Ann. Inst. Statist. Math.
Vol. 47, No. 1, 99–104 (1995)

A NOTE ON ACCELERATED SEQUENTIAL ESTIMATION OF THE MEAN OF NEF-PVF DISTRIBUTIONS

ARUP BOSE[1] AND NITIS MUKHOPADHYAY[2]

[1] *Stat-Math Unit, Indian Statistical Institute, 203 Barrackpore Trunk Road, Calcutta 700 035, India*
[2] *Department of Statistics, University of Connecticut, UBox 120, 196 Auditorium Road, Storrs, CT 06269, U.S.A.*

(Received December 20, 1993)

Abstract. The minimum risk point estimation for the mean is addressed for a natural exponential family (NEF) that also has a power variance function (PVF) under a loss function given by the squared error plus linear cost. An appropriate accelerated version of the full purely sequential methodology of Bose and Boukai (1993*b*, submitted) is proposed along the lines of Mukhopadhyay (1993*a*, Tech. Report, No. 93-27, Department of Statistics, University of Connecticut) in order to achieve operational savings. The main result provides the asymptotic second-order expansion of the regret function associated with the accelerated sequential estimator of the population mean.

Key words and phrases: Natural exponential family, power variance function, mean estimation, minimum risk, acceleration, regret expansion, operational savings.

1. Introduction

Let $\mathcal{F} = \{F_\theta : \theta \in \Theta\}$ be the class of natural exponential family (NEF) of distributions. That is, $\mathcal{F}$ is minimal NEF of order 1 whose members are of the form

$$F_\theta(dx) = \exp\{\theta x + c(\theta)\}\Delta(dx), \qquad \theta \in \Theta \tag{1.1}$$

where Δ is a sigma-finite measure on the Borel sets of R and the parameter space Θ consists of all $\theta \in R$ for which $\int \exp(\theta x)\Delta(dx)$ is finite. It is well-known (see Barndorff-Nielsen (1978)) that F_θ has finite moments of all order. For $\theta \in \Theta$, we let $\mu = \mu(\theta) = -dc(\theta)/d\theta$ and $\Omega = \mu(\operatorname{int}\Theta)$ respectively denote the mean value of F_θ and the mean parameter space. We also write $V(\mu)$ for the variance function corresponding to (1.1). Let us suppose that the members of the NEF $\mathcal{F}$ have a power variance function (PVF) so that one can write

$$V(\mu) = \alpha\mu^\gamma, \qquad \mu \in \Omega \tag{1.2}$$

for some *known constants* $\alpha \neq 0$ and γ. In what follows, we will write

$$\nu_0^2(\mu) = V(\mu)/\mu^2 = \alpha\mu^{\gamma-2}, \tag{1.3}$$

for the square of the coefficient of variation of F_θ.

This class of NEF-PVF is known to be a rich family that possesses many interesting properties. Bar-Lev and Enis (1986) have shown, aside from a reproducibility property, that all NEF-PVF's are infinitely divisible with self generating property. It is convenient to identify the members of $\mathcal{F}$ by their γ values. The characterization of NEF by means of their variance function was initiated by Morris (1982) where he considered a subclass of NEF having a quadratic variance function (QVF) and showed that the NEF-QVF class consists of only six members, some of which are also members of NEF-PVF class. From Bar-Lev and Enis (1986), it is clear that the NEF-PVF family consists of many interesting distributions. See also Morris (1983) for other related details. Here we focus only on *non-lattice members* from the NEF-PVF family for which $\gamma \neq 0$. In this case, it is known that either $\Omega = R^-$ or $\Omega = R^+$. We will assume without any loss of generality that the support of (1.1) is $[0, \infty)$ and that $\Omega = R^+$ with $\gamma > 1$ in which case α is obviously positive. For a comprehensive discussion, see Bar-Lev and Enis (1986).

Bose and Boukai (1993*b*) considered point estimation problems for μ via $\bar{X}_n(= n^{-1}\sum_{i=1}^n X_i)$ under the loss function

$$L_n = A(\bar{X}_n - \mu)^2 + n \tag{1.4}$$

where $X_1, X_2, \ldots$ are i.i.d. having a distribution $F_\theta \in \mathcal{F}$, defined in (1.1). The associated risk is $R_n(A) = E(L_n) = A\sigma^2 n^{-1} + n$ where $\sigma^2 = \sigma^2(\theta)$ is the variance of the distribution F_θ. The fixed-sample size risk is minimized if $n = n^* = A^{1/2}\sigma$. In the situation when θ is unknown, Bose and Boukai (1993*b*) proposed a purely sequential stopping time N and the corresponding estimator $\bar{X}_N$ for μ along with the second-order approximation of $E(L_N) - R_{n*}(A)$, that is the *regret function*. In order to review various aspects of sequential estimation, one may refer to Woodroofe (1977, 1982), Sen (1981), Martinsek (1983), Mukhopadhyay (1988, 1991), and Bose and Boukai (1993*a*, *b*), among others.

It is, however, well-known that one by one purely sequential sampling schemes are operationally inconvenient. Hall (1983) came up with an acceleration technique of the original purely sequential process in the context of obtaining a fixed-width confidence interval for the mean of a normal distribution having an unknown variance, and that particular idea was later generalized in Mukhopadhyay and Solanky (1991). In the present context, for the Bose-Boukai (1993*b*) stopping rule though, $I(N = n)$ and $\bar{X}_n$ turns out to be dependent for all fixed $n \geq 1$, and hence the acceleration techniques of Hall (1983) and Mukhopadhyay and Solanky (1991) do not hold much promise. On the other hand, in Section 2 we devise an appropriate accelerated version of the Bose-Boukai (1993*b*) procedure along the lines of recent modifications from Mukhopadhyay (1993*a*, *b*) and provide the second-order expansion of the associated risk function in Theorem 2.1. We provide a brief justification of Theorem 2.1 in Section 3.

2. The acceleration technique

Under the loss function (1.4), recall that $n^* = A^{1/2}\sigma$ where $\sigma^2 = \alpha\mu^\gamma$. The purely sequential sampling process of Bose and Boukai (1993*b*) goes like this:

One starts with $X_1, \ldots, X_m$ for $m \geq 1$ and then proceeds by taking one additional observation at a time according to the stopping rule

$$N = \inf\{n \geq m : n^2 a(n) \geq A\alpha \bar{X}_n^\gamma\}, \tag{2.1}$$

where $a(n)$ is positive, nonincreasing and $a(n) = 1 + a_0 n^{-1} + o(n^{-1})$ as $n \to \infty$, $a_0 \in R$. Since $P(N < \infty) = 1$, one then estimates μ by means of $\bar{X}_N$. Observe that N asymptotically gets near n^*, in a sense to be made specific later, in approximately n^* steps. On the other hand, we wish to accelerate the process by finally arriving at the same n^* via sequential sampling part of the way first, followed by batch sampling in one single step. Such a modification will make the sampling methdology operationally whole lot more convenient as well as attractive.

First choose and fix $\rho \in (0,1)$ *such that* ρ^{-1} *is an integer* and start sampling with $X_1, \ldots, X_m$. Define

$$\begin{aligned} t &= \inf\{n \geq m(\geq 1) : n^2 a(n) \geq A\rho^2 \alpha \bar{X}_n^\gamma\}, \\ N &= \rho^{-1} t. \end{aligned} \tag{2.2}$$

Observe that t estimates ρn^*, a fraction of n^*, and then one determines N. At this step, one samples the difference $(N - t)$, namely $(1-\rho)\rho^{-1}t$, in one single batch. In other words, $X_1, \ldots, X_m, \ldots, X_t$ is augmented by $X_{t+1}, \ldots, X_N$, all *in one single batch*, and μ is finally estimated by the corresponding $\bar{X}_N$. One should note that the accelerated sequential estimation procedure (2.2) saves approximately $100(1-\rho)\%$ of sampling operations compared with the full purely sequential scheme of Bose and Boukai (1993*b*), given by (2.1). The following is an immediate consequence of Lemma 1 in Bose and Boukai (1993*b*).

LEMMA 2.1. *Let* $s \geq 1$ *be fixed. Then as* $A \to \infty$, *we have for* $0 < \varepsilon < \rho$, $n^{*s} P\{m \leq N \leq \varepsilon n^*\} \to 0$ *provided that one of the following holds*:

a) $\gamma > 2$ *and* $m \geq 1$;

b) $\gamma = 2$ *and* $m > s/\alpha$;

c) $\gamma < 2$ *and* $m > (1+s)(2-\gamma)\log n^*$.

Here, N *comes from* (2.2).

For $1 < \gamma < 2$, $\mathcal{F}$ is a family of compound Poisson generated by a random sum of gamma variates. These distributions have positive probability mass at zero. See Bar-Lev and Enis (1986) for details. Condition (c) of Lemma 2.1 indicates that in order to handle the situation $1 < \gamma < 2$, one additionally needs an appropriate "growth" condition on the starting sample size m. The following results can be derived from Bose and Boukai (1993*b*) as $A \to \infty$:

$$(t - \rho n^*)/(\rho n^*)^{1/2} \xrightarrow{\mathcal{L}} N(0, \tau^2) \quad \text{if} \quad m \geq 1; \tag{2.3}$$

(2.4) $(t-\rho n^*)^2/(\rho n^*)$ is uniformly integrable under the conditions of Lemma 2.1 with $s=1$;

$$E(t)=\rho n^*+\frac{\nu\gamma}{2\mu}-\frac{1}{2}a_0-\frac{1}{8}\gamma(\gamma+2)\nu_0^2(\mu)+o(1) \tag{2.5}$$

under the conditions of Lemma 2.1 with $s=1$;

where $\tau^2=(\delta\mu)^{-2}V(\mu)=\nu_0^2(\mu)/\delta^2$ and

$$\nu=\frac{1}{2}\frac{[\delta^2\mu^2+V(\mu)]}{\delta\mu}-\sum_{n=1}^{\infty}n^{-1}E[(S_n-(1+\delta)n\mu)^+] \tag{2.6}$$

with $S_n=\sum_{i=1}^n X_i$, $x^+=\max(0,x)$, $\delta=2/\gamma$.

The results (2.3)–(2.5) merely verify the Assumptions A1–A3 of Mukhopadhyay (1993*a*). Hence, the following proposition easily follows from Theorem 2.1 of Mukhopadhyay (1993*a*).

PROPOSITION 2.1. *For the accelerated stopping time N given by* (2.2), *we have as $A\to\infty$:*

(i) $n^{*-1/2}(N-n^*)\xrightarrow{\mathcal{L}}N(0,\tau^2/\rho)$ *if* $m\geq 1$;

(ii) *Under the conditions of Lemma* 2.1 *with* $s=1$,

a) $n^{*-1}(N-n^*)^2$ *is uniformly integrable*;

b) $E(N)=n^*+\rho^{-1}\{\frac{\nu\gamma}{2\mu}-\frac{1}{2}a_0-\frac{1}{8}\gamma(\gamma+2)\nu_0^2(\mu)\}+o(1)$.

2.1 *The main result*

The risk associated with $\bar{X}_N$ under the loss function (1.4) and the accelerated stopping time (2.2) is given by

$$R(A)=E(L_N)=AE[(\bar{X}_N-\mu)^2]+E(N). \tag{2.7}$$

On the other hand, the optimal fixed sample size risk is given by $R_{n*}(A)$ where $n^*=\{A^2V(\mu)\}^{1/2}$, and one defines the regret

$$w(A)=R(A)-R_{n*}(A). \tag{2.8}$$

The following result provides the asymptotic second-order expansion of $w(A)$.

THEOREM 2.1. *Under the conditions of Lemma* 2.1 *with* $s=2+\varepsilon$ *with arbitrary* $\varepsilon>0$, *for the accelerated sequential procedure* (2.2), *we have as* $A\to\infty$:

$$w(A)=\nu_0^2(\mu)\left\{\frac{1}{4}\gamma(\gamma+4)+(1-\rho)(\rho\delta^2)^{-1}\right\}+o(1)$$

where $\nu_0^2=V(\mu)/\mu^2$ *and* $\delta=2/\gamma$.

When $\rho = 1$, the regret expansion provided in Theorem 2.1 reduces to that in Bose and Boukai (1993*b*) as expected. Let us denote the Bose and Boukai (1993*b*) regret as $w^*(A)$. Write $\rho = 1/k$ where k is an integer. Then, asymptotically one has

$$w(A) = w^*(A) + (k-1)\delta^{-2}\nu_0^2(\mu) + o(1). \tag{2.9}$$

In this case, the accelerated sequential procedure (2.2) saves about $100(k-1)k^{-1}\%$ sampling operations when compared with the full purely sequential scheme of Bose and Boukai (1993*b*), but this is achieved at the expense of the increase in the regret function by the amount $(k-1)\delta^{-2}\nu_0^2(\mu)$ up to the order $o(1)$. Given some prior knowledge about the coefficient of variation $\nu_0^2(\mu)$, the experimenter would then "balance" the operational savings on the face of some regret increment. In the cited references one will however notice that the operational savings obtained through acceleration quite often outweights the slight increase in the regret function associated with it.

3. Proof of Theorem 2.1

Along the lines of (4.4) in Mukhopadhyay (1993*a*) one obtains

$$E\{A(\bar{X}_N - \mu)^2\} = A\rho^2 E[(\bar{X}_t - \mu)^2] + A\rho(1-\rho)V(\mu)E(t^{-1}). \tag{3.1}$$

In other words, one has with $n_1^* = \rho n^*$,

$$\begin{aligned} R(A) &= A\rho^2 E[(\bar{X}_t - \mu)^2] + E(t) \\ &\quad + (1-\rho)n^*\{E(n_1^*/t) + E(t/n_1^*)\} \\ &= E[L_t^*] + \rho^{-1}(1-\rho)E\{(t-n_1^*)^2/t\} + 2(1-\rho)n^*, \end{aligned} \tag{3.2}$$

where

$$L_n^* = A\rho^2(\bar{X}_n - \mu)^2 + n. \tag{3.3}$$

Suppose now that one pretends to obtain a minimum risk point estimator of μ via $\bar{X}_n$ under the pseudo loss function (3.3). Under this framework, the "optimal" fixed sample size is indeed $n_1^* = \rho A^{1/2}\sigma$ and hence the corresponding purely sequential rule of Bose and Boukai (1993*b*) will coincide with t given in (2.2). Thus, under appropriate assumed conditions, one immediately obtains the following from Theorem 2 of Bose and Boukai (1993*b*):

$$E(L_t^*) = 2n_1^* - a_0 + \frac{1}{4}\gamma(\gamma+4)\nu_0^2(\mu) + o(1). \tag{3.4}$$

Combining (2.3)–(2.4) and Lemma 2.1, one also obtains

$$E\{(t-n_1^*)^2/t\} = \tau^2 + o(1). \tag{3.5}$$

Hence from (3.2), (3.4) and (3.5) we get

$$\begin{aligned} R(A) &= 2\rho n^* - a_0 + \frac{1}{4}\gamma(\gamma+4)\nu_0^2(\mu) + \rho^{-1}(1-\rho)\tau^2 + 2(1-\rho)n^* \\ &= 2n^* - a_0 + \nu_0^2(\mu)\left\{\frac{1}{4}\gamma(\gamma+4) + (1-\rho)(\rho\delta^2)^{-1}\right\} + o(1). \end{aligned}$$

The proof is now complete.

REFERENCES

Bar-Lev, S. K. and Enis, P. (1986). Reproducibility and natural exponential families with power variance function, *Ann. Statist.*, **14**, 1507–1522.

Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory*, Wiley, New York.

Bose, A. and Boukai, B. (1993*a*). Sequential estimation results for a two-parameter exponential family of distributions, *Ann. Statist.*, **21**, 484–502.

Bose, A. and Boukai, B. (1993*b*). Sequential estimation of the mean of NEF-PVF distributions (submitted).

Hall, P. (1983). Sequential estimation saving sampling operations, *J. Roy. Statist. Soc. Ser. B*, **45**, 219–223.

Martinsek, A. T. (1983). Second order approximations to the risk of a sequential procedure, *Ann. Statist.*, **11**, 827–836.

Morris, C. N. (1982). Natural exponential families with quadratic variance functions, *Ann. Statist.*, **10**, 65–80.

Morris, C. N. (1983). Natural exponential families with quadratic variance functions: Statistical theory, *Ann. Statist.*, **11**, 515–529.

Mukhopadhyay, N. (1988). Sequential estimation problems for negative exponential populations, *Comm. Statist.* (Reviews Section), *Theory Methods*, **17**, 2471–2506.

Mukhopadhyay, N. (1991). Parametric sequential point estimation, Chapter 10, *Handbook of Sequential Analysis* (ed. B. K. Ghosh and P. K. Sen), 245–267, Marcel Dekker, New York.

Mukhopadhyay, N. (1993*a*). An alternative formulation of accelerated sequential procedures with applications, Tech. Report, No. 93-27, Department of Statistics, University of Connecticut, Storrs.

Mukhopadhyay, N. (1993*b*). On accelerated sequential point estimation of means when the distributions are unspecified, Tech. Report, No. 93-28, Department of Statistics, University of Connecticut, Storrs.

Mukhopadhyay, N. and Solanky, T. K. S. (1991). Second order properties of accelerated stopping times with applications in sequential estimation, *Sequential Anal.*, **10**, 99–123.

Sen, P. K. (1981). *Sequential Nonparametrics*, Wiley, New York.

Woodroofe, M. (1977). Second order approximations for sequential point and interval estimation, *Ann. Statist.*, **5**, 985–995.

Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis*, Society for Industrial and Applied Mathematics, Philadelphia.