Available online at www.elixirpublishers.com (Elixir International Journal)

Discrete Mathematics

Elixir Dis. Math. 88 (2015) 36412-36414

ISSN: 2229-712X

Efficient Minimum Mean Square Error (MMSE) Estimation Using Sample Standard Deviation Improving Searls' Usual MMSE Mean Estimator for Normal Population with Known Coefficient of Variation

Raghunadh M. Acharya[1] and Ashok Sahai[2,*]

[1]Director; Decision Tree Pvt. Ltd; Hyderabad. India.

[2]Department of Mathematics & Statistics; Faculty of Science & Technology. St. Augustine Campus; The University of The West Indies. Trinidad & Tobago; West Indies.

ARTICLE INFO

Article history:

Received: 08 August 2015;
Received in revised form:
10 November 2015;
Accepted: 16 November 2015;

Keywords

MVUE,
MMSE,
Complete Sufficient Statistic,
Simulated Empirical Numerical Study.

ABSTRACT

This paper is aimed at finding an optimal estimator of the normal population mean when the coefficient of variation. The paper proposes an "Efficient Minimum Mean Square Error (MMSE) Estimation Approach [Using 'Sample Standard Deviation', for improving the well-known Searls' Usual MMSE Mean Estimator (1964)] for Normal Population with Known Coefficient of Variation". The 'Relative Efficiencies [as compared to the usual unbiased sample-mean estimator $\bar{y}$] estimator, per the proposed approach, 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 the relative efficiency of our estimator with respect to the usual unbiased estimator $\bar{y}$, using an illustrative simulation study with high replication. *MATLAB R2013a* is used in programming this illustrative "Simulated Empirical Numerical Study".

Introduction

This paper is aimed at finding an optimal estimator of the normal population mean when the coefficient of variation is known. This might well happen to be the case particularly if the population mean 'θ' is assumed to be positive. Let the population standard deviation be denoted as 'σ' (≡ a.θ; 'a' being the known population coefficient of variation). 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 = $\bar{y}/(1+a^2/n) \equiv \bar{y}/D$; wherein, $(1+a^2/n) = D$; Say. (1.1)

In (1.1), 'a' is the Known Coefficient of Variation (σ/θ) > {1/2} & 'n' is the size of the random sample [> 30] from the Normal population N $(\theta, a^2.\theta^2)$.

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,

$$S^2 = \sum_{i-1}^{i=n} (y_i - \bar{y})^2 / (n-1) \equiv \textit{SampleVariance}$$

{Well-known Estimator of 'σ^2'} based on the random sample of size 'n'; from the Normal population N $(\theta, a^2.\theta^2)$ is sufficient but NOT complete for $\{\theta, \sigma^2\}$; 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!

The Proposition of an "Efficient Minimum Mean Square Error (MMSE) Estimation Approach Using Sample Standard Deviation improving the well-known Searls' (1964) MMSE Mean Estimator for Normal Population with Known Coefficient of Variation"

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

The Usual (well-known) Searles Estimator (1964) of the 'Normal mean' with known coefficient of variation 'a' is as follows:

SE = $\bar{y}/(1+a^2/n)$ ≡ **OSE** ≡ **USEARLESESTROFMEAN; Say** ≡ $\bar{y}$/D; D = $(1+a^2/n)$. (2.1)

Wherein,

$$S^2 = \sum_{i-1}^{i=n} (y_i - \bar{y})^2 / (n-1) \equiv SampleVariance \sim \{\text{Well-known Estimator of 'population variance'}\}. \quad (2.2).$$

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

Reff (•) = [V ($\bar{x}$)/MSE (•)] x100% [Defined in %]. (2.3)

To prepare for the proposition of our **"Efficient Minimum Mean Square Error (MMSE) Estimation Approach Using 'Sample Standard Deviation' improving the well-known Searls' (1964) MMSE Mean Estimator (Say PPDSSDMMSEESTROFMEAN) for Normal Population with Known Coefficient of Variation",** we note some elementary results:

Lemma 2.1

For a random sample of size 'n' from the Normal Population N (θ, $a^2.\theta^2$); say, $y_1, y_2, y_3, \ldots y_n$ we have: (i) E [1/S] = UCC/σ ≡ UCC/ (a. θ); wherein, UCC = $\sqrt{\{(n-1)/2\}} * \Gamma\{n/2-1\}/\Gamma\{(n-1)/2\}$. (2.4)

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

AND (ii) E ($1/S^2$) = CC2/ (a^2. θ^2); wherein, CC2 = (n-1)/ (n-3). (2.5)

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

Q.E.D.

Lemma 2.2

For a random sample of size 'n' from the Normal Population N (θ, $a^2.\theta^2$); say, $y_1, y_2, y_3, \ldots y_n$ we have: (i) E [$(\bar{y})^2$] = (1 + a^2/n).θ^2 (2.6)

AND (ii) E [$(\bar{y})^4$] = (1 + 6. (a^2/n) + 3. (a^4/n^2)). θ^4 ≡ CCC. θ^4; Say. (2.7)

Proof: It follows from the well-known fact that $\bar{y} \approx$ N (θ, a^2. θ^2/n).

Q.E.D.

In view of the above simple results our **"Efficient Minimum Mean Square Error (MMSE) Estimation [Approach Significantly Improving over the Searls' Usual Mean Estimator for Normal Population with Known Coefficient of Variation] of Normal Mean (PPDSSDMMSEESTROFMEAN)" gets to be derived as follows.**

In the class of estimators "M. {$(\bar{y})^2$/S}" [S =√ (S^2) ≡ 'Sample Standard Deviation'], the 'MMSE' estimator gets to be the one with the 'Optimal Value; Say M* of 'M', as derived below, using (2.4) to (2.7):

Let, Ø (M) = E [M. {$(\bar{y})^2$/S} – θ]^2 ≡ MSE (M.{$(\bar{y})^2$/S})

The "Normal Equation" for the Stationary point of Ø (M): δ Ø (M)/δM= 0 yields: →
M* = {CCC*UCC*a}/CC2; wherein a = σ/θ; 'CCC', 'UCC' & 'CC2' are as in (2.4), (2.5) & (2.7). (2.8)

The aforesaid is seminal to our proposed MMSEESTROFMEAN of the Normal mean 'θ', using design-parameter 'M':
Say, PPDSSDMMSEESTROFMEAN = M*. {$(\bar{y})^2$/S}; (2.9)

Wherein; a = σ/θ,

$$S^2 = \sum_{i-1}^{i=n} (y_i - \bar{y})^2 / (n-1) [S \equiv Sqrt(S^2)]$$

and, 'CCC', 'UCC' & 'CC2' are as in (2.4), (2.5) & (2.7)

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 **"Efficient Minimum Mean Square Error (MMSE) Estimation [Approach using 'Sample Standard Deviation' improving the well-known Searls' Mean Estimator for Normal Population with Known Coefficient of Variation] of Normal Mean (PPDSSDMMSEESTROFMEAN)" is pursued** 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 '$\bar{y}$') for the TWO estimators, namely **'USEARLESESTROFMEAN' in (2.1) & for our proposed estimator 'PPDSSDMMSEESTROFMEAN' in (2.9).**

Say **Reff (•)'s** have been calculated for ***SIX*** illustrative values of the '**sample size n': 35, 50, 75, 100, 150 & 200 & *FIVE*** illustrative values of the '**Population Standard Deviation σ ': 4, 5, 6, 7 & 8.**

We could assume that, without any loss of generality [Using the translation of the Parent Normal Population Data by **"6-$\bar{y}$"** without changing the Population Standard Deviation **'σ ≡ a.θ'**], & for the simplicity of the illustration, the NORMAL population mean **θ = 6** [i.e. positive]. The values of the **actual MSE's** are calculated by considering the random samples of size "n" using ***'55,555', a***

large number of replications (pseudo-random normal samples of the size 'n') for the TWO estimators as also for the usual unbiased estimator, namely the sample mean "$\bar{y}$".

Hence values of **Reff (•)'s are calculated, as per (2.3).** These **Reff (•)'s** are reported to the closest ***SIX*** 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".**

Conclusions

As expected, the "Relative Efficiency" of the proposed **"Efficient Minimum Mean Square Error (MMSE) Estimation [Approach Significantly Improving over the Searls' Usual Mean Estimator for Normal Population with Known Coefficient of Variation] of Normal Mean (PPDSSDMMSEESTROFMEAN)"** estimator of the "Normal Population Mean" is way above that of the "**Usual Minimum Mean Square (MMSE)/Ordinary Searle's Estimator (OSE)/USEARLESESTROFMEAN" for all sample sizes, as also for all the illustrative value-combinations:'{θ, σ}'!**

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, $a^2.\theta^2$)". We had a fulfilling success, but it could NOT be ruled out that the betterment is of our proposition is achievable and it is an **open problem** to work out, if at all, an **"UMVU"/UMMSE" estimator of 'θ',** and we are working on it!!

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 '$\bar{y}$'] OF SERALES' & OUR IMPROVED MMSE MEAN ESTIMATORS {n = 35; μ = 6}.					
USEARLESESTROFMEAN	101.141573	101.871658	102.545554	104.022546	105.060523
PPDSSDMMSEESTROFMEAN	162.430521	182.129340	188.713634	190.350640	187.793413
RELATIVE EFFICIENCIES [W.R.T '$\bar{y}$'] OF SERALES' & OUR IMPROVED MMSE MEAN ESTIMATORS {n = 50; μ = 6}.					
USEARLESESTRORMEAN	100.910467	101.389732	102.362448	102.704135	103.822909
PPDSSDMMSEESTROFMEAN	168.448583	191.744097	200.331512	203.216192	205.262145
RELATIVE EFFICIENCIES [W.R.T '$\bar{y}$'] OF SERALES' & OUR IMPROVED MMSE MEAN ESTIMATORS {n = 75; μ = 6}.					
USEARLESESTRORMEAN	100.648765	100.927734	101.342330	101.773545	102.423469
PPDSSDMMSEESTROFMEAN	172.829340	198.855408	212.922654	219.954282	219.607345
RELATIVE EFFICIENCIES [W.R.T '$\bar{y}$'] OF SERALES' & OUR IMPROVED MMSE MEAN ESTIMATORS {n =100; μ = 6}.					
USEARLESESTRORMEAN	100.445578	100.768111	101.033125	101.538568	101.716395
PPDSSDMMSEESTROFMEAN	177.059633	205.901868	218.590794	228.156940	230.319525
RELATIVE EFFICIENCIES [W.R.T '$\bar{y}$'] OF SERALES' & OUR IMPROVED MMSE MEAN ESTIMATORS {n = 150; μ = 6}.					
USEARLESESTRORMEAN	100.287921	100.364340	100.776003	100.908517	101.253608
PPDSSDMMSEESTROFMEAN	182.795983	208.705748	225.983823	237.065900	242.873966
RELATIVE EFFICIENCIES [W.R.T '$\bar{y}$'] OF SERALES' & OUR IMPROVED MMSE MEAN ESTIMATORS {n = 200; μ = 6}.					
USEARLESESTRORMEAN	100.263177	100.334355	100.447803	100.580389	100.796859
PPDSSDMMSEESTROFMEAN	183.363022	214.230063	230.926299	242.702044	249.459199
σ's: →→→	*σ =4* ↑↑↑	*σ =5* ↑↑↑	*σ =6* ↑↑↑	*σ =7* ↑↑↑	*σ =8* ↑↑↑