The Annals of Statistics
1993, Vol. 21, No. 1, 484–502

SEQUENTIAL ESTIMATION RESULTS FOR A TWO-PARAMETER EXPONENTIAL FAMILY OF DISTRIBUTIONS

BY ARUP BOSE AND BENZION BOUKAI

Purdue University

We consider the problem of sequentially estimating one parameter in a class of two-parameter exponential family of distributions. We assume a weighted 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 asymptotic normality of the stopping variable is established and approximations to its mean and to the regret associated with it are also provided. The general results are exemplified by the normal, gamma and the inverse Gaussian densities.

1. Introduction.

Let

$$f(x;\theta) = a(x)\exp\{\theta_1 U_1(x) + \theta_2 U_2(x) + c(\theta)\}, \qquad \theta = (\theta_1, \theta_2), \tag{1.1}$$

be a density function (w.r.t. Lebesgue measure on $\mathbb{R}$), of a *regular* two-parameter exponential family of distributions [see Brown (1986)]. The natural parameter space Θ is defined by

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

so that $\Theta \equiv \operatorname{int}\Theta \neq \varnothing$. 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

$$E_\theta(\mathbf{U}) = (\mu_1, \mu_2), \qquad \mu_i = -\partial c(\theta)/\partial\theta_i, \qquad i = 1,2 \tag{1.2}$$

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 (positive definite) variance–covariance matrix.

Let $X_1, \dots, X_n$, $n > 1$, be n independent identically distributed r.v.'s having a common density of the form (1.1). Let $T_{i:n} = \sum_{j=1}^n U_i(X_j)$ and denote by $\overline{T}_{i:n}$, $i = 1,2$ the usual averages. The joint distribution of $\mathbf{T} = (T_{1:n}, T_{2:n})$ is a member of the two-parameter exponential family and

$$E_\theta(\mathbf{T}) = (n\mu_1, n\mu_2), \qquad V_\theta(\mathbf{T}) = (n\sigma_{ij}), \qquad i,j = 1,2. \tag{1.3}$$

Received February 1990; revised August 1990.

AMS 1991 *subject classifications*. Primary 62L12; secondary 62E99.

Key words and phrases. Stopping rule, nuisance parameter, risk function, second order asymptotics.

In the present paper we consider a particular subfamily of (1.1), which was studied in detail by Bar-Lev and Reiser (1982) (henceforth referred to as BLR), in the context of construction of UMPU tests. This subfamily is characterized by the following two conditions which we assume to hold throughout the paper:

ASSUMPTION A.1. The parameter θ_2 can be represented as $\theta_2 = -\theta_1\psi'(\mu_2)$, where $\psi'(\mu_2) = d\psi(\mu_2)/d\mu_2$, for some function ψ.

ASSUMPTION A.2. $U_2(x) = h(x)$, where $h(x)$ is a $1-1$ function on the support of (1.1).

By using the mixed parametrization $(\theta_1, \theta_2) \to (\theta_1, \mu_2)$ which is a homeomorphism with $(\theta_1, \mu_2) \in \Theta_1 \times \mathscr{N}_2$ (varying independently, respectively), it can be shown that under Assumptions A.1 and A.2 the following relations hold (see BLR):

(a) The variance of U_2 is given by

$$\sigma_{22}(\theta) \equiv \frac{\partial\mu_2}{\partial\theta_2} = \frac{-1}{\theta_1\psi''(\mu_2)} \qquad (>0). \tag{1.4}$$

(b) The functions $c(\theta)$ and $\mu_1(\theta)$ when expressed by θ_1 and μ_2, have the following form

$$\begin{aligned} c(\theta_1, \mu_2) &= \theta_1[\mu_2\psi'(\mu_2) - \psi(\mu_2)] - G(\theta_1), \\ \mu_1 &= \psi(\mu_2) + G'(\theta_1), \end{aligned} \tag{1.5}$$

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.

An immediate consequence of (1.4) and Assumptions A.1–A.2 (see BLR) is that either $\Theta_1 \subset \mathbb{R}^-$ or $\Theta_1 \subset \mathbb{R}^+$.

Suppose that on the basis of n independent observations $x_1, \dots, x_n$ from (1.1), we wish to estimate $\mu_2 \equiv E_\theta(U_2)$ in the presence of the nuisance parameter θ_1. Let $\hat{\theta}_1$ and $\hat{\mu}_2$ denote the maximum likelihood estimators of θ_1 and μ_2, respectively. It is easy to show that under the preceding assumptions, $\hat{\mu}_2 = \overline{T}_{2:n}$ and that $\hat{\theta}_1$ satisfies the equation

$$nG'(\hat{\theta}_1) = T_{1:n} - n\psi(\overline{T}_{2:n}) \equiv Z_n. \tag{1.6}$$

The distributional properties of the statistic Z_n have been studied thoroughly by BLR. It has been shown there (see BLR, Theorem 3.2) that the distribution of the statistic Z_n is a member of the one-parameter exponential family of distributions indexed by θ_1. Denoting by $M_{Z_n}(\cdot)$ the moment generating function of Z_n, BLR have shown that

$$M_{Z_n}(t) = e^{H_n(t+\theta_1) - H_n(\theta_1)} \qquad t + \theta_1 \in \Theta_1, \tag{1.7}$$

where for all $\theta_1 \in \Theta_1$

$$H_n(\theta_1) = nG(\theta_1) - G(n\theta_1). \tag{1.8}$$

It follows (by Basu's Theorem) that for any fixed n and for all $\theta \in \Theta$, the statistics Z_n and $T_{2:n}$ are stochastically independent.

For the estimation problem considered here, we assume that the loss incurred by using $\overline{T}_{2:n}$ as an estimate for μ_2 is

$$L_\rho(\overline{T}_{2:n}) = \rho|\psi''(\mu_2)|(\overline{T}_{2:n} - \mu_2)^2 + n,$$

where $\rho > 0$. The factor $\rho|\psi''(\mu_2)|$ represents the importance of the estimation error relative to the cost of one observation.

From (1.3) and (1.4) it follows that for a fixed $\theta_1 \in \Theta_1$ the corresponding risk is

$$R_\rho(n) = E_\theta\left[L_\rho(\overline{T}_{2:n})\right] = \frac{\rho}{n|\theta_1|} + n.$$

The optimal sample size which minimizes the risk is obtained by choosing an integer adjacent to $n_0 = (\rho/|\theta_1|)^{1/2}$ at which $R_\rho(n_0) = 2n_0$. However, since θ_1 is an unknown nuisance parameter, this optimal sample size cannot be determined. So one may use a random sample size $\tilde{N}_\rho$ based on the following stopping rule,

$$\tilde{N}_\rho = \inf\{n \geq m_0; |\hat{\theta}_1| > \rho/n^2\},$$

for some initial sample size m_0 ($m_0 \geq 2$).

Since the function $G'(\theta_1)$ is strictly increasing on Θ_1, it follows from (1.6) that the stopping rule $\tilde{N}_\rho$ has one of the following forms:

(i) If $\Theta_1 \subset \mathbb{R}^-$, then $\tilde{N}_\rho = \inf\{n \geq m_0;\ Z_n < nG'(-\rho/n^2)\}$.
(ii) If $\Theta_1 \subset \mathbb{R}^+$, then $\tilde{N}_\rho = \inf\{n \geq m_0;\ Z_n > nG'(\rho/n^2)\}$.

By Lemma 1 these two cases are symmetrical and hence with no loss of generality we may assume from now on that $\Theta_1 \subset \mathbb{R}^-$.

In the sequel, we consider a simple modification of the stopping rule $\tilde{N}_\rho$ in (i). This modification is intended to reduce bias incurred by underestimating n_0 using $\tilde{N}_\rho$. The modified stopping variable is defined as

$$N_\rho = \inf\left\{n \geq m_0;\ Z_n a_n < nG'\left(\frac{-\rho}{n^2}\right)\right\}, \tag{1.9}$$

where we suppose that $a_n > 1$ are of the form $a_n = 1 + a_0/n + \delta_n$ with $\delta_n = o(1/n)$ as $n \to \infty$. By Lemmas 1 and 2, the function G' is strictly increasing and positive on Θ_1 and $\overline{Z}_n$ converges a.s. to $G'(\theta_1)$. Hence it follows that for each fixed ρ the stopping rule N_ρ is finite w.p.1. It is also easy to see that $\lim_{\rho\to\infty} N_\rho = \infty$ w.p.1.

In Section 2 we present the main results concerning some of the asymptotic properties of the stopping variable N_ρ (as $\rho \to \infty$) and of the risk $R_\rho(N_\rho)$

associated with it. In particular, we establish the asymptotic normality of (appropriately normalized) N_ρ and obtain approximation to $E_\theta(N_\rho)$. Denoting by $\mathscr{R}(\rho, \theta_1) = R_\rho(N_\rho) - R_\rho(n_0)$, the so-called regret (i.e., the additional risk incurred by the sequential estimation procedure based on N_ρ instead of n_0), we also show that

$$\lim_{\rho \to \infty} \mathscr{R}(\rho, \theta_1) = \left[4\theta_1^2 G''(\theta_1)\right]^{-1} > 0.$$

These types of results are usually referred to as second order approximations. These approximations are based on results from nonlinear renewal theory developed mainly by Lai and Siegmund [(1977), (1979)] and Woodroofe (1977). For a detailed exposition of this theory see Woodroofe (1982). A crucial key in proving our results is an independence result presented in Theorem 1. This result provides, in the general case discussed here, the independence of the event $\{N_\rho = n\}$ with the terminal estimator. Although we focus attention here on the point estimation problem, this result may also be used to study fixed width confidence interval problems.

Sequential estimation procedures similar to the one considered here have been discussed by several researchers. For the estimation of the mean of a normal population, Robbins (1959) suggested a stopping rule N based on the successive estimates of the population variance by the sample variance. The estimate $\bar{x}_n$ of the mean and the event $\{N = n\}$ are then independent for every n. This property was heavily exploited by most researchers who worked on the normal problem. Starr (1966) showed that for the normal case the sequential estimation procedure is *risk efficient* [i.e., $\lim_{\rho \to \infty} R_\rho(N)/R_\rho(n_0) = 1$], if and only if $m_0 \geq 3$. Starr and Woodroofe (1969) found that the regret is bounded. Woodroofe (1977) used second order approximations to study the regret and proved that $\mathscr{R}(\rho) \to 1/2$, as $\rho \to \infty$ if $m_0 \geq 6$.

Extensions of this procedure to nonnormal cases have also been discussed in the literature. Starr and Woodroofe (1972) dealt with the negative exponential distribution and proved the boundedness of the regret. Vardi (1979) provided similar results for his stopping time in the Poisson case. A survey of results concerning sequential estimation procedures for the negative exponential distribution, with and without a truncation parameter, can be found in Mukhopadhyay (1988). Aras [(1987), (1989)] provided first and second order results for the case of censored data from negative exponential distribution. To estimate the mean with a "distribution free" approach, Ghosh and Mukhopadhyay (1979) and Chow and Yu (1981) allowed the initial sample size m_0 to be a function of ρ, which tends to ∞ as $\rho \to \infty$. They show the risk efficiency of the estimation procedure when the distributions are unspecified and $m_0 \to \infty$. Similar results were proved by Sen and Ghosh (1981) for estimation of symmetric parametric functions using U-statistics. In the distribution free case, Chow and Martinsek (1982) have shown that the regret is bounded. With stronger conditions, Martinsek (1983) obtained an expression for the regret for nonlattice variables and bounds on the regret when the variables are lattice.

2. Main results. This section contains all the main results of this paper. To simplify the presentation, we provide their proofs separately in Section 3. Whenever possible we will omit the parameter subscript from probabilistic statements. We begin this section with the independence result mentioned in the introduction. Recall that Assumptions A.1 and A.2 are in force throughout the paper.

THEOREM 1. *For all* $n \geq 2$ *and* $\theta \in \Theta$, *the random variables* $(Z_2, \dots, Z_n)$ *are jointly independent of* $T_{2:n}$.

REMARK 1. Clearly [see (1.9)] the event $\{N_\rho = n\}$ is determined only by $(Z_{m_0}, \dots, Z_n)$ and therefore by Theorem 1 it is independent of $\overline{T}_{2:n}$.

THEOREM 2. *Let* N_ρ *be the stopping time defined in* (1.9). *Then for all* $\theta \in \Theta$ *the following hold*:

(a) $\lim_{\rho \to \infty} N_\rho / n_0 = 1$ *w.p.*1.
(b) $\lim_{\rho \to \infty} E(N_\rho / n_0) = 1$.

The next result provides the asymptotic normality of N_ρ (appropriately standardized), as the cost of the sampling error relative to the cost of one observation tends to infinity. It is understood that the two statements $\{\rho \to \infty\}$ and $\{n_0 \to \infty\}$ are equivalent.

PROPOSITION 1. *As* $n_0 \to \infty$,

$$N_\rho^* \equiv \frac{(N_\rho - n_0)}{\sqrt{n_0}} \to_{\mathscr{D}} \mathscr{N}\left(0, \frac{1}{4\theta_1^2 G''(\theta_1)}\right). \tag{2.1}$$

As was shown by Starr (1966) and by Woodroofe [(1977), (1982)], the initial sample size m_0 plays a crucial role in an attempt to analyze the *risk* as well as the *regret* associated with N_ρ. It was also shown [see Woodroofe (1977), page 987], that the left tail behavior of the underlying c.d.f. is crucial in the risk's assessments. For the general case discussed here, we impose the following two natural conditions on the model. The first condition is imposed on the function G'. Note that the function G' determines the boundary for the stopping rule N_ρ, as well as the moments of Z_n [see (1.7) and (1.8)]. The second condition is imposed to ensure an appropriate initial sample size m_0.

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

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

Proposition 2 shows the risk efficiency of our procedure under these weak conditions. The condition on β in A.4 will be strengthened subsequently in order to establish the second order results.

PROPOSITION 2. *If G' and m_0 satisfy Assumptions* A.3 *and* A.4, *then*

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

THEOREM 3. *If Assumptions* A.3 *and* A.4 *with $\beta > 3/(2\gamma - 1)$ hold, then as $n_0 \to \infty$,*

$$E(N_\rho) = n_0 + b_0 - \frac{1}{4\theta_1^2 G''(\theta_1)}\left[\frac{1}{2} + 2\theta_1 a_0 G'(\theta_1) - \frac{\theta_1 G'''(\theta_1)}{G''(\theta_1)}\right] + o(1),$$

where b_0 is

$$b_0 = \frac{1}{2} + \frac{1}{8\theta_1^2 G''(\theta_1)} - \sum_{k=1}^{\infty} \frac{1}{k} E\left(\tilde{S}_k I\left[\tilde{S}_k < 0\right]\right)$$

and $\tilde{S}_k$, $k \geq 1$, are partial sums of i.i.d. r.v.'s. defined in (3.19).

THEOREM 4. *If Assumptions* A.3 *and* A.4 *with $\beta > 5/(2\gamma - 1)$ hold, then*

$$\lim_{\rho\to\infty} \mathscr{R}(\rho, \theta_1) = \frac{1}{4\theta_1^2 G''(\theta_1)}.$$

It is interesting to note that this regret is always positive. This is in contrast to the distribution free case where arbitrary large negative values of the regret are possible [see Martinsek (1983)].

The following are the only examples of distributions known to us which satisfy Assumptions A.1 and A.2.

EXAMPLE 1 [The normal distribution, $N(\mu, \sigma^2)$].

(a) $\theta_1 = -1/2\sigma^2$, $\theta_2 = \mu/\sigma^2$, $\Theta = \mathbb{R}^- \times \mathbb{R}$.
(b) $U_1(X) = X^2$, $U_2(X) = X$, $T_{1:n} = \Sigma_{i=1}^n X_i^2$, T_i^2, $T_{2:n} = \Sigma_{i=1}^n X_i$.
(c) $\mu_2 = -\theta_2/2\theta_1$, $\theta_2 = -2\theta_1\mu_2$.
(d) $\mu_1 = \mu_2^2 - 1/2\theta_1$, $\psi(\mu_2) = \mu_2^2$, $G'(\theta_1) = -1/2\theta_1$.
(e) $Z_n = T_{1:n} - n\psi(\overline{T}_{2:n}) = \Sigma_{i=1}^n (X_i - \overline{X}_n)^2 > 0$ a.s.

EXAMPLE 2 [The gamma distribution, $\mathscr{G}(\alpha, \lambda)$].

(a) $\theta_1 = \alpha$, $\theta_2 = -\lambda$, $\Theta = \mathbb{R}^+ \times \mathbb{R}^-$.
(b) $U_1(X) = \log(X)$, $U_2(X) = X$, $T_{1:n} = \Sigma_{i=1}^n \log(X_i)$, $T_{2:n} = \Sigma_{i=1}^n X_i$.
(c) $\mu_2 = \alpha/\lambda$, $\psi(\mu_2) = \log(\mu_2)$, $G'(\alpha) = \Gamma'(\alpha)/\Gamma(\alpha) - \log(\alpha)$.
(d) $Z_n = T_{1:n} - n\psi(\overline{T}_{2:n}) = \Sigma_{i=1}^n \log(X_i/\overline{X}_n) < 0$ a.s.

EXAMPLE 3 (Inverse Gaussian distribution).

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

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

(c) $\mu_2 = -(\theta_1/\theta_2)^{1/2}$, $\psi(\mu_2) = 1/\mu_2$, $G'(\theta_1) = -1/2\theta_1$.

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

REMARK 2. (a) In all of these cases explicit expressions for the statistic Z_n and the function G are available. It can be shown that in these three cases Assumption A.3 holds with $\gamma = 1$ and Assumption A.4 is satisfied with $m_0 > 1 + 2\beta$. Hence, the second order result of Theorem 4 requires an initial sample size $m_0 \geq 12$. For specific cases the initial sample size may be reduced. For instance Woodroofe (1977) has shown that in the normal case $m_0 \geq 6$ suffices.

(b) It is interesting to observe that for the normal case and for the inverse Gaussian case the asymptotic regret $[4\theta_1^2 G''(\theta_1)]^{-1}$ is independent of θ_1 and is equal to $1/2$. Thus, it is easily verified that in these two cases our estimation procedure has *asymptotic local minimax regret* in the sense of Woodroofe [(1985), page 678].

3. Auxiliary results and proofs. In this section we provide the proofs of the main results along with some auxiliary results needed in the sequel. We begin with the proof of Theorem 1 which requires only the validity of Assumptions A.1 and A.2.

PROOF OF THEOREM 1. By Basu's theorem, it suffices to show that the distribution of $(Z_2, \dots, Z_n)$ does not depend on θ_2. To that end, we let $M_Z(\alpha)$, $\alpha = (\alpha_2, \dots, \alpha_n)$ denote the (joint) moment generating function of $Z = (Z_2, \dots, Z_n)$,

$$M_Z(\alpha) = E_{\theta_1, \theta_2}\left(\exp\left\{\sum_{j=2}^{n} \alpha_j Z_j\right\}\right). \tag{3.1}$$

We will show by induction, that (3.1) is independent of θ_2. This is carried out in two steps.

STEP 1. Show that $(Z_{j-1}, Z_j) \perp T_{2:j}$ for all $2 < j \leq n$.

STEP 2. Show that if $(Z_2, \dots, Z_k) \perp T_{2:k}$ for all $k \leq n-1$, then $(Z_2, \dots, Z_n) \perp T_{2:n}$.

PROOF OF STEP 1. Let $j > 2$ and $E^j_{\theta_1, \theta_2}(\cdot)$ denote expectation under the (product) probability measure of $X_1, \dots, X_j$. Then by (1.1) and (1.6),

$$\begin{aligned} &E^j_{\theta_1,\theta_2}\big(\exp\{\alpha_{j-1}Z_{j-1} + \alpha_j Z_j\}\big) \\ &\quad = \exp\{jc(\theta_1,\theta_2) - jc(\theta_1+\alpha_j,\theta_2)\} \\ &\qquad \times E^j_{\theta_1+\alpha_j,\theta_2}\Big(\exp\{\alpha_{j-1}Z_{j-1}\} \times \exp\Big\{-j\alpha_j\psi\big(\overline{T}_{2:j}\big)\Big\}\Big). \end{aligned} \tag{3.2}$$

By Theorem 3.2(iii) of BLR, $Z_{j-1} \perp T_{j-1}$ and hence $Z_{j-1} \perp T_{2:j-1} + U_2(X_j) \equiv T_{2:j}$. Hence, the expectation on the r.h.s. of (3.2) factors and the first expectation term does not involve $T_{2:j}$ nor θ_2 and therefore we can write

$$E^j_{\theta_1+\alpha_j,\theta_2}\big(\exp\{\alpha_{j-1}Z_{j-1}\}\big) \equiv Q_j(\theta_1+\alpha_j), \quad \text{say}. \tag{3.3}$$

For the evaluation of the second expectation term in the r.h.s. of (3.2) we use Theorems 3.1 and 3.2(i) of BLR, to obtain that under Assumptions A.1 and A.2,

$$E^j_{\theta_1+\alpha_j,\theta_2}\Big(\exp\Big\{-j\alpha_j\psi\big(\overline{T}_{2:j}\big)\Big\}\Big) = \exp\Big\{c\big(j(\theta_1+\alpha_j), j\theta_2\big) - c\big(j\theta_1, j\theta_2\big)\Big\}. \tag{3.4}$$

By substitution of (3.3) and (3.4) in (3.2) we get

$$\begin{aligned} E^j_{\theta_1,\theta_2}\big(\exp\{\alpha_{j-1}Z_{j-1} + \alpha_j Z_j\}\big) &= Q_j(\theta_1+\alpha_j)\exp\{jc(\theta_1,\theta_2) - jc(\theta_1+\alpha_j,\theta_2)\} \\ &\quad \times \exp\Big\{c\big(j(\theta_1+\alpha_j), j\theta_2\big) - c\big(j\theta_1, j\theta_2\big)\Big\}. \end{aligned}$$

Finally, we let μ_2, μ_2^* denote the mean of U_2 under (θ_1, θ_2) and $(\theta_1 + \alpha_j, \theta_2)$ respectively. By Theorem 3.1 of BLR (under A.1),

$$\begin{aligned} \mu_2 &\equiv \mu_2(\theta_1,\theta_2) = \mu_2(j\theta_1, j\theta_2), \\ \mu_2^* &\equiv \mu_2(\theta_1+\alpha_j,\theta_2) = \mu_2\big(j(\theta_1+\alpha_j), j\theta_2\big). \end{aligned} \tag{3.5}$$

Then by using the identities (3.5) in conjunction with relation (1.5) of the $c(\cdot)$ function, it is easy to verify that

$$E^j_{\theta_1,\theta_2}\big(\exp\{\alpha_{j-1}Z_{j-1} + \alpha_j Z_j\}\big) = Q_j(\theta_1+\alpha_j)\exp\{H_j(\theta_1+\alpha_j) - H_j(\theta_1)\},$$

where $H_j(\cdot)$ is defined in (1.8). This completes the proof of Step 1. The proof of Step 2 is similar and therefore omitted. □

The following Lemmas are needed for the proof of our main results.

LEMMA 1. *If $\Theta_1 \subset \mathbb{R}^-$ (if $\Theta_1 \subset \mathbb{R}^+$), then:*

(a) *ψ is strictly convex (concave) function on $\mathscr{N}_2$.*
(b) *$Z_1 = 0$ and $Z_n > Z_{n-1}$ a.s. ($Z_n < Z_{n-1}$ a.s.).*
(c) *G' is positive (negative) on Θ_1.*

PROOF. The proof is similar to that of Lemma 4.2 in BLR and is therefore omitted. □

LEMMA 2. *For each $\theta_1 \in \Theta_1$, as $n \to \infty$,*

(a) $\bar{Z}_n \equiv Z_n/n \to G'(\theta_1)$ *a.s.*
(b) $\sqrt{n}\,(\bar{Z}_n - G'(\theta_1)) \to_{\mathscr{D}} N(0, G''(\theta_1))$.

PROOF. (a) follows by the strong law of large numbers and (1.5). To prove (b) we expand $\psi(\bar{T}_{2:n})$ about μ_2 to get

$$\psi(\bar{T}_{2:n}) = \psi(\mu_2) + (\bar{T}_{2:n} - \mu_2)\psi'(\mu_2) + \xi_n/n,$$

where $\xi_n = n(\bar{T}_{2:n} - \mu_2)^2\psi''(\mu_n)/2$ and μ_n satisfies $|\mu_n - \mu_2| \le |\bar{T}_{2:n} - \mu_2|$. Accordingly, Z_n can be rewritten as

$$Z_n = \sum_{j=1}^{n} Y_j - \xi_n, \tag{3.6}$$

where

$$Y_j = U_1(X_j) - \psi'(\mu_2)U_2(X_j) + (\psi'(\mu_2)\mu_2 - \psi(\mu_2)), \quad j = 1,\dots,n. \tag{3.7}$$

Clearly $Y_1,\dots,Y_n$ are i.i.d. r.v.'s. Using (1.2)–(1.5) it follows that $E(Y_1) = G'(\theta_1)$ and $\mathrm{Var}(Y_1) = G''(\theta_1)$. Next, observe that $\psi''(\mu_n) \xrightarrow{\text{a.s.}} \psi''(\mu_2)$ as $n \to \infty$ and that $\sqrt{n}\,(\bar{T}_{2:n} - \mu_2)^2 \overset{\mathscr{P}}{\to} 0$ so that $\xi_n/\sqrt{n} \overset{\mathscr{P}}{\to} 0$, as $n \to \infty$. The proof of (b) now follows from Slutsky's Theorem and the C.L.T. □

REMARK 3. It can be verified that the sequence ξ_n in (3.6) satisfies conditions 4.1 and 4.2 of Woodroofe (1982) [see Example 4.1(ii) there] and thus is said to be *slowly changing*.

LEMMA 3. *Let $n_0 = (\rho/|\theta_1|)^{1/2}$ and $\varepsilon > 1$ be fixed. Then for all $n > n_0\varepsilon$ there exists a constant C depending on ε and G such that*

$$P(N_\rho > n) \le P\left(a_n Z_n > nG'\left(\frac{-\rho}{n^2}\right)\right) \le \exp\{-(n - n_0)C\}.$$

PROOF. The first inequality follows from the definition (1.9) of N_ρ. To verify the second inequality, let $\varepsilon_n = (n_0/n)^2 < 1$ and $t_n = \theta_1(\varepsilon_n - 1)$. Clearly $t_n \in [0, -\theta_1)$. By Markov inequality and (1.7),

$$P(a_n Z_n > nG'(\theta_1\varepsilon_n)) \le e^{-t_n a_n^{-1} G'(\theta_1\varepsilon_n)} M_{Z_n}(t_n) \equiv \exp\{\varphi_n(t_n) + nt_n\beta_n G'(\theta_1\varepsilon_n)\},$$

where we have put $\varphi_n(t) = H_n(t + \theta_1) - H_n(\theta_1) - ntG'(\theta_1\varepsilon_n)$ and $\beta_n = a_0/n + o(1/n)$. By using the definition of $H_n(\cdot)$ [see (1.8)],

$$\begin{aligned} \varphi_n(t_n) &= H_n(t_n + \theta_1) - H_n(\theta_1) - t_n nG'(\theta\varepsilon_n) \\ &= n[G(\theta_1\varepsilon_n) - G(\theta_1)] - [G(n\theta_1\varepsilon_n) - G(n\theta_1)] \\ &\quad + \theta_1(1 - \varepsilon_n)nG'(\theta_1\varepsilon_n). \end{aligned} \tag{3.8}$$

Since $G(n\theta_1\varepsilon_n) - G(n\theta_1) > 0$, and $G''(\cdot) > 0$, the last equality in (3.8) implies that for some ε_n^* between 1 and ε_n,

$$\varphi_n(t_n) \leq -n\theta_1^2(1-\varepsilon_n)^2 G''(\theta_1\varepsilon_n^*)/2. \tag{3.9}$$

Note that there exists a constant $C_0 > 0$ such that $G''(x) \geq C_0$ for all $x \in [\theta_1, 0]$. Since $n > n_0\varepsilon$ we have

$$(1-\varepsilon_n)^2 \geq (1-1/\varepsilon)(1-\varepsilon_n) \geq (1-1/\varepsilon)(1-n_0/n). \tag{3.10}$$

It is also easy to see that for arbitrarily small C_1 (positive) and large n_0,

$$nt_n\beta_n G'(\theta_1\varepsilon_n) \leq (n-n_0)C_1. \tag{3.11}$$

The lemma then follows by combining (3.9)–(3.11). □

Proof of Theorem 2. (a) We make use of Lemma 1(b) along with the definition (1.9) of N_ρ to obtain the inequalities

$$(N_\rho - 1)G'\left(\frac{-\rho}{(N_\rho-1)^2}\right) \leq Z_{N_\rho-1}a_{N_\rho-1} < Z_{N_\rho}a_{N_\rho-1} < N_\rho G'\left(\frac{-\rho}{N_\rho^2}\right)\frac{a_{N_\rho-1}}{a_{N_\rho}}, \quad \text{a.s.}$$

Since $a_{N_\rho} \to 1$ a.s. and $Z_{N_\rho}a_{N_\rho}/N_\rho \to G'(\theta_1)$ a.s. as $\rho \to \infty$, it follows that $\lim_{\rho\to\infty} G'(-\rho/N_\rho^2) = G'(\theta_1)$. By using the relation $-\rho = \theta_1 n_0^2$ the required result follows.

(b) For $\varepsilon > 1$,

$$\begin{aligned} m_0 \leq E(N_\rho) &\leq n_0 + \sum_{n=n_0+1}^{\infty} P(N_\rho > n) \\ &\leq n_0 + (n_0+1)(\varepsilon-1) + \sum_{n=K}^{\infty} P\left(Z_n a_n > nG'\left(\frac{-\rho}{n^2}\right)\right), \end{aligned} \tag{3.12}$$

where $K = [(n_0+1)\varepsilon] + 1$ and $[x]$ denotes the integer part of x. By Lemma 3, for all $n \geq K$ and for some constant $C > 0$,

$$P\left(Z_n a_n > nG'\left(\frac{-\rho}{n^2}\right)\right) \leq e^{-(n-n_0)C}.$$

Hence, the last inequality in (3.12) implies that

$$\begin{aligned} E(N_\rho) &\leq n_0 + (n_0+1)(\varepsilon-1) + \sum_{n=K}^{\infty} e^{-(n-n_0)C} \\ &\leq n_0 + (n_0+1)(\varepsilon-1) + \frac{e^{-Cn_0(\varepsilon-1)}}{1-e^{-C}} < \infty. \end{aligned} \tag{3.13}$$

From part (a) and Fatou's Lemma $\liminf_{\rho\to\infty} E(N_\rho/n_0) \geq 1$.

By (3.13), $\limsup_{\rho\to\infty} E(N_\rho/n_0) \leq 1 + (\varepsilon - 1)$. Finally by letting $\varepsilon \to 1$ the proof of (b) is complete. □

LEMMA 4. *If G' and m_0 satisfy Assumptions* A.3 *and* A.4, *then*

(a) $$E\left(\frac{n_0}{N_\rho} I[N_\rho \le n_0/2]\right) \to 0 \quad \text{as } \rho \to \infty.$$

If in addition Assumption A.4 *holds with* $\beta > 3/(2\gamma - 1)$, *then*

(b) $$n_0 P(N_\rho \le n_0/2) \to 0 \quad \text{as } \rho \to \infty$$

and

(c) $$E\left(\left(\frac{n_0}{N_\rho}\right)^2 I[N_\rho \le n_0/2]\right) \to 0 \quad \text{as } \rho \to \infty.$$

PROOF. We will only prove the first part since the proofs of the other parts are similar. Let $1/2 < \alpha < 1$ be fixed (to be chosen later) and let C be a generic constant. Then

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

For the first term I_1,

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

where the last two inequalities were obtained using Assumptions A.3 and A.4 (with $\beta > 2/(2\gamma - 1)$ and Lemma 1(b). Accordingly,

$$I_1 \le E\left(Z_{m_0}^{-\beta}\right) C n_0^{(1 - 2\gamma\beta + \alpha\beta(1+2\gamma))} \to 0$$

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

Clearly, $I_2 \le n_0^{(1-\alpha)} P[Z_k a_k < k G'(\theta_1 (n_0/k)^2)$, for some $k \in (n_0^\alpha, n_0/2]]$. Define $L_{1:k} = T_{1:k} - k\mu_1$, $L_{2:k} = \varphi(\overline{T}_{2:k}) - \varphi(\mu_2)$. Note that $Z_k = L_{1:k} - k L_{2:k}$

and by (1.5), $\mu_1 - \varphi(\mu_2) = G'(\theta_1)$. Hence

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

with $\Delta_k = a_k[G'(\theta_1(n_0/k)^2) - G'(\theta_1)]$. Since $G'(\cdot)$ is increasing and $k < n_0/2$,

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

Thus,

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

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

$$\begin{aligned} I_{21} &\le n_0^{(1-\alpha)} P\left[\max_{n_0^\alpha < k \le n_0/2} |L_{1:k}| > n_0^\alpha \varepsilon \right] \\ &\le n_0^{(1-\alpha)} n_0^{-\alpha r} \varepsilon^r E\left(\left(L_{1:[n_0/2]}\right)^r\right), \qquad r > 0 \\ &= O\left(n_0^{1-\alpha-r(1/2-\alpha)}\right). \end{aligned} \tag{3.14}$$

For the second term I_{22}, it follows by the continuity of $\varphi(\cdot)$ that there is $\delta(\varepsilon) > 0$ such that $|x - \mu_2| < \delta(\varepsilon) \Rightarrow |\varphi(x) - \varphi(\mu_2)| < \varepsilon$. Thus, by using submartingale inequality as in (3.14),

$$\begin{aligned} I_{22} &\le n_0^{(1-\alpha)} P\left[|\overline{T}_{2:k} - \mu_2| > \delta(\varepsilon) \text{ for some } k \in \left(n_0^\alpha, \frac{n_0}{2}\right]\right] \\ &\le n_0^{(1-\alpha)} P\left[|T_{2:k} - k\mu_2| > n_0^\alpha \delta(\varepsilon) \text{ for some } k \in \left(n_0^\alpha, \frac{n_0}{2}\right]\right] \\ &= O\left(n_0^{1-\alpha+r(1/2-\alpha)}\right). \end{aligned} \tag{3.15}$$

Finally, by combining (3.14) and (3.15) and choosing r large and $\alpha > 1/2$, we obtain

$$I_2 \le O\left(n_0^{1-\alpha+r(1/2-\alpha)}\right) \to 0.$$

Finally, upon choosing $1/2 < \alpha < (2\gamma\beta - 1)/[\beta(1 + 2\gamma)]$ with $\beta > 2/(2\gamma - 1)$ as required, the proof is complete. □

REMARK 4. By using the same arguments as in the proof of Lemma 4(a), it can be easily verified that for $k \geq 1$,

$$n_0^k P(N_\rho \leq n_0/2) \to 0 \quad \text{as } \rho \to \infty$$

provided that the condition on β in A.4 is replaced with $\beta > (1 + 2k)/(2\gamma - 1)$.

PROOF OF PROPOSITION 2. By definition

$$R_\rho(N_\rho) = E\Big(\rho|\psi''(\mu_2)|\big(\overline{T}_{2:N_\rho} - \mu_2\big)^2 + N_\rho\Big).$$

Thus

$$R_\rho(N_\rho) = E\left(\frac{\rho}{N_\rho|\theta_1|}\right) + E(N_\rho)$$

$$= E\left(\frac{n_0^2}{N_\rho}\right) + E(N_\rho)$$

and therefore

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

In view of Theorem 2 it suffices to show that $\limsup_{\rho\to\infty} E(n_0/N_\rho) \leq 1$. Fix $0 < \varepsilon < 1/2$. Then

$$\begin{aligned} E\left(\frac{n_0}{N_\rho}\right) &= E\left(\frac{n_0}{N_\rho}I[N_\rho \leq n_0/2]\right) + E\left(\frac{n_0}{N_\rho}I\left[\frac{n_0}{2} \leq N_\rho < n_0(1-\varepsilon)\right]\right) \\ &\quad + E\left(\frac{n_0}{N_\rho}I\big[n_0(1-\varepsilon) \leq N_\rho < n_0(1+\varepsilon)\big]\right) \\ &\quad + E\left(\frac{n_0}{N_\rho}I\big[N_\rho \geq n_0(1+\varepsilon)\big]\right) \\ &= B_1 + B_2 + B_3 + B_4, \quad \text{say.} \end{aligned}$$

By Lemma 4(a), $B_1 \to 0$ as $\rho \to \infty$.

As immediate consequences of Theorem 2 and Lemma 3,

$$B_2 \leq 2P\left(\frac{1}{2} \leq \frac{N_\rho}{n_0} < 1 - \varepsilon\right) \to 0,$$

$$B_4 \leq \frac{1}{(1+\varepsilon)}P\left(\frac{N_\rho}{n_0} > 1 + \varepsilon\right) \to 0.$$

By using the dominated convergence theorem $B_3 \to 1$ as $\rho \to \infty$. This completes the proof. □

To derive additional properties of the stopping variable N_ρ, we will present its definition (1.9) in a different form. Since $G'(\cdot)$ is monotone increasing on Θ_1, we rewrite N_ρ as

$$N_\rho = \inf\left\{n \geq m_0;\, n\left(-g\left(\bar{Z}_n a_n\right)\right)^{1/2} > \rho^{1/2}\right\}, \tag{3.16}$$

with $g(u) = G'^{-1}(u)$. By using the relation $g(G'(\theta_1)) = \theta_1$ and Taylor's expansion of g about $G'(\theta_1)$ we obtain

$$\begin{aligned} n\left(-g\left(\bar{Z}_n a_n\right)\right)^{1/2} &= n(-\theta_1)^{1/2} - \frac{n\left(\bar{Z}_n a_n - G'(\theta_1)\right)}{2(-\theta_1)^{1/2} G''(\theta_1)} \\ &\quad + \frac{n\left(\bar{Z}_n a_n - G'(\theta_1)\right)^2}{2} Q(\gamma_n), \end{aligned} \tag{3.17}$$

where $Q(\gamma_n) = d^2[(-g(\theta)^{1/2}]/d\theta^2|_{\theta=\gamma_n}$ and γ_n satisfies $|\gamma_n - G'(\theta_1)| \leq |Z_n a_n - G'(\theta_1)|$. Using this and expression (3.6) for Z_n in (3.17) we immediately obtain

$$\frac{n\left(-g\left(\bar{Z}_n a_n\right)\right)^{1/2}}{(-\theta_1)^{1/2}} \equiv \tilde{Z}_n = \tilde{S}_n + \tilde{\xi}_n, \tag{3.18}$$

where with $\xi_n = n(\bar{T}_{2:n} - \mu_2)^2 \psi''(\mu_n)/2$ as in (3.6) and Y_i as in (3.7) and with $\tau(\theta_1) = -2\theta_1 G''(\theta_1)$,

$$\begin{aligned} \tilde{S}_n &= \sum_{i=1}^{n} \tilde{Y}_i, \qquad \tilde{Y}_i = 1 - \frac{(Y_i - G'(\theta_1))}{\tau(\theta_1)} \qquad i \geq 1, \\ \tilde{\xi}_n &= \frac{\xi_n}{\tau(\theta_1)} - \frac{\bar{Z}_n(a_0 + n\delta_n)}{\tau(\theta_1)} + \frac{n\left(\bar{Z}_n a_n - G'(\theta_1)\right)^2}{2(-\theta_1)^{1/2}} Q(\gamma_n). \end{aligned} \tag{3.19}$$

So that by (3.16), (3.18) [and with $\rho^{1/2} = n_0(-\theta_1)^{1/2}$],

$$N_\rho = \inf\left\{n \geq m_0;\, \tilde{Z}_n > n_0\right\}. \tag{3.20}$$

Clearly $\tilde{S}_n$, $n \geq 1$, are partial sums of i.i.d. r.v.'s. with $E(\tilde{Y}_i) = 1$ and $V(\tilde{Y}_i) = G''(\theta_1)/\tau^2(\theta_1)$. Also, by following Example 4.1(ii) and Lemma 1.4 in Woodroofe (1982), it is easily seen that $\tilde{\xi}_n$ are slowly changing. Further, in view of Lemma 2 and the independence of $T_{2:n}$ and Z_n, it follows that $\tilde{\xi}_n \to_{\mathscr{D}} V$ as $n \to \infty$, where

$$V = \frac{1}{\tau^2(\theta_1)}\left[G''(\theta_1)V_1 - \frac{G''(\theta_1)}{2}V_2 - \theta_1 G'''(\theta_1)V_2 - \tau(\theta_1)G'(\theta_1)a_0\right] \tag{3.21}$$

with V_1 and V_2 being two i.i.d. $\chi^2_{(1)}$ random variables. In particular it follows that $\tilde{\xi}_n/\sqrt{n} \to_{\mathscr{P}} 0$ as $n \to \infty$.

PROOF OF PROPOSITION 1. Since (3.18) and (3.20) hold, $\tilde{\xi}_n/\sqrt{n} \to_{\mathscr{P}} 0$ and $\tilde{\xi}_n$ are slowly changing, the result is an immediate consequence of Lemma 4.2 in Woodroofe (1982). □

Let $\varepsilon > 0$, $A_n = \{|\overline{T}_{1:n} - \mu_1| < \varepsilon$ and $|\overline{T}_{2:n} - \mu_2| < \varepsilon\}$ and $V_n = \tilde{\xi}_n I[A_n]$. In the following lemma we show that conditions 4.10–4.15 in Woodroofe (1982) are satisfied by V_n and the sets A_n, $n \geq 1$. These conditions together with the result of Lemma 4(b) are required to establish Theorem 3.

LEMMA 5. *Let A_n and V_n be as above. Then*

(a) $$\sum_{n=1}^{\infty} P\Big(\bigcup_{k \geq n} A'_k\Big) < \infty.$$

(b) $$\sum_{n=1}^{\infty} P(V_n < -n\delta) < \infty, \quad \textit{for some } \delta, 0 < \delta < 1.$$

(c) $$\max_{0 \leq k \leq n} |V_{n+k}|, n \geq 1 \textit{ is uniformly integrable}.$$

(d) $$V_n \to_{\mathscr{D}} V \quad \textit{as } n \to \infty.$$

PROOF. Since $I[A_n] \to 1$ w.p.1. and $\tilde{\xi}_n \to_{\mathscr{D}} V$, Lemma 5(d) follows trivially. Since ψ is twice differentiable, for a fixed μ_2,

$$\psi(x) - \psi(\mu_2) = (x - \mu_2)\psi'(x^*)$$

for some intermediate point x^* between x and μ_2. On $|x - \mu_2| < \varepsilon$, $|\psi'(x^*)| < c(\varepsilon)$ for some constant $c(\varepsilon)$. Also on A_n the functions ψ and Q are bounded. Thus

$$|V_n| \leq c_1 n\big(\overline{T}_{1:n} - \mu_1\big)^2 + c_2 n\big(\overline{T}_{2:n} - \mu_2\big)^2 + c_3,$$

for some constants c_i, $i = 1, \dots, 3$ depending only on ε, ψ and Q. Now Lemmas 5(b) and (c) follow as in Example 4.3 of Woodroofe (1982) and relation (2.14) there. To prove Lemma 5(a) note that

$$\sum_{n=1}^{\infty} P\Big(\bigcup_{k \geq n} A'_k\Big) \leq \sum_{n=1}^{\infty} P\Big(\max_{k \geq n}|\overline{T}_{1:k} - \mu_1| > \frac{\epsilon}{2}\Big)$$

$$+ \sum_{n=1}^{\infty} P\Big(\max_{k \geq n}|\overline{T}_{2:k} - \mu_2| > \frac{\varepsilon}{2}\Big)$$

and each of these last two sums is finite by applying the reverse submartingale inequality to the sequences $\{|\overline{T}_{i:n} - \mu_i|;\ n \geq 1\}$, $i = 1, 2$.

This completes the proof of the lemma. □

PROOF OF THEOREM 3. The assertion follows immediately from Lemma 4(b), (3.21), Lemma 5 and Theorem 4.5 in Woodroofe (1982). □

LEMMA 6. *Let N_ρ^* be as defined in* (2.1). *If Assumptions* A.3 *and* A.4 *with* $\beta > 3/(2\gamma - 1)$ *hold, then as* $n_0 \to \infty$,

(a) $$E\big(N_\rho^{*2} I[N_\rho \leq n_0/2]\big) + E\big(N_\rho^{*2} I[N_\rho \geq 2n_0]\big) \to 0.$$

(b) $N_\rho^{*2} I[n_0/2 < N_\rho \leq 2n_0]$, $n_0 \geq 1$, *are uniformly integrable and*

$$\lim_{n_0 \to \infty} E\big(N_\rho^{*2}\big) = \big[4\theta_1^2 G''(\theta_1)\big]^{-1}.$$

PROOF. (a) Clearly, on the set $\{N_\rho \leq n_0/2\}$, $N_\rho^{*2} \leq cn_0$, a.s. for some constant c. Therefore by Lemma 4(b) as $n_0 \to \infty$,

$$E\big(N_\rho^{*2} I[N_\rho \leq n_0/2]\big) \leq cn_0 P(N_\rho \leq n_0/2) \to 0.$$

Also, since $N_\rho^{*2} \leq (N_\rho^2/n_0 + n_0)$, we have

$$E\big(N_\rho^{*2} I[N_\rho \geq 2n_0]\big) \leq \frac{1}{n_0} E\big(N_\rho^2 I[N_\rho \geq 2n_0]\big) + n_0 P(N_\rho \geq 2n_0).$$

By using Lemma 3 it follows that

$$E\big(N_\rho^{*2} I[N_\rho \geq 2n_0]\big) \leq \frac{2}{n_0} \sum_{k=2n_0}^{\infty} k e^{-(k-2n_0)c} + n_0 e^{-n_0 c} \to 0.$$

(b) The second assertion is an immediate consequence of the first assertion, Proposition 1 and part (a) above. To prove the first assertion it suffices to show that there exists a function $A(x)$ such that $xA(x)$ is integrable [w.r. to Lebesgue measure on $(0, \infty)$] and

$$P\big(n_0/2 < N_\rho \leq 2n_0,\ |N_\rho^*| > x\big) \leq A(x)$$

for all x and n_0 (sufficiently large).

If $x \geq \sqrt{n_0}/2$, then clearly $P(n_0/2 < N_\rho \leq 2n_0,\ N_\rho^* < -x) = 0$.

If $0 \leq x < \sqrt{n}/2$, then $N_\rho > n_0/2$ and $N_\rho^* < -x$ imply that $n_0/2 < N_\rho \leq n_0 - \sqrt{n_0}x$.

Define $I_{n_0:x} = \{k\colon n_0/2 < k \leq n_0 - \sqrt{n_0}x\}$. Since $a_k > 1$,

$$P\Big(N_\rho > \frac{n_0}{2},\ N_\rho^* < -x\Big) \leq P\Big(Z_k < kG'\Big(\frac{-\rho}{k^2}\Big), \text{ for some } k \in I_{n_0:x}\Big). \tag{3.22}$$

By using expression (1.5) for $G'(\theta_1)$ and (1.6) for Z_k and the fact that G'' is positive and continuous, it follows that the event $Z_k < kG'(-\rho/k^2)$ implies that for sufficiently large x and n_0,

$$T_{1:k} - k\mu_1 - k\big(\psi(\overline{T}_{2:k}) - \psi(\mu_2)\big) \le k\left(G'\left(\frac{-\rho}{k^2}\right) - G'(\theta_1)\right) \le -|\theta_1| c\sqrt{n_0}\, x.$$

Thus for some constants c_1 and c_2, the right-hand side of (3.22) is bounded by

$$\begin{aligned} &P\Big[T_{1:k} - k\mu_1 \le -\sqrt{n_0}\, x c_1, \text{ for some } k \in I_{n_0:x}\Big] \\ &\quad + P\Big[-k\big(\psi(\overline{T}_{2:k}) - \psi(\mu_2)\big) \le -\sqrt{n_0}\, x c_2, \text{ for some } k \in I_{n_0:x}\Big] \\ &\quad \equiv I_1 + I_2. \end{aligned}$$

By the submartingale inequality,

$$I_1 \le P\left[\max_{n_0/2 < k \le n_0} |T_{1:k} - k\mu_1| > \sqrt{n_0}\, x c_1\right] \le Cx^{-4}$$

for some constant C. Since ψ is twice differentiable, for any $\delta > 0$ there exists a constant $\lambda(\delta)$ such that $|x - \mu_2| < \delta$ implies $|\psi(x) - \psi(\mu_2)| < \lambda(\delta)|x - \mu_2|$. Let

$$A = \left\{|\overline{T}_{2:k} - \mu_2| \le \lambda \frac{\sqrt{n_0}\, x}{k}, \text{ for all } k \in I_{n_0:x}\right\}.$$

Notice that $\sqrt{n_0}\, x/k \le 1$ for $k \in I_{n_0:x}$. By choosing λ small, on the set A, for all $k \in I_{n_0:x}$,

$$|\psi(\overline{T}_{2:k}) - \psi(\mu_2)| \le \lambda(1)\lambda \frac{\sqrt{n_0}\, x}{k} < \frac{\sqrt{n_0}\, x c_2}{k}.$$

Again, by using the submartingale inequality (as in I_1),

$$I_2 \le P\left[\max_{n_0/2 < k \le n_0} |T_{2:k} - k\mu_2| > \sqrt{n_0}\, x\beta\right] \le Cx^{-4}.$$

The same bound can be obtained for $P(n_0/2 < N_\rho \le 2n_0,\ N_\rho^* > x)$, $x > 0$, by similar arguments. This completes the proof of the lemma. □

PROOF OF THEOREM 4. Let $R_\rho(N_\rho)$ denote the risk associated with the stopping time N_ρ. Then

$$R_\rho(N_\rho) = \rho|\psi''(\mu_2)| E\left[\left(\overline{T}_{2:N_\rho} - \mu_2\right)^2 + N_\rho\right].$$

By using the relation $n_0^2 = \rho/|\theta_1|$ and (1.4) along with the independence result stated in Theorem 1, we obtain that $R_\rho(N_\rho) = E(n_0^2/N_\rho^* + N_\rho)$. Accordingly the regret $\mathscr{R}$ may be written as

$$\mathscr{R}(\rho, \theta_1) = E\left(\frac{n_0^2}{N_\rho} + N_\rho\right) - 2n_0 = n_0 E\left(u\left(\frac{N_\rho}{n_0}\right) - u(1)\right),$$

where $u(x) = x + 1/x$. By using a second order Taylor series expansion of

$u(x)$ about 1, we obtain that

$$u\left(\frac{N_\rho}{n_0}\right) - u(1) = \left(\frac{N_\rho}{n_0} - 1\right)^2 \left(\frac{1}{b}\right)^3,$$

for some intermediate point b satisfying $|b - 1| \le |N_\rho/n_0 - 1|$, $b \to 1$ a.s. and $b \ge 1/2$ on the set $\{N_\rho > n_0/2\}$. Hence by Lemma 1 and Lemma 6

$$E\left(n_0\left(u\left(\frac{N_\rho}{n_0}\right) - u(1)\right) I\left[N_\rho > n_0/2\right]\right)$$
$$= E\left(N_\rho^{*2}\left(\frac{1}{b}\right)^3 I\left[N_\rho > n_0/2\right]\right) \to \frac{G''(\theta_1)}{\tau^2(\theta_1)}.$$

On the set $\{N_\rho \le n_0/2\}$ with $n_0 > 1$, $0 \le (u(N_\rho/n_0) - u(1)) \le cn_0$, for some constant $c > 0$. It follows that

$$E\left(n_0\left(u\left(\frac{N_\rho}{n_0}\right) - u(1)\right) I\left[N_\rho \le n_0/2\right]\right) \le cn_0^2 P(N_\rho \le n_0/2).$$

By Lemma 4(c) and Remark 4, the right side of the above inequality tends to 0 as $n_0 \to \infty$ provided that $\beta > 5/(2\gamma - 1)$. This completes the proof. □

Acknowledgments. The authors wish to thank the referee for prompt and useful suggestions made on an earlier version of this work. We are grateful to an Associate Editor for his constructive comments and in particular for suggesting an alternative, simpler proof of Theorem 1. Thanks are also due to S. Lalley for stimulating discussions and to A. T. Martinsek for his useful comments.

REFERENCES

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

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

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

Brown, L. D. (1986). *Fundamentals of Statistical Exponential Families* IMS, Hayward, CA.

Chow, Y. S. and Martinsek, A. T. (1982). Bounded regret of a sequential procedure for estimation of the mean. *Ann. Statist.* **10** 909–914.

Chow, Y. S. and Yu, K. F. (1981). The performance of a sequential procedure for the estimation of the mean. *Ann. Statist.* **9** 189–198.

Ghosh, M. and Mukhopadhyay, N. (1979). Sequential point estimation of the mean when the distribution is unspecified. *Comm. Statist. Theory Methods* **8** 637–652.

Lai, T. L. and Siegmund, D. (1977). A nonlinear renewal theory with applications to sequential analysis I. *Ann. Statist.* **5** 946–954.

Lai, T. L. and Siegmund, D. (1979). A nonlinear renewal theory with applications to sequential analysis II. *Ann. Statist.* **7** 60–76.

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

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

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

SEN, P. K. and GHOSH, M. (1981). Sequential point estimation of estimable parameters based on U-statistics. *Sankhyā Ser. A* **43** 331–344.

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

STARR, N. and WOODROOFE, M. (1969). Remarks on sequential point estimation. *Proc. Nat. Acad. Sci. U.S.A.* **66** 285–288.

STARR, N. and WOODROOFE, M. (1972). Further remarks on sequential estimation: The exponential case. *Ann. Math. Statist.* **43** 1147–1154.

VARDI, Y. (1979). Asymptotic optimal sequential estimation: the Poisson case. *Ann. Statist.* **7** 1040–1051.

WOODROOFE, M. (1977). Second order approximations for sequential point and interval estimation. *Ann. Statist.* **5** 984–995.

WOODROOFE, M. (1982). *Nonlinear Renewal Theory In Sequential Analysis*. SIAM, Philadelphia.

WOODROOFE, M. (1985). Asymptotic local minimaxity in sequential point estimation. *Ann. Statist.* **13** 676–687.

DIVISION OF THEORETICAL STATISTICS
AND MATHEMATICS
INDIAN STATISTICAL INSTITUTE
203 B. T. ROAD
CALCUTTA 700035
INDIA

DEPARTMENT OF MATHEMATICAL SCIENCES
AT IUPUI
PURDUE UNIVERSITY
1125 EAST 38TH STREET
INDIANAPOLIS, INDIANA 46205-2810