

ON SEQUENTIAL FIXED-WIDTH CONFIDENCE INTERVALS FOR THE MEAN AND SECOND-ORDER EXPANSIONS OF THE ASSOCIATED COVERAGE PROBABILITIES

NITIS MUKHOPADHYAY AND SUJAY DATTA*

Department of Statistics, University of Connecticut, Storrs, CT 06269, U.S.A.

(Received August 15, 1994; revised February 13, 1995)

Abstract. In order to construct fixed-width ($2d$) confidence intervals for the mean of an unknown distribution function F , a new purely sequential sampling strategy is proposed first. The approach is quite different from the more traditional methodology of Chow and Robbins (1965, *Ann. Math. Statist.*, **36**, 457–462). However, for this new procedure, the coverage probability is shown (Theorem 2.1) to be at least $(1 - \alpha) + Ad^2 + o(d^2)$ as $d \rightarrow 0$ where $(1 - \alpha)$ is the preassigned level of confidence and A is an appropriate functional of F , under some regularity conditions on F . The rates of convergence of the coverage probability to $(1 - \alpha)$ obtained by Csenki (1980, *Scand. Actuar. J.*, 107–111) and Mukhopadhyay (1981, *Comm. Statist. Theory Methods*, **10**, 2231–2244) were merely $O(d^{1/2-q})$, with $0 < q < 1/2$, under the Chow-Robbins stopping time τ^* . It is to be noted that such considerable sharpening of the rate of convergence of the coverage probability is achieved even though the new stopping variable is $O_P(\tau^*)$. An accelerated version of the stopping rule is also provided together with the analogous second-order characteristics. In the end, an example is given for the mean estimation problem of an exponential distribution.

Key words and phrases: Distribution-free, fixed-width confidence intervals, confidence level, second-order expansions, purely sequential, accelerated sequential, Markov inequality.

1. Introduction

Let X_1, X_2, \dots be a sequence of independent and identically distributed (i.i.d.) random variables having an unknown distribution function (d.f.) $F(x)$, $x \in \mathcal{R}$. We write $\theta = \int_{\mathcal{R}} x dF(x)$ and $\sigma^2 = \int_{\mathcal{R}} (x - \theta)^2 dF(x)$ respectively for the mean and variance of the d.f. F . Having obtained X_1, \dots, X_n and given a preassigned number $d(> 0)$, we consider the fixed-width confidence interval $I_n = [\bar{X}_n - d, \bar{X}_n + d]$ for θ where $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$. For large n , $P_F\{\theta \in I_n\} \approx 2\Phi(n^{1/2}d/\sigma) - 1$, which will be at least $(1 - \alpha)$ with a preassigned $\alpha \in (0, 1)$ provided that n is the

* Now at Department of Statistics, University of Michigan, Ann Arbor, MI 48109, U.S.A.

smallest integer $\geq a_0^2 \sigma^2 / d^2 = C$, say, where $\Phi(a_0) = 1 - \frac{1}{2}\alpha$. Here and elsewhere, one writes $\Phi(x) = \int_{-\infty}^x (2\pi)^{-1/2} \exp(-\frac{1}{2}y^2) dy$. Of course, C is unknown since $\sigma^2 = \sigma^2(F)$ is unknown. In fact no fixed-sample-size procedure will provide a fixed-width confidence interval for θ having a prescribed coverage probability at the same time.

Chow and Robbins (1965) proposed the following ingenious purely sequential sampling strategy. One starts with X_1, \dots, X_m where $m (\geq 2)$ is the initial sample size and then proceeds by taking one X at a time according to the stopping rule

$$(1.1) \quad N = N(d) = \inf\{n \geq m : n \geq a_0^2(S_n^2 + n^{-1})/d^2\},$$

where $S_n^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$. Having obtained $X_1, \dots, X_m, \dots, X_N$, one proposes the fixed-width confidence interval $I_N = [\bar{X}_N \pm d]$ for θ . Under the assumption that $0 < \sigma^2 < \infty$, Chow and Robbins (1965) proved that

$$(1.2) \quad \begin{aligned} N/C &\rightarrow 1 \quad \text{a.s.}, & E(N/C) &\rightarrow 1, & \text{and} \\ P\{\theta \in I_N\} &\rightarrow 1 - \alpha, & \text{as } d &\rightarrow 0. \end{aligned}$$

From this point onward, let us write $P(\cdot)$ and $E(\cdot)$ instead of $P_F(\cdot)$ and $E_F(\cdot)$ respectively. Csenki (1980) used the Berry-Esseen rate for the random central limit theorem obtained by Landers and Rogge (1976) to verify that

$$(1.3) \quad P\{\theta \in I_N\} = (1 - \alpha) + O(d^{1/2-q}),$$

for the Chow-Robbins stopping rule (1.1) with $0 < q < \frac{1}{2}$, under appropriate moment conditions on F . Mukhopadhyay (1981) also obtained (1.3) under more economical moment conditions. One is referred to Mukhopadhyay and Vik (1985) for some related details and other references.

When $F(x) = \Phi((x - \theta)/\sigma)$, the result in (1.3) can be strengthened considerably by replacing it with the asymptotic second-order expansion, namely,

$$(1.4) \quad P\{\theta \in I_N\} = (1 - \alpha) + A\sigma^{-2}d^2 + o(d^2),$$

where A is a computable real number, $m \geq 7$, and $(S_n^2 + n^{-1})$ is replaced by S_n^2 in (1.1). Such an expansion was provided by Woodroffe (1977). In the case of two-parameter negative exponential distributions, analogous second-order expansions have been derived. One is referred to Mukhopadhyay (1988) for a review. But, we are not aware of results such as (1.4) when F is unknown.

Our intention in this paper is to propose purely sequential and accelerated sequential stopping times with a generic notation N for which we can claim the following second-order expansion of the coverage probability: As $d \rightarrow 0$,

$$(1.5) \quad P\{\theta \in I_N\} \geq (1 - \alpha) + A^*d^2 + o(d^2),$$

under the assumption that $E[|X_1|^6] < \infty$, where $A^* = A^*(F)$ is an appropriate functional. We obtain such a strong result in spite of the fact that in this

case $N/C \rightarrow k$ a.s. and $E(N/C) \rightarrow k$ as $d \rightarrow 0$ where $k(> 1)$ is a known constant, that is our newly proposed stopping variables N and the Chow-Robbins stopping variable have similar rates of convergence and they are asymptotically in the close proximity of each other in some sense. Section 2 is concerned with these aspects associated with the newly proposed purely sequential stopping rule and the main results are summarized in Theorem 2.1. In Section 3, we introduce the acceleration technique in this setup in order to reduce sampling operations, and yet obtain an asymptotic second-order expansion of the coverage probability that resembles (1.5), among other characteristics (Theorem 3.1). Section 4 deals with the analogous problems for the mean of an exponential distribution from the perspective of second-order asymptotics, briefly indicating the “fine-tuning” ideas proposed earlier in Mukhopadhyay and Datta (1994) (referred to in the sequel as MD (1994)).

2. A purely sequential methodology

Recall that $I_n = [\bar{X}_n \pm d]$. Now, we have

$$P\{\theta \notin I_n\} \leq E[(\bar{X}_n - \theta)^2]/d^2 = \sigma^2(nd^2)^{-1} \leq \alpha,$$

if n is the smallest integer $\geq \sigma^2(\alpha d^2)^{-1} = n_0$, say. We shall pretend that n_0 is the optimal fixed sample size, had σ been known, for our purposes instead of the $C(= a_0^2\sigma^2/d^2)$ in Chow and Robbins (1965) where $\Phi(a_0) = 1 - \frac{1}{2}\alpha$. Note that $n_0/C = (\alpha a_0^2)^{-1}$, which is fixed for all $d(> 0)$, but it is true that this ratio is larger than unity. However, the point is that $n_0 = O(C)$ and we shall mimic this n_0 in proposing the purely sequential sampling scheme for which a second-order expansion such as (1.5) would be provided.

Let $m = m(d) = \max\{2, [(\alpha d^2)^{-1/2}]^* + 1\}$ where $[u]^* =$ largest integer $< u$. Now, one starts with X_1, \dots, X_m and then proceeds by taking one X at a time according to the stopping rule [with $S_n^2 = n^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$]

$$(2.1) \quad N = N(d) = \inf\{n \geq m(d) : n\alpha d^2 \geq S_n^2\}.$$

We have $P(N < \infty) = 1$ for all fixed $\alpha \in (0, 1)$ and $d(> 0)$ when $0 < \sigma < \infty$. In the end, one thus proposes the fixed-width confidence interval $I_N = [\bar{X}_N \pm d]$ for θ . Along the lines of Chow and Robbins (1965), we can show that $N \rightarrow \infty$ a.s., $N/n_0 \rightarrow 1$ a.s., and $E(N/n_0) \rightarrow 1$ as $d \rightarrow 0$.

Let us write $m_i = E[(X_1 - \theta)^i]$, $i = 3, 4$, for the third and fourth central moments of F , and $\eta = E(R)$ where the probability distribution of R is given by

$$(2.2) \quad P(r \leq R \leq r + dr) = E^{-1}(\tau)P\left\{2\tau - \sum_{k=1}^{\tau} \frac{(X_k - \theta)^2}{\sigma^2} > r\right\} dr$$

for $0 < r < \infty$ with $\tau = \inf\{n \geq 1 : 2n - \sum_{k=1}^n \frac{(X_k - \theta)^2}{\sigma^2} > 0\}$. With these notations in mind, we now state and prove the following results.

THEOREM 2.1. *Let $E[|X_1|^6] < \infty$. Then, under the purely sequential procedure (2.1), we have as $d \rightarrow 0$:*

- (i) $E(N - n_0) = \eta - m_4\sigma^{-4} + o(1)$;
 - (ii) $P\{\theta \in I_N\} \geq (1 - \alpha) - \alpha^2\sigma^{-2}d^2\{5 + 6(m_3/\sigma^3)^2 - \eta\} + o(d^2)$;
- if $(X_1 - \theta)^2/\sigma^2$ has a nonarithmetic distribution.*

PROOF. We exploit the results and tools from Aras and Woodroffe (1993), henceforth abbreviated as AW (1993), after verifying their conditions (C1)–(C6). Let us write $\mathbf{Y}_i = (X_i - \theta, (X_i - \theta)^2 - \sigma^2)$, $i = 1, 2, \dots$, so that these \mathbf{Y} 's are i.i.d. having a d.f., say, G whose mean vector is $\mathbf{0}$ and the covariance matrix is $\begin{pmatrix} \sigma^2 & m_3 \\ m_3 & m_4 - \sigma^4 \end{pmatrix}$. We define functions $g_n : R^2 \rightarrow R$ as

$$g_n(u, v) = \begin{cases} \sigma^2/(\sigma^2 + v - u^2) & \text{if } n \leq \sigma^{-2} \\ \sigma^2/\max\{n^{-1}, \sigma^2 + v - u^2\} & \text{if } n > \sigma^{-2}. \end{cases}$$

If we now define $g(u, v) = \sigma^2(\sigma^2 + v - u^2)^{-1}$ for all $(u, v) \in R^2$ such that $\sigma^2 + v - u^2 \neq 0$, then it is easy to verify that g is twice continuously differentiable on some neighborhood of $(0, 0)$, $\mathbf{c} = D(g) |_{(0,0)} = (0, -\sigma^{-2})$, $g_n = g$ for all $n \geq 1$ on some neighborhood of $(0, 0)$, and obviously $g(0, 0) = 1$.

Next, one notes that the stopping variable N from (2.1) has the same representation as in (2) of AW(1993) with $a = n_0$, $\mathbf{c} = (0, -\sigma^{-2})$, $\Sigma_n = \sum_{i=1}^n \mathbf{Y}_i$ and $Z_n = n + \langle \mathbf{c}, \Sigma_n \rangle + \xi_n$ for $n \geq 1$, where $\xi_n = ng_n(\bar{\mathbf{Y}}_n) + \sum_{i=1}^n (X_i - \theta)^2\sigma^{-2} - 2n$, and let $\xi = \frac{1}{2}\langle \mathbf{W}, D^2(g) |_{(0,0)} \mathbf{W} \rangle$ with $\mathbf{W} \sim N_2\left(\mathbf{0}, \begin{pmatrix} \sigma^2 & m_3 \\ m_3 & m_4 - \sigma^4 \end{pmatrix}\right)$, $D^2(g) |_{(0,0)} = \text{Diag}(2\sigma^{-2}, 2\sigma^{-4})$.

As per suggestions from one of the referees, we briefly state conditions (C1)–(C6) from AW (1993) for completeness. For some $3 \leq p < \infty$, and $0 < \varepsilon_0, \varepsilon_1 < 1$, AW (1993) assumes:

- (C1) $\int_{R^2} \langle \mathbf{y}, \mathbf{y} \rangle F(d\mathbf{y}) < \infty$ and $\int_{R^2} |\langle \mathbf{c}, \mathbf{y} \rangle|^p F(d\mathbf{y}) < \infty$;
- (C2) $[\max(Z_n - n\varepsilon_0^{-1}, 0)]^p, n \geq 1$, are uniformly integrable;
- (C3) $\sum_{n=1}^\infty nP\{\xi_n < -(1 - \varepsilon_1)n\} < \infty$;
- (C4) $\lim_{\delta \searrow 0} \sup_{n \geq 1} P\{\max_{k \leq n\delta} |\xi_{n+k} - \xi_n| > \varepsilon\} = 0$ for all $0 < \varepsilon < \infty$;
- (C5) There exist events $A_n, n = 1, 2, \dots$, and $\beta \in [\frac{3}{2}, \infty)$ for which $\sum_{n=1}^\infty nP(\bigcup_{k=n}^\infty A'_k) < \infty, \max_{k \leq n} |\xi_{n+k} I(A_{n+k})|^\beta, n \geq 1$, are uniformly integrable;
- (C6) $(n^{-1/2}\Sigma_n, \xi_n) \xrightarrow{\mathcal{L}} (\mathbf{W}, \xi)$ as $n \rightarrow \infty$.

Since equations (12) and (17) of AW (1993) hold in our case with $q = 3$, their Proposition 4 leads to the verification of conditions (C4)–(C6) with $\beta = \frac{3}{2}$. Next, (C1) of AW (1993) holds since G has mean $\mathbf{0}$, F has finite 6th moment, so that $\int_{R^2} \|\mathbf{y}\|^2 dG < \infty$ and $\int_{R^2} |\langle \mathbf{c}, \mathbf{y} \rangle|^3 dG < \infty$. In order to verify (C2) of AW (1993), first notice that $Z_n = ng_n(\bar{\mathbf{Y}}_n)$ and thus, $Z_n \leq n^2\sigma^2$ for all but a few small values of n , and $Z_n \leq 2n$ on the set $\{S_n^2 \geq \frac{1}{2}\}$ w.p. 1, for all $n \geq 2$. So, by analogous arguments from Example 2 of AW (1993), we claim that (C2) holds in this situation. As far as the verification of (C3) goes, Proposition 5 of AW (1993) seems to be of no help. In the Appendix, we show directly that (C3) indeed holds here. One notes at this point that $\nu = E(\xi) = \frac{1}{2}E[\langle \mathbf{W}, D^2(g) |_{(0,0)}$

$\mathbf{W}] = \frac{1}{2}E[2W_1^2\sigma^{-2} + 2W_2^2\sigma^{-4}] = m_4\sigma^{-4}$, and by Proposition 3 of AW (1993), $(N^{-1/2}\Sigma_N, \xi_N, R_N) \xrightarrow{L} (\mathbf{W}, \xi, R)$ as $d \rightarrow 0$ with $R_N = Z_N - n_0$, the ‘‘overshoot’’ at the stopped stage and R given by (2.2). Also, R and (\mathbf{W}, ξ) are independent.

Part (i) now follows immediately from Theorem 1 of AW (1993). For part (ii), we use Corollary 1 of Theorem 2 of AW (1993) with $\mathbf{b} = (\sigma^{-1}, 0)$ and write

$$(2.3) \quad \sigma^4(\alpha^2 d^4)^{-1} E\{(\bar{X}_N - \theta)^2/\sigma^2\} + E(N) - 2\sigma^2(\alpha d^2)^{-1} \\ = 2[E(\xi W_1^2/\sigma^2) - \nu] + 2E\left[\left(\frac{X_1 - \theta}{\sigma}\right)^2 \left(1 - \left(\frac{X_1 - \theta}{\sigma}\right)^2\right)\right] \\ + m_4\sigma^{-4} - 1 + 2(m_3\sigma^{-3})^2 + o(1).$$

Now, use the expression of ξ in terms of W_1 and W_2 as well as the facts that $W_1/\sigma \sim N(0, 1)$ and the distribution of W_1W_2 is the same as that of the random variable $\frac{1}{2}[(m_3 + \sigma(m_4 - \sigma^4)^{1/2})U_1 + (m_3 - \sigma(m_4 - \sigma^4)^{1/2})U_2]$ where U_1, U_2 are i.i.d. χ_1^2 . In other words, $E(W_1^4/\sigma^4) = 3$ and $E[W_1^2W_2^2/\sigma^6] = 2(m_3/\sigma^3)^2 + (m_4/\sigma^4) - 1$. Hence, (2.3) leads to

$$(2.4) \quad E[(\bar{X}_N - \theta)^2/d^2] = \alpha + \alpha^2 d^2 \sigma^{-2} \{5 + 6(m_3/\sigma^3)^2 - \eta\} + o(d^2),$$

in view of part (i). Now, observe that

$$P\{\theta \in I_N\} \geq 1 - E\{(\bar{X}_N - \theta)^2/d^2\},$$

and hence (2.4) completes the proof.

Remark 2.1. If the distribution of $(X_1 - \theta)^2/\sigma^2$ is arithmetic, then we can claim that $E(N - n_0) = O(1)$ and $P\{\theta \in I_N\} \geq (1 - \alpha) + O(d^2)$ under the same conditions as in Theorem 2.1. The necessary modifications in the proof will be routine in view of AW (1993).

Remark 2.2. We have used a special case of Markov inequality to bound $P\{\theta \in I_N\}$. Instead, one may be tempted to use a sharper bound such as $P\{\theta \notin I_N\} \leq \inf_{p>0} E\{|\bar{X}_N - \theta|^p/d^p\}$. But, first of all this infimum may not be attained for any p and even if it is attained for some p , such ‘‘optimal’’ choice of p will then depend on the unknown distribution F . However, for F belonging to a certain specific class, say, \mathcal{F} , such ‘‘optimal’’ p may not depend on F , and this may lead to a sharper bound. But then one would require second-order expansions of $E\{|\bar{X}_N - \theta|^p/d^p\}$. This direction of potential improvement in the bound may perhaps materialize in the future in view of AW (1993). The present work provides the impetus and opens up possibilities for further investigations.

Remark 2.3. For fixed n , one may write down the Edgeworth expansion of $P\{\theta \in I_n\}$, and determine the ‘‘optimal’’ fixed-sample n via that route. But, once the stopping rule is proposed, then the problem is one of evaluating $P\{\theta \in I_n \mid N = n\}$ for all fixed n . At this time, it is quite unclear as to how the Edgeworth expansion of $P\{\theta \in I_n\}$ is supposed to help here. On top of this, the analysis

would also depend largely on the extent of control one would have on $P(N = n)$ for all n . At present, we do not have many clues.

Remark 2.4. In the case where F is given by a one-parameter exponential family, Woodroffe (1986) obtained very weak expansions of certain sequential confidence intervals. The impetus of his work in the general case of ours with unknown F remains a matter of guesswork at this time.

3. An accelerated sequential methodology

Instead of one by one sequential sampling as in (2.1), let us now pursue the idea of acceleration. In order to reduce the number of sampling operations, one starts out purely sequentially and goes part of the way, followed by augmentation via batch sampling of an appropriate size. The original development in the normal fixed-width confidence interval problem for the mean was due to Hall (1983). The associated general theory was put forth in Mukhopadhyay and Solanky (1991). Here, we develop the analog of the new and improved acceleration idea of Mukhopadhyay (1993) in the context of the present problem and study the associated rates.

Let us first choose and fix $\rho \in (0, 1)$ such that ρ^{-1} is an integer and define $m = m(d) = \max\{2, [(\rho^{-1}\alpha d^2)^{-1/2}]^* + 1\}$. One starts with X_1, \dots, X_m and then proceeds by taking one X at a time according to the stopping rule [with $S_n^2 = n^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$]

$$(3.1) \quad t = \inf\{n \geq m(d) : n\alpha d^2 \geq \rho S_n^2\}.$$

One has $P(t < \infty) = 1$ for all fixed $\alpha \in (0, 1)$, $d(> 0)$, and $\sigma \in (0, \infty)$. Note that t estimates ρn_0 , a fraction of n_0 . Let

$$(3.2) \quad Q = \rho^{-1}t,$$

and one then samples the difference of $(Q - t)$ observations, all in one single batch. Based on X_1, \dots, X_Q , we propose to estimate θ by means of the fixed-width confidence interval $I_Q = [\bar{X}_Q \pm d]$.

THEOREM 3.1. *Let $E[|X_1|^6] < \infty$. Then, under the accelerated sequential procedure (3.1)–(3.2), we have as $d \rightarrow 0$:*

- (i) $E(Q - n_0) = \rho^{-1}(\eta - m_4\sigma^{-4}) + o(1)$;
- (ii) $P\{\theta \in I_Q\} \geq (1 - \alpha) - \alpha^2\sigma^{-2}d^2\{5 + 6(m_3/\sigma^3)^2 - \eta + (\rho^{-1} - 1)(2m_4\sigma^{-4} - 1 - \eta)\} + o(d^2)$;

if $(X_1 - \theta)^2/\sigma^2$ has a nonarithmetic distribution and η, m_3, m_4 are the same as in Theorem 2.1.

Before we proceed to prove this result, let us first prove the following lemma.

LEMMA 3.1. *Suppose that $(X_1 - \theta)^2/\sigma^2$ has a nonarithmetic distribution, $E[|X_1|^6] < \infty$ and let $V = n_0^{*-1/2}(t - n_0^*)$ where $n_0^* = \rho n_0$. Then, we have as $d \rightarrow 0$:*

- (i) $V \xrightarrow{\mathcal{L}} N(0, m_4\sigma^{-4} - 1)$;
- (ii) V^2 is uniformly integrable;
- (iii) $E(t^{-1}) = n_0^{*-1} + (2m_4\sigma^{-4} - 1 - \eta)n_0^{*-2} + o(n_0^{*-2})$.

PROOF. In order to prove part (i), first note that $S_n^2 = (1 - n^{-1})S_n^{*2}$ where S_n^{*2} is a U -statistic and hence from Sproule (1969, 1974), one immediately concludes that $n_0^{*1/2}(S_t^2 - \sigma^2)$ and $n_0^{*1/2}(S_{t-1}^2 - \sigma^2)$ both converge in distribution to $N(0, m_4 - \sigma^4)$ as $d \rightarrow 0$. One may also refer to Lee (1990) in this context. Now, from (3.1) one writes

$$(3.3) \quad n_0^{*1/2}(S_t^2 - \sigma^2) \leq \sigma^2 V \leq n_0^{*1/2}(S_{t-1}^2 - \sigma^2) + \sigma^2 m n_0^{*-1/2} I(t = m),$$

where $I(\cdot)$ stands for the indicator function of (\cdot) . Now, part (i) follows if we can show that $I(t = m) \xrightarrow{P} 0$ as $d \rightarrow 0$. Observe that we can make $m\alpha d^2 \leq \frac{1}{2}\rho\sigma^2$ for sufficiently small d . Now, for arbitrary $\varepsilon (> 0)$, we write, for sufficiently small d ,

$$\begin{aligned} P\{I(t = m) > \varepsilon\} &\leq \varepsilon^{-1} P(t = m) = \varepsilon^{-1} P(m\alpha d^2 \geq \rho S_m^2) \\ &\leq \varepsilon^{-1} P\left\{|S_m^2 - \sigma^2| \geq \frac{1}{2}\sigma^2\right\} \rightarrow 0 \quad \text{as } d \rightarrow 0. \end{aligned}$$

Part (ii) follows from the Corollary to Proposition 8 of AW (1993). For part (iii), first note that

$$(3.4) \quad E(n_0^*/t) = n_0^{*-1}\{E[(t - n_0^*)^2/t] - E(t - n_0^*) + n_0^*\},$$

$$(3.5) \quad E(t - n_0^*) = \eta - m_4\sigma^{-4} + o(1),$$

in view of Theorem 2.1(i). Now, $(t - n_0^*)^2 t^{-1} I(t > \frac{1}{4}n_0^*) \leq 4V^2$, that is the l.h.s. is uniformly integrable. Also, the l.h.s. $\xrightarrow{\mathcal{L}} (m_4\sigma^{-4} - 1)\chi_1^2$, in other words,

$$(3.6) \quad E\left[(t - n_0^*)^2 t^{-1} I\left(t > \frac{1}{4}n_0^*\right)\right] = (m_4\sigma^{-4} - 1) + o(1).$$

Also, $E[(t - n_0^*)^2 t^{-1} I(t \leq \frac{1}{4}n_0^*)] \leq (n_0^* + \frac{n_0^{*2}}{m})P(t \leq \frac{1}{4}n_0^*) = o(1)$, by applying Lemma 1 of AW (1993) with $p = 3$. Combine this with (3.6) to write

$$E[(t - n_0^*)^2 t^{-1}] = (m_4\sigma^{-4} - 1) + o(1),$$

and the result follows from (3.4)–(3.5).

PROOF OF THEOREM 3.1. Part (i) is immediate from (3.5) since $Q = \rho^{-1}t$ and $n_0^* = \rho n_0$. In order to verify part (ii), first note that

$$(3.7) \quad E[(\bar{X}_Q - \theta)^2] = \rho^2 E[(\bar{X}_t - \theta)^2] + \rho(1 - \rho)\sigma^2 E(t^{-1}).$$

One may refer to Mukhopadhyay (1993) in this regard. Now, replacing N and d by t and $d\rho^{-1/2}$ respectively, we obtain from (2.4),

$$(3.8) \quad \rho E[(\bar{X}_t - \theta)^2/d^2] = \alpha + \alpha^2 d^2 \rho^{-1} \sigma^{-2} \{5 + 6(m_3/\sigma^3)^2 - \eta\} + o(d^2).$$

Combine (3.7)–(3.8) with Lemma 3.1 (iii) to write

$$\begin{aligned}
 (3.9) \quad E[(\bar{X}_Q - \theta)^2/d^2] &= \rho\alpha + \alpha^2 d^2 \sigma^{-2} [\{5 + 6(m_3/\sigma^3)^2 - \eta\} \\
 &\quad + (\rho^{-1} - 1)(2m_4\sigma^{-4} - 1 - \eta)] + (1 - \rho)\alpha + o(d^2) \\
 &= \alpha + \alpha^2 d^2 \sigma^{-2} [\{5 + 6(m_3/\sigma^3)^2 - \eta\} \\
 &\quad + (\rho^{-1} - 1)(2m_4\sigma^{-4} - 1 - \eta)] + o(d^2),
 \end{aligned}$$

parallel to what we had in (2.4). The proof is now complete since $P\{\theta \in I_Q\} \geq 1 - E\{(\bar{X}_Q - \theta)^2/d^2\}$.

Remark 3.1. Observe that the accelerated sequential methodology (3.1)–(3.2) requires approximately 100ρ% of the sampling operations associated with the full sequential procedure (2.1). In other words, the amount of operational savings can be truly substantial via (3.1)–(3.2), whereas the second-order expansions provided in Theorems 2.1 and 3.1 are quite similar indeed.

Remark 3.2. In the case where $(X_1 - \theta)^2/\sigma^2$ is arithmetic, we can only claim that $E(Q - n_0) = O(1)$ and $P\{\theta \in I_Q\} \geq (1 - \alpha) + O(d^2)$ under the same conditions as in Theorem 3.1.

4. Exponential mean: an example

Now, we consider a special d.f. given by $F(x) = \{1 - \exp(-x/\lambda)\} I(x > 0)$ with the mean of the distribution, $\lambda \in (0, \infty)$. Having recorded X_1, \dots, X_n , and given $d(> 0)$ as well as $0 < \alpha < 1$, we consider the bounded length confidence interval $I_n^* = [(\bar{X}_n - d) \vee 0, \bar{X}_n + d]$ for λ . However, note that

$$\begin{aligned}
 (4.1) \quad P\{\lambda \in I_n^*\} &= P\{\lambda \in I_n \cap \bar{X}_n \geq d\} + P\{\lambda \in [0, \bar{X}_n + d] \cap \bar{X}_n < d\} \\
 &= P\{\lambda \in I_n \cap \bar{X}_n \geq d\} + P\{\lambda - d \leq \bar{X}_n < d\} \\
 &= P\{\lambda \in I_n \cap \bar{X}_n \geq d\} + P\{\lambda \in I_n \cap \bar{X}_n < d\} \\
 &= P\{\lambda \in I_n\} \geq 1 - \alpha,
 \end{aligned}$$

if $E\{(\bar{X}_n - \lambda)^2/d^2\} \leq \alpha$, that is if $n \geq n_0 = \lambda^2(\alpha d^2)^{-1}$. In this case, of course, $\sigma^2 = \lambda^2$. Since n_0 is unknown, the purely sequential procedure (2.1) and its accelerated sequential version (3.1)–(3.2) can be put forth in a straightforward manner, and hence the Theorems 2.1 and 3.1 would certainly provide the associated second-order rates of convergences.

The full sequential procedure (2.1) is in order. The notations in the proof of Theorem 2.1 simplify to $a = n_0^* = \lambda^2(\alpha d^2)^{-1}$, $Z_n = n + \langle c, \Sigma_n \rangle + \xi_n$ where $c = (0, -\lambda^{-2})$, $\Sigma_n = \sum_{i=1}^n \mathbf{Y}_i$ with $\mathbf{Y}_i = (X_i - \lambda, (X_i - \lambda)^2 - \lambda^2)$. Here $\mathbf{W} = (W_1, W_2) \sim N_2\left(\mathbf{0}, \begin{pmatrix} \lambda^2 & 2\lambda^3 \\ 2\lambda^3 & 8\lambda^4 \end{pmatrix}\right)$ and $D^2g|_{(0,0)} = \begin{pmatrix} 2\lambda^{-2} & 0 \\ 0 & 2\lambda^{-4} \end{pmatrix}$, so that $\nu = 9$. That is, $P\{\lambda \in I_n^*\} \geq (1 - \alpha) - \alpha^2 d^2 \lambda^{-2}(29 - \eta) + o(d^2)$. Note, however, that in this methodology, one plugs in the sample variance S_n^2 in the stopping rules while replacing λ^2 by an estimator. But, in this specific instance we have a simple

parametric family, and it will perhaps make more sense to replace λ^2 by \bar{X}_n^2 in the stopping rules from sufficiency considerations.

Hence, the following purely sequential stopping rule is proposed. Let $m(\geq 1)$ be the starting sample size and define

$$(4.2) \quad N_0 = \inf\{n \geq m : n\alpha d^2 \geq \bar{X}_n^2\}.$$

Finally, we estimate λ by means of $I_{N_0}^*$. Then, if we choose to use a fixed m , from Theorem 2.1 of MD (1994), it follows that

$$(4.3) \quad P\{\lambda \in I_{N_0}^*\} \geq (1 - \alpha) - \alpha^2 d^2 \lambda^{-2} (16.5 + 2D) + o(d^2),$$

for $m \geq 7$, with $D = \sum_{n=1}^{\infty} \frac{1}{2} n^{-1} E\{(\chi_{2n}^2 - 3n) \vee 0\}$. Also, for $m \geq 3$, one obtains

$$(4.4) \quad E(N_0 - n_0) = -\frac{1}{2} - 2D + o(1).$$

One may independently derive (4.3)–(4.4) by using the results of AW (1993) once one allows m to grow as in our (2.1). It will suffice to note that $\mathbf{Y}_i = (X_i - \lambda, 2\lambda(X_i - \lambda))$, $g(y_1, y_2) = \lambda^2(\lambda^2 + y_1^2 + y_2)^{-1}$, $\mathbf{c} = Dg|_{(0,0)} = (0, -\lambda^{-2})$, $D^2g|_{(0,0)} = \begin{pmatrix} -2\lambda^{-2} & 0 \\ 0 & 2\lambda^{-4} \end{pmatrix}$, and $\xi = \lambda^{-4}W_2^2 - \lambda^{-2}W_1^2$ where $(W_1, W_2) \sim N_2\left(\mathbf{0}, \begin{pmatrix} \lambda^2 & 2\lambda^3 \\ 2\lambda^3 & 4\lambda^4 \end{pmatrix}\right)$. Other details are omitted.

Along the lines of MD (1994), we may “fine-tune” (4.2) so that in the second-order expansion of the confidence coefficient, one obtains $(1 - \alpha) + o(d^2)$ instead of (4.3). Define

$$(4.5) \quad N^* = \inf\{n \geq m(\geq 1) : (n + \varepsilon)\alpha d^2 \geq \bar{X}_n^2\}$$

with $\varepsilon = -(16.5 + 2D)$ where m is held fixed. Now, one proposes the interval $I_{N^*}^*$ for λ and from Theorem 2.2 of MD (1994), one claims that $P\{\lambda \in I_{N^*}^*\} \geq (1 - \alpha) + o(d^2)$ if $m \geq 7$.

4.1 The accelerated methodology

The accelerated estimation technique proposed in (3.1)–(3.2) does indeed work here without any difficulty. Theorem 3.1 then provides the second-order expansion of $E(Q - n_0)$ as well as the lower bound of $P\{\lambda \in I_Q^*\}$ which is $(1 - \alpha) - \alpha^2 d^2 \lambda^{-2} \{(29 - \eta) + (\rho^{-1} - 1)(17 - \eta)\} + o(d^2)$. But, then one would be interested to accelerate the purely sequential procedure (4.2) in order to curtail sampling operations.

Let $m(\geq 1)$ be the fixed starting sample size, and choose and fix $0 < \rho < 1$ where ρ^{-1} is an integer. Define

$$(4.6) \quad t_0 = \inf\{n \geq m : n\alpha d^2 > \rho \bar{X}_n^2\},$$

$$(4.7) \quad M = \rho^{-1} t_0.$$

One samples the difference $(M - t_0)$ in one single batch. Finally, we propose to estimate λ by means of I_M^* . Then, from Theorem 3.1 of MD (1994), we obtain

$$(4.8) \quad P\{\lambda \in I_M^*\} \geq (1 - \alpha) - \alpha^2 d^2 \lambda^{-2} \{12 + \rho^{-1}(4.5 + 2D)\} + o(d^2)$$

for $m \geq 7$. It would be desirable to obtain the fine-tuned version of (4.6)–(4.7). However, in view of Section 3.2 of MD (1994), such a development would be fairly straightforward. Hence, this is omitted for brevity.

Acknowledgements

We thank two referees for their helpful remarks.

Appendix

In the setup of Theorem 2.1, we need to verify that the condition (C3) of AW (1993) holds. Note that

$$(A.1) \quad n^{-1}\xi_n = \frac{\sigma^2}{S_n^2} + n^{-1} \sum_{i=1}^n \frac{(X_i - \theta)^2}{\sigma^2} - 2 = T_n, \quad \text{say,}$$

and we have to show that

$$(A.2) \quad \sum_{n=1}^{\infty} nP\{\xi_n < -\varepsilon_1 n\} < \infty$$

with $0 < \varepsilon_1 < 1$. Now, for $0 < \varepsilon < \varepsilon_1 < 1$, we write

$$(A.3) \quad P\{\xi_n \leq -\varepsilon_1 n\} \leq P\left\{T_n \leq -\varepsilon_1 \cap \frac{\sigma^2}{S_n^2} \geq 1 - \varepsilon\right\} + P\left\{\frac{\sigma^2}{S_n^2} < 1 - \varepsilon\right\} \\ = I_n + II_n, \quad \text{say.}$$

Then, one has (with $\varepsilon' = \varepsilon_1 - \varepsilon > 0$)

$$I_n \leq P\left\{(1 - \varepsilon) + \frac{1}{n} \sum_{i=1}^n \frac{(X_i - \theta)^2}{\sigma^2} - 2 \leq -\varepsilon_1\right\} \\ \leq P\left\{\left|\frac{1}{n} \sum_{i=1}^n \frac{(X_i - \theta)^2}{\sigma^2} - 1\right| \geq \varepsilon'\right\},$$

and hence using Theorem 1 of Katz (1963) with $t = 3$, we have $\sum_{n=1}^{\infty} nI_n < \infty$ if $E[|X_1|^6] < \infty$. Also, with some $\varepsilon'' > 0$, we can write for large enough n ,

$$II_n \leq P\left\{\frac{S_n^2}{\sigma^2} \geq 1 + \varepsilon''\right\} \\ \leq P\left\{(n-1)^{-1} \sum_{i=1}^n \frac{(X_i - \theta)^2}{\sigma^2} \geq 1 + \varepsilon''\right\} \\ \leq P\left\{\frac{1}{n} \sum_{i=1}^n \frac{(X_i - \theta)^2}{\sigma^2} \geq 1 + \varepsilon''\right\},$$

and by a similar reasoning as before, we claim that $\sum_{n=1}^{\infty} nII_n < \infty$ if $E[|X_1|^6] < \infty$. Hence, (A.3) leads to (A.2).

REFERENCES

- Aras, G. and Woodroffe, M. (1993). Asymptotic expansions for the moments of randomly stopped average, *Ann. Statist.*, **21**, 503–519.
- Chow, Y. S. and Robbins, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean, *Ann. Math. Statist.*, **36**, 457–462.
- Csenki, A. (1980). On the convergence rate of fixed-width sequential confidence intervals, *Scand. Actuar. J.*, 107–111.
- Hall, P. (1983). Sequential estimation saving sampling operations, *J. Roy. Statist. Soc. Ser. B*, **45**, 219–223.
- Katz, M. L. (1963). The probability in the tail of distribution, *Ann. Math. Statist.*, **34**, 312–318.
- Landers, D. and Rogge, L. (1976). The exact approximation order in the central-limit-theorem for random summation, *Zeitschrift für Wahrscheinlichkeit und Verwandte Gebiete*, **36**, 269–283.
- Lee, A. J. (1990). *U-Statistics: Theory and Practice*, Marcel Dekker, New York.
- Mukhopadhyay, N. (1981). Convergence rates of sequential confidence intervals and tests for the mean of a U -statistic, *Comm. Statist. Theory Methods*, **10**, 2231–2244.
- Mukhopadhyay, N. (1988). Sequential estimation problems for negative exponential populations, *Comm. Statist. Theory Methods*, (Reviews Section), **17**, 2471–2506.
- Mukhopadhyay, N. (1993). An alternative formulation of accelerated sequential procedures with applications, Tech. Report, No. 93-27, Department of Statistics, University of Connecticut, Storrs (*Sequential Anal.*, (1996), **15**, in press).
- Mukhopadhyay, N. and Datta, S. (1994). On fine-tuned bounded risk sequential point estimation of the mean of an exponential distribution, Tech. Report, No. 94-06, Department of Statistics, University of Connecticut, Storrs (*South African Statist. J.*, (1995), **29**, 9–27).
- Mukhopadhyay, N. and Solanky, T. K. S. (1991). Second order properties of accelerated stopping times with applications in sequential estimation, *Sequential Anal.*, **10**, 99–123.
- Mukhopadhyay, N. and Vik, G. (1985). Asymptotic results for stopping times based on U -statistics, *Sequential Anal.*, **4**, 83–109.
- Sproule, R. N. (1969). A sequential fixed-width confidence interval for the mean of a U -statistic, Ph.D. Dissertation, Department of Statistics, University of North Carolina, Chapel Hill.
- Sproule, R. N. (1974). Asymptotic properties of U -statistics, *Trans. Amer. Math. Soc.*, **199**, 55–64.
- Woodroffe, M. (1977). Second order approximations for sequential point and interval estimation, *Ann. Statist.*, **5**, 985–995.
- Woodroffe, M. (1986). Very weak expansions for sequential confidence levels, *Ann. Statist.*, **14**, 1049–1067.