J. Appl. Probab. 1–23 (2024)
*doi:*10.1017/jpr.2024.100

ACCELERATED GROUP SEQUENTIAL SAMPLING

JUN HU,* *Oakland University*
YAN ZHUANG,** *Connecticut College*

Abstract

We propose a novel and unified sampling scheme, called the accelerated group sequential sampling scheme, which incorporates four different types of sampling scheme: (i) the classic Anscombe–Chow–Robbins purely sequential sampling scheme; (ii) the accelerated sequential sampling scheme; (iii) the relatively new k-at-a-time group sequential sampling scheme; and (iv) the new k-at-a-time accelerated group sequential sampling scheme. The first-order and second-order properties of this unified sequential sampling scheme are fully investigated with two illustrations on minimum risk point estimation for the mean of a normal distribution and on bounded variance point estimation for the location parameter of a negative exponential distribution. We also provide extensive Monte Carlo simulation studies and real data analyses for each illustration.

Keywords: Minimum risk point estimation; bounded variance point estimation; Monte Carlo simulations; real data analyses

2010 Mathematics Subject Classification: Primary 62L12
Secondary 62L05; 62L10

1. Introduction

In statistical inference problems no fixed-sample-size procedure exists, sequential sampling schemes have been developed and widely used with efficiency properties proved in terms of the sample size required. The fundamental theories of sequential estimation are largely based on the ground-breaking papers [2, 3], in which purely sequential sampling methodologies were developed for the problem of constructing fixed-width confidence intervals (FWCIs).

In a parallel path, [27] originally formulated the minimum risk point estimation (MRPE) problem. Under the absolute error loss plus linear cost of sampling, a purely sequential stopping rule was proposed to estimate an unknown normal mean μ when the variance σ^2 was assumed unknown. Then, [31, 32] considered a more general loss function and proved a number of interesting asymptotic first- and second-order properties of the purely sequential MRPE methodology. Using nonlinear renewal theoretical tools developed in [13, 14], [34] further developed explicit second-order approximations associated with *efficiency*, *risk efficiency*, and

Received 5 July 2022; accepted 19 September 2024.
* Postal address: Department of Mathematics and Statistics, Oakland University, 146 Library Drive, Rochester, MI 48309, USA. Email: junhu@oakland.edu
** Postal address: Department of Mathematics and Statistics, Connecticut College, 270 Mohegan Avenue, New London, CT 06320, USA. Email: yzhuang@conncoll.edu

regret. Moreover, [5] provided a different method to evaluate the expression for regret, which could generalize the corresponding result in [34].

Let us begin with a sequence of independent and identically distributed (i.i.d.) positive and continuous random variables $\{W_n, n \geq 1\}$. For simplicity, we assume that these random variables have all positive moments finite, with mean $\mathbb{E}[W_1] = \theta$ and variance $\mathbb{V}[W_1] = \tau^2$ specified in particular. In addition, the distribution function of W_1 satisfies the condition $\mathbb{P}(W_1 \leq x) \leq Bx^{\alpha}$ for all $x > 0$ and for some $B > 0$ and $\alpha > 0$, both free from x. In the light of [34, (2.5)], we further assume this condition. One may wonder why this condition is seemingly irrelevant or unused anywhere, but it is a necessary condition for the results in Theorem 1. For more details, see [34]. Similar to [34, (1.1)], all the stopping times arising from the aforementioned inference problems can be written in a general form given by

$$t_0 = \inf\left\{n \geq m\colon n^{-1}\sum_{i=1}^{n} W_i \leq \theta(n/n^*)^{\delta} l_1(n)\right\}, \tag{1}$$

where $\delta > 0$ is a positive constant, $l_1(n) = 1 + l_0 n^{-1} + o(n^{-1})$ as $n \to \infty$ with $-\infty < l_0 < \infty$ a convergent sequence of numbers, $m \geq 1$ indicates a pilot sample size, and n^* is called an *optimal fixed sample size* whose expression is to be determined in specific problems. See [18, Section A.4] for more details.

The stopping rule (1) is implemented as follows. After an initial sample of size m is gathered, one observation is taken at a time as needed successively. Each time when a new observation is recorded, the sample data is evaluated to check with the stopping rule and sampling terminates at the first time that the stopping rule is satisfied. Therefore, this is a purely sequential sampling scheme; we denote it by $\mathcal{M}_0$.

To introduce the properties of the stopping time t_0 and relate its expected value to the optimal fixed sample size, let us define a general function of $\eta(k)$ as follows. For each integer $k \geq 1$,

$$\eta(k) = \frac{k}{2} - \frac{1}{2}\delta^{-1}\theta^2\tau^2 - \delta^{-1}l_0 - (\delta\theta)^{-1}\sum_{n=1}^{\infty} n^{-1}\mathbb{E}\left[\left\{\sum_{i=1}^{kn} W_i - kn(\delta+1)\theta\right\}^{+}\right], \tag{2}$$

where $\{u\}^+ = \max\{0, u\}$. Under certain conditions, [34] fully studied the properties of t_0, which are summarized in the following theorem.

Theorem 1. *For the purely sequential sampling scheme $\mathcal{M}_0$ and the stopping time t_0 given in (1), as $n^* \to \infty$, if $m > (\alpha\delta)^{-1}$, $\mathbb{E}_{\theta,\tau}[t_0 - n^*] = \eta(1) + o(1)$.*

In the spirit of [7], we define the term 'sampling operation' as the procedure of collecting new observations and evaluating the sample data to make a decision. Let φ denote the number of sampling operations. For the purely sequential sampling scheme $\mathcal{M}_0$ associated with the stopping time t_0 given in (1), we have $\varphi_{\mathcal{M}_0} = t_0 - m + 1$ and

$$\mathbb{E}_{\theta,\tau}[\varphi_{\mathcal{M}_0}] = n^* + \eta(1) - m + 1 + o(1). \tag{3}$$

Not surprisingly, the purely sequential sampling scheme $\mathcal{M}_0$ requires a lot of sampling operations. In view of this, [7] proposed an accelerated sequential estimation methodology saving sampling operations. Furthermore, [22] and [16] developed alternative formulations of the accelerated sequential sampling technique. The purely sequential sampling scheme of Anscombe–Chow–Robbins was improved in [15], which proposed a new sequential methodology requiring substantially fewer sampling operations. Accelerated sequential sampling

methodologies first draw samples sequentially part of the way and then augment with sampling in one single batch. As per the discussions from [18, p. 228]: 'An accelerated sequential strategy would always be operationally much more convenient in practical implementation than its sequential counterpart!'

On the other hand, [8] considered sampling in bulk or groups, rather than one at a time, and proposed group sequential sampling with variable group sizes. A concept of *sequential planning* was presented in [29] as an extension and generalization of group sequential procedures. Implementation of these sampling schemes has been proved to require only a few groups to cross the stopping boundary, leading to only a moderate increase in sample size.

Most recently, [24] first brought up a new type of sequential sampling scheme which considers recording k observations at a time, given the thoughts that in real life packaged items purchased in bulk often cost less per unit than the cost of an individual item. This new sampling strategy was discussed in both FWCI and MRPE problems for the mean of a normal population. These problems were revisited in [25] which incorporated the newly constructed estimators under permutations within each group for the stopping boundaries, leading to tighter estimation of required sample sizes. A double-sequential sampling scheme was developed in [9] defined as k-at-a-time part of the way, and then one-at-a-time sequentially, which requires similar sample sizes to the purely sequential strategies but saves sampling operations. Sequential estimation strategies for big data science with minimal computational complexities were further proposed in [21], with the idea of k-tuples instead of one single observation.

In this paper, we incorporate these path-breaking ideas with modification to accelerate the purely sequential sampling scheme without sacrificing the first- and second-order efficiency: (i) using the purely sequential methodology to determine only a proportion $\rho(0 < \rho < 1)$ of the desired final sample, and then augmenting with sampling in one single batch; and (ii) drawing a fixed number $k(k \geq 2)$ observations at a time successively until termination in the sequential sampling portion. In this way, we expect to save roughly $100(1 - k^{-1}\rho)\%$ of sampling operations for any given combination of k and ρ values compared with a purely sequential strategy where $k = 1$ and $\rho = 1$. Taking all possible combinations of $k \geq 1$ and $0 < \rho \leq 1$ values into consideration, we propose a novel and unified accelerated group sequential sampling scheme, denoted by $\mathcal{M}(\rho, k)$, in Section 2 along with a number of desirable properties. The sampling scheme $\mathcal{M}(\rho, k)$ encompasses a wide range of sampling procedures with different selections of k and ρ values:

(i) $k = 1$, $\rho = 1$, the purely sequential sampling scheme, which was originally established in [2, 3];

(ii) $k = 1$, $0 < \rho < 1$, the accelerated sequential sampling scheme first established in [7], with a unified version proposed in [22, 16];

(iii) $k > 1$, $\rho = 1$, the k-at-a-time group sequential sampling scheme first brought up in [24, 25];

(iv) $k \geq 2$, $0 < \rho < 1$, the k-at-a-time accelerated group sequential sampling scheme proposed as a new sampling scheme in this paper.

Remark 1. This work provides a sampling scheme that is both unified and novel because it incorporates the traditional sampling strategies (purely sequential and accelerated sequential schemes), the relatively new k-at-a-time group sequential sampling scheme, and the new k-at-a-time accelerated sequential sampling scheme, all under one big umbrella. We also provide first- and second-order properties for this unified sampling scheme in general.

The rest of this paper is organized as follows. Section 2 proposes the accelerated group sequential sampling scheme $\mathcal{M}(\rho, k)$ and explores its appealing first- and second-order properties, with a special focus on the k-at-a-time accelerated group sequential sampling scheme. In Section 3, we construct MRPE for an unknown normal mean μ with the variance σ^2 also assumed unknown as a possible illustration of the newly proposed sequential sampling scheme $\mathcal{M}(\rho, k)$. In Section 4, we construct bounded variance point estimation for an unknown location parameter μ from a negative exponential distribution. Simulated performances and real data analyses are included to support and supplement our theory for both methodologies in Sections 3 and 4. Section 5 shares some brief concluding thoughts.

2. The accelerated group sequential sampling scheme $\mathcal{M}(\rho, k)$

In this section, we provide a general framework of accelerated group sequential sampling in detail. We start with the stopping rules of sampling, and we also include the appealing properties under this sampling scheme.

Under the assumptions given in Section 1, we propose a novel and unified sequential sampling scheme $\mathcal{M}(\rho, k)$ associated with the following stopping times modified in view of (1):

$$\begin{aligned} t_1 &\equiv t_1(\rho, k) = \inf\left\{ n \geq m \colon (kn)^{-1} \sum_{i=1}^{n} U_i \leq \theta [kn/(\rho n^*)]^{\delta} l_k(n) \right\}, \\ t_2 &\equiv t_2(\rho, k) = \lfloor \rho^{-1} k t_1(\rho, k) \rfloor + 1. \end{aligned} \tag{4}$$

In addition to the notation in the stopping rule (1), $k \geq 1$ is a prefixed integer, $0 < \rho \leq 1$ is a prefixed proportion, $l_k(n) = 1 + l_0(kn)^{-1} + o(n^{-1})$ as $n \to \infty$ with $-\infty < l_0 < \infty$ is a convergent sequence of numbers, $U_i \equiv \sum_{j=(i-1)k+1}^{ik} W_j$, $i = 1, 2, \ldots$, are i.i.d. random variables, and $\lfloor u \rfloor$ denotes the largest integer that is strictly smaller than u ($< u$).

The implementation of the stopping rule (4) can be interpreted as follows. Starting with km pilot observations, we sample k observations at a time as needed successively and determine a preliminary sample size of $kt_1(\rho, k)$. Then, we continue to sample $t_2(\rho, k) - kt_1(\rho, k)$ additional observations if needed, all in one batch. Obviously, $\mathbb{P}_{\theta,\tau}(t_2(\rho, k) < \infty) = 1$ and $t_2(\rho, k) \uparrow \infty$ with probability 1 as $n^* \uparrow \infty$. If both ρ and k are chosen to be 1, then our newly developed sampling scheme $\mathcal{M}(1, 1)$ will be the purely sequential sampling scheme $\mathcal{M}_0$ associated with the stopping rule (1). If $\rho = 1$ and $k \geq 2$, the new sampling scheme $\mathcal{M}(1, k)$ is group sequential, but taking multiple (k) observations at a time. If $0 < \rho < 1$ and $k = 1$, the new sampling scheme $\mathcal{M}(\rho, 1)$ is an accelerated sequential sampling scheme, similar to the one proposed in [16].

Compared with the purely sequential sampling scheme $\mathcal{M}_0$, our newly developed sequential sampling scheme $\mathcal{M}(\rho, k)$ can reduce approximately $100(1 - k^{-1}\rho)\%$ of the operational time, which makes it flexible in real practice. One is allowed to choose the values of k and ρ to optimize the sampling process, time limitations, and cost considerations under different situations. We incorporate a brief discussion here to illustrate the flexibility of our sampling scheme.

If the cost of taking more observations is high, we can determine a smaller k value and/or choose ρ to be closer to 1; and if the process is more time sensitive, we can use the methodology with a larger k and/or choose ρ to be closer to 0. As a direct result of the operational convenience, the new sequential sampling scheme $\mathcal{M}(0 < \rho < 1, k \geq 2)$ tends to oversample, but the increase in the expected sample size is bounded by an amount depending on ρ and

k. Moreover, we have a thorough investigation of the efficiency properties for the sampling scheme $\mathcal{M}(\rho, k)$ in what follows.

We now establish an appealing property around this unified sampling scheme. We suppose that the following limit operation holds in the spirit of [6]:

$$m \to \infty,\ n^* = O(m^r),\ \text{and}\ \limsup \frac{m}{n^*} < \rho, \tag{5}$$

where $r \geq 1$ is a fixed constant. In the light of (1) and Theorem 1, we are now in a position to state the major results of this paper as Theorem 2. See [34, Theorem 2.4] for more details.

Theorem 2. *For the accelerated group sequential sampling scheme $\mathcal{M}(\rho, k)$ and the stopping times $t_1(\rho, k)$ and $t_2(\rho, k)$ given in (4), for fixed $0 < \rho < 1$ and k, with $\eta(k)$ defined in (2), under the limit operation (5):*

$$\begin{aligned} &\mathbb{E}_{\theta,\tau}[kt_1(\rho, k) - \rho n^*] = \eta(k) + o(1), \\ &\rho^{-1}\eta(k) + o(1) \leq \mathbb{E}_{\theta,\tau}[t_2(\rho, k) - n^*] \leq \rho^{-1}\eta(k) + 1 + o(1). \end{aligned} \tag{6}$$

And if $\rho = 1$, then $\mathbb{E}_{\theta,\tau}[t_2(1, k) - n^*] = \mathbb{E}_{\theta,\tau}[kt_1(1, k) - n^*] = \eta(k) + o(1)$.

It is clear that when $0 < \rho < 1$ and $k \geq 2$, our new sampling scheme $\mathcal{M}(\rho, k)$ is expected to oversample up to $\rho^{-1}\eta(k) + 1 + o(1)$ observations. In terms of the number of sampling operations, it is not hard to obtain that $\varphi_{\mathcal{M}(\rho,k)} = t_1 - m + 1 + \mathbf{1}(\rho < 1)$, where $\mathbf{1}(D)$ stands for the indicator function of an event D. We also have

$$\mathbb{E}_{\theta,\tau}[\varphi_{\mathcal{M}(\rho,k)}] = k^{-1}[\rho n^* + \eta(k) - km] + 1 + \mathbf{1}(\rho < 1) + o(1). \tag{7}$$

Comparing (3) and (7), the accelerated group sequential sampling scheme $\mathcal{M}(\rho, k)$ requires roughly $100(1 - k^{-1}\rho)\%$ fewer sampling operations than those of the purely sequential sampling scheme $\mathcal{M}_0$, based on the actual choices of k and ρ. Therefore, it enjoys great operational convenience with the cost of only a slight increase in the projected final sample size.

Remark 2. According to Theorem 2, as the optimal sample size n^* tends to infinity, the extra sample size is expected to be a finite number around $\rho^{-1}\eta(k)$, while the saving in the number of sampling operations is expected to be $(1 - k^{-1}\rho)n^* + O(1)$. Given that the cost per sampled unit and the cost per sampled group are both positive constants, this indicates that the extra cost due to oversampling tends to be a finite number, while the saving due to acceleration tends to infinity. In this sense, accelerated group sequential sampling is still advantageous despite the possible overshoot. Moreover, as seen from the Monte Carlo simulation studies in Tables 2 and 6, the oversampling under discussion is within 5 for all the scenarios we simulated.

Hereafter, we mainly focus on the sampling scheme $\mathcal{M}(\rho, k)$ with $0 < \rho < 1$ and $k \geq 2$, which makes it specifically the k-at-a-time accelerated group sequential sampling scheme. Nevertheless, note that all the theories and methodologies we discuss generally work for the sequential sampling scheme with $0 < \rho \leq 1$ and/or $k \geq 1$.

3. Minimum risk point estimation for a normal mean

In this section, we discuss minimum risk point estimation (MRPE) for a normal mean as an illustration of our accelerated group sequential sampling scheme $\mathcal{M}(\rho, k)$. Having recorded a sequence of independent observations $X_1, \ldots, X_n$, $n \geq 2$, from an $N(\mu, \sigma^2)$ population where

both $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$ are unknown, we denote the sample mean, the sample variance, and the sample standard deviation as follows:

$$\begin{aligned} \text{Sample mean:} \quad & \bar{X}_n = n^{-1} \textstyle\sum_{i=1}^n X_i, \\ \text{Sample variance:} \quad & S_n^2 = (n-1)^{-1} \textstyle\sum_{i=1}^n (X_i - \bar{X}_n)^2, \\ \text{Sample standard deviation:} \quad & S_n = \sqrt{S_n^2}. \end{aligned}$$

According to [27], the MRPE for μ under the squared-error loss plus linear cost of sampling can be formulated as follows. Define the loss function by

$$L_n \equiv L_n(\mu, \bar{X}_n) = A(\bar{X}_n - \mu)^2 + cn, \tag{8}$$

where $A\,(>0)$ is a known weight function and $c\ (>0)$ is the known unit cost of each observation. Associated with the loss function in (8), we have the following risk function:

$$R_n \equiv \mathbb{E}_{\mu,\sigma}[L_n(\mu, \bar{X}_n)] = A\sigma^2 n^{-1} + cn,$$

which is minimized at

$$n^* \equiv n^*(c) = \sigma\sqrt{A/c}, \tag{9}$$

with the resulting minimum risk

$$R_{n^*} = 2cn^*. \tag{10}$$

Note no fixed-sample-size procedure exists that achieves the exact minimum risk due to the fact that σ is unknown. A fundamental solution is due to the purely sequential MRPE methodology in the light of [27, 31, 32], briefly introduced below.

Since the population standard deviation σ remains unknown, it is essential for us to estimate it customarily using the sample standard deviation S_n, and update its value at every stage. We can start with m pilot observations, $X_1, \ldots, X_m$, $m \geq 2$, and then sample one additional observation at a time as needed until the following stopping rule is satisfied:

$$\text{Methodology } \mathcal{P}_0\colon N_{\mathcal{P}_0} \equiv N_{\mathcal{P}_0}(c) = \inf\{n \geq m\colon n \geq S_n\sqrt{A/c}\}. \tag{11}$$

It is clear that $\mathbb{P}_{\mu,\sigma}\{N_{\mathcal{P}_0} < \infty\} = 1$ and $N_{\mathcal{P}_0} \uparrow \infty$ with probability 1 as $c \downarrow 0$. Upon termination with the accrued data $\{N_{\mathcal{P}_0}, X_1, \ldots, X_m, \ldots, X_{N_{\mathcal{P}_0}}\}$, we estimate the unknown normal mean μ with $\bar{X}_{N_{\mathcal{P}_0}} \equiv N_{\mathcal{P}_0}^{-1} \sum_{i=1}^{N_{\mathcal{P}_0}} X_i$. The achieved risk is then given by

$$R_{N_{\mathcal{P}_0}}(c) \equiv \mathbb{E}_{\mu,\sigma}\left[L_{N_{\mathcal{P}_0}}\left(\mu, \bar{X}_{N_{\mathcal{P}_0}}\right)\right] = A\mathbb{E}_{\mu,\sigma}\left[\left(\bar{X}_{N_{\mathcal{P}_0}} - \mu\right)^2\right] + c\mathbb{E}_{\mu,\sigma}[N_{\mathcal{P}_0}]. \tag{12}$$

To measure the closeness between the achieved risk in (12) and the minimum risk in (10), [27] and [31] respectively constructed the following two crucial notions, namely, the *risk efficiency* and *regret*:

$$\begin{aligned} \text{Risk efficiency:} \quad & \xi_{\mathcal{P}_0}(c) \equiv R_{N_{\mathcal{P}_0}}(c)/R_{n^*}(c) = \tfrac{1}{2}\mathbb{E}_{\mu,\sigma}[N_{\mathcal{P}_0}/n^*] + \tfrac{1}{2}\mathbb{E}_{\mu,\sigma}[n^*/N_{\mathcal{P}_0}]; \\ \text{Regret:} \quad & \omega_{\mathcal{P}_0}(c) \equiv R_{N_{\mathcal{P}_0}}(c) - R_{n^*}(c) = c\mathbb{E}_{\mu,\sigma}\left[N_{\mathcal{P}_0}^{-1}(N_{\mathcal{P}_0} - n^*)^2\right]. \end{aligned}$$

Alternatively, the stopping rule (11) can be rewritten in the way we presented (1). By using the Helmert transformation, we express $N_{\mathcal{P}_0} = N'_{\mathcal{P}_0} + 1$ with probability 1, where the new stopping time $N'_{\mathcal{P}_0}$ is defined as follows:

$$N'_{\mathcal{P}_0} = \inf\left\{n \geq m-1\colon n^{-1}\sum_{i=1}^n W_i \leq (n/n^*)^2(1 + 2n^{-1} + n^{-2})\right\}, \tag{13}$$

with $\delta = 2$, $l_0 = 2$, and $W_1, W_2, \ldots$ being i.i.d. χ_1^2 random variables such that $\theta = 1$, $\tau^2 = 1$, and $\alpha = \frac{1}{2}$. From [27], [31], and [34], we conclude the following theorem to address the asymptotic first- and second-order properties that the purely sequential MRPE methodology $\mathcal{P}_0$ enjoys. See [18] for more details.

Theorem 3. *For the purely sequential MRPE methodology $\mathcal{P}_0$ given in (11), for all fixed μ, σ, m, and A, as $c \to 0$:*

(i) *Asymptotic first-order efficiency:* $\mathbb{E}_{\mu,\sigma}[N_{\mathcal{P}_0}/n^*] \to 1$ *if* $m \geq 2$;

(ii) *Asymptotic second-order efficiency:* $\mathbb{E}_{\mu,\sigma}[N_{\mathcal{P}_0} - n^*] = \eta_1(1) + o(1)$ *if* $m \geq 3$, *where* $\eta_1(1) = -\frac{1}{2}\sum_{n=1}^{\infty} n^{-1}\mathbb{E}\big[\{\chi_n^2 - 3n\}^+\big]$;

(iii) *Asymptotic first-order risk efficiency:* $\xi_{\mathcal{P}_0}(c) \to 1$ *if* $m \geq 3$;

(iv) *Asymptotic second-order risk efficiency:* $\omega_{\mathcal{P}_0}(c) = \frac{1}{2}c + o(c)$ *if* $m \geq 4$.

3.1. The accelerated group sequential MRPE methodology

Following (6), we propose an accelerated group sequential MRPE methodology $\mathcal{P}(\rho, k)$:

$$\begin{aligned} T_{\mathcal{P}(\rho,k)} &\equiv T_{\mathcal{P}(\rho,k)}(c) = \inf\{n \geq 0 \colon m + kn \geq \rho S_{m+kn}\sqrt{A/c}\}, \\ N_{\mathcal{P}(\rho,k)} &\equiv N_{\mathcal{P}(\rho,k)}(c) = \lfloor \rho^{-1}(m + kT_{\mathcal{P}(\rho,k)})\rfloor + 1. \end{aligned} \tag{14}$$

Here, $0 < \rho \leq 1$ is a prefixed proportion, $k \geq 1$ is a prefixed positive integer, $m \geq 2$ again indicates a pilot sample size but picked such that $m - 1 \equiv 0 \pmod{k}$, and $\lfloor u \rfloor$ continues to denote the largest integer that is strictly smaller than u. Writing $m - 1 = m_0 k$ for some integer $m_0 \geq 1$, we further assume that the following limit operation holds:

$$m_0 \to \infty,\ m = m_0 k + 1 \to \infty,\ c \equiv c(m) = O(m^{-2r}),\ n^* = O(m^r),\ \text{and}\ \limsup \frac{m}{n^*} < \rho, \tag{15}$$

where $r \geq 1$ is a fixed constant. The new methodology $\mathcal{P}(\rho, k)$ is implemented as follows.

Starting with $m(= m_0 k + 1)$ pilot observations, $X_1, \ldots, X_m$, we sample k observations at a time as needed and determine $T_{\mathcal{P}(\rho,k)}$, which indicates the number of sequential sampling operations according to the stopping rule (14). Next, we continue to sample $(N_{\mathcal{P}(\rho,k)} - m - kT_{\mathcal{P}(\rho,k)})$ additional observations all in one batch. Upon termination, based on the fully gathered data

$$\{T_{\mathcal{P}(\rho,k)}, N_{\mathcal{P}(\rho,k)}, X_1, \ldots, X_m, \ldots, X_{m+kT_{\mathcal{P}(\rho,k)}}, \ldots, X_{N_{\mathcal{P}(\rho,k)}}\},$$

we construct the minimum risk point estimator $\bar{X}_{N_{\mathcal{P}(\rho,k)}} = N_{\mathcal{P}(\rho,k)}^{-1}\sum_{i=1}^{N_{\mathcal{P}(\rho,k)}} X_i$ for μ, and derive

$$\begin{aligned} \text{Risk efficiency:}\quad & \xi_{\mathcal{P}(\rho,k)}(c) \equiv R_{N_{\mathcal{P}(\rho,k)}}(c)/R_{n^*}(c) = \tfrac{1}{2}\mathbb{E}_{\mu,\sigma}[N_{\mathcal{P}(\rho,k)}/n^*] + \tfrac{1}{2}\mathbb{E}_{\mu,\sigma}[n^*/N_{\mathcal{P}(\rho,k)}]; \\ \text{Regret:}\quad & \omega_{\mathcal{P}(\rho,k)}(c) \equiv R_{N_{\mathcal{P}(\rho,k)}}(c) - R_{n^*}(c) = c\mathbb{E}_{\mu,\sigma}\big[N_{\mathcal{P}(\rho,k)}^{-1}(N_{\mathcal{P}(\rho,k)} - n^*)^2\big]. \end{aligned}$$

Obviously, $\mathbb{P}_{\mu,\sigma}(N_{\mathcal{P}(\rho,k)} < \infty) = 1$ and $N_{\mathcal{P}(\rho,k)} \uparrow \infty$ with probability 1 as $c \downarrow 0$. If both ρ and k are chosen to be 1, then the sequential MRPE methodology $\mathcal{P}(1, 1)$ will be the purely sequential MRPE methodology $\mathcal{P}_0$ as per (11). That is, $\mathcal{P}(1, 1) \equiv \mathcal{P}_0$.

Along the lines of (13), we can similarly express the stopping time $T_{\mathcal{P}(\rho,k)}$ from (14) in the general form provided in (4). Define $T_{\mathcal{P}(\rho,k)} = T'_{\mathcal{P}(\rho,k)} - m_0$ with probability 1. Then $T'_{\mathcal{P}(\rho,k)}$ is a new stopping time that can be rewritten as

$$\begin{aligned} T'_{\mathcal{P}(\rho,k)} &= \inf\{n \geq m_0 \colon kn + 1 \geq \rho S_{kn+1}\sqrt{A/c}\} \\ &= \inf\left\{n \geq m_0 \colon kn(kn+1)^2 \geq (\rho\sigma\sqrt{A/c})^2 \frac{knS_{kn+1}^2}{\sigma^2}\right\} \\ &= \inf\{n \geq m_0 \colon kn(kn+1)^2 \geq (\rho n^*)^2 \chi_{kn}^2\} \\ &= \inf\left\{n \geq m_0 \colon (kn)^{-1} \sum_{i=1}^{n} U_i \leq [kn/(\rho n^*]^2 (1 + 2(kn)^{-1})(kn)^{-2}\right\}, \end{aligned}$$

where $U_i = \sum_{j=(i-1)k+1}^{ik} W_j$, $i = 1, 2, \ldots$, with $W_1, W_2, \ldots$ being i.i.d. χ_1^2 random variables. We list the simplified version of the equation for $T'_{\mathcal{P}(\rho,k)}$ in (16) for easy reference:

$$T'_{\mathcal{P}(\rho,k)} = \inf\left\{n \geq m_0 \colon (kn)^{-1} \sum_{i=1}^{n} U_i \leq [kn/(\rho n^*)]^2 (1 + 2(kn)^{-1} + (kn)^{-2})\right\}, \tag{16}$$

where $\delta = 2$, $l_0 = 2$, and $U_i = \sum_{j=(i-1)k+1}^{ik} W_j$, $i = 1, 2, \ldots$, with $W_1, W_2, \ldots$ being i.i.d. χ_1^2 random variables such that $\theta = 1$, $\tau^2 = 2$, and $\alpha = \frac{1}{2}$. Therefore, $U_1, U_2, \ldots$ are i.i.d. χ_k^2 random variables.

Now we state a number of asymptotic first- and second-order properties of the accelerated group sequential MRPE methodology $\mathcal{P}(\rho, k)$, summarized in the following theorem.

Theorem 4. *For the accelerated group sequential MRPE methodology $\mathcal{P}(\rho, k)$ given in (14), for all fixed μ, σ, A, k and $0 < \rho < 1$, under the limit operation (15):*

(i) *Asymptotic first-order efficiency:* $\mathbb{E}_{\mu,\sigma}[N_{\mathcal{P}(\rho,k)}/n^*] \to 1$;

(ii) *Asymptotic second-order efficiency:* $\rho^{-1}\eta_1(k) + o(1) \leq \mathbb{E}_{\mu,\sigma}[N_{\mathcal{P}(\rho,k)} - n^*] \leq \rho^{-1}\eta_1(k) + 1 + o(1)$, *where* $\eta_1(k) = (k-1)/2 - \frac{1}{2}\sum_{n=1}^{\infty} n^{-1}\mathbb{E}\big[\{\chi_{kn}^2 - 3kn\}^+\big]$;

(iii) *Asymptotic first-order risk efficiency:* $\xi_{\mathcal{P}(\rho,k)}(c) \to 1$;

(iv) *Asymptotic second-order risk efficiency:* $\omega_{\mathcal{P}(\rho,k)}(c) = \frac{1}{2}\rho^{-1}c + o(c)$.

Again, when $\rho = 1$, we have the exact expression $\mathbb{E}_{\mu,\sigma}[N_{\mathcal{P}(1,k)} - n^*] = \eta_1(k) + o(1)$ instead of the inequality in Theorem 4(ii). The number of sampling operations for the accelerated group sequential MRPE methodology $\mathcal{P}(\rho, k)$ is

$$\varphi_{\mathcal{P}(\rho,k)} = T_{\mathcal{P}(\rho,k)} + 1 + \mathbf{1}(\rho < 1), \tag{17}$$

and

$$\mathbb{E}_{\mu,\sigma}[\varphi_{\mathcal{P}(\rho,k)}] = k^{-1}[\rho n^* - m + \eta_1(k)] + 1 + \mathbf{1}(\rho < 1) + o(1). \tag{18}$$

For any integer $k \geq 1$, $\eta_1(k) = (k-1)/2 - \frac{1}{2}\Sigma_{n=1}^{\infty} n^{-1}\mathbb{E}\big[\{\chi_{kn}^2 - 3kn\}^+\big]$ is computable. In order to obtain numerical approximations, we wrote our own R code and provide the values in Table 1. In the spirit of [23, Table 3.8.1], any term smaller than 10^{-15} in magnitude was

TABLE 1. $\eta_1(k)$ approximations in Theorem 4(ii).

k	$\eta_1(k)$	$k^{-1}\eta_1(k)$	k	$\eta_1(k)$	$k^{-1}\eta_1(k)$
1	−0.1165	−0.1165	11	4.9993	0.4545
2	0.4367	0.2183	12	5.4996	0.4583
3	0.9636	0.3212	13	5.9997	0.4615
4	1.4785	0.3696	14	6.4998	0.4643
5	1.9872	0.3974	15	6.9999	0.4667
6	2.4922	0.4154	16	7.4999	0.4687
7	2.9952	0.4279	17	8.0000	0.4706
8	3.4971	0.4371	18	8.5000	0.4722
9	3.9982	0.4442	19	9.0000	0.4737
10	4.4989	0.4499	20	9.5000	0.4750

excluded in the infinite sum with regard to $\eta_1(k)$. Intuitively, the infinite sum and $k^{-1}\eta_1(k)$ converge to zero and $\frac{1}{2}$ respectively as $k \to \infty$. However, by looking at the columns for $\eta_1(k)$ and $k^{-1}\eta_1(k)$, we can see that the infinite sum converges very quickly, while $k^{-1}\eta_1(k)$ converges at a rather slow rate.

Remark 3. The sign of $\eta(k)$ may imply whether a sequential sampling procedure leads to fewer (negative) or more (positive) observations than the optimal sample size on average upon termination. However, the sign depends on the specific inference problem and population distribution, and can be either positive or negative even when $k = 1$. This is because the magnitude of $\eta(k)$ involves problem-specific parameters δ, θ, τ^2, and l_0.

Remark 4. Readers may have found that our loss function (8), while classical in sequential analysis, only combined estimation error and the cost per sampled unit. The cost per sampled group, however, is not included. In fact, it can be of significance to account for both the cost per sampled group and the cost per sample unit in the loss function. For example, [29] considered a loss function given by $L_N \equiv L_N(\lambda, \hat{\lambda}_N) + \sum_{j=1}^{T}(cN_j + a)$, where c is the cost per sampled unit, a is the cost per sampled group, λ is the parameter under estimation, and $L_N(\lambda, \hat{\lambda}_N)$ represents the incurred loss due to the estimator $\hat{\lambda}_N$ and the total sample size $N = \sum_{j=1}^{T} N_j$, where T indicates the total number of sampled groups and N_j is the sample size of the jth group, $j = 1, \ldots, T$. Under this loss function, both the optimal group sizes N_j and the optimal number of groups T will need to be determined sequentially, which means they are both stopping variables. In our proposed accelerated group sequential sampling scheme, however, the group size k is a prefixed constant. This difference will complicate the sampling scheme, so we have neglected the cost per sampled group in the loss function (8). However, this would be a very interesting future project to work on.

3.2. Simulated performance

To investigate the appealing properties of the accelerated group sequential MRPE methodology $\mathcal{P}(\rho, k)$, and illustrate how it saves sampling operations with $0 < \rho < 1$ and/or $k \geq 2$, we conducted extensive sets of Monte Carlo simulations under the normal case in the spirit of [19]. To be specific, we generated pseudorandom samples from an $N(5, 2^2)$ population. While fixing the weight function $A = 100$, the pilot sample size $m = 21$, we selected a wide range

of values of c, the unit cost of sampling, including 0.04, 0.01, and 0.0025, so that the optimal fixed sample size n^* turned out to be 100, 200, and 400 accordingly. We also considered various combinations of $\rho = (1, 0.8, 0.5)$ and $k = (1, 2, 5)$ to compare the number of sampling operations under different possible scenarios. The findings are summarized in Table 2. For each methodology $\mathcal{P}(\rho, k)$, we computed the average total final sample size $\bar{n}$ with the associated standard error $s(\bar{n})$, the difference between $\bar{n}$ and n^* to be compared with the second-order efficiency term in Theorem 4(ii), the estimated risk efficiency $\widehat{\xi}$ to be compared with 1, the estimated regret in terms of unit cost $\widehat{\omega}/c$ to be compared with $\frac{1}{2}\rho^{-1}$ from Theorem 4(iv), and the average number of sampling operations $\bar{\varphi}$ to be compared with the expected number of sampling operations $\mathbb{E}(\varphi)$ from (18).

It is clear that across the board, $\bar{n} - n^*$ is close to the second-order approximation $\rho^{-1}\eta_1(k)$, $\hat{\xi}$ is close to 1, and $\hat{\omega}/c$ is close to the coefficient $\frac{1}{2}\rho^{-1}$. These empirically verify Theorem 4. Focusing on the last two columns, we can also easily find that the average number of sampling operations needed, $\bar{\varphi}$, is almost the same as the theoretical value $\mathbb{E}(\varphi)$, and the accelerated group sequential MRPE procedure $\mathcal{P}(\rho, k)$ reduces by approximately $100(1 - k^{-1}\rho)\%$ sampling operations in comparison to the Anscombe–Chow–Robbins purely sequential procedure $\mathcal{P}(1, 1)$. For example, when $n^* = 400$, $\mathcal{P}(1, 1)$ requires around 380 sampling operations on average, while $\mathcal{P}(0.8, 5)$ requires 62. So about $100(1 - 62/380)\% = 83.7\%$ sampling operations are saved, which is close to $100(1 - 0.8/5)\% = 84\%$.

3.3. Real data analysis

Next, to illustrate the applicability of our newly developed accelerated group sequential MRPE methodology $\mathcal{P}(\rho, k)$, we proceeded to analyze a real-life dataset on hospital infection data from [12]. This data is from 113 hospitals in the United States for the 1975–76 study period. Each line of the data set has an identification number and provides information on 11 other variables for a single hospital. One of these variables is the infection risk, which records the average estimated probability of acquiring an infection in hospital (in percent). With the cost of observations taken into consideration, it is of great interest to propose an MRPE for the infection risk.

We treated the real dataset on the infection risk, which seemed to follow a normal distribution, confirmed via the Shapiro–Wilk normality test with the associated p-value of 0.1339. The simple descriptive statistics from the whole dataset of infection risk are summarized in Table 3.

For illustrative purposes, we treated this dataset of infection risk with size 113 from [12] as our population, with both mean and variance assumed unknown. Then, we performed our accelerated group sequential MRPE methodologies to obtain minimum risk point estimators for the infection risk. To start, we first randomly picked $m = 11$ observations as a pilot sample, based upon which we proceeded with sampling according to the methodologies $\mathcal{P}(\rho, k)$ with $A = 100$, $c = 0.04$, $\rho = (1, 0.8, 0.5)$, $k = (1, 2, 5)$. We summarize the terminated sample sizes as well as the associated numbers of sampling operations under each setting in Table 4, where $\mathcal{P}(\rho, k)$ denotes a certain sampling procedure with fixed values of ρ and k, and $n_{\mathcal{P}(\rho,k)}$ and $\varphi_{\mathcal{P}(\rho,k)}$ indicate the respective terminated sample size and number of sampling operations performing $\mathcal{P}(\rho, k)$ accordingly.

From Table 4, we can see that our terminated sample size ranges from 54 to 77. Apparently, many fewer sampling operations are needed when we fix $k = 2, 5$, without increasing to a significant number of observations. Also, we need the fewest observations with the fewest sampling operations when we use the methodology $\mathcal{P}(\rho, k)$ with $\rho = 0.5$, compared with larger

TABLE 2. Simulations from $N(5, 2^2)$ with $A = 100$ and $m = 21$ from 10 000 runs implementing $\mathcal{P}(\rho, k)$ from (14).

$\mathcal{P}(\rho, k)$	$\bar{n}$	$s(\bar{n})$	$\bar{n} - n^*$	$\rho^{-1}\eta_1(k)$	$\widehat{\xi}$	$\frac{1}{2}\rho^{-1}$	$\widehat{\omega}/c$	$\bar{\varphi}$	$\mathbb{E}(\varphi)$
$n^* = 100$, $c = 0.04$									
$\mathcal{P}(1, 1)$	99.8528	0.071 82	−0.1472	−0.1165	0.992 24	0.5	0.533 04	79.853	79.883
$\mathcal{P}(1, 2)$	100.4296	0.072 31	0.4296	0.4367	0.992 39	0.5	0.533 90	40.715	40.718
$\mathcal{P}(1, 5)$	102.0110	0.073 40	2.0110	1.9872	0.993 14	0.5	0.562 11	17.202	17.197
$\mathcal{P}(0.8, 1)$	100.1738	0.081 24	0.1738	−0.1456	0.992 91	0.625	0.682 28	60.836	60.883
$\mathcal{P}(0.8, 2)$	101.0466	0.080 49	1.0466	0.5459	0.993 39	0.625	0.663 20	31.718	31.718
$\mathcal{P}(0.8, 5)$	102.8798	0.083 68	2.8798	2.4840	0.994 09	0.625	0.748 73	14.195	14.197
$\mathcal{P}(0.5, 1)$	99.8002	0.103 24	−0.1998	−0.2330	0.995 13	1	1.142 53	30.900	30.884
$\mathcal{P}(0.5, 2)$	100.8176	0.104 05	0.8176	0.8734	0.995 24	1	1.124 74	16.704	16.718
$\mathcal{P}(0.5, 5)$	103.8990	0.106 64	3.8990	3.9744	0.996 97	1	1.215 88	8.190	8.197
$n^* = 200$, $c = 0.01$									
$\mathcal{P}(1, 1)$	199.9278	0.100 18	−0.0722	−0.1165	0.9965 2	0.5	0.508 80	179.928	179.884
$\mathcal{P}(1, 2)$	200.4926	0.100 56	0.4926	0.4367	0.996 46	0.5	0.509 28	90.746	90.718
$\mathcal{P}(1, 5)$	202.0495	0.101 08	2.0495	1.9872	0.996 64	0.5	0.523 08	37.210	37.197
$\mathcal{P}(0.8, 1)$	200.3638	0.112 58	0.3638	−0.1456	0.996 80	0.625	0.641 17	140.991	140.884
$\mathcal{P}(0.8, 2)$	201.0546	0.112 08	1.0546	0.5459	0.996 86	0.625	0.633 42	71.721	71.718
$\mathcal{P}(0.8, 5)$	202.8111	0.113 21	2.8111	2.4840	0.997 09	0.625	0.663 53	30.189	30.197
$\mathcal{P}(0.5, 1)$	199.7988	0.143 67	−0.2012	−0.2330	0.997 76	1	1.061 74	80.899	80.884
$\mathcal{P}(0.5, 2)$	200.9172	0.143 57	0.9172	0.8734	0.997 78	1	1.048 09	41.729	41.718
$\mathcal{P}(0.5, 5)$	203.9490	0.145 31	3.9490	3.9744	0.998 12	1	1.097 37	18.195	18.197

TABLE 2. Continued.

$\mathcal{P}(\rho, k)$	$\bar{n}$	$s(\bar{n})$	$\bar{n} - n^*$	$\rho^{-1}\eta_1(k)$	$\widehat{\xi}$	$\frac{1}{2}\rho^{-1}$	$\widehat{\omega}/c$	$\bar{\varphi}$	$\mathbb{E}(\varphi)$
$n^* = 400$, $c = 0.0025$									
$\mathcal{P}(1, 1)$	399.8625	0.140 74	−0.1375	−0.1165	0.998 05	0.5	0.497 88	379.863	379.883
$\mathcal{P}(1, 2)$	400.4010	0.141 55	0.4010	0.4367	0.998 03	0.5	0.502 78	190.701	190.718
$\mathcal{P}(1, 5)$	401.9310	0.141 69	1.9310	1.9872	0.998 06	0.5	0.506 81	77.186	77.197
$\mathcal{P}(0.8, 1)$	400.1549	0.159 43	0.1549	−0.1456	0.998 15	0.625	0.639 31	300.824	300.883
$\mathcal{P}(0.8, 2)$	401.0467	0.157 57	1.0467	0.5459	0.998 21	0.625	0.622 97	151.718	151.718
$\mathcal{P}(0.8, 5)$	402.8376	0.159 05	2.8376	2.4840	0.998 24	0.625	0.644 02	62.194	62.197
$\mathcal{P}(0.5, 1)$	399.5954	0.203 07	−0.4046	−0.2330	0.998 65	1	1.045 84	180.798	180.883
$\mathcal{P}(0.5, 2)$	400.9764	0.202 88	0.9764	0.8734	0.998 70	1	1.035 36	91.744	97.718
$\mathcal{P}(0.5, 5)$	403.9610	0.201 71	3.9610	3.9744	0.998 75	1	1.036 29	38.196	38.197

TABLE 3. Descriptive statistics for the infection risk.

n	$\bar{x}$	s	Min	Q_1	Med	Q_3	Max
113	4.355	1.341	1.300	3.700	4.400	5.200	7.800

TABLE 4. Terminated sample size associated with number of sampling operations using $\mathcal{P}(\rho, k)$ as per (14).

$\mathcal{P}(\rho, k)$	$n_{\mathcal{P}(\rho,k)}$	$\varphi_{\mathcal{P}(\rho,k)}$	$\hat{\mu}$
$\mathcal{P}(1, 1)$	70	60	4.4471
$\mathcal{P}(1, 2)$	73	32	4.4342
$\mathcal{P}(1, 5)$	76	14	4.4526
$\mathcal{P}(0.8, 1)$	72	48	4.4306
$\mathcal{P}(0.8, 2)$	72	25	4.4306
$\mathcal{P}(0.8, 5)$	77	12	4.4532
$\mathcal{P}(0.5, 1)$	54	18	4.4519
$\mathcal{P}(0.5, 2)$	54	10	4.4519
$\mathcal{P}(0.5, 5)$	62	6	4.4565

$\rho = 0.8$ or 1, for a fixed k value. The point estimates constructed from each sampling procedure are listed in the last column, and they were close to each other. Finally, we should reiterate that each row in Table 4 was obtained from one single run, but shows the practical applicability of our accelerated group sequential MRPE methodology $\mathcal{P}(\rho, k)$. We have indeed repeated similar implementations, but no obvious difference appeared. Consequently, we have left out many details for brevity.

4. Bounded variance point estimation for negative exponential location

For a fixed sample size, however large it is, the variance of an estimator can be larger than a prescribed level to an arbitrary extent. This problem was addressed in [10], where the authors focused on estimating the pure premium in actuarial science. Here, our newly proposed accelerated group sequential sampling scheme $\mathcal{M}(\rho, k)$ can be implemented to guarantee that the variance of our estimator is close to all small predetermined levels. In this section, therefore, we include another illustration: the bounded variance point estimation (BVPE) for the location parameter μ of a negative exponential distribution $\mathrm{NExp}(\mu, \sigma)$ with the probability density function

$$f(y;\mu, \sigma) = \frac{1}{\sigma} \exp\left\{-\frac{y-\mu}{\sigma}\right\} \mathbf{1}(y > \mu),$$

where both $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$ remain unknown. Having recorded a random sample $Y_1, \ldots, Y_n$, $n \geq 2$, we write $Y_{n:1} = \min\{Y_1, \ldots, Y_n\}$, which is the maximum likelihood estimator (MLE) of μ, and $V_n = (n-1)^{-1} \sum_{i=1}^{n} (Y_i - Y_{n:1})$, which is the uniformly minimum variance unbiased estimator (UMVUE) of σ. As a standard approach, we estimate μ using its MLE $Y_{n:1}$, which is a consistent estimator.

It is well known that (i) $n(Y_{n:1}-\mu)/\sigma \sim \text{NExp}(0,1)$; (ii) $2(n-1)V_n/\sigma \sim \chi^2_{2n-2}$; and (iii) $Y_{n:1}$ and $(V_2,\ldots,V_n)$, $n\geq 2$, are independent. Hence, the variance of the proposed point estimator $Y_{n:1}$ is $\mathbb{V}_{\mu,\sigma}[Y_{n:1}]=\sigma^2/n^2$. Now, our goal is to make $\mathbb{V}_{\mu,\sigma}[Y_{n:1}]$ fall below (or be close to) a predetermined level b^2, $b>0$, for all $0<\sigma<\infty$. Then, it is clear that we have $n\geq \sigma/b$. The optimal fixed sample size is therefore given by

$$n^*=\frac{\sigma}{b}. \tag{19}$$

See [18, p. 183] for more information.

Since σ is unknown to us, we estimate it by updating its UMVUE V_n at every stage as needed, and implement the following accelerated group sequential BVPE methodology $\mathcal{Q}(\rho,k)$:

$$\begin{aligned} T_{\mathcal{Q}(\rho,k)} &\equiv T_{\mathcal{Q}(\rho,k)}(c)=\inf\{n\geq 0\colon m+kn\geq \rho V_{m+kn}/b\},\\ N_{\mathcal{Q}(\rho,k)} &\equiv N_{\mathcal{Q}(\rho,k)}(c)=\lfloor \rho^{-1}(m+kT_{\mathcal{Q}(\rho,k)})\rfloor+1. \end{aligned} \tag{20}$$

Here, $0<\rho\leq 1$ is a prefixed proportion, $k\geq 1$ is a prefixed positive integer, and the pilot sample size $m=m_0k+1$ for some m_0. We further assume that the following limit operation holds:

$$m_0\to\infty,\ m=m_0k+1\to\infty,\ b\equiv b(m)=O(m^{-r}),\ n^*=O(m^r),\ \text{and}\ \limsup\frac{m}{n^*}<\rho, \tag{21}$$

where $r\geq 1$ is a fixed constant. The methodology $\mathcal{Q}(\rho,k)$ is conducted analogously to the methodology $\mathcal{P}(\rho,k)$ introduced in Section 3.

Again, it is clear that $\mathbb{P}_{\mu,\sigma}(N_{\mathcal{Q}(\rho,k)}<\infty)=1$ and $N_{\mathcal{Q}(\rho,k)}\uparrow\infty$ with probability 1 as $b\downarrow 0$. Upon termination with the fully gathered data

$$\{T_{\mathcal{Q}(\rho,k)},N_{\mathcal{Q}(\rho,k)},Y_1,\ldots,Y_m,\ldots,Y_{m+kT_{\mathcal{Q}(\rho,k)}},\ldots,Y_{N_{\mathcal{Q}(\rho,k)}}\},$$

we construct the bounded variance point estimator $Y_{N_{\mathcal{Q}(\rho,k)}:1}=\min\{Y_1,\ldots,Y_{N_{\mathcal{Q}}(\rho,k)}\}$ for μ.

Along the lines of (16), we define $T_{\mathcal{Q}(\rho,k)}=T'_{\mathcal{Q}(\rho,k)}-m_0$ with probability 1. Then $T'_{\mathcal{Q}(\rho,k)}$ is a new stopping time that can be rewritten as

$$T'_{\mathcal{Q}(\rho,k)}=\inf\left\{n\geq m_0\colon (kn)^{-1}\sum_{i=1}^{n}U_i\leq 2[kn/(\rho n^*)](1+(kn)^{-1})\right\}, \tag{22}$$

where $\delta=1$, $l_0=1$, and $U_i=\sum_{j=(i-1)k+1}^{ik}W_j$, $i=1,2,\ldots$, with $W_1,W_2,\ldots$ being i.i.d. χ^2_2 random variables such that $\theta=2$, $\tau^2=4$, and $\alpha=1$. Therefore, $U_1,U_2,\ldots$ are i.i.d. χ^2_{2k} random variables. Now we state a number of asymptotic first- and second-order properties of the accelerated group sequential BVPE methodology $\mathcal{Q}(\rho,k)$, summarized in the following theorem.

Theorem 5. *For the accelerated group sequential BVPE methodology $\mathcal{Q}(\rho,k)$ given in (20), for all fixed μ, σ, k and $0<\rho<1$, under the limit operations (21):*

TABLE 5. $\eta_2(k)$ approximations in Theorem 5(ii).

k	$\eta_2(k)$	$k^{-1}\eta_2(k)$	k	$\eta_2(k)$	$k^{-1}\eta_2(k)$
1	−0.2552	−0.2552	11	4.9940	0.4540
2	0.3433	0.1717	12	5.4957	0.4580
3	0.8976	0.2992	13	5.9970	0.4613
4	1.4308	0.3577	14	6.4978	0.4641
5	1.9523	0.3905	15	6.9984	0.4666
6	2.4667	0.4111	16	7.4988	0.4687
7	2.9765	0.4252	17	7.9992	0.4705
8	3.4834	0.4354	18	8.4994	0.4722
9	3.9883	0.4431	19	8.9996	0.4737
10	4.4916	0.4492	20	9.4997	0.4750

(i) *Asymptotic first-order efficiency:* $\mathbb{E}_{\mu,\sigma}[N_{\mathcal{Q}(\rho,k)}/n^*] \to 1$;

(ii) *Asymptotic second-order efficiency:* $\rho^{-1}\eta_2(k) + o(1) \le \mathbb{E}_{\mu,\sigma}[N_{\mathcal{Q}(\rho,k)} - n^*] \le \rho^{-1}\eta_2(k) + 1 + o(1)$, *where* $\eta_2(k) = (k-1)/2 - \frac{1}{2}\sum_{n=1}^{\infty} n^{-1}\mathbb{E}\big[\{\chi^2_{2kn} - 4kn\}^+\big]$;

(iii) *Asymptotic variance:* $\mathbb{V}_{\mu,\sigma}[Y_{N_{\mathcal{Q}(\rho,k)}:1}] = b^2 + o(b^2)$.

When $\rho = 1$, we have the exact expression $\mathbb{E}_{\mu,\sigma}[N_{\mathcal{Q}(1,k)} - n^*] = \eta_2(k) + o(1)$ instead of the inequality in Theorem 4(ii). The number of sampling operations for the accelerated group sequential BVPE methodology $\mathcal{Q}(\rho, k)$ is

$$\varphi_{\mathcal{Q}(\rho,k)} = T_{\mathcal{Q}(\rho,k)} + 1 + \mathbf{1}(\rho < 1), \tag{23}$$

and

$$\mathbb{E}_{\mu,\sigma}[\varphi_{\mathcal{Q}(\rho,k)}] = k^{-1}[\rho n^* - m + \eta_2(k)] + 1 + \mathbf{1}(\rho < 1) + o(1). \tag{24}$$

For any integer $k \ge 1$, $\eta_2(k) = (k-1)/2 - \frac{1}{2}\sum_{n=1}^{\infty} n^{-1}\mathbb{E}\big[\{\chi^2_{2kn} - 4kn\}^+\big]$ is also computable. Table 5 provides some numerical approximations in the same fashion as Table 1.

4.1. Simulated performance

In this section, we summarize selective Monte Carlo simulation results to demonstrate the appealing properties, including both first and second order, of the accelerated group sequential BVPE methodologies that we provided in (20). We investigated a wide range of scenarios in terms of the location and scale parameters of the negative exponential population (NExp), as well as the prespecified parameters: the parameter b^2 for the bounded variance, and the two parameters, ρ and k, of $\mathcal{Q}(\rho, k)$ for using different sampling schemes. For brevity, we summarize the results from pseudorandom samples of an NExp(5,2) population in Table 6. We specified $b^2 = 0.0004$, 0.0001, and 0.000 025, and the optimal fixed sample size n^* turned out to be 100, 200, and 400 accordingly. To compare the number of sampling operations under different possible scenarios, we considered the combinations of $\rho = (1.0, 0.8, 0.5)$ and $k = (1, 2, 5)$. For each sampling scheme of $\mathcal{Q}(\rho, k)$, we included the average total final sample

TABLE 6. Simulations from NExp(5,2) from 10 000 runs implementing $\mathcal{Q}(\rho, k)$ from (20).

$\mathcal{Q}(\rho, k)$	$\bar{n}$	$s(\bar{n})$	$\bar{n} - n^*$	$\rho^{-1}\eta_2(k)$	$\bar{\varphi}$	$\mathbb{E}(\varphi)$	$\mathbb{V}(Y_{N:1})$
$n^* = 100$, $b = 0.0004$							
$\mathcal{Q}(1, 1)$	99.7232	0.102 14	−0.2768	−0.2552	95.723	97.952	0.000 426
$\mathcal{Q}(1, 2)$	100.3324	0.102 52	0.3324	0.3433	46.666	47.476	0.000 421
$\mathcal{Q}(1, 5)$	101.9885	0.103 33	1.9885	1.9523	17.198	17.190	0.000 393
$\mathcal{Q}(0.8, 1)$	100.0483	0.116 79	0.0483	−0.3190	76.737	78.952	0.000 441
$\mathcal{Q}(0.8, 2)$	100.8849	0.115 55	0.8849	0.4291	37.654	38.476	0.000 412
$\mathcal{Q}(0.8, 5)$	102.8252	0.116 37	2.8252	2.4404	14.191	14.190	0.000 387
$\mathcal{Q}(0.5, 1)$	99.4418	0.149 38	−0.5582	−0.5104	46.721	48.952	0.000 478
$\mathcal{Q}(0.5, 2)$	100.7288	0.147 63	0.7288	0.6866	22.682	23.476	0.000 428
$\mathcal{Q}(0.5, 5)$	103.7720	0.149 44	3.7720	3.9046	8.177	8.190	0.000 399
$n^* = 200$, $b = 0.0001$							
$\mathcal{Q}(1, 1)$	199.8580	0.144 26	−0.1420	−0.2552	195.858	197.952	0.000 105
$\mathcal{Q}(1, 2)$	200.4932	0.142 29	0.4932	0.3433	96.747	97.476	0.000 101
$\mathcal{Q}(1, 5)$	202.1355	0.143 24	2.1355	1.9523	37.227	37.190	0.000 096
$\mathcal{Q}(0.8, 1)$	200.1291	0.159 78	0.1291	−0.3190	156.805	158.952	0.000 100
$\mathcal{Q}(0.8, 2)$	201.0824	0.158 91	1.0824	0.4291	77.733	78.476	0.000 097
$\mathcal{Q}(0.8, 5)$	202.9941	0.159 44	2.9941	2.4404	30.219	30.190	0.000 095
$\mathcal{Q}(0.5, 1)$	199.2108	0.205 46	−0.7892	−0.5104	96.605	98.952	0.000 103
$\mathcal{Q}(0.5, 2)$	200.7224	0.206 82	0.7224	0.6866	47.681	48.476	0.000 100
$\mathcal{Q}(0.5, 5)$	204.0600	0.204 89	4.0600	3.9046	18.206	18.190	0.000 098
$n^* = 400$, $b = 0.000\,025$							
$\mathcal{Q}(1, 1)$	399.8497	0.200 26	−0.1503	−0.2552	395.850	397.952	0.000 025
$\mathcal{Q}(1, 2)$	400.4404	0.199 41	0.4404	0.3433	196.720	197.476	0.000 025
$\mathcal{Q}(1, 5)$	402.0665	0.200 82	2.0665	1.9523	77.213	77.190	0.000 025
$\mathcal{Q}(0.8, 1)$	400.2258	0.225 04	0.2258	−0.3190	316.882	318.952	0.000 025
$\mathcal{Q}(0.8, 2)$	401.0488	0.224 46	1.0488	0.4291	157.718	158.476	0.000 024
$\mathcal{Q}(0.8, 5)$	403.0076	0.225 92	3.0076	2.4404	62.221	62.190	0.000 025
$\mathcal{Q}(0.5, 1)$	399.5034	0.284 95	−0.4966	−0.5104	196.752	198.952	0.000 026
$\mathcal{Q}(0.5, 2)$	400.9348	0.286 33	0.9348	0.6866	97.734	98.476	0.000 026
$\mathcal{Q}(0.5, 5)$	403.9800	0.290 34	3.9800	3.9046	38.198	38.190	0.000 025

size $\bar{n}$ with the associated standard error $s(\bar{n})$, the difference between $\bar{n}$ and n^* to be compared with the second-order efficiency term, $\rho^{-1}\eta_2(k)$, in Theorem 4(ii), $\mathbb{V}(Y_{N:1})$, which should be close to the asymptotic variance as listed in Theorem 4(iii), and the average number of sampling operations $\bar{\varphi}$ to be compared with the expected number of sampling operations $\mathbb{E}(\varphi)$ from (24).

From Table 6, it is obvious that $\bar{n} - n^*$ hangs tightly around each of the second-order approximations $\rho^{-1}\eta_2(k)$. From the sixth and seventh columns, we can also easily see that the average number of sampling operations needed, $\bar{\varphi}$, is very close to its theoretical value $\mathbb{E}(\varphi)$. Moreover, the sampling operations for $\mathcal{Q}(\rho, k)$ are significantly reduced when $0 < \rho < 1$ and/or $k > 1$, compared to the Anscombe–Chow–Robbins purely sequential procedure $\mathcal{Q}(1, 1)$ under the same b. And the reductions are approximately $100(1 - k^{-1}\rho)\%$. The last column of

TABLE 7. Terminated sample size associated with the number of sampling operations using $\mathcal{Q}(\rho, k)$ as per (20).

$\mathcal{Q}(\rho, k)$	$n_{\mathcal{Q}(\rho,k)}$	$\varphi_{\mathcal{Q}(\rho,k)}$	$\hat{\mu}$
$\mathcal{Q}(1, 1)$	28	18	6.53
$\mathcal{Q}(1, 2)$	29	10	6.53
$\mathcal{Q}(1, 5)$	31	5	6.53
$\mathcal{Q}(0.8, 1)$	30	15	6.53
$\mathcal{Q}(0.8, 2)$	32	9	6.53
$\mathcal{Q}(0.8, 5)$	33	5	6.53
$\mathcal{Q}(0.5, 1)$	34	8	6.53
$\mathcal{Q}(0.5, 2)$	34	5	6.53
$\mathcal{Q}(0.5, 5)$	42	4	6.53

Table 6 shows that the variance of the smallest observations is approximately b^2 across the board.

4.2. Real data analysis

In this section, we implement the accelerated group sequential BVPE methodology $\mathcal{Q}(\rho, k)$ as per (20) on a real dataset about survival times of a group of patients suffering from head and neck cancer who were treated using a combination of radiotherapy and chemotherapy; the dataset has been presented in multiple research articles [4, 30, 35].

It is fair to assume that the survival time data follows a negative exponential distribution as claimed in [35]; a Kolmogorov–Smirnov test yielded a test statistic value of 0.156 86 with a p-value of 0.5572. Assuming that researchers in this study want to use the smallest observation of sample data to estimate the location parameter μ, and they also want to restrict the variance of the estimator to be $b^2 = 100$, we implement the sampling scheme $\mathcal{Q}(\rho, k)$ in this investigation with a combination of $\rho = (1, 0.8, 0.5)$ and $k = (1, 2, 5)$.

For each sampling scheme $\mathcal{Q}_{\rho,k}$, the sampling procedure is the following: we randomize all of the observations but pretend these observations are not known to us. We start with $m = 11$ observations, and proceed with the sampling following the procedures as per (20). We also assume the data comes in the order after randomization. We summarize the terminated sample size and the corresponding number of sampling operations in Table 7. The columns are defined similarly to Table 4.

We can see from Table 7 that the terminated sample size ranges from 28 to 42, with the least number of observations when using the sampling scheme $\mathcal{Q}(\rho = 1, k = 1)$ and the most number of observations when using $\mathcal{Q}(\rho = 0.5, k = 5)$. Moreover, the sampling operations are reduced the most when using $\mathcal{Q}(\rho = 0.5, k = 5)$. Also, with the same ρ, larger k means fewer sampling operations; with the same k, smaller ρ means fewer sampling operations. The last column records the point estimates, which were the minimum survival times observed in each sampling procedure. We should emphasize that all of these results were obtained from one single run, but we have indeed repeated similar implementations and there was little to no difference. We also want to emphasize that the real data example is only for illustration purposes. The example shows how our developed methodology can be used in real research problems.

5. Proofs

Note that Theorems 1 and 2 follow immediately from [34, Theorem 2.4], (3) follows from Theorem 1 and the definition of $\varphi_{\mathcal{M}_0}$, (7) follows from Theorem 2, and Theorem 3 is paraphrased from [18, (6.4.14)]. In this section, therefore, we only prove Theorems 4 and 5.

5.1. Proof of Theorem 4 and (15)

By the stopping rule defined in (14), we have the following inequality:

$$N_{\mathcal{P}(\rho,k)} \le \rho^{-1}(m + kT_{\mathcal{P}(\rho,k)}) + 1, \tag{25}$$

where $m + kT_{\mathcal{P}(\rho,k)} < m + k + \rho S_{m+k(T_{\mathcal{P}(\rho,k)}-1)}\sqrt{A/c}$. Therefore,

$$\frac{N_{\mathcal{P}(\rho,k)}}{n^*} \le \frac{S_{m+k(T_{\mathcal{P}(\rho,k)}-1)}}{\sigma} + \frac{\rho^{-1}(m+k)+1}{n^*} \xrightarrow{P_{\mu,\sigma}} 1 \text{ as } c \to 0. \tag{26}$$

On the other hand, we also have the inequality that

$$N_{\mathcal{P}(\rho,k)} \ge \rho^{-1}(m + kT_{\mathcal{P}(\rho,k)}) \ge S_{m+kT_{\mathcal{P}(\rho,k)}}\sqrt{A/c}, \tag{27}$$

from which we conclude that

$$\frac{N_{\mathcal{P}(\rho,k)}}{n^*} \ge \frac{S_{m+kT_{\mathcal{P}(\rho,k)}}}{\sigma} \xrightarrow{P_{\mu,\sigma}} 1 \text{ as } c \to 0.$$

Combining this with (26), it is clear that

$$N_{\mathcal{P}(\rho,k)}/n^* \xrightarrow{P_{\mu,\sigma}} 1 \text{ as } c \to 0. \tag{28}$$

Note that, for sufficiently small c, with the limit operation given in (15), we have (with probability 1)

$$\frac{N_{\mathcal{P}(\rho,k)}}{n^*} \le \sigma^{-1} \sup_{n\ge 2} S_n + 2.$$

Since $(\mathbb{E}_{\mu,\sigma}[\sup_{n\ge 2} S_n])^2 \le \mathbb{E}_{\mu,\sigma}[(\sup_{n\ge 2} S_n)^2] \le \mathbb{E}_{\mu,\sigma}[\sup_{n\ge 2} S_n^2]$, and Wiener's ergodic theorem [33, Theorem IV] leads to $\mathbb{E}_{\mu,\sigma}[\sup_{n\ge 2} S_n^2] < \infty$, combined with (28) it follows by the dominated convergence theorem that $\mathbb{E}_{\mu,\sigma}[N_{\mathcal{P}(\rho,k)}/n^*] \to 1$ as $c \to 0$. Since $c = A\sigma^2/n^{*2}$ from (9), Theorem 4(i) holds under the limit operations (15).

Recall that $T'_{\mathcal{P}(\rho,k)}$ defined in (16) is of the same form as t_0 from (1). Then, referring to [34, (1.1)] or [18, Section A.4], we have, as $c \to 0$, for $m_0 \ge 2$,

$$\mathbb{E}_{\mu,\sigma}\left[T'_{\mathcal{P}(\rho,k)}\right] = k^{-1}\rho n^* + \frac{1}{2} - \frac{3}{2k} - \frac{1}{2k}\sum_{n=1}^{\infty} n^{-1}\mathbb{E}\left[\left\{\chi^2_{kn} - 3kn\right\}^+\right] + o(1).$$

Therefore, with $T'_{\mathcal{P}(\rho,k)} = T_{\mathcal{P}(\rho,k)} + m$ with probability 1 and $m = m_0 k + 1$, we have

$$\mathbb{E}_{\mu,\sigma}[m + kT_{\mathcal{P}(\rho,k)}] = \mathbb{E}_{\mu,\sigma}\left[1 + kT'_{\mathcal{P}(\rho,k)}\right] = \rho n^* + \eta_1(k) + o(1). \tag{29}$$

Putting together (17) and (29), we obtain (18). And under the limit operations (15), Theorem 4(ii) follows immediately from (29) and inequalities given in (25) and (27).

Next, we state the following lemmas to derive the desirable results in Theorem 4(iii) and (iv).

Lemma 1. *For the accelerated group sequential MRPE methodology $\mathcal{P}(\rho, k)$ given in (14), under the limit operations (15), for any arbitrary $0 < \varepsilon < 1$, with some $\gamma \geq 2$,*

$$\mathbb{P}_{\mu,\sigma}(N_{\mathcal{P}(\rho,k)} \leq \varepsilon n^*) = O\big(n^{*^{-\gamma/2r}}\big).$$

Proof. Recall that $\lfloor u \rfloor$ denotes the largest integer that is smaller than u, and define

$$t_u = \lfloor k^{-1}\rho\varepsilon n^* \rfloor + 1.$$

It should be obvious that $0 \leq T_{\mathcal{P}(\rho,k)} \leq t_u$. Then, the rate at which $\mathbb{P}_{\mu,\sigma}\{N_{\mathcal{P}(\rho,k)} \leq \varepsilon n^*\}$ may converge to zero under the limit operations (15) is given by

$$\begin{aligned}
\mathbb{P}_{\mu,\sigma}\{N_{\mathcal{P}(\rho,k)} \leq \varepsilon n^*\} &\leq \mathbb{P}_{\mu,\sigma}\{\rho^{-1}(m + kT_{\mathcal{P}(\rho,k)}) \leq \varepsilon n^*\} \\
&\leq \mathbb{P}_{\mu,\sigma}\{S_{m+kt} \leq \varepsilon\sigma \text{ for some } t \text{ such that } 0 \leq t \leq t_u\} \\
&\leq \mathbb{P}_{\mu,\sigma}\big\{ \max_{0\leq t\leq t_u} |S_{m+kt} - \sigma| \geq (1-\varepsilon)\sigma \big\} \\
&\leq \{(1-\varepsilon)\sigma\}^{-\gamma}\mathbb{E}_{\mu,\sigma}|S_m - \sigma|^{\gamma} = O(m^{-\gamma/2}) = O\big(n^{*^{-\gamma/2r}}\big),
\end{aligned}$$

where the last inequality comes from Kolmogorov's inequality for reversed martingales. See [11, Section 5.1] for more details.

Lemma 2. *For the accelerated group sequential MRPE methodology $\mathcal{P}(\rho, k)$ given in (14), under the limit operations (15),*

$$\text{(i)} \quad \frac{N_{\mathcal{P}(\rho,k)} - n^*}{n^{*^{1/2}}} \xrightarrow{\mathrm{d}} N\big(0, \tfrac{1}{2}\rho^{-1}\big); \qquad \text{(ii)} \quad \frac{N_{\mathcal{P}(\rho,k)} - n^*}{N_{\mathcal{P}(\rho,k)}^{1/2}} \xrightarrow{\mathrm{d}} N\big(0, \tfrac{1}{2}\rho^{-1}\big).$$

Proof. First, we prove that

$$\frac{m + kT_{\mathcal{P}(\rho,k)} - \rho n^*}{\sqrt{\rho n^*}} \xrightarrow{\mathrm{d}} N\big(0, \tfrac{1}{2}\big) \text{ as } c \to 0 \tag{30}$$

based on the inequalities

$$\frac{\rho S_{m+kT_{\mathcal{P}(\rho,k)}}\sqrt{A/c} - \rho n^*}{\sqrt{\rho n^*}} \leq \frac{m + kT_{\mathcal{P}(\rho,k)} - \rho n^*}{\sqrt{\rho n^*}} \leq \frac{\rho S_{m+k(T_{\mathcal{P}(\rho,k)}-1)}\sqrt{A/c} - \rho n^* + m + k}{\sqrt{\rho n^*}}.$$

It is not hard to see that $(\lfloor \rho n^* \rfloor + 1)^{1/2}\big(S_{\lfloor \rho n^* \rfloor + 1}/\sigma - 1\big) \xrightarrow{\mathrm{d}} N\big(0, \tfrac{1}{2}\big)$ as $c \to 0$, and the sequence $\{S_{\lfloor \rho n^* \rfloor + 1}\}$ is *uniformly continuous in probability* [1, 2]. From previous results, we can easily show that

$$\frac{m + kT_{\mathcal{P}(\rho,k)}}{\lfloor \rho n^* \rfloor + 1} \xrightarrow{P_{\mu,\sigma}} 1 \text{ as } c \to 0.$$

Now, Anscombe's random central limit theorem [1] leads to

$$\frac{\rho S_{m+kT_{\mathcal{P}(\rho,k)}}\sqrt{A/c} - \rho n^*}{\sqrt{\rho n^*}} \xrightarrow{\mathrm{d}} N\big(0, \tfrac{1}{2}\big) \text{ and } \frac{\rho S_{m+k(T_{\mathcal{P}(\rho,k)}-1)}\sqrt{A/c} - \rho n^* + m + k}{\sqrt{\rho n^*}} \xrightarrow{\mathrm{d}} N\big(0, \tfrac{1}{2}\big)$$

as $c \to 0$. Hence, (30) holds. Next, with the inequalities that

$$\rho^{-1/2}\frac{m+kT_{\mathcal{P}(\rho,k)}-\rho n^*}{\sqrt{\rho n^*}} \le \frac{N_{\mathcal{P}(\rho,k)}-n^*}{\sqrt{n^*}} \le \rho^{-1/2}\frac{m+kT_{\mathcal{P}(\rho,k)}-\rho n^*}{\sqrt{\rho n^*}} + \frac{1}{\sqrt{n^*}}, \tag{31}$$

Lemma 2(i) follows immediately, and Slutsky's theorem provides Lemma 2(ii) under the limit operations (15).

Lemma 3. *For the accelerated group sequential MRPE methodology $\mathcal{P}(\rho, k)$ given in (14), under the limit operations (15), $\left(N_{\mathcal{P}(\rho,k)} - n^*\right)^2/n^*$ is uniformly integrable.*

Proof. In the light of [11, Theorem 3.4], we can prove that, for sufficiently small $c \le c_0$ by choosing some c_0 (> 0) appropriately, $(\rho n^*)^{-1}\left(m + kT_{\mathcal{P}(\rho,k)} - \rho n^*\right)^2$ is uniformly integrable. Therefore, under the limit operations (15), we have Lemma 3 by applying the inequalities given in (31).

Now, Theorem 4(iii) and (iv) follow from Lemmas 1–3. Alternatively, appealing to nonlinear renewal theory, we can also prove the same results in the spirit of [34]. Many details are left out for brevity.

5.2. Proof of Theorem 5 and (23)

In the same fashion as we proved Theorem 4(i), we have

$$\frac{V_{m+kT}}{\sigma} \le \frac{N_{\mathcal{Q}(\rho,k)}}{n^*} \le \frac{V_{m+k(T_{\mathcal{Q}(\rho,k)}-1)}}{\sigma} + \frac{\rho^{-1}(m+k)+1}{n^*}.$$

As $b \to 0$, the two bounds of these inequalities both tend to 1 in probability, so

$$\frac{N_{\mathcal{Q}(\rho,k)}}{n^*} \xrightarrow{P_{\mu,\sigma}} 1.$$

For sufficiently small b, with the limit operation (21), we have (with probability 1)

$$\frac{N_{\mathcal{Q}(\rho,k)}}{n^*} \le \sigma^{-1} \sup_{n\ge 2} V_n + 2,$$

where $2(n-1)V_n/\sigma \sim \chi^2_{2n-2}$. Similarly, $\mathbb{E}_{\mu,\sigma}[\sup_{n\ge 2} V_n] < \infty$ follows from Wiener's ergodic theorem [33, Theorem IV] so that $\mathbb{E}_{\mu,\sigma}[N_{\mathcal{Q}(\rho,k)}/n^*] \to 1$ as $b \to 0$. The proof of Theorem 4(i) is complete.

Then, recall that $T'_{\mathcal{Q}(\rho,k)}$ defined in (22) is of the same form as t_0 from (1). So, referring to [34, (1.1)] or [18, Section A.4], we have, as $b \to 0$, for $m_0 \ge 2$,

$$\mathbb{E}_{\mu,\sigma}\left[T'_{\mathcal{Q}(\rho,k)}\right] = k^{-1}\rho n^* + \frac{1}{2} - \frac{3}{2k} - \frac{1}{2k}\sum_{n=1}^{\infty} n^{-1}\mathbb{E}\left[\left\{\chi^2_{2kn} - 4kn\right\}^+\right] + o(1).$$

Therefore, with $T'_{\mathcal{Q}(\rho,k)} = T_{\mathcal{Q}(\rho,k)} + m$ with probability 1 and $m = m_0 k + 1$, we have

$$\mathbb{E}_{\mu,\sigma}[m + kT_{\mathcal{Q}(\rho,k)}] = \mathbb{E}_{\mu,\sigma}\left[1 + kT'_{\mathcal{Q}(\rho,k)}\right] = \rho n^* + \eta_2(k) + o(1). \tag{32}$$

Putting together (23) and (32), we obtain (24). And under the limit operations (21), Theorem 4(ii) follows immediately from (32).

To evaluate the asymptotic variance in Theorem 4(iii), we utilize the law of total variance and obtain

$$\begin{aligned}\mathbb{V}_{\mu,\sigma}\left[Y_{N_{\mathcal{Q}(\rho,k)}:1}\right] &= \mathbb{E}_{\mu,\sigma}\left[\mathbb{V}\left(Y_{N_{\mathcal{Q}(\rho,k)}:1} \mid N_{\mathcal{Q}(\rho,k)}\right)\right] + \mathbb{V}_{\mu,\sigma}\left[\mathbb{E}\left(Y_{N_{\mathcal{Q}(\rho,k)}:1} \mid N_{\mathcal{Q}(\rho,k)}\right)\right] \\ &= \sum_{n=m}^{\infty} \mathbb{V}\left(Y_{N_{\mathcal{Q}(\rho,k)}:1} \mid N_{\mathcal{Q}(\rho,k)} = n\right)\mathbb{P}_{\mu,\sigma}(N_{\mathcal{Q}(\rho,k)} = n) + \mathbb{V}_{\mu,\sigma}\left[\frac{\sigma^2}{N_{\mathcal{Q}(\rho,k)}} + \mu\right] \\ &= \sum_{n=m}^{\infty} \frac{\sigma^2}{n^2}\mathbb{P}_{\mu,\sigma}(N_{\mathcal{Q}(\rho,k)} = n) + \mathbb{E}_{\mu,\sigma}\left[\frac{\sigma^2}{N^2_{\mathcal{Q}(\rho,k)}}\right] - \mathbb{E}^2_{\mu,\sigma}\left[\frac{\sigma}{N_{\mathcal{Q}(\rho,k)}}\right] \\ &= 2\mathbb{E}_{\mu,\sigma}\left[\frac{\sigma^2}{N^2_{\mathcal{Q}(\rho,k)}}\right] - \mathbb{E}^2_{\mu,\sigma}\left[\frac{\sigma}{N_{\mathcal{Q}(\rho,k)}}\right], \end{aligned} \tag{33}$$

since the event $\{N_{\mathcal{Q}(\rho,k)} = n\}$ depends on V_n alone and is therefore independent of $Y_{n:1}$.

Applying Taylor's theorem to expand $N^{-j}_{\mathcal{Q}(\rho,k)}$, $j \geq 1$, around n^*, we have

$$N^{-j}_{\mathcal{Q}(\rho,k)} = n^{*-j} - j\lambda^{-j-1}(N_{\mathcal{Q}(\rho,k)} - n^*), \tag{34}$$

where λ is a random variable lying between $N_{\mathcal{Q}(\rho,k)}$ and n^*. Combining (32), (34), (19), and Theorem 4(ii) yields $\mathbb{V}_{\mu,\sigma}[Y_{N_{\mathcal{Q}(\rho,k)}:1}] = b^2 + O(b^3) = b^2 + o(b^2)$, which completes the proof.

6. Concluding remarks

We have proposed a novel accelerated group sequential sampling scheme with the motivation of saving sampling operations while retaining efficiency. Following the idea of drawing multiple observations at a time sequentially to determine a preliminary sample, and then gathering the rest of the observations all in one batch, we demonstrated the MRPE and BVPE problems under the new sampling scheme as possible illustrations. Furthermore, the new sequential sampling scheme can be applied to deal with other statistical inference problems, including, but not limited to, sequential analogues of Behrens–Fisher problems (see, e.g., [28]), fixed-width confidence intervals (see, e.g., [7]), ranking and selection (see, e.g., [23]), bounded-risk point estimation (see, e.g., [17]), treatment means comparison (see, e.g., [20]), etc.

Due to the appealing properties of our newly developed methodology and the reality of substantial sampling operation savings, it will be of great interest for further investigations into the problems that researchers have recently been working on. A full list would keep going for a while, so we just list a couple here to demonstrate the possible directions: (i) [29] proposed a *sequentially planned probability ratio test* as a sequentially planned extension of the famous Wald *sequential probability ratio test*, and (ii) [26] worked on two sample mean comparisons of normal distributions with unknown and unequal variances, where they developed both purely sequential and two-stage methodologies. Our sequential sampling design introduced here can be directly applied to their problem settings, and is expected to save sampling operations significantly.

On another note, recall from Section 3.1 that our loss function for estimating a normal mean has only included the cost per sampled unit. In certain situations, we may have to consider both the cost per sampled group and the cost per sampled unit. Then, a different type of (accelerated) group sequential sampling scheme will be desired. We think it would be a very

interesting future research problem to explore, perhaps developing two stopping variables for statistical inference with a minimal cost: one for determining the size of a group, and the other for determining the number of groups needed.

Acknowledgement

The authors thank the editor and the three anonymous reviewers for their valuable comments that helped improve this paper.

Funding information

There are no funding bodies to thank relating to the creation of this article.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] ANSCOMBE, F. J. (1952). Large-sample theory of sequential estimation. *Proc. Camb. Phil. Soc*. 48, 600–607.
[2] ANSCOMBE, F. J. (1953). Sequential estimation. *J. R. Statist. Soc. B* **15**, 1–29.
[3] 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.
[4] EFRON, B. (1988). Logistic regression, survival analysis, and the Kaplan–Meier curve. *J. Amer. Statist. Assoc.* **83**, 414–425.
[5] GHOSH, M. AND MUKHOPADHYAY, N. (1980). Sequential point estimation of the difference of two normal means. *Ann. Statist.* **8**, 221–225.
[6] HALL, P. (1981). Asymptotic theory of triple sampling for sequential estimation of a mean. *Ann. Statist.* **9**, 1229–1238.
[7] HALL, P. (1983). Sequential estimation saving sampling operations. *J. R. Statist. Soc. B* **45**, 219–223.
[8] HAYRE, L. S. (1985). Group sequential sampling with variable group sizes. *J. R. Statist. Soc. B* **47**, 90–97.
[9] HU, J. (2022). A double-sequential sampling scheme. *Commun. Statist. Theory Meth.* **51**, 6319–6333.
[10] HU, J. AND HONG, L. (2022). A nonparametric sequential learning procedure for estimating the pure premium. *Europ. Actuarial J.* **12**, 485–502.
[11] HU, J. AND MUKHOPADHYAY, N. (2019). Second-order asymptotics in a class of purely sequential minimum risk point estimation (MRPE) methodologies. *Japanese J. Statist. Data Sci.* **2**, 81–104.
[12] KUTNER, M. H., NACHTSHEIM, C., NETER, J. AND LI, W. (2005). *Applied Linear Statistical Models*, Vol. **5**. McGraw-Hill Irwin, Boston, MA.
[13] LAI, T. L. AND SIEGMUND, D. (1977). A nonlinear renewal theory with applications to sequential analysis I. *Ann. Statist.* **5**, 946–954.
[14] LAI, T. L. AND SIEGMUND, D. (1979). A nonlinear renewal theory with applications to sequential analysis II. *Ann. Statist.* **7**, 60–76.
[15] LIU, W. (1997). Improving the fully sequential sampling scheme of Anscombe–Chow–Robbins. *Ann. Statist.* **25**, 2164–2171.
[16] MUKHOPADHYAY, N. (1996). An alternative formulation of accelerated sequential procedures with applications to parametric and nonparametric estimation. *Sequent. Anal.* **15**, 253–269.
[17] MUKHOPADHYAY, N. AND BAPAT, S. R. (2018). Purely sequential bounded-risk point estimation of the negative binomial mean under various loss functions: One-sample problem. *Ann. Inst. Statist. Math.* **70**, 1049–1075.
[18] MUKHOPADHYAY, N. AND DE SILVA, B. M. (2009). *Sequential Methods and their Applications*. CRC, Boca Raton, FL.
[19] MUKHOPADHYAY, N. AND HU, J. (2017). Confidence intervals and point estimators for a normal mean under purely sequential strategies involving Gini's mean difference and mean absolute deviation. *Sequent. Anal.* **36**, 210–239.

[20] MUKHOPADHYAY, N., HU, J. AND WANG, Z. (2022). Second-order asymptotics for comparing treatment means from purely sequential estimation strategies under possible outlying observations. *Commun. Statist. Simul. Comput.* **53**, 1330–1355.
[21] MUKHOPADHYAY, N. AND SENGUPTA, P. P. (2021). *Gini Inequality Index: Methods and Applications*. CRC, Boca Raton, FL.
[22] MUKHOPADHYAY, N. AND SOLANKY, T. K. S. (1991). Second order properties of accelerated stopping times with applications in sequential estimation. *Sequent. Anal.* **10**, 99–123.
[23] MUKHOPADHYAY, N. AND SOLANKY, T. K. S. (1994). *Multistage Selection and Ranking Procedures: Second Order Asymptotics* (Statist.: Ser. Textbooks Monographs 142). Marcel Dekker, New York.
[24] MUKHOPADHYAY, N. AND WANG, Z. (2020). Purely sequential FWCI and MRPE problems for the mean of a normal population by sampling in groups with illustrations using breast cancer data. *Sequent. Anal.* **39**, 176–213.
[25] MUKHOPADHYAY, N. AND WANG, Z. (2020). Purely sequential estimation problems for the mean of a normal population by sampling in groups under permutations within each group and illustrations. *Sequent. Anal.* **39**, 484–519.
[26] MUKHOPADHYAY, N. AND ZHUANG, Y. (2019). Two-sample two-stage and purely sequential methodologies for tests of hypotheses with applications: Comparing normal means when the two variances are unknown and unequal. *Sequent. Anal.* **38**, 70–115.
[27] ROBBINS, H. (1959). Sequential estimation of the mean of a normal population. In *Probability and Statistics (Harold Cramér volume)*, ed. ULF GRENANDER. Almquist & Wiksell, Uppsala, pp. 235–245.
[28] ROBBINS, H., SIMONS, G. AND STARR, N. (1967). A sequential analogue of the Behrens–Fisher problem. *Ann. Math. Statist.* **38**, 1384–1391.
[29] SCHMEGNER, C. AND BARON, M. I. (2004). Principles of optimal sequential planning. *Sequent. Anal.* **23**, 11–32.
[30] SHANKER, R., FESSHAYE, H. AND SELVARAJ, S. (2016). On modeling of lifetime data using one parameter Akash, Lindley and exponential distributions. *Biom. Biostat. Int. J.* **2**, 1–10.
[31] STARR, N. (1966). On the asymptotic efficiency of a sequential procedure for estimating the mean. *Ann. Math. Statist.* **37**, 1173–1185.
[32] STARR, N. AND WOODROOFE, M. (1969). Remarks on sequential point estimation. In *Proc. Nat. Acad. Sci.* 63, 285–288.
[33] WIENER, N. (1939). The ergodic theorem. *Duke Math. J.* **5**, 1–18.
[34] WOODROOFE, M. (1977). Second order approximations for sequential point and interval estimation. *Ann. Statist.* **5**, 984–995.
[35] ZHUANG, Y. AND BAPAT, R. S. (2022). On comparing locations of two-parameter exponential distributions using sequential sampling with applications in cancer research. *Commun. Statist. Simul. Comput.* **51**, 6114–6135.