International Journal of Statistics and Applied Mathematics 2017; 2(6): 241-249

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2017; 2(6): 241-249

www.mathsjournal.com
Received: 01-09-2017
Accepted: 02-10-2017

K Sivasakthi
Department of Statistics, Tagore College of Arts and Science, Chennai, Tamil Nadu, India

R Sakthivel
Department of Statistics, Presidency College, Chennai, Tamil Nadu, India

Martin L William
Department of Statistics, Loyola College, Chennai, Tamil Nadu, India

Correspondence
K Sivasakthi
Department of Statistics, Tagore College of Arts and Science, Chennai, Tamil Nadu, India

Preference for a class of super-efficient estimators of the normal mean: A study on sample size requirement

K Sivasakthi, R Sakthivel and Martin L William

Abstract
A class of super-efficient estimators of the mean of a normal population with unit variance has been recently constructed by Sivasakthi *et al.* (2016) through the 'Delta Method'. Theoretically, a super-efficient estimator is preferable to the asymptotically efficient estimator (could be the maximum likelihood estimator) in a large-sample context. In this paper, we address the super-efficient estimation of the normal mean when the population variance is known. The important question on the sample size required for a super-efficient estimator to be preferred over the (asymptotically) efficient estimator / maximum likelihood estimator is addressed through a numerical study. The answer to the question is sought for a chosen subset of the class of super-efficient estimators under consideration.

Keywords: Maximum Likelihood Estimators, Super-efficiency

Introduction

The notion of 'super-efficient estimator' flowed out of an unpublished work of Hodges (1951) [4] which was followed up with theoretical developments by Basu (1952) [2], Le Cam (1953, 1956, 1960, 1972) [5, 6, 7, 8] and Stein (1956) [11]. There has not been much development on this topic ever since, except for a few sporadic research articles by Bahadur (1983) [1], Sethuraman (2004) [9] and Durairajan (2012) [3]. Even though the above-mentioned authors addressed the problem of obtaining super-efficient estimators in different contexts, there were no systematic approaches taken towards constructing super-efficient estimators. Recently, Sivasakthi *et. al.* (2017) [10] addressed this problem in the context of estimating a real parameter through the 'Delta Method' of asymptotic inference theory and discussed a number of illustrative examples.

The property of super-efficiency is asymptotic and a super-efficient estimator (SEE) compares favourably with the asymptotically efficient estimator for adequately large samples. The question that naturally arises in this context is about the sample size that would be required to prefer a super-efficient estimator over the efficient estimator. It is found that there has been no attempt till now to address this question. Algebraic closed-form expressions are not available for the 'Mean-Square Error' or 'Variance' of the super-efficient estimators and, so comparisons are to be done numerically. This paper is an exploration in the above direction.

In the recent work of Sivasakthi *et al.* (2017) [10], the super-efficient estimation of the mean of a normal population with unit variance, N(θ, 1), was considered as one of the applications of the 'Delta Method' of deriving super-efficient estimators. For this situation, a class of super-efficient estimators was constructed by Sivasakthi *et. al.* (2017) [10] with super-efficiency at θ= 0. But, the question on the sample size required for the SEE to overtake the MLE remains unanswered.

In this paper, the important question on sample size needed to prefer the SEE over the MLE is addressed by a numerical study. As SEE's are only asymptotically unbiased, the comparison is based on mean-square errors instead of variances of the estimators. For larger sample sizes, however, the mean-square error is approximately same as variance owing to asymptotic unbiasedness.

This paper contains a total of five sections including the present introduction section. Section 2 reviews the class of super-efficient estimators under consideration for estimating the mean of a normal distribution with unit variance.

Section 3 provides the theoretical results concerning the mean-square error of the SEE. Section 4, presents the results of the numerical work which have been carried out using 'R' language. Finally, Section 5 contains concluding remarks with recommendations on the sample sizes needed for the preferential use of the class of SEE's over MLE.

2. A class of super-efficient estimators of the normal mean

As stated earlier, in a recent work, Sivasakthi *et. al.* (2017) [10] applied the 'Delta Method' to obtain super-efficient estimators for the mean of a normal distribution with unit variance, *viz*N(θ, 1), $\theta \in R.$ One class of super-efficient estimators considered in that paper is given by

$$\theta_n^*(k) = \begin{cases} \overline{X}^k & if \ \sqrt{n}\,\left|\overline{X}\right| \le n^{1/4} \\ \overline{X} & if \ \sqrt{n}\,\left|\overline{X}\right| > n^{1/4} \end{cases} \qquad (2.1)$$

where k> 1 is any known positive integer.

This class is generated by the choice of the function g(θ) = θ^k on the parameter space. The estimators in this class are super-efficient at θ= 0 and we refer to Sivasakthi *et. al.* (2017) [10] for other details. Other classes of super-efficient estimators for the same family of normal distributions have also been proposed in that paper. One class is specified by

$$g(\theta) = a_1(\theta-\theta_0)+a_2(\theta-\theta_0)^2+\cdots+a_m(\theta-\theta_0)^m+\theta_0,\ \theta\in\Omega$$

where $\theta_0 \in \Omega$ is a chosen point and $\left|a_1\right|$< 1.This is a generic choice applicable to any distribution, not just N(θ, 1) alone. Yet another class is generated by

$$g(\theta) = c\sin\left(\theta-\theta_0\right)+\theta_0 \quad \text{where } |c|<1$$

But the class of estimators in (2.1) is interesting in that, when the sample mean is closer to zero, which happens with higher likelihood when the true parameter is zero or close to zero, the super-efficient estimator 'improves' the estimate by moving closer to zero by raising it to an integer power 'k'. Thus, there is a rationale in considering the class of estimators in (2.1). The use of a super-efficient estimator from the class instead of the efficient estimator is therefore recommended for adequately large samples. The interesting question of how large the sample should be to prefer a super-efficient estimator is sought to be addressed in the subsequent sections of this paper.

3. Some Theoretical Results

We provide some theoretical results on the class of S.E.E.'s $\theta_n^*(k)$ in this section towards deriving an expression for the Mean Square Error of the estimators.

3.1. Bias of SEE's $\theta_n^*(k)$ in estimating θ

Denote $\overline{X}$ = T and the event A = $[\,|T|>n^{-1/4}\,]$

Consider $E_\theta\left(\theta_n^*(k)\right) = E_\theta\left(T \mid |T|>n^{-1/4}\right)P_\theta\left(|T|>n^{-1/4}\right) + E_\theta\left(T^k \mid |T|\le n^{-1/4}\right)P_\theta\left(|T|\le n^{-1/4}\right)$

$$E_\theta\left(\theta_n^*(k)\right) = a_\theta + b_\theta(k) \qquad (3.1)$$

where, $a_\theta = E_\theta\left(T \mid A\right)P_\theta(A)$, $b_\theta(k) = E_\theta\left(T^k \mid A^c\right)P_\theta\left(A^c\right)$

Noting that T ~ $N[\theta, 1/n]$, we get the conditional pdf of T given the event A as

$$g(t) = \frac{\sqrt{\frac{n}{2\pi}}e^{-n(t-\theta)^2/2}}{P_\theta(A)},\ t\in A$$

So, we get, $a_\theta = E_\theta\left(T \mid A\right)P_\theta(A) = \int_{t\in A}\sqrt{\frac{n}{2\pi}}\ t\ e^{-n(t-\theta)^2/2}dt$

$$= \int_{-\infty}^{\infty}\sqrt{\frac{n}{2\pi}}\ t\,e^{-n(t-\theta)^2/2}dt - \int_{|t|\le n^{-1/4}}\sqrt{\frac{n}{2\pi}}\,t\,e^{-n(t-\theta)^2/2}dt$$

$$= \theta - \sqrt{\frac{n}{2\pi}}\int_{-n^{-1/4}}^{n^{-1/4}} t\,e^{-n(t-\theta)^2/2}dt \qquad (3.2)$$

and $$b_\theta(k) = E_\theta\left(T^k \mid A^c\right)P_\theta\left(A^c\right) = \sqrt{\frac{n}{2\pi}}\int_{-n^{-1/4}}^{n^{-1/4}} t^k\,e^{-n(t-\theta)^2/2}\,dt \qquad (3.3)$$

Using (3.2) and (3.3) in (3.1), we get $E_\theta\left(\theta_n^*(k)\right)$.

Hence, $Bias_\theta\left(\theta_n^*(k)\right) = \sqrt{\frac{n}{2\pi}} \int_{-n^{-1/4}}^{n^{-1/4}} (t^k - t)\, e^{-n(t-\theta)^2/2}\, dt$ (3.4)

3.2: Comparison of SEE's with MLE

Consider $E_\theta\left(\theta_n^{*2}(k)\right) = E_\theta\left(T^2 \mid A\right) P_\theta(A) + E_\theta\left(T^{2k} \mid A^c\right) P_\theta(A^c) = c_\theta + d_\theta(k)$ (3.5)

Here, $c_\theta = E_\theta\left(T^2 \mid A\right) P_\theta(A) = \int_{t\in A} \sqrt{\frac{n}{2\pi}}\, t^2 e^{-n(t-\theta)^2/2} dt$

$$= \int_{-\infty}^{\infty} \sqrt{\frac{n}{2\pi}}\, t^2 e^{-n(t-\theta)^2/2} dt - \int_{-n^{-1/4}}^{n^{-1/4}} \sqrt{\frac{n}{2\pi}}\, t^2 e^{-n(t-\theta)^2/2} dt$$

$$= E_\theta(T^2) - \int_{-n^{-1/4}}^{n^{-1/4}} \sqrt{\frac{n}{2\pi}}\, t^2 e^{-n(t-\theta)^2/2} dt$$

$$= \theta^2 + \frac{1}{n} - \sqrt{\frac{n}{2\pi}} \int_{-n^{-1/4}}^{n^{-1/4}} t^2 e^{-n(t-\theta)^2/2} dt \quad (3.6)$$

and $d_\theta(k) = E_\theta\left(T^{2k} \mid A^c\right) P_\theta(A^c) = \int_{-n^{-1/4}}^{n^{-1/4}} t^{2k} \sqrt{\frac{n}{2\pi}}\, e^{-n(t-\theta)^2/2} dt$ (3.7)

Using (3.6) and (3.7) in (3.5) we get $E_\theta\left(\theta_n^{*2}(k)\right)$

Hence, the mean square error of $\theta_n^*(k)$ is given by

$$MSE_\theta\left(\theta_n^*(k)\right) = E_\theta\left(\theta_n^*(k) - \theta\right)^2$$

$$= \frac{1}{n} - \sqrt{\frac{n}{2\pi}} \int_{-n^{-1/4}}^{n^{-1/4}} (t^{2k} - t^2) e^{-n(t-\theta)^2/2} dt + 2\theta \sqrt{\frac{n}{2\pi}} \int_{-n^{-1/4}}^{n^{-1/4}} (t - t^k) e^{-n(t-\theta)^2/2} dt \quad (3.8)$$

3.3 Gain in preferring the SEE's over the MLE

The MLE of θ is $\overline{X}$ which is efficient with variance 1/n. The gain or loss in preferring $\theta_n^*(k)$ over the MLE $\overline{X}$ in estimating θ is given by

$$Gain_\theta\left(\theta_n^*(k)\right) = Var_\theta\left(\overline{X}\right) - MSE_\theta\left(\theta_n^*(k)\right)$$

$$= \sqrt{\frac{n}{2\pi}} \int_{-n^{-1/4}}^{n^{-1/4}} (t^{2k} - t^2) e^{-n(t-\theta)^2/2} dt - 2\theta \sqrt{\frac{n}{2\pi}} \int_{-n^{-1/4}}^{n^{-1/4}} (t - t^k) e^{-n(t-\theta)^2/2} dt \quad (3.9)$$

4. Numerical Investigation on the performance of S.E.E. over M.L.E.

As seen in the previous section, $Gain_\theta\left(\theta_n^*(k)\right)$ does not have a closed form expression and the process of finding the required sample sizes 'n' cannot be resolved mathematically, but only numerically. Further, even though the class (2.1) itself is infinitely large, we take a small subset of the class by choosing k from 2 to 5 and do the numerical investigation in the sequel.

As pointed out earlier, the comparison of the SEE's $\theta_n^*(k)$ with the Asymptotic Efficient estimator (MLE) $\overline{X}$ is through the 'Gain' function found in (3.9). We seek to know the 'n' values for which the 'Gain Function' is non-negative. Along with this, the Bias in $\theta_n^*(k)$ is also of interest because the estimator is not exactly unbiased and is only asymptotically so. Thus, in this section, we shall present the MSE, Bias and the Gain in $\theta_n^*(k)$.

We have carried out the computations of these measures for a very wide choice of sample sizes, namely, n = 1 (1) 10 (10) 100 (100) 1000 (1000) 10000 (10000) 100000 and for a good range of θ values, namely, θ = –10 (1) –1 (0.1) 1 (1) 10. However, in order to keep the length of this article reasonably moderate, we present the results for k = 2 only with a few selected choices of θ values, namely θ = –10, –5, –1, –0.5, –0.1, 0, 0.5, 1, 3, 6. The sample sizes 'n' required for preferring the SEE $\theta_n^*(2)$, corresponding to the selected θ values are shaded in the Tables 4.1 to 4.10. We then give an abridged Table 4.11 for k = 2, 3, 4, 5 indicating the sample sizes required for preferring the SEE's corresponding to the wider choice of θ values originally included in our study.

It would be specially interesting to observe the performance of $\theta_n^*(k)$ over the MLE for small sample sizes also, especially for values of θ close to zero and values at a good distance away from zero, because the class $\theta_n^*(k)$ itself has been constructed with super-efficiency at θ = 0.

We need to mention here that the results have been rounded off to 5 decimal places as the largest sample size that we have considered is 100000 and the variance of the MLE $\overline{X}$ corresponding to this sample size is 10^{-5}.

θ =−10	Table	4.1	k = 2	θ =−5	Table	4.2	k = 2
N	**MSE**	**Bias**	**Gain**	**n**	**MSE**	**Bias**	**Gain**
1	1.00000	0.00000	0.00000	1	1.00044	0.00004	-0.00045
2	0.50000	0.00000	0.00000	2	0.50000	0.00000	0.00000
3	0.33333	0.00000	0.00000	3	0.33333	0.00000	0.00000
4	0.25000	0.00000	0.00000	4	0.25000	0.00000	0.00000
5	0.20000	0.00000	0.00000	5	0.20000	0.00000	0.00000
6	0.16667	0.00000	0.00000	6	0.16667	0.00000	0.00000
7	0.14286	0.00000	0.00000	7	0.14286	0.00000	0.00000
8	0.12500	0.00000	0.00000	8	0.12500	0.00000	0.00000
9	0.11111	0.00000	0.00000	9	0.11111	0.00000	0.00000
10	0.10000	0.00000	0.00000	10	0.10000	0.00000	0.00000
20	0.05000	0.00000	0.00000	20	0.05000	0.00000	0.00000
30	0.03333	0.00000	0.00000	30	0.03333	0.00000	0.00000
40	0.02500	0.00000	0.00000	40	0.02500	0.00000	0.00000
50	0.02000	0.00000	0.00000	50	0.02000	0.00000	0.00000
60	0.01667	0.00000	0.00000	60	0.01667	0.00000	0.00000
70	0.01429	0.00000	0.00000	70	0.01429	0.00000	0.00000
80	0.01250	0.00000	0.00000	80	0.01250	0.00000	0.00000
90	0.01111	0.00000	0.00000	90	0.01111	0.00000	0.00000
100	0.01000	0.00000	0.00000	100	0.01000	0.00000	0.00000
200	0.00500	0.00000	0.00000	200	0.00500	0.00000	0.00000
300	0.00333	0.00000	0.00000	300	0.00333	0.00000	0.00000
400	0.00250	0.00000	0.00000	400	0.00250	0.00000	0.00000
500	0.00200	0.00000	0.00000	500	0.00200	0.00000	0.00000
600	0.00167	0.00000	0.00000	600	0.00167	0.00000	0.00000
700	0.00143	0.00000	0.00000	700	0.00143	0.00000	0.00000
800	0.00125	0.00000	0.00000	800	0.00125	0.00000	0.00000
900	0.00111	0.00000	0.00000	900	0.00111	0.00000	0.00000
1000	0.00100	0.00000	0.00000	1000	0.00100	0.00000	0.00000
2000	0.00050	0.00000	0.00000	2000	0.00050	0.00000	0.00000
3000	0.00033	0.00000	0.00000	3000	0.00033	0.00000	0.00000
4000	0.00025	0.00000	0.00000	4000	0.00025	0.00000	0.00000
5000	0.00020	0.00000	0.00000	5000	0.00020	0.00000	0.00000
6000	0.00017	0.00000	0.00000	6000	0.00017	0.00000	0.00000
7000	0.00014	0.00000	0.00000	7000	0.00014	0.00000	0.00000
8000	0.00012	0.00000	0.00000	8000	0.00013	0.00000	0.00000
9000	0.00011	0.00000	0.00000	9000	0.00011	0.00000	0.00000
10000	0.00010	0.00000	0.00000	10000	0.00010	0.00000	0.00000
20000	0.00005	0.00000	0.00000	20000	0.00005	0.00000	0.00000
30000	0.00003	0.00000	0.00000	30000	0.00003	0.00000	0.00000
40000	0.00002	0.00000	0.00000	40000	0.00003	0.00000	0.00000
50000	0.00002	0.00000	0.00000	50000	0.00002	0.00000	0.00000
60000	0.00002	0.00000	0.00000	60000	0.00002	0.00000	0.00000
70000	0.00001	0.00000	0.00000	70000	0.00001	0.00000	0.00000
80000	0.00001	0.00000	0.00000	80000	0.00001	0.00000	0.00000
90000	0.00001	0.00000	0.00000	90000	0.00001	0.00000	0.00000
100000	0.00001	0.00000	0.00000	100000	0.00001	0.00000	0.00000

θ =−1	Table	4.3	k = 2	θ =−0.5	Table	4.4	k = 2
N	**MSE**	**Bias**	**Gain**	**n**	**MSE**	**Bias**	**Gain**
1	1.51428	0.28886	-0.51428	1	1.19754	0.27769	-0.19754
2	0.93211	0.24555	-0.43211	2	0.68010	0.26305	-0.18010
3	0.69830	0.20736	-0.36497	3	0.51488	0.25938	-0.18155
4	0.55443	0.17292	-0.30443	4	0.43416	0.25722	-0.18416
5	0.45144	0.14284	-0.25144	5	0.38573	0.25490	-0.18573
6	0.37285	0.11715	-0.20618	6	0.35277	0.25207	-0.18610
7	0.31100	0.09556	-0.16814	7	0.32832	0.24871	-0.18546
8	0.26151	0.07760	-0.13651	8	0.30902	0.24490	-0.18402
9	0.22152	0.06277	-0.11041	9	0.29308	0.24074	-0.18197
10	0.18902	0.05061	-0.08902	10	0.27945	0.23629	-0.17945

20	0.05908	0.00516	-0.00908	20	0.19463	0.18664	-0.14463
30	0.03412	0.00045	-0.00079	30	0.14315	0.14072	-0.10982
40	0.02506	0.00003	-0.00006	40	0.10606	0.10339	-0.08107
50	0.02000	0.00000	0.00000	50	0.07864	0.07451	-0.05864
60	0.01667	0.00000	0.00000	60	0.05839	0.05283	-0.04172
70	0.01429	0.00000	0.00000	70	0.04354	0.03693	-0.02925
80	0.01250	0.00000	0.00000	80	0.03274	0.02549	-0.02024
90	0.01111	0.00000	0.00000	90	0.02496	0.01738	-0.01384
100	0.01000	0.00000	0.00000	100	0.01936	0.01173	-0.00936
200	0.00500	0.00000	0.00000	200	0.00512	0.00014	-0.00012
300	0.00333	0.00000	0.00000	300	0.00333	0.00000	0.00000
400	0.00250	0.00000	0.00000	400	0.00250	0.00000	0.00000
500	0.00200	0.00000	0.00000	500	0.00200	0.00000	0.00000
600	0.00167	0.00000	0.00000	600	0.00167	0.00000	0.00000
700	0.00143	0.00000	0.00000	700	0.00143	0.00000	0.00000
800	0.00125	0.00000	0.00000	800	0.00125	0.00000	0.00000
900	0.00111	0.00000	0.00000	900	0.00111	0.00000	0.00000
1000	0.00100	0.00000	0.00000	1000	0.00100	0.00000	0.00000
2000	0.00050	0.00000	0.00000	2000	0.00050	0.00000	0.00000
3000	0.00033	0.00000	0.00000	3000	0.00033	0.00000	0.00000
4000	0.00025	0.00000	0.00000	4000	0.00025	0.00000	0.00000
5000	0.00020	0.00000	0.00000	5000	0.00020	0.00000	0.00000
6000	0.00017	0.00000	0.00000	6000	0.00017	0.00000	0.00000
7000	0.00014	0.00000	0.00000	7000	0.00014	0.00000	0.00000
8000	0.00012	0.00000	0.00000	8000	0.00012	0.00000	0.00000
9000	0.00011	0.00000	0.00000	9000	0.00011	0.00000	0.00000
10000	0.00010	0.00000	0.00000	10000	0.00010	0.00000	0.00000
20000	0.00005	0.00000	0.00000	20000	0.00005	0.00000	0.00000
30000	0.00003	0.00000	0.00000	30000	0.00003	0.00000	0.00000
40000	0.00002	0.00000	0.00000	40000	0.00002	0.00000	0.00000
50000	0.00002	0.00000	0.00000	50000	0.00002	0.00000	0.00000
60000	0.00002	0.00000	0.00000	60000	0.00002	0.00000	0.00000
70000	0.00001	0.00000	0.00000	70000	0.00001	0.00000	0.00000
80000	0.00001	0.00000	0.00000	80000	0.00001	0.00000	0.00000
90000	0.00001	0.00000	0.00000	90000	0.00001	0.00000	0.00000
100000	0.00001	0.00000	0.00000	100000	0.00001	0.00000	0.00000

θ = −0.1	**Table**	**4.5**	**k = 2**	**θ = 0**	**Table**	**4.6**	**k = 2**
N	**MSE**	**Bias**	**Gain**	**n**	**MSE**	**Bias**	**Gain**
1	0.95744	0.21807	0.04256	1	0.91357	0.19871	0.08643
2	0.44498	0.17809	0.05502	2	0.40906	0.14887	0.09094
3	0.28054	0.15986	0.05279	3	0.24835	0.12336	0.08499
4	0.20131	0.14919	0.04869	4	0.17140	0.10688	0.07860
5	0.15543	0.14215	0.04457	5	0.12712	0.09501	0.07288
6	0.12588	0.13718	0.04079	6	0.09875	0.08590	0.06791
7	0.10544	0.13349	0.03742	7	0.07927	0.07862	0.06359
8	0.09059	0.13066	0.03441	8	0.06519	0.07263	0.05981
9	0.07937	0.12842	0.03174	9	0.05464	0.06758	0.05647
10	0.07066	0.12662	0.02934	10	0.04649	0.06326	0.05351
20	0.03538	0.11868	0.01462	20	0.01462	0.03925	0.03538
30	0.02577	0.11625	0.00757	30	0.00680	0.02866	0.02653
40	0.02155	0.11508	0.00345	40	0.00378	0.02257	0.02122
50	0.01925	0.11437	0.00075	50	0.00234	0.01860	0.01766
60	0.01782	0.11387	-0.00115	60	0.00155	0.01580	0.01511
70	0.01684	0.11348	-0.00256	70	0.00109	0.01373	0.01320
80	0.01614	0.11317	-0.00364	80	0.00079	0.01212	0.01171
90	0.01561	0.11291	-0.00450	90	0.00060	0.01085	0.01051
100	0.01520	0.11269	-0.00521	100	0.00047	0.00981	0.00954
200	0.01349	0.11143	-0.00850	200	0.00009	0.00499	0.00491
300	0.01298	0.11084	-0.00965	300	0.00004	0.00333	0.00330
400	0.01273	0.11048	-0.01023	400	0.00002	0.00250	0.00248
500	0.01259	0.11022	-0.01059	500	0.00001	0.00200	0.00199
600	0.01249	0.11002	-0.01082	600	0.00001	0.00167	0.00166
700	0.01242	0.10985	-0.01099	700	0.00001	0.00143	0.00142
800	0.01237	0.10970	-0.01112	800	0.00000	0.00125	0.00125
900	0.01232	0.10956	-0.01121	900	0.00000	0.00111	0.00111
1000	0.01229	0.10943	-0.01129	1000	0.00000	0.00100	0.00100
2000	0.01206	0.10804	-0.01157	2000	0.00000	0.00050	0.00050
3000	0.01184	0.10590	-0.01151	3000	0.00000	0.00033	0.00033

4000	0.01150	0.10247	-0.01125	4000	0.00000	0.00025	0.00025
5000	0.01100	0.09740	-0.01080	5000	0.00000	0.00020	0.00020
6000	0.01030	0.09052	-0.01013	6000	0.00000	0.00017	0.00017
7000	0.00939	0.08191	-0.00925	7000	0.00000	0.00014	0.00014
8000	0.00833	0.07193	-0.00820	8000	0.00000	0.00012	0.00013
9000	0.00715	0.06115	-0.00704	9000	0.00000	0.00011	0.00011
10000	0.00594	0.05025	-0.00584	10000	0.00000	0.00010	0.00010
20000	0.00018	0.00108	-0.00014	20000	0.00000	0.00005	0.00005
30000	0.00003	0.00000	0.00000	30000	0.00000	0.00003	0.00003
40000	0.00003	0.00000	0.00000	40000	0.00000	0.00002	0.00003
50000	0.00002	0.00000	0.00000	50000	0.00000	0.00002	0.00002
60000	0.00002	0.00000	0.00000	60000	0.00000	0.00002	0.00002
70000	0.00001	0.00000	0.00000	70000	0.00000	0.00001	0.00001
80000	0.00001	0.00000	0.00000	80000	0.00000	0.00001	0.00001
90000	0.00001	0.00000	0.00000	90000	0.00000	0.00001	0.00001
100000	0.00001	0.00000	0.00000	100000	0.00000	0.00001	0.00001

θ = 0.5	**Table**	**4.7**	**k = 2**	**θ = 1**	**Table**	**4.8**	**k = 2**
N	**MSE**	**Bias**	**Gain**	**n**	**MSE**	**Bias**	**Gain**
1	0.82169	0.09816	0.17831	1	0.88795	0.02431	0.11205
2	0.40127	0.01577	0.09873	2	0.51321	-0.03609	-0.01321
3	0.28149	-0.02599	0.05184	3	0.38289	-0.04966	-0.04956
4	0.22790	-0.05097	0.02210	4	0.30767	-0.04954	-0.05767
5	0.19784	-0.06701	0.00216	5	0.25562	-0.04492	-0.05562
6	0.17838	-0.07767	-0.01171	6	0.21662	-0.03904	-0.04996
7	0.16448	-0.08487	-0.02162	7	0.18617	-0.03314	-0.04331
8	0.15382	-0.08970	-0.02882	8	0.16179	-0.02774	-0.03679
9	0.14521	-0.09287	-0.03410	9	0.14195	-0.02298	-0.03084
10	0.13799	-0.09483	-0.03799	10	0.12560	-0.01890	-0.02560
20	0.09592	-0.08793	-0.04592	20	0.05309	-0.00216	-0.00309
30	0.07233	-0.06989	-0.03899	30	0.03363	-0.00020	-0.00029
40	0.05563	-0.05296	-0.03063	40	0.02502	-0.00002	-0.00002
50	0.04316	-0.03902	-0.02316	50	0.02000	0.00000	0.00000
60	0.03372	-0.02817	-0.01706	60	0.01667	0.00000	0.00000
70	0.02660	-0.01999	-0.01231	70	0.01429	0.00000	0.00000
80	0.02124	-0.01398	-0.00874	80	0.01250	0.00000	0.00000
90	0.01722	-0.00965	-0.00611	90	0.01111	0.00000	0.00000
100	0.01421	-0.00658	-0.00421	100	0.01000	0.00000	0.00000
200	0.00506	-0.00009	-0.00006	200	0.00500	0.00000	0.00000
300	0.00333	0.00000	0.00000	300	0.00333	0.00000	0.00000
400	0.00250	0.00000	0.00000	400	0.00250	0.00000	0.00000
500	0.00200	0.00000	0.00000	500	0.00200	0.00000	0.00000
600	0.00167	0.00000	0.00000	600	0.00167	0.00000	0.00000
700	0.00143	0.00000	0.00000	700	0.00143	0.00000	0.00000
800	0.00125	0.00000	0.00000	800	0.00125	0.00000	0.00000
900	0.00111	0.00000	0.00000	900	0.00111	0.00000	0.00000
1000	0.00100	0.00000	0.00000	1000	0.00100	0.00000	0.00000
2000	0.00050	0.00000	0.00000	2000	0.00050	0.00000	0.00000
3000	0.00033	0.00000	0.00000	3000	0.00033	0.00000	0.00000
4000	0.00025	0.00000	0.00000	4000	0.00025	0.00000	0.00000
5000	0.00020	0.00000	0.00000	5000	0.00020	0.00000	0.00000
6000	0.00017	0.00000	0.00000	6000	0.00017	0.00000	0.00000
7000	0.00014	0.00000	0.00000	7000	0.00014	0.00000	0.00000
8000	0.00012	0.00000	0.00000	8000	0.00012	0.00000	0.00000
9000	0.00011	0.00000	0.00000	9000	0.00011	0.00000	0.00000
10000	0.00010	0.00000	0.00000	10000	0.00010	0.00000	0.00000
20000	0.00005	0.00000	0.00000	20000	0.00005	0.00000	0.00000
30000	0.00003	0.00000	0.00000	30000	0.00003	0.00000	0.00000
40000	0.00002	0.00000	0.00000	40000	0.00002	0.00000	0.00000
50000	0.00002	0.00000	0.00000	50000	0.00002	0.00000	0.00000
60000	0.00002	0.00000	0.00000	60000	0.00002	0.00000	0.00000
70000	0.00001	0.00000	0.00000	70000	0.00001	0.00000	0.00000
80000	0.00001	0.00000	0.00000	80000	0.00001	0.00000	0.00000
90000	0.00001	0.00000	0.00000	90000	0.00001	0.00000	0.00000
100000	0.00001	0.00000	0.00000	100000	0.00001	0.00000	0.00000

Θ = 3	Table	4.9	k = 2	Θ = 6	Table	4.10	k = 2
N	**MSE**	**Bias**	**Gain**	**n**	**MSE**	**Bias**	**Gain**
1	1.01332	-0.00281	-0.01332	1	1.00000	0.00000	0.00000
2	0.50108	-0.00022	-0.00109	2	0.50000	0.00000	0.00000
3	0.33339	-0.00001	-0.00006	3	0.33333	0.00000	0.00000
4	0.25000	0.00000	0.00000	4	0.25000	0.00000	0.00000
5	0.20000	0.00000	0.00000	5	0.20000	0.00000	0.00000
6	0.16667	0.00000	0.00000	6	0.16667	0.00000	0.00000
7	0.14286	0.00000	0.00000	7	0.14286	0.00000	0.00000
8	0.12500	0.00000	0.00000	8	0.12500	0.00000	0.00000
9	0.11111	0.00000	0.00000	9	0.11111	0.00000	0.00000
10	0.10000	0.00000	0.00000	10	0.10000	0.00000	0.00000
20	0.05000	0.00000	0.00000	20	0.05000	0.00000	0.00000
30	0.03333	0.00000	0.00000	30	0.03333	0.00000	0.00000
40	0.02500	0.00000	0.00000	40	0.02500	0.00000	0.00000
50	0.02000	0.00000	0.00000	50	0.02000	0.00000	0.00000
60	0.01667	0.00000	0.00000	60	0.01667	0.00000	0.00000
70	0.01429	0.00000	0.00000	70	0.01429	0.00000	0.00000
80	0.01250	0.00000	0.00000	80	0.01250	0.00000	0.00000
90	0.01111	0.00000	0.00000	90	0.01111	0.00000	0.00000
100	0.01000	0.00000	0.00000	100	0.01000	0.00000	0.00000
200	0.00500	0.00000	0.00000	200	0.00500	0.00000	0.00000
300	0.00333	0.00000	0.00000	300	0.00333	0.00000	0.00000
400	0.00250	0.00000	0.00000	400	0.00250	0.00000	0.00000
500	0.00200	0.00000	0.00000	500	0.00200	0.00000	0.00000
600	0.00167	0.00000	0.00000	600	0.00167	0.00000	0.00000
700	0.00143	0.00000	0.00000	700	0.00143	0.00000	0.00000
800	0.00125	0.00000	0.00000	800	0.00125	0.00000	0.00000
900	0.00111	0.00000	0.00000	900	0.00111	0.00000	0.00000
1000	0.00100	0.00000	0.00000	1000	0.00100	0.00000	0.00000
2000	0.00050	0.00000	0.00000	2000	0.00050	0.00000	0.00000
3000	0.00033	0.00000	0.00000	3000	0.00033	0.00000	0.00000
4000	0.00025	0.00000	0.00000	4000	0.00025	0.00000	0.00000
5000	0.00020	0.00000	0.00000	5000	0.00020	0.00000	0.00000
6000	0.00017	0.00000	0.00000	6000	0.00017	0.00000	0.00000
7000	0.00014	0.00000	0.00000	7000	0.00014	0.00000	0.00000
8000	0.00013	0.00000	0.00000	8000	0.00012	0.00000	0.00000
9000	0.00011	0.00000	0.00000	9000	0.00011	0.00000	0.00000
10000	0.00010	0.00000	0.00000	10000	0.00010	0.00000	0.00000
20000	0.00005	0.00000	0.00000	20000	0.00005	0.00000	0.00000
30000	0.00003	0.00000	0.00000	30000	0.00003	0.00000	0.00000
40000	0.00003	0.00000	0.00000	40000	0.00003	0.00000	0.00000
50000	0.00002	0.00000	0.00000	50000	0.00002	0.00000	0.00000
60000	0.00002	0.00000	0.00000	60000	0.00002	0.00000	0.00000
70000	0.00001	0.00000	0.00000	70000	0.00001	0.00000	0.00000
80000	0.00001	0.00000	0.00000	80000	0.00001	0.00000	0.00000
90000	0.00001	0.00000	0.00000	90000	0.00001	0.00000	0.00000
100000	0.00001	0.00000	0.00000	100000	0.00001	0.00000	0.00000

Table 4.11: Suitable Range of Sample Sizes (n) to prefer $\theta_n^*(k)$

θ	k = 2	k = 3	k = 4	k = 5
-10	≥ 1	≥ 1	≥ 1	≥ 1
-9	≥ 1	≥ 1	≥ 1	≥ 1
-8	≥ 1	≥ 1	≥ 1	≥ 1
-7	≥ 1	≥ 1	≥ 1	≥ 1
-6	≥ 2	≥ 2	≥ 2	≥ 2
-5	≥ 2	≥ 2	≥ 2	≥ 2
-4	≥ 3	≥ 3	≥ 3	≥ 3
-3	≥ 5	≥ 5	≥ 5	≥ 5
-2	≥ 20	≥ 20	≥ 20	≥ 20
-1	≥ 50	≥ 60	≥ 60	≥ 60
-0.9	≥ 70	≥ 80	≥ 80	≥ 80
-0.8	≥ 90	≥ 100	≥ 100	≥ 100
-0.7	≥ 200	≤1, ≥200	≥ 200	≤1, ≥200
-0.6	≥ 200	≤2, ≥200	≥ 200	≤1, ≥200
-0.5	≥ 300	≤3, ≥400	≥ 400	≤2, ≥400
-0.4	≥ 500	≤4, ≥600	≥ 600	≤4, ≥600
-0.3	≥ 2000	≤8, ≥2000	≤5, ≥2000	≤7, ≥2000

-0.2	≥ 4000	≤20, ≥4000	≤10, ≥4000	≤10, ≥4000
-0.1	≤50, ≥30000	≤90, ≥30000	≤80, ≥30000	≤90, ≥30000
0	≥ 1	≥ 1	≥ 1	≥ 1
0.1	≤100, ≥30000	≤90, ≥30000	≤90, ≥30000	≤90, ≥30000
0.2	≤20, ≥4000	≤20, ≥4000	≤10, ≥4000	≤10, ≥4000
0.3	≤10, ≥2000	≤8, ≥2000	≤8, ≥2000	≤7, ≥2000
0.4	≤7, ≥600	≤4, ≥600	≤4, ≥600	≤4, ≥600
0.5	≤5, ≥300	≤3, ≥400	≤3, ≥400	≤2, ≥400
0.6	≤3, ≥200	≤2, ≥200	≤2, ≥200	≤1, ≥200
0.7	≤3, ≥200	≤1, ≥200	≤1, ≥200	≤1, ≥200
0.8	≤2, ≥100	≤1, ≥100	≤1, ≥100	≤1, ≥100
0.9	≤2, ≥80	≥ 80	≤1, ≥80	≥ 80
1	≤1, ≥50	≥ 60	≤1, ≥60	≥ 60
2	≥ 20	≥ 20	≥ 20	≥ 20
3	≥ 4	≥ 5	≥ 5	≥ 5
4	≥ 3	≥ 3	≥ 3	≥ 3
5	≥ 2	≥ 2	≥ 2	≥ 2
6	≥ 1	≥ 2	≥ 2	≥ 2
7	≥ 1	≥ 1	≥ 1	≥ 1
8	≥ 1	≥ 1	≥ 1	≥ 1
9	≥ 1	≥ 1	≥ 1	≥ 1
10	≥ 1	≥ 1	≥ 1	≥ 1

5. Concluding Remarks

It may be noted that, for absolute values of θ not less than 1, $\theta_n^*(k)$ perform well compared to the MLE, even for moderate sample sizes. This is true for k = 2, 3, 4, 5. In fact, the performance of $\theta_n^*(k)$ and the MLE are very similar when the true θ shifts away from 0 to more than one standard deviation units, because under such ‘large’ shifts, the SEE and MLE would most likely be almost equal in value. This is evident in the way the SEE’s $\theta_n^*(k)$ have been defined in (2.1).

For smaller absolute values of θ considered in the interval (–1, 1), it is found that when ‘shifts’ in θ from 0 are not very small (to the magnitude of 0.3 standard deviation units and above), $\theta_n^*(k)$ perform well for smaller sample sizes and sizes exceeding 2000. For still smaller magnitude shifts, like 0.2 standard deviation units, again $\theta_n^*(k)$ perform well for small sizes like upto 10 or 20 and sizes exceeding 4000. However, for shifts of magnitude 0.1, sample size should exceed 30000 for $\theta_n^*(k)$ to perform better than the MLE or the sample size should be moderate around 50 to 100. The latter result gives some satisfaction that one need not go for very large sample sizes like 30000 when there are very small shifts in θ and can handle the situation even with samples of size upto 100 and use $\theta_n^*(k)$ to estimate θ.

As a prospective application of using SEE’s, we consider a manufacturing process with a normally distributed quality characteristic. In this situation, suppose shifts in the process average of magnitudes exceeding 0.3 standard deviation units need to be correctly detected. This is required in various situations because these shifts, though not very large, will lead to the process not meeting the quality specifications. In such a scenario, the numerical study carried out in this research article suggests that drawing a sample of size above 2000 is recommended to get a closer estimate of the process average using the SEE’s instead of the MLE.

A perusal of the ‘Bias’ columns in Tables 4.1 to 4.10 shows that the bias in $\theta_n^*(k)$ approaches zero in the right-tail of ‘n’ values [that is, larger ‘n’ values] revealing the ‘asymptotic unbiasedness’ of $\theta_n^*(k)$. We note that, for the left-tail of ‘n’ values [that is, smaller ‘n’ values] for which the ‘Gain function’ is positive, the Bias in $\theta_n^*(k)$ is non-zero (to the approximation of the order of 10^{-5}). That is, for smaller sample sizes, we do get biased estimators but with a lower mean-square error compared to the mean-square error of the MLE.

A closer examination of Table 4.11 also brings out the fact that the SEE $\theta_n^*(3)$ shows a better performance over the other SEE’s considered, in terms of the range of sample sizes wherein it overtakes the MLE. Moreover, if $\theta_n^*(2)$ or $\theta_n^*(4)$ are used, we see that negative estimate values will be made positive when $\overline{X}$ takes negative values closer to zero. Thus, $\theta_n^*(3)$ or $\theta_n^*(5)$ are preferable over $\theta_n^*(2)$ or $\theta_n^*(4)$.

However, as pointed above, the SEE $\theta_n^*(3)$ is preferable over a wider range of sample sizes compared to the MLE $\overline{X}$ and the other SEE’s as well.

A study on the performance of $\theta_n^*(k)$ for values of k > 4 and the question on the range of sample sizes for which the corresponding SEE’s may be preferred over the MLE remains to be addressed in a separate article.

References

1. Bahadur RR. Hodges Super-Efficiency. In: *Encyclopedia of Statistical Sciences*. John-Wiley, 1983; 3:645-646.
2. Basu D. Unpublished dissertation. 1952.
3. Durairajan TM. Sub-score and Super- Efficient Estimator. In: Navin Chandra and Gopal, G., eds. Applications of Reliability Theory and Survival Analysis, Bonfring, 2012, 120-129.
4. Hodges JL. Unpublished. 1951.
5. Le Cam, L. On some asymptotic properties of maximum likelihood estimates and related Baye's estimates. University of California Publications in Statistics. 1953; 1:277-330.
6. Le Cam L. On the asymptotic theory of estimation and testing hypotheses. In: Neyman, J. ed. Proceedings of the Third Berkeley Symposium on Mathematics and Probability. University of California Press, Berkeley. 1956; 1:129-156.
7. Le Cam L. Locally asymptotically normal families of distributions. University of California Publications in Statistics. 1960; 3:37-98.
8. Le Cam L. Limits of experiments. In*:* Proceedings ofthe Sixth Berkeley Symposium on Mathematics and Probability. University of California Press, Berkeley. 1972; 1:245-261.
9. Sethuraman J. Are Super- Efficient Estimators Super-powerful?. Communications in Statistics - Theory and Methods. 2004; 33(9):2003-2013.
10. Sivasakthi K, Durairajan TM, Martin William L. Obtaining Super-Efficient Estimators for a Real-Valued Parameter: Delta Method. International Journal of Applied Mathematics and Statistical Sciences. 2017; 6(1):81-88.
11. Stein, C. Inadmissibility of the usual estimator for the mean of a normal distribution. In: Neyman, J. ed. Proceedings of the Third Berkeley Symposium on Mathematics and Probability. University of California Press, Berkeley. 1956; 1:197-206.