International Journal of Statistics and Probability; Vol. 1, No. 2; 2012
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

On Second-order Approximations to the Risk in Estimating the Exponential Mean by a Two-stage Procedure

Eiichi Isogai[1] & Chikara Uno[2]

[1] Department of Mathematics, Niigata University, Niigata, Japan

[2] Department of Mathematics, Akita University, Akita, Japan

Correspondence: Eiichi Isogai, Department of Mathematics, Niigata University, 8050 Ikarashi 2-No-Cho, Nishi-Ku, Niigata 950-2181, Japan. Tel: 81-25-262-6121. E-mail: isogai@math.sc.niigata-u.ac.jp

Received: June 25, 2012 Accepted: July 11, 2012 Online Published: July 30, 2012

doi:10.5539/ijsp.v1n2p47 URL: http://dx.doi.org/10.5539/ijsp.v1n2p47

The research of the first and second authors is financed by the Japan Society for the Promotion of Science through its Grant-in-Aid for Science Research (C), under contract numbers 23540128 and 24540107, respectively. The authors are grateful to the referees for their helpful suggestions

Abstract

We consider the problem of minimum risk point estimation of the mean of an exponential distribution under the assumption that the mean exceeds some positive known number. For this problem Mukhopadhyay and Duggan (2001) proposed a two-stage procedure and provided second-order approximations to the lower and upper bounds for the regret. Under the same set up we give second-order approximations to the regret and compare our approximations with those of them. It turns out that our bounds for the regret are sharper. We also propose a bias-corrected procedure which reduces the risk.

Keywords: regret, exponential, second-order approximation, bias-correction, two-stage procedure

1. Introduction

Let $X_1, X_2, X_3, \dots$ be a sequence of independent and identically distributed random variables from an exponential distribution having the probability density function

$$f(x;\lambda) = \lambda^{-1}\exp(-x/\lambda)I(x>0),$$

where the mean $\lambda \in (0,\infty)$ of the distribution is assumed unknown parameter. Here and elsewhere, $I(\cdot)$ would stand for the indicator function of $(\cdot)$. As an application, consider the lifetime of a system component which can be usefully represented by an exponential random variable. Exponential distributions have been widely used in many reliability and life testing experiments, and so investigated by many authors (see Balakrishnan & Basu (1995), for instance). Under the assumption that the mean λ exceeds some number λ_L where $\lambda_L\,(>0)$ is known to the experimenter, Mukhopadhyay and Duggan (2001) considered the problem of minimum risk point estimation for λ via a two-stage procedure and derived second-order lower and upper bounds of the regret function. In this paper we consider the same problem under the same setup as Mukhopadhyay and Duggan (2001). For a review of sequential estimation problems one may refer to Mukhopadhyay (1988), Gosh, Mukhopadhyay and Sen (1997) and Mukhopadhyay and de Silva (2009).

On the basis of the random sample $X_1, \dots, X_n$ of size n, we want to estimate the mean λ by the sample mean $\overline{X}_n = n^{-1}\sum_{i=1}^{n} X_i$ under the squared error loss plus linear cost

$$L_n(\overline{X}_n;\lambda) = (\overline{X}_n-\lambda)^2 + cn, \tag{1}$$

where $c\,(>0)$ is the known cost per unit sample. Then, the risk is given by $R_n(c) = E\{L_n(\overline{X}_n;\lambda)\} = \lambda^2 n^{-1} + cn$ which is minimized when

$$n = n_0 = c^{-1/2}\lambda. \tag{2}$$

The associated minimum risk is $R_{n_0}(c) = 2cn_0$. The goal is to achieve this minimum risk as closely as possible. Unfortunately λ is unknown, so we cannot use the optimal fixed sample size n_0, thus making it necessary to find a sequential sampling rule. Mukhopadhyay and Duggan (2001) dealt with this minimum risk point estimation problem under the assumption that $\lambda > \lambda_L$ and explored a two-stage estimation methodology under the loss function (1). They then developed second-order bounds for the associated regret. In this paper we use the two-stage procedure below proposed by Mukhopadhyay and Duggan (2001) under the assumption that $\lambda > \lambda_L$. The initial sample size is defined by

$$m \equiv m(c) = \max\left\{m_0,\ \left[c^{-1/2}\lambda_L\right]^* + 1\right\}, \tag{3}$$

where $m_0\ (\geq 1)$ is a preassigned integer and $[x]^*$ denotes the largest integer less than x. Based on the pilot sample $X_1, \ldots, X_m$, we calculate the sample mean $\overline{X}_m$ and define

$$N \equiv N(c) = \max\left\{m,\ \left[c^{-1/2}\overline{X}_m\right]^* + 1\right\}. \tag{4}$$

If $N > m$, then one takes the second sample $X_{m+1}, \ldots, X_N$. By using the total observations $X_1, \cdots, X_N$, we estimate λ by $\overline{X}_N$. The risk is given by $E\{L_N(\overline{X}_N; \lambda)\} = E\{(\overline{X}_N - \lambda)^2 + cN\}$ and the regret is defined by

$$\omega(c) = E[L_N(\overline{X}_N; \lambda)] - R_{n_0}(c) = E[(\overline{X}_N - \lambda)^2] + cE(N) - 2cn_0. \tag{5}$$

The purpose of this paper is to provide second-order approximations to the regret $\omega(c)$ as c tends to zero and compare them with the results of Theorem 3.2 of Mukhopadhyay and Duggan (2001). Our bounds for the regret are proved to be sharper than those of them. We can also show that the purely sequential procedure in Woodroofe (1977) is more efficient than the two-stage procedure in regret under a certain condition. In order to reduce the risk we propose a bias-corrected procedure. In Section 2 we present the main results with second-order approximations to the regret and compare them with those of Mukhopadhyay and Duggan (2001). Section 3 gives brief simulation results. In Section 4 we describe conclusions. In the appendix we supply all the proofs of the theorems in Section 2.

2. Second-order Approximations

In this section we provide the main results with second-order approximations to the regret for the two-stage and bias-corrected procedures. The following theorem gives a second-order approximation to the regret.

Theorem 1 *We have as $c \to 0$*

$$\frac{\omega(c)}{c} = 2 + \frac{\lambda}{\lambda_L} + \varepsilon_c,$$

where

$$|\varepsilon_c| \leq \frac{\lambda_L}{\lambda} + o(1).$$

Comparison. (i) We compare our bounds for the regret with those of Mukhopadhyay and Duggan (2001) who provided the following lower and upper bounds for the regret:

$$B_1 + o(1) \leq \omega(c)/c \leq B_2 + o(1) \quad \text{as } c \to 0,$$

where $B_1 = -12 + (\lambda/\lambda_L)$ and $B_2 = 8 + (\lambda/\lambda_L)$. Let

$$A_1 = 2 + (\lambda/\lambda_L) - (\lambda_L/\lambda) \quad \text{and} \quad A_2 = 2 + (\lambda/\lambda_L) + (\lambda_L/\lambda).$$

Then from Theorem 1 we get the bounds for the regret

$$A_1 + o(1) \leq \omega(c)/c \leq A_2 + o(1) \quad \text{as } c \to 0 \tag{6}$$

and

$$A_1 - B_1 = 14 - (\lambda_L/\lambda) > 13 \quad \text{and} \quad B_2 - A_2 = 6 - (\lambda_L/\lambda) > 5,$$

since $0 < \lambda_L/\lambda < 1$. Therefore our bounds are sharper than those of Mukhopadhyay and Duggan (2001) for sufficiently small c.

(ii) For the purely sequential procedure

$$N^* = \inf\left\{n \geq m :\ n \geq c^{-1/2}\overline{X}_n\right\} \tag{7}$$

without the condition $\lambda > \lambda_L$, Woodroofe (1977) showed that $\{E[L_{N^*}(\overline{X}_{N^*};\lambda)] - R_{n_0}(c)\}/c = 3 + o(1)$ as $c \to 0$, provided $m \geq 3$. It follows from (6) that if $A_1 > 3$, or equivalently, $(\lambda/\lambda_L) - (\lambda_L/\lambda) > 1$, then the purely sequential procedure N^* is superior to the two-stage procedure N in regret for sufficiently small c. Obviously, if $(\lambda/\lambda_L) > (\sqrt{5}+1)/2$ then it holds that $(\lambda/\lambda_L) - (\lambda_L/\lambda) > 1$. Thus if λ_L is sufficiently small compared to λ then the purely sequential procedure should be used.

Now we shall consider the bias of $\overline{X}_N$.

Theorem 2 *We have as $c \to 0$*

$$E(\overline{X}_N) = \lambda - c^{1/2} + O(c^{3/4}).$$

Taking the bias of $\overline{X}_N$ into account, we propose the following bias-corrected procedure:

$$\delta_N = \overline{X}_N + c^{1/2}.$$

Then the associated risk is given by $E[L_N(\delta_N;\lambda)] = E[(\delta_N - \lambda)^2 + cN]$. The following theorem shows that the bias-corrected procedure reduces one sample cost compared with the two-stage procedure (3) and (4). Thus the bias-correction is a little bit more effective in reducing the risk for sufficiently small cost.

Theorem 3 *As $c \to 0$ we have*

$$E[L_N(\overline{X}_N;\lambda)] - E[L_N(\delta_N;\lambda)] = c + O(c^{5/4}).$$

3. Simulation Results

In this section we shall present brief simulation results. We consider the case $\lambda = 1$ and let $m_0 = 3, 10$ in (3). The two-stage procedure N defined by (3) and (4) and the purely sequential procedure N^* in (7) was carried out with 1,000,000 independent replications for $\lambda_L = 0.2,\ 0.4,\ 0.6$ when $n_0 = 30, 50$ and 100, namely $c = 0.0011, 0.0004$ and 0.0001, respectively. In six tables, $\bar{N}$, $\bar{X}_N$, $\bar{\delta}_N$, $\bar{\omega}_1(c)/c$, $\bar{\omega}_2(c)/c$, $\bar{\omega}_3(c)/c$ are the averages of N, $\overline{X}_N$, δ_N, $L_1(c)/c$, $L_2(c)/c$, $L_3(c)/c$, respectively, where

$$L_1(c) = L_N(\overline{X}_N;\lambda) - R_{n_0}(c), \quad L_2(c) = L_N(\delta_N;\lambda) - R_{n_0}(c) \quad \text{and} \quad L_3(c) = L_{N^*}(\overline{X}_{N^*};\lambda) - R_{n_0}(c)$$

while $s(\cdot)$ denotes the standard error of the estimator of $(\cdot)$. For $0 < \lambda_L < (\sqrt{5}-1)/2 = 0.618$ the inequality $(\lambda/\lambda_L) - (\lambda_L/\lambda) > 1$ holds. The tables seem to show that (i) the regret becomes smaller as λ_L grows larger, (ii) our bias-corrected procedure reduces one sample cost in risk, (iii) m_0 in (3) almost does not affect the regret and (iv) our theorems and the results in comparison in Section 2 are verified.

Table 1. Two-stage ($m_0 = 3$) and purely sequential procedures for $n_0 = 30$

λ_L	0.2	0.4	0.6			
m ($m_0 = 3$)	6	12	18	m	3	18
$\bar{N}$	30.484217	30.513074	30.531488	$\bar{N}^*$	29.443295	29.722876
$[s(\bar{N})]$	0.012231	0.008641	0.006987	$[s(\bar{N}^*)]$	0.006161	0.005539
$\bar{X}_N$	0.961975	0.965453	0.967265	$\bar{X}_{N^*}$	0.957040	0.965231
$[s(\bar{X}_N)]$	0.000201	0.000191	0.000186	$[s(\bar{X}_{N^*})]$	0.000204	0.000186
bias of $\bar{X}_N$	-0.038025	-0.034547	-0.032735			
bias of $\bar{\delta}_N$	-0.004692	-0.001213	0.000599			
$\bar{\omega}_1(c)/c$	8.313601	4.437137	2.473775	$\bar{\omega}_3(c)/c$	8.732682	2.075925
$[s(\bar{\omega}_1(c)/c)]$	0.060634	0.049632	0.043377	$[s(\bar{\omega}_3(c)/c)]$	0.075117	0.043303
A_1	6.8	4.1	3.07			
A_2	7.2	4.9	4.27			
$\bar{\omega}_2(c)/c$	7.032105	3.364334	1.509687			
$[s(\bar{\omega}_2(c)/c)]$	0.057991	0.047971	0.042621			

Table 2. Two-stage ($m_0 = 3$) and purely sequential procedures for $n_0 = 50$

λ_L	0.2	0.4	0.6			
m ($m_0 = 3$)	10	20	30	m	3	30
$\bar{N}$	50.494496	50.495585	50.506280	$\bar{N}^*$	49.595943	49.677777
$[s(\bar{N})]$	0.015784	0.011158	0.009106	$[s(\bar{N}^*)]$	0.007490	0.007220
$\bar{X}_N$	0.977877	0.979337	0.979838	$\bar{X}_{N^*}$	0.977173	0.978651
$[s(\bar{X}_N)]$	0.000150	0.000146	0.000144	$[s(\bar{X}_{N^*})]$	0.000149	0.000144
bias of $\bar{X}_N$	-0.022123	-0.020663	-0.020162			
bias of $\bar{\delta}_N$	-0.002123	-0.000663	-0.000162			
$\bar{\omega}_1(c)/c$	7.818174	4.612660	3.222004	$\bar{\omega}_3(c)/c$	6.704822	3.010618
$[s(\bar{\omega}_1(c)/c)]$	0.088550	0.079614	0.074541	$[s(\bar{\omega}_3(c)/c)]$	0.110638	0.074978
A_1	6.8	4.1	3.07			
A_2	7.2	4.9	4.27			
$\bar{\omega}_2(c)/c$	6.605852	3.546366	2.205815			
$[s(\bar{\omega}_2(c)/c)]$	0.086159	0.077817	0.073200			

Table 3. Two-stage ($m_0 = 3$) and purely sequential procedures for $n_0 = 100$

λ_L	0.2	0.4	0.6			
m ($m_0 = 3$)	20	40	60	m	3	60
$\bar{N}$	100.468660	100.476463	100.466908	$\bar{N}^*$	99.688633	99.706543
$[s(\bar{N})]$	0.022352	0.015800	0.012905	$[s(\bar{N}^*)]$	0.010175	0.010120
$\bar{X}_N$	0.989363	0.989761	0.989630	$\bar{X}_{N^*}$	0.989473	0.989654
$[s(\bar{X}_N)]$	0.000103	0.000101	0.000101	$[s(\bar{X}_{N^*})]$	0.000102	0.000101
bias of $\bar{X}_N$	-0.010637	-0.010239	-0.010370			
bias of $\bar{\delta}_N$	-0.000637	-0.000239	-0.000370			
$\bar{\omega}_1(c)/c$	7.118705	4.334002	3.448066	$\bar{\omega}_3(c)/c$	4.070238	2.953309
$[s(\bar{\omega}_1(c)/c)]$	0.178955	0.149970	0.147216	$[s(\bar{\omega}_3(c)/c)]$	0.178955	0.147767
A_1	6.8	4.1	3.07			
A_2	7.2	4.9	4.27			
$\bar{\omega}_2(c)/c$	5.991331	3.286131	2.373969			
$[s(\bar{\omega}_2(c)/c)]$	0.154811	0.148287	0.145660			

Table 4. Two-stage ($m_0 = 10$) procedure for $n_0 = 30$

λ_L	0.2	0.4	0.6
m ($m_0 = 10$)	10	12	18
$\bar{N}$	30.502143	30.493566	30.539402
$[s(\bar{N})]$	0.009483	0.008645	0.006996
$\bar{X}_N$	0.964419	0.965011	0.967180
$[s(\bar{X}_N)]$	0.000193	0.000191	0.000185
bias of $\bar{X}_N$	-0.035581	-0.034989	-0.032820
bias of $\bar{\delta}_N$	-0.002247	-0.001656	0.000513
$\bar{\omega}_1(c)/c$	5.326909	4.517480	2.447595
$[s(\bar{\omega}_1(c)/c)]$	0.052342	0.049847	0.043378
A_1	6.8	4.1	3.07
A_2	7.2	4.9	4.27
$\bar{\omega}_2(c)/c$	4.192069	3.418111	1.478373
$[s(\bar{\omega}_2(c)/c)]$	0.050344	0.048110	0.042650

Table 5. Two-stage ($m_0 = 10$) procedure for $n_0 = 50$

λ_L	0.2	0.4	0.6
m ($m_0 = 10$)	10	20	30
$\bar{N}$	50.487837	50.500755	50.500892
$[s(\bar{N})]$	0.015808	0.011166	0.009109
$\bar{X}_N$	0.977925	0.979272	0.979923
$[s(\bar{X}_N)]$	0.000150	0.000146	0.000144
bias of $\bar{X}_N$	-0.022075	-0.020728	-0.020077
bias of $\bar{\delta}_N$	-0.002075	-0.000728	-0.000077
$\bar{\omega}_1(c)/c$	8.001983	4.553595	3.414406
$[s(\bar{\omega}_1(c)/c)]$	0.088908	0.079646	0.075093
A_1	6.8	4.1	3.07
A_2	7.2	4.9	4.27
$\bar{\omega}_2(c)/c$	6.794460	3.480803	2.406726
$[s(\bar{\omega}_2(c)/c)]$	0.086547	0.077820	0.073753

Table 6. Two-stage ($m_0 = 10$) procedure for $n_0 = 100$

λ_L	0.2	0.4	0.6
m ($m_0 = 10$)	20	40	60
$\bar{N}$	100.519676	100.479207	100.485107
$[s(\bar{N})]$	0.022326	0.015825	0.012899
$\bar{X}_N$	0.989482	0.989591	0.989861
$[s(\bar{X}_N)]$	0.000103	0.000102	0.000101
bias of $\bar{X}_N$	-0.010518	-0.010409	-0.010139
bias of $\bar{\delta}_N$	-0.000518	-0.000409	-0.000139
$\bar{\omega}_1(c)/c$	7.159623	4.759837	3.649018
$[s(\bar{\omega}_1(c)/c)]$	0.156130	0.150685	0.148065
A_1	6.8	4.1	3.07
A_2	7.2	4.9	4.27
$\bar{\omega}_2(c)/c$	6.056022	3.678078	2.621164
$[s(\bar{\omega}_2(c)/c)]$	0.154323	0.148963	0.146534

4. Conclusions

For the problem of minimum risk point estimation of the mean of an exponential distribution under the assumption that the mean exceeds some positive known number, we used the two-stage procedure proposed by Mukhopadhyay and Duggan (2001) and provided the second-order approximation to the regret as cost tends to zero. We found that this approximation gives the sharper lower and upper bounds for the regret than those of Mukhopadhyay and Duggan (2001). We also proposed the bias-corrected procedure and showed that this procedure is a little bit more effective than the two-stage procedure in reducing the risk for sufficiently small cost. Furthermore, it turned out that the regret decreases as the lower bound λ_L for the true value λ increases.

5. Appendix

In this appendix we shall give all the proofs of the results in Section 2. Let

$$S_m = \sum_{i=1}^{m} X_i/\lambda, \quad T = c^{-1/2}\overline{X}_m = (n_0/m)S_m \quad \text{and} \quad U_c = [T]^* + 1 - T. \tag{8}$$

Then (4) becomes $N = \max\{m, T + U_c\}$. The two-stage procedure defined by (3) and (4) belongs to the general procedure of Mukhopadhyay and Duggan (1999). In the notations of Uno and Isogai (2012), set

$$\theta = \lambda, \quad h = c^{1/2}, \quad q_m^* = q = 1, \quad U(m) = \overline{X}_m, \quad p_m = 2m \quad \text{and} \quad c_3 = 0.$$

Then Theorem 1 of Uno and Isogai (2012) yields

$$E(N - n_0) = 2^{-1} + O(n_0^{-1}) \quad \text{as } c \to 0. \tag{9}$$

Throughout this appendix K denotes a generic positive constant, not depending on c.

A.1. ***Three Lemmas.*** We shall provide three lemmas in order to prove the theorems. Lemma 1 can be obtained from Mukhopadhyay and Duggan (1999; 2001), Uno and Isogai (2012), (2) and (3).

Lemma 1 *The following statements hold.*

(i) For $h = (\lambda_L/\lambda)\exp\{1 - (\lambda_L/\lambda)\} \in (0, 1)$

$$P(N = m) = O(h^m) \quad as \quad c \to 0$$

and hence, $m^s P(N = m) \to 0$ *as* $c \to 0$ *for all fixed* $s > 0$.

(ii) $N/n_0 \xrightarrow{P} 1$ *as* $c \to 0$, *where* "$\xrightarrow{P}$" *stands for convergence in probability.*

(iii) $m \to \infty$ *and* $m/n_0 = (\lambda_L/\lambda) + O(c^{1/2})$ *as* $c \to 0$.

(iv) $m^{-1/2}(S_m - m) \xrightarrow{\mathcal{D}} N(0, 1)$ *as* $c \to 0$, *where* "$\xrightarrow{\mathcal{D}}$" *stands for convergence in distribution and* $N(0, 1)$ *denotes a standard normal random variable.*

(v) For any fixed $s > 0$, $\{|m^{-1/2}(S_m - m)|^s;\ c > 0\}$ *is uniformly integrable.*

By using Lemma 2.2 (i) of Mukhopadhyay and Duggan (1999) and Lemma 4 (ii) of Uno and Isogai (2012), we have

Lemma 2 *(i)* $n_0^{-1/2}(N - n_0) \xrightarrow{\mathcal{D}} N(0, \lambda/\lambda_L)$ *as* $c \to 0$.

(ii) For any fixed $s \ge 1$, $\{|n_0^{-1/2}(N - n_0)|^s;\ 0 < c \le c_0\}$ *is uniformly integrable for some sufficiently small* $c_0 > 0$.

Lemma 3 is needed to show the second-order approximation to the regret. Let $Y_i = 2X_i/\lambda$. Then $Y_1, Y_2, \ldots$ are independent and identically distributed random variables according to the chi-square distribution χ^2_2 with two degrees of freedom and $2S_m = \sum_{i=1}^m Y_i \sim \chi^2_{2m}$.

Lemma 3 *Let*

$$\Delta_c = -2n_0^{-1} E\left[\left(U_c - 2^{-1}\right)(S_m - m)^2\right].$$

Then we have as $c \to 0$

$$\begin{aligned}
&n_0^{-1} E[(N - n_0)(S_m - m)] = 1 + O(c^{1/4}),\\
&n_0^{-1} E[(N - n_0)(S_m - m)^2] = 2 - 2^{-1}\Delta_c + 2^{-1}(\lambda_L/\lambda) + O(c^{1/2}) \quad \textit{and}\\
&n_0^{-2} E[(N - n_0)^2(S_m - m)^2] = 3 + O(c^{1/4}).
\end{aligned}$$

Proof. Note that

$$\begin{aligned}
N - n_0 &= (T - n_0 + U_c)I(N > m) + (m - n_0)I(N = m)\\
&= (T - n_0 + U_c) - (T - n_0 + U_c)I(N = m) + (m - n_0)I(N = m) \quad \text{and}\\
T - n_0 &= (n_0/m)(S_m - m).
\end{aligned}$$

It follows from the Marcinkiewicz-Zygmund inequality (Gut (2005)) that

$$E[|S_m - m|^{2p}] \le Km^p \quad \text{for any fixed} \quad p \ge 1, \tag{10}$$

where K is a positive constant, depending only on p. The Cauchy-Schwarz inequality, (10) and Lemma 1 (iii) imply

$$E[|T - n_0|^p] \le (n_0/m)^p \left\{E[|S_m - m|^{2p}]\right\}^{1/2} \le Km^{p/2} \quad \text{for all fixed} \quad p \ge 1. \tag{11}$$

We shall show the first assertion. By using Lemma 1 (i), $0 \le U_c \le 1$, (10), (11) and Cauchy-Schwarz's inequality, we get

$$\begin{aligned}
&n_0^{-1} E[(N - n_0)(S_m - m)]\\
&= n_0^{-1} E[(T - n_0)(S_m - m)] + n_0^{-1} E[U_c(S_m - m)]\\
&\quad - n_0^{-1} E[(T - n_0 + U_c)(S_m - m)I(N = m)] + \{(m/n_0) - 1\}E[(S_m - m)I(N = m)]\\
&= m^{-1} E[(S_m - m)^2] + O(c^{1/4}) = 1 + O(c^{1/4}),
\end{aligned}$$

which shows the first assertion. Next we shall prove the second part. Lemma 3 of Uno and Isogai (2012) shows that $E(U_c) = 2^{-1} + O(c^{1/2})$ as $c \to 0$. Taking this result into account, we have

$$\begin{aligned}
&n_0^{-1}E[(N-n_0)(S_m-m)^2] \\
&= n_0^{-1}E[(T-n_0)(S_m-m)^2] + n_0^{-1}E[(U_c-2^{-1})(S_m-m)^2] + 2^{-1}n_0^{-1}E[(S_m-m)^2] \\
&\quad -n_0^{-1}E[(T-n_0+U_c)(S_m-m)^2 I(N=m)] + \{(m/n_0)-1\}E[(S_m-m)^2 I(N=m)]. \qquad (12)
\end{aligned}$$

Lemma 1 (iii) implies that $2^{-1}n_0^{-1}E[(S_m-m)^2] = 2^{-1}(m/n_0) = 2^{-1}(\lambda_L/\lambda) + O(c^{1/2})$. The definition of Δ_c gives that $n_0^{-1}E[(U_c-2^{-1})(S_m-m)^2] = -\Delta_c/2$. By using Lemma 1 (i), (10), (11) and $0 \le U_c \le 1$, we have as $c \to 0$

$$\begin{aligned}
&|n_0^{-1}E[(T-n_0+U_c)(S_m-m)^2 I(N=m)]| \\
&\le n_0^{-1}\{P(N=m)\}^{1/2}\left\{E[(T-n_0)^4]\right\}^{1/4}\left\{E[(S_m-m)^8]\right\}^{1/4} \\
&\quad +n_0^{-1}\{P(N=m)\}^{1/2}\left\{E[(S_m-m)^4]\right\}^{1/2} = o(c^{1/2}) \quad \text{and} \\
&\{(m/n_0)-1\}E[(S_m-m)^2 I(N=m)] = o(c^{1/2}).
\end{aligned}$$

Since $2S_m \sim \chi^2_{2m}$, we have that $n_0^{-1}E[(T-n_0)(S_m-m)^2] = 8^{-1}m^{-1}E[(2S_m-2m)^3] = 2$. Thus, combining the above results with (12) yields the second part. Finally we shall show the third statement. In the same way as the second part we have

$$\begin{aligned}
&n_0^{-2}E[(N-n_0)^2(S_m-m)^2] \\
&= n_0^{-2}E[(T-n_0)^2(S_m-m)^2] + 2n_0^{-2}E[(T-n_0)U_c(S_m-m)^2] \\
&\quad +n_0^{-2}E[U_c^2(S_m-m)^2] + o(c^{1/2}) \\
&\equiv I_1 + 2I_2 + I_3 + o(c^{1/2}), \quad \text{say}. \qquad (13)
\end{aligned}$$

It is easy to see

$$I_1 = m^{-2}E[(S_m-m)^4] = 3m^{-1}(m+2) = 3 + O(c^{1/2}).$$

Also we get as $c \to 0$

$$\begin{aligned}
&|I_2| \le (n_0 m)^{-1}E[|S_m-m|^3] \le K\, n_0^{-1}m^{1/2} = O(c^{1/4}) \quad \text{and} \\
&|I_3| \le n_0^{-2}E[(S_m-m)^2] = n_0^{-2}m = O(c^{1/2}).
\end{aligned}$$

Hence from the above relations and (13) we have the third statement. This completes the proof.

A.2. ***Proof of Theorem 1.*** Along the lines of the proof of Mukhopadhyay and Duggan (2001) we shall show the theorem. Since $E(N) \le c^{-1/2}\lambda + m < \infty$ for all $c > 0$ by (4), Wald's lemma with (2) and (8) implies

$$\begin{aligned}
E[(\overline{X}_N-\lambda)^2] &= cE\left[\left\{\left(\frac{N}{n_0}\right)^{-2}-1\right\}(S_N-N)^2\right] + cE(S_N-N)^2 \\
&= cE\left[\left\{\left(\frac{N}{n_0}\right)^{-2}-1\right\}(S_N-N)^2\right] + cE(N).
\end{aligned}$$

Hence we have from (5) and (9)

$$\begin{aligned}
\frac{\omega(c)}{c} &= E\left[\left\{\left(\frac{N}{n_0}\right)^{-2}-1\right\}(S_N-N)^2\right] + 2E(N-n_0) \\
&= E\left[\left\{\left(\frac{N}{n_0}\right)^{-2}-1\right\}(S_N-N)^2\right] + 1 + O(c^{1/2}). \qquad (14)
\end{aligned}$$

Using conditioning arguments such as (3.12)–(3.13) of Mukhopadhyay and Duggan (2001), we have

$$\begin{aligned}
&E\left[\left\{\left(\frac{N}{n_0}\right)^{-2}-1\right\}(S_N-N)^2\right] \\
&= E\left[\left\{\left(\frac{N}{n_0}\right)^{-2}-1\right\}(S_m-m)^2\right] + E\left[\left\{\left(\frac{N}{n_0}\right)^{-2}-1\right\}(N-m)\right]. \qquad (15)
\end{aligned}$$

By Taylor's theorem, we get

$$\left(\frac{N}{n_0}\right)^{-2} - 1 = -2n_0^{-1}(N - n_0) + 3n_0^{-2}(N - n_0)^2 W^{-4}, \tag{16}$$

where W is a random variable lying between 1 and N/n_0. It follows from Lemma 3 that

$$\begin{aligned}
&E\left[\left\{\left(\frac{N}{n_0}\right)^{-2} - 1\right\}(S_m - m)^2\right] \\
&= -2n_0^{-1}E\left[(N - n_0)(S_m - m)^2\right] + 3n_0^{-2}E\left[(N - n_0)^2(S_m - m)^2\right] \\
&\quad +3n_0^{-2}E\left[(N - n_0)^2(W^{-4} - 1)(S_m - m)^2\right] \\
&= 5 + \Delta_c - (\lambda_L/\lambda) + 3n_0^{-2}E\left[(N - n_0)^2(W^{-4} - 1)(S_m - m)^2\right] + O(c^{1/4}).
\end{aligned} \tag{17}$$

Since for any fixed $r > 0$ $g(x) = x^{-r}$ $(x > 0)$ is a convex function, we get that $W^{-r} \leq 1 + (N/n_0)^{-r}$. (4) implies that $N \geq T = (n_0/m)S_m$. It follows from (3) that $m \geq c^{-1/2}\lambda_L$ which gives that $m > r$ if $0 < c < c_1 \equiv (\lambda_L/\lambda)^2$. Hence we have for any fixed $r > 0$

$$\begin{aligned}
E(W^{-r}) &\leq 1 + E[(N/n_0)^{-r}] \leq 1 + m^r E(S_m^{-r}) \leq 1 + m^r\Gamma(m - r)/\Gamma(m) \\
&\leq K \quad \text{for all } 0 < c < c_1.
\end{aligned} \tag{18}$$

Lemma 1 (ii) and (iv) imply that $W \xrightarrow{P} 1$ and $(W^{-4} - 1)^2 m^{-2}(S_m - m)^4 \xrightarrow{P} 0$ as $c \to 0$. Also Lemma 1 (v) and (18) yield

$$\begin{aligned}
&\sup_{0<c<c_1} E\left[\left\{(W^{-4} - 1)^2 m^{-2}(S_m - m)^4\right\}^2\right] \\
&\leq K\left\{\sup_{0<c<c_1} E(W^{-32}) + 1\right\}^{1/2}\left\{\sup_{0<c<c_1} E\left[\{m^{-1/2}(S_m - m)\}^{16}\right]\right\}^{1/2} \leq K,
\end{aligned}$$

which provides the uniform integrability of $\{(W^{-4} - 1)^2 m^{-2}(S_m - m)^4;\ 0 < c < c_1\}$. Hence we have that $E[(W^{-4} - 1)^2 m^{-2}(S_m - m)^4] = o(1)$ as $c \to 0$. By using Lemmas 1 and 2 we get as $c \to 0$

$$\begin{aligned}
&n_0^{-2}E\left[(N - n_0)^2(W^{-4} - 1)(S_m - m)^2\right] \\
&\leq (m/n_0)\left\{E[n_0^{-2}(N - n_0)^4]\right\}^{1/2}\left\{E\left[(W^{-4} - 1)^2 m^{-2}(S_m - m)^4\right]\right\}^{1/2} = o(1).
\end{aligned}$$

Combining this result with (17), we have as $c \to 0$

$$E\left[\left\{\left(\frac{N}{n_0}\right)^{-2} - 1\right\}(S_m - m)^2\right] = 5 + \Delta_c - (\lambda_L/\lambda) + o(1). \tag{19}$$

We obtain from (9), (16), (18) and Lemmas 1 and 2

$$\begin{aligned}
&E\left[\left\{\left(\frac{N}{n_0}\right)^{-2} - 1\right\}(N - m)\right] \\
&= -2n_0^{-1}E[(N - n_0)(N - m)] + 3n_0^{-2}E[(N - n_0)^2 W^{-4}(N - m)] \\
&= -2E\left\{n_0^{-1}(N - n_0)^2\right\} + 2\{(m/n_0) - 1\}E(N - n_0) \\
&\quad +3n_0^{-1/2}E\left[n_0^{-3/2}(N - n_0)^3 W^{-4}\right] + 3E[n_0^{-1}(N - n_0)^2 W^{-4}]\{1 - (m/n_0)\} \\
&= -2(\lambda/\lambda_L) + \{(\lambda_L/\lambda) - 1\} + 3(\lambda/\lambda_L)\{1 - (\lambda_L/\lambda)\} + o(1) \\
&= -4 + (\lambda/\lambda_L) + (\lambda_L/\lambda) + o(1),
\end{aligned}$$

which, together with (14), (15) and (19), yields

$$\frac{\omega(c)}{c} = 2 + (\lambda/\lambda_L) + \Delta_c + o(1) \quad \text{as} \quad c \to 0.$$

Since $\varepsilon_c = (\omega(c)/c) - (2 + (\lambda/\lambda_L))$ we get that $\varepsilon_c = \Delta_c + o(1)$. We shall prove the inequality for Δ_c in Lemma 3. Using that $|U_c - 2^{-1}| \le 2^{-1}$, we have from Lemma 1 (iii)

$$|\Delta_c| \le 2n_0^{-1}E[2^{-1}(S_m - m)^2] = m/n_0 = \lambda_L/\lambda + O(c^{1/2}).$$

Therefore we get that $|\varepsilon_c| \le |\Delta_c| + o(1) \le \lambda_L/\lambda + o(1)$ as $c \to 0$. This completes the proof of Theorem 1.

A.3. ***Proof of Theorem 2.*** By conditioning arguments,

$$E(\overline{X}_N) = E\left\{\frac{m}{N}\overline{X}_m + \frac{\lambda(N-m)}{N}\right\} = \lambda + c^{1/2}E\left[\left(\frac{N}{n_0}\right)^{-1}(S_m - m)\right]. \tag{20}$$

By Taylor's theorem,

$$\left(\frac{N}{n_0}\right)^{-1} - 1 = -n_0^{-1}(N - n_0) + n_0^{-2}(N - n_0)^2 W^{-3},$$

where W is a random variable lying between 1 and N/n_0. We have from Lemmas 1 and 3, (10), (11) and (18) and the fact that $E(S_m - m) = 0$

$$\begin{aligned}
&E\left[\left(\frac{N}{n_0}\right)^{-1}(S_m - m)\right] = E\left[\left\{\left(\frac{N}{n_0}\right)^{-1} - 1\right\}(S_m - m)\right] \\
&= -n_0^{-1}E[(N - n_0)(S_m - m)] + n_0^{-2}E[(N - n_0)^2 W^{-3}(S_m - m)] \\
&= -1 + m^{-1/2}E[W^{-3}m^{-3/2}(S_m - m)^3] \\
&\quad +2n_0^{-2}E[(T - n_0)U_c W^{-3}(S_m - m)] + n_0^{-2}E[U_c^2 W^{-3}(S_m - m)] + O(c^{1/4}) \\
&= -1 + O(c^{1/4}),
\end{aligned}$$

which, together with (20), yields the theorem. Therefore the proof of Theorem 2 is complete.

A.4. ***Proof of Theorem 3.*** From Theorem 2,

$$\begin{aligned}
E\{L_N(\delta_N;\lambda)\} &= E[(\overline{X}_N - \lambda + c^{1/2})^2] + cE(N) \\
&= E\{L_N(\overline{X}_N;\lambda)\} + 2c^{1/2}E(\overline{X}_N - \lambda) + c \\
&= E\{L_N(\overline{X}_N;\lambda)\} + 2c^{1/2}\left\{-c^{1/2} + O(c^{3/4})\right\} + c \\
&= E\{L_N(\overline{X}_N;\lambda)\} - c + O(c^{5/4}),
\end{aligned}$$

which implies the theorem. This completes the proof.

References

Balakrishnan, N., & Basu, A. P. (1995). *The Exponential Distributions: Theory, Methods and Applications*. Amsterdam: Gordon and Breach Publishers.

Ghosh, M., Mukhopadhyay, N., & Sen, P. K. (1997). *Sequential Estimation*. New York: Wiley. http://dx.doi.org/10.1002/9781118165928

Gut, A. (2005). *Probability: A Graduate Course*. New York, NY: Springer.

Mukhopadhyay, N. (1988). Sequential estimation problems for negative exponential populations. *Communications in Statistics-Theory & Methods, 17*, 2471-2506. http://dx.doi.org/10.1080/03610928808829758

Mukhopadhyay, N., & de Silva, B. M. (2009). *Sequential Methods and Their Applications*. Chapman & Hall/CRC, Boca Raton, FL.

Mukhopadhyay, N., & Duggan, W. (1999). On a two-stage procedure having second-order properties with applications. *Annals of the Institute of Statistical Mathematics, 51*, 621-636. http://dx.doi.org/10.1023/A:1004074912105

Mukhopadhyay, N., & Duggan, W. T. (2001). A two-stage point estimation procedure for the mean of an exponential distribution and second-order results. *Statistics & Decisions, 19*, 155-171.

Uno, C., & Isogai, E. (2012). Asymptotic properties of a two-stage procedure and their applications. *submitted to Far East J. Theor. Statist*.

Woodroofe, M. (1977). Second order approximations for sequential point and interval estimation. *The Annals of Statistics, 5*, 984-995. http://dx.doi.org/10.1214/aos/1176343953