Ann. Inst. Statist. Math.
Vol. 46, No. 3, 509–519 (1994)

IMPROVED SEQUENTIAL ESTIMATION OF MEANS OF EXPONENTIAL DISTRIBUTIONS

N. MUKHOPADHYAY

Department of Statistics, U Box 120, University of Connecticut, 196 Auditorium Road, Storrs, CT 06269, U.S.A.

(Received March 8, 1993; revised September 20, 1993)

Abstract. Both one-sample and multi-sample estimation problems for the means of one parameter exponential distributions are addressed. In the one-sample case, for the existing purely sequential and recently obtained piecewise sequential estimation methodologies, we follow and extend the development in Isogai and Uno (1993, *Ann. Inst. Statist. Math.* (in press)) in order to obtain a class of estimators that provides asymptotic second-order risk improvement. In the multi-sample problem, we address the analogous aspects for the existing purely sequential methodology as well as the newly developed piecewise methodology.

Key words and phrases: Risk improvement, second-order asymptotics, exponential distributions, sequential methods, piecewise sequential methods, one-sample problems, multi-sample problems.

1. Introduction

Let there be a sequence $X_1, X_2, \ldots$ of independent and identically distributed (i.i.d.) random variables from a population having the probability density function (p.d.f.) given by

$$f(x;\lambda) = \lambda^{-1}\exp(-x/\lambda), \tag{1.1}$$

for $x \in R^+$ and $\lambda \in R^+$. The unknown parameter λ represents the mean of the population and we first address *minimum risk point estimation* problems for λ via sequential sampling.

Starr and Woodroofe (1972) proposed the following formulation of the problem. Having recorded $X_1, \ldots, X_n$, one estimates λ by means of $\bar{X}_n$ $(= n^{-1} \cdot \sum_{i=1}^{n} X_i)$ and assumes that the loss function is given by

$$L_n = A(\bar{X}_n - \lambda)^2 + cn, \tag{1.2}$$

where A (> 0) and c (> 0) are known numbers, respectively reflecting a "weight" constant and the "cost" per observation. The risk associated with (1.2) is given by $R_n(c) = E(L_n) = A\lambda^2 n^{-1} + cn$, which is minimized when $n = n^* = (A/c)^{1/2}\lambda$. Since n^* is unknown, Starr and Woodroofe (1972) proposed a purely sequential estimator and studied the *asymptotic* (as $c \to 0$) *behavior* of the *regret* function. Later, Woodroofe (1977) obtained the asymptotic second-order expansion of the associated regret function. Recently, Isogai and Uno (1993) provided an *improvement over* the *sequential estimators* of the percentiles of a two-parameter negative exponential distribution, first proposed in Ghosh and Mukhopadhyay (1989). As a corollary, Isogai and Uno (1993) also provides an improved sequential estimator of the corresponding population mean as compared with the original version suggested in Mukhopadhyay (1987).

For the population distribution given by (1.1), we first obtain a *class of sequential estimators* of λ which *improve* upon the Starr and Woodroofe (1972) estimator up to the *second-order approximation* (that is, up to $o(c)$) of the regret function. This is included in Subsection 2.1. In Subsection 2.2, we pursue similar aspects in the multi-sample set up as developed in Mukhopadhyay and Chattopadhyay (1991).

Along the lines of Mukhopadhyay and Sen (1993), recently Bose and Mukhopadhyay (1993) have introduced a piecewise sequential version of the Starr and Woodroofe (1972) estimator of λ which is asymptotically equivalent to the original purely sequential estimator in terms of the regret expansion up to the second-order term. But, in terms of operational convenience, the piecewise estimator is indeed significantly superior. In Subsection 3.1, we first incorporate the idea of risk improvement for the piecewise sequential estimator of λ, considered in Bose and Mukhopadhyay (1993). Then, in Subsection 3.2, we introduce a new piecewise solution of the multi-sample problem of Mukhopadhyay and Chattopadhyay (1991), followed by related analogous aspects on risk improvement.

2. Sequential estimation of means

First we consider the one-sample problem followed by the multi-sample methodology. In either situation, we put forth a class of estimators that improve upon the existing sequential estimators.

2.1 *The one-sample problem*

Under the loss function (1.2), recall that $n^* = (A/c)^{1/2}\lambda$. Let m (≥ 1) be the starting sample size and define

$$N = \inf\{n \geq m(\geq 1) : n \geq (A/c)^{1/2}\bar{X}_n\}. \tag{2.1}$$

Starr and Woodroofe (1972) estimated λ by means of $\bar{X}_N$. Woodroofe (1977) showed the $E(L_N) - R_{n^*}(c) = 3c + o(c)$ if $m \geq 3$. Now, suppose we estimate λ by $\hat{\lambda}_a(N)$ defined by

$$\hat{\lambda}_a(N) = \bar{X}_N + a(c/A)^{1/2}, \tag{2.2}$$

for $a \in R$ such that $a^2 < 2a$. For the stopping variable (2.1), let the overshoot be $\mathcal{R}_c = N^2 n^{*-1} - S_N$ where $S_N = \sum_{i=1}^N Y_i$, the Y's being i.i.d. standard exponentials. Now,

$$E[\bar{X}_N - \lambda] = \lambda\{E[(N - n^*)n^{*-1}] - E(\mathcal{R}_c N^{-1})\} \\ = -\lambda n^{*-1} + o(c^{1/2}), \tag{2.3}$$

if $m \geq 2$, since (i) $E(N - n^*) = (\nu - 1) + o(1)$, (ii) $E(\mathcal{R}_c) = \nu + o(1)$, and (iii) $n^* N^{-1}$ is uniformly integrable, where "ν" comes from equation (2.4) in Woodroofe (1977). Now, the risk associated with $\hat{\lambda}_a(N)$ is given by (for $m \geq 3$):

$$\begin{aligned} &E\{A(\hat{\lambda}_a(N) - \lambda)^2 + cN\} \\ &= E\{A(\bar{X}_N - \lambda)^2 + cN\} + 2a(Ac)^{1/2}E[\bar{X}_N - \lambda] + a^2 c \\ &= R_{n^*}(c) + 3c + ca^2 - 2ac + o(c). \end{aligned} \tag{2.4}$$

That is, $\hat{\lambda}_a(N)$ will *asymptotically dominate* $\bar{X}_N$ up to the $o(c)$ term if $2a > a^2$. The asymptotic "gain" in risk is obviously maximized if $a = 1$. In otherwords, (2.2) provides a class of dominating estimators while the best one turns out to be $\hat{\lambda}_1(N)$ in the class $\{\hat{\lambda}_a(N) : a \in R, 2a > a^2\}$.

We remark that the result given in equation (2.4) also follows from Theorem 2.1 of Isogai and Uno (1992) obtained independently in the context of estimation of the mean of a gamma distribution.

2.2 *The multi-sample problem*

Along the lines of Mukhopadhyay and Chattopadhyay (1991), let us consider k (≥ 2) independent populations and suppose that we have $X_{i1}, X_{i2}, \ldots$ i.i.d. from the i-th population having the p.d.f. $f(x; \lambda_i)$ given in (1.1), $i = 1, \ldots, k$. Having recorded $X_{i1}, \ldots, X_{in_i}$ from the i-th population, we estimate λ_i by means of $\bar{X}_{in_i} = n_i^{-1} \sum_{j=1}^{n_i} X_{ij}$. The goal is to estimate $\theta = \sum_{i=1}^k b_i \lambda_i$ where $b_1, \ldots, b_k$ are known, but arbitrary and fixed, non-zero real numbers. The parameter θ is then estimated by $T(\underset{\sim}{n}) = \sum_{i=1}^k b_i \bar{X}_{in_i}$ where $\underset{\sim}{n} = (n_1, \ldots, n_k)$. In the paper of Mukhopadhyay and Chattopadhyay (1991), the following loss function was used. Let

$$L(\underset{\sim}{n}) = A[T(\underset{\sim}{n}) - \theta]^2 + c\sum_{i=1}^k n_i, \tag{2.5}$$

where A and c are known positive numbers. The associated risk is given by $R(\underset{\sim}{n}, c) = E(L(\underset{\sim}{n})) = A\sum_{i=1}^k b_i^2 \lambda_i^2 n_i^{-1} + c\sum_{i=1}^k n_i$, which is minimized for $\underset{\sim}{n} = \underset{\sim}{n}^* = (n_1^*, \ldots, n_k^*)$, $n_i^* = (A/c)^{1/2}|b_i|\lambda_i$, $i = 1, \ldots, k$. Since $\underset{\sim}{n}^*$ is unknown, Mukhopadhyay and Chattopadhyay (1991) proposed the following sequential estimation procedure. Suppose m_i (≥ 1) is the starting sample size from the i-th population and let

$$N_i = \inf\{n \geq m_i : n \geq (A/c)^{1/2}|b_i|\bar{X}_{in}\}, \tag{2.6}$$

for $i = 1, \ldots, k$. Finally, one estimates θ by $T(\underset{\sim}{N})$ where $\underset{\sim}{N} = (N_1, \ldots, N_k)$. In the theorem of Mukhopadhyay and Chattopadhyay (1991), it was shown that

$$E[L(\underset{\sim}{N})] - R(\underset{\sim}{n}^*, c) = 3kc + 2c \sum\sum_{1 \le i < j \le k} d_i d_j + o(c), \tag{2.7}$$

if $\min\{m_1, \ldots, m_k\} \ge 3$, where $d_i = b_i/|b_i|$, $i = 1, \ldots, k$.

Now, having given the stopping variables $N_1, \ldots, N_k$, let us consider the following class of estimators for θ. Choose $\underset{\sim}{a} = (a_1, \ldots, a_k) \in R^k$ such that $2q\sum_{i=1}^k d_i > q^2$ where $q = \sum_{i=1}^k a_i b_i$. Now define $\hat{\theta}_{\underset{\sim}{a}}(\underset{\sim}{N}) = \sum_{i=1}^k b_i\{\bar{X}_{iN_i} + a_i(c/A)^{1/2}\} = T(\underset{\sim}{N}) + q(c/A)^{1/2}$. The risk associated with the estimator $\hat{\theta}_{\underset{\sim}{a}}(\underset{\sim}{N})$ of θ under the loss function (2.5) is given by (in view of (2.7) and (2.3))

$$\begin{aligned} &E[A(\hat{\theta}_{\underset{\sim}{a}}(\underset{\sim}{N}) - \theta)^2 + cN] \\ &\quad = E[A(T(\underset{\sim}{N}) - \theta)^2 + cN] + 2q(cA)^{1/2}E[(T(\underset{\sim}{N}) - \theta)] + q^2 c \\ &\quad = R(\underset{\sim}{n}^*, c) + 3kc + 2c\sum\sum_{1 \le i < j \le k} d_i d_j - c\left(2q\sum_{i=1}^k d_i - q^2\right) + o(c), \end{aligned} \tag{2.8}$$

where one may recall that $N = \sum_{i=1}^k N_i$. That is, $\hat{\theta}_{\underset{\sim}{a}}(\underset{\sim}{N})$ will asymptotically dominate $T(\underset{\sim}{N})$ up to the $o(c)$ term if $2q\sum_{i=1}^k d_i - q^2 > 0$. From (2.8) one notes that the asymptotic "gain" in risk is obviously maximized if $\underset{\sim}{a} \in R^k$ is chosen in such a way that $2q\sum_{i=1}^k d_i - q^2$ is maximized. One should recall that $q = \sum_{i=1}^k a_i b_i$. In general, it does not appear feasible to give an explicit expression of such "optimal" choice for $\underset{\sim}{a}$. However, in a given scenario, one should be able to find improved estimators in the class, as suggested.

Example 1. Consider $k = 2$ and $\theta = \lambda_1 + \lambda_2$. Then, one gets improved estimators $\hat{\theta}_{\underset{\sim}{a}}(\underset{\sim}{N})$ where, for example, $\underset{\sim}{a} = (1, 1)$ or $(1, \frac{1}{2})$ among many other choices. In this case, $\underset{\sim}{a} = (1, 1)$ is "optimal" in our sense in the subclass, when $a_1 = a_2$.

Example 2. Consider $k = 3$ and $\theta = 2\lambda_1 + \lambda_2 - \lambda_3$. Then, one gets improved estimators $\hat{\theta}_{\underset{\sim}{a}}(\underset{\sim}{N})$ where, for example, $\underset{\sim}{a} = (\frac{1}{10}, \frac{1}{10}, \frac{1}{10})$ or $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ or $(\frac{1}{10}, \frac{1}{5}, \frac{1}{10})$ among many other possibilities. In this case, $\underset{\sim}{a} = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ is "optimal" in our sense in the subclass when $a_1 = a_2 = a_3$.

In fact, the following result is in order. Let $B = \sum_{i=1}^k b_i$ and assume that B is positive. Now, let us restrict the choice of "$\underset{\sim}{a}$" so that $a_1 = \cdots = a_k = a$, say, and we also write $D = \sum_{i=1}^{k} d_i$. Then, in order to find the optimal choice for "a", one should maximize $2aDB - a^2B^2$ with respect to $a \in R$. The solution obviously turns out to be $a = D/B$.

Example 3. Consider $k = 3$ and $\theta = \lambda_1 - \lambda_2 - \lambda_3$, so that $B = -1$ and $D = -1$. Then, $\hat{\theta}_{\underset{\sim}{a}}(\underset{\sim}{N})$ will dominate $T(\underset{\sim}{N})$ in our sense if $\underset{\sim}{a} = a\underset{\sim}{1}$ where, for example, $a = 1$ or $\frac{1}{2}$. In fact, any "a" such that $2a > a^2$ will provide an improved version $\hat{\theta}_{\underset{\sim}{a}}(\underset{\sim}{N})$. In this subclass, the "optimal" choice works out when $a = 1$. In other words, the result quoted right before this example also holds if *both B and D are negative.*

3. Piecewise sequential estimation of means

Recently, Bose and Mukhopadhyay (1993) proposed a *piecewise estimation methodology* along the lines of Mukhopadhyay and Sen (1993) for the problem discussed in Subsection 2.1. Suppose there are p (≥ 2) locations or individuals, and each is going to estimate λ in the one-sample case simultaneously in a parallel network and finally obtaining an estimator of λ by combining resources obtained from all the components together. First, we discuss the aspects of improved estimators of λ in the framework of Bose and Mukhopadhyay (1993). Then, we introduce the piecewise methodology for the *multi-sample problem* of Subsection 2.2, followed by similar, but corresponding discussions of associated improvement aspects.

3.1 *The one-sample problem*

Let $X_{i1}, X_{i2}, \ldots, X_{in_i}, \ldots$ be the sequence of i.i.d. random variables having the p.d.f. given by (1.1) associated with the i-th individual, $i = 1, \ldots, p$. One writes $\bar{X}_i(n_i) = n_i^{-1} \sum_{j=1}^{n_i} X_{ij}$, $\underset{\sim}{n} = (n_1, \ldots, n_p)$, $n = \sum_{i=1}^{p} n_i$, $i = 1, \ldots, p$ and the *pooled estimator* of λ is

$$T_p(\underset{\sim}{n}) = n^{-1} \sum_{i=1}^{p} n_i \bar{X}_i(n_i). \tag{3.1}$$

Along the lines of (1.2), suppose that the loss function in estimating λ by $T_p(\underset{\sim}{n})$ is given by

$$L(\underset{\sim}{n}) = A[T_p(\underset{\sim}{n}) - \lambda]^2 + cn, \tag{3.2}$$

where A and c are the same known positive numbers as before. The risk associated with (3.2) is then $R(\underset{\sim}{n}, c) = E[L(\underset{\sim}{n})] = A\lambda^2 n^{-1} + cn$, which is minimized if $n = n^* = (A/c)^{1/2}\lambda$. Instead of estimating n^* as a whole, Bose and Mukhopadhyay (1993) proposed that the i-th individual estimates $n_i^* = p^{-1}n^*$, $i = 1, \ldots, p$ and they considered p separate and independent stopping times. Each of the p individuals starts with m (≥ 1) samples and proceeds purely sequentially according to the stopping variable,

$$N_i = \inf\{n \geq m : n \geq p^{-1}(A/c)^{1/2}\bar{X}_i(n)\}, \tag{3.3}$$

$i = 1, \ldots, p$. Finally, one combines all the samples thus obtained and estimates λ by means of $T_p(\underset{\sim}{N})$. Bose and Mukhopadhyay (1993) showed that

$$E[L(\underset{\sim}{N})] = R(\underset{\sim}{n}^*, c) + 3c + o(c), \tag{3.4}$$

if $m \geq 3$, where $\underset{\sim}{n}^* = (n_1^*, \ldots, n_p^*)$.

Along the lines of (2.2), let us consider a class of estimators for λ given by

$$\hat{\lambda}_a(p, \underset{\sim}{N}) = T_p(\underset{\sim}{N}) + a(c/A)^{1/2}, \tag{3.5}$$

for $a \in R$ such that $2a > a^2$.

One combines (2.3) and uses tools from Bose and Mukhopadhyay (1993) to write (if $m \geq 3$),

$$\begin{aligned} E[T_p(\underset{\sim}{N}) - \lambda] &= pE[N_1(\bar{X}_1(N_1) - \lambda)/N] \\ &= -\lambda n^{*-1} + o(c^{1/2}). \end{aligned} \tag{3.6}$$

Thus, for $m \geq 3$, one combines (3.1)–(3.6) to write

$$\begin{aligned} &E[A(\hat{\lambda}_a(p, \underset{\sim}{N}) - \lambda)^2 + cN] \\ &\quad = E[A\{T_p(\underset{\sim}{N}) - \lambda\}^2 + cN] + 2a(Ac)^{1/2}E[T_p(\underset{\sim}{N}) - \lambda] + a^2c \\ &\quad = R(\underset{\sim}{n}^*, c) + 3c + ca^2 - 2ac + o(c). \end{aligned} \tag{3.7}$$

That is, $\hat{\lambda}_a(p, \underset{\sim}{N})$ will *asymptotically dominate* $T_p(\underset{\sim}{N})$ up to the $o(c)$ term if $2a > a^2$. The asymptotic "gain" in risk is obviously maximized if $a = 1$. This is clear from (3.7). In other words, (3.5) provides a class of estimators dominating the Bose and Mukhopadhyay (1993) estimator $T_p(\underset{\sim}{N})$, while the "best" one in this class turns out to be $\hat{\lambda}_1(p, \underset{\sim}{N})$.

3.2 *The multi-sample problem*

First we propose a piecewise sequential analogue of Mukhopadhyay and Chattopadhyay's (1991) estimation procedure. Then, we address the associated improvement aspects. Suppose we have $X_{ij1}, X_{ij2}, \ldots, X_{ijn_{ij}}, \ldots$ i.i.d. from the

j-th "location", corresponding to the i-th treatment (population), say, characterized by the p.d.f. $f(x;\lambda_i)$, as in Subsection 2.2, $j = 1,\dots,p_i$ (≥ 2), $i = 1,\dots,k$ (≥ 2). Let $\theta = \sum_{i=1}^k b_i\lambda_i$ where $b_1,\dots,b_k$ are known and non-zero real numbers as before. Let us write $\underset{\sim}{p} = (p_1,\dots,p_k)$, $\underset{\sim}{n}_i = (n_{i1},\dots,n_{ip_i})$ and $\underset{\sim}{n} = (\underset{\sim}{n}_1,\dots,\underset{\sim}{n}_k)$ for $i = 1,\dots,k$. Now, the parameter λ_i will be estimated by

$$(3.8) \qquad T_{ip_i}(\underset{\sim}{n}_i) = n_i^{-1}\sum_{j=1}^{p_i} n_{ij}\bar{X}_j(n_{ij}),$$

where $\bar{X}_j(n_{ij}) = n_{ij}^{-1}\sum_{l=1}^{n_{ij}} X_{ijl}$, $n_i = \sum_{j=1}^{p_i} n_{ij}$, $i = 1,\dots,k$. Thus, θ is estimated by means of

$$(3.9) \qquad T(\underset{\sim}{p},\underset{\sim}{n}) = \sum_{i=1}^k b_i T_{ip_i}(\underset{\sim}{n}_i),$$

which is the *piecewise version* of $T(\underset{\sim}{n})$ considered in Subsection 2.2. Suppose that the loss function in estimating θ by $T(\underset{\sim}{p},\underset{\sim}{n})$ is given by

$$(3.10) \qquad L(\underset{\sim}{n}) = A[T(\underset{\sim}{p},\underset{\sim}{n}) - \theta]^2 + c\sum_{i=1}^k n_i,$$

along the lines of (2.5). The associated risk is given by $R(\underset{\sim}{n},c) = E[L(\underset{\sim}{n})] = \sum_{i=1}^k[Ab_i^2\lambda_i^2 n_i^{-1} + cn_i]$, which is minimized if $n_i = n_i^* = (A/c)^{1/2}|b_i|\lambda_i$, for $i = 1,\dots,k$. Now, each "location" within the i-th population starts sampling with m_i (≥ 1) observations and let

$$(3.11) \qquad N_{ij} = \inf\{n \geq m_i : n \geq p_i^{-1}(A/c)^{1/2}|b_i|\bar{X}_j(n)\},$$

for $j = 1,\dots,p_i$ and $i = 1,\dots,k$. In other words $N_{i1},\dots,N_{ip_i}$ are i.i.d. estimators of $n_{ij}^* = p_i^{-1}n_i^*$, that is, all p_i pieces in the j-th population, when combined, estimate n_i^*. Finally, we propose to estimate θ by means of $T(\underset{\sim}{p},\underset{\sim}{N})$. Now, after utilizing (3.8)–(3.11), the associated risk is given by

$$(3.12) \qquad \begin{aligned} E[L(\underset{\sim}{N})] &= \sum_{i=1}^k E[b_i^2 A\{T_{ip_i}(\underset{\sim}{N}_i) - \lambda_i\}^2 + cN_i] \\ &\quad + 2A\sum\sum_{1\leq i<l\leq k} b_i b_l E[T_{ip_i}(\underset{\sim}{N}_i) - \lambda_i]E[T_{lp_l}(\underset{\sim}{N}_l) - \lambda_l] \\ &= \mathrm{I} + \mathrm{II}, \quad \text{say.} \end{aligned}$$

Here, we obviously write $N_i = \sum_{j=1}^{p_i} N_{ij}$. Now, from the theorem of Bose and Mukhopadhyay (1993), it follows that

$$(3.13) \qquad \mathrm{I} = R(\underset{\sim}{n}^*,c) + 3kc + o(c),$$

if $\min\{m_1, \ldots, m_k\} \geq 3$. Next, utilize (3.6) to evaluate II and verify that

$$E[T_{ip_i}(\underset{\sim}{N}_i) - \lambda_i] = -\lambda_i n_i^{*-1} + o(c^{1/2}), \tag{3.14}$$

if $\min\{m_1, \ldots, m_k\} \geq 3$. Hence, from (3.12)–(3.14), one can write

$$E[L(\underset{\sim}{N})] = R(\underset{\sim}{n}^*, c) + 3kc + 2c \sum\sum_{1 \leq i < l \leq k} d_i d_l + o(c), \tag{3.15}$$

if $\min\{m_1, \ldots, m_k\} \geq 3$ where $d_i = b_i/|b_i|$ for $i = 1, \ldots, k$. The second-order expansion of the associated regret is then same (up to $o(c)$) as that for the original purely sequential estimator of θ proposed in Mukhopadhyay and Chattopadhyay (1991). This is clear from (3.15). However, the piecewise methodology (3.11) has significantly *more operational convenience* than the Mukhopadhyay-Chattopadhyay procedure. On top of that we can unbiasedly estimate the $\mathrm{Var}(N_i)$ by

$$\Delta_i = p_i(p_i - 1)^{-1} \sum_{j=1}^{p_i} (N_{ij} - p_i^{-1} N_i)^2, \tag{3.16}$$

where recall that $N_i = \sum_{j=1}^{p_i} N_{ij}$, $i = 1, \ldots, k$. The type of estimators proposed in (3.16) was originally put forth in Mukhopadhyay and Sen (1993) for the one-sample normal problem.

Now, consider the following class of analogous improved version of $T(\underset{\sim}{p}, \underset{\sim}{N})$ in order to estimate θ. Let us define

$$\hat{\theta}_a(\underset{\sim}{p}, \underset{\sim}{N}) = T(\underset{\sim}{p}, \underset{\sim}{N}) + a(c/A)^{1/2}, \tag{3.17}$$

where $a \in R$ such that $2a \sum_{i=1}^k d_i > a^2$. Obviously, for $\min\{m_1, \ldots, m_k\} \geq 3$, from (3.6) one obtains

$$E[T(\underset{\sim}{p}, \underset{\sim}{N}) - \theta] = -(c/A)^{1/2} \sum_{i=1}^k d_i + o(c^{1/2}), \tag{3.18}$$

and hence from (3.17)–(3.18),

$$\begin{aligned} &E[A(\hat{\theta}_a(\underset{\sim}{p}, \underset{\sim}{N}) - \theta)^2 + cN] \\ &\quad = E[A(T(\underset{\sim}{p}, \underset{\sim}{N}) - \theta)^2 + cN] + ca^2 - 2ac \sum_{i=1}^k d_i + o(c). \end{aligned} \tag{3.19}$$

That is, in view of (3.15) and (3.19), $\hat{\theta}_a(\underset{\sim}{p}, \underset{\sim}{N})$ *asymptotically dominates* $T(\underset{\sim}{p}, \underset{\sim}{N})$ up to the $o(c)$ term if $2a \sum_{i=1}^k d_i > a^2$. The asymptotic "gain" in risk is obviously maximized if $a = \sum_{i=1}^k d_i$. In other words, (3.17) provides a class of estimators dominating $T(\underset{\sim}{p}, \underset{\sim}{N})$, while the best one in this class is obtained when $a = \sum_{i=1}^k d_i$.

Acknowledgements

I thank the two referees for several helpful comments. I am particularly grateful to the referee who had questioned an earlier incorrect version of equation (3.6) and provided the reference to the paper of Isogai and Uno (1992).

Appendix

Here we provide a proof of (3.6). We write $I(\cdot)$ for the indicator function of $(\cdot)$. Observe that

$$E[N_1(\bar{X}_1(N_1)-\lambda)/N] = \lambda E[(S_1(N_1)-N_1)/N] \tag{A.1}$$

where one writes $S_1(n)=\sum_{i=1}^{n} Y_i$ and $\mathcal{R}=N_1^2 n_1^{*-1} - S_1(N_1)$, the Y's being i.i.d. standard exponentials. Note that $\mathcal{R}$ is the overshoot for the stopping variable N_1 defined in (3.3).

Now, $n^*N^{-1} = n^{*-1}\{(N-n^*)^2N^{-1} - N + 2n^*\}$ and hence (A.1) leads to

$$\begin{aligned}(S_1(N_1)-N_1)/N &= n^{*-1}\left\{\frac{(N_1-n_1^*)^2}{n_1^*}\frac{n^*}{N} + (N_1-n_1^*)\frac{n^*}{N} - \mathcal{R}\frac{n^*}{N}\right\}\\ &= n^{*-1}\{\mathrm{I}_1+\mathrm{I}_2-\mathrm{I}_3\}, \quad \text{say.}\end{aligned} \tag{A.2}$$

From Lemma 2.3 and Theorem 2.3 of Woodroofe (1977), one respectively obtains

$$P\left(N_1 \le \frac{1}{2}n_1^*\right) = O(n_1^{*-m}), \qquad P\left(N \le \frac{1}{2}n^*\right) = O(n^{*-m}), \qquad \text{and} \tag{A.3}$$

$$|N_1^*|^s \text{ is uniformly integrable if } m > \frac{1}{2}s \text{ for } s>0, \tag{A.4}$$

where $N_1^* = (N_1-n_1^*)/n_1^{*1/2}$ which converges to $N(0,1)$ in distribution.

Let us evaluate the term I_1 in (A.2). Since $N_1^{*2}n^*N^{-1}I(N\ge \frac{1}{2}n^*) \le 2N_1^{*2}$, the l.h.s. is uniformly integrable if $m\ge 2$. Also, the l.h.s. $\xrightarrow{\mathcal{L}} \chi_1^2$, hence $E(\mathrm{I}_1 I(N\ge \frac{1}{2}n^*)) = 1+o(1)$ if $m\ge 2$. But $E[N_1^{*2}n^*N^{-1}I(N<\frac{1}{2}n^*)] \le n^*E^{1/2}(N_1^{*4})\cdot P^{1/2}(N<\frac{1}{2}n^*) = n^*O(1)O(n^{*-m/2}) = o(1)$ if $m\ge 3$. Thus,

$$E(\mathrm{I}_1) = 1+o(1) \quad \text{if} \quad m\ge 3. \tag{A.5}$$

This follows after repeated use of (A.3)–(A.4).

Now, we consider I_3 in (A.2). Let $D=\sum_{n=1}^{\infty} n^{-1}E[S_1(n)-2n]^+$ and $\nu = 1-D$, where x^+ stands for $\max(0,x)$. Since $\mathcal{R}n^*N^{-1}I(N\ge\frac{1}{2}n^*) \le 2\mathcal{R}$, from Lemma 2.1 of Woodroofe (1977) it immediately follows that $E[\mathcal{R}n^*N^{-1}I(N\ge \frac{1}{2}n^*)] = \nu+o(1)$ if $m\ge 2$. On the other hand, note that $E[\mathcal{R}n^*N^{-1}I(N<\frac{1}{2}n^*)] \le n^*E^{1/2}(\mathcal{R}^2)P^{1/2}(N<\frac{1}{2}n^*) = n^*O(1)O(n^{*-m/2}) = o(1)$ if $m\ge 3$, in view of (A.3) and Lemma 2.1 of Woodroofe (1977). Thus,

$$E(\mathrm{I}_3) = \nu + o(1) \quad \text{if} \quad m\ge 3. \tag{A.6}$$

Let us now evaluate the term I_2 in (A.2). One notes that

$$\begin{aligned}\text{(A.7)}\qquad \mathrm{I}_2 &= (N_1 - n_1^*)n^{*-1}\{(N-n^*)^2N^{-1} - (N-n^*) + n^*\}\\ &= n^{*-1}(N_1 - n_1^*)(N-n^*)^2N^{-1} - n^{*-1}(N_1-n_1^*)(N-n^*)\\ &\quad + (N_1 - n_1^*)\\ &= \mathrm{I}_{21} - \mathrm{I}_{22} + \mathrm{I}_{23}, \qquad \text{say.}\end{aligned}$$

Observe that

$$\begin{aligned}\text{(A.8)}\ |E(\mathrm{I}_{21})| &\le n_1^{*1/2}n^{*-1}E[\{|N_1 - n_1^*|/n_1^{*1/2}\}\{(N-n^*)^2/N\}]\\ &\le O(n^{*-1/2})E^{1/r}[\{|N_1-n_1^*|/n_1^{*1/2}\}^r]E^{1/s}[\{(N-n^*)^2/N\}^s],\end{aligned}$$

by Holder's inequality where $r > 1$, $s > 1$ and $r^{-1} + s^{-1} = 1$. Now, by (A.4), $E[\{|N_1 - n_1^*|/n_1^{*1/2}\}^r] < \infty$ if $m > \frac{1}{2}r$. On the other hand, $\{(N-n^*)^2/N\}^s I(N \ge \frac{1}{2}n^*) \le 2^s\{|N-n^*|/n^{*1/2}\}^{2s}$, and the r.h.s. is uniformly integrable if $m > \frac{1}{2}(2s) = s$. This can be verified along the lines of part (ii) in Theorem 3.2 of Mukhopadhyay and Sen (1993). Thus, $E[\{(N-n^*)^2/N\}^s I(N \ge \frac{1}{2}n^*)] < \infty$ if $m > s$. Also, $E[\{(N-n^*)^2/N\}^s I(N < \frac{1}{2}n^*)] \le E[\{(N-n^*)^2\}^s I(N < \frac{1}{2}n^*)] \le O(n^{*2s})P(N < \frac{1}{2}n^*) = O(n^{*2s-m}) = O(1)$ if $m \ge 2s$, by (A.3). That is, $E[\{(N-n^*)^2/N\}^s] < \infty$ if $m \ge 2s$. In other words, from (A.8) one obtains $E(\mathrm{I}_{21}) = o(1)$, if $m > \frac{1}{2}r$ and $m \ge 2s$. Now, plug in $s = \frac{5}{4}$ and $r = 5$, and hence $E(\mathrm{I}_{21}) = o(1)$ if $m \ge 3$. For the term I_{22} in (A.7), let us first write that $\mathrm{I}_{22} = n^{*-1}(N_1 - n_1^*)^2 + \sum_{i=2}^{p}(N_1 - n_1^*)(N_i - n_i^*)n^{*-1}$, and hence for $m \ge 2$,

$$\begin{aligned}E(\mathrm{I}_{22}) &= (n_1^*/n^*)E[N_1^{*2}] + (p-1)n^{*-1}E[(N_1-n_1^*)(N_2-n_2^*)]\\ &= p^{-1}(1+o(1)) + (p-1)n^{*-1}\{(\nu-1)+o(1)\}^2\\ &= p^{-1} + o(1),\end{aligned}$$

in view of (A.4). Obviously, $E(\mathrm{I}_{23}) = (\nu - 1) + o(1)$ for $m \ge 2$. Thus, from (A.7), one obtains

$$\text{(A.9)}\qquad E(\mathrm{I}_2) = -p^{-1} + (\nu - 1) + o(1) \quad \text{if} \quad m \ge 3.$$

Now, combine (A.2), (A.5), (A.6) and (A.9) for $m \ge 3$ to conclude that

$$\begin{aligned}&E[(S_1(N_1) - N_1)/N]\\ &\quad = n^{*-1}\{1 - p^{-1} + (\nu-1) - \nu + o(1)\}\\ &\quad = -(pn^*)^{-1} + o(n^{*-1}).\end{aligned}$$

This leads to (3.6).

References

Bose, A. and Mukhopadhyay, N. (1993). Sequential estimation of the mean of an exponential distribution via replicated piecewise stopping number, Tech. Report, No. 93-06, Department of Statist., University of Connecticut, Storrs.

Ghosh, M. and Mukhopadhyay, N. (1989). Sequential estimation of the percentiles of exponential and normal distributions, *South African Statist. J.*, **23**, 251–268.

Isogai, E. and Uno, C. (1992). On the sequential point estimation of the mean of a gamma distribution (submitted).

Isogai, E. and Uno, C. (1993). Sequential estimation of a parameter of an exponential distribution, *Ann. Inst. Statist. Math.* (in press).

Mukhopadhyay, N. (1987). Minimum risk point estimation of the mean of a negative exponential distribution, *Sankhyā Ser. A*, **49**, 105–112.

Mukhopadhyay, N. and Chattopadhyay, S. (1991). Sequential methodologies for comparing exponential mean survival times, *Sequential Anal.*, **10**, 139–148.

Mukhopadhyay, N. and Sen, P. K. (1993). Replicated piecewise stopping numbers and sequential analysis, *Sequential Anal.*, **12**, 179–197.

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

Woodroofe, M. (1977). Second order approximations for sequential point and interval estimation, *Ann. Statist.*, **5**, 985–995.