The Annals of Statistics
1996, Vol. 24, No. 4, 1792–1803

ESTIMATION WITH PRESCRIBED PROPORTIONAL ACCURACY FOR A TWO-PARAMETER EXPONENTIAL FAMILY OF DISTRIBUTIONS

BY ARUP BOSE[1] AND BENZION BOUKAI

Indian Statistical Institute and
Indiana University–Purdue University

We propose a sequential procedure for estimating with prescribed proportional accuracy one parameter in a class of two-parameter exponential family of distributions. We study the properties of the resulting stopping time and provide second-order analysis of the coverage probability associated with it by using Edgeworth expansion.

1. Introduction. Let $x_1, x_2, \dots$ be a sequence of independent observations from a model $f(\ ;\theta)$ with $\theta \in \Theta$ being an unknown parameter (possibly a vector) and let μ and σ^2 denote the mean and variance of $f(\ ;\theta)$, respectively. Consider the problem of constructing a sequential procedure for estimating the unknown mean μ which achieves a fixed-proportional accuracy with a preassigned probability. That is, for $\alpha < 1/2$ and $h > 0$, we seek a sequential procedure with a stopping time t such that

$$P_\theta\Big(|\hat{\mu}_t - \mu| \le h\sqrt{\Delta(\theta)}\Big) \approx 1 - 2\alpha \qquad \forall\, \theta \in \Theta, \tag{1.1}$$

where $\hat{\mu}_n$, $n = 1, 2, \dots,$ is the sample estimate of μ and Δ is some proportionality function. Here, $1 - 2\alpha$ is the desired coverage probability and by $\approx$ we mean equality up to terms of $O(h^2)$ as $h \to 0$. When $\Delta \equiv 1$, this procedure leads to a fixed-width confidence interval for μ of the form $\mathscr{C}_t = (\hat{\mu}_t - h, \hat{\mu}_t + h)$. Much of the interest in such a sequential procedure was motivated by Stein's (1945) two-state procedure, the purely sequential procedure of Anscombe (1953) [see also Chow and Robbins (1965) and Starr (1966)] and Hall's (1981) three-stage procedure for fixed-width interval estimation in the normal case with unknown σ^2. In the normal case, the independence of the sample mean and variance (which in turn implies the independence of the event $\{t = n\}$ and $\hat{\mu}_n \equiv \bar{x}_n$) plays a crucial role. It allows a second-order asymptotic expansion of the coverage probability which utilizes the first two moments of the stopping time t [see Woodroofe (1977, 1982)]. These procedures were developed further to include proportional accuracy (in purely sequential and three-stage schemes) by Woodroofe (1987, 1988), who consid-

Received January 1993; revised August 1995.

[1]Research conducted while the author was visiting the Department of Mathematical Sciences at Indiana University–Purdue University at Indianapolis.

AMS 1991 *subject classifications*. Primary 62L12; secondary 62F12.

Key words and phrases. Stopping time, confidence interval, coverage probability, Edgeworth expansion, second order asymptotic.

ered the normal case with known σ^2 and with $\Delta \equiv \Delta(\mu)$ in (1.1). In practice, of course, the unknown Δ is replaced by its appropriate estimate to obtain a confidence interval for μ. Woodroofe (1987) provides a weak expansion of the average coverage probability of such a confidence interval for the normal case. To a great extent, Woodroofe's (1987) work demonstrates the difficulties encountered in providing higher order expansions of the coverage probability in cases lacking the independence property.

In this paper we develop a sequential estimation procedure as described in (1.1), for the following class of two-parameter exponential family of distributions.

Let $\mathscr{F} = \{F_\theta,\ \theta \in \Theta\}$ be a minimal regular exponential family of order 2 characterized by densities of the form

$$(1.2)\quad f(x;\theta) = a(x)\exp\{\theta_1 u_1(x) + \theta_2 u_2(x) + c(\theta)\}, \qquad \theta = (\theta_1, \theta_2) \in \Theta.$$

Here $\Theta = \{\theta \in \mathbb{R}^2;\ e^{-c(\theta)} < \infty\}$ is the natural parameter space. For any $\theta \in \Theta$ the r.v. $\mathbf{u} = (u_1, u_2)$ has moments of all orders. In particular, for $i = 1, 2$, we denote by $\mu_i = -\partial c(\theta)/\partial\theta_i$ and $\sigma_i^2 = -\partial^2 c(\theta)/\partial\theta_i^2$ the mean and variance of u_i, respectively. We further assume that the density (1.2) satisfies the following assumption.

ASSUMPTION A. For some function ψ, $\theta_2 = -\theta_1\psi'(\mu_2)$, where $\psi'(\mu_2) = d\psi(\mu_2)/d\mu_2$ and u_2 is a 1–1 function on the support of (1.2).

The class $\mathscr{F}$ includes the *normal*, *gamma* and *inverse Gaussian* families and was studied in details by Bar-Lev and Reiser (1982) [henceforth referred to as BLR (1982)] in the context of construction of UMPU tests and by Barndorff–Nielsen and Blæsild (1983) for its reproductive properties. With the homeomorphic reparametrization $(\theta_1, \theta_2) \to (\theta_1, \mu_2) \in \Theta_1 \times \mathscr{N}_2$ (varying independently), it can be shown that there exists an infinitely differentiable function G on Θ_1 with $G''(\theta_1) > 0$, such that $\mu_1 = \psi(\mu_2) + G'(\theta_1)$ and

$$(1.3)\qquad \sigma_2^2(\theta) \equiv \partial\mu_2/\partial\theta_2 = \left[|\theta_1|\psi''(\mu_2)\right]^{-1}.$$

By Assumption A, either $\Theta_1 \subset \mathbb{R}^-$ or $\Theta_1 \subset \mathbb{R}^+$ [see BLR (1982) for details], and without loss of generality we assume the former.

Let $x_1, \dots, x_n, \dots,\ n > 1$, be independent observations from (1.2). For each n and $i = 1, 2$, we let $u_{i:n} = \sum_{j=1}^n u_i(x_j)$ and let $\bar{u}_{i:n} = u_{i:n}/n$. The maximum likelihood estimators $\hat{\theta}_{1:n}$ and $\hat{\mu}_2$ of θ_1 and μ_2 satisfy $\hat{\mu}_2 = \bar{u}_{2:n}$ and

$$(1.4)\qquad nG'(\hat{\theta}_{1:n}) = u_{1:n} - n\psi(\bar{u}_{2:n}) \equiv z_n.$$

Bose and Boukai (1993) [henceforth abbreviated here as BB (1993)] established certain second-order results on the properties of a sequential *point estimation* procedure for $\mu_2 \equiv E(u_2)$. It was shown that the stopping time, being based on the MLE $\hat{\theta}_{1:n}$ of the nuisance parameter θ_1, is independent of the terminal estimate for μ_2. In the present paper we apply this independence result to the construction of a sequential estimation procedure for the mean μ_2 which achieves, in similarity to (1.1), prescribed proportional accuracy with a preassigned probability. Following the suggestion of an Associate Editor of BB (1993), we also allow the proportionality function Δ to depend on the nuisance parameter θ_1. More precisely, let q be some positive, twice continuously differentiable and strictly increasing function on $\mathbb{R}^+$ and let

$$\Delta(\theta) \equiv \Delta(\theta_1, \mu_2) = q(|\theta_1|)/|\theta_1|\psi''(\mu_2) \tag{1.5}$$

in (1.1). It may be noted that if $q(x) = x$, then the *length* of the interval is free of θ_1. If in addition ψ'' is a constant, the interval is of fixed width. We further assume that this function satisfies the following condition.

ASSUMPTION B1. For any $\theta_1 \in \Theta_1$ and $0 < x < 1$, q satisfies $xq(|\theta_1|) \le q(x|\theta_1|)$.

With a Δ as in (1.5), it follows from (1.3) and the CLT that the (nonrandom) sample size required to achieve

$$P_\theta\Big(|\bar{u}_{2:n} - \mu_2| \le h\sqrt{\Delta(\theta_1, \mu_2)}\Big) \ge 1 - 2\alpha$$

(asymptotically as $h \to 0$) would have to exceed the nominal sample size

$$a = \eta^2/h^2 q(|\theta_1|), \tag{1.6}$$

where $\eta = \Phi^{-1}(\alpha)$. Here Φ stands for the standard normal distribution whose p.d.f. is denoted by ϕ. Since θ_1 is unknown, we estimate a by using $\hat{\theta}_{1:n}$ in (1.6) and consequently stop sampling as soon as $n \ge \hat{a}$. Accordingly we consider the stopping time

$$\begin{aligned}\tilde{t}_h &= \inf\Big\{n \ge m_0;\, q\big(|\hat{\theta}_{1:n}|\big) > \eta^2/h^2 n\Big\}\\ &= \inf\big\{n \ge m_0;\, z_n < nG'\big(-q^{-1}(\eta^2/h^2 n)\big)\big\},\end{aligned}$$

where the last equality follows from (1.4). In order to reduce bias, we consider a modified stopping rule

$$t_h = \inf\big\{n \ge m_0;\, z_n l_n < nG'\big(-q^{-1}(\eta^2/h^2 n)\big)\big\}, \tag{1.7}$$

where $l_n > 1$ are constants of the form $l_n = 1 + l_0/n + \delta_n$ with $\delta_n = o(1/n)$ as $n \to \infty$. Since G' and q are strictly increasing and $\bar{z}_n \equiv z_n/n$ converges a.s. to $G'(\theta_1)$ (see Lemma 2), it follows that for each fixed h the stopping rule t_h is finite a.s. and $\lim_{h\to 0} t_h = \infty$ a.s. Let $\mathbb{X}_n = \sqrt{n}(\bar{u}_{2:n} - \mu_2)\sqrt{|\theta_1|\psi''(\mu_2)}$. By relations (1.3), (1.5) and (1.6), the coverage probability in (1.1) may be written as

$$\mathscr{P}(h, \theta) \equiv P_\theta\Big(|\bar{u}_{2:t} - \mu_2| \le h\sqrt{\Delta(\theta_1, \mu_2)}\Big) = P_\theta\Big(|\mathbb{X}_{t_h}| \le \eta\sqrt{t_h/a}\Big).$$

The closely related problem of constructing confidence sets for μ_2 can be formulated similarly. The unknown nuisance parameter θ_1 in (1.5) can be estimated by some consistent estimator $\hat{\theta}^*_{1:t}$ in order to obtain such confidence sets. The coverage probability of such a set is

$$\mathscr{P}^*(h,\theta) \equiv P_\theta\Big(|\bar{u}_{2:t} - \mu_2| \le h\sqrt{q\big(|\hat{\theta}^*_{1:t}|\big)/|\hat{\theta}^*_{1:t}|\psi''(\mu_2)}\,\Big). \tag{1.8}$$

Alternatively, both θ_1 and μ_2 can be estimated in (1.5) leading to a confidence interval for μ_2 of the form $\mathscr{C}_{\Delta_t} = (\bar{u}_{2:t} - h\sqrt{\Delta_t},\ \bar{u}_{2:t} + h\sqrt{\Delta_t})$, with $\Delta_t \equiv \Delta(\hat{\theta}^*_{1:t}, \bar{u}_{2:t})$. We discuss these procedures further in the next section. In Section 2 we present the asymptotic properties of the stopping variable t_h (Proposition 2 and Theorems 1 and 2) and provide second-order asymptotic expansion of the coverage probabilities $\mathscr{P}$ and $\mathscr{P}^*$ as the width factor h shrinks to zero (Theorems 3 and 4). Section 3 is devoted to proofs.

2. Main results.

This section contains all the main results of this paper. We provide their proofs separately in Section 3. Throughout, we write $I[\mathscr{A}]$ for the indicator function of the set $\mathscr{A}$.

PROPOSITION 1 [BB (1993)]. *For all $n \ge 2$ and $\theta \in \Theta$, the random variable $I[t_h = n]$ is independent of $\bar{u}_{2:n}$.*

THEOREM 1. *If q satisfies B1, then $\lim_{h\to 0}(t_h/a) = 1$ a.s. and $\lim_{h\to 0} E(t_h/a) = 1$.*

To keep our presentation simple, we strengthen Assumption B1 by the following assumption.

ASSUMPTION B2. $q(x) = x^\lambda$ for some $\lambda \equiv 1/\delta$ with $\delta \ge 1$.

Clearly with such a q, $a = \eta^2/h^2|\theta_1|^\lambda$in (1.6) and t_h in (1.7) takes the form

$$t_h = \inf\Big\{n \ge m_0;\ z_n l_n < nG'\big(\theta_1(a/n)^\delta\big)\Big\}. \tag{2.1}$$

The next result provides the asymptotic normality of t_h as $h \to 0$.

PROPOSITION 2. *Under Assumption B2, $t^*_h \equiv (t_h - a)/\sqrt{a} \to_{\mathscr{D}} \mathscr{N}(0,\tau^2)$ as $h \to 0$, where $\tau^2 \equiv \tau^2(\theta_1) = [\delta^2|\theta_1|^2 G''(\theta_1)]^{-1}$.*

The initial sample size m_0 and the left tail behavior of the underlying c.d.f. play a crucial role in any secondary-order analysis [Woodroofe (1977, 1982)]. We address these issues in the following two lemmas.

LEMMA 1. *Let $s \ge 1$ be fixed. If $G(x) \sim -\frac{1}{2}\log|x|$ as $|x| \to \infty$, then as $h \to 0$,*

(i) $a^s P(t_h \le a/2) \to 0$, *if* $m_0 > 1 + 2s/\delta$,

(ii) $aE\big((a/t_h)^s I[t_h \le a/2]\big) \to 0$, *if* $m_0 > 1 + 2(1+s)/\delta$.

LEMMA 1a. *Let $\delta > 1$ and $s \geq 1$ be fixed. Suppose that m_0 and G satisfy the following set of conditions*:

C1. *for some* $\gamma > 1/\delta$, $\sup_{x \geq 4|\theta_1|} x^{\gamma} G'(-x) \leq M < \infty$.
C2. m_0 *is such that for some* $\beta > 0$, $E_{\theta_1}(z_{m_0}^{-\beta}) < \infty$ (*for all* $\theta_1 \in \Theta_1$).

Then $a^s P(t_h \leq a/2) \to 0$, *if* $\beta > (1+2s)/(\delta\gamma - 1)$, *and* $aE((a/t_h)^s I[t_h \leq a/2]) \to 0$, *if* $\beta > (3+s)/(\delta\gamma - 1)$.

To state the second-order results we use in the sequel the notation

$$v_0 = \tau(\theta_1)\sqrt{G''(\theta_1)}\left[\frac{G'''(\theta_1)}{2(G''(\theta_1))^2} - \frac{l_0 G'(\theta_1)}{G''(\theta_1)}\right]. \tag{2.2}$$

THEOREM 2. *Suppose that m_0 and G satisfy either the conditions of Lemma 1 with $m_0 > 1 + 2/\delta$ or those of Lemma 1a with $\beta > 3/(\delta\gamma - 1)$. Then as $h \to 0$,*

$$E(t_h) = a + \rho - v_0 + \tau^2/2 + o(1),$$

where $\rho = ((1+\tau^2)/2) - \sum_{k=1}^{\infty}(1/k)E(\tilde{S}_k I[\tilde{S}_k < 0])$ *is the expected value of the asymptotic overshoot and* $\tilde{S}_k$, $k \geq 1$, *are defined in* (3.3).

The proof of Theorem 2 is similar to that of Theorem 3 in BB (1993) and therefore is omitted.

THEOREM 3. *Suppose that m_0 and G satisfy either the conditions of Lemma 1 with $m_0 > 1 + 5/\delta$ or those of Lemma 1a with $\beta > 9/2(\delta\gamma - 1)$. Then as $h \to 0$,*

$$\mathscr{P}(h,\theta) = (1 - 2\alpha) + \frac{h^2|\theta_1|^{\lambda}\phi(\eta)}{\eta}\left[\frac{2}{\eta}p_2(\eta) + \rho - v_0 - \frac{\tau^2}{4}(\eta^2 - 1)\right] + o(h^2),$$

where $p_2(\cdot)$ *is the second Edgeworth polynomial*. (*See the proof of Theorem* 3.)

REMARK 1. The three most important classes of distributions that satisfy our conditions are the two-parameter normal distribution $\mathscr{N}(\mu, \sigma^2)$ with $\mu_2 = \mu$, $\theta_1 = -1/2\sigma^2$ and $\psi(\mu_2) = \mu_2^2$; the gamma distribution $\mathscr{G}(\alpha, \lambda)$ with $\mu_2 = \alpha/\lambda$, $\theta_1 = \alpha$ and $\psi(\mu_2) = \log(\mu_2)$; and the inverse Gaussian distribution $\mathscr{IN}(\lambda, \alpha)$ with $\mu_2 = \sqrt{\lambda/\alpha}$, $\theta_1 = -\lambda/2$ and $\psi(\mu_2) = 1/\mu_2$ [see BLR (1982) or BB (1993) for details]. In all these cases $G(x) \sim -\frac{1}{2}\log|x|$ as $|x| \to \infty$. It follows that when $\delta = 1$, Theorem 2 holds with $m_0 \geq 4$ and Theorem 3 holds with $m_0 \geq 7$. This agrees with Woodroofe's (1977) result for the normal distribution case. Note that in some of the case s, l_0 in (2.2) can be chosen so that $\mathscr{P}(h,\theta) \geq (1 - 2\alpha) + o(h^2)$ as $h \to \infty$.

We now turn to the confidence estimation problem. Consider the estimator $\hat{\theta}^*_{1:n}$ of θ_1 which satisfies

$$G'(\hat{\theta}^*_{1:n}) = G'(\hat{\theta}_{1:n}) l_n \equiv \bar{z}_n l_n. \tag{2.3}$$

Clearly, $\hat{\theta}^*_{1:n} \to \theta_1$ a.s., $\hat{\theta}^*_{1:n}$ may be viewed as a bias-corrected estimator for θ_1. By using relations (1.3) and (1.5), we rewrite the coverage probability (1.8) as

$$\mathscr{P}^*(h, \theta) = P_\theta\Big(|\mathbb{X}_{t_h}| \le \eta\sqrt{t_n/a}\,\big(\hat{\theta}^*_{1:t}/\theta_1\big)^{(\lambda-1)/2}\Big). \tag{2.4}$$

The next theorem exhibits the effect that $\hat{\theta}^*_{1:t}$ has on the coverage probability.

THEOREM 4. *Under the conditions of Theorem* 3 *we have as* $h \to 0$,

$$\begin{aligned}\mathscr{P}^*(h, \theta) &= \mathscr{P}(h, \theta) + (1-\delta)\frac{h^2|\theta_1|^\lambda \phi(\eta)}{\eta} \\ &\quad \times \left[v_0 - \frac{\tau^2}{4}(1+\delta)(\eta^2 - 1)\right] + o(h^2),\end{aligned} \tag{2.5}$$

where $\mathscr{P}(h, \theta)$ *is as given in Theorem* 3.

REMARK 2. It is easy to verify that the coverage probability of the confidence interval $\mathscr{C}_{\Delta_t}$, with $\Delta_t = |\hat{\theta}^*_{1:t}|^{\lambda-1}/\psi''(\bar{u}_{2:n})$, may be written as

$$P_\theta\Big(|\bar{u}_{2:t} - \mu_2| \le h\sqrt{\Delta_t}\Big) = P_\theta\Big(\sqrt{t_h}\,|w(\bar{u}_{2:t})| \le \eta\sqrt{t_h/a}\,\big(\hat{\theta}^*_{1:t}/\theta_1\big)^{(\lambda-1)/2}\Big),$$

where $w(x) = (x - \mu_2)[\psi''(x)|\theta_1|]^{1/2}$. It can be shown, by using the same arguments given in the proof of Theorem 4 along with the formal Edgeworth expansion of Bhattacharya and Ghosh (1978) for functions of sample means, that

$$\begin{aligned}&P_\theta\Big(|\bar{u}_{2:t} - \mu_2| \le h\sqrt{\Delta_t}\Big) \\ &\quad = \tilde{P}(h, \theta) + (1-\delta)\frac{h^2|\theta_1|^\lambda \phi(\eta)}{\eta}\left[v_0 - \frac{\tau^2}{4}(1+\delta)(\eta^2 - 1)\right] + o(h^2),\end{aligned}$$

where $\tilde{\mathscr{P}}(h, \theta)$ is as given in Theorem 3 but with a different second Edgeworth polynomial. That new polynomial $\tilde{p}_2(x)$ (say) has coefficients which now depend on the moments of (1.2) as well as on the function w. For sake of brevity, we omit the details.

3. Proofs.

We begin with some basic properties of G and z_n.

LEMMA 2 [BB (1993)]. *For each* $\theta_1 \in \Theta_1$, *we have*:

(a) $z_1 = 0$ *and* $z_n > z_{n-1}$ *a.s.*;
(b) G' *is positive on* Θ_1;
(c) $\bar{z}_n \equiv z_n/n \to G'(\theta_1)$ *a.s. as* $n \to \infty$;
(d) $\sqrt{n}(\bar{z}_n - G'(\theta_1)) \to_{\mathscr{D}} N(0, G''(\theta_1))$, *as* $n \to \infty$.

BLR (1982) have shown that the distribution of z_n is a member of the one-parameter exponential family of distributions with moment generating function

$$M_{z_n}(s) = \exp\big(H_n(s+\theta_1) - H_n(\theta_1)\big), \qquad s+\theta_1 \in \Theta_1, \tag{3.1}$$

where for all $\theta_1 \in \Theta_1$, $H_n(\theta_1) = nG(\theta_1) - G(n\Theta_1)$. We will use relation (3.1) repeatedly in the proofs to follow. For later use, we also note that $z_n = \sum_{j=1}^n Y_j - \xi_n$, where [see BB (1993)] $Y_1, \dots, Y_n$ are i.i.d. r.v.s. with $E(Y_1) = G'(\theta_1)$, $\mathrm{Var}(Y_1) = G''(\theta_1)$ and $\xi_n \equiv n(\bar{u}_{2:n} - \mu_2)^2\psi''(\mu_2)/2$ is slowly changing with $\psi''(\mu_n) \to \psi''(\mu_2)$ a.s. Since G' is monotonically increasing on Θ_1, by putting $g(u) = G'^{-1}(u)$, we may rewrite t_h in (2.1) as

$$\begin{aligned} t \equiv t_h &= \inf\Big\{n \ge m_0;\, n\big(-g(\bar{z}_n l_n)\big)^{\lambda} > |\theta_1|^{\lambda} a\Big\} \\ &= \inf\Big\{n \ge m_0;\, \tilde{S}_n + \tilde{\xi}_n > a\Big\}. \end{aligned} \tag{3.2}$$

The last equality in (3.2) was obtained by a Taylor's series expansion of g about $G'(\theta_1)$, which yields $|\theta_1|^{-\lambda} n(-g(\bar{z}_n l_n))^{\lambda} \equiv \tilde{S}_n + \tilde{\xi}_n$, where with ξ_n and Y_i as before,

$$\begin{aligned} \tilde{S}_n &= \sum_{i=1}^n \tilde{Y}_i, \qquad \tilde{Y}_i = 1 - \frac{\lambda(Y_i - G'(\theta_1))}{|\theta_1| G''(\theta_1)}, \qquad i \ge 1, \\ \tilde{\xi}_n &= \frac{\lambda \xi_n}{|\theta_1| G''(\theta_1)} - \frac{\lambda \bar{z}_n (l_0 + n\delta_n)}{|\theta_1| G''(\theta_1)} + \frac{n(\bar{z}_n l_n - G'(\theta_1))^2}{2|\theta_1|^{\lambda}} D(\gamma_n). \end{aligned} \tag{3.3}$$

Here $D(\gamma_n) \equiv (d^2[(-g(\theta)^{1/2}])/d\theta^2|_{\theta=\gamma_n}$ and γ_n satisfies $|\gamma_n - G'(\theta_1)| \le |z_n l_n - G'(\theta_1|$. Note that $E(\tilde{Y}_i) = 1$ and $\mathrm{Var}(\tilde{Y}_i) = \tau^2$. Following Example 4.1(ii) and Lemma 1.4 in Woodroofe (1982) it is easily seen that $\tilde{\xi}_n$ are slowly changing. By Lemma 2 and the independence of $\bar{u}_{2:n}$ and z_n it follows that $\tilde{\xi}_n \to_{\mathscr{D}} V$, where

$$\begin{aligned} V = {} & \frac{\lambda}{2|\theta_1| G''(\theta_1)} \left[\frac{(V_1 - V_2)}{|\theta_1|} + \frac{G'''(\theta_1)}{G''(\theta_1)} V_2 - 2 l_0 G'(\theta_1) \right] \\ & + \frac{\lambda^2}{2|\theta_1|^2 G''(\theta_1)} V_2, \end{aligned} \tag{3.4}$$

with V_1 and V_2 being two i.i.d. $\chi^2_{(1)}$ random variables. Note that $\tilde{\xi}_n/\sqrt{n} \to_{\mathscr{P}} 0$ and that $E(V) = v_0 + \tau^2/2$, where v_0 is as given in (2.2). It can be easily verified that with $\hat{\theta}^*_{1:n}$ as defined in (2.3), the overshoot of t_h in (3.2) is -3.6$R_a \equiv \tilde{S}_t + \tilde{\xi}_t - a = t_h(\hat{\theta}^*_{1:t}/\theta_1)^{\lambda} - a$. We use this fact later toward the proof of Theorem 4.

PROOF OF PROPOSITION 2. Since (3.2) holds, $\tilde{\xi}_n/\sqrt{n} \to_{\mathscr{P}} 0$ and $\tilde{\xi}_n$ are slowly changing, the result follows from Lemma 4.2 in Woodroofe (1982). □

The next lemma is on the right tail behavior of t_h and is analogous to Lemma 3 of BB (1993). There was, however, an oversight in its proof. The proof of Lemma 3 given here serves also as a correct proof to that lemma.

LEMMA 3. *Suppose q satisfies Assumption B1 and let $\varepsilon > 1$ be fixed. Then for all $n > a\varepsilon$, there exists a constant $C > 0$ depending on ε, q and G such that*

$$P(t_h > n) \le P\left(z_n l_n > nG'\left(-q^{-1}\left(\frac{a}{n}q(|\theta_1|)\right)\right)\right) \le \exp\{-C(n-a)\}.$$

PROOF. The first inequality follows directly from (1.7). By Assumption B1,

$$P\left[z_n l_n > nG'\left(-q^{-1}\left(\frac{a}{n}q(|\theta_1|)\right)\right)\right] \le P\left(z_n l_n > nG'\left(\frac{a\theta_1}{n}\right)\right).$$

To verify the second inequality, define $\varepsilon_n = (a/n) < 1$ and let $s > 0$ be small (to be chosen). By Markov's inequality and (3.1),

$$P(z_n l_n > nG'(\theta_1\varepsilon_n)) \le \exp(-snG'(\theta_1\varepsilon_n))M_{z_n}(sl_n) \equiv \exp\{\varphi_n(s)\},$$

where we have put $\varphi_n(s) = H_n(sl_n + \theta_1) - H_n(\theta_1) - snG'(\theta_1\varepsilon_n)$. By using the definition (3.1) of $H_n(\cdot)$, we rewrite $\varphi_n(s)$ as

$$\begin{aligned}\varphi_n(s) &= n[G(\theta_1 + sl_n) - G(\theta_1)] \\ &\quad -[G(n(\theta_1 + sl_n)) - G(n\theta_1)] - snG'(\theta_1\varepsilon_n).\end{aligned} \tag{3.5}$$

Since $G(n(\theta_1 + sl_n)) - G(n\theta_1) > 0$ and $G'' > 0$, (3.5) implies that for some ε_n^* between 1 and ε_n and some θ_1^* between θ_1 and $\theta_1 + sl_n$,

$$\begin{aligned}\varphi_n(s) &\le -ns\theta_1(\varepsilon_n - 1)G''(\theta_1\varepsilon_n^*) \\ &\quad + ns^2 l_n^2 G''(\theta_1^*)/2 + s(l_0 + \delta_n)G'(\theta_1).\end{aligned} \tag{3.6}$$

Note that $G''(x) \ge C_0$ for all $x \in [\theta_1, 0]$ for some constant $C_0 > 0$, and in a small neighborhood of θ_1, G'' is bounded above. Thus for a small s, (3.6) gives $\varphi_n(s) \le -ns\theta_1(\varepsilon_n - 1)C_1$, for some constant $C_1 > 0$ and the lemma follows. □

PROOF OF THEOREM 1. The first assertion follows from Lemma 2 and (1.7). The second assertion follows from Lemma 3 and is similar to Theorem 2 of BB (1993). We omit the details. □

PROOF OF LEMMAS 1 AND 1a. Let $1/2 < \alpha < 1$ be fixed, and let C denote a generic constant. Then for (ii) we have

$$\begin{aligned}aE\left(\left(\frac{a}{t_h}\right)^s I\left[t_h \le \frac{a}{2}\right]\right) &\le aE\left(\left(\frac{a}{t_h}\right)^s I[m_0 \le t_h \le a^\alpha]\right) \\ &\quad + a^{1+s(1-\alpha)}P\left(a^\alpha < t_h \le \frac{a}{2}\right) \\ &= a^{s+1}I_1 + I_2 \quad (\text{say}).\end{aligned}$$

Now, by (2.1),

$$I_1 = \sum_{k=m_0}^{[a^\alpha]} \frac{1}{k^s} P(t_h = k) \le \sum_{k=m_0}^{[a^\alpha]} \frac{1}{k^s} P\left(z_k l_k \le kG'\left(\left(\frac{a}{k} \right)^\delta \theta_1 \right) \right).$$

For $m_0 \le k \le a^\alpha$, let $\varepsilon_k = (a/k)^\delta > 1$, let $\nu = \theta_1(\varepsilon_k - 1)$ and note that $\nu < 0$. Since $l_k > 1$, by Markov's inequality and (3.1),

$$P(z_k l_k < kG'(\theta_1 \varepsilon_k)) \le \exp(-\nu k G'(\theta_1 \varepsilon_k)) M_{z_k}(\nu) \equiv \exp\{\varphi_k(\nu)\},$$

where we have put $\varphi_k(\nu) = H_k(\nu + \theta_1) - H_k(\theta_1) - \nu k G'(\theta_1 \varepsilon_k)$. By (3.1),

$$\varphi_k(\nu) = k[G(\theta_1 \varepsilon_k) - G(\theta_1)] - \nu k G'(\theta_1 \varepsilon_k) - [G(k\theta_1 \varepsilon_k) - G(k\theta_1)].$$

Note that $\sup_k |G(k\theta_1)|/k \le C$ and hence $k[G(k\theta_1)/k - G(\theta_1)] \le kC$. Moreover, since $\inf_k \varepsilon_k \to \infty$ we have, $-G(k\theta_1\varepsilon_k) \sim \frac{1}{2}\log(k) + \frac{1}{2}\log(\varepsilon_k) + \frac{1}{2}\log|\theta_1|$ and $G(\theta_1\varepsilon_k) \sim -\frac{1}{2}\log(\varepsilon_k) - \frac{1}{2}\log|\theta_1|$. It is also easy to verify that $|\nu G'(\theta_1\varepsilon_k)| \le C|\theta_1|$. Hence we obtain

$$\begin{aligned} \varphi_k(\nu) &\le k\left(C - \frac{1}{2}\log(\varepsilon_k) \right) + \frac{1}{2}\log(k) + \frac{1}{2}\log(\varepsilon_k) + \frac{1}{2}\log|\theta_1| \\ &\le -\frac{(k-1)}{2}(C + \log(\varepsilon_k)). \end{aligned}$$

It follows that for any $\varepsilon > 0$, arbitrary small, $P(z_k l_k < kG'(\theta_1\varepsilon_k)) \le (k/a)^{\delta(k-1)/2-\varepsilon}$. Hence, by arguments similar to those given in Woodroofe [(1982), page 107],

$$a^{s+1} I_1 \le a \sum_{k=m_0}^{[a^\alpha]} \left(\frac{k}{a} \right)^{\delta(k-1)/2-\varepsilon-s} \le Ca^{(1+s-\delta(m_0-1)/2+\varepsilon)} \to 0. \tag{3.7}$$

It can be easily shown, using the same arguments as in Lemma 4 in BB (1993), that for some arbitrary large r and $\alpha > 1/2$,

$$I_2 \le O(a^{1+s(1-\alpha)+r(1/2-\alpha)}) \to 0. \tag{3.8}$$

The second part of Lemma 1 is now obtained by combining (3.7) and (3.8). The proof of (i) is similar. Lemma 1a may be proved along the lines of Lemma 4 in BB (1993). The details are omitted. □

The following lemma establishes the uniform integrability of t_h^* as defined in Proposition 2. Its proof is similar to that of Lemma 6 of BB (1993) and is therefore omitted.

LEMMA 4. *Suppose m_0 and G satisfy the conditions of Lemma 1 with $m_0 > 1 + 2/\delta$ or of Lemma 1a with $\beta > 3/(\delta\gamma - 1)$. Then*:

(a) $E(t_h^{*2} I[t_h \le a/2]) + E(t_h^{*2} I[t_h \ge 2a]) \to 0$, *as* $h \to 0$;
(b) $t_h^{*2} I[a/2 < t_h \le 2a]$ *are uniformly integrable and* $\lim_{h\to 0} E(t_h^{*2}) = \tau^2$.

PROOF OF THEOREM 3. As in Section 1, we let $\mathbb{X}_n = \sqrt{n}(\bar{u}_{2:n} - \mu_2) \times \sqrt{|\theta_1|\psi''(\mu_2)}$ and recall that the covrage probability is $\mathscr{P}(h, \theta) \equiv P_\theta(|\mathbb{X}_{t_h}| \le \eta\sqrt{t_h/a})$. By Proposition 1,

$$\mathscr{P}(h, \theta) \equiv \mathscr{P}(h, \theta_1) = E\Big[P_\theta\Big(|\mathbb{X}_{t_h}| \le \eta\sqrt{t_h/a}\Big)\Big], \tag{3.9}$$

where E denotes expectation with respect to t_h. Note that $\mathscr{P}(h, \theta)$ depends only on θ_1. Since $\mathbb{X}_n$ is a partial sum of the i.i.d. r.v.'s $u_j^* = (u_2(x_j) - \mu_2)\sqrt{|\theta_1|\psi''(\mu_2)}$ $(j = 1, \dots, n)$, we obtain by an Edgeworth expansion of the probability in the right side of (3.9),

$$\begin{aligned}\mathscr{P}(h, \theta_1) &= E\big[(2\Phi(\eta_t) - 1) + 2t_h^{-1}p_2(\eta_t)\phi(\eta_t) + t_h^{-2}O(1)\big] \\ &= E_1 + E_2 + E_3 \quad (\text{say}),\end{aligned} \tag{3.10}$$

where $\eta_t \equiv \eta\sqrt{t_h/a}$ and

$$p_2(y) = -y\big[(\kappa_4/24)(y^2 - 3) + \big(\kappa_3^2/72\big)\big(y^4 - 10y^2 + 15\big)\big],$$

with κ_i, $i = 3, 4$, being the ith cumulant of the standardized random variable u_1^*. The $O(1)$ term in (3.10) is bounded uniformly over all sample paths. Hence it immediately follows from Lemma 1 (or Lemma 1a) that $E_3 = o(a^{-1})$.

Let $\Psi(x) = 2\Phi(\sqrt{x}) - 1$ and let Ψ' and Ψ'' be its first and second derivatives. The arguments of Woodroofe [(1982), page 111] together with Lemma 4 yield

$$E_1 = \Psi(\eta^2) + \frac{\eta^2}{a}\Psi'(\eta^2)E(t_h - a) + \frac{\tau^2\eta^4}{2a}\Psi''(\eta^2) + o(a^{-1}). \tag{3.11}$$

Since $p_2(x)\phi(x)$ is bounded and continuous, it follows (via one-step expansion) from Theorem 1 and Lemma 1 (or 1a) that

$$E_2 = E\big[2t_h^{-1}p_2(\eta_t)\phi(\eta_t)\big] = \frac{2}{a}p_2(\eta)\phi(\eta) + o(a^{-1}). \tag{3.12}$$

The proof is completed by combining (3.9)–(3.12) and Theorem 2. □

REMARK 3. A crucial step in the preceding proof is to show that $E[(a/t_h)^{3/2}I[t_h \le a/2]] = o(a^{-1})$, which is guaranteed by Lemma 1 (or 1a). Any other set of conditions which ensures this would yield all results of the present paper.

PROOF OF THEOREM 4. Since z_t is independent of $\bar{u}_{2:t}$ and G' is injective, it follows from (2.3) that $\hat{\theta}_{1:t}^*$ is also independent of $\bar{u}_{2;t}$. Hence, by an Edgeworth expansion (as before), we may rewrite $\mathscr{P}^*$ in (2.4) as

$$\begin{aligned}\mathscr{P}^*(h, \theta) &= E\big[\Psi\big(x_t^2\big) = 2t_h^{-1}p_2(x_t)\phi(x_t) + t_h^{-2}O(1)\big] \\ &= E_1 + E_2 + E_3,\end{aligned} \tag{3.13}$$

where we have put $x_t \equiv \eta\sqrt{t_h/a}\,(\hat{\theta}^*_{1:t}/\theta_1)^{(\lambda-1)/2}$. Note that since the overshoot of t_h is $R_a = t_h(\hat{\theta}^*_{1:t}/\theta_1)^{\lambda} - a$, we may rewrite x_t^2 in (3.13) as $x_t^2 \equiv \eta^2 + \eta^2 r_t$, with

$$r_t = \frac{t_h}{a}\left(\frac{\hat{\theta}^*_{1:t}}{\theta_1}\right)^{(\lambda-1)} - 1 \equiv \left(\frac{t_h}{a}\right)^{\delta}\left(1 + \frac{R_a}{a}\right)^{1-\delta} - 1, \tag{3.14}$$

where $\delta = 1/\lambda$. As in the proof of Theorem 3, we have $E_3 = o(a^{-1})$ and $E_2 = (2/a)p_2(\eta)\phi(\eta) + o(a^{-1})$. To evaluate the term E_1, define

$$\mathscr{A} = \{a/2 \le t_h \le 2a\} \quad \text{and} \quad \mathscr{B} = \left\{\left(\frac{\hat{\theta}^*_{1:t}}{\theta_1}\right)^{\lambda} \le 2\right\}.$$

From Lemma 1 (or 1a) and Lemma 3, $P(\mathscr{A}^c) = o(a^{-1})$ and hence $P(\mathscr{A}^c \cap \mathscr{B}^c) = o(a^{-1})$. Also, by using relation (2.3) and arguments similar to those of Lemma 6 in BB (1993), it can be easily shown that $P(\mathscr{A} \cap \mathscr{B}^c) = o(a^{-1})$. Thus, since Ψ is a bounded function,

$$E\big(\Psi\big(x_t^2\big)I[\mathscr{A}^c \cup \mathscr{B}^c]\big) = o(a^{-1}).$$

On the set $\mathscr{A} \cap \mathscr{B}$, we first expand $\Psi(x_t^2)$ about $\Psi(\eta^2)$ and then utilize relation (3.14) to expand $(t_h/a)^{\delta}$ and $(1 + R_a/a)^{1-\delta}$ about 1. From these expansions, which are omitted for the sake of brevity, it is clear that the asymptotic expansion of $E(\Psi(x_t^2)I[\mathscr{A} \cap \mathscr{B}])$ will be established, provided that $|t_h^*|^4$ and $(R_a/\sqrt{a})^4$ are uniformly integrable on the set $\mathscr{A} \cap \mathscr{B}$. Both of these can indeed be easily established by following the lines of the proof of Lemma 6 of BB (1993). We omit the details. □

Acknowledgments. We gratefully acknowledge the comments of the Editor, Professor M. Woodroofe, an Associate Editor and the referee which have led to a significant improvement in the manuscript. We thank the referee for his suggestion leading to a simpler proof of Theorem 3.

REFERENCES

ANSCOMBE, F. (1953). Sequential estimation. *J. Roy. Statist. Soc. Ser. B* **15** 1–21.

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.

BARNDORFF-NIELSEN, O. and BLÆSILD, P. (1983). Reproductive exponential families. *Ann. Statist.* **11** 770–782.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

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

CHOW, Y. S. and ROBBINS, H. (1965). Asymptotic theory of fixed width confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

HALL, P. (1981). Asymptotic theory of triple sampling for sequential estimation of a normal mean. *Ann. Statist.* **9** 1229–1238.

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

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

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. (1987). Confidence interval with fixed proportional accuracy. *J. Statist. Plann. Inference* **15** 131–146.

WOODROOFE, M. (1988). Fixed proportional accuracy in three stages. In *Statistical Decision Theory and Related Topics IV* (S. S. Gupta and J. O. Berger, eds.) 209–221. Springer, New York.

INDIAN STATISTICAL INSTITUTE
203 B. T. ROAD
CALCUTTA 700035
INDIA

INDIANA UNIVERSITY–PURDUE UNIVERSITY
AT INDIANAPOLIS
402 N. BLACKFORD STREET
INDIANAPOLIS, INDIANA 46202-3216