International Journal of Scientific Research in Mathematical and Statistical Sciences **Research Paper**
Volume-5, Issue-6, pp.212-221, December (2018) **E-ISSN:** 2348-4519

A Class of Super-Efficient Estimators of the Normal Variance: A Study on Sample Size Preference

K. Sivasakthi[1*] , Martin L. William[2]

[1]Department of Statistics, Tagore College of Arts and Science, Chennai, India
[2]Department of Statistics, Loyola College, Chennai, India

Available online at: www.isroset.org

Received: 09/Dec/2018, Accepted: 21/Dec/2018, Online: 31/Dec/2018

Abstract**-** A class of super-efficient estimators of the variance of a normal population with known mean has been recently constructed by Sivasakthi, Durairajan, and William (2017) through the 'Delta Method'. The preference for a super-efficient estimator over the asymptotically efficient estimator (say, maximum likelihood estimator) is for 'large samples'. In this paper, we address the super-efficient estimation of the normal variance when the population mean is known. The issue that is taken up in this paper is on the sample size required for a super-efficient estimator to be preferred over the maximum likelihood estimator and is addressed through a numerical study. The answer is sought for a subset of the class of super-efficient estimators of the normal variance.

Mathematics Subject Classification: 62F10, 62F12
Keywords and Phrases: Maximum Likelihood Estimators, Super-efficiency

I. INTRODUCTION

The concept of 'Super-Efficient Estimator' was brought about in an unpublished work of Hodges (1951). This was followed up with theoretical developments by Basu (1952), Le Cam (1953, 1956, 1960, 1972) and Stein (1956). With such developments, a few interesting articles on super-efficiency were authored by Bahadur (1983), Sethuraman (2004) and Durairajan (2012). These investigations on super-efficiency have been quite interesting but have happened at a slow pace and, further, there has been no systematic approaches towards constructing super-efficient estimators. Recently, Sivasakthi, Durairajan, and William (2017) considered the construction of the same using the 'Delta Method' of asymptotic inference theory and applied it to an estimation of the normal variance besides other examples.

The property of super-efficiency is asymptotic and a super-efficient estimator (SEE) is expected to perform favorably over the asymptotically efficient estimator / maximum likelihood estimator for adequately large samples. In a practical situation, the question is on the sample size required to prefer an SEE over the asymptotic-efficient estimator / maximum likelihood estimator. This question has not been addressed until the recent past, but an interesting work on sample size required for preferring the SEEs over the MLE to estimate the normal mean has been carried out by Sivasakthi, Sakthivel, and William (2017). In the same spirit, the present work aims at addressing the same question for estimating the normal variance. Algebraic closed-form expressions are not available for the 'Mean-Square Error' or 'Variance' of super-efficient estimators and so, comparisons ought to be carried out numerically.

In Sivasakthi, Durairajan, and William (2017), the super-efficient estimation of the variance of a normal population with zero mean, N(0, θ), 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, with super-efficiency at θ= 1. But, the question on the sample size required for the SEE to overtake the MLE remains unanswered.

In the present paper, the answer to the question on sample size needed to prefer the SEE against the MLE is obtained through a numerical study. As SEE's are only asymptotically unbiased, the comparison is through the mean-square errors instead of variances of the estimators. For larger sample sizes, the mean-square error is approximately equal to variance owing to asymptotic unbiasedness.

This paper has five sections including the present introductory section. Section 2 reviews the class of super-efficient estimators under consideration. Section 3 provides the theoretical results concerning the mean-square error and bias of the SEE's for small sample situation. In Section 4, the same results for large sample situation are derived. Section 5 presents the results of the numerical work carried out to compare the mean-square errors of the SEE's with the variance of the MLE. Finally, Section 6 contains concluding remarks with recommendations on the sample sizes needed for the preferential use of the class of SEE's over the MLE of the variance.

II. A CLASS OF SUPER-EFFICIENT ESTIMATORS OF THE NORMAL VARIANCE

In a recent work, Sivasakthi, Durairajan, and William (2017) applied the 'Delta Method' to obtain super-efficient estimators for the variance of a normal distribution whose mean is known, assumed to be zero without loss of generality, *viz* N(0, θ), $\theta > 0$. The class of super-efficient estimators considered in that paper is

$$\breve{\theta}_n(d) = \begin{cases} (S^2)^d & if \ \sqrt{n}\,\left|S^2 - 1\right| \le n^{1/4} \\ S^2 & if \ \sqrt{n}\,\left|S^2 - 1\right| > n^{1/4} \end{cases} \tag{2.1}$$

where $S^2 = \sum_{i=1}^{n} X_i^2 / n$ is the MLE of θ and d is any known number in the interval (0, 1).

This class is obtained by choosing the function g(θ) = θ^d on the parameter space, where d is any known number with 0<d< 1. The estimators in this class are super-efficient at θ= 1. This class is interesting because when the estimator S^2 is closer to 1, which occurs with higher likelihood when the true variance is unity or close to unity, the SEE 'improves' the estimate by moving closer to unity by raising it to a power between 0 and 1. Hence, there is a rationale in considering the class (2.1). The use of an SEE from the class rather than the MLE is logically appealing for 'large' samples. The question we address in the subsequent sections of this paper is: how large should the sample be to prefer an SEE from the class (2.1) over the MLE? The answer is got through numerical methods by deriving the required theoretical expressions in Sections 3 and 4.

III. SOME THEORETICAL RESULTS

In this section, we provide some theoretical results on the class of S.E.E.'s $\breve{\theta}_n(d)$ to derive expressions for the Mean Square Error and Bias of the estimators.

3.1 Bias of SEE's $\breve{\theta}_n(d)$ *in* estimating θ

Denote S^2 = T and the event A = $[\,|T-1| > n^{-1/4}\,]$

Now, $E_\theta\left(\breve{\theta}_n(d)\right) = E_\theta\left(T \mid |T-1| > n^{-1/4}\right) P_\theta\left(|T-1| > n^{-1/4}\right) + E_\theta\left(T^d \mid |T-1| \le n^{-1/4}\right) P_\theta\left(|T-1| \le n^{-1/4}\right)$

$$= a_\theta + b_\theta(d) \tag{3.1}$$

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

Let U = n T / θ. Using the fact that U follows $\chi^2(n)$, the p.d.f. of 'T' is given by

$$f(t) = \left(\frac{n}{2\theta}\right)^{n/2} \frac{t^{n/2-1} e^{-nt/(2\theta)}}{\Gamma(n/2)}, \quad t > 0$$

The conditional p.d.f. of T given event A is $g_A(t) = \begin{cases} f(t)/P(A), & t \in A \\ 0, & otherwise \end{cases}$

Now, $E_\theta\left(T \mid A\right) = \int_{t \in A} t \, g_A(t) \, dt$

And, $a_\theta = E_\theta\left(T \mid A\right) P_\theta(A) = \left(\frac{n}{2\theta}\right)^{n/2} \frac{1}{\Gamma(n/2)} \int_A t^{n/2} e^{-nt/(2\theta)} \, dt$

$$= \left(\frac{n}{2\theta}\right)^{n/2} \frac{1}{\Gamma(n/2)} \left[\frac{\Gamma(n/2+1)\ (2\theta)^{n/2+1}}{n^{n/2+1}} - \int_{A^C} t^{n/2}\, e^{-nt/(2\theta)}\ dt \right]$$

$$= \theta - \left(\frac{n}{2\theta}\right)^{n/2} \frac{1}{\Gamma(n/2)} \int_{1-n^{-1/4}}^{1+n^{-1/4}} t^{n/2} e^{-nt/(2\theta)} dt \tag{3.2}$$

Also, $b_\theta(d) = \mathrm{E}_\theta\left(\mathrm{T}^d \mid A^c\right) \mathrm{P}_\theta\left(A^c\right) = \left(\frac{n}{2\theta}\right)^{\frac{n}{2}} \frac{1}{\Gamma(n/2)} \int_{1-n^{-1/4}}^{1+n^{-1/4}} t^{n/2+d-1} e^{-nt/(2\theta)} dt$ (3.3)

Using (3.2) and (3.3) in (3.1), we get $\mathrm{E}_\theta\left(\breve{\theta}_n(d)\right)$.

Hence, $Bias_\theta\left(\breve{\theta}_n(d)\right) = \left(\frac{n}{2\theta}\right)^{\frac{n}{2}} \frac{1}{\Gamma(n/2)} \int_{1-n^{-1/4}}^{1+n^{-1/4}} (t^{n/2+d-1} - t^{n/2})\ e^{-nt/(2\theta)} dt$ (3.4)

3.2 Mean-Square Error of SEE's

Consider $E_\theta\left(\breve{\theta}_n^{\,2}(d)\right) = E_\theta\left(T^2 \mid A\right) P_\theta(A) + E_\theta\left(T^{2d} \mid A^c\right) P_\theta(A^c) = u_\theta + v_\theta(d)$ (3.5)

With some computations, we get

$$u_\theta = E_\theta\left(T^2 \mid A\right) P_\theta(A) = \frac{\theta^2 (n+2)}{n} - \left(\frac{n}{2\theta}\right)^{\frac{n}{2}} \frac{1}{\Gamma(n/2)} \int_{1-n^{-1/4}}^{1+n^{-1/4}} t^{n/2+1} e^{-nt/(2\theta)} dt \tag{3.6}$$

and $v_\theta(d) = E_\theta\left(T^{2d} \mid A^c\right) P_\theta(A^c) = \left(\frac{n}{2\theta}\right)^{\frac{n}{2}} \frac{1}{\Gamma(n/2)} \int_{1-n^{-1/4}}^{1+n^{-1/4}} t^{n/2+2d-1}\, e^{-nt/(2\theta)}\ dt$ (3.7)

Using (3.6) and (3.7) in (3.5) we get $E_\theta\left(\breve{\theta}_n^{\,2}(d)\right)$.

Hence, the mean square error of $\breve{\theta}_n(d)$ is given by

$$MSE_\theta\left(\breve{\theta}_n(d)\right) = E_\theta\left(\breve{\theta}_n(d) - \theta\right)^2 = u_\theta + v_\theta(d) - 2\,\theta\,(a_\theta + b_\theta(d)) + \theta^2$$

$$= \frac{2\theta^2}{n} - \left(\frac{n}{2\theta}\right)^{\frac{n}{2}} \frac{1}{\Gamma(n/2)} \left(\int_{1-n^{-1/4}}^{1+n^{-1/4}} t^{n/2-1}\, (t^2 - t^{2d})\ e^{-nt/(2\theta)}\ dt - 2\,\theta \int_{1-n^{-1/4}}^{1+n^{-1/4}} t^{n/2-1}\, (t - t^d)\ e^{-nt/(2\theta)}\ dt \right) \tag{3.8}$$

3.3 Gain in preferring the SEE's over the MLE

The MLE of θ is S^2 which is efficient with variance $2\theta^2/n$. The gain or loss in preferring $\breve{\theta}_n(d)$ over the MLE S^2 in estimating θ is given by

$$Gain_\theta\left(\breve{\theta}_n(d)\right) = Var_\theta\left(S^2\right) - MSE_\theta\left(\breve{\theta}_n(d)\right)$$

$$= \left(\frac{n}{2\theta}\right)^{\frac{n}{2}} \frac{1}{\Gamma(n/2)} \left(\int_{1-n^{-1/4}}^{1+n^{-1/4}} t^{n/2-1}\, (t^2 - t^{2d})\ e^{-nt/(2\theta)}\ dt - 2\,\theta \int_{1-n^{-1/4}}^{1+n^{-1/4}} t^{n/2-1}\, (t - t^d)\ e^{-nt/(2\theta)}\ dt \right) \tag{3.9}$$

Non-negative values of the Gain function indicate the higher performance of the SEE over the MLE and we seek to know the values of 'n' for which this occurs.

IV. RESULTS USING THE NORMAL APPROXIMATION TO GAMMA

For large 'n', $\frac{U-n}{\sqrt{2n}}$ ~ N (0,1). Consider $A^C = \left[\frac{n-n^{3/4}}{\theta} \le U \le \frac{n+n^{3/4}}{\theta}\right]$

$$=\left[\frac{n-n^{3/4}-n\theta}{\theta\sqrt{2n}} \le \frac{U-n}{\sqrt{2n}} \le \frac{n+n^{3/4}-n\theta}{\theta\sqrt{2n}}\right] = A^{*C}, \text{ say}$$

Now, E_θ [T|A] $= \frac{\theta}{n} E_\theta (U \mid A) = \frac{\theta}{n}\int_A \frac{u\,h(u)}{P(A)} du$ where h(u) is pdf of N(n, 2n)

So, E_θ [T|A] P(A)$= \frac{\theta}{n}\int_A u \frac{e^{-\frac{1}{2.2n}(u-n)^2}}{\sqrt{2\pi}\sqrt{2n}} du$

$= \frac{\theta}{n}\int_{A*}\left(z\sqrt{2n}+n\right)\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\, dz$

$=\frac{\theta\sqrt{2}}{\sqrt{n}}\int_{A*} z\,\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\, dz + \theta\int_{A*}\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz$

$=\frac{\theta\sqrt{2}}{\sqrt{n}}\left[-\int_{A^{*c}} z\,\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz\right] + \theta\left[1-\int_{A^{*c}}\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz\right] = a_\theta$, (say) (4.1)

And, E_θ [T^d|A^C] P(A^C) $=\frac{\theta^d}{n^d}\int_{A^c} u^d \frac{1}{\sqrt{2\pi}\sqrt{2n}} e^{-\frac{(u-n)^2}{2.2n}} du$

$= \frac{\theta^d}{n^d\sqrt{2\pi}}\int_{A^{*c}}\left(n+z\sqrt{2n}\right)^d e^{-\frac{z^2}{2}} dz = b_\theta$(d) , (say) (4.2)

Thus, $E\left(\breve{\theta}_n(d)\right) = a_\theta + b_\theta$(d)

$= \theta - \theta\int_{A^{*C}}\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz - \frac{\theta\sqrt{2}}{\sqrt{n}}\int_{A^{*C}}\frac{z}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz + \frac{\theta^d}{n^d\sqrt{2\pi}}\int_{A^{*C}}\left(n+z\sqrt{2n}\right)^d e^{-\frac{z^2}{2}} dz$

Therefore,

$Bias_\theta\left(\breve{\theta}_n(d)\right) = \frac{\theta^d}{n^d\sqrt{2\pi}}\int_{A^{*C}}\left(n+z\sqrt{2n}\right)^d e^{-\frac{z^2}{2}} dz - \theta\int_{A^{*C}}\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz - \frac{\theta\sqrt{2}}{\sqrt{n}}\int_{A^{*C}}\frac{z}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$ (4.3)

Now, E(T^2|A).P(A)$= \frac{\theta^2}{n^2} E(U^2|A).P(A) = \frac{\theta^2}{n^2}\int_{A^*}\left(\sqrt{2n}.z+n\right)^2\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$

$=\frac{2\theta^2}{n}\int_{A^*}\frac{z^2}{\sqrt{2\pi}} e^{-z^2} dz + \theta^2\int_{A^*}\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz + \frac{2^{-\frac{3}{2}}\theta^2}{\sqrt{n}}\int_{A^*}\frac{z}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$

$$= \frac{2\theta^2}{n}\left[\int_{-\infty}^{\infty}\frac{z^2}{\sqrt{2\pi}} e^{-z^2} dz - \int_{A^{*C}}\frac{z^2}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz\right] + \theta^2\left[\int_{-\infty}^{\infty}\frac{z^2}{\sqrt{2\pi}} e^{-z^2} dz - \int_{A^{*C}}\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz\right] +$$

$$\frac{2^{\frac{3}{2}}\theta^2}{\sqrt{n}}\left[\int_{-\infty}^{\infty}\frac{z}{\sqrt{2\pi}} e^{-z^2} dz - \int_{A^{*C}}\frac{z}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz\right]$$

$$= \frac{2\theta^2}{n} + \theta^2 - \frac{2\theta^2}{n}\int_{A^{*C}}\frac{z^2}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz - \theta^2\int_{A^{*C}}\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz - \frac{2^{\frac{3}{2}}\theta^2}{\sqrt{n}}\int_{A^{*C}}\frac{z}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = u_\theta \text{ (say)}$$

(4.4)

And, E(T^{2d}|A^C).P(A^C)$=\frac{\theta^{2d}}{n^{2d}}\int_{A^C} u^{2d}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2n}} e^{-\frac{(u-n)^2}{2.2n}} du$

$= \frac{\theta^{2d}}{n^{2d}}\int_{A^{*c}}\left(z\sqrt{2n}+n\right)^{2d}\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = v_\theta(d)$ (say) (4.5)

Now, $E\left(\breve{\theta}_n^{\,2}(d)\right) = \frac{2\theta^2}{n} + \theta^2 - \frac{2\theta^2}{n}\int_{A^{*C}}\frac{z^2}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz - \theta^2\int_{A^{*C}}\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$

$$- \frac{2^{\frac{3}{2}}\theta^2}{\sqrt{n}}\int_{A^{*C}}\frac{z}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz + \frac{\theta^{2d}}{n^{2d}}\int_{A^C}\left(z\sqrt{2n}+n\right)^{2d}\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad (4.6)$$

Therefore, $MSE\left(\breve{\theta}_{n(d)}\right) = E\left(\breve{\theta}_n^{\,2}(\mathrm{d})\right) - 2\theta\ E\left(\breve{\theta}_n(\mathrm{d})\right) + \theta^2$ (4.7)

and $Gain_\theta\left(\breve{\theta}_n(d)\right) = \dfrac{2\theta^2}{n} - MSE\left(\breve{\theta}_{n(d)}\right)$ (4.8)

V. NUMERICAL INVESTIGATION OF THE PERFORMANCE OF S.E.E. OVER M.L.E.

As observed in Sections 3 and 4, there is no closed algebraic expression for $Gain_\theta\left(\breve{\theta}_n(d)\right)$ and the solution to the question on the sample sizes required to prefer the SEE cannot be resolved mathematically but only numerically. Also, the class of SEEs in (2.1) is infinitely large and therefore, we consider a subset of the class by choosing d = 0.1 (0.1) 0.9 and carry out the numerical study. For the problem of estimating the normal mean with known variance, Sivasakthi, Sakthivel, and William (2017) carried out a similar numerical investigation for comparing the performance of a class SEE's with that of the MLE and we refer the reader to that paper for further details.

As stated earlier, the comparison of the SEE's $\breve{\theta}_n(d)$ with the (asymptotic) efficient estimator [namely, the MLE, $S^2 = \sum_{i=1}^{n} X_i^2/n$] is through the 'Gain Function' given in (3.9) for the small sample case and (4.8) for a large sample case. We wish to know the values of 'n' for which this Gain Function is non-negative. In addition to the gain function, the Bias in $\breve{\theta}_n(d)$ is of interest as the SEE's are not exactly unbiased but only asymptotically. Hence, in this section, we present the Mean Square Error, Bias and the Gain in $\breve{\theta}_n(d)$.

We compute the Mean Square Error, Bias and Gain for a wide choice of sample sizes and a wide range of θ values but for the economy of space, we report the results for θ = 0.1, 0.5, 0.9, 1.0,1.5, 2.0, 3.0, 5.0,10.0 corresponding to d = 0.5 only. The sample sizes 'n' for which the SEE is gainful over the MLE are indicated with shaded cells in the Tables 5.1 to 5.9. We then give an abridged Table 5.10 for the d = 0.1 (0.1) 0.9, indicating the sample sizes for which the SEE's are preferable over the MLE corresponding to the range of θ values considered. We round off the results to five decimal places.

θ = 0.1 Table 5.1 d = 0.5

n	MSE	Bias	Gain
2	0.03720	0.04934	-0.02720
3	0.01726	0.01596	-0.01059
4	0.00846	0.00470	-0.00346
5	0.00500	0.00127	-0.00100
6	0.00360	0.00032	-0.00027
7	0.00292	0.00008	-0.00007
8	0.00252	0.00002	-0.00002
9	0.00223	0	0
10	0.00200	0	0
20	0.00100	0	0
30	0.00067	0	0
40	0.00050	0	0
50	0.00040	0	0
60	0.00033	0	0
70	0.00029	0	0
80	0.00025	0	0
90	0.00022	0	0
100	0.00020	0	0
200	0.00010	0	0

θ = 0.5 Table 5.2 d = 0.5

n	MSE	Bias	Gain
2	0.21927	0.09764	0.03073
3	0.15727	0.09501	0.00939
4	0.12807	0.09381	-0.00307
5	0.11112	0.09258	-0.01112
6	0.09987	0.09109	-0.01653
7	0.09168	0.08932	-0.02025
8	0.08531	0.08733	-0.02281
9	0.08011	0.08517	-0.02456
10	0.07571	0.08289	-0.02571
20	0.04935	0.05913	-0.02435
30	0.03456	0.03960	-0.01790
40	0.02469	0.02555	-0.01219
50	0.01795	0.01602	-0.00795
60	0.01334	0.00980	-0.00501
70	0.01022	0.00587	-0.00308
80	0.00810	0.00345	-0.00185
90	0.00664	0.00199	-0.00109
100	0.00563	0.00113	-0.00063
200	0.00250	0	0

θ = 0.9 Table 5.1 d = 0.5

n	MSE	Bias	Gain
2	0.70400	0.05368	0.10600
3	0.45340	0.04540	0.08660
4	0.32973	0.04114	0.07527
5	0.25656	0.03856	0.06744
6	0.20845	0.03688	0.06155
7	0.17456	0.03572	0.05686
8	0.14949	0.03492	0.05301
9	0.13024	0.03435	0.04976
10	0.11505	0.03395	0.04695
20	0.05021	0.03357	0.03079
30	0.03095	0.03482	0.02305
40	0.02219	0.03615	0.01831
50	0.01734	0.03733	0.01506
60	0.01432	0.03834	0.01268
70	0.01228	0.03920	0.01086
80	0.01082	0.03993	0.00943
90	0.00974	0.04055	0.00826
100	0.00889	0.04109	0.00731
200	0.00543	0.04397	0.00267

300	0.00007	0	0
400	0.00005	0	0
500	0.00004	0	0
600	0.00003	0	0
700	0.00003	0	0
800	0.00003	0	0
900	0.00002	0	0
1000	0.00002	0	0
2000	0.00001	0	0
3000	0.00001	0	0
4000	0.00001	0	0
5000	0	0	0
6000	0	0	0
7000	0	0	0
8000	0	0	0
9000	0	0	0
10000	0	0	0
20000	0	0	0
30000	0	0	0
40000	0	0	0
50000	0	0	0
60000	0	0	0
70000	0	0	0
80000	0	0	0
90000	0	0	0
100000	0	0	0

300	0.00167	0	0
400	0.00125	0	0
500	0.00100	0	0
600	0.00083	0	0
700	0.00071	0	0
800	0.00063	0	0
900	0.00056	0	0
1000	0.00050	0	0
2000	0.00025	0	0
3000	0.00017	0	0
4000	0.00013	0	0
5000	0.00010	0	0
6000	0.00008	0	0
7000	0.00007	0	0
8000	0.00006	0	0
9000	0.00006	0	0
10000	0.00005	0	0
20000	0.00003	0	0
30000	0.00002	0	0
40000	0.00001	0	0
50000	0.00001	0	0
60000	0.00001	0	0
70000	0.00001	0	0
80000	0.00001	0	0
90000	0.00001	0	0
100000	0.00001	0	0

300	-0.08339	0.04441	0.08879
400	-0.08405	0.04509	0.08810
500	-0.08441	0.04548	0.08765
600	-0.08462	0.04572	0.08732
700	-0.08475	0.04586	0.08706
800	-0.08481	0.04595	0.08683
900	-0.08483	0.04600	0.08663
1000	-0.08482	0.04602	0.08644
2000	-0.08384	0.04542	0.08465
3000	-0.08180	0.04404	0.08234
4000	-0.07880	0.04208	0.07921
5000	-0.07481	0.03956	0.07513
6000	-0.06986	0.03655	0.07013
7000	-0.06408	0.03315	0.06431
8000	-0.05767	0.02947	0.05787
9000	-0.05087	0.02569	0.05105
10000	-0.04396	0.02193	0.04412
20000	-0.00334	0.00152	0.00343
30000	0.00001	0.00002	0.00005
40000	0.00004	0	0
50000	0.00003	0	0
60000	0.00003	0	0
70000	0.00002	0	0
80000	0.00002	0	0
90000	0.00002	0	0
100000	0.00002	0	0

θ = 1 Table 5.4 d = 0.5

n	MSE	Bias	Gain
2	0.88658	0.04587	0.11342
3	0.57396	0.03567	0.09271
4	0.41925	0.02976	0.08075
5	0.32743	0.02572	0.07257
6	0.26687	0.02273	0.06646
7	0.22406	0.02039	0.06165
8	0.19228	0.01849	0.05772
9	0.16780	0.01691	0.05442
10	0.14840	0.01557	0.05160
20	0.06438	0.00833	0.03562
30	0.03858	0.00525	0.02809
40	0.02654	0.00353	0.02346
50	0.01975	0.00245	0.02025

θ = 1.5 Table 5.5 d = 0.5

n	MSE	Bias	Gain
2	2.13125	0.02112	0.11875
3	1.41632	0.00613	0.08368
4	1.06249	-0.00356	0.06251
5	0.85227	-0.01063	0.04773
6	0.71337	-0.01610	0.03663
7	0.61494	-0.02047	0.02791
8	0.54163	-0.02405	0.02087
9	0.48495	-0.02703	0.01505
10	0.43983	-0.02954	0.01017
20	0.23870	-0.04103	-0.01370
30	0.17056	-0.04250	-0.02056
40	0.13474	-0.04077	-0.02224
50	0.11189	-0.03783	-0.02189

θ = 2 Table 5.6 d = 0.5

n	MSE	Bias	Gain
2	3.89642	0.00952	0.10358
3	2.61321	-0.00546	0.05345
4	1.97581	-0.01433	0.02419
5	1.59500	-0.02007	0.00500
6	1.34158	-0.02389	-0.00825
7	1.16050	-0.02644	-0.01764
8	1.02437	-0.02809	-0.02437
9	0.91808	-0.02908	-0.02919
10	0.83259	-0.02960	-0.03259
20	0.43482	-0.02425	-0.03482
30	0.29104	-0.01611	-0.02437
40	0.21561	-0.01010	-0.01561
50	0.16965	-0.00617	-0.00965

60	0.01546	0.00172	0.01787
70	0.01256	0.00120	0.01602
80	0.01047	0.00082	0.01453
90	0.00892	0.00053	0.01330
100	0.00773	0.00031	0.01227
200	0.00308	-0.00044	0.00692
300	0.00185	-0.00081	0.00482
400	0.00133	-0.00062	0.00367
500	0.00104	-0.00050	0.00296
600	0.00085	-0.00042	0.00248
700	0.00073	-0.00036	0.00213
800	0.00063	-0.00031	0.00187
900	0.00056	-0.00028	0.00166
1000	0.00050	-0.00025	0.00150
2000	0.00025	-0.00013	0.00075
3000	0.00017	-0.00008	0.00050
4000	0.00013	-0.00006	0.00037
5000	0.00010	-0.00005	0.00030
6000	0.00008	-0.00004	0.00025
7000	0.00007	-0.00004	0.00021
8000	0.00006	-0.00003	0.00019
9000	0.00006	-0.00003	0.00017
10000	0.00005	-0.00003	0.00015
20000	0.00003	-0.00001	0.00007
30000	0.00002	-0.00001	0.00005
40000	0.00001	-0.00001	0.00004
50000	0.00001	-0.00001	0.00003
60000	0.00001	0	0.00002
70000	0.00001	0	0.00002
80000	0.00001	0	0.00002
90000	0.00001	0	0.00002
100000	0.00001	0	0.00001

60	0.09566	-0.03450	-0.02066
70	0.08336	-0.03112	-0.01907
80	0.07362	-0.02787	-0.01737
90	0.06567	-0.02482	-0.01567
100	0.05904	-0.02202	-0.01404
200	0.02645	-0.00587	-0.00395
300	0.02637	-0.00174	-0.01137
400	0.01430	-0.00046	-0.00305
500	0.00978	-0.00011	-0.00078
600	0.00769	-0.00003	-0.00019
700	0.00647	-0.00001	-0.00005
800	0.00564	0	-0.00001
900	0.00500	0	0
1000	0.00450	0	0
2000	0.00225	0	0
3000	0.00150	0	0
4000	0.00113	0	0
5000	0.00090	0	0
6000	0.00075	0	0
7000	0.00064	0	0
8000	0.00056	0	0
9000	0.00050	0	0
10000	0.00045	0	0
20000	0.00023	0	0
30000	0.00015	0	0
40000	0.00011	0	0
50000	0.00009	0	0
60000	0.00008	0	0
70000	0.00006	0	0
80000	0.00006	0	0
90000	0.00005	0	0
100000	0.00005	0	0

60	0.13918	-0.00371	-0.00585
70	0.11778	-0.00220	-0.00350
80	0.10207	-0.00130	-0.00207
90	0.09011	-0.00076	-0.00122
100	0.08071	-0.00044	-0.00071
200	0.04000	0	0
300	0.02667	0	0
400	0.02000	0	0
500	0.01600	0	0
600	0.01333	0	0
700	0.01143	0	0
800	0.01000	0	0
900	0.00889	0	0
1000	0.00800	0	0
2000	0.00400	0	0
3000	0.00267	0	0
4000	0.00200	0	0
5000	0.00160	0	0
6000	0.00133	0	0
7000	0.00114	0	0
8000	0.00100	0	0
9000	0.00089	0	0
10000	0.00080	0	0
20000	0.00040	0	0
30000	0.00027	0	0
40000	0.00020	0	0
50000	0.00016	0	0
60000	0.00013	0	0
70000	0.00011	0	0
80000	0.00010	0	0
90000	0.00009	0	0
100000	0.00008	0	0

θ = 3 Table 5.7 d = 0.5

n	MSE	Bias	Gain
2	8.93284	0.00034	0.06716
3	5.99949	-0.01130	0.00051
4	4.53097	-0.01635	-0.03097
5	3.64610	-0.01829	-0.04610
6	3.05250	-0.01859	-0.05250
7	2.62542	-0.01799	-0.05400
8	2.30270	-0.01691	-0.05270

θ = 5 Table 5.8 d = 0.5

n	MSE	Bias	Gain
2	24.98450	-0.00369	0.01550
3	16.71931	-0.00991	-0.05264
4	12.56819	-0.01067	-0.06819
5	10.06528	-0.00961	-0.06528
6	8.38955	-0.00803	-0.05622
7	7.18882	-0.00645	-0.04597
8	6.28647	-0.00506	-0.03647

θ = 10 Table 5.9 d = 0.5

n	MSE	Bias	Gain
2	100.04123	-0.00368	-0.04123
3	66.74876	-0.00519	-0.08209
4	50.06622	-0.00399	-0.06622
5	40.04476	-0.00265	-0.04476
6	33.36146	-0.00165	-0.02812
7	28.58843	-0.00099	-0.01700
8	25.01005	-0.00058	-0.01005

9	2.04986	-0.01560	-0.04986
10	1.84623	-0.01421	-0.04623
20	0.91441	-0.00417	-0.01441
30	0.60363	-0.00104	-0.00363
40	0.45086	-0.00024	-0.00086
50	0.36020	-0.00006	-0.00020
60	0.30004	-0.00001	-0.00004
70	0.25715	0	-0.00001
80	0.22500	0	0
90	0.20000	0	0
100	0.18000	0	0
200	0.09000	0	0
300	0.06000	0	0
400	0.04500	0	0
500	0.03600	0	0
600	0.03000	0	0
700	0.02571	0	0
800	0.02250	0	0
900	0.02000	0	0
1000	0.01800	0	0
2000	0.00900	0	0
3000	0.00600	0	0
4000	0.00450	0	0
5000	0.00360	0	0
6000	0.00300	0	0
7000	0.00257	0	0
8000	0.00225	0	0
9000	0.00200	0	0
10000	0.00180	0	0
20000	0.00090	0	0
30000	0.00060	0	0
40000	0.00045	0	0
50000	0.00036	0	0
60000	0.00030	0	0
70000	0.00026	0	0
80000	0.00023	0	0
90000	0.00020	0	0
100000	0.00018	0	0

9	5.58394	-0.00391	-0.02838
10	5.02180	-0.00298	-0.02180
20	2.50117	-0.00016	-0.00117
30	1.66672	-0.00001	-0.00005
40	1.25000	0	0
50	1.00000	0	0
60	0.83333	0	0
70	0.71429	0	0
80	0.62500	0	0
90	0.55556	0	0
100	0.50000	0	0
200	0.25000	0	0
300	0.16667	0	0
400	0.12500	0	0
500	0.10000	0	0
600	0.08333	0	0
700	0.07143	0	0
800	0.06250	0	0
900	0.05556	0	0
1000	0.05000	0	0
2000	0.02500	0	0
3000	0.01667	0	0
4000	0.01250	0	0
5000	0.01000	0	0
6000	0.00833	0	0
7000	0.00714	0	0
8000	0.00625	0	0
9000	0.00556	0	0
10000	0.00500	0	0
20000	0.00250	0	0
30000	0.00167	0	0
40000	0.00125	0	0
50000	0.00100	0	0
60000	0.00083	0	0
70000	0.00071	0	0
80000	0.00063	0	0
90000	0.00056	0	0
100000	0.00050	0	0

9	22.22807	-0.00034	-0.00585
10	20.00337	-0.00019	-0.00337
20	10.00001	0	-0.00001
30	6.66667	0	0
40	5.00000	0	0
50	4.00000	0	0
60	3.33333	0	0
70	2.85714	0	0
80	2.50000	0	0
90	2.22222	0	0
100	2.00000	0	0
200	1.00000	0	0
300	0.66667	0	0
400	0.50000	0	0
500	0.40000	0	0
600	0.33333	0	0
700	0.28571	0	0
800	0.25000	0	0
900	0.22222	0	0
1000	0.20000	0	0
2000	0.10000	0	0
3000	0.06667	0	0
4000	0.05000	0	0
5000	0.04000	0	0
6000	0.03333	0	0
7000	0.02857	0	0
8000	0.02500	0	0
9000	0.02222	0	0
10000	0.02000	0	0
20000	0.01000	0	0
30000	0.00667	0	0
40000	0.00500	0	0
50000	0.00400	0	0
60000	0.00333	0	0
70000	0.00286	0	0
80000	0.00250	0	0
90000	0.00222	0	0
100000	0.00200	0	0

It is interesting to observe the performance of $\breve{\theta}_n(d)$ over the MLE for small sizes also, especially for θ near 1 and away from 1, since the class $\breve{\theta}_n(d)$ itself is constructed with super-efficiency at θ = 1.

Table 5.10
Suitable Range of Sample Sizes (n) to
prefer $\breve{\theta}_n(d)$

θ	d = 0.9	d = 0.8	d = 0.7	d = 0.6	d = 0.5	d = 0.4	d = 0.3	d = 0.2	d = 0.1
0.1	≥ 7	≥ 8	≥ 8	≥ 8	≥ 8	≥ 8	≥ 8	≥ 8	≥ 8
0.2	≥ 20	≥ 20	≥ 20	≥ 20	≥ 20	≥ 20	≥ 20	≥ 20	≥ 20
0.5	≤ 6, ≥ 100	≤ 6, ≥ 100	≤ 4, ≥ 200	≤ 4, ≥ 200	≤ 3, ≥ 200	≤ 3, ≥ 200	≤ 2, ≥ 200	≤ 2, ≥ 200	≥ 200
0.7	≤ 30, ≥ 300	≤ 30, ≥ 300	≤ 20, ≥ 300	≤ 20, ≥ 300	≤ 10, ≥ 400	≤ 10, ≥ 400	≤ 10, ≥ 400	≤ 10, ≥ 500	≤ 8, ≥ 500
0.9	≥ 2	≥ 2	≥ 2	≥ 2	≥ 2	≥ 2	≥ 2	≤200, ≥400	≤200, ≥400
1	≥ 2	≥ 2	≥ 2	≥ 2	≥ 2	≥ 2	≥ 2	≥ 2	≥ 2
1.1	≥ 2	≥ 2	≥ 2	≥ 2	≥ 2	≥ 2	≥ 2	≤200, ≥400	≤200, ≥400
1.5	≤ 10, ≥ 300	≤ 10, ≥ 400	≤ 10, ≥ 400	≤ 10, ≥ 400	≤ 10, ≥ 400	≤ 10, ≥ 400	≤ 10, ≥ 400	≤ 10, ≥ 400	≤ 10, ≥ 400
2	≤ 5, ≥ 100	≤ 5, ≥ 200	≤ 5, ≥ 200	≤ 5, ≥ 200	≤ 5, ≥ 200	≤ 5, ≥ 200	≤ 5, ≥ 200	≤ 5, ≥ 200	≤ 5, ≥ 200
3	≤ 2, ≥ 50	≤ 2, ≥ 50	≤ 2, ≥ 60	≤ 2, ≥ 60	≤ 2, ≥ 60	≤ 2, ≥ 70	≤ 3, ≥ 70	≤ 3, ≥ 70	≤ 3, ≥ 70
5	≥ 30	≥ 30	≥ 30	≥ 30	≥ 40	≥ 40	≥ 40	≥ 40	≥ 40
8	≥ 20	≥ 30	≥ 30	≥ 30	≥ 30	≥ 30	≥ 30	≥ 30	≥ 30
10	≥ 20	≥ 20	≥ 20	≥ 20	≥ 20	≥ 20	≥ 20	≥ 20	≥ 20

VI. CONCLUDING REMARKS

It is observed from the results given in Section 4 that, for values of θ not close to the point of super-efficiency, namely θ = 1, $\breve{\theta}_n(d)$ performs well compared to the MLE, even for moderate sample sizes like 50 and above. This is true for d = 0.1 to 0.9 taken at intervals of 0.1 difference. Actually, the performance of $\breve{\theta}_n(d)$ and the MLE are very similar when the true θ shifts away from unity, because under such 'large' shifts, the SEE and MLE would most likely be almost equal in value. This is behavior is natural because of the way in which the SEE's $\breve{\theta}_n(d)$ have been defined in (2.1).

For values of θ closer to the point of super-efficiency (θ = 1) which are within 10% deviation from 1, we find that the SEE's $\breve{\theta}_n(d)$ perform quite well for any sample of size not less than two for 'd' ranging from 0.3 to 0.9. Even the other two choices namely d = 0.1 and 0.2, the SEE's overtake the MLE for all sample sizes except possibly over the range 200 to 400. Even for other θ values, the sample size required to prefer the SEE's is not very huge as evident in Table 4.10. This is an encouraging phenomenon as an investigator does not need a very large sample even when the value of θ is 'far away' from the point of super-efficiency namely 1. The use of $\breve{\theta}_n(d)$ instead of the MLE is thus found to be rewarding even for moderately large samples of size about 500.

As a prospective application of using SEE's, consider a manufacturing process with a normally distributed quality characteristic. Even when the process mean is in control, if the process dispersion is not in control, especially when there is an upward shift in the process dispersion parameter, it presents a case of inability to meet quality specifications. A 10% to 50% upward deviation in the process variance from an existing long-term variance needs a sample of size exceeding 300 or 400 depending on the choice of'd'. This is not too prohibitive a requirement.

A study of the 'bias' columns in Tables 4.1 to 4.9, shows that generally, the bias in $\breve{\theta}_n(d)$ approaches zero as 'n' increases. We also note that, for smaller sample sizes for which the gain function is positive, the bias in the SEE's are non-zero. That is, for smaller sample sizes, we get biased super-efficient estimators but with a lower mean-square error compared to that of the MLE.

A closer examination of Table 4.10 brings out that the SEE $\breve{\theta}_n(0.5)$ is preferable over the other SEE's considered in terms of the range of sample sizes wherein it overtakes the MLE and in its moderate 'tuning' of the estimate of θ unlike other choices of'd'. For instance, when the MLE turns out to be close to the super-efficiency point '1', which is where the SEE comes into play, the choice of d = 0.1 induces a 'big' difference in the estimate and moves it too close to 1 while, d = 0.9 does not make 'much' difference with the MLE and effectively presents a scenario where the MLE itself can be taken instead of the SEE. Thus, $\breve{\theta}_n(0.5)$ is recommended as the SEE preferable over the MLE and other SEE's considered.

REFERENCES

[1] Bahadur, R. R. (1983). Hodges Super-Efficiency. In: *Encyclopedia of Statistical Sciences*. 3, John-Wiley, 645-646

[2] Basu, D. (1952). Unpublished Thesis

[3] Durairajan, T.M. (2012).Sub-score and Super- Efficient Estimator. In: Navin Chandra and Gopal, G., eds. *Applications of Reliability Theory and Survival Analysis*, Bonfring, 120-129

[4] Hodges, J. L. (1951). *Unpublished*

[5] Le Cam, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Baye's estimates. *University of California Publications in Statistics*. 1, 277-330.

[6] Le Cam, L. (1956). On the asymptotic theory of estimation and testing hypotheses. In: Neyman, J. ed. *Proceedings of the Third Berkeley Symposium on Mathematics and Probability*. 1, University of California Press, Berkeley, 129-156

[7] Le Cam, L. (1960). Locally asymptotically normal families of distributions. *University of California Publications in Statistics*. 3, 37-98

[8] Le Cam, L. (1972). Limits of experiments. In: *Proceedings of the Sixth Berkeley Symposium on Mathematics and Probability*. 1, University of California Press, Berkeley, 245-261

[9] Sethuraman, J. (2004). Are Super- Efficient Estimators Super-powerful?. *Communications in Statistics-Theory and Methods*. 33(9), 2003-2013

[10] Sivasakthi, K., Durairajan, T. M., and William, M. L. (2017), Obtaining Super-Efficient Estimators for a Real-Valued Parameter: Delta Method. *International Journal of Applied Mathematics and Statistical Sciences*, 6(1), 81-88.

[11] Sivasakthi, K., Sakthivel, R. and William, M. L. (2017). Preference for a class of Super-Efficient Estimators of the normal mean: A study on sample size requirement. *International Journal of Statistics and Applied Mathematics*, 2(6), 241-249

[12] Stein, C. (1956). 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*. 1, University of California Press, Berkeley, 197-206

