Sequential Estimation Procedure for the Two-Parameter Exponential Family of Distributions

by

Arup Bose
Purdue University

Benzion Boukai
Purdue University

Technical Report #89-34

Department of Statistics
Purdue University

1989

A SEQUENTIAL ESTIMATION PROCEDURE FOR A TWO-PARAMETER EXPONENTIAL FAMILY OF DISTRIBUTIONS; FIRST ORDER RESULTS

by
Arup Bose and Benzion Boukai

Department of Statistics
Purdue University
West Lafayette, Indiana 47907

Key words and phrases: Stopping rule; Nuisance parameter; Risk function; First order asymptotic.

AMS 1985 Subject classifications: Primary 62L12; Secondary 62E99.

ABSTRACT

We consider the problem of sequentially estimating one parameter in a class of two-parameter exponential family of distributions. We assume a squared error loss with a fixed cost of estimation error. The stopping rule, based on the maximum likelihood estimate of the nuisance parameter, is shown to be independent of the terminal estimate. The first order asymptotic properties of the risk function are investigated. It is shown that the suggested procedure is an asymptotically *risk efficient* procedure. This procedure is exemplified for the normal, gamma and the inverse Gaussian densities, which follow as special cases of our general results.

1. Introduction. Consider a model in which the underlying distribution of a sequence of ($i.i.d.$) random variables $X_1, X_2, \ldots\ldots$ depends on two unknown parameters θ_1 and θ_2. For a fixed (and finite) sample size, measures of accuracy of an estimate for the parameter of interest θ_2, say, typically depend on the unknown value of the nuisance parameter θ_1. Thus to achieve a given level of accuracy on has to proceed sequentially: to determine the final (random) sample size based on an estimate of θ_1, and then the terminal estimate of θ_2 is determined based on this sample.

Procedures of this nature were discussed initially by Stein (1945, 1949), as two-stage procedures for estimation and interval estimation of prescribed accuracy, for the mean of normally distributed r.v.'s , when the variance σ^2 is unknown. Later, this problem (for the normal mean), was tackled by purely sequential procedures proposed by Robbins (1959), (for point estimation) and Chow and Robbins (1965), (for fixed width interval estimation). Although the normal case has been studied extensively, there are, to the best of our knowledge, only few studies (cited below), dealing with other cases of underlying distributions.

To illustrate the general problem on hand, consider the following point estimation problem. Let $X_1, X_2, \ldots\ldots$, be i.i.d. random variables with unknown mean μ (the parameter of interest) and variance σ^2. Having recorded the first n observations $x_1, \ldots, x_n$, let the loss incurred in estimating μ by $\hat{\mu}_n = \sum x_i/n$ be:

$$L_\rho(\hat{\mu}_n) = \rho(\hat{\mu}_n - \mu)^2 + n \tag{1.1}$$

where ρ ,(> 0), is the known weight of the estimation error relative to the sampling cost. The objective is to minimize the associated risk;

$$R_\rho(n) = E_{\mu,\sigma}(L_\rho(\hat{\mu}_n)) = \frac{\rho\sigma^2}{n} + n\ . \tag{1.2}$$

with respect to the sample size n. When σ is known, the expected loss, $R_\rho(n)$ is minimized by taking a sample of size (an integer adjacent to) $n_0 = \rho^{\frac{1}{2}}\sigma$, with corresponding risk, $R_\rho(n_0) = 2n_0$. However, when σ is an unknown nuisance parameter, no fixed sample size minimizes $R_\rho(n)$ simultaneously for all σ. This motivates the following choice of a random sample size N_ρ, when σ^2 is unknown,

$$N_\rho = \inf\{\ n \geq m_0,\ \hat{\sigma}_n^2 \leq a_n n^2/\rho\} \tag{1.3}$$

where m_0 (≥ 2) is the initial sample size, $\hat{\sigma}_n^2$ is a suitable estimate of σ^2, (for example; $\hat{\sigma}_n^2 = s_n^2 \equiv \sum(x_i - \bar{x})^2/(n-1)$), and a_n is some nonincreasing sequence, ($a_n \to 1$ as

$n \to \infty$). According to this procedure, the parameter μ is estimated *at termination*, by $\hat{\mu}_{N_\rho}$. Clearly then, the study of the stopping variable N_ρ and the *risk* associated with it, become important.

For normal random variables, the statistic s_n^2 is ancillary to μ and it can be shown (by using Helmert orthogonal transformation), that the event $\{N_\rho = n\}$ and $\hat{\mu}_n$ are independent. This property was heavily exploited by most researcher who worked on the normal problem. Robbins (1959) studied the properties of N_ρ and provided a recursive formula for its distribution. Later this problem was studied extensively by Starr (1966), and Starr and Woodroofe (1969). Woodroofe (1977) has used second order approximations to study this procedure and to analyze the *regret* in the expected loss incurred upon using the sample size N_ρ as compared to n_0.

Extensions of this procedure to nonnormal cases were considered by several authors. Starr and Woodroofe (1972) deal with the negative exponential distribution and provide results on the regret. Here, even though there is one parameter, the problem becomes interesting since the variance of the m.l.e. depends on the unknown parameter. Ghosh and Mukhopadhyay (1979) with a 'distribution free' approach allowed the initial sample size m_0 to be a function of ρ and to $\to \infty$ as $\rho \to \infty$. They proved first order result for the risk, showing that the ratio of the risk associated with N_ρ to that associated with the hypothetical fixed sample size n_0 converges to 1, as $\rho \to \infty$. Mukhopadhyay (1988) surveyed results concerning sequential estimation procedures for the negative exponential distribution, with and without a truncation parameter. Related studies are those of Aras (1987, 1989), dealing with sequential estimation procedure based on censored data from negative exponential distribution. He provided first and second order results, also by allowing the initial sample size $m_0 \to \infty$.

In the present paper, we consider a sequential point estimation problem in a class of two-parameter exponential family of distributions. The model considered here will be restricted by assumptions on its natural parameters (θ_1, θ_2), but is general enough to include the normal, the gamma, and the inverse Gaussian distributions, as special cases. This exponential subfamily was first introduced by Bar-Lev and Rieser (1982) in context of UMPU tests based on single test statistics. A description and the basic properties of such an exponential subfamily are provided in section 2. We also present a new independence result, analogous to the one discussed for the normal case, which provides in the general case discussed, the independence of the event $\{N_\rho = n\}$ with the terminal estimator. Finally, in Section 3 we present an appropriate stopping rule along with the estimation procedure for the parameter of interest. We then discuss the properties of the suggested stopping rule and provide the first order properties of its associated risk as $\rho \to \infty$, under a loss function similar to (1.1) (and to (1) of Woodroofe (1985)).

2. Preliminaries; the exponential family and an independence result.

Let

$$f(x;\theta) = a(x)exp\{\theta_1 U_1(x) + \theta_2 U_2(x) + c(\theta)\}, \quad \theta = (\theta_1, \theta_2), \tag{2.1}$$

be a density function (w.r.t. Lebesgue measure on **R**), which characterizes a *regular* two-parameter exponential family of distributions, (see Brown (1986)), i.e.; the natural parameter space Θ is defined by;

$$\Theta = \{\theta \in \mathbf{R}^2 \; ; e^{-c(\theta)} = \int a(x)exp\{\theta_1 U_1(x) + \theta_2 U_2(x)\}\, dx \; < \infty\},$$

so that $\Theta \equiv int\Theta \neq \emptyset$. It is well known that for any $\theta \in \Theta$ the r.v. $\mathbf{U} = (U_1, U_2)$ has moments of all orders. In particular, we denote;

$$E_\theta(\mathbf{U}) = (\nu_1, \nu_2), \qquad \nu_i = -\partial c(\theta)/\partial\theta_i, \qquad i = 1,2 \tag{2.2}$$

and

$$V_\theta(\mathbf{U}) = (\,\sigma_{ij}\,), \qquad \sigma_{ij} = -\partial^2 c(\theta)/\partial\theta_i\partial\theta_j \qquad i,j = 1,2,$$

where $V_\theta(\mathbf{U})$ is the corresponding covariance matrix, (positive definite).

Let $X_1, \ldots, X_n$, $n > 1$, be independent *r.v.'s* having a common density of the form (2.1). We set $T_{i:n} = \sum_{j=1}^{n} U_i(X_j)$ and denote by $\bar{T}_{i:n}$, $i = 1,2$ the usual average. The joint distribution of $\mathbf{T} = (T_{1:n}, T_{2:n})$ is a member of the two-parameter exponential family, and

$$E_\theta(\mathbf{T}) = (n\nu_1, n\nu_2), \qquad V_\theta(\mathbf{T}) = (n\sigma_{ij}) \quad i,j = 1,2\;. \tag{2.3}$$

The results stated in the following theorem were proved by Bar-Lev and Reiser (1982), and are concerned with a characterization of (2.1) which admits a single ancillary statistic for θ_2 in the presence of θ_1, (that is, its distribution depends only on θ_1). For additional applications of this result see also Brown (1986 pp. 44-48). However, that characterization requires the following two assumptions:

ASSUMPTION A.1. *The parameter* θ_2 *can be represented as;* $\theta_2 = -\theta_1\psi'(\nu_2)$, *where* $\psi'(\nu_2) = d\psi(\nu_2)/d\nu_2$, *for some function* ψ.

ASSUMPTION A.2. $U_2(x) = h(x)$, *where* $h(x)$ *is* $1-1$ *function on the support of (2.1).*

THEOREM 2.1. *(Bar-Lev and Reiser, (1982)) Under the above assumptions, the following hold:*

a) $U_1[h^{-1}(\bar{T}_{2:n})] = \psi(\bar{T}_{2:n})$ *a.s.* $,(n \geq 1)$.

b) The distribution of the statistics

$$Z_n = T_{1:n} - n\psi(\bar{T}_{2:n}), \tag{2.4}$$

belongs to the one parameter exponential family with natural parameter θ_1 *and density of the form,*

$$f_{Z_n}(z_n, \theta_1) = q(z_n)\exp\{\theta_1 z_n - H_n(\theta_1)\}, \quad \theta_1 \in \Theta_1. \tag{2.5}$$

c) For each $n \geq 2$ *and for any* $\theta \in \Theta$, *the r.v.'s* Z_n *and* $T_{2:n}$ *are independent.*

By Theorem 2.1, the statistic Z_n is ancillary to θ_2 in the presence of θ_1, and therefore may be used in fixed sample estimation procedures. However, in the context of sequential estimation, we need a result stronger than that of part (c) of Theorem 2.1. This is given below.

THEOREM 2.2. *Under the above assumptions, for all* $n \geq 2$ *and* $\theta \in \Theta$, *the random variables* $(Z_2, \ldots, Z_n)$ *are jointly independent of* $T_{2:n}$, *i.e.;*

$$(Z_2, \ldots, Z_n) \perp T_{2:n} \ .$$

Since the proof of this theorem is rather technical, it is deferred to the Appendix. As was mentioned in Section 1, the result of Theorem 2.2, will enable us to obtain an independence property analogous to the one discussed for the normal case, and thus is of great importance in context of sequential estimation. In light of this, *we assume from now on that the two conditions; A.1 and A.2, hold without further reference.*

Using (2.2), one can introduce a parameterization of the exponential family by means of the mapping $(\theta_1, \theta_2) \to (\theta_1, \nu_2)$, which is a homeomorphism, and has its components θ_1, ν_2 varying independently, (see Barndoff-Nielsen (1978), Theorem 8.4). Accordingly, $(\theta_1, \nu_2) \in \Theta_1 \times \mathcal{N}_2$ where $\mathcal{N}_2$ is connected and open. With such parameterization, and under the above assumptions, the following results can be easily shown to hold, (see Bar-Lev and Reiser (1982)).

LEMMA 2.1.

a) $\psi'(\nu_2)$ *is not identically constant.*

b) The variance of U_2 *is given by:*

$$\sigma_{22}(\theta) \equiv \frac{\partial \nu_2}{\partial \theta_2} = \frac{-1}{\theta_1 \psi''(\nu_2)}, (> 0), \tag{2.6}$$

c) The functions $c(\theta)$ *and* $\nu_1(\theta)$ *when expressed by means of the mixed parameters* θ_1 *and* ν_2, *have the following form:*

$$\begin{cases} c(\theta_1, \nu_2) = \theta_1[\nu_2\psi'(\nu_2) - \psi(\nu_2)] - G(\theta_1) \\ \\ \nu_1 = \psi(\nu_2) + G'(\theta_1) \end{cases} \tag{2.7}$$

where $G(\theta_1)$ *is an infinitely differentiable function on* Θ_1 *for which* $G''(\theta_1) > 0$, *for all* $\theta_1 \in \Theta_1$.

Here G' and G'' denote the first and second derivatives of G, respectively. In fact, using the above results it can be shown, (see Bar-Lev and Reiser (1982)), that the function H_n in (2.5) is given by:

$$H_n(\theta_1) = nG(\theta_1) - G(n\theta_1), \tag{2.8}$$

so that $E_{\theta_1}(Z_n) \equiv H_n'(\theta_1) = n(G'(\theta_1) - G'(n\theta_1))$ and $V_{\theta_1}(Z_n) \equiv H_n''(\theta_1) = nG''(\theta_1) - n^2G''(n\theta_1))$. Moreover, since $H_n(\theta_1) > 0$ and $G''(\theta_1) > 0$, it follows that $G'(\theta_1) > G'(n\theta_1)$, for all $\theta_1 \in \Theta_1$ and $n > 1$. Furthermore, parts (a)-(b) of Lemma 2.1 suggest that either $\Theta_1 \subset \mathbf{R}^-$ or $\Theta_1 \subset \mathbf{R}^+$.

LEMMA 2.2. *If* $\Theta_1 \subset \mathbf{R}^-$ *(if* $\Theta_1 \subset \mathbf{R}^+$*), then:*

a) ψ *is strictly convex (concave) function on* $\mathcal{N}_2$,

b) $Z_1 = 0$ *and* $Z_n > Z_{n-1}$ *a.s. ,* $(Z_n < Z_{n-1}$ *a.s.)* $, n \geq 2$

c) G' *is positive (negative) on* Θ_1.

Proof: We will prove only the case $\Theta \subset \mathbf{R}^-$ of the lemma, since the proof of the other case is similar. That ψ is strictly convex on $\mathcal{N}_2$, follows immediately from Lemma 2.1(a-b) and that $Z_1 = 0$ *a.s.*, follows from part (a) of Theorem 2.1. In fact, since $X_1, \ldots, X_n$ are identically distributed, we have by Theorem 2.1 (a) that $U_1(X_j) = \psi(U_2(X_j))$, *a.s.*, for all $j = 1, \ldots, n$. Now, ψ is convex and thus:

$$\psi(\bar{T}_{2:n}) < \frac{n-1}{n}\psi(\bar{T}_{2:n-1}) + \frac{1}{n}U_1(X_n) \qquad a.s.,$$

which in turn, implies that:

$$Z_n = T_{1:n} - n\psi(\bar{T}_{2:n}) > T_{1:n-1} - (n-1)\psi(\bar{T}_{2:n-1})) = Z_{n-1} \qquad a.s. .$$

Furthermore, since for $n > 1$, $Z_n > 0$, $a.s.$, it follows from (2.3), (2.7) and Jensen's inequality that for each $\theta_1 \in \Theta_1$:

$$0 < E(\bar{Z}_n) = \nu_1 - E[\psi(\bar{T}_{2:n})] < \nu_1 - \psi(\nu_2) = G'(\theta_1), \qquad \blacksquare$$

<u>LEMMA 2.3</u>. *For each* $\theta_1 \in \Theta_1$, $\bar{Z}_n \equiv Z_n/n \xrightarrow{a.s.} G'(\theta_1)$,

<u>Proof</u>: Clearly, $T_{i:n}$ $, i = 1, 2$ are partial sums of $i.i.d.$ random variables having finite moments (of all orders). So that $\bar{T}_{i:n} \xrightarrow{a.s.} \nu_i$, $i = 1, 2$. Since ψ is continuous, we have by (2.4) and (2.7), that

$$\frac{Z_n}{n} = \bar{T}_{1:n} - \psi(\bar{T}_{2:n}) \xrightarrow{a.s} \nu_1 - \psi(\nu_2) = G'(\theta_1).$$

$\blacksquare$

The following are some examples illustrating the construction of the statistic Z_n.

<u>*Example 1*</u>: The Normal distribution, $N(\mu, \sigma^2)$.

(i) $\theta_1 = -1/2\sigma^2$, $\theta_2 = \mu/\sigma^2$, $\Theta = \mathrm{R}^- \times \mathrm{R}$

(ii) $U_1(X) = X^2$, $U_2(X) = X$, $T_{1:n} = \sum_{i=1}^n X_i^2$, $T_{2:n} = \sum_{i=1}^n X_i$

(iii) $\nu_2 = -\theta_2/2\theta_1$, $\theta_2 = -2\theta_1\nu_2$

(iv) $\nu_1 = \nu_2^2 - 1/2\theta_1$, $\psi(\nu_2) = \nu_2^2$, $G'(\theta_1) = -1/2\theta_1$

(v) $Z_n = T_{1:n} - n\psi(\bar{T}_{2:n}) = \sum_{i=1}^n (X_i - \bar{X}_n)^2 > 0$ $a.s.$

<u>*Example 2*</u>: The Gamma distribution, $\mathcal{G}(\alpha, \lambda)$.

(i) $\theta_1 = \alpha$, $\theta_2 = -\lambda$, $\Theta = \mathrm{R}^+ \times \mathrm{R}^-$

(ii) $U_1(X) = \log(X)$, $U_2(X) = X$, $T_{1:n} = \sum_{i=1}^n \log(X_i)$, $T_{2:n} = \sum_{i=1}^n X_i$.

(iii) $\nu_2 = \alpha/\lambda$, $\psi(\nu_2) = \log(\nu_2)$, $G'(\alpha) = \Gamma'(\alpha)/\Gamma(\alpha) - \log(\alpha)$

(iv) $Z_n = T_{1:n} - n\psi(\bar{T}_{2:n}) = \sum_{i=1}^n \log(X_i/\bar{X}_n) < 0$ $a.s.$.

<u>*Example 3*</u>: The Inverse Gaussian distribution.

(i) $f(x : \lambda, \alpha) = (2\pi)^{1/2} x^{-3/2} \lambda^{1/2} \exp\{-\alpha x/2 - \lambda/2x + (\alpha\lambda)^{1/2}\}$, $x, \lambda \in \mathrm{R}^+$, $\alpha \in \mathrm{R}^+ \cup \{0\}$.

(ii) $U_1(X) = 1/X$, $U_2(X) = X$, $\theta_1 = -\lambda/2$, $\theta_2 = -\alpha/2$, $\Theta = \mathrm{R}^- \times (\mathrm{R}^- \cup \{0\})$.

(iii) $\nu_2 = -(\theta_1/\theta_2)^{1/2}$, $\psi(\nu_2) = 1/\nu_2$, $G'(\theta_1) = -1/2\theta_1$.

(v) $Z_n = \sum_{i=1}^n (1/X_i) - (n/\bar{X}) > 0$ $a.s.$; although this model is *steep*, all results stated above for a *regular* model, hold for $\theta \in int(\Theta)$. For further discussion, see Bar-Lev and Reiser (1982).

3. The sequential estimation procedure. Suppose that on the basis of n independent observations $x_1, \ldots, x_n$ from (2.1), we wish to estimate $\nu_2 \equiv E_\theta(U_2)$ in the presence of the nuisance parameter θ_1. Let $\hat{\theta}_1$ and $\hat{\nu}_2$ denote the maximum likelihood estimators of θ_1 and ν_2, respectively. So that by (2.2), $\hat{\theta}_1$ and $\hat{\nu}_2$ are the simultaneous solutions of the (log-likelihood derivatives) equations:

$$\begin{cases} T_{1:n} - n\nu_1 = 0 \\ \\ T_{2:n} - n\nu_2 = 0 \end{cases} \tag{3.1}$$

Hence, by using (2.7) in (3.1) we immediately obtain that $\hat{\nu}_2 = \bar{T}_{2:n}$ and that $\hat{\theta}_1$ satisfies the equation:

$$G'(\hat{\theta}_1) = \bar{T}_{1:n} - \psi(\bar{T}_{2:n}) = Z_n/n \ . \tag{3.2}$$

Further suppose that the loss incurred by using $\bar{T}_{2:n}$ as an estimate for ν_2 is:

$$L_\rho(\bar{T}_{2:n}) = \rho|\psi''(\nu_2)|(\bar{T}_{2:n} - \nu_2)^2 + n \ ,$$

where $\rho > 0$. The factor $\rho|\psi''(\nu_2)|$ represents the importance of the estimation error relative to the cost of one observation. ¿From (2.3) and (2.6) it follows that the for a fixed $\theta_1 \in \Theta_1$ the corresponding risk is:

$$R_\rho(n) = E_\theta[L_\rho(\bar{T}_{2:n})] = \frac{\rho}{n|\theta_1|} + n \ ,$$

which is minimized (w.r.t. n) at integer adjacent to $n_0 = (\rho/|\theta_1|)^{\frac{1}{2}}$, at which $R_\rho(n_0) = 2n_0$. However, since θ_1 is unknown, the estimation procedure has to be conducted sequentially, and to be terminated according to the stopping rule N_ρ where:

$$N_\rho = \inf\{n \geq m_0 \ ; |\hat{\theta}_1| > \rho/n^2\} \tag{3.3}$$

for some initial sample size m_0 , $(m_0 \geq 2)$. Moreover, since by Lemma 2.1.c, the function $G'(\theta_1)$ is strictly increasing on Θ_1, it follows from (3.2) and Lemma 2.2, that the stopping rule (3.3) has the following forms:

(i) If $\Theta_1 \subset \mathbf{R}^-$ then; $\quad N_\rho = \inf\{n \geq m_0 \ ; Z_n < nG'(\frac{-\rho}{n^2})\}$,

(ii) If $\Theta_1 \subset \mathbf{R}^+$ then; $\quad N_\rho = \inf\{n \geq m_0 \ ; Z_n > nG'(\frac{\rho}{n^2})\}$.

Remark: In either case, the event $\{N_\rho = n\}$ is determined only by $(Z_{m_0}, \ldots, Z_n)$, and therefore by Theorem 2.2 is independet of $\bar{T}_{2:n}$.

By Lemma 2.2, the symmetry of the two cases; $\Theta_1 \subset \mathrm{R}^+$ and $\Theta_1 \subset \mathrm{R}^-$ is evident and in view of (2.5) and proposition 1.6 of Brown (1986), there is no loss of generality by assuming (conveniently) that $\Theta_1 \subset \mathrm{R}^-$. Accordingly, we let $\Theta_1 \subset \mathrm{R}^-$ (so that $\theta_1 < 0$), and consider the stopping rule N_ρ as defined in (i) above.

Since the function G' is strictly increasing (and positive) on Θ_1 and $\bar{Z}_n$ converges *a.s.* to the finite limit $G'(\theta_1)$, it follows that for each fixed ρ, the stopping rule N_ρ is finite *w.p.1.* Moreover, since $G'(\frac{-\rho}{n^2})$ is decreasing as a function of ρ, N_ρ is stochastically increasing in ρ *w.p.1*, i.e.; for $0 < \rho_1 < \rho_2$, $N_{\rho_1} < N_{\rho_2}$ *w.p.1*, and hence, $\lim_{\rho\to\infty} N_\rho = \infty$ *w.p.1.*

The main results of this section are presented in the following two theorems.

<u>Theorem 3.1</u>. *Let N_ρ be the stopping time as defined in (i) above, then for all $\theta \in \Theta$ the following properties hold:*

a) *For each fixed ρ,* $E_\theta(N_\rho) < \infty$

b) $\lim_{\rho\to\infty} \frac{N_\rho}{n_0} = 1$ *w.p.1*

c) $\lim_{\rho\to\infty} E_\theta(\frac{N_\rho}{n_0}) = 1$

As was shown by Starr (1966) and by Woodroofe (1977, 1982), the initial sample size m_0 plays a crucial role in any attempt to analyze the *risk* (as well as the *regret*) associated with N_ρ. Moreover, it was shown, (see Woodroofe (1977), pp. 987), that the left tail behavior of the underlying c.d.f. is also crucial in the risk's assessments. For the general case, we have a need to impose the following conditions on the model at hand. The first condition pertains to G'. Notice that G' determines both; the boundary for the stopping rule N_ρ, as well as the moments of Z_n. The second condition is imposed to ensure an appropriate initial sample size m_0.

<u>Assumption A.3</u>. *For some $\gamma > 1/2$, $\sup_{x\geq 4|\theta_1|} x^\gamma G'(-x) \leq M < \infty$.*

<u>Assumption A.4</u>. . *The initial sample size m_0 is such that $\forall\, \theta_1 \in \Theta_1$, $E_{\theta_1}(Z_{m_0}^{-\beta}) < \infty$ for some $\beta > \frac{2}{(2\gamma-1)}$.*

<u>Theorem 3.2</u>. *Let $R_\rho(N)$ denote the risk associated with the stopping rule $N \equiv N_\rho$, then under Assumptions A.1-A.4:*

$$\lim_{\rho\to\infty} \frac{R_\rho(N)}{R_\rho(n_0)} = 1 .$$

Theorem 3.2 asserts that the proposed estimation procedure is asymptotically *risk efficient.* That is, the risk incurred by the sequential estimation procedure based on N_ρ, is asymptotically equivalent to the risk incurred by estimation procedure based on the optimal (and hypothetical) fixed sample size n_0. Note that for the normal case, Assumption A.3 holds with $\gamma = 1$ and Assumption A.4, is satisfied for $\beta > 2$ and $m_0 \geq 6$, (see in comparison: Woodroofe (1977)). However, it should be noted that it may be possible to relax the above requirements on m_0 and β in particular cases.

Proof of *Theorem* 3.1:

a) Let ρ be fixed so that $n_0 < \infty$. Fix $\epsilon > 1$, clearly

$$
\begin{aligned}
m_0 \leq E_\theta(N_\rho) &\leq n_0 + \sum_{n=n_0+1}^{\infty} P_\theta(N_\rho > n) \\
&\leq n_0 + (n_0+1)(\epsilon - 1) + \sum_{n=K}^{\infty} P_{\theta_1}(Z_n > nG'(\frac{-\rho}{n^2})),
\end{aligned} \tag{3.4}
$$

where $K = [(n_0+1)\epsilon] + 1$ and $[x]$ denotes the integer part of x. But according to Lemma A in Appendix, for all $n \geq K$

$$
P_{\theta_1}(Z_n > nG'(\frac{-\rho}{n^2})) \leq e^{-(n-n_0)C},
$$

for some constant $C > 0$. Hence, the last inequality in (3.4), implies that

$$
\begin{aligned}
E_\theta(N_\rho) &\leq n_0 + (n_0+1)(\epsilon - 1) + \sum_{n=K}^{\infty} e^{-(n-n_0)C} \\
&\leq n_0 + (n_0+1)(\epsilon - 1) + \frac{e^{-Cn_0(\epsilon-1)}}{1-e^{-C}} < \infty.
\end{aligned} \tag{3.5}
$$

b) For this part, we make use of Lemma 2.2(b) along with the definition of N_ρ, to obtain the inequalities

$$
(N_\rho - 1)G'(\frac{-\rho}{(N_\rho - 1)^2}) \leq Z_{N_\rho - 1} < Z_{N_\rho} < N_\rho G'(\frac{-\rho}{{N_\rho}^2}),
$$

which hold $w.p.1$. Since $Z_{N_\rho}/N_\rho \xrightarrow{a.s} G'(\theta_1)$, as $\rho \to \infty$, it follows that $\lim_{\rho\to\infty} G'(\frac{-\rho}{N_\rho^2}) = G'(\theta_1)$. Then, by using the relation $-\rho = \theta_1 n_0^2$, the required result follows.

c) From part (b) and Fatou's Lemma $\liminf_{\rho\to\infty} E_\theta(N_\rho/n_0) \geq 1$ Also, by (3.5) above, $\limsup_{\rho\to\infty} E_\theta(N_\rho/n_0) \leq 1 + (\epsilon - 1)$. Finally, by letting $\epsilon \to 1$, the proof of (c) is completed. ∎

Proof of *Theorem* 3.2: By definition,

$$R_\rho(N) = E_\theta[\rho|\psi''(\nu_2)|(\bar{T}_{2:N_\rho} - \nu_2)^2 + N_\rho].$$

Note that by the definition of N_ρ, the event $\{N_\rho = n\}$, depends only on $Z_1 \dots Z_n$ and hence by Theorem 2.2 is independent on $\bar{T}_{2:n}$. Thus:

$$\begin{aligned} R_\rho(N_\rho) &= E_\theta\left(\frac{\rho}{N_\rho|\theta_1|}\right) + E_\theta(N_\rho) \\ &= E_\theta\left(\frac{n_0^2}{N_\rho}\right) + E_\theta(N_\rho), \end{aligned}$$

and therefore:

$$\frac{R_\rho(N_\rho)}{R_\rho(n_0)} = \frac{1}{2}E_\theta\left(\frac{n_0}{N_\rho}\right) + \frac{1}{2}E_\theta\left(\frac{N_\rho}{n_o}\right).$$

In view of Theorem 3.1 it suffices to show that $\limsup_{\rho\to\infty} E_\theta(\frac{n_0}{N_\rho}) \le 1$.

Fix $0 < \varepsilon < 1/2$, then

$$\begin{aligned} E_\theta\Big(\frac{n_0}{N_\rho}\Big) =& E_\theta\Big(\frac{n_0}{N_\rho}I(N_\rho \le n_0/2)\Big) \\ &+ E_\theta\Big(\frac{n_0}{N_\rho}I\Big(\frac{n_0}{2} \le N_\rho < n_0(1-\varepsilon)\Big) \\ &+ E_\theta\Big(\frac{n_0}{N_\rho}I(n_0(1-\varepsilon) \le N_\rho < n_0(1+\varepsilon)\Big)\Big) \\ &+ E_\theta\Big(\frac{n_0}{N_\rho}I\Big(N_\rho \ge n_0(1+\varepsilon)\Big) \\ =& B_1 + B_2 + B_3 + B_4, \quad \text{say}. \end{aligned}$$

By Lemma B in appendix, $B_1 \to 0$, as $\rho \to \infty$.
As immediate consequences of Theorem 3.1 and Lemma A in appendix;

$$\begin{aligned} B_2 &\le 2P\Big[\frac{1}{2} \le \frac{N_\rho}{n_0} < 1-\varepsilon\Big] \to 0 \\ B_4 &\le \frac{1}{(1+\varepsilon)}P\Big[\frac{N_\rho}{n_0} > 1+\varepsilon\Big] \to 0. \end{aligned}$$

Finally, by using the dominated convergence theorem it is easy to show that $B_3 \to 1$ as $\rho \to \infty$, which completes the proof. ■

Concluding Remark: It is evident that the independence result, presented in Theorem 2.2, is a crucial key in the risk assessments. It will also turn out to be an important tool in other sequential problems concerning the family of distributions we have discussed here.

These problems include the second order properties of the risk, as well as problems similar to those discussed in Siegmund (1985). Currently, we are studying some of these problems and the results will appear in a future paper.

APPENDIX

Proof of *Theorem* 2.2. The proof will be carried in two steps.

Step 1. Show that $(Z_{j-1}, Z_j) \perp T_{2:j}$ for all $j > 2$.

Step 2. Show that if $(Z_2, \ldots, Z_i) \perp T_{2:i}$ for all $i \leq k$ then $(Z_2, \ldots, Z_{k+1}) \perp T_{2:k+1}$.

We will use the following notations in the proof. For $j > 2$, $\mathbf{Z}_j = (Z_2, \ldots, Z_j)$, $\mathbf{0}_j = \underbrace{(0, \ldots, 0)}_{\text{(j-1) times}}$, $\boldsymbol{\alpha}_j = (\alpha_2, \ldots, \alpha_j)$, $\boldsymbol{\beta}_j = (\beta_2, \ldots, \beta_j)$ where α_i's are and β_i's are complex numbers to be specified later.

Note that the joint density of $(\mathbf{Z}_{j-1}, T_{1:j}, T_{2:j})$ is of the form:

$$f(\mathbf{z}_{j-1}, t_{1:j}, t_{2:j}) = K_j(\mathbf{z}_{j-1}, t_{1:j}, t_{2:j}) \exp(\theta_1 t_{1:j} + \theta_2 t_{2:j} + j c(\theta))$$

for some function $K_j(\cdot) > 0$.

For each j, $(j > 2)$, define the functions ϕ_j and b_j as follows, whenever they exist.

$$\phi_j(\boldsymbol{\alpha}_j, t_{2:j}, \theta_1) = E\left[\exp\left(i \sum_{k=2}^{j} \alpha_k Z_k\right) \middle| T_{2:j} = t_{2:j}\right]$$

$$b_j(\boldsymbol{\beta}_j, t_{2:j}) = \int K_j(\mathbf{z}_{j-1}, t_{1:j}, t_{2:j}) \exp\left(\sum_{k=2}^{j-1} \beta_k z_k + \beta_j t_{1:j}\right) dz_2 \ldots dz_{j-1} dt_{1:j}$$

Note that the conditional density of $(Z_2, \ldots, Z_{j-1}, T_{1:j})$ given $T_{2:j} = t_{2:j}$ is given by

$$f(\mathbf{z}_{j-1}, t_{1:j} | T_{2:j} = t_{2:j}) = \frac{K_j(\mathbf{z}_{j-1}, t_{1:j}, t_{2:j}) \exp(\theta_1 t_{1:j})}{b_j(\mathbf{0}_{j-1}, \theta_1, t_{2:j})} \tag{A.1}$$

To prove Step 1, we need to show that $\phi_j(\mathbf{0}_{j-2}, \alpha_{j-1}, \alpha_j, t_{2:j}, \theta_1) \equiv \phi(\alpha_{j-1}, \alpha_j, t_{2:j}, \theta_1)$, which is the conditional characteristic function of (Z_{j-1}, Z_j) given $T_{2:j} = t_{2:j}$, does not involve $t_{2:j}$ for (α_{j-1}, α_j) in a neighborhood of $\mathbf{0}$ in $\mathbf{R}^2$.

Using (A.1),

$$\phi(\alpha_{j-1}, \alpha_j, t_{2:j}, \theta_1) = \frac{e^{-i\alpha_j j \psi(\bar{t}_{2:j})} b_j(\mathbf{0}_{j-2}, i\alpha_{j-1}, i\alpha_j + \theta_1, t_{2:j})}{b_j(\mathbf{0}_{j-1}, \theta_1, t_{2:j})}$$

where we have used the fact that $Z_j = T_{1:j} - j\psi(\overline{T}_{2:j})$.
Thus

$$(A.2) \qquad e^{-i\alpha_j j\psi(\bar{t}_{2:j})} b_j(\mathbf{0}_{j-2}, i\alpha_{j-1}, i\alpha_j + \theta_1, t_{2:j}) = b_j(\mathbf{0}_{j-1}, \theta_1, t_{2:j})\phi(\alpha_{j-1}, \alpha_j, t_{2:j}, \theta_1)$$

Note that $Z_{j-1} \perp T_{2:j-1}$ and hence $Z_{j-1} \perp T_{2:j-1}+U_2(X_j) = T_{2:j}$. Thus $\phi(\alpha_{j-1}, 0, t_{2:j}, \theta_1)$ does not involve $t_{2:j}$. Call this function $\phi(\alpha_{j-1}, 0, \theta_1)$.
Thus, by substituting $\alpha_j = 0$ in (A.2), we get:

$$(A.3) \qquad b_j(\mathbf{0}_{j-2}, i\alpha_{j-1}, \theta_1, t_{2:j}) = b_j(\mathbf{0}_{j-1}, \theta_1, t_{2:j})\phi(\alpha_{j-1}, 0, \theta_1).$$

By extending the parameter space to the complex plane, it is easy to extend the definition of $\phi(\alpha_{j-1}, 0, \theta_1)$ to $\phi(\alpha_{j-1}, 0, i\alpha_j+\theta_1)$. Using analytic continuation, equation (A.3) continues to hold when θ_1 is replaced by $i\alpha_j + \theta_1$. Accordingly

$$(A.4) \qquad b_j(\mathbf{0}_{j-2}, i\alpha_{j-1}, i\alpha_j + \theta_1, t_{2:j}) = b_j(\mathbf{0}_{j-1}, i\alpha_j + \theta_1, t_{2:j})\phi(\alpha_{j-1}, 0, i\alpha_j + \theta_1).$$

Hence, using (A.2)–(A.4),

$$(A.5) \qquad \begin{aligned} \phi(\alpha_{j-1}, \alpha_j, t_{2:j}, \theta_1) &= e^{-i\alpha_j j\psi(\bar{t}_{2:j})} \frac{b_j(\mathbf{0}_{j-2}, i\alpha_{j-1}, i\alpha_j + \theta_1, t_{2:j})}{b_j(\mathbf{0}_{j-1}, \theta_1, t_{2:j})} \\ &= \phi(\alpha_{j-1}, 0, i\alpha_j + \theta_1) e^{-i\alpha_j j\psi(\bar{t}_{2:j})} \frac{b_j(\mathbf{0}_{j-1}, i\alpha_j + \theta_1, t_{2:j})}{b_j(\mathbf{0}_{j-1}, \theta_1, t_{2:j})}. \end{aligned}$$

Comparing with equations (3.18)–(3.19) of Bar–Lev and Rieser (1982), (note that their $b(s, t_2) \equiv b_j(\mathbf{0}_{j-1}, s, t_{2:j})$ and their $U_1 g^{-1}(\bar{t}_2) \equiv \psi(\bar{t}_{2:j})$), we get that for $s = \theta_1 + i\alpha, \alpha \in \mathbb{R}$, $\log b_j(\cdot)$ is of the form;

$$\log b_j(\mathbf{0}_{j-1}, s, t_{2:j}) = js\psi(\bar{t}_{2:j}) + R_j(\bar{t}_{2:j}) + H_j(s).$$

Using this, we immediately obtain that

$$\frac{b_j(\mathbf{0}_{j-1}, i\alpha_j + \theta_1, t_{2:j})}{b_j(\mathbf{0}_{j-1}, \theta_1, t_{2:j})} = e^{i\alpha_j j\psi(\bar{t}_{2:j}) - [H_j(i\alpha_j+\theta_1) - H_j(\theta_1)]}.$$

This shows that (A.5) does not involve $t_{2:j}$, completing the proof of Step 1.

To prove Step 2, we need to show that $\phi_{k+1}(\alpha_{k+1}, t_{2:k+1}, \theta_1)$ does not involve $t_{2:k+1}$ for α_{k+1} in a neighborhood of $\mathbf{0}$ in $\mathbb{R}^k$.
Proceeding as in the proof of Step 1,

$$(A.6) \quad \phi_{k+1}(\alpha_{k+1}, t_{2:k+1}, \theta_1) = \frac{e^{-i(k+1)\alpha_{k+1}\psi(\bar{t}_{2:k+1})} b_{k+1}(i\alpha_1, \ldots, i\alpha_k, i\alpha_{k+1} + \theta_1, t_{2:k+1})}{b_{k+1}(\mathbf{0}_k, \theta_1, t_{2:k+1})}$$

However, by hypothesis, $(Z_1, \ldots, Z_k) \perp T_{2:k}$ and hence $(Z_1, \ldots, Z_k) \perp T_{2:k+1}$. Thus

$$\phi_{k+1}(\alpha_k, 0, t_{2:k+1}, \theta_1) \equiv \phi_{k+1}(\alpha_k, 0, \theta_1),$$

is independent of $t_{2:k+1}$. Accordingly,

$$(A.7) \qquad b_{k+1}(i\alpha_1, \ldots, i\alpha_k, \theta_1, t_{2:k+1}) = b_{k+1}(\mathbf{0}_k, \theta_1, t_{2:k+1})\phi_{k+1}(\alpha_k, 0, \theta_1)$$

Again, arguing as in Step 1,

$$(A.8) \qquad \begin{aligned} b_{k+1}(i\alpha_1, \ldots, i\alpha_k, i\alpha_{k+1} + \theta_1, t_{2:j+1}) = b_{k+1}(\mathbf{0}_k, i\alpha_{k+1} + \theta_1, t_{2:k+1}) \times \\ \times\, \phi_{k+1}(\alpha_k, 0, i\alpha_{k+1} + \theta_1). \end{aligned}$$

Using (A.6), (A.7), (A.8) and equations (3.18)–(3.19) of Bar–Lev and Reiser (1982) the proof of Step 2 can now be completed exactly as in Step 1. ∎

<u>LEMMA A</u>. *Let* $n_0 = (\rho/|\theta_1|)^{1/2}$ *,*$(\theta_1 < 0)$ *and* $\epsilon > 1$*, be fixed and let* Z_n *be as defined in (2.4) with the p.d.f (2.5). Then for all* $n > n_0\epsilon$ *there exits a constant* C_0 *such that:*

$$P_{\theta_1}(Z_n > nG'(\frac{-\rho}{n^2})) \leq \exp\{-(n - n_0)C_0(\epsilon - 1)/2\epsilon\}.$$

<u>Proof:</u> By (2.5), the moment generating functions, $M_{Z_n}(t)$ of Z_n, exists for all $t < -\theta_1$, and is given by:

$$M_{Z_n}(t) = \exp\{H_n(t + \theta_1) - H_n(\theta_1)\}, \quad t < -\theta_1,$$

with $H_n(\cdot)$ as defined in (2.8). Let $\varepsilon_n = (\frac{n_0}{n})^2 < 1$ and let $t_n = \theta_1(\varepsilon_n - 1)$. Clearly, $t_n \in [0, \ -\theta_1)$. It can be easily verified that

$$p_n \equiv P_{\theta_1}[Z_n > nG'(\theta_1\varepsilon_n)] \leq e^{-t_n nG'(\theta_1\varepsilon_n)}M_{Z_n}(t_n) \equiv \exp\{\varphi_n(t_n)\},$$

where we have put: $\varphi_n(t) = H_n(t + \theta_1) - H_n(\theta_1) - tnG'(\theta_1\varepsilon_n), \ \ t \geq 0$. However, by the definition of $H_n(\cdot)$,

$$(A.9) \qquad \begin{aligned} \varphi_n(t_n) &= H_n(t_n + \theta_1) - H_n(\theta_1) - t_n nG'(\theta\varepsilon_n) \\ &= n[G(\theta_1\varepsilon_n) - G(\theta_1)] - [G(n\theta_1\varepsilon_n) - G(n\theta_1)] + \theta_1(1 - \varepsilon_n)nG'(\theta_1\varepsilon_n). \end{aligned}$$

Since $G(n\theta_1\varepsilon_n) - G(n\theta_1) > 0$, and $G''(\cdot) > 0$, the last equality in A.9 implies that

$$\varphi_n(t_n) \leq -n\theta_1^2(1 - \varepsilon_n)^2 G''(\theta_1\varepsilon_n^*)/2\ ,$$

for some ε_n^* between 1 and ε_n. Note that $G''(x) \geq C_0 > 0$ for all $x \in [\theta_1, 0]$, (see also the discussion following (2.8)). In addition, since $n > n_0\epsilon$, we have:

$$(1-\varepsilon_n)^2 \geq (1-1/\epsilon)(1-\varepsilon_n) \geq (1-1/\epsilon)(1-n_0/n).$$

Accordingly:

$$\begin{aligned} p_n \leq \exp\{\varphi_n(t_n)\} &\leq \exp\{-nC_0(1-\varepsilon_n)^2/2\} \\ &\leq \exp\{-(n-n_0)C_0(\epsilon-1)/2\epsilon\}, \end{aligned}$$

which completes the proof. ■

<u>LEMMA B</u>. *Suppose that $G'(\cdot)$ and m_0 satisfy Assumptions A.3 and A.4, then:*

$$E(\frac{n_0}{N_\rho} I[N_\rho \leq n_0/2]) \to 0 \text{ as } \rho \to \infty.$$

<u>Proof</u> Let $1/2 < \alpha < 1$ be fixed, (to be chosen later), and let C be a generic constant. Then

$$\begin{aligned} &E(\frac{n_0}{N_\rho} I[m_0 \leq N_\rho \leq n_0/2] \\ &\leq n_0 E(\frac{1}{N_\rho} I[m_0 \leq N_\rho \leq n_0^\alpha] \\ &+ n_0^{(1-\alpha)} P(n_0^\alpha < N_\rho \leq n_0/2) \\ &= I_1 + I_2 \quad \text{(say)}. \end{aligned}$$

Now, for the first term I_1,

$$\begin{aligned} E(\frac{1}{N_\rho} I[m_0 \leq N_\rho \leq n_0^\alpha]) &= \sum_{k=m_0}^{[n_0^\alpha]} \frac{1}{k} P(N_\rho = k) \\ &\leq \sum_{k=m_0}^{[n_0^\alpha]} \frac{1}{k} P(Z_k \leq kG'(\theta_1(\frac{n_0}{k}^2)) \\ &\leq \sum_{k=m_0}^{[n_0^\alpha]} \frac{1}{k} P(Z_k < \frac{k^{1+2\gamma}}{n_0^{2\gamma}|\theta_1|^\gamma} M) \\ &\leq \sum_{k=m_0}^{[n_0^\alpha]} \frac{1}{k} P(Z_{m_0} < \frac{k^{1+2\gamma}}{n_0^{2\gamma}} C) \\ &\leq E(Z_{m_0}^{-\beta}) C n_0^{-2\gamma\beta} \sum_{k=m_0}^{[n_0^\alpha]} k^{(1+2\gamma)\beta - 1}, \end{aligned}$$

where the last two inequalities following Assumptions A.3-A.4, and Lemma 2.2 (b). Accordingly

$$I_1 \le E\left(Z_{m_0}^{-\beta}\right) C n_0^{(1-2\gamma\beta+\alpha\beta(1+2\gamma))},$$

which $\to 0$ for $\alpha < (2\gamma\beta - 1)/\beta(1+2\gamma)$.

Clearly, $I_2 \le n_0^{(1-\alpha)} P\Big[Z_k < kG'\big(\theta_1\big(\frac{n_0}{k}\big)^2\big)$, for some $k \in \big(n_0^\alpha, \frac{n_0}{2}\big]\Big]$.

Define: $L_{1:k} = T_{1:k} - k\nu_1$, $L_{2:k} = \varphi(\overline{T}_{2:k}) - \varphi(\nu_2)$, note that $Z_k = L_{1:k} - kL_{2:k}$, and by (2.7), $\nu_1 - \varphi(\nu_2) = G'(\theta_1)$, hence;

$$I_2 \le n_0^{(1-\alpha)} P\big[L_{1:k} - kL_{2:k} < k\Delta_k, \text{ for some } k \in (n_0^\alpha, \frac{n_0}{2}]\big]$$

with $\Delta_k = G'\big(\theta_1\big(\frac{n_0}{k}\big)^2\big) - G'(\theta_1)$. Since $G'(\cdot)$ is increasing and $k < \frac{n_0}{2}$,

$$\Delta_k \le G'(4\theta_1) - G'(\theta_1) \equiv -2\varepsilon(<0), \text{ say.}$$

Thus,

$$\begin{aligned} I_2 &\le n_0^{(1-\alpha)} P\big[L_{1:k} - kL_{2:k} < -k\varepsilon, \text{ for some } k \in (n_0^\alpha, \frac{n_0}{2}]\big] \\ &\le n_0^{(1-\alpha)} P\big[|L_{1:k}| > k\varepsilon, \text{ for some } k \in (n_0^\alpha, \frac{n_0}{2}]\big] + \\ &+ n_0^{(1-\alpha)} P\big[\ k|L_{2:k}| > k\varepsilon, \text{ for some } k \in (n_0^\alpha, \frac{n_0}{2}]\big] \\ &= I_{21} + I_{22}, \text{ (say)}. \end{aligned}$$

Since $T_{1:k}$ has moments of all orders, it follows immediately, using submartingale inequality that,

$$\begin{aligned} I_{21} &\le n_0^{(1-\alpha)} P\big[\max_{n_0^\alpha < k \le n_0/2} |L_{1:k}| > n_0^\alpha \varepsilon\big] \\ &\le n_0^{(1-\alpha)} n_0^{-\alpha r} \varepsilon^r E\big(\big(L_{1:[\frac{n_0}{2}]}\big)^r\big), \ r > 0 \\ &= 0\big(n_0^{1-\alpha+r(\frac{1}{2}-\alpha)}\big). \end{aligned} \tag{A.10}$$

As for the second term I_{22}, it follows by the continuity of $\varphi(\cdot)$ that there is $\delta(\varepsilon) > 0$ such that $|x - \nu_2| < \delta(\varepsilon) \Rightarrow |\varphi(x) - \varphi(\nu_2)| < \varepsilon$. Thus,

$$\begin{aligned} I_{22} &\le n_0^{(1-\alpha)} P\big[|\overline{T}_{2:k} - \nu_2| > \delta(\varepsilon), \text{ for some } k \in (n_0^\alpha, \frac{n_0}{2}]\big] \\ &\le n_0^{(1-\alpha)} P\big[|T_{2:k} - \nu_2| > n_0^\alpha \delta(\varepsilon), \text{ for some } k \in (n_0^\alpha, \frac{n_0}{2}]\big] \\ &= 0\big(n_0^{1-\alpha+r(\frac{1}{2}-\alpha)}\big), \end{aligned} \tag{A.11}$$

again, by using submartingale inequality as in (A.10). Finally, by combining (A.10) and (A.11) together, we obtain;

$$I_2 \le 0(n_0^{1-\alpha+r(\frac{1}{2}-\alpha)}) \to 0,$$

for r large and $\alpha > 1/2$.

Hence, upon choosing $1/2 < \alpha < \frac{(2\gamma\beta-1)}{\beta(1+2\gamma)}$ with $\beta > 2/(2\gamma-1)$ as required, the proof is completed. ■

Bibliography

1. Aras, G. (1987). Sequential estimation of the mean exponential survival time under random censoring, *J. Statis. Plann. Inf.* **16**, 147–158.

2. Aras, G. (1989). Second order sequential estimation of the mean exponential survival time under random censoring, *J. Statis. Plann. Inf.* **21**, 3–17.

3. Bar–Lev, S.K. and Reiser, B. (1982). An exponential subfamily which admits UMPU test based on a single test statistic, *Ann. of Stat.* **10** (3), 979–989.

4. Barndorff–Nielson, O., *"Information and The Exponential Families in Statistical Theory"*, John Wiley, New–York, 1978.

5. Brown, L.D. (1986). *Fundamentals of Statistical Exponential Families*, IMS Lecture Notes, Vol 9, Editor: S. S. Gupta.

6. Chow, Y.S. and Robbins H. (1965). Asymptotic theory of fixed width confidence intervals for the mean, *Ann. Math. Stat.* **36**, 457–462.

7. Ghosh, M. and Mukhopadhyay, N. (1979). Sequential point estimation of the mean when the distribution is unspecified. *Commun. Statist. – Theor. Meth.* **8** (7), 637–652.

8. Mukhopadhyay, N. (1988). Sequential estimation problems for negative exponential populations, *Commun. Statist.-Theor. Meth.* **17** (8), 2471–2506.

9. Robbins, H. (1959). Sequential estimation of the mean of a normal population, *Probability and Statistics*, Almquist and Wiksell, Stockholm.

10. Siegmund, D. (1985). *"Sequential Analysis, Tests and Confidence Intervals"*, Springer-Verlag, New-York.

11. Starr, N. (1966). The performance of a sequential procedure for the fixed width interval estimation of the mean, *Ann. Math. Stat.* **37**, 36–50.

12. Starr, N. and Woodroofe, M. (1969). Remarks on sequential point estimation, *Proc. Nat. Acad. Sci.* **66**, 285–288.

13. Starr, N. and Woodroofe, M. (1972). Further remarks on sequential estimation: The exponential case, *Ann. of Mathematical Statistics* **43** (4), 1147–1154.

14. Stein, C. (1945). A two-sample test for a linear hypothesis whose power is independent of the variance, *Ann. Math. Stat.* **16**, 243–258.

15. Stein, C. (1949), Some problems in sequential estimation, *Econometrica* **17**, 77–78.

16. Woodroofe, M. (1977). Second order approximations for sequential point and interval estimation, *Ann. of Statistics*, **5** (5), 984–995.

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

18. Woodroofe, M. (1985). Asymptotic local minimaxity in sequential point estimation, *Ann. of Stattistics*, 13 (2), 676-687.