Ann. Inst. Statist. Math.
Vol. 40, No. 4, 785-797 (1988)

TRIPLE STAGE POINT ESTIMATION FOR THE EXPONENTIAL LOCATION PARAMETER

H. I. HAMDY[1], N. MUKHOPADHYAY[2], M. C. COSTANZA[3] AND M. S. SON[1]

[1]*Department of Mathematics and Statistics, University of Vermont, 16 Colchester Avenue, Burlington, VT 05405, U.S.A.*
[2]*Department of Statistics, The College of Liberal Arts and Sciences, The University of Connecticut, 196 Auditorium Road, Storrs, CT 06268, U.S.A.*
[3]*University of Vermont, Medical Biostatistics, College of Medicine, Given Building, Burlington, VT 05405, U.S.A.*

(Received August 19, 1987; revised January 5, 1988)

Abstract. This paper deals with the problem of estimating the minimum lifetime (guarantee time) of the two parameter exponential distribution through a three-stage sampling procedure. Several forms of loss functions are considered. The regret associated with each loss function is determined. The results in this paper generalize the basic results of Hall (1981, *Ann. Statist.*, **9**, 1229–1238).

Key words and phrases: Two parameter (negative) exponential distribution, minimum lifetime, three-stage sampling, regret, Taylor expansion, uniform integrability, minimum risk, bounded risk.

1. Introduction

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables with the two parameter exponential probability density function (p.d.f.)

$$g(x;\mu,\sigma)=\sigma^{-1}\exp\{-(x-\mu)\sigma^{-1}\}I\{x>\mu\}, \qquad \sigma>0\,, \tag{1.1}$$

where $I\{\cdot\}$ is the indicator function. We assume that both the location parameter μ and the scale parameter σ are unknown. The exponential distribution provides a useful model for data from reliability and life testing experiments for which the assumptions of a constant failure rate (σ^{-1}) and of a minimum lifetime before which no items fail (μ) are reasonable (e.g., Grubbs (1971)). Our emphasis is on point estimation of μ. For example, a good estimate of μ would be crucial to a manufacturer of electronic components who wishes to offer a minimum warranty period for the lifetime of items produced as part of a cost-effective marketing strategy.

Based on a random sample $X_1, X_2, \ldots, X_n$ $(n \geq 2)$ from the p.d.f. in (1.1), we consider

$$X_{n(1)} = \min\{X_1, \ldots, X_n\} \quad \text{and} \quad \hat{\sigma}_n = (n-1)^{-1} \sum_{i=1}^{n} (X_i - X_{n(1)}),$$

the usual estimators of μ and σ, respectively. We propose to estimate μ by $X_{n(1)}$ under a general loss function, say L_n. However, we impose the condition that the process of controlling the risk (in a sense to be made precise later in specific problems), $E(L_n)$, requires the sample size to be the least integer

$$n \geq \alpha\sigma = n^* \quad \text{(say)}, \tag{1.2}$$

where α is a given positive constant.

Since σ is unknown, the required sample size n^* is indeed unknown. Purely sequential sampling procedures were studied by Basu (1971) and Mukhopadhyay (1974, 1982) to estimate quantities like n^*. Both Hall (1981) and Woodroofe (1985) mentioned that triple stage sampling could be used in point estimation problems. However, Mukhopadhyay (1985) and Mukhopadhyay *et al.* (1987) indeed explored this area. Here, we resort to a three-stage sampling scheme to estimate μ via estimation of n^*. Such a group sampling procedure would appear to provide a feasible framework for life testing mass-produced items such as component parts and, since it is based on only three sampling operations, it has obvious advantages from an implementational standpoint over a one-by-one, purely sequential sampling scheme, as noted by Hall (1981). Moreover, as shown by Mukhopadhyay (1985) and by Mukhopadhyay *et al.* (1987), three-stage sampling procedures are competitive with purely sequential schemes on theoretical grounds as well. Here, we mention that a closely related problem of reducing the number of sampling operations using accelerated sequential schemes was discussed by Hall (1983).

In Section 2 we present a triple stage sampling and point estimation procedure along the lines of Hall (1981), Mukhopadhyay (1985), Mukhopadhyay *et al.* (1987), and Woodroofe (1987). We derive some intermediate results that are useful for determining the positive and negative integer moments of any order of the associated stopping variable. Also, we stress that results like those in our Theorems 2.1 and 2.2 are not available elsewhere in these forms. In Section 3 we give some applications corresponding to explicit forms of the loss functions; in particular, asymptotic expansions of the associated regret functions are provided. We also emphasize that these analyses are simpler and more direct than others currently available. We close with a few remarks concerning the moderate sample size performance of the proposed procedures based on a series of simulation studies.

2. Triple stage sampling procedure

We start the estimation process with a random pilot sample of size k_0 (≥ 2) from the p.d.f. in (1.1), say $X_1, \ldots, X_{k_0}$. We also choose a real number $\gamma \in (0, 1)$, to determine the proportion of n^* which will be estimated from the pilot data. Let $S^* = [\gamma\alpha\hat{\sigma}_{k_0}] + 1$, where $[x]$ denotes the largest integer less than or equal to x, and define

$$(2.1) \qquad N_1^* = \max\{k_0, S^*\}.$$

We terminate the sampling if $k_0 \geq S^*$; otherwise, we obtain the additional observations $X_{k_0+1}, \ldots, X_{S^*}$. Then compute $T^* = [\alpha\hat{\sigma}_{N_1^*}] + 1$. The final sample size is determined by

$$(2.2) \qquad N^* = \max\{N_1^*, T^*\}.$$

If $T^* > N_1^*$, we obtain $T^* - N_1^*$ more observations, $X_{S^*+1}, \ldots, X_{T^*}$. Otherwise, we stop the sampling at the second stage. When we stop sampling, we propose $X_{N^*(1)}$ as the point estimator of μ.

Determination of the large sample properties of our proposed estimator will eventually require us to obtain Taylor expansions for the moments of N^*. This is greatly simplified by employing the following techniques which permit us to replace the stopping variable N^* by, say, N, which is defined similarly but only in terms of sample averages of positive i.i.d. random variables.

Specifically, let $Y_1, Y_2, \ldots$ be i.i.d. random variables with the p.d.f. $g(y; 0, \sigma)$. Define $\bar{Y}_n = \sum_{i=1}^{n-1} Y_i/(n-1)$ and $S = [\gamma\alpha\bar{Y}_{k_0}] + 1$, and then take

$$(2.3) \qquad N_1 = \max\{k_0, S\},$$

and

$$(2.4) \qquad N = \max\{N_1, T\},$$

where $T = [\alpha\bar{Y}_{N_1}] + 1$. From the results in Lombard and Swanepoel (1978), we conclude that the random variables N_1^* and N_1 (as well as N^* and N) are identically distributed. From now on we will use the expressions in (2.3) and (2.4) instead of those in (2.1) and (2.2), respectively.

Throughout the following sections we assume that

(2.5) $$n^* = O(k_0^r), \qquad r > 1 ,$$

and that

(2.6) $$\limsup (k_0/n^*) < \gamma \quad \text{as} \quad k_0 \to \infty .$$

The main result of this section is given in the following theorem.

THEOREM 2.1. *For every positive integer $m \geq 1$, and as $k_0 \to \infty$, we have*

$$E(\bar{Y}_{N_1}^m) = \sigma^m + \sigma^m m(m-3)/2\gamma n^* + o(\alpha^{-1}) .$$

PROOF. First we take the expectation conditional on $Y_1, Y_2, \ldots, Y_{k_0-1}$. We write

$$\begin{aligned} E(\bar{Y}_{N_1}^m) &= E\{E(\bar{Y}_{N_1}^m | Y_1, \ldots, Y_{k_0-1})\} \\ &= E\left\{ (N_1-1)^{-m} E\left\{ \left(\sum_{i=1}^{k_0-1} Y_i + \sum_{i=k_0}^{N_1-1} Y_i \right)^m \,\middle|\, Y_1, \ldots, Y_{k_0-1} \right\}\right\} \\ &= E\left\{ (N_1-1)^{-m} \sum_{j=0}^{m} \binom{m}{j} \left(\sum_{i=1}^{k_0-1} Y_i \right)^j E\left\{ \left(\sum_{i=k_0}^{N_1-1} Y_i \right)^{m-j} \,\middle|\, Y_1, \ldots, Y_{k_0-1} \right\}\right\} . \end{aligned}$$

Given $Y_1, \ldots, Y_{k_0-1}$, the random sum $\sum_{i=k_0}^{N_1-1} 2\sigma^{-1} Y_i \sim \chi^2$ with $2(N_1 - k_0)$ degrees of freedom. Therefore,

$$E\left\{ \left(\sum_{i=k_0}^{N_1-1} Y_i \right)^{m-j} \,\middle|\, Y_1, \ldots, Y_{k_0-1} \right\} = \sigma^{m-j}(N_1 - k_0)^{m-j}\{1 + O(N_1^{-1})\} ,$$

for large k_0, using factorial formulas. After some algebra based on repeated use of the binomial formula, we get

(2.7) $$E(\bar{Y}_{N_1}^m) = \sigma^m \sum_{j=0}^{m} \binom{m}{j} E\left\{ \sum_{i=1}^{k_0-1} W_i \Big/ (N_1 - 1) \right\}^j + o(\alpha^{-1}) ,$$

where $W_i = (Y_i - \sigma)\sigma^{-1}$ are i.i.d. random variables with $E(W_i) = 0$ and $V(W_i) = 1$ for all i.

Now, expanding $(N_1 - 1)^{-j}$ as $k_0 \to \infty$ in a Taylor series around γn^*, we write

(2.8) $$(N_1 - 1)^{-j} = (\gamma n^*)^{-j} - j(\gamma\alpha\bar{Y}_{k_0} - \gamma n^*)(\gamma n^*)^{-(j+1)} + R(n^*) .$$

Here, we considered $N_1 \approx \gamma\alpha\bar{Y}_{k_0}$ for large k_0. However, even for moderate values of k_0, the expected difference between N_1 and its continuous version, (say) $N_{1c} = \max\{k_0, \gamma\alpha\bar{Y}_{k_0}\}$, is less than 0.58 observation. In order to see that, we write

$$\begin{aligned} N_1 &= ([\gamma\alpha\bar{Y}_{k_0}] + 1) I\{\gamma\alpha\bar{Y}_{k_0} > k_0\} + k_0 I\{\gamma\alpha\bar{Y}_{k_0} \le k_0\} \\ &= \{\gamma\alpha\bar{Y}_{k_0} + 1 - (\gamma\alpha\bar{Y}_{k_0} - [\gamma\alpha\bar{Y}_{k_0}])\} I\{\gamma\alpha\bar{Y}_{k_0} > k_0\} + k_0 I\{\gamma\alpha\bar{Y}_{k_0} \le k_0\} \\ &= \gamma\alpha\bar{Y}_{k_0} I\{\gamma\alpha\bar{Y}_{k_0} > k_0\} + k_0 I\{\gamma\alpha\bar{Y}_{k_0} \le k_0\} + \beta I\{\gamma\alpha\bar{Y}_{k_0} > k_0\}\,, \end{aligned}$$

where $\beta = 1 - (\gamma\alpha\bar{Y}_{k_0} - [\gamma\alpha\bar{Y}_{k_0}]) \sim U(0,1)$. Now $N_1 - N_{1c} = \beta I\{\gamma\alpha\bar{Y}_{k_0} > k_0\}$, and so

$$E(N_1 - N_{1c}) \le \{E(\beta^2) P(\gamma\alpha\bar{Y}_{k_0} > k_0)\}^{1/2} \le 0.58\,,$$

where we have used the Cauchy-Schwarz inequality.

Next, substitution of the standardized i.i.d. W_i's into (2.8) leads to

$$(N_1 - 1)^{-j} = (\gamma n^*)^{-j} - j\left(\sum_{i=1}^{k_0-1} W_i\right)\Big/(\gamma n^*)^j(k_0 - 1) + R(n^*)\,.$$

Hence, substitution into (2.7) yields

$$\begin{aligned} E(\bar{Y}_{N_1}^m) &= \sigma^m \sum_{j=0}^{m} \binom{m}{j} (\gamma n^*)^{-j} E\left(\sum_{i=1}^{k_0-1} W_i\right)^j \\ &\quad - \{\sigma^m/(k_0 - 1)\} \sum_{j=0}^{m} \binom{m}{j} (\gamma n^*)^{-j} j E\left(\sum_{i=1}^{k_0-1} W_i\right)^{j+1} + o(\alpha^{-1})\,, \end{aligned}$$

since $E\left\{\left(\sum_{i=1}^{k_0-1} W_i\right)^j R(n^*)\right\} = o(\alpha^{-1})$. Finally, we have

$$E(\bar{Y}_{N_1}^m) = \sigma^m + \sigma^m m(m-1)(k_0 - 1)/2(\gamma n^*)^2 - m\sigma^m/\gamma n^* + o(\alpha^{-1})\,.$$

Now use the fact that $(k_0/n^*) \approx \gamma$ by (2.6), combine terms, and Theorem 2.1 is immediate.

Remark 1. Similar results for $m = 1$ in the normal case were obtained by Hall (1981), Mukhopadhyay (1985) and Mukhopadhyay *et al.* (1987). We are unaware of any such expansion of $E(\bar{Y}_{N_1}^m)$ for $m > 1$. Theorem 2.1 will be useful in computing the moments of the stopping variable N (or N^*) as we shall see in Theorem 2.2 which follows. The following Lemma 2.1 is used in the proof of Theorem 2.2 in Appendix.

LEMMA 2.1. *For the three-stage procedure* (2.4), *as* $\alpha \to \infty$, *we have*

$$P(N \leq \xi n^*) = O(\alpha^{-m^*/r}) ,$$

for every positive integer m^* *and real number* $\xi \in (0, 1)$.

PROOF. We consider the events $\{N \leq \xi n^*\} \subseteq \{\alpha \bar{Y}_{N_1} \leq \xi n^*\}$, and thus

$$\begin{aligned} P(N \leq \xi n^*) &\leq P(\alpha \bar{Y}_{N_1} \leq \xi n^*) \\ &\leq P(|\bar{Y}_{N_1} - \sigma| > (1 - \xi)\sigma) \\ &\leq P\left(\max_{k_0 \leq n \leq [\xi n^*]} |\bar{Y}_n - \sigma| > (1 - \xi)\sigma \right) \\ &= O(k_0^{-m^*}) , \end{aligned}$$

by the Hajek-Renyi inequality. Hence, (2.5) gives the statement of Lemma 2.1.

THEOREM 2.2. *For* N *defined in* (2.4) *and integer* k, *as* $\alpha \to \infty$, *we have*

(i) $E(N/n^*)^k = 1 + k(k - 3 + \gamma)/2\gamma n^* + o(\alpha^{-1})$,

(ii) $E(\ln N) = (\ln n^*) - (3 - \gamma)/2\gamma n^* + o(\alpha^{-1})$.

The proof of Theorem 2.2 is given in Appendix.

3. Some applications

In this section we consider some typical situations in which the process of controlling the risk associated with our proposed triple stage procedure leads to the same form of n^* given in (1.2). As in Robbins (1959), we define the regret of the three-stage procedure as $\omega(\alpha) = E(L_N(\alpha)) - R_{n^*}(\alpha)$, where $R_{n^*}(\alpha)$ is the optimal risk corresponding to the fixed sample size procedure based on n^* in (1.2).

3.1 *The bounded risk problem*

Let the loss function associated with the estimation process be given by

$$L_n(A) = A(X_{n(1)} - \mu)^t , \tag{3.1}$$

where A is a known positive constant and t is a positive integer. Given W (> 0), which is related to the preassigned overhead cost, we require $E(L_n(A)) \leq W$, which provides

$$n^* = \alpha\sigma, \quad \text{where} \quad \alpha = (A\Gamma(t+1)/W)^{1/t},$$

($\Gamma(\cdot)$ is the gamma function). The associated optimal risk is then $R_{n^*}(A) = W$.

THEOREM 3.1. *In the context of the loss function* (3.1), *the regret of the triple stage procedure* (2.4) *is given by*

$$\omega(A) = Wt(t+3-\gamma)/2\gamma n^* + o(A^{-1/t}) \quad \textit{as} \quad A \to \infty .$$

PROOF. The risk of the triple stage procedure is

$$E(L_N(A)) = A\sigma^t\Gamma(t+1)n^{*-t}E(n^*/N)^t ,$$

and by (i) of Theorem 2.2 with $k = -t$ the proof is completed.

3.2 *The minimum risk problem*

Here we consider two examples with different loss functions. The estimation cost is assumed to be similar in both cases, while the sampling costs are different.

First, suppose the loss incurred in estimating μ by $X_{n(1)}$ takes the form

$$L_n(c) = A\sigma^m(X_{n(1)} - \mu)^t + cn^m , \tag{3.2}$$

where A and c are known positive constants and m and t are positive integers. A loss function similar to (3.2) was studied by Chow and Martinsek (1982) for the normal case. The risk associated with (3.2) is

$$E(L_n(c)) = (A\sigma^{m+t}\Gamma(t+1)/n^t) + cn^m . \tag{3.3}$$

Treating n as a continuous variable, we differentiate (3.3) with respect to n to obtain the optimal sample size

$$n^* = \alpha\sigma, \quad \text{where} \quad \alpha = (tA\Gamma(t+1)/cm)^{1/(m+t)} .$$

The optimal risk is therefore $R_{n^*}(c) = cn^{*m}(m+t)/t$.

THEOREM 3.2. *In the context of the loss function* (3.2), *the regret of the triple stage procedure* (2.4) *is given by*

$$\omega(c) = cmn^{*m-1}(t+m)/2\gamma + o(c^{(t+1)/(m+t)}) \quad \textit{as} \quad c \to 0 .$$

PROOF. If we write

Table 1. Moderate sample size performance of triple stage procedure: Bounded risk[+].

		$\gamma = 0.30$					$\gamma = 0.50$					$\gamma = 0.70$				
k_0	n^*	$\hat{\mu}$	$\hat{se}_{\hat{\mu}}$	$\bar{N}$	$\hat{se}_{\bar{N}}$	$\hat{\omega}$	$\hat{\mu}$	$\hat{se}_{\hat{\mu}}$	$\bar{N}$	$\hat{se}_{\bar{N}}$	$\hat{\omega}$	$\hat{\mu}$	$\hat{se}_{\hat{\mu}}$	$\bar{N}$	$\hat{se}_{\bar{N}}$	$\hat{\omega}$
	10	.183	.006	6	0.1	2.58	.134	.005	9	0.1	1.52	.137	.005	9	0.1	0.95
	25	.085	.004	21	0.3	6.71	.059	.003	23	0.3	2.33	.044	.002	25	0.3	0.95
	50	.029	.002	47	0.5	4.22	.024	.001	47	0.4	1.72	.021	.001	50	0.4	0.60
	100	.012	.001	97	0.7	1.94	.011	.000	99	0.6	0.64	.011	.000	103	0.7	0.27
5	150	.008	.000	146	0.9	2.01	.007	.000	149	0.6	0.08	.007	.000	157	1.0	−.00
	200	.005	.000	197	0.9	0.16	.005	.000	200	0.8	1.65	.005	.000	209	1.2	0.03
	250	.004	.000	246	1.0	0.10	.004	.000	249	0.9	0.05	.004	.000	262	1.4	−.03
	500	.002	.000	496	1.5	0.05	.002	.000	502	1.3	0.01	.002	.000	522	2.6	−.04
	1000	.001	.000	1002	2.1	0.01	.001	.000	1007	2.6	0.00	.001	.000	1054	5.4	−.06
	10	.100	.003	10	0.0	0.00	.096	.003	10	0.0	−.01	.098	.003	11	0.1	−.06
	25	.086	.003	14	0.3	4.32	.056	.002	22	0.3	1.61	.045	.002	24	0.2	0.57
	50	.031	.002	45	0.5	3.44	.022	.001	48	0.3	0.58	.021	.001	49	0.3	0.25
	100	.011	.000	97	0.6	0.38	.010	.000	99	0.5	0.10	.010	.000	101	0.4	0.05
10	150	.007	.000	147	0.8	0.16	.006	.000	148	0.6	0.11	.007	.000	151	0.6	0.03
	200	.005	.000	196	0.9	0.14	.005	.000	200	0.7	0.04	.005	.000	203	0.7	0.01
	250	.004	.000	248	0.9	0.06	.004	.000	248	0.8	0.05	.004	.000	252	0.7	0.01
	500	.002	.000	496	1.4	0.04	.002	.000	498	1.1	0.02	.002	.000	507	1.4	−.01
	1000	.001	.000	998	1.9	0.02	.001	.000	1000	1.5	0.01	.001	.000	1014	2.3	−.02
	10	.048	.002	20	0.0	−.75	.055	.002	20	0.0	−.75	.051	.002	20	0.0	−.75
	25	.047	.002	20	0.0	0.56	.047	.002	20	0.1	0.55	.046	.002	23	0.2	0.33
	50	.047	.002	25	0.5	4.73	.026	.001	45	0.5	1.19	.022	.001	49	0.3	0.21
	100	.013	.001	95	0.8	1.28	.011	.000	98	0.4	0.12	.011	.000	99	0.4	0.07
20	150	.007	.000	146	0.7	0.19	.007	.000	149	0.6	0.07	.006	.000	150	0.5	0.04
	200	.005	.000	196	0.8	0.10	.005	.000	198	0.7	0.06	.005	.000	200	0.6	0.03
	250	.004	.000	248	1.0	0.08	.004	.000	247	0.7	0.05	.004	.000	249	0.7	0.03
	500	.002	.000	496	1.3	0.04	.002	.000	500	1.0	0.01	.002	.000	501	0.9	0.01
	1000	.001	.000	997	1.9	0.02	.001	.000	1001	1.5	0.01	.001	.000	999	1.4	0.01

[+]Loss function (3.1) with $t = 2$, $\mu = 0$, $\sigma = 1$, $W = 1$. Each entry based on 1000 simulations.

$$E(L_N(c)) = A\Gamma(t+1)\sigma^{m+t} n^{*-t} E(n^*/N)^t + n^{*m} c E(N/n^*)^m ,$$

and recall (i) of Theorem 2.2 with $k = -t$, and $k = m$, then the statement of Theorem 3.2 is immediate.

The second example assumes the loss function

$$L_n(c) = A(X_{n(1)} - \mu)^t + c \ln(n) . \tag{3.4}$$

Thus the cost of sampling increases at a slower rate than n, which is likely when bulk sampling rates apply. The optimal sample size is

$$n^* = \alpha\sigma, \qquad \text{where} \qquad \alpha = (tA\Gamma(t+1)/c)^{1/t} ,$$

and the corresponding optimal risk is $R_{n^*}(c) = c \ln(n^{*t} e)$.

Table 2. Moderate sample size performance of triple stage procedure: Minimum risk[††].

		$\gamma = 0.30$					$\gamma = 0.50$					$\gamma = 0.70$				
k_0	n^*	$\hat{\mu}$	$\hat{se}_{\hat{\mu}}$	$\bar{N}$	$\hat{se}_{\bar{N}}$	$\hat{\omega}$	$\hat{\mu}$	$\hat{se}_{\hat{\mu}}$	$\bar{N}$	$\hat{se}_{\bar{N}}$	$\hat{\omega}$	$\hat{\mu}$	$\hat{se}_{\hat{\mu}}$	$\bar{N}$	$\hat{se}_{\bar{N}}$	$\hat{\omega}$
	10	.189	.006	6	0.1	0.04	.154	.006	8	0.1	0.03	.134	.005	9	0.1	0.02
	25	.078	.004	21	0.3	0.02	.055	.002	23	0.3	0.01	.052	.002	24	0.3	0.00
	50	.034	.002	46	0.5	0.00	.024	.001	48	0.4	0.00	.021	.001	51	0.4	0.00
	100	.013	.001	95	0.7	0.00	.011	.000	97	0.6	0.00	.010	.000	103	0.7	0.00
5	150	.007	.000	146	0.9	0.00	.007	.000	149	0.6	0.00	.006	.000	156	0.9	0.00
	200	.005	.000	196	1.0	0.00	.005	.000	198	0.8	0.00	.005	.000	207	1.1	0.00
	250	.004	.000	245	1.1	0.00	.004	.000	251	1.0	0.00	.004	.000	262	1.4	0.00
	500	.002	.000	497	1.5	0.00	.002	.000	501	1.4	0.00	.002	.000	517	2.2	0.00
	1000	.001	.000	997	2.1	0.00	.001	.000	1005	2.3	0.00	.001	.000	1056	5.3	0.00
	10	.097	.003	10	0.0	0.00	.104	.003	10	0.0	0.00	.097	.003	11	0.1	0.00
	25	.083	.003	15	0.3	0.01	.054	.002	23	0.3	0.00	.050	.002	24	0.2	0.00
	50	.035	.002	45	0.5	0.00	.023	.001	48	0.3	0.00	.021	.001	49	0.3	0.00
	100	.011	.000	96	0.6	0.00	.010	.000	98	0.5	0.00	.010	.000	100	0.4	0.00
10	150	.007	.000	146	0.8	0.00	.007	.000	148	0.6	0.00	.007	.000	152	0.6	0.00
	200	.005	.000	197	0.9	0.00	.005	.000	198	0.7	0.00	.005	.000	202	0.7	0.00
	250	.004	.000	247	0.9	0.00	.004	.000	248	0.7	0.00	.004	.000	252	0.8	0.00
	500	.002	.000	498	1.4	0.00	.002	.000	499	1.1	0.00	.002	.000	507	1.3	0.00
	1000	.001	.000	996	1.9	0.00	.001	.000	997	1.5	0.00	.001	.000	1014	2.4	0.00
	10	.050	.002	20	0.0	0.03	.050	.002	20	0.0	0.03	.050	.002	20	0.0	0.03
	25	.050	.002	20	0.0	0.00	.049	.002	20	0.1	0.00	.046	.002	23	0.2	0.00
	50	.045	.002	24	0.4	0.00	.027	.001	46	0.5	0.00	.021	.001	49	0.3	0.00
	100	.013	.001	95	0.8	0.00	.010	.000	98	0.5	0.00	.010	.000	100	0.4	0.00
20	150	.007	.000	146	0.7	0.00	.007	.000	149	0.6	0.00	.007	.000	149	0.5	0.00
	200	.005	.000	197	0.9	0.00	.005	.000	197	0.7	0.00	.005	.000	200	0.6	0.00
	250	.004	.000	247	1.0	0.00	.004	.000	248	0.7	0.00	.004	.000	250	0.6	0.00
	500	.002	.000	497	1.4	0.00	.002	.000	500	1.0	0.00	.002	.000	500	0.9	0.00
	1000	.001	.000	998	1.9	0.00	.001	.000	999	1.5	0.00	.001	.000	999	1.3	0.00

[††]Loss function (3.2) with $t = 2$, $m = 1$, $\mu = 0$, $\sigma = 1$, $A = 1$. Each entry based on 1000 simulations.

THEOREM 3.3. *In the context of the loss function* (3.4), *the regret of the triple stage procedure* (2.4) *is given by*

$$\omega(c) = ct/2\gamma n^* + o(c^{(t+1)/t}) \quad as \quad c \to 0 .$$

PROOF. The proof is based on (ii) of Theorem 2.2 and is similar to those of Theorems 3.1 and 3.2. Further details are omitted.

4. The performance of the triple stage procedures

Since the results in Sections 2 and 3 are asymptotic in nature, it is of interest to examine the performance of the triple stage procedures for the case of moderate n^*. We conducted simulations with $k_0 = 5$, 10, 20 and $n^* = 10$, 25, 50(50)250, 500, 1000 under squared error loss (i.e., $t = 2$) in (3.1), (3.2) and (3.4) with $W = 1$ in (3.1) and $m = 1$ in (3.2) and $A = 1$ in both (3.2) and (3.4). We set $\mu = 0$ and $\sigma = 1$ in all cases and did 1000 repetitions in each study. We considered the values of $\gamma = 0.3$, 0.5 and 0.7. We note that the numerical results for the loss function in (3.4) had very similar patterns to those for the loss function in (3.2). For brevity, however, we report numerical findings only for the loss functions in (3.1) and (3.2).

Detailed findings from the numerical studies are presented in Tables 1 and 2. For each row of our tables we give summary measures including the estimates $\hat{\mu}$ of μ and its estimated standard error $\hat{se}_{\hat{\mu}}$, the estimate $\bar{N}$ of N and its estimated standard error $\hat{se}_{\bar{N}}$, and the observed regret $\hat{\omega}$ associated with the triple sampling procedure.

Overall, for $\gamma < 1/2$ the three-stage procedure tends to undersample for the smaller values of n^*, which generally results in bad estimates of μ. On the contrary, for $\gamma > 1/2$ substantial oversampling for the larger values of n^* occurs. It appears that $k_0 = 10$ or 20 represents a reasonable pilot sample size in applications. The numerical results agree well with the large sample theory; the approach to the asymptotics occurs remarkably quickly and the goal of controlling the risk associated with each loss function is always met. For practical implementations we strongly recommend using the triple sampling procedure with $\gamma = 1/2$. This recommendation essentially follows from general patterns in Tables 1 and 2.

Acknowledgements

The authors are thankful to the referees for their comments, which led to considerable improvement of an earlier version. This work was supported in part by a Vermont EPSCoR grant from the National Science Foundation (HIH, MSS), by contract #N-00014-85-K-0548 with the Office of Naval Research (NM), and by grant #P30-CA22435 from the National Cancer Institute (MCC).

Appendix

PROOF OF THEOREM 2.2. To prove (i), first consider the case $k = m$, say where m is a positive integer. We notice that $N = T$ except perhaps on a set

$$\psi = \{S < k_0\} \cup \{\alpha \bar{Y}_{N_1} < \gamma \alpha \bar{Y}_{k_0} + 1\},$$

where

$$\int_{\psi} N^m dP = o(\alpha^{m-1}).$$

Therefore,

$$\begin{aligned} E(N^m) &= E(T^m) + o(\alpha^{m-1}) \\ &= E([\alpha\bar{Y}_{N_1}] + 1)^m + o(\alpha^{m-1}) \\ &= E(\alpha\bar{Y}_{N_1} - (\alpha\bar{Y}_{N_1} - [\alpha\bar{Y}_{N_1}]) + 1)^m + o(\alpha^{m-1}) . \end{aligned}$$

It follows from Hall (1981) that $\delta_{N_1} = \alpha\bar{Y}_{N_1} - [\alpha\bar{Y}_{N_1}]$ is asymptotically $\sim U(0,1)$. Let $\beta_{N_1} = 1 - \delta_{N_1}$, also asymptotically $\sim U(0,1)$. Thus,

$$\begin{aligned} E(N^m) &= E(\alpha\bar{Y}_{N_1} + \beta_{N_1})^m + o(\alpha^{m-1}) \\ &= \sum_{j=0}^{m} \binom{m}{j} E\{(\alpha\bar{Y}_{N_1})^{m-j}\beta_{N_1}^j\} + o(\alpha^{m-1}) \end{aligned}$$

$$= \alpha^m E(\bar{Y}_{N_1}^m) + m\alpha^{m-1}E(\bar{Y}_{N_1}^{m-1}\beta_{N_1}) + o(\alpha^{m-1}) . \tag{A.1}$$

To estimate the quantity $E(\bar{Y}_{N_1}^{m-1}\beta_{N_1})$, we write $\beta_{N_1} = 1/2 + (\beta_{N_1} - 1/2)$, and then

$$\begin{aligned} E(\bar{Y}_{N_1}^{m-1}\beta_{N_1}) &= 1/2 \cdot E(\bar{Y}_{N_1}^{m-1}) + E\{\bar{Y}_{N_1}^{m-1}(\beta_{N_1} - 1/2)\} \\ &= 1/2 \cdot \alpha^{m-1} + E\{\bar{Y}_{N_1}^{m-1}(\beta_{N_1} - 1/2)\} . \end{aligned}$$

To evaluate the expression $E(\{\bar{Y}_{N_1}^{m-1}(\beta_{N_1} - 1/2)\}$, first we recall Theorem 2.1 to obtain

$$\operatorname{Var}(\bar{Y}_{N_1}^{m-1}) = (m-1)^2\sigma^{2(m-1)}/2\gamma n^* + o(\alpha^{-1}) . \tag{A.2}$$

On the other hand, the well known Cauchy-Schwarz inequality provides

$$\operatorname{Cov}^2(\bar{Y}_{N_1}^{m-1}, (\beta_{N_1} - 1/2)) \le \operatorname{Var}(\bar{Y}_{N_1}^{m-1}) \operatorname{Var}(\beta_{N_1}) = o(1) ,$$

by (A.2) as $n^* \to \infty$. It follows that $\bar{Y}_{N_1}^{m-1}$ and $(\beta_{N_1} - 1/2)$ are asymptotically uncorrelated. Thus,

$$E\{\bar{Y}_{N_1}^{m-1}(\beta_{N_1} - 1/2)\} = o(1) .$$

Hence,

$$E(N^m) = \alpha^m E(\bar{Y}_{N_1}^m) + 1/2 \cdot m\sigma^{m-1}\alpha^{m-1} + o(\alpha^{m-1}) ,$$

and by Theorem 2.1, (A.2) yields

$$E(N^m) = n^{*m} + 1/2 \cdot mn^{*m}(m - 3 + \gamma)/\gamma n^* + o(\alpha^{m-1}) . \tag{A.3}$$

Next, if k is a negative integer, say $k = -m$, where m is a positive integer, we need to prove that $E(n^*/N)^m = 1 + m(m+3-\gamma\ )/2\gamma n^* + o(\alpha^{-1})$. Write

$$N^{-m} = n^{*-m} - m(N-n^*)n^{*-(m+1)} + 1/2 \cdot m(m+1)(N-n^*)^2 v^{-(m+2)} ,$$

for a suitable random variable v between N and n^*. Consequently, (A.3) with $m = 1$ gives

$$E(N^{-m}) = n^{*-m} - m(1/2 - \gamma + o(1))n^{*-(m+1)} \\ + 1/2 \cdot m(m+1)n^{*-(m+1)}E(Q) ,$$

where $Q = n^{*m+1}(N-n^*)^2 v^{-(m+2)}$. Let $U = (N-n^*)^2/n^*$. We conclude from the results of Anscombe (1952) (see also Ghosh and Mukhopadhyay (1975)) that $U \xrightarrow{L} \gamma^{-1}\chi^2_{(1)}$. Also, $E(U) = \gamma^{-1} + o(1)$, which can be obtained from (A.3) with $m = 1$ and 2, hence U is uniformly integrable.

We write

$$Q = QI\{N > 1/2 \cdot n^*\} + QI\{N \leq 1/2 \cdot n^*\} . \tag{A.4}$$

Now,

$$QI\{N > 1/2 \cdot n^*\} \leq 2^{m+2}(N-n^*)^2/n^* ,$$

and hence $QI\{N > 1/2 \cdot n^*\}$ is uniformly integrable. Thus,

$$E(QI\{N > 1/2 \cdot n^*\}) = \gamma^{-1} + o(1) .$$

Also,

$$QI\{N \leq 1/2 \cdot n^*\} \leq (n^{*m+1}/N^m)(1 + n^{*2}/N^2)I\{N \leq 1/2 \cdot n^*\} \\ \leq (n^{*m+1}/k_0^m)(1 + n^{*2}/k_0^2)I\{N \leq 1/2 \cdot n^*\} .$$

Thus,

$$E(QI\{N \leq 1/2 \cdot n^*\}) \leq (n^{*m+1}/k_0^m)(1 + n^{*2}/k_0^2)P(N \leq 1/2 \cdot n^*) ,$$

and by Lemma 2.1 this can be made $o(1)$ if we make m^* large enough. Hence,

$$E(QI\{N \leq 1/2 \cdot n^*\}) = o(1) .$$

So, (A.4) leads to

$$E(Q) = \gamma^{-1} + o(1) ,$$

which completes the proof of part (i).

To prove (ii) of Theorem 2.2, first write ($\ln n$) in Taylor series expansion

$$E \ln (N) = L_n(n^*) + E(N - n^*)n^{*-1} - n^{*-1}E(Q_1) ,$$

where $Q_1 = n^*(N - n^*)^2\zeta^{-2}$ for a suitable random variable ζ between N and n^*. Then $E(Q_1) = \gamma^{-1} + o(1)$ by arguments similar to those leading to (A.1). Further arguments similar to those above lead to (ii) of Theorem 2.2, and the proof of Theorem 2.2 is now complete.

REFERENCES

Anscombe, F. J. (1952). Large sample theory of sequential estimation, *Proc. Cambridge Philos. Soc.*, **48**, 600–607.

Basu, A. P. (1971). On a sequential rule for estimating the location parameter of an exponential distribution, *Naval Res. Logist. Quart.*, **18**, 329–337.

Chow, Y. S. and Martinsek, A. T. (1982). Bounded regret of a sequential procedure for estimation of the mean, *Ann. Statist.*, **10**, 909–914.

Ghosh, M. and Mukhopadhyay, N. (1975). Asymptotic normality of stopping times in sequential analysis (unpublished manuscript).

Grubbs, F. E. (1971). Approximate fiducial bounds on reliability for the two parameter negative exponential distribution, *Technometrics*, **13**, 873–876.

Hall, P. (1981). Asymptotic theory of triple sampling for sequential estimation of a mean, *Ann. Statist.*, **9**, 1229–1238.

Hall, P. (1983). Sequential estimation saving sampling operations, *J. Roy. Statist. Soc. Ser. B*, **45**, 219–223.

Lombard, F. and Swanepoel, J. W. H. (1978). On finite and infinite confidence sequences, *South African. Statist. J.*, **12**, 1–24.

Mukhopadhyay, N. (1974). Sequential estimation of location parameter in exponential distributions, *Calcutta Statist. Assoc. Bull.*, **23**, 85–95.

Mukhopadhyay, N. (1982). On the asymptotic regret while estimating the location parameter of an exponential distribution, *Calcutta Statist. Assoc. Bull.*, **31**, 207–213.

Mukhopadhyay, N. (1985). A note on three-stage and sequential point estimation procedures for a normal mean, *Sequential Anal.*, **4**, 311–319.

Mukhopadhyay, N., Hamdy, H. I., Al-Mahmeed, M. and Costanza, M. C. (1987). Three-stage point estimation procedures for a normal mean, *Sequential Anal.*, **6**, 21–36.

Robbins, H. (1959). Sequential estimation of the mean of a normal population, *Probability and Statistics*, (ed. H. Cramer), 235–245, Almquist and Wiksell, Uppsala.

Woodroofe, M. (1985). Asymptotic local minimaxity in sequential point estimation, *Ann. Statist.*, **13**, 676–688.

Woodroofe, M. (1987). Asymptotically optimal sequential point estimation in three stages, *New Perspectives in Theoretical and Applied Statistics*, 397–411, Wiley, New York.