An Efficient Estimator Improving the Searls' Normal Mean Estimator for Known Coefficient of Variation

Ashok Sahai

Department of Mathematics & Statistics, Faculty of Science & Technology, St. Augustine Campus The University of the West Indies, Trinidad & Tobago; WEST INDIES E-mail: sahai.ashok@gmail.com

Abstract: This paper addresses the issue of finding an optimal estimator of the normal population mean when the coefficient of variation is known and is expected to be rather high, as per the pilot surveys of the population at hand. The paper proposes an "Efficient Estimation Improving the well-known Searles' Normal Mean Estimator". The 'Relative Efficiencies [as compared to the usual unbiased sample-mean estimator \overline{y}] estimator per the proposed strategy has no simple algebraic form, and hence is not amenable to an analytical study determining its relative gainfulness, as compared to the usual unbiased sample mean estimator. Nevertheless, we examine these relative efficiencies of our estimator with respect to the usual unbiased estimator \overline{y} , using an illustrative simulation study with high replication. *MATLAB R2013a* is used in programming this illustrative "Simulated Empirical Numerical Study".

Keywords: MVUE, MMSE, Complete Sufficient Statistic, Simulated Empirical Numerical Study

I. INTRODUCTION

This paper addresses the issue of finding an optimal estimator of the normal population mean when the coefficient of variation is known and is expected to be rather high, as per the pilot surveys of the population at hand. This might well happen to be the case particularly if the population mean ' θ ' is expected to be positive even though the population standard deviation, say ' σ ' happens to be relatively large. Such cases are known to be relevant practically in the statistical modeling applications in the areas of astronomy, stock-markets, biodiversity, and environmental sciences, etc. Searls (1964) [7], Khan (1968) [3], Gleser & Healy (1976) [2], Arnholt & Hebert (1995) [1], Miodrag M Lovric & Sahai, Ashok (2011), [4], Sahai, Ashok (2011) [5], Winston A. Richards, Robin Antoine, Ashok Sahai, and M. Raghunadh Acharya (2010) [9], Skrepnek, Grant H; Sahai, Ashok; Antoine, Robin (2015) [10], and Wright, Kimberly & Raghunadh M. Acharya (2009) [11] considered the problem of estimating the normal population mean and variance, when its coefficient of variation (c. v.) is known/unknown. In the context of this paper, the reader would benefit by perusing [6] Samuel-Cahn, E. (1994), "Combining Unbiased Estimators", The American Statistician, 48, 34-36 & Searls, Donald T. & Intarapanich, P. (1990) [8]. It is very well known that the Searles (1964) estimator for the normal population mean gets to be:

Say, SE =
$$\overline{y} / (1 + a^2/n)$$

Wherein, 'a' is the Known Coefficient of Variation (σ/θ) [≥ 1] & 'n' is the size of the random sample [> 30] from the Normal population N (θ , $a^2.\theta^2$).

Vol. – 1, Issue – 4 August – 2015

Incidentally, Winston A. Richards, Robin Antoine, Ashok Sahai, and M. Raghunadh Acharya (2010) considered the sample counterpart of "Searls' Estimator of efficient estimation. Also, Miodrag M Lovric & Ashok Sahai (2011) considered using the sample coefficient of variation for efficient estimation of Normal population variance.

The major fact is that the statistic (\bar{y}, S^2) ; Wherein, $S^2 = \sum_{i=1}^{i=n} (y_i - \bar{y})^2 / (n-1) \equiv \text{Sample-Variance}$

{Well-known Estimator of ' σ^2 '} based on the random sample of size 'n; from the Normal population N (θ , a^2 . θ^2) is sufficient but NOT complete for { θ , σ^2 }; rendering the unavailability of the "Rao-Blackwellization". Hence, essentially, we are groping in dark for the 'UMVU' & 'MMSE' estimators for ' θ ', or any function of ' θ ', as such.

II. THE PROPOSITION OF "PROPOSED UNBIASED (EFFICIENT) ESTIMATOR OF NORMAL MEAN (PPDUEFFESTROMEAN)" FOR INCREASING THE RELATIVE EFFICIENCY" OF THE ORDINARY SEARLE'S ESTIMATOR (OSE)

We start by considering the "Minimum Mean Square (MMSE)/Ordinary Searle's Estimator (OSE)", using the known Coefficient of Variation ('C.V.') 'a' \equiv ' σ/μ ' \equiv ' σ/θ ', and the Usual Unbiased Estimator Sample Mean ' \overline{y} '.

The well-known Searles Estimator of the 'Normal mean' with known coefficient of variation 'a' is as follows:

$$SE = \overline{y} / (1 + a^2/n) \equiv OSE \equiv SEARLESESTROMEAN; Say.$$
 (2.1)

Wherein, $S^2 = \sum_{i=1}^{i=n} (y_i - \overline{y})^2 / (n-1) \equiv \text{Sample-Variance } \{\text{Well-known Estimator of '}\sigma^2^2 \}.$ (2.2)

The 'Relative Efficiency of any Estimator "•" (Relative to the Usual Unbiased Estimator Sample Mean ' \overline{y} ')' are defined as:

 $\operatorname{Reff}(\bullet) = [\operatorname{MSE}(\bullet)/\operatorname{V}(\overline{x})] \times 100\% \text{ [Defined in \%]}.$ (2.3)

Preparatory to the proposition of our "Proposed Unbiased (Efficient) Estimator Of Normal Mean (PPDUEFFESTROMEAN)", we note an elementary result:

Lemma 2.1: For a random sample of size 'n' from the Normal Population N (θ , a^2 , θ^2); say, $y_1, y_2, y_3, \dots, y_n$ we have E [S] = C* $\sigma \equiv C$ *(a. θ); wherein, C = $\sqrt{\frac{2}{(n-1)}} * \Gamma\{n/2\} / \Gamma\{(n-1)/2\}$. (2.4)

Hence, an 'Unbiased Efficient Estimator' of ' θ ': "S/{C.a)}".

Proof: It follows from the well-known fact that $\{(n-1) * S^2 / \sigma^2\} \approx \chi^2_{(n-1)}$.

Q. E. D.

In view of the above simple result our "Proposed Unbiased (Efficient) Estimator of Normal Mean (PPDUEFFESTROMEAN)" gets to be as follows:

Vol. – 1, Issue – 4 August – 2015

PPDUEFFESTROMEAN = S/{C.a}. Wherein; $a = \sigma/\theta$, $S^2 = \sum_{i=1}^{i=n} (y_i - \overline{y})^2 / (n-1)$ [S = Sqrt (S²)], and C = $\sqrt{\{2/(n-1)\}} * \Gamma\{n/2\} / \Gamma\{(n-1)/2\}$. (2.5)

III. THE SIMULATION EMPIRICAL NUMERICAL STUDY

In the preceding section, it is apparent, that the extent of 'Relative Gain in Efficiency of Estimation' will be algebraically rather too intricate, as the issue will depend on the values of the parameters like 'n', 'a', ' θ ', and ' σ '. Consequently, the answer to the question as to what is the relative gain/achievement in pursuing the "Proposed Unbiased (Efficient) Estimator Of Normal Mean (PPDUEFFESTROMEAN)" of the normal population mean lies in trying to know it through an illustrative "Simulation Empirical Numerical Study", as is attempted in this section. These 'Relative Efficiencies (Relative to the Usual Unbiased Estimator Sample Mean ' \overline{y} '); Say Reff (•)'s have been calculated for *SIX* illustrative values of the 'sample size n': 35, 50, 75, 100, 150 & 200 & *THREE* illustrative values of the 'Population Standard Deviation σ ': 8, 11 & 15.

We could assume that, without any loss of generality [Using the translation of the Parent Normal Population Data by "10- \overline{y} " without changing the Population Standard Deviation ' $\sigma \equiv a.\theta$ '], & for the simplicity of the illustration, the NORMAL population mean $\theta = 6$ [i.e. positive]. The values of the actual MSE's are calculated by considering the random samples of size "n" using '55,555' replications (pseudo-random normal samples of the size 'n') for the TWO estimators as also for the usual unbiased estimator, namely the sample mean " \overline{y} ". Hence values of Reff (•)'s calculated as per (2.3). These Reff (•)'s reported to the closest three decimal places of their respective actual values in three tables in the APPENDIX. *MATLAB R2013a* is used in programming the calculations in this illustrative "Simulated Empirical Numerical Study".

IV. CONCLUSIONS

As expected, the "Relative Efficiency" of the proposed "Proposed Unbiased (Efficient) Estimator Of Normal Mean (PPDUEFFESTROMEAN)" estimator of the "Normal Population Mean" is way above that of the "Minimum Mean Square (MMSE)/Ordinary Searle's Estimator (OSE)" for all sample sizes, as also for all the illustrative value-combinations: $\{\theta, \sigma\}$ "!

As noted earlier, it is very significant to note again that, the fact that the "Rao-Blackwellization" is unavailable in the absence of a complete-sufficient statistic for ' θ ', and hence, essentially, we are groping in dark for the 'UMVU' & 'MMSE' estimators for ' θ ', or any function of ' θ ', as such. This fact has been seminal to the motivational zeal for us to look for a more efficient estimator of the "Normal Population Mean when Y ~ N (θ , a². θ ²)". Our proposed estimator is unbiased, but NOT UMVUE. We had a fulfilling success, but it could NOT be ruled out that the betterment is of our proposition is ruled and it is an open problem to work out, if at all, an "UMVU"/UMMSE" estimator of ' θ ', and we are working on it.

V. REFERENCES

- [1] Arnholt, A. T. & Hebert, J. E. (1995), "Estimating the Mean with known Coefficient of Variation", The American Statistician, 49, 367-369.
- [2] Gleser, L. J., & Healy, J.D. (1976), "Estimating the Mean of a Normal Distribution with Known Coefficient of variation", Journal of the American Statistical Association, 71, 977-981.
- [3] Khan, R. A. (1968), "A Note on Estimating the Mean of a Normal Distribution with Known Coefficient of Variation", Journal of the American Statistical Association, 63, 1039-1041.
- [4] Miodrag M Lovric & Ashok Sahai (2011). "An Iterative Algorithm for Efficient Estimation of Normal Variance Using Sample Coefficient of Variation", InterStat. # 001.
- [5] Sahai, Ashok (2011), "Efficient Estimator of Population Variance of Normal Distribution with Known Coefficient of Variation", InterStat. # 001.
- [6] Samuel-Cahn, E. (1994), "Combining Unbiased Estimators", The American Statistician, 48, 34-36.
- [7] Searls, Donald T. (1964), "The Utilization of a known Coefficient of Variation in the Estimation Procedure", Journal of the American Statistical Association, 59, 1225-1226.
- [8] Searls, Donald T. & Intarapanich, P. (1990), "A Note on an Estimator for the variance that Utilizes the Kurtosis", The American Statistician, 44, 295-29.
- [9] Winston A. Richards, Robin Antoine, Ashok Sahai, and M. Raghunadh Acharya (2010),"On Efficient Iterative Estimation Algorithm Using Sample Counterpart of the Searles' Normal Mean Estimator with Exceptionally Large But Unknown Coefficient of Variation", InterStat. # 003.
- [10] Skrepnek, Grant H; Sahai, Ashok; Antoine, Robin (2015), "An Improved Minimum Mean Squared Error Estimate of the Square of the Normal Population Variance Using Computational Intelligence", Journal of Applied Mathematics and Bioinformatics 5.1, pp: 53-65.
- [11] Wright, Kimberly & Raghunadh M. Acharya (2009), "On Efficient Variance Estimation for Normal Populations", InterStat; November.

APPENDIX: TABLE A.1

RELATIVE EFFICIENCIES [W.R.T ' \overline{y} '] OF SERALES' & OUR PROPOSED MEAN ESTIMATOR {n = 35; $\mu = 6$ }.					
SEARLESESTROMEAN	102.842812	103.604273	105.553373	107.678882	111.014761
PPDUEFFESTROMEAN	184.053381	251.894264	328.847293	515.505123	753.336835
RELATIVE EFFICIENCIES [W.R.T ' \overline{y} '] OF SERALES' & OUR PROPOSED MEAN ESTIMATOR {n = 50; $\mu = 6$ }.					
SEARLESESTROMEAN	101.915640	102.340544	103.324674	105.496001	107.799412
PPDUEFFESTROMEAN	190.672473	253.956034	340.339732	522.976796	776.477966
RELATIVE EFFICIENCIES [W.R.T ' \overline{y} '] OF SERALES' & OUR PROPOSED MEAN ESTIMATOR {n = 75; $\mu = 6$ }.					
PPDUEFFESTROMEAN	101.333736	102.082503	102.269609	103.796095	105.304558
PPDUEFFESTROMEAN	192.054502	261.810433	342.939065	531.333762	770.242196
RELATIVE EFFICIENCIES [W.R.T ' \overline{y} '] OF SERALES' & OUR PROPOSED MEAN ESTIMATOR {n = 100; $\mu = 6$ }.					
SEARLESESTROMEAN	100.958292	101.334953	101.560984	102.857766	104.031777
PPDUEFFESTROMEAN	193.556375	262.575924	345.240080	539.295122	783.812514
RELATIVE EFFICIENCIES [W.R.T ' \overline{y} '] OF SERALES' & OUR PROPOSED MEAN ESTIMATOR {n = 150; $\mu = 6$ }.					
SEARLESESTROMEAN	100.545548	100.773835	101.232775	101.878231	102.846701
PPDUEFFESTROMEAN	199.009632	268.382665	343.241153	544.173063	796.635144
RELATIVE EFFICIENCIES [W.R.T ' \overline{y} '] OF SERALES' & OUR PROPOSED MEAN ESTIMATOR {n = 200; $\mu = 6$ }.					
SEARLESESTROMEAN	100.413472	100.583801	100.880124	101.276540	101.923958
PPDUEFFESTROMEAN	196.813933	265.836192	355.146131	542.909294	791.269984
σ 's: $\rightarrow \rightarrow \rightarrow$	$\sigma = 6 \uparrow \uparrow \uparrow$	$\sigma = 7 \uparrow \uparrow \uparrow$	$\sigma = 8 \uparrow \uparrow \uparrow$	$\sigma = 10 \uparrow \uparrow \uparrow$	$\sigma = 12 \uparrow \uparrow \uparrow$