

A class of accelerated sequential procedures with applications to estimation problems for some distributions useful in reliability theory

Neeraj Joshi^{1,a}, Sudeep R. Bapat^b, Ashish Kumar Shukla^c

^aDepartment of Statistics, University of Delhi, India;

^bDepartment of Operations Management and Quantitative Techniques, IIM, India;

^cDepartment of Statistics, Ramanujan College (University of Delhi), India

Abstract

This paper deals with developing a general class of accelerated sequential procedures and obtaining the associated second-order approximations for the expected sample size and ‘regret’ (difference between the risks of the proposed accelerated sequential procedure and the optimum fixed sample size procedure) function. We establish that the estimation problems based on various lifetime distributions can be tackled with the help of the proposed class of accelerated sequential procedures. Extensive simulation analysis is presented in support of the accuracy of our proposed methodology using the Pareto distribution and a real data set on carbon fibers is also analyzed to demonstrate the practical utility. We also provide the brief details of some other inferential problems which can be seen as the applications of the proposed class of accelerated sequential procedures.

Keywords: accelerated sequential, carbon fibers, inverse Gaussian, minimum risk, multivariate normal, negative exponential, normal, Pareto, regret, second-order approximations

1. Introduction

Dantzig (1940) proved the non-existence of fixed sample size procedures to construct the confidence interval of preassigned width and coverage probability for a normal mean in the ignorance of any knowledge about the variance. In order to deal with this estimation problem, Stein (1945) proposed a two-stage procedure. Starr (1966) and Woodroffe (1977) adopted purely sequential procedures for the same estimation problem. Two-stage and purely sequential procedures for estimating the parameters involved with different probabilistic models have been developed and studied by various authors. For a brief review, one may refer to the monographs of Ghosh *et al.* (1997) and Mukhopadhyay and de Silva (2009).

Hall (1983) deduced that both the two-stage and purely sequential estimation procedures have some drawbacks. On one side, two-stage procedure is easy to operate as it requires only two stages and achieves the exact coverage probability but the difference between the average sample number and the ‘optimal’ fixed-sample size does not remain asymptotically bounded. As a result, two-stage procedure leads us to considerable oversampling. On the other side, purely sequential procedure is complicated in nature to apply and achieves the target value of coverage probability only asymptotically. Hall (1983) developed a sampling scheme in order to construct fixed-width confidence interval for a normal

¹ Corresponding author: Department of Statistics, University of Delhi, Delhi-110007, India.
E-mail: stats.joshi@gmail.com

mean, which could combine the advantages of two-stage and sequential procedures. He named it as an “accelerated” sequential procedure. In his procedure, sampling stages can be reduced by a predetermined factor at the cost of finite number of observations with very nearly the desired coverage and difference between the average sample number and ‘optimal’ fixed sample size remains bounded.

Mukhopadhyay and Solanky (1991) developed the second-order asymptotic theory for accelerated sequential stopping rules and provided several interesting applications of their general setup. They have also provided solutions to various interesting ranking and selection problems by using the accelerated sequential sampling design (Mukhopadhyay and Solanky, 1992a,b, 1993). Mukhopadhyay (1996) provided an alternative formulation of accelerated sequential procedures having applications to various parametric and non parametric estimation problems. Datta and Mukhopadhyay (1998) proposed the accelerated sequential procedures in partitioning a set of normal populations. Chattopadhyay (2000) and Chattopadhyay and Sengupta (2006) have developed accelerated sequential procedures for the parameters of exponential and normal populations under asymmetric loss structures. Recently, Hu (2021) provided an improved accelerated sequential procedure to tackle the problem of fixed-width confidence interval estimation of an unknown normal mean in the absence of any knowledge about the variance. Similar procedures to deal with some other estimation problems have been developed and studied by numerous authors. To cite a few, one may refer to Son and Hamdy (1990), Hamdy and Son (1991), Chaturvedi and Tomer (2003), Chaturvedi *et al.* (2019a,b), Joshi and Bapat (2020a), Mukhopadhyay and Khariton (2020), Zhuang and Bhattacharjee (2021), Joshi *et al.* (2021) and the references therein.

Taking into consideration the relationship among the ‘optimal’ fixed sample size solutions of estimation and ranking and selection problems related to normal populations, Chaturvedi and Gupta (1994) developed a class of sequential procedures. The confidence interval and ranking and selection problems were linked with the bounded risk point estimation problems under zero-one loss function exploiting the common functional form of the risks associated with these problems and second-order approximations were obtained for the ‘regret’ (difference between the risks of optimal and fixed sample size procedures). It was shown that the results obtained under this general set-up provided second-order approximations to estimation and ranking and selection problems and, as such, no separate dealing was required.

In the present paper, motivated by the work of Chaturvedi and Gupta (1994), a general class of accelerated sequential procedures is developed. Second-order approximations are obtained for the expected sample size and ‘regret’ respectively. By means of examples, it is established that the estimation problems based on various probability distributions can be handled with the help of the proposed class of accelerated sequential procedures. Note that the work of Chaturvedi and Gupta (1994) covers various inferential problems only for normal distribution whereas our present paper provides a more generalized approach covering the wide range of problems (the problems based on normal distribution are one among them). Further, in the present work, we propose a general class of accelerated sequential procedures which reduces the number of sampling operations substantially, whereas Chaturvedi and Gupta (1994) dealt with a class of sequential procedures under the normal distribution case.

The rest of the paper is organized as follows. In Section 2, we give the set-up of the estimation problem and establish the failure of fixed sample size procedure to deal with ‘it’. In Section 3, we propose the accelerated sequential procedure to handle this estimation problem. Section 4 contains some important results associated with the proposed accelerated sequential procedure. In Section 5, we derive the associated second-order approximations. Section 6 deals with applications of the proposed methodology to tackle the inferential problems related to various distributions useful in

reliability theory. Section 7 presents extensive simulation and real data analyses, to complement our proposed procedures. Finally, in Section 8, we provide brief set of conclusions and future scope of the proposed work.

2. The estimation problem

We know that the ‘optimal’ fixed sample size required by a ‘given precision problem’ often takes the form,

$$n^* = a\theta^b, \tag{2.1}$$

where $\theta (> 0)$ is an unknown scalar parameter, a and b are known positive numbers and ‘ a ’ may be allowed to approach infinity. Given a random sample $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ of size $n (\geq t + 1)$ from a t -variate absolutely continuous population, let $\hat{\theta}_n = \hat{\theta}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$ be the estimator of θ such that, for all $n \geq s + 1$,

$$q(n - s) \frac{\hat{\theta}_n}{\theta} = \sum_{j=1}^{n-s} Z_j^{(q)}, \tag{2.2}$$

where q and s are positive integers with $Z_j^{(q)}$ following χ^2 -distribution with q degrees of freedom. We now provide a fundamental problem as an example, to gauge as to how the ‘optimal’ sample size takes the form equation (2.1) and how the assumption equation (2.2) is useful in obtaining the desired asymptotic results.

Suppose we are interested in estimating the mean $\mu (-\infty < \mu < \infty)$ of a normal population $N(\mu, \sigma^2)$, when the variance $\sigma^2 (0 < \sigma < \infty)$ remains unknown. Given two preassigned numbers $d (> 0)$ and $0 < \alpha < 1$, we wish to construct a confidence interval $J = [\mu - d, \mu + d]$ of length $2d$ for μ such that, $P_{\mu, \sigma}(\mu \in J) > 1 - \alpha$ for all μ, σ . Since there does not exist any fixed-sample size methodology to solve this problem, we proceed as follows. Starting with a sequence of independent observations X_1, X_2, \dots, X_n from $N(\mu, \sigma^2)$, denote $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$ and $S_n^2 = (n - 1)^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ as the sample mean and sample variance, respectively. We propose the interval $J_n = [\bar{X}_n - d, \bar{X}_n + d]$ to estimate μ which satisfies $P_{\mu, \sigma}(\mu \in J_n) > 1 - \alpha$. The confidence coefficient can be obtained as follows

$$P_{\mu, \sigma}(\mu \in J_n) = P\left\{|\bar{X}_n - \mu| \leq d\right\} = 2\Phi\left(\frac{\sqrt{nd}}{\sigma}\right) - 1,$$

where $\Phi(\cdot)$ is the distribution function of a standard normal variate. Since, the associated confidence coefficient must be $1 - \alpha$, we can write

$$2\Phi\left(\frac{\sqrt{nd}}{\sigma}\right) - 1 \geq 1 - \alpha = 2\Phi(z) - 1,$$

where z is the upper 50% point of a standard normal distribution. Now, the optimal fixed sample size n^* can be obtained as follows

$$n \geq z^2 \frac{\sigma^2}{d^2} = n^*$$

It is obvious that the above optimal fixed sample size takes the form equation (2.1) with $a = z^2/d^2$, $b = 1$ and $\theta = \sigma^2$. Moreover, we have

$$\frac{(n - 1)S_n^2}{\sigma^2} = \sum_{i=1}^n Z_j$$

where Z_j follows χ^2 -distribution with 1 degree of freedom. We can see that this result takes the form in equation (2.2) with $q = 1$, $s = 1$, $\hat{\theta}_n = S_n^2$ and $\theta = \sigma^2$.

In a similar manner, one can observe the assumptions equations (2.1)–(2.2) for several other estimation problems. Obviously, the values of $a, b, q, s, \hat{\theta}_n$ and θ would be different for each estimation problem. At this point, we leave out many such examples for brevity alone. However, we will discuss some specific problems in Section 6.

Coming back to equation (2.1), since θ is unknown, no fixed sample size procedure provides solution to the problem and hence we adopt the following accelerated sequential procedure to deal with it.

3. Accelerated sequential procedure

Taking $m \geq \max(s + 1, t + 1)$ be the initial sample size, where, as in Hall (1983), m is chosen to satisfy $m = o(a^{-1})$ as $a \rightarrow \infty$ and $\lim_{a \rightarrow \infty} \sup(m/n^*) < 1$. Let $\eta \in (0, 1)$ be specified. We start sampling sequentially with the stopping time L defined by,

$$L = \inf \{n \geq m; n \geq \eta a \hat{\theta}_n^b\}. \quad (3.1)$$

Based on these L observations, we compute $\hat{\theta}_L$. Then, denoting by $\lfloor u \rfloor$ – the largest integer smaller than u , we jump ahead and take $N - L$ observations, where,

$$N = \max \{L, \lfloor a \hat{\theta}_L^b \rfloor + 1\}. \quad (3.2)$$

The stopping rule N estimates n^* from equation (2.1). Based on these N observations, we propose $\hat{\theta}_N$ as the estimator of θ . It is worth mentioning here that, we are using a purely sequential strategy (which takes observations one by one until the process terminates) only to determine a preliminary sample and then jumping ahead to take a batch of remaining observations in one step. So, somehow we are accelerating our sequential procedure and that's why this process is named as an ‘‘accelerated’’ sequential procedure. Suppose that the risk associated with the accelerated sequential procedure equations (3.1)–(3.2) is of the form,

$$R_N(a) = E \left[g \left(\frac{N}{n^*} \right) \right], \quad (3.3)$$

where $g(x)$ is a continuous, twice differentiable function and for the random variable ‘ U ’ defined by $|U - 1| \leq |(N/n^*) - 1|$,

$$g''(U) \leq k_1 a^{-u} [k_2 U^v + k_3 U^{-w}], \quad (3.4)$$

where $u \in [0, \infty)$, $v \in (-\infty, \infty)$, $w \in [0, \infty)$ and $|k_i| < \infty, i = 1, 2, 3$.

Obviously, the ‘optimal’ fixed sample size risk can be obtained by putting $N = n^*$ in equation (3.3) and is given by

$$R_{n^*}(a) = g(1), \quad (3.5)$$

and the ‘regret’ associated with the accelerated sequential procedure equations (3.1)–(3.2) is defined by

$$R_g(a) = R_N(a) - R_{n^*}(a). \quad (3.6)$$

4. Some important results

In this section, we provide some important lemmas associated with the accelerated sequential procedure equations (3.1)–(3.2). We will use these lemmas to obtain the second-order approximations in Section 5.

Lemma 1. *For the class of accelerated sequential procedures equations (3.1)–(3.2), we have,*

$$P(N < \infty) = 1, \quad (4.1)$$

$$\lim_{a \rightarrow \infty} N = \infty, \quad (4.2)$$

$$\lim_{a \rightarrow \infty} \left(\frac{N}{n^*} \right) = 1. \quad (4.3)$$

Proof: To prove the result equation (4.1), first let us write $P(N = \infty) = \lim_{n \rightarrow \infty} P(N > n)$. Note that the event $\{N > n\}$ means, we did not stop with n observations. Thus, at the stage when we have $N(\geq m)$ observations, we must have all the events $k < \eta a \hat{\theta}_k$ satisfied, where, $k = m, \dots, n$. Thus, we can write,

$$\begin{aligned} P(N > n) &= P\{k < \eta a \hat{\theta}_k, k = m, \dots, n\} \\ &\leq P\{n < \eta a \hat{\theta}_n^b\} \\ &= P\left\{\hat{\theta}_n - \theta > \left(\frac{n}{\eta a}\right)^{\frac{1}{b}} - \theta\right\}. \end{aligned}$$

We may consider ‘ n ’ sufficiently large, i.e., when $(n/\eta a)^{1/b} > \theta$, using Markov’s inequality, we obtain,

$$\begin{aligned} P(N > n) &\leq P\left\{|\hat{\theta}_n - \theta| > \left(\frac{n}{\eta a}\right)^{\frac{1}{b}} - \theta\right\} \\ &\leq E\{|\hat{\theta}_n - \theta|^2\} \left\{\left(\frac{n}{\eta a}\right)^{\frac{1}{b}} - \theta\right\}^{-2} \\ &= \text{Var}(\hat{\theta}_n) \left\{\left(\frac{n}{\eta a}\right)^{\frac{1}{b}} - \theta\right\}^{-2}. \end{aligned} \quad (4.4)$$

Since from equation (2.2), we have,

$$\begin{aligned} \hat{\theta}_n &\sim \frac{\theta}{q(n-s)} \chi_{q(n-s)}^2, \\ \Rightarrow \text{Var}(\hat{\theta}_n) &\equiv \frac{2\theta^2}{q(n-s)}, \end{aligned}$$

we can rewrite equation (4.4) as,

$$P(N > n) = \frac{2\theta^2}{q(n-s)} \left\{\left(\frac{n}{\eta a}\right)^{\frac{1}{b}} - \theta\right\}^{-2}.$$

Thus, we have,

$$\lim_{n \rightarrow \infty} P(N > n) \leq \lim_{n \rightarrow \infty} \left[\frac{2\theta^2}{q(n-s)} \left\{ \left(\frac{n}{\eta a} \right)^{\frac{1}{b}} - \theta \right\}^{-2} \right] = 0,$$

and equation (4.1) follows.

Equation (4.2) is an immediate consequence of the definition of N given in equation (3.2). Further, from equation (3.1), we observe the inequality

$$\left(\frac{\hat{\theta}_L}{\theta} \right)^b \leq \frac{N}{n^*} \leq \left(\frac{\hat{\theta}_L}{\theta} \right)^b + \frac{LI(N=L) + 1}{n^*},$$

which can further be written as

$$\left(\frac{\hat{\theta}_L}{\theta} \right)^b \leq \frac{N}{n^*} \leq \left(\frac{\hat{\theta}_L}{\theta} \right)^b + O(a^{-1}), \quad (4.5)$$

where $I(E)$ denotes the indicator function of an event E . Moreover, from equation (2.2) and Kolmogorov's strong law of large numbers (Bhat, 1981), $\hat{\theta}_L \xrightarrow{a.s.} \theta$ as $L \rightarrow \infty$. Equation (4.3) now follows from equation (4.5) on taking the limit as $a \rightarrow \infty$. \square

Lemma 2. For $0 < \epsilon < 1$, as $a \rightarrow \infty$, we have,

$$P(N \leq \epsilon n^*) = O\left(a^{-\frac{q(m-s)}{2b}}\right). \quad (4.6)$$

Proof: We have,

$$\begin{aligned} P(N \leq \epsilon n^*) &\leq P(L \leq \epsilon n^*) \\ &= P(L = m) + P(m+1 \leq L \leq \epsilon n^*). \end{aligned} \quad (4.7)$$

Firstly, we focus on $P(L = m)$. We have,

$$\begin{aligned} P(L = m) &= P\left[m \geq \eta a \hat{\theta}_m^b\right] \\ &= P\left[V_m \leq q(m-s) \left(\frac{m}{\eta n^*}\right)^{\frac{1}{b}}\right] \\ &= \left[2^{\frac{q(m-s)}{2}} \Gamma\left(\frac{q(m-s)}{2}\right)\right]^{-1} \int_0^{q(m-s)\left(\frac{m}{\eta n^*}\right)^{\frac{1}{b}}} e^{-\frac{y}{2}} y^{\left(\frac{q(m-s)}{2}\right)-1} dy \\ &\leq \left[2^{\frac{q(m-s)}{2}} \Gamma\left(\frac{q(m-s)}{2}\right)\right]^{-1} \int_0^{q(m-s)\left(\frac{m}{\eta n^*}\right)^{\frac{1}{b}}} y^{\left(\frac{q(m-s)}{2}\right)-1} dy \\ &= \left[2^{\frac{q(m-s)}{2}} \Gamma\left(\frac{q(m-s)}{2}\right)\right]^{-1} \left[q(m-s) \left(\frac{m}{\eta n^*}\right)^{\frac{1}{b}}\right]^{\frac{q(m-s)}{2}} \\ &= O\left(a^{-\frac{q(m-s)}{2b}}\right). \end{aligned} \quad (4.8)$$

We focus on $P(m + 1 \leq L \leq \epsilon n^*)$. Without loss of generality, let us suppose that ϵn^* is an integer. Using the Bernstein-Chernoff inequality, we can write,

$$\begin{aligned}
 P(m + 1 \leq L \leq \epsilon n^*) &\leq \sum_{n=m+1}^{\epsilon n^*} P \left[V_n \leq q(n-s) \left(\frac{n}{\eta n^*} \right)^{\frac{1}{b}} \right] \\
 &\leq \sum_{n=m+1}^{\epsilon n^*} \inf_{h>0} P \left[-hV_n \geq -hq(n-s) \left(\frac{n}{\eta n^*} \right)^{\frac{1}{b}} \right] \\
 &\leq \sum_{n=m+1}^{\epsilon n^*} \inf_{h>0} \left[(1+2h)^{-\frac{q(n-s)}{2}} \exp \left\{ hq(n-s) \left(\frac{n}{\eta n^*} \right)^{\frac{1}{b}} \right\} \right] \\
 &\leq \sum_{n=m+1}^{\epsilon n^*} (1+2h_0)^{-\frac{q(n-s)}{2}} \exp \left\{ h_0 q(n-s) \left(\frac{n}{\eta n^*} \right)^{\frac{1}{b}} \right\} \tag{4.9}
 \end{aligned}$$

with

$$h_0 = \frac{1}{2} \left\{ \left(\frac{\eta n^*}{n} \right)^{\frac{1}{b}} - 1 \right\}. \tag{4.10}$$

This is positive obviously when $n < \epsilon n^*$. Combining equations (4.9)–(4.10), we obtain,

$$\begin{aligned}
 P(m + 1 \leq L \leq \epsilon n^*) &\leq \sum_{n=m+1}^{\epsilon n^*} \left(\frac{\eta n^*}{n} \right)^{-\frac{q(n-s)}{2}b} \exp \left\{ -\frac{q(n-s)}{2} \left(\frac{n}{\eta n^*} \right)^{\frac{1}{b}} + \frac{q(n-s)}{2} \right\} \\
 &= \sum_{n=m+1}^{\epsilon n^*} \left[\left(\frac{n}{\eta n^*} \right)^{\frac{1}{b}} \exp \left\{ 1 - \left(\frac{n}{\eta n^*} \right)^{\frac{1}{b}} \right\} \right]^{\frac{q(n-s)}{2}} \tag{4.11}
 \end{aligned}$$

We observe that for all $n \leq \epsilon n^*$, we have $(n/\eta n^*)^{1/b} < \epsilon^{1/b} = k^* (< 1)$, say. But $xe^{1-x} \uparrow x$, for $0 < x < 1$, which means $\max_{0 < a < 1} xe^{1-x}$ must be ae^{1-a} . Now we can rewrite equation (4.11) as,

$$P(m + 1 \leq L \leq \epsilon n^*) \leq \sum_{n=m+1}^{\epsilon n^*} k^{* \frac{q(n-m)}{2}} \left(\frac{n}{\eta n^*} \right)^{\frac{q(m-s)}{2}b} \left[\exp \left\{ 1 - \left(\frac{n}{\eta n^*} \right)^{\frac{1}{b}} \right\} \right]^{\frac{q(m-s)}{2}}. \tag{4.12}$$

Since $e^{1-x} \leq e$ for all real $x > 0$, equation (4.12) leads to,

$$\begin{aligned}
 P(m + 1 \leq L \leq \epsilon n^*) &\leq e^{\frac{q(m-s)}{2}} \sum_{n=m+1}^{\epsilon n^*} k^{* \frac{q(m-s)}{2}} \left(\frac{n}{\eta n^*} \right)^{\frac{q(m-s)}{2}b} \\
 &= n^{*-\frac{q(m-s)}{2}b} e^{\frac{q(m-s)}{2}} \sum_{n=m+1}^{\epsilon n^*} b_n. \tag{4.13}
 \end{aligned}$$

The infinite sum $\sum_{n=m+1}^{\epsilon n^*} b_n$ in equation (4.13) is convergent by the ‘‘root’’ test once we note that $b_n^{1/n} \rightarrow k, 0 < k < 1$, as $n \rightarrow \infty$. Thus, we have,

$$P(m + 1 \leq L \leq \epsilon n^*) = O \left(a^{-\frac{q(m-s)}{2}b} \right). \tag{4.14}$$

The lemma now follows on making substitutions from equation (4.8) and equation (4.14) in equation (4.7). \square

Lemma 3. $g''(U)$ is uniformly integrable for all

$$m > \max\{s + 2q^{-1}b(v^* - u), s + 2q^{-1}b(w - u)\},$$

where, for $v < 0$, $v = -v^*$, ($v^* > 0$).

Proof: On the event ' $N > \epsilon n^*$ ', we have $|U - 1| \leq |(N/n^*) - 1| \leq 1 - \epsilon$, i.e., $U \geq 2 - \epsilon$ and $U^{-1} \leq \eta^{-1}$. Hence, from equation (3.4)

$$E[g''(U)I(N > \epsilon n^*)] \leq Ka^{-u}P(N > \epsilon n^*),$$

which on using equation (4.1) gives that, as $a \rightarrow \infty$

$$E[g''(U)I(N > \epsilon n^*)] = o(1). \quad (4.15)$$

Moreover, on the event ' $N \leq \epsilon n^*$ ', we have, $|U - 1| \leq 1 - m/n^*$. Thus, $U < 2$ and $U^{-1} \leq n^*/m$. From equation (3.4)

$$E[g''(U)I(N \leq \epsilon n^*)] \leq k_1 a^{-u} [2^v k_2 + k_3 \left(\frac{n^*}{m}\right)^w] P(N \leq \epsilon n^*), \quad (4.16)$$

which on applying Lemma 2 gives that

$$\begin{aligned} E[g''(U)I(N \leq \epsilon n^*)] &= O\left(a^{-\frac{u-q(m-s)}{2}b}\right) + O\left(a^{\frac{(w-u)-q(m-s)}{2}b}\right), \\ &= o(1), \text{ for all } m > s + 2bq^{-1}(w - u). \end{aligned} \quad (4.17)$$

Further, when $v < 0$, say $v = -v^*$, $v^* > 0$, we have from equation (3.4)

$$\begin{aligned} E[g''(U)I(N \leq \epsilon n^*)] &\leq k_1 a^{-u} \left[k_2 \left(\frac{n^*}{m}\right)^{v^*} + k_3 \left(\frac{n^*}{m}\right)^w \right] P(N \leq \epsilon n^*), \\ &= O\left(a^{v^* \frac{-u-q(m-s)}{2}b}\right) + O\left(a^{(w-u) \frac{-q(m-s)}{2}b}\right), \\ &= o(1), \end{aligned} \quad (4.18)$$

for all $m > \max\{s + 2bq^{-1}(v^* - u), s + 2bq^{-1}(w - u)\}$. The lemma now follows on combining equation (4.15), equation (4.16) and equation (4.18). \square

5. Second-order approximations

In this section, we provide second-order approximations for the expected sample size (theorem 1) and 'regret' (theorem 2) associated with the proposed accelerated sequential procedure equation (3.1)–(3.2).

Theorem 1. For all $m > s + 2q^{-1}b$, as $a \rightarrow \infty$, we have,

$$E(N) = n^* - (\eta q)^{-1}b(b + 1) + \frac{1}{2} + o(1), \quad (5.1)$$

$$\text{Var}(N) = 2(\eta q)^{-1}n^*b^2 + o(a). \quad (5.2)$$

Proof: Denoting by, $\psi_L = 1 - \{a\hat{\theta}_L^b - \lfloor a\hat{\theta}_L^b \rfloor\}$, we can write,

$$E(N) = I + II, \tag{5.3}$$

where,

$$\begin{aligned} I &= E \left\{ L \cdot I \left(L \geq \lfloor a\hat{\theta}_L^b \rfloor + 1 \right) \right\}, \\ II &= aE \left(\hat{\theta}_L^b \right) + E(\psi_L). \end{aligned} \tag{5.4}$$

It follows from Hall (1983) that, as $a \rightarrow \infty, I = o(1)$. Further, from equation (2.2) one can see that the distribution of $\hat{\theta}_n$ is equivalent to $\theta/q(n-s)\chi_{q(n-s)}^2$, whose Fourier transform becomes,

$$\frac{\theta}{q(n-s)} \left(\frac{1}{1-2iu} \right)^{q(n-s)},$$

which clearly tends to 0 as $u \rightarrow \infty$. Hence, from Tukey (1938) one can get that the fractional part of $a\hat{\theta}_L^b$ tends to a Uniform(0, 1) distribution and thus, ψ_L is also uniform over (0, 1). So, we obtain from equation (5.3) and equation (5.4) that

$$E(N) = aE(\hat{\theta}_L^b) + \frac{1}{2} + o(1). \tag{5.5}$$

We evaluate $E(\hat{\theta}_L^b)$. To this end, we write equation (3.1) as,

$$\begin{aligned} L &= \inf \left\{ n \geq m; q(n-s) \frac{\hat{\theta}_n}{\theta} \leq q(n-s) \left(\frac{n}{\eta n^*} \right)^{\frac{1}{b}} \right\}, \\ &= \inf \left\{ n \geq m; \sum_{j=1}^{n-s} Z_j^{(q)} \leq q(n-s) \left(\frac{n}{\eta n^*} \right)^{\frac{1}{b}} \right\}. \end{aligned}$$

Let us define another stopping rule $L^* = L - s$ by,

$$L^* = \inf \left\{ n \geq m - s; \sum_{j=1}^n Z_j^{(q)} \leq \frac{qn^{1+b^{-1}}(1+sn^{-1})^{\frac{1}{b}}}{(\eta n^*)^{\frac{1}{b}}} \right\}. \tag{5.6}$$

Comparing equation (5.6) with equation (1.1) of Woodroffe (1977), we obtain in his notations,

$$\begin{aligned} S_n &= \sum_{j=1}^n Z_j^{(q)}, \quad \alpha = 1 + b^{-1}, \quad \beta = b, \quad \lambda = \eta n^*, \quad a = \frac{q}{2}, \\ L_n &= \left(1 + sn^{-1} \right)^{\frac{1}{b}}, \quad L_0 = \frac{s}{b}, \quad \mu = q, \quad \tau^2 = 2q. \end{aligned}$$

It now follows from this theorem 2.4 that, for all $m > s + 2q^{-1}b$, as $a \rightarrow \infty$,

$$E(L^*) = \eta n^* + bv - s - bq^{-1}(b+1) + o(1),$$

where the value of ν is given by,

$$\nu = \frac{b}{2} \left[\frac{q}{b^2} + 2 \right] - \sum_{n=1}^{\infty} n^{-1} E \left[\max \left(0, \chi_{qn}^2 - n \left(1 + \frac{1}{b} \right) q \right) \right].$$

Since $L^* = L - s$, we have,

$$E(L) = \eta n^* + b\nu - bq^{-1}(b+1) + o(1).$$

Let us consider the difference,

$$D_a = L - \eta a \hat{\theta}_L^b. \quad (5.7)$$

It follows from Woodroffe (1977) that the mean of the asymptotic distribution of D_a is $b\nu$. From equation (5.5) and equation (5.7), we obtain,

$$\begin{aligned} \eta a E(\hat{\theta}_L^b) &= \eta n^* - q^{-1}b(b+1) + o(1), \\ \Rightarrow a E(\hat{\theta}_L^b) &= n^* - (\eta q)^{-1}b(b+1) + o(1). \end{aligned} \quad (5.8)$$

Equation (5.1) follows on combining equation (5.5) and equation (5.8). Let $h(L) = |(L - \eta n^*) / (\eta n^*)^{1/2}|$. It follows from a result of Bhattacharya and Mallik (1973) that $h(L) \xrightarrow{L} N(0, 2q^{-1}b^2)$ as $a \rightarrow \infty$ and from Theorem 2.3 of Woodroffe (1977), $h^2(L)$ is uniformly integrable for all $m > s + 2q^{-1}b$. Hence,

$$E[h^2(L)] = \frac{1}{2} + o(1). \quad (5.9)$$

Using (5.9), we have,

$$\begin{aligned} \text{Var}(N) &= \eta^{-2} \text{Var}(L) \\ &= \eta^{-2} \left[\eta n^* \left(2q^{-1}b^2 + o(1) \right) \right], \end{aligned}$$

and equation (5.2) holds. Proof of theorem 1 is now complete. \square

Theorem 2. For all $m > \max\{s + 2bq^{-1}, t, s + 2bq^{-1}(v^* - u), s + 2bq^{-1}(w - u)\}$, as $a \rightarrow \infty$, we have,

$$R_g(a) = n^{*-1} \left[\left\{ -(\eta q)^{-1}b(b+1) + \frac{1}{2} \right\} g'(1) + 2(\eta q)^{-1}b^2 g''(1) \right] + o(a^{-1}). \quad (5.10)$$

Proof: Substituting equation (3.3) and equation (3.5) in equation (3.6), we obtain

$$R_g(a) = E \left[g \left(\frac{N}{n} \right) \right] - g(1). \quad (5.11)$$

Expanding $g(\cdot)$ by Taylor's series, we obtain for $|U - 1| \leq |(N/n^*) - 1|$,

$$R_g(a) = n^{*-1} E[(N - n^*)g'(1)] + (2n^*)^{-1} E \left[\frac{(N - n^*)^2}{n^*} g''(U) \right]. \quad (5.12)$$

It immediately follows from equation (4.3) and the definition of U that, $U \xrightarrow{a.s.} 1$ as $a \rightarrow \infty$. Utilizing this result and the Lemma 3, we obtain from equation (5.12) that, for all

$$m > \max \left\{ s + 2q^{-1}b, t, s + 2q^{-1}b(v^* - u), s + 2q^{-1}b(w - u) \right\},$$

as $a \rightarrow \infty$,

$$R_g(a) = n^{*-1} \left[\left\{ -(\eta q)^{-1}b(b + 1) + \frac{1}{2} + o(1) \right\} g'(1) \right] + (n^*)^{-1} \left[2(\eta q)^{-1}n^*b^2 + o(a) \right] g''(1),$$

and the theorem follows. □

6. Applications of the proposed class of accelerated sequential procedures

In this section, we discuss some inferential problems based on various distributions useful in reliability theory and provide their solutions by using the proposed class of accelerated sequential procedures given in equation (3.1)–(3.2).

6.1. Minimum risk point estimation of the mean of a normal distribution

Let $X_i, i = 1, 2, \dots$, be a sequence of independent and identically distributed (i.i.d.) random variables from a normal distribution $N(\mu, \sigma^2)$, where $\mu \in (-\infty, \infty)$ and $\sigma^2 \in (0, \infty)$ are the unknown mean and variance respectively. Given a random sample X_1, X_2, \dots, X_n of size $n \geq 2$, let us define $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$ (sample mean) and $S_n^2 = (n - 1)^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ (sample variance) as the estimators of μ and σ^2 respectively. The problem is to estimate the population mean μ . Suppose the loss incurred in estimating μ by \bar{X}_n be

$$L(\mu, \bar{X}_n) = A (\bar{X}_n - \mu)^2 + cn, \tag{6.1}$$

where A is the known positive weight and c is known cost per unit observations. The risk corresponding to the loss function (6.1) comes out to be

$$R_n(c) = \frac{A\sigma^2}{n} + cn. \tag{6.2}$$

The value $n = n^*$ which minimizes the risk $R_n(c)$ is

$$n^* = \left(\frac{A}{c} \right)^{\frac{1}{2}} \sigma. \tag{6.3}$$

Since σ^2 is unknown, there does not exist any fixed sample size procedure which minimizes the risk $R_n(c)$ for all values of σ . Thus, motivated by equation (6.3), we adopt the following accelerated sequential procedure. Let us start with a sample of size $m \geq 2$, where $m = o(c^{1/2})$ and $\lim_{c \rightarrow 0} \sup(m/n^*) < 1$ and sampling sequentially with the stopping time L defined by

$$L = \inf \left\{ n \geq m; n \geq \eta \left(\frac{A}{c} \right)^{\frac{1}{2}} S_n \right\}. \tag{6.4}$$

Then, we move forward ahead and take $N - L$ observations, where

$$N = \max \left\{ L, \left\lceil \left(\frac{A}{c} \right)^{\frac{1}{2}} S_L \right\rceil + 1 \right\}. \tag{6.5}$$

After stopping, we estimate μ by \bar{X}_N . Comparing equation (6.3) with equation (2.1), we obtain $\theta = \sigma^2$, $a = (A/c)^{1/2}$, $b = 1/2$ and letting $c \rightarrow 0$ is equivalent to $a \rightarrow \infty$. Moreover, equation (2.2) holds for $q = 1$, $s = 1$, $\hat{\theta}_n = S_n^2$ and $\theta = \sigma^2$. The risk associated with the accelerated sequential procedure equations (6.4)–(6.5) is

$$R_N(c) = cn^* \left[E\left(\frac{n^*}{N}\right) + E\left(\frac{N}{n^*}\right) \right]. \quad (6.6)$$

Comparing equation (6.6) with equation (3.3), we obtain

$$g(x) = cn^* [x + x^{-1}], \quad g'(x) = cn^* [1 - x^{-2}], \quad g''(x) = 2cn^* x^{-3},$$

where x stands for N/n^* . Further, we have

$$R_{n^*}(c) = g(1) = 2cn^*.$$

Comparing $g''(x)$ with equation (3.4), we get $u = -1$, $v = 0$ and $w = 3$. Second-order approximations for the expected sample size and 'regret' can be obtained by making appropriate substitutions in theorems 1 and 2 and given by

$$E(N) = n^* - \frac{3}{4}\eta + \frac{1}{2} + o(1) \quad \text{and} \quad R_g(c) = \frac{c}{\eta} + o(c). \quad (6.7)$$

6.2. Minimum risk point estimation of the mean vector of a multivariate normal distribution

Let X_i , $i = 1, 2, \dots$, be a sequence of independent observations from a p -variate normal population $N_p(\mu, \sigma^2 \Sigma)$, where μ is a $p \times 1$ unknown mean vector, σ^2 is an unknown scalar and Σ is a known $p \times p$ positive definite matrix. Having recorded a sample X_i , $i = 1, 2, \dots, n$ of size $n \geq 2$, let us define

$$\bar{X}_n = n^{-1} \sum_{i=1}^n X_i,$$

$$\hat{\sigma}_n^2 = \{p(n-1)\}^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^T \Sigma^{-1} (X_i - \bar{X}_n),$$

as the unbiased and consistent estimators of μ and σ^2 , respectively. The problem is to estimate the mean vector μ . Let the loss incurred in estimating μ by \bar{X}_n be

$$L_n(\mu, \bar{X}_n) = A \left[(\bar{X}_n - \mu)^T \Sigma^{-1} (\bar{X}_n - \mu) \right] + cn, \quad (6.8)$$

where A and c are known positive constants. The risk comes out to be

$$R_n(c) = \frac{Ap\sigma^2}{n} + cn. \quad (6.9)$$

The sample size n^* which minimizes the risk is

$$n^* = \left(\frac{Ap}{c} \right)^{\frac{1}{2}} \sigma. \quad (6.10)$$

Since σ^2 is unknown, we propose the following accelerated sequential strategy to solve the problem

$$L = \inf \left\{ n \geq m; n \geq \eta \left(\frac{Ap}{c} \right)^{\frac{1}{2}} \hat{\sigma}_n \right\}, \tag{6.11}$$

$$N = \max \left\{ L, \left[\left(\frac{Ap}{c} \right)^{\frac{1}{2}} \hat{\sigma}_L \right] + 1 \right\}. \tag{6.12}$$

After stopping, we estimate μ by \bar{X}_N . Along with the lines of minimum risk estimation of normal mean, one can easily obtain the associated second-order approximations on making appropriate comparisons with the proposed class of accelerated sequential procedures. These expressions are given by

$$E(N) = n^* - \frac{3}{4\eta p} + \frac{1}{2} + o(1),$$

$$R_g(c) = \frac{c}{\eta p} + o(c^{\frac{1}{2}}).$$

6.3. Minimum risk point estimation of the location of a negative exponential distribution

Let $X_i, i = 1, 2, \dots$, be a sequence of *i.i.d.* random variables from the two parameter exponential distribution having *p.d.f.*,

$$f(x; \mu, \sigma) = \frac{1}{\sigma} \exp \left\{ -\frac{x - \mu}{\sigma} \right\}; \quad x > \mu, \sigma > 0, -\infty < \mu < \infty,$$

where μ and σ are unknown location and scale parameters, respectively. Given a random sample X_1, X_2, \dots, X_n from the distribution given above, let us define

$$X_{n(1)} = \min (X_1, X_2, \dots, X_n),$$

$$\hat{\sigma}_n = (n - 1)^{-1} \sum_{i=1}^n (X_i - X_{n(1)}),$$

as the estimators of μ and σ . The problem is to estimate μ . Let the loss incurred in estimating μ by $X_{n(1)}$ be

$$L_n (\mu, X_{n(1)}) = A (X_{n(1)} - \mu)^2 + cn, \tag{6.13}$$

where A and c are known positive constants. The risk comes out to be

$$R_n(c) = \frac{4A\sigma^2}{2n^2} + cn. \tag{6.14}$$

The sample size n^* which minimizes the risk is

$$n^* = \left(\frac{4A}{c} \right)^{\frac{1}{3}} \sigma^{\frac{1}{3}}. \tag{6.15}$$

Since σ is unknown, we propose the following accelerated sequential strategy to solve the problem

$$L = \inf \left\{ n \geq m; n \geq \eta \left(\frac{4A}{c} \right)^{\frac{1}{3}} \hat{\sigma}_n^{\frac{1}{3}} \right\}, \quad (6.16)$$

$$N = \max \left\{ L, \left\lceil \left(\frac{4A}{c} \right)^{\frac{1}{3}} \hat{\sigma}_L^{\frac{1}{3}} + 1 \right\rceil \right\}. \quad (6.17)$$

After stopping, we estimate μ by $X_{n(1)}$. On making the appropriate comparisons, we can easily obtain the values of average sample number and regret as

$$E(N) = n^* - \frac{3}{9\eta} + \frac{1}{2} + o(1),$$

$$R_g(c) = \frac{4c}{3\eta} + o(c^{\frac{1}{3}}).$$

6.4. Minimum risk estimation of the mean of an inverse Gaussian distribution

Let X_1, X_2, \dots, X_n be an *i.i.d.* sequence of inverse Gaussian random variables with a common density function,

$$f(x; \mu, \lambda) = \left\{ \frac{\lambda}{2\pi x^3} \right\}^{\frac{1}{2}} \exp \left\{ \frac{-\lambda(x - \mu)^2}{2x\mu^2} \right\}, \quad x > 0,$$

where $\mu > 0$ and $\lambda > 0$ are the mean and scale parameters, respectively. Having observed a random sample X_1, X_2, \dots, X_n of size $n(\geq 2)$ from the distribution given above, let us define, $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$ and $\hat{\lambda}_n^{-1} = n^{-1} \sum_{i=1}^n (X_i^{-1} - \bar{X}_n^{-1})$ as the estimators of μ and λ^{-1} , respectively. The problem is to estimation μ when λ is unknown. Suppose the loss incurred in estimating μ by \bar{X}_n be

$$L(\mu, \bar{X}_n) = \frac{A(\bar{X}_n - \mu)^2}{\mu^3} + cn\lambda^k, \quad (6.18)$$

where A, c and k are fixed known positive constants (Chaturvedi *et al.*, 2019a). The corresponding risk is given by

$$R_n(c) = \frac{A}{n\lambda} + cn\lambda^k. \quad (6.19)$$

The value $n = n^*$ minimizing the risk is given by

$$n^* = \left(\frac{A}{c} \right)^{\frac{1}{2}} \left(\frac{1}{\lambda} \right)^{\frac{k+1}{2}}, \quad (6.20)$$

and the corresponding minimum risk is

$$R_{n^*}(c) = 2c\lambda^k n^*. \quad (6.21)$$

However, in the absence of any knowledge about λ , no fixed sample size procedure minimizes the risk for all values of λ . In such situation, we adopt the following accelerated sequential procedure

$$L = \inf \left\{ n \geq m; n \geq \eta \left(\frac{A}{c} \right)^{\frac{1}{2}} \left(\frac{1}{\hat{\lambda}_n} \right)^{\frac{k+1}{2}} \right\}, \tag{6.22}$$

$$N = \max \left\{ L, \left\lceil \left(\frac{A}{c} \right)^{\frac{1}{2}} \left(\frac{1}{\hat{\lambda}_L} \right)^{\frac{k+1}{2}} \right\rceil + 1 \right\}. \tag{6.23}$$

After stopping, we estimate μ by \bar{X}_N . On making the appropriate comparisons, second-order approximations for the expected sample size and regret are given by

$$E(N) = n^* - \frac{(k+1)}{(k+3)} 2\eta + \frac{1}{2} + o(1),$$

$$R_g(c) = -\frac{c\lambda^k}{4\eta} (k^2 + 3) + o(c^{\frac{1}{2}}).$$

6.5. Minimum risk point estimation of the scale parameter of Pareto distribution

Let $X_i, i = 1, 2, \dots$, be a sequence of *i.i.d.* random variables from the Pareto distribution having *p.d.f.*,

$$f(x; \beta, \sigma) = \sigma^{-1} \beta^{\frac{1}{\sigma}} x^{-(1+\sigma^{-1})}; \quad x \geq \beta > 0, \quad \sigma > 0, \tag{6.24}$$

where β and σ are unknown scale and shape parameters respectively. We consider the transformation $Y_i = \log X_i, i = 1, 2, \dots, n$. Now the random variables Y_1, Y_2, \dots, Y_n are *i.i.d.* with *p.d.f.*

$$g(y; \log \beta, \sigma) = \sigma^{-1} \exp \left[-\frac{(y - \log \beta)}{\sigma} \right]; \quad y > \log \beta.$$

Let $X_{n(1)} = \min(X_1, X_2, \dots, X_n)$ and $Y_{n(1)} = \min(Y_1, Y_2, \dots, Y_n)$. We propose to estimate β by $X_{n(1)}$, which may be considered as equivalent to estimating $\log \beta$ by $Y_{n(1)}$. Let $\hat{\sigma}_n = (n-1)^{-1} \sum_{i=1}^n (Y_i - Y_{n(1)})$ be the estimator of σ . Suppose the loss incurred in estimating β by $X_{n(1)}$ be

$$L(\beta, X_{n(1)}) = A \left(\frac{X_{n(1)}}{\beta} - 1 \right)^2 + cn, \tag{6.25}$$

where A and c are known positive constants. The above loss function can be approximated by (Mukhopadhyay and Ekwo, 1987),

$$L(\log \beta, Y_{n(1)}) = A (Y_{n(1)} - \log \beta)^2 + cn. \tag{6.26}$$

The corresponding risk is

$$R_n = \frac{2A}{n^2} \sigma^2 + cn. \tag{6.27}$$

The value $n = n^*$ minimizing the risk is given by

$$n^* = \left(\frac{4A}{c} \right)^{\frac{1}{3}} (\sigma)^{\frac{2}{3}}, \tag{6.28}$$

and the corresponding minimum risk is

$$R_{n^*}(c) = \frac{3}{2}cn^*. \tag{6.29}$$

However, in the absence of any knowledge about σ , no fixed sample size procedure minimizes the risk for all values of σ . In such situation, we adopt the following accelerated sequential procedure

$$L = \inf \left\{ n \geq m; n \geq \eta \left(\frac{4A}{c} \right)^{\frac{1}{3}} (\hat{\sigma}_n)^{\frac{2}{3}} \right\}, \tag{6.30}$$

$$N = \max \left\{ L, \left\lceil \left(\frac{4A}{c} \right)^{\frac{1}{3}} (\hat{\sigma}_L)^{\frac{2}{3}} \right\rceil + 1 \right\}. \tag{6.31}$$

After stopping, we estimate $\log\beta$ by $Y_{N(1)}$. By making appropriate comparisons with the proposed class of accelerated sequential procedures, the expressions of expected sample size and regret are given by

$$E(N) = n^* - \frac{5}{9\eta} + \frac{1}{2} + o(1), \tag{6.32}$$

$$R_g(c) = \frac{4c}{3\eta} + o\left(c^{\frac{1}{3}}\right). \tag{6.33}$$

7. Analysis from simulations and real data

We will now showcase our proposed accelerated sequential procedures using extensive simulation and real data analyses. For brevity alone, we will only present results for estimating the scale parameter of a Pareto distribution, which precisely corresponds to Section 6.5.

7.1. Analysis from simulations

The following simulation results are for the accelerated sequential procedure outlined in equations (6.30)–(6.31). Table 1 contains these results. We first generated a set of pseudo random observations at-a-time from a Pareto distribution which is as given in equation (6.24). Each row in Table 1 corresponds to the results from 10,000 replications where results are tabulated for different combinations of β, σ, A, c, m and n^* , where n^* comes from equation (6.28). We first pick η ($0 < \eta < 1$) randomly from a Uniform(0, 1) distribution and run the sequential procedure.

Each block in Table 1 shows n^* (column 1), c (column 2), the values \bar{x}_{\min} along with their standard errors $s_{\bar{x}_{\min}}$ (column 3), the values $\bar{n}, s_{\bar{n}}$ (column 4), the ratio \bar{n}/n^* (column 5), the values \bar{r}_g (column 6) where \bar{r}_g denotes the “regret” function, κ (column 7), where $\kappa = 4c/3\eta$, as seen on the right side of equation (6.33). One should hence expect \bar{r}_g to be close to κ . The values $\bar{n} - n^*$ (column 8) and δ (column 9), where $\delta = -5/9\eta + 1/2$ as seen on the right side of equation (6.32). One should hence expect $\bar{n} - n^*$ to be close to δ .

As one can note, the values of \bar{n} seem to estimate n^* reasonably well across the rows, which become closer for larger values of n^* , indicating a superior first-order efficiency. The $\bar{x}_{\min}(= e^{\bar{y}_{\min}})$ values are also seen to be very good at estimating the true β value of 3 under each case. All the regret values are also seen to be close to their corresponding κ values, whereas all the $\bar{n} - n^*$ values are also seen to be close to the respective δ values, suggesting a good second-order efficiency.

Table 1: Simulation results from 10,000 replications of the accelerated sequential methodology equations (6.30)–(6.31) with $m = 10, A = 100$

$\beta = 3, \sigma = 0.5, A = 100$								
n^*	c	$\bar{x}_{\min}, s_{\bar{x}_{\min}}$	$\bar{n}, s_{\bar{n}}$	\bar{n}/n^*	\bar{r}_g	κ	$\bar{n} - n^*$	δ
50	8e-4	3.0299, 0.0009	49.54, 0.1339	0.99	0.0011	0.0004	-0.46	-0.76
100	1e-4	3.0161, 0.0005	99.27, 0.1793	0.99	1.4e-4	5.3e-4	-0.73	-0.75
250	6.4e-6	3.0060, 0.0002	249.85, 0.6742	0.99	3.3e-5	1.8e-5	-0.15	-0.31
500	8e-7	3.0030, 0.0001	499.61, 1.2893	0.99	7.7e-6	4.1e-6	-0.39	-0.59
650	3.7e-7	3.0023, 7.3e-5	649.36, 0.5544	0.99	5.3e-7	3e-7	-0.64	-0.44
800	2e-7	3.0018, 6e-5	799.49, 0.6218	0.99	3.2e-7	1.4e-7	-0.51	-0.58

Table 2: Analysis of breaking stress data using accelerated sequential procedure equations (6.30)–(6.31) with $m = 5, A = 100, \hat{\beta} = 0.39, \hat{\sigma} = 1.8188$

n^*	c	x_{\min}	n	n/n^*	r_g	ξ
30	0.0490	0.32	28	0.93	0.0107	0.1025
50	0.0105	0.31	53	1.06	0.0026	0.0221
60	0.0061	0.35	62	1.03	0.0050	0.0128
70	0.0038	0.34	75	1.07	0.0018	0.0080
80	0.0025	0.38	82	1.03	0.0008	0.0054

7.2. Analysis from real data

We now present analysis using a real data, by implementing our proposed accelerated sequential strategy. The dataset consists of breaking stress on carbon fibers (in Gba). This dataset has been studied and analyzed by many researchers, Fatima and Roohi (2015) who fitted a transmuted exponentiated Pareto distribution, Aljarrah *et al.* (2015) who fitted a Weibull-Pareto distribution to the data, Jayakumar *et al.* (2018) or Nichols and Padgett (2006). A regular Pareto distribution also turned out to be a viable model, with a p -value of 0.17 using a Kolmogorov-Smirnov test.

The full data consists of breaking stress of 100 carbon fibers. Treating these data as the universe, the maximum likelihood estimates were $\hat{\beta} = 0.39$ and $\hat{\sigma} = 1.8188$. An accelerated sequential procedure equations (6.30)–(6.31) was then applied to draw observations from the full set of data as needed. Table 2 contains these results, which are from a single run.

8. Concluding remarks and future research directions

We have developed a general class of accelerated sequential procedures and obtained the associated second-order approximations for the expected sample size and ‘regret’ function. We have discussed the applications of the proposed class to estimate the parameters of various distributions such as normal, exponential, Pareto, inverse Gaussian, multivariate normal etc. Considering the special case of Pareto distribution, we have also studied the associated properties via simulations and presented a real data set on carbon fibers in support of the practical aspect of the proposed methodology. In application part, we deliberately focus on one particular topic (minimum risk point estimation) for brevity alone. One can easily tackle the problems of bounded risk point estimation and confidence interval (region) estimation respectively using the proposed class of accelerated sequential procedures. One may also consider these problems for some other models under various loss structures.

Moreover, there is a wide scope to extend the idea of this paper because several other problems fall under the proposed set-up. One such area is ‘‘ranking and selection’’ where the problems like selection of the largest of k normal means (Chaturvedi and Gupta, 1996) or selection of the best exponential population (Mukhopadhyay and Solanky, 1992a) can be considered. Besides this, the problems of

estimation of parameters in an intraclass model, random one-way model, linear regression model (Mukhopadhyay, 1974; Chaturvedi 1987), multiple comparison procedure, estimation of contrasts in the means (Bhargava and Srivastava, 1973). are the suitable choices to be tackled using the proposed class of accelerated sequential procedures.

One may also think about the development of other multi-stage sampling designs (two-stage, three-stage or k -stage) in a similar manner. A sequential sampling design has been already developed by Joshi and Bapat (2020b).

Acknowledgments

The authors are extremely grateful to the anonymous reviewers for their comments and suggestions on the original version of this manuscript leading to an improved presentation. Moreover, the first author Neeraj Joshi is indebted to the Department of Science and Technology, Government of India for providing financial support for this research work under the INSPIRE fellowship program (Grant No. - IF170889).

References

- Aljarrah MA, Famoye F, and Lee C (2015). A new Weibull-Pareto distribution, *Communications in Statistics - Theory and Methods*, **44**, 4077–4095.
- Bhargava RP and Srivastava MS (1973). On Tukey's confidence intervals for the contrasts in the means of the intraclass correlation model, *Journal of the Royal Statistical Society, Series B*, **35**, 147–152.
- Bhat BR (1981). *Modern Probability Theory*, Wiley Eastern Limited, New York.
- Bhattacharya PK and Mallik A (1973). Asymptotic normality of the stopping times of some sequential procedures, *Annals of Statistics*, **1**, 1203–1211.
- Chattopadhyay S (2000). Sequential estimation of exponential location parameter using an asymmetric loss function, *Communications in Statistics - Theory and Methods*, **29**, 783–795.
- Chattopadhyay S and Sengupta RN (2006). Three-stage and accelerated sequential point estimation of the normal mean using LINEX loss function, *Statistics*, **40**, 39–49.
- Chaturvedi A (1987). Sequential point estimation of regression parameters in a linear model, *Annals of the Institute of Statistical Mathematics*, **39**, 55–67.
- Chaturvedi A and Gupta R (1994). Linking the estimation and ranking and selection problems through sequential procedures: the normal case, *Statistics and Probability Letters*, **20**, 273–285.
- Chaturvedi A and Gupta R (1996). On multi-stage procedures for estimating the largest mean of k normal populations having unequal and unknown variances, *Scandinavian Actuarial Journal*, **1**, 64–78.
- Chaturvedi A and Tomar SK (2003). Three-stage and 'accelerated' sequential procedures for the mean of a normal population with known coefficient of variation, *Statistics*, **37**, 51–64.
- Chaturvedi A, Bapat SR, and Joshi N (2019a). Multi-stage point estimation of the mean of an inverse Gaussian distribution, *Sequential Analysis*, **38**, 1–25.
- Chaturvedi A, Bapat SR, and Joshi N (2019b). Multi-stage procedures for the minimum risk and bounded risk point estimation of the location of negative exponential distribution under the modified LINEX loss function, *Sequential Analysis*, **38**, 135–162.
- Dantzig GB (1940). On the non-existence of tests of Student's hypothesis having power functions independent of σ , *Annals of Mathematical Statistics*, **11**, 186–192.
- Datta S and Mukhopadhyay N (1998). Second-order asymptotics for multistage methodologies in

- partitioning a set of normal populations having a common unknown variance, *Statistics and Risk Modeling*, **16**, 191–206.
- Fatima A and Roohi A (2015). Transmuted exponentiated Pareto-I distribution, *Pakistan Journal of Statistics*, **32**, 63–80.
- Ghosh M, Mukhopadhyay N, and Sen PK (1997). *Sequential Estimation*, Wiley, New York.
- Hall P (1983). Sequential estimation saving sampling operations, *Journal of the Royal Statistical Society, Series B*, **45**, 19–223.
- Hamdy HI and Son MS (1991). On accelerating sequential procedures for point estimation: the normal case, *Statistica*, **51**, 437–446.
- Hu J (2021). Improving Hall's accelerated sequential procedure: generalized multistage fixed-width confidence intervals for a normal mean, *Methodology and Computing in Applied Probability*, **23**, 823–835.
- Jayakumar K, Krishnan B, and Hamedani GG (2018). On a new generalization of Pareto distribution and its applications, *Communications in Statistics-Simulation and Computation*, **49**, 1264–1284.
- Joshi N and Bapat SR (2020a). On improved accelerated sequential estimation of the mean of an inverse Gaussian distribution, *Communications in Statistics-Theory and Methods*.
- Joshi N and Bapat SR (2020b). On a Class of Purely Sequential Procedures with Applications to Estimation and Ranking and Selection Problems, *Submitted*.
- Joshi N, Bapat SR, and Shukla AK (2021). Multi-stage estimation methodologies for an inverse Gaussian mean with known coefficient of variation, *American Journal of Mathematical and Management Sciences*.
- Mukhopadhyay N (1974). Sequential estimation of regression parameters in Gauss-Markoff setup, *Journal of Indian Statistical Association*, **12**, 39–43.
- Mukhopadhyay N (1996). An alternative formulation of accelerated sequential procedures with applications to parametric and nonparametric estimation, *Sequential Analysis*, **15**, 253–269.
- Mukhopadhyay N and de Silva BM (2009). *Sequential Methods and Their Applications*, Boca Raton: CRC.
- Mukhopadhyay N and Ekwo ME (1987). Sequential estimation problems for the scale parameter of a Pareto distribution, *Scandinavian Actuarial Journal*, **1**, 83–103.
- Mukhopadhyay N and Khariton Y (2020). Minimum risk point estimation (MRPE) of the mean in an exponential distribution under powered absolute error loss (PAEL) due to estimation plus cost of sampling, *Sequential Analysis*, **39**, 241–268.
- Mukhopadhyay N and Solanky TKS (1991). Second order properties of accelerated stopping times with applications in sequential estimation, *Sequential Analysis*, **10**, 99–123.
- Mukhopadhyay N and Solanky TKS (1992a). Accelerated sequential procedure for selecting the best exponential population, *Journal of Statistical Planning and Inference*, **32**, 347–361.
- Mukhopadhyay N and Solanky TKS (1992b). Accelerated sequential procedure for selecting the largest mean, *Sequential Analysis*, **11**, 137–148.
- Mukhopadhyay N and Solanky TKS (1993). A nonparametric accelerated sequential procedure for selecting the largest center of symmetry, *Journal of Nonparametric Statistics*, **3**, 155–166.
- Nichols MD and Padgett WJ (2006). A bootstrap control chart for Weibull percentiles, *Quality and Reliability Engineering International*, **22**, 141–51.
- Son MS and Hamdy HI (1990). On accelerating sequential procedures for estimating exponential distributions, *Communications in Statistics-Theory and Methods*, **9**, 3315–3324.
- Starr N (1966). On the asymptotic efficiency of a sequential procedure for estimating the mean, *Annals of Mathematical Statistics*, **37**, 1173–1185.

- Stein C (1945). A two-sample test for a linear hypothesis whose power is independent of the variance, *Annals of Mathematical Statistics*, **16**, 243–258.
- Tukey JW (1938). On the distribution of the fractional part of a statistical variable, *Matematicheskii Sbornik*, **4**, 561–562.
- Woodroofe M (1977). Second-order approximation for sequential point and interval estimation, *Annals of Statistics*, **5**, 984–995.
- Zhuang Y and Bhattacharjee D (2021). Minimum risk point estimation of the size of a finite population under mark-recapture strategy, *Sequential Analysis*, **40**, 243–258.

Received May 07, 2021; Revised June 26, 2021; Accepted June 26, 2021