The Annals of Mathematical Statistics
1970, Vol. 41, No. 4, 1217–1226

THE CONSTRUCTION OF UNIFORMLY MINIMUM VARIANCE UNBIASED ESTIMATORS FOR EXPONENTIAL DISTRIBUTIONS[1]

BY J. L. ABBEY[2] AND H. T. DAVID

Iowa State University

1. Summary and Introduction. Consider a sample $(x_1, x_2, \cdots, x_N)$ from a population with a distribution function $F_\theta(x)$, $(\theta \in \Omega)$ for which a complete sufficient statistic, $s(x)$, exists. Then any parametric function $g(\theta)$ possesses a unique minimum variance unbiased estimator U.M.V.U.E., which may be obtained by the Rao–Blackwell theorem provided an unbiased estimator of $g(\theta)$ with finite variance for each $\theta \in \Omega$ is available. In this paper we will consider the Koopman–Darmois class of exponential densities and develop a method for obtaining the U.M.V.U.E., t_g, of $g(\theta)$ without explicit knowledge of any unbiased estimator of $g(\theta)$. The U.M.V.U.E. t_g is given as the limit in the mean (l.i.m.) of a series and a convergent series is also given for the variance.

For any arbitrary but fixed $\theta_0 \in \Omega$, it can be verified that the complete sufficient statistic $s(x)$ has moments of all orders and that these moments determine its distribution function. Hence the set of polynomials in $s(x)$ is dense in the Hilbert space, V (with the usual inner product), of Borel measurable functions of $s(x)$. Since t_g is an element of V, we may obtain a generalized Fourier series for it by constructing a complete orthonormal set $\{\varphi_n\}$ for V. Such a set $\{\varphi_n\}$ may be obtained from the density function and its derivatives with respect to θ. For a subclass of the exponential family, Seth [18] has obtained $\{\varphi_n\}$ in a form which is convenient for our purposes. We will study this case in Section 3 and use Seth's results to give an explicit construction of t_g. Criteria for the pointwise convergence of the series will also be given. In Section 4 examples illustrating the use of the method are given and some related results are discussed.

The general theory for the representation of minimum variance unbiased estimates, both local and uniform, has been developed in depth, for example in [5], [18], [19], [16], [3], and [4]. The present remarks, though founded in the general theory (in particular [3] and [4]), are tailored specifically to the exponential family.

2. Preliminaries and construction. Let x be a sample from sample space $(X, \mathscr{B})$; where X is a Borel subset of N-dimensional Euclidean space and $\mathscr{B}$ is the Borel field of subsets of X. Let $\mathscr{P} = \{p_\theta(x) \mid \theta \in \Omega\}$ be a set of probability measures admitting exponential densities $\{f(x;\theta) \mid \theta \in \Omega\}$ with respect to a fixed σ-finite measure $\lambda(x)$, (either Lebesgue or counting measure). Following Lehmann [13] we will write

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

with Ω, an open interval, the natural parameter range.

Received April 1, 1968; revised November 24, 1969.

[1] This research was partially supported by National Science Foundation Grant GP-6013.

[2] Now at the University of Ghana.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics*. www.jstor.org

DEFINITION 2.1. A parametric function $g(\theta)$ will be said to be estimable if there exists at least one $\mathscr{B}$-measurable, square integrable $(\mathscr{P})$ real-valued function $h(x)$ such that

$$\int h(x)\,dp_\theta(x) = g(\theta) \qquad \text{for all} \quad \theta\in\Omega. \tag{2.2}$$

Let Y denote the range of s; $\mathscr{A}$ the Borel field of subsets of Y and $\mathscr{B}_0$ the subfield of $\mathscr{B}$ induced by s. Then $\mathscr{B}_0$ is a sufficient and complete subfield. Let θ_0 be arbitrary but fixed, $\mu(x) = p_{\theta_0}(x)$ and $v(s)$ the corresponding measure on $\mathscr{A}$ induced by s. By Theorem 9, Chapter 2, of [13] it can be verified that $\gamma(\theta) = (1/\beta(\theta))$ is analytic at all points of Ω (indeed, a slight modification of the argument shows $g(\theta)$ analytic there) and that the moments, $\{u(n)\}$, of $s(x)$ at θ_0 exist and are given by

$$u(n) = \int s^n\,dv(s) = \beta(\theta_0)\gamma^{(n)}(\theta_0). \tag{2.3}$$

It can be further verified from (2.3), the analytic properties of $\gamma(\theta)$ and Feller [8] or Kendall [11], [12], that the set of moments $\{u(n)\}$ determines the distribution function corresponding to $v(s)$. Hence the set of polynomials $\{s^n\}$ is dense in the Hilbert space $W = \mathscr{L}_2(Y, \mathscr{A}, v)$; Akhiezer [2]. However the $\mathscr{A}$-measurable, square integrable (v) functions of s are precisely the $\mathscr{B}_0$-measurable, square integrable (μ) functions of x; Halmos [10]; hence the set of functions $\{s(x)^n\}$ is dense in $V = \mathscr{L}_2(X, \mathscr{B}_0, \mu)$ which is a subspace of $H = \mathscr{L}_2(X, \mathscr{B}, \mu)$. Further if P is the orthogonal projection operator from H onto V, then from Bahadur ([3], [4]), P is also the conditional expectation operator given $\mathscr{B}_0$ and hence projection onto V yields the U.M.V.U.E. when applied to any $h\in H$ satisfying (2.2).

Differentiating w.r.t. any strictly monotone analytic function of θ, let

$$\psi_n(x) = f^{(n)}(x;\theta_0)/f(x;\theta_0) \qquad n = 1, 2, \cdots \tag{2.4}$$

and $\psi_0(x) = 1$. Then it is easily seen that the set $\{\psi_n(x) \mid n = 0, 1, 2, \cdots\}$ is a set of polynomials in $s(x)$ dense in V in view of the above remarks, and for any unbiased estimator, $h(x)$, of $g(\theta)$

$$(h, \psi_n) = \int h(x)\psi_n(x)\,d\mu = g^{(n)}(\theta_0). \tag{2.5}$$

Hence if $\{\varphi_n(x)\}$ is the complete orthonormal set obtained from $\{\psi_n(x)\}$ then (h, φ_n) for any n is a linear combination of the derivatives of $g(\theta)$ and since, by the above remarks, the projection onto V is the U.M.V.U.E., t_g, of $g(\theta)$ the latter is given by

$$t_g = \text{l.i.m.} \textstyle\sum_{n=0}^{\infty} (h, \varphi_n)\varphi_n \tag{2.6}$$

with variance

$$V(t_g) = \textstyle\sum_{n=1}^{\infty} (h, \varphi_n)^2 \tag{2.7}$$

since $(h, \varphi_0) = g(\theta_0)$. Clearly then the expressions in (2.6) and (2.7) may be obtained without explicit knowledge of any unbiased estimator, $h(x)$, of $g(\theta)$. Further, the convergence of (2.7) for all $\theta\in\Omega$ is necessary and sufficient for the estimability of g.

In the next section we will use the results obtained by Seth [18] and display more explicitly the series (2.6) and (2.7) for the case where the density belongs to a subclass of the exponential family which includes the normal, gamma, exponential, Poisson, binomial and negative binomial distributions.

3. Pointwise convergence. Consider the subclass of exponential densities satisfying, for θ a strictly monotone analytic function of the natural parameter,

$$n^{-1}s(x)-\theta = K(\theta)\psi_1(x;\theta) \tag{3.1}$$

where $K(\theta)$ may depend only on θ and is a quadratic in θ, that is

$$K(\theta) = a\theta^2+b\theta+c, \tag{3.2}$$

a, b, c real constants, as well as the regularity conditions on page 21 of [18].

Seth [18] has shown that for this subclass $\{\psi_n\}$ is a set of orthogonal polynomails in ψ_1 (and hence in $s(x)$) with respect the $p_\theta(x)$. He also obtained the following results.

$$\psi_{n+1} = (\psi_1-A_n)\psi_n-B_n\psi_{n-1} \tag{3.3}$$

where

$$A_n = nK'(\theta)/K \tag{3.4}$$

$$B_n = n(n-1)K''(\theta)/2K+n/K \tag{3.5}$$

and

$$||\psi_n||_\theta^2 = \textstyle\prod_{i=1}^n B_i. \tag{3.6}$$

For this subclass we have the following theorem.

THEOREM 3.1. *For the subclass of the exponential densities satisfying* (3.1) *and* (3.2), *any estimable parametric function* $g(\theta)$ *possesses a unique* U.M.V.U.E., t_g, *given by*

$$t_g = \text{l.i.m.} \textstyle\sum_{n=0}^\infty g^{(n)}(\theta)f^{(n)}(x;\theta)/||\psi_n||^2 f(x;\theta) \tag{3.7}$$

with variance

$$V(t_g) = \textstyle\sum_{n=1}^\infty [g^{(n)}(\theta)]^2/||\psi_n||^2 < \infty. \tag{3.8}$$

PROOF. Follows from (2.6), (2.7) and the fact that for this subclass $\varphi_n(x) = \psi_n(x)/||\psi_n(x)||$.

If in (3.2) $a<0$ then $\psi_n(x) = 0$ a.e. ($\mathscr{P}$) for all $n>N$ where N is the greatest integer less than $(a-1)/a$. In this case the series has only a finite number of terms and is therefore convergent. It appears from the construction that in such a case no $g(\theta)$ having non-vanishing derivatives of higher order than N is estimable. An example of this occurs with the binomial distribution where

$$K(\theta) = \theta(1-\theta)/N \tag{3.9}$$

and N is the number of trials. It is easily seen that

$$\psi_n(x) = 0 \qquad \text{for} \quad n>N. \tag{3.10}$$

Hence only polynomials in θ of degree $\leqq N$ may be estimable. This well-known result is also given in DeGroot [7].

Suppose now that, $a > 0$. Let

(3.11) $$d = 1/a.$$ Then

(3.12) $$\|\psi_n\|^2 = n!(n+d-1)!/(d-1)!(dK)^n.$$

And a sufficient condition for the pointwise convergence of the series (3.7) is that

(3.13) $$\{(d-1)!\}^{\frac{1}{2}} \sum_{n=0}^{\infty} |g^{(n)}(\theta)| (dK)^{n/2}/\{n!(n+d-1)!\}^{\frac{1}{2}} < \infty.$$

However, (3.13) holds if N is large enough to ensure $(dK)^{\frac{1}{2}} < \delta_0$ where δ_0 is the radius of convergence of the Taylor series expansion of g at θ; in particular, (3.13) is guaranteed of convergence for all θ and N when Ω is the entire line.

Finally consider the case $a = 0$. It follows from Loève [14] that the series (3.7) converges pointwise provided

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

since in this case

(3.15) $$\|\psi_n\|^2 = n!K^{-n}.$$

However from (3.8), $\sum_{n=0}^{\infty} [g^{(n)}(\theta)]^2 K^n/n! < \infty$. Hence (3.14) is also convergent.

4. Applications and some related results. It is clear from the discussion in Bahadur [4] on Bhattacharyya bounds that the variance of the U.M.V.U.E. given here equals the limiting Bhattacharyya bound. We will illustrate this with the negative binomial

(4.1) $$P(X = x) = pq^x, \quad 0 < p < 1, \quad q = 1-p, \quad x = 0, 1, 2, \cdots.$$

Taking p as the parameter, Murty [15] considered the Bhattacharyya bounds for estimating $g(p) = p$. He obtained the kth bound as

(4.2) $$L_k = p^2 q(q^{k-1} + q^{k-2} + \cdots + 1).$$ Hence

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

which will now be shown to be the variance of the U.M.V.U.E. using (3.8).

With the parameterization $\theta = p^{-1}$ we find

(4.4) $$P_\theta(X = x) = (\theta - 1)^x \theta^{-(x+1)}.$$ Hence

(4.5) $$(x+1) - \theta = K(\theta)\psi_1(x)$$ where

(4.6) $$K(\theta) = \theta(\theta - 1).$$

Thus from (3.12)

(4.7) $$\|\psi_n\|_\theta^2 = \{\theta(\theta-1)\}^{-n}(n!)^2$$ and

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

Thus, $g^{(n)}(\theta) = (-1)^n n!\theta^{-(n+1)}$. Hence from (3.8)

$$V(t_g) = \sum_{n=1}^{\infty} (n!)^2\theta^{-2(n+1)}/(n!)^2\{\theta(\theta-1)\}^{-n}$$
$$= \theta^{-2}\sum_{n=1}^{\infty}\{(\theta-1)/\theta\}^n = \theta^{-2}(\theta-1) = pq.$$

DeGroot [7] has considered the more general problem

$$P_p(X = x+c) = \binom{x+c-1}{x}p^c q^x, \qquad x = 0, 1, 2, \cdots.$$

Using the above reparameterization we obtain

$$\frac{x+c}{c} - \theta = \frac{\theta(\theta-1)}{c}\psi_1$$

and $$\|\psi_n\|^2 = \{\theta(\theta-1)\}^{-n}n!(n+c-1)!/(c-1)!.$$

Hence, with $g(\theta) = \theta^{-1}$, (3.8) yields $V(t_g) = \theta^{-2}\sum_{n=1}^{\infty}\{(\theta-1)\theta^{-1}\}/\binom{n+c-1}{n}$. Again, under the conditions of Theorem 4.1 in [7], the series representation of any estimable $g(\theta)$ converges pointwise a.e. μ. In fact the necessary and sufficient condition for estimability given in that theorem is the condition required for the convergence of the series.

Fend [9] and Rao [15] have shown that if the mth Bhattacharyya bound is achieved when estimating θ, then the U.M.V.U.E. for θ is a polynomial of degree m in the sufficient statistic. Conversely if the density function satisfies conditions (3.1) and (3.2) then any estimable polynomial in θ of degree m has a U.M.V.U.E. which is a polynomial of the same degree in the sufficient statistic. We will illustrate this with the following Example 1 of Fend [9].

Consider $f(x;\alpha) = \alpha^{-(1/n)}\exp[-x\alpha^{-(1/n)}]$ $0 < x$, $0 < \alpha$. Let $\theta = \sigma^{1/n}$. Fend [9] gives the U.M.V.U.E. of $\alpha = \theta^n$ as $x^n/n!$. We will verify this result and give the variance using the method developed here.

Thus consider estimating $g(\theta) = \theta^n$ with density

$$f(x;\theta) = \theta^{-1}\exp -x\theta^{-1} \qquad 0 < x, \quad 0 < \theta. \tag{4.9}$$

It can be verified that the density (4.9) belongs to the subclass of the exponential family considered in Section 3 and that

$$\psi_1\psi_j = \psi_{j+1} + (2j/\theta)\psi_j + (j/\theta)^2\psi_{j-1} \tag{4.10}$$

$$\|\psi_j\|^2 = (j!)^2/\alpha^{2j}. \tag{4.11}$$

Hence from the theorem in Section 3 with $g(\theta) = \theta^n$

$$t_g = \sum_{j=0}^{n}\binom{n}{j}\theta^{n+j}\psi_j/j! \qquad \text{and} \tag{4.12}$$

$$V(t_g) = \theta^{2n}\sum_{j=1}^{n}\binom{n}{j}^2 = \theta^{2n}\left[\binom{2n}{n}-1\right].$$

It can also be verified by induction that $t_g = x^n/n!$ which is the result given by Fend [9].

Fend also gives the following Example 2 to show that the kth Bhattacharyya bound (k finite) may be unachievable. Consider $f(x;\alpha) = \alpha^{-n}\exp[-x\alpha^{-n}]$ $0 < x$,

$0 < \alpha$. Let $\theta = \alpha^n$. Then $f(x;\theta) = \theta^{-1}\exp[-x\theta^{-1}]\ 0 < x,\ 0 < \theta$ with $g(\theta) = \theta^{1/n}$. Since $g^{(i)}(\theta)$ does not vanish it follows from the results here that if $g(\theta)$ is estimable, the U.M.V.U.E. cannot be a finite polynomial and hence no finite Bhattacharyya is achievable.

Acknowledgment. We have benefited from conversations with Barry C. Arnold and B. K. Kale.

REFERENCES

[1] ABBEY, JOSEPH L. (1967). Series representation of uniformly minimum variance unbiased estimators. Unpublished M.S. Thesis, Library, Iowa State Univ.

[2] AKHIEZER, N. I. (1965). *The Classical Moment Problem* (1st ed.). Oliver and Boyd, London.

[3] BAHADUR, R. R. (1957). Measurable subspaces and subalgebras. *Proc. Amer. Math. Soc.* **6** 565–570.

[4] BAHADUR, R. R. (1957). On unbiased estimation of uniformly minimum variance. *Sankhyā* **18** 211–224.

[5] BARANKIN, E. W. (1949). Locally best unbiased estimates. *Ann. Math. Statist.* **20** 477–501.

[6] CHENEY, E. W. (1966). *Introduction to Approximation Theory*. McGraw-Hill, New York.

[7] DEGROOT, M. H. (1959). Unbiased sequential estimation for binomial populations. *Ann. Math. Statist.* **30** 80–101.

[8] FELLER, WILLIAM (1966). *An Introduction to Probability Theory and its Applications*, **2** (1st ed.). Wiley, New York.

[9] FEND, A. V. (1959). On the attainment of Cramér–Rao and Bhattacharyya bounds for the variance of an estimate. *Ann. Math. Statist.* **30** 381–388.

[10] HALMOS, P. R. (1950). *Measure Theory*. Van Nostrand, New York.

[11] KENDALL, M. G. (1940). Conditions for uniqueness in the problem of moments. *Ann. Math. Statist.* **11** 402–409.

[12] KENDALL, M. G. (1941). Corrections to a paper on the uniqueness problem of moments. *Ann. Math. Statist.* **12** 464–465.

[13] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

[14] LOÈVE, M. (1955). *Probability Theory*. Van Nostrand, New York.

[15] MURTY, V. N. (1956). A note on Bhattacharyya bounds for the negative binomial distribution. *Ann. Math. Statist.* **27** 1182–1183.

[16] RAO, C. R. (1952). Some theorems on minimum variance estimation. *Sankhyā* **12** 27–42.

[17] RIESZ, F. and SZ-NAGY, B. (1955). *Functional Analysis*. Frederick Ungar, New York.

[18] SETH, G. R. (1949). On the variance of estimates. *Ann. Math. Statist.* **20** 1–27.

[19] STEIN, C. M. (1950). Unbiased estimates with minimum variance. *Ann. Math. Statist.* **21** 406–415.