quad , \tag{2.76}$$

which we will now show to be the variance of the U.M.V.U.E. of p.

Using the reparameterization $\theta = p^{-1}$ we find

$$f(x; \theta) = (\theta - 1)^x \theta^{-(x+1)} \quad . \tag{2.77}$$

Hence

$$(x+1)-\theta = K(\theta)\,\psi_1(x) \tag{2.78}$$

where

$$K(\theta) = \theta(\theta-1) = \theta^2-\theta \ . \tag{2.79}$$

Clearly Equations 2.16, 2.17 and 2.18 are satisfied and hence from Equation 2.58

$$\|\psi_n\|^2 = [\theta(\theta-1)]^{-n}(n!)^2 \tag{2.80}$$

since from Equation 2.56

$$\begin{aligned} B_m &= m(m-1)/K + m/K \\ &= m^2/K \ . \end{aligned} \tag{2.81}$$

And with $g(\theta) = \theta^{-1}$, we have

$$g^{(n)}(\theta) = (-1)^n n!\ \theta^{-(n+1)} \ . \tag{2.82}$$

Hence from Corollary 2.1, the variance of the U.M.V.U.E., $h_c(x)$, of $g(\theta)$

$$\begin{aligned} \mathrm{Var}(h_c(x)) &= \sum_{n=1}^{\infty}(n!)^2\,\theta^{-2(n+1)}/(n!)^2[\theta(\theta-1)]^{-n} \\ &= \theta^{-2}(\theta-1) \\ &= pq \ . \end{aligned} \tag{2.83}$$

Consider now the following more general negative binomial problem

$$f(x,\ p) = \binom{X+N-1}{x} p^N q^X \qquad x = 0,\ 1,\ 2, \ldots \tag{2.84}$$

where $N \geq 1$, is a finite positive integer.

Using the reparameterization of the above special case we have

$$(X+N) - N\theta = \theta(\theta-1)\psi_1 \tag{2.85}$$

and

$$\|\psi_n\|^2 = [\theta(\theta-1)]^{-n} n!(n+N-1)!/(N-1)! \tag{2.86}$$

Again taking the case $g(\theta) = \theta^{-1}$ we obtain, from Corollary 2.1, the variance of the U.M.V.U.E., $h_c(x)$ to be

$$\begin{aligned} \mathrm{Var}(h_c(x)) &= (N-1)! \sum_{n=1}^{\infty} (n!)^2 \theta^{-2(n+1)} / n!(n+N-1)! [\theta(\theta-1)] \\ &= \theta^{-2} \sum_{n=1}^{\infty} [(\theta-1)/\theta]^n / \binom{n+N-1}{n} \quad . \end{aligned} \tag{2.87}$$

Since the above expression is finite for all $N \geq 1$, it follows that the series

$$\sum_{n=1}^{\infty} n(n-1)[(\theta-1)/\theta]^n / \binom{n+N-1}{n} = (N-1)! \sum_{n=1}^{\infty} [(\theta-1)/\theta]^n (n!)^2 / (n-2)!(n+N-1)! \tag{2.88}$$

converges. This implies by Lemma 2.8 that the series for $h_c(x)$ also converges pointwise. It is interesting to note that from Theorem 4.1 of DeGroot (10), which gives necessary and sufficient conditions for estimability of a given $g(\theta)$ for the negative binomial, we also obtain a sufficient condition for the convergence of the series representation of $h_c(x)$.

It is possible to regard the general problem of estimation as the inversion of certain transforms. This implies that in specific situations more direct methods may be available for the construction of the U.M.V.U.E. than the more general method given here. DeGroot (10) has done this for the negative binomial. Fend (12) has also given theorems which in specific situations may yield the U.M.V.U.E. for some subclasses of the exponential family. Fend (12) and Rao (21) have shown that if for the fixed sample situation the m^{th} Bhattacharyya bound is achieved when estimating θ in exponential families, then the U.M.V.U.E. for θ is a polynomial of degree m in the complete sufficient statistic. Similarly if the density function satisfies Equations 2.16, 2.17 and 2.18 then any estimable polynomial in θ of degree m has a U.M.V.U.E. which is itself a polynomial of degree m in the complete sufficient statistic. We will illustrate this with the following Example 1 of Fend (12).

Consider the density

$$f(x, \alpha) = \alpha^{-(1/n)} \exp[-x\alpha^{-(1/n)}] \quad 0<x, \ 0<\alpha \ . \tag{2.89}$$

Let

$$\theta = \alpha^{1/n} \ .$$

Then

$$f(x,\ \theta) = \theta^{-1} \exp - x\theta^{-1} \quad 0 < x,\ 0 < \theta \quad .$$

Hence we have

$$x - \theta = \theta^2 \psi_1 \tag{2.90}$$

and

$$\|\psi_m\|^2 = (m!)^2 \theta^{-2m} \quad . \tag{2.91}$$

Let

$$g(\theta) = \alpha = \theta^n \quad . \tag{2.92}$$

Then from a theorem of which this is an illustrative example, Fend (12) gives the U.M.V.U.E., $h_c(x)$ of $g(\theta)$ as

$$h_c(x) = x^n/n! \tag{2.93}$$

We will verify this by induction.

Now from Corollary 2.1

$$h_c(x) = \sum_{m=0}^{n} \psi_m \binom{n}{m} \theta^{(n+m)}/m! \tag{2.94}$$

It is easily verified that $h_c(x) = x^n/n!$ for $n = 1, 2$. Assume it holds for $n-1$, then

$$\sum_{m=0}^{n-1} \binom{n-1}{m} \theta^{n-1+m} \psi_m/m! = x^{n-1}/(n-1)! \tag{2.95}$$

Using Equation 2.90 it is clear from Equation 2.95 that

$$\theta^2\psi_1 \sum_{m=0}^{n-1}\binom{n-1}{m}\theta^{n-1+m}\psi_m/m! + \theta n^{-1}\sum_{m=0}^{n-1}\binom{n-1}{m}\theta^{n-1+m}\psi_m/m! = \frac{x^n}{n!} \quad . \quad (2.96)$$

Hence it is sufficient to show that the left hand side of Equation 2.96 equals the right hand side of Equation 2.94. However substituting in Equation 2.54 we have

$$\psi_1\psi_m = \psi_{m+1} + 2m\,\psi_m/\theta + m^2\psi_{m-1}/\theta^2 \quad . \quad (2.97)$$

Substituting from Equation 2.97 in Equation 2.96, the left hand side reduces to

$$\binom{n}{o}\theta^n\psi_o + n^{-1}\sum_{m=1}^{n-2}\left[\binom{n-1}{m-1}/(m-1)! + \binom{n-1}{m}(2m+1)/m! + \binom{n-1}{m+1}(m+1)/m!\right]\psi_m\theta^{n+m}$$

$$+ \binom{n}{n-1}\psi_{n-1}\theta^{2n-1}/(n-1)! + \binom{n}{n}\psi_n\theta^{2n}/n!$$

and the expression in square brackets can be shown to be equal to $\binom{n}{m}/m!$ This proves the result.

It is of some interest that the expression given for the variance in Corollary 2.1 may be useful in deciding the estimability of a given function. Thus if a given parametric function $g(\theta)$ is estimable the variance expression is finite, hence whenever that expression is not finite, $g(\theta)$ is not estimable for the exponential family of densities. We will illustrate this with the following example of estimating the inverse of the parameter in a Poisson distribution.

Let

$$f(x, \theta) = e^{-N\theta}\, \theta^{\sum_{i=1}^{N} x_i} / \prod_{i}^{N} x_i! \qquad \theta > 0 \quad x_i = 1, 2, \ldots$$

and

$$g(\theta) = \theta^{-1} \quad .$$

Then

$$\Sigma x_i - N\theta = \theta \psi_1 \quad .$$

Here

$$\|\psi_n\|^2 = n!(N/\theta)^n \quad .$$

Consider the expression for the variance in Corollary 2.1

$$\sum_{n=1}^{\infty} (n!)^2 \theta^{-2(n+1)} / n!(N/\theta)^n = \theta^{-2} \sum_{n=1}^{\infty} n!(\theta N)^{-n}$$

which clearly diverges for any θ and N.

The converse to the above situation is clearly false since we may obtain finite expressions for the variance even where $g(\theta)$ is not estimable. Any non-estimable polynomial in θ of finite degree may be used as an example. In particular the binomial distribution is particularly useful in providing examples of non-estimable polynomials.

It is also clear from the negative binomial example that in cases where the estimable parametric function possesses non-vanishing derivatives of all orders, no finite Bhattacharyya bound may be achievable since these bounds increase monotonically and the U.M.V.U.E. achieves the limiting bound.

So far we have limited attention to the case where θ is a scalar variable. The general case $\Omega \subseteq E^p$ $(p > 1)$ may be obtained using the same techniques as above. In some parts, the details are different and in the next chapter we will consider this problem for the fixed sample situation.

III. THE MULTIPARAMETER CASE

A. Introduction

This chapter deals with the general case of the fixed sample estimation problem where the parameter $\theta = (\theta_1, \theta_2, \ldots, \theta_p)$ is a p-dimensional vector $(p \geq 1)$. The special case $p = 1$ of this was dealt with as a special case of the sequential problem and this more general case can also be dealt with in much the same way. The treatment given here, whilst, in general follows the same lines as before, is more direct and is closer in details to that in Abbey and David (2) for the special case $p = 1$. As noted however the general ideas remain the same and wherever possible the results of Chapter II will be used.

Let $x = (x_1, x_2, \ldots, x_N)$ be a sample from the sample space $(X, \mathcal{A}, \mathcal{P})$ where X is a Borel subset of E^N, $\mathcal{A}$ is the Borel field of subsets of X and $\mathcal{P} = \{p_\theta(x) \mid \theta \in \Omega\}$ is a set of probability measures on $\mathcal{A}$ such that each $p_\theta(x)$ admits an exponential probability density function $f(x, \theta)$ with respect to a fixed σ-finite measure $\mu(x)$ which may be Lebesgue or counting measure. [In some applications it is convenient to regard the n^{th} coordinate x_n of x $(n = 1, 2, \ldots, N)$ itself as a vector.]

It follows from Lehmann (17) that we may assume the density function $f(x, \theta)$ to be of form

$$f(x, \theta) = \alpha(x)\,\beta(\theta) \exp \sum_{i=1}^{p} \theta_i T_i(x) \quad ; \quad \theta = (\theta_1, \theta_2, \ldots, \theta_p) \in \Omega \tag{3.1}$$

with Ω, an open p-dimensional rectangle, the natural parameter range.

Each $T_i(x)$, $(i = 1, 2, \ldots, p)$ is a real valued $\mathcal{A}$-measurable function and the vector statistic $T(x)$ defined by

$$T(x) = (T_1(x), T_2(x), \ldots, T_p(x)) \tag{3.2}$$

is sufficient and complete. *In terms of subfields this implies, the subfield* $\mathcal{A}_o^-$ of $\mathcal{A}$ induced on X by $T(x)$ is sufficient and complete.

By a theorem of Lehmann (17) it can be verified that $\gamma(\theta) = (1/\beta(\theta))$, with $\beta(\theta)$ as in Equation 3.1, is an analytic in each of the coordinates of θ at all interior points of Ω, and further that if $u_i(n)$ $i = 1, 2, \ldots, p$ denotes the n^{th} moment of $T_i(x)$ at θ, that is,

$$u_i(n) = \int_X T_i(x)^n \, dp_\theta(x) \quad , \tag{3.3}$$

then $u_i(n)$ exists for each i and for all n and that

$$u_i(n) = \beta(\theta)\, \gamma_i^{(n)}(\theta) \tag{3.4}$$

where $\gamma_i^{(n)}(\theta)$ is the n^{th} partial derivative of $\gamma(\theta)$ with respect to θ_i, the i^{th} component of θ. That is

$$\gamma_i^{(n)} = \frac{\partial^n}{\partial \theta_i^n} \gamma(\theta) \qquad i = 1, 2, \ldots, p \quad n = 1, 2, \ldots . \tag{3.5}$$

Using the above results we will show that the moments of $T(x)$ determine its distribution function and hence from Akhiezer (3) it would follow that the polynomials in $T(x)$ are dense in the Hilbert space V_θ of square integrable $(p_\theta(x))$ functions of $T(x)$.

Definition 3.1. By a polynomial in a vector $T(x) = (T_1(x), T_2(x), \ldots, T_p(x))$ we mean a polynomial in the components $T_i(x)$ $(i = 1, 2, \ldots, p)$.

Formally this definition leads to the following: If

$$J(x) = \sum_{i=1}^{p} \lambda_i T_i(x) \tag{3.6}$$

where the λ_i are arbitrary real constants; then a polynomial in $T(x)$ is a polynomial in $J(x)$.

Lemma 3.1. The moments of $T(x)$ determine its distribution function at each $\theta \in \Omega$.

Proof. From Feller (11) it is sufficient to prove the corresponding result for the function $J(x)$ given by Equation 3.6.

For any set of real constants $\lambda_1, \lambda_2, \ldots, \lambda_p$ we have

$$|u(n)|^{\frac{1}{n}} \leq \sum_{i=1}^{p} |\lambda_i| \, |u_i(n)|^{\frac{1}{n}} \tag{3.7}$$

where $u(n)$ is the n^{th} moment of $J(x)$. Hence

$$|u(n)/n!|^{\frac{1}{n}} \leq \sum_{i=1}^{p} |\lambda_i| \, |u_i(n)/n!|^{\frac{1}{n}} \tag{3.8}$$

and therefore

$$\begin{aligned} \ell\text{im sup}\,|u(n)/n!|^{\frac{1}{n}} &\leq \sum_{i=1}^{p} \ell\text{im sup}\,|\lambda_i| \, |u_i(n)/n!|^{\frac{1}{n}} \\ &= \sum_{i=1}^{p} \lambda_i \, \ell\text{im sup}\,|u_i(n)/n!|^{\frac{1}{n}} \quad . \end{aligned} \tag{3.9}$$

However from Equation 3.4

$$|u_i(n)/n!|^{\frac{1}{n}} = |\beta(\theta)\,\gamma_i^{(n)}(\theta)/n!|^{\frac{1}{n}}$$

$$= |\gamma_i^{(n)}(\theta)/n!|^{\frac{1}{n}}\,|\beta(\theta)|^{\frac{1}{n}} \quad .$$

And since $\beta(\theta)$ is finite for all $\theta \in \Omega$, it follows from the analytic properties of $\gamma(\theta)$ that

$$\ell\text{im sup}|u_i(n)/n!|^{\frac{1}{n}} = \ell\text{im sup}|\gamma_i^{(n)}(\theta)/n!| < \infty \quad \text{for each } i \quad .$$

Hence from Equation 3.9

$$\ell\text{im sup}|u(n)/n!|^{\frac{1}{n}} < \infty$$

which from Feller (11) or Kendall (15) and (16) establishes the lemma.

It now follows from Akhiezer (3) and Halmos (14) that the polynomials in $T(x)$, regarded as functions of x, are dense in the Hilbert space V_θ given by

$$V_\theta = \mathcal{L}_2(X, \mathcal{A}_o, p_\theta(x)) \quad . \tag{3.10}$$

Let H_θ be the Hilbert space defined by

$$H_\theta = \mathcal{L}_2(X, \mathcal{A}, p_\theta(x)) \quad . \tag{3.11}$$

Then V_θ is a subspace of H_θ and from Bahadur (4) the orthogonal projection operator P from H_θ onto V_θ is precisely the conditional expectation operator given the subfield $\mathcal{A}_o$ of $\mathcal{A}$. Hence since $\mathcal{A}_o$ is sufficient and complete, projection onto V_θ yields the U.M.V.U.E. when applied to any $h \in H_\theta$ which is an unbiased estimator of a given estimable parametric function $g(\theta)$. In the next section, we will use the results obtained in this to, obtain a limit in the mean representation of the U.M.V.U.E.

B. Construction of U.M.V.U.E.

Let $g(\theta)$ be estimable and let U_g be the class of unbiased estimators of $g(\theta)$ in H_θ. It is clear from the arguments of the preceding section that

$$U_g \cap V_\theta = \{h_c(x)\} \tag{3.12}$$

where $h_c(x)$ is the U.M.V.U.E. of $g(\theta)$. Further if any arbitrary element $h(x)$ of U_g is available, $h_c(x)$ is obtainable by the method of Rao-Blackwellisation. We will define a complete orthonormal set for V_θ and then use the characterization of $h_c(x)$ given by Equation 3.12 to obtain a representation of $h_c(x)$ and of its variance without knowledge of any $h(x)$ in U_g whatever. This construction is possible because of certain analytic properties of $g(\theta)$ which are given by the next lemma, and which have been established in Lemma 2.5 for the case $p = 1$.

Lemma 3.2. Let $h(x)$ be an unbiased estimator of $g(\theta)$. Then

(i) the parametric function

$$g(\theta) = \int_X h(x)\,\alpha(x)\,\beta(\theta) \exp \sum_{i=1}^{p} \theta_i T_i(x)\, d\mu(x)$$

considered as a function of the complex variables $\theta_j = \xi_j + i\nu_j$ is an analytic function in each of these variables in the region R of parameter points for which $(\xi_1, \xi_2, \ldots, \xi_p)$ is an interior point of Ω.

(ii) the derivatives of all orders with respect to the θ's of $g(\theta)$ can be computed under the integral sign.

Proof. Since $\beta(\theta)$ is analytic in each of the θ's, it is sufficient to consider the integral $I(\theta)$ defined by ,

$$I(\theta) = \int_X h(x)\,\alpha(x) \exp \Sigma\, \theta_i T_i \, d\mu(x) \quad .$$

Now from Lehmann (17), it is sufficient to prove the result for integrals of form

$$I(\theta_1) = \int h(x)\,\alpha(x) \exp \theta_1 T_1(x)\, d\mu(x) \quad .$$

However the proof of this follows directly that of Lemma 2.5.

Let $n \geq 0$ be any integer and let $\{n_i | i = 1, 2, \ldots, p\}$ be any set of non-negative integers such that

$$\sum_{i=1}^{p} n_i = n \qquad n_i = 0, 1, 2, \ldots, n \text{ for each } i \quad . \tag{3.13}$$

Let

$$\psi_n(x; n_1, \ldots, n_p) = \left[\frac{\partial^n}{\partial\theta_1^{n_1}, \ldots, \partial\theta_p^{n_p}} f(x, \theta)\right] / f(x, \theta) \quad n_i \text{ not all zero} \tag{3.14}$$

and

$$\psi_o(x) = \psi_n(x, n_1, \ldots, n_p) = 1 \qquad n_i = 0 \text{ for all } i \text{ and any } n \tag{3.15}$$

since we may take $f(x, \theta) > 0$ $x \in X$, $\psi_n(x, n_1, \ldots, n_p)$ is well defined for each $n \geq 0$.

It is easily verified that for any n and any set $\{n_i\}$ satisfying Equation 3.13, $\psi_n(x, n_1, \ldots, n_p)$ is a polynomial in $T(x)$ of exact degree n. Moreover since p is finite, there is only a countably infinite number of these polynomials. Let $\{\psi_n(x)\}_{n=0}^{\infty}$ be any enumeration of them such that $\psi_o(x) = 1$. It follows from Cheney (9) that the elements of $\{\psi_n\}$ are linearly independent and further that any polynomial in $T(x)$ may be obtained as a linear combination of these elements. Using the Gram-Schmidt orthonormalization process, therefore, we may obtain a complete orthonormal set $\{\varphi_n\}$ for V_θ with each φ_n a linear combination of $\psi_o, \psi_1, \psi_2, \ldots, \psi_n$; since the set of polynomials $\{\psi_n\}$ is dense in V_θ. Clearly we have $\varphi_o = 1$ since ψ_o is orthogonal to ψ_n for any $n > 0$.

Now if $h(x)$ is any unbiased estimator of $g(\theta)$ then from Lemma 3.2 we have

$$\int_X h(x)\psi_n(x; n_1, \ldots, n_p)dp_\theta(x) = \frac{\partial^n}{\partial\theta_1^{n_1}, \ldots, \partial\theta_p^{n_p}} g(\theta) \quad . \tag{3.16}$$

And in particular

$$\int_X h(x)\psi_o(x)\, dp_\theta(x) = g(\theta) \quad . \tag{3.17}$$

Let

$$(h, \varphi_n)_\theta = \int h(x)\,\varphi_n(x)\, dp_\theta(x) \quad . \tag{3.18}$$

Then since φ_n is a linear combination of $\{\psi_n\}$ it follows that from Equations 3.16, 3.17 and 3.18 that the inner product $(h, \varphi_n)_\theta$ is a linear combination of $g(\theta)$ and its partial derivatives and is independent of the choice of h in U_g.

We therefore have the following theorem which gives the limit in the mean representation of the U.M.V.U.E., $h_c(x)$ of $g(\theta)$, as well as an expression for its variance.

Theorem 3.1. Any estimable parametric function, $g(\theta)$, possesses a unique U.M.V.U.E., $h_c(x)$, which is given by

$$h_c(x) = \ell.\mathrm{i.m.} \sum_{n=0}^{\infty} (h, \varphi_n)_\theta \, \varphi_n$$

with variance

$$\mathrm{Var}(h_c(x)) = \sum_{n=1}^{\infty} (h, \varphi_n)_\theta^2$$

where (h, φ_n) is a linear combination of $g(\theta)$ and its partial derivatives and h is any unbiased estimator of $g(\theta)$.

Proof. Follows from the fact that $\{\varphi_n\}$ is a complete orthonormal set for V_θ.

The series for $h_c(x)$ given in Theorem 3.1 is too general for any criteria for pointwise convergence to be given. In specific applications or special cases it may be possible to examine the possibility of convergence either by the use of Lemmas 2.6 and 2.8 and their possible generalizations or by the techniques available in the general area of pointwise convergence of generalized Fourier series.

In the next section we will consider some applications and wherever possible discuss the problem of convergence.

C. Some Special Cases and Conclusions

For the exponential class of densities considered in this chapter it is readily verified from Fend (12) and Seth (22) that the density function, $f(x, \theta)$ satisfies the following set of p equations

$$T_i(x) - N\theta_i = \sum_{n_1, n_2, \ldots, n_p} {}^{(i)}K_{n_1, n_2, \ldots, n_p} \psi(n_1, n_2, \ldots, n_p) \quad i=1,2,\ldots,p \tag{3.19}$$

where the K's depend only on θ and in this section we are writing

$$\psi(n_1, n_2, \ldots, n_p) = \psi_n(n_1, n_2, \ldots, n_p); \; \sum^{p} n_i = n \quad . \tag{3.20}$$

We will consider two special cases of Equation 3.19 both with $p = 2$. The first generalizes the scalar parameter case dealt with in sections C and D of the previous chapter. Thus consider the case in which the density function satisfies the two equations

$$T_1(x) - N\theta_1 = K_1(\theta_1)\, \psi(1, 0) \tag{3.21}$$

and

$$T_2(x) - N\theta_2 = K_2(\theta_2)\, \psi(0, 1) \tag{3.22}$$

where $K_i(\theta_i)$, is a quadratic in θ_i; $(i = 1, 2)$ with constant coefficients independent of x.

Using Equation 3.21, it follows from Seth (22) that

$$\psi(n+1, 0) = \psi(1, 0)\, \psi(n, 0) - A_{1n}\psi(n, 0) - B_{1n}\psi(n-1, 0) \tag{3.23}$$

$$n = 1, 2, \ldots$$

where

$$A_{1n} = nK_1'/K \tag{3.24}$$

and

$$B_{1n} = n(n-1)K''/2K + nN/K \quad . \tag{3.25}$$

Further

$$\|\psi(n, 0)\|^2 = \prod_{i=1}^{n} B_{1i} \quad . \tag{3.26}$$

It is easily verified that the corresponding equations for $\psi(0, n)$ are valid and further by induction that

$$\psi(n, m) = \psi(n, 0)\,\psi(0, m) \quad . \tag{3.27}$$

Hence taking expectations of both sides of the above equation,

$$(\psi(n, 0), \psi(0, m))_\theta = 0 \quad n, m \geq 0 \quad \theta \in \Omega \quad . \tag{3.28}$$

Again from Seth (22)

$$\psi(n, 0) = \sum_{\ell=0}^{n} a_\ell \psi(1, 0)^\ell \tag{3.29}$$

where the a's are real constants.

Let $j \geq 0$ be an arbitrary but fixed integer, then from Equation 3.27 and successive application of Equation 3.23, we have

$$\psi(1, 0)^j \psi(n, m) = \sum_{\ell=0}^{n+j} b_\ell \psi(\ell, 0)\,\psi(0, m) \tag{3.30}$$

where the b's are functions of the B's in Equation 3.25 and $b_o = 0$ unless $j=n$.

Now from Equations 3.29 and 3.30 we have

$$\psi(j,\ 0)\,\psi(n,\ m) \ = \ \sum_{\ell=0}^{n+j} d_\ell \psi(\ell,\ 0)\,\psi(0,\ m) \tag{3.31}$$

where the d's are constants and $d_o = 0$ unless $j = n$. Using the corresponding equations for $\psi(0,\ n)$ we immediately obtain, by repeating the arguments above,

$$\psi(0,\ \ell)\,\psi(j,\ 0)\,\psi(n,\ m) \ = \ \sum_{i=0}^{n+j} \sum_{u=0}^{m+\ell} c_{iu} \psi(i,\ 0)\,\psi(0,\ u) \tag{3.32}$$

with $c_{oo} = 0$ unless $j = n$ and $\ell = m$.

And substituting from Equation 3.27 in Equation 3.32 we have

$$\psi(j,\ \ell)\,\psi(n,\ m) \ = \ \sum_{i=0}^{n+j} \sum_{u=0}^{m+\ell} c_{iu} \psi(i,\ 0)\,\psi(0,\ u) \quad . \tag{3.33}$$

From Equation 3.28, therefore, we have upon taking expectations

$$(\psi(j,\ \ell),\ \psi(n,\ m))_\theta \ = \ 0 \quad \text{unless } j = n \text{ and } \ell = m \quad . \tag{3.34}$$

And hence from Equation 3.26 and the corresponding one for $\psi(0,\ m)$,

$$(\psi(n,\ m),\psi(n,\ m))_\theta \ = \ \|\psi(n,\ m)\|_\theta^2 = (\prod_{i=1}^{n} B_{1i})(\prod_{j=1}^{m} B_{2j}) \quad . \tag{3.35}$$

Clearly then, the elements of $\{\psi(n,\ m)\}$ are mutually orthogonal with norms obtainable from Equation 3.35.

We therefore have the following corollary to Theorem 3.1.

Corollary 3.1. For the exponential class of densities satisfying Equations 3.21 and 3.22, such that $K_i(\theta_i)$ is a quadratic in θ_i $(i = 1,\ 2)$, any estimable parametric function, $g(\theta)$, possesses a unique U.M.V.U.E., $h_c(x)$, which is given by

$$h_c(x) = \ell.\text{i.m.} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} g_{nm}(\theta)\, \psi(n, m)/\|\psi(n, m)\|^2$$

with variance

$$\text{Var}(h_c(x)) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} g_{nm}(\theta)^2/\|\psi(n, m)\|^2 - g(\theta)^2 \quad ,$$

where

$$g_{nm}(\theta) = \frac{\partial^{n+m}}{\partial\theta_1^n \, \partial\theta_2^m} g(\theta)$$

and $\psi(n, m)$ is given by Equations 3.14 and 3.20.

Proof. Follows directly from Theorem 3.1.

For the case $p = 1$, Corollary 3.1 reduces to Corollary 2.1. In particular if $T_1(x)$ and $T_2(x)$ are independent so that the density function $f(x, \theta)$ factors into a product $F_1(x, \theta_1)F_2(x, \theta_2)$ then we may obtain pointwise convergence of the series for $h_c(x)$ given in Corollary 3.1 at least for $g(\theta)$ a polynomial in θ_1, θ_2 from the lemmas given in Chapter II; since in this case it is easy to show that the U.M.V.U.E. for $g(\theta)$ will be a sum of products of U.M.V.U.E.'s each of which has a pointwise convergent series.

Again if $g(\theta)$ is a function of θ_i $(i = 1, 2)$ only then it is possible to obtain pointwise convergence criteria from those given in Chapter II. It is also easy to see that the restriction $p = 2$ is purely for convenience and the considerations here are valid for $p \geq 2$ but finite.

Example 1. Consider a random sample of size N from a bi-parameter normal population with known variances σ_1^2, σ_2^2, zero correlation and an unknown vector $\theta = (\theta_1, \theta_2)$ of means. Let $x = (x_{11}x_{21}, x_{12}x_{22}, \ldots, x_{1N}x_{2N})$. Then by direct calculation we have

$$\sum^{N} x_{1i} - N\theta_1 = \sigma_1^2 \psi(1, 0)$$

$$\sum^{N} x_{2i} - N\theta_2 = \sigma_2^2 \psi(0, 1)$$

and

$$\|\psi(n, m)\|^2 = N^{n+m} n!\, m!$$

Clearly the conditions for Corollary 3.1 all hold and hence all the conclusions arrived at above hold for estimation problems involving θ.

Seth (22) has considered the following example of the more general form of Equation 2.19 $(p = 2)$.

Example 2. Let $x = (x_1, \ldots, x_N)$ be a random sample from a normal distribution with mean θ_2 and variance θ_1. Then Seth (22) has shown that the density function satisfies the equations

$$T_1(x) - N\theta_1 = 2N\theta_1^2(N-1)^{-1}\psi(1, 0) - \theta_1^2(N-1)\psi(0, 2) \qquad (3.37)$$

and

$$T_2(x) - N\theta_2 = \theta_1 \psi(0, 1) \qquad (3.38)$$

where

$$T_1(x) = N(N-1)^{-1} \sum^{N} (x_i - \bar{x})^2 \qquad (3.39)$$

and

$$T_2(x) = \Sigma x_i \quad . \qquad (3.40)$$

In this case the elements of $\{\psi(n, m)\}$ are not all mutually orthogonal. However from Equation 3.38 it is easily verified that the elements of the subset $\{\psi(n, 0)\}$ are mutually orthogonal and also that $\psi(0, 1)$ is orthogonal to every other element of $\{\psi(n, m)\}$. It also follows from Equation 3.37 that $\psi(1, 0)$ and $\psi(0, 2)$ are not orthogonal. Indeed

$$(\psi(1, 0), \psi(0, 2)) = N/\theta_1^2 = \lambda_{21} \quad . \qquad (3.41)$$

The following results may also be verified:

$$\|\psi(1, 0)\|^2 = N/2\theta_1^2 \qquad (3.42)$$

$$\|\psi(0, 1)\|^2 = N/\theta_1 \qquad (3.43)$$

$$\|\psi(0, 2)\|^2 = 2N^2/\theta_1^2 \qquad (3.44)$$

$$(\psi(1, 0), \psi(n, 0))_\theta = 0 \qquad (3.45)$$

$$(\psi(0, 2), \psi(n, 0))_\theta = 0 \quad . \qquad (3.46)$$

Hence if $g(\theta) = \theta_2^2$, then from Theorem 3.1, the U.M.V.U.E. $h_c(x)$ is given by

$$h_c(x) = \theta_2^2 + 2\theta_2 \psi(0,1)/\|\psi(0,1)\|^2 + 2\psi(0,2)/\|\psi(0,2)\|^2 \tag{3.47}$$

$$+ \begin{vmatrix} 2 & 0 \\ \|\psi(0,2)\|^2 & \lambda_{21} \end{vmatrix} \begin{vmatrix} \psi(0,2) & \psi(1,0) \\ \|\psi(0,2)\|^2 & \lambda_{21} \end{vmatrix} \Bigg/ \|\psi(0,2)\|^2 \begin{vmatrix} \|\psi(0,2)\|^2 & \lambda_{21} \\ \lambda_{21} & \|\psi(1,0)\|^2 \end{vmatrix} .$$

Substituting from Equations 3.41, 3.42, 3.43, 3.44 we obtain

$$h_c(x) = \theta_2^2 + 2N\theta_2\theta_1^{-1}\psi(0,1) + \theta_1^2 N^{-2}\psi(0,2) + \theta_1^2 N^{-2}(N-1)^{-1}[\psi(0,2) - 2N\psi(1,0)] . \tag{3.48}$$

Using Equations 3.37 and 3.38 this further reduces to

$$h_c(x) = [T_2(x)/N]^2 - T_1(x)/N^2 \tag{3.49}$$

which is known to be the U.M.V.U.E. of θ_2^2. Again from Theorem 3.1 the variance is given by

$$\mathrm{Var}(h_c(x)) = 4\theta_2^2/\|\psi(0,1)\|^2 + 4/\|\psi(0,2)\|^2$$

$$+ \begin{vmatrix} 2 & 0 \\ 2N^2/\theta_1^2 & N/\theta_1^2 \end{vmatrix}^2 \Bigg/ (2N^2/\theta_1^2) \begin{vmatrix} 2N^2/\theta_1^2 & N/\theta_1^2 \\ N/\theta_1^2 & N/2\theta_1^2 \end{vmatrix}$$

$$= 4\theta_2^2\theta_1/N + 2\theta_1^2/N^2 + 2\theta_1^2/N^2(N-1)$$

$$= 4\theta_2^2\theta_1/N + 2\theta_1^2/N(N-1) .$$

It is interesting to compare the results here with those obtained by Abbey (1) for estimating θ_2^2 when θ_1 is assumed known. In that case

$\psi(n, 0)$ are all zero and hence do not enter into the construction. The results here indicate that even where the parametric function $g(\theta)$ depends only on one coordinate, its U.M.V.U.E. will usually involve derivatives of the density function with respect to the other coordinates of θ if these are not orthogonal to all the derivatives with respect to the relevant coordinate. In some cases it may be possible by reparameterization to obtain mutual orthogonality of the $\{\psi(n, m)\}$. From the above discussion, this is desirable since it leads to less computations in the construction of the U.M.V.U.E. as well as making the criteria for pointwise convergence developed in Chapter II applicable.

IV. LITERATURE CITED

1. Abbey, Joseph L. Series representation of uniformly minimum variance unbiased estimators. Unpublished M.S. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology. 1967.

2. Abbey, Joseph L. and David, H. T. Series representation of uniformly minimum variance unbiased estimators for exponential distributions I. To be published. Annals of Mathematical Statistics. ca. 1968.

3. Akhiezer, N. I. The classical moment problem. 1st ed. London, England, Oliver and Boyd. 1965.

4. Bahadur, R. R. Measurable subspaces and subalgebras. American Mathematical Society Proceedings 6: 565-570. 1955.

5. Bahadur, R. R. On unbiased estimation of uniformly minimum variance. Sankhyā 18, Parts 3 and 4: 211-224. 1957.

6. Bahadur, R. R. Sufficiency and statistical decision functions. Annals of Mathematical Statistics 25: 423-462. 1954.

7. Blackwell, David. Conditional expectation and unbiased sequential expectation. Annals of Mathematical Statistics 18: 105-110. 1947.

8. Blackwell, David and Girshick, M. A. A lower bound for the variance of some unbiased sequential estimates. Annals of Mathematical Statistics 18: 277-280. 1947.

9. Cheney, E. W. Introduction to approximation theory. New York, New York, McGraw-Hill Book Company, Inc. 1966.

10. DeGroot, M. H. Unbiased sequential estimation for binomial populations. Annals of Mathematical Statistics 30: 80-101. 1959.

11. Feller, William. An introduction to probability theory and its applications. Vol. 2. New York, New York, John Wiley and Sons, Inc. 1966.

12. Fend, A. V. On the attainment of Cramer-Rao and Bhattacharyya bounds for the variance of an estimate. Annals of Mathematical Statistics 30: 381-388. 1959.

13. Girshick, M. A., Mosteller, Frederick and Savage, L. J. Unbiased estimates for certain binomial sampling problems. Annals of Mathematical Statistics 17: 13-23. 1947.

14. Halmos, P. R. Measure theory. New York, New York, Van Nostrand Company, Inc. 1950.

15. Kendall, M. G. Conditions for uniqueness in the problem of moments. Annals of Mathematical Statistics 11: 402-409. 1940.

16. Kendall, M. G. Corrections to a paper on the uniqueness problem of moments. Annals of Mathematical Statistics 12: 464-465. 1941.

17. Lehmann, E. L. Testing statistical hypotheses. New York, New York, John Wiley and Sons, Inc. 1959.

18. Lehmann, E. L. and Stein, C. Completeness in the sequential case. Annals of Mathematical Statistics 21: 376-385. 1950.

19. Loève, M. Probability theory. New York, New York, Van Nostrand Company, Inc. 1955.

20. Murty, V. N. A note on Bhattacharyya bounds for the negative binomial distribution. Annals of Mathematical Statistics 27: 1182-1183. 1956.

21. Rao, C. R. Some theorems on minimum variance estimation. Sankhyā 12: 27-42. 1952.

22. Seth, G. R. On the variance of estimates. Annals of Mathematical Statistics 20: 1-27. 1949.

23. Wolfowitz, J. The efficiency of sequential estimates and Wald's equation for sequential processes. Annals of Mathematical Statistics 18: 215-230. 1947.

V. ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to Dr. H. T. David for his assistance and encouragement in the planning and preparation of this dissertation.

The author is also indebted to all his committee members and to Dr. B. K. Kale for suggestions given through numerous conversations.

Finally the author wishes to thank the Government of Ghana for financial assistance and Dr. T. A. Bancroft who was instrumental in making the last year of his graduate study financially possible. Part of the research on this thesis was sponsored by National Science Foundation Grant GP-6013.