International Journal of Inventive Engineering and Sciences (IJIES)
ISSN: 2319–9598, Volume-2, Issue-6, May 2014

Improved of Stein-Type Estimator of the Variance of Normal Distribution via Shrinkage Estimation

Abbas Najim Salman, Bayda Atiya Khalaf, Haydar Sabeeh Kalash

Abstract: *This paper concerned with pre- test single stage shrinkage estimator for estimating the variance (σ^2) of normal distribution N(μ,σ^2), when a prior estimate (σ_o^2) about (σ^2) is available from the past experiences or similar cases as well as the mean is known (say μ_o) , by using Stein-type estimator , shrinkage weight factor ψ(.) and pre-test region R. Expressions for Bias , Mean Squared Error and Relative Efficiency of the proposed estimator are derived. Conclusions and numerical results are presented for Relative Efficiency and Bias Ratio. Comparisons were made with the existing estimators.*

Index Terms: *Normal Distribution, Stein-Type Estimator, Single Stage Shrinkage Estimator, Prior Estimate, Bias Ratio, Mean Squared Error and Relative Efficiency.*

I. INTRODUCTION

The normal distribution plays a very important role in statistical theory and methods. The problem of estimating variance plays a significant role in solving the allocation problem in stratified random sampling, particularly in Neyman allocation, giving a quite good fit for the failure time data in life testing and reliability problems and many more[10]. Assume that $x_1, x_2, \ldots, x_n$ be a random sample of size (n) from a normal population with known mean μ(say μ_o) and unknown variance (σ^2). In conventional notation, we write X~N(μ,σ^2). The probability density function (p. d. f.) of X is given by:

$$f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{for } -\infty<x<\infty, -\infty<\mu<\infty, \sigma^2>0 \qquad \ldots(1)$$

where μ being the population mean acts as a location parameter and σ^2 being the population variance acts as a scale parameter. For a complete sample $S^2=\sum_{i=1}^{n}(x_i-\bar{x})^2/(n-1)$, is the minimum variance unbiased estimator (MVUE) of σ^2, with variance $\operatorname{var}(S^2)=2\sigma^4/(n-1)$, and $\bar{x}=\sum_{i=1}^{n}x_i/n$ is the sample mean of these observations. The aim of this paper was to estimate the variance of Normal distribution when the mean is known using pre- test single stage shrinkage estimation(PTSSS) technique via study the performance of Bias, Mean Squared Error and Relative Efficiency expressions of the proposed estimator when we set up a selection of shrinkage weight factor $\psi(\sigma^2)$ and suitable region R and create comparisons of the numerical results with Stein-type estimator $\hat{\sigma}_s^2$ and with some existing studies. A numerical study is carried out to appraise these effects of proposed estimators. Shrinkage technique was introduced by Thompson [12] as follows

$$\tilde{\sigma}^2=\psi(\hat{\sigma}^2)(\hat{\sigma}^2-\sigma_o^2)+\sigma_o^2 \qquad \ldots(2)$$

where σ_o^2 is a prior estimate(initial value) about (σ^2) from the past experiences and $0\le\psi(\hat{\sigma}^2)\le1$, represents a shrinkage weight factor specifying the degree of belief in $\hat{\sigma}^2$ and 1-$\psi(\hat{\sigma}^2)$ specifying the degree of belief in σ_o^2. We used the form (2) above to estimate the variance parameter σ^2 of Normal distribution in case $\psi(\hat{\sigma}^2)$ is chosen as follows

$$\psi(\hat{\sigma}^2)=\begin{cases}\varphi_1(\hat{\sigma}^2) & ,\text{if } \hat{\sigma}^2\in R\\ \varphi_2(\hat{\sigma}^2) & ,\text{if } \hat{\sigma}^2\notin R\end{cases} \qquad \ldots..(3)$$

where R is the pre- test region of acceptance of size α for testing the hypothesis $H_0:\sigma^2=\sigma_o^2$ against the hypothesis $H_A:\sigma^2\neq\sigma_o^2$ using the test statistic $T=(n-1)\hat{\sigma}^2/\sigma_o^2$ and $\hat{\sigma}^2$ is the classical estimator of σ^2 (MLE or MVUE), then the estimator which is defined in (4) will be written as below

$$\tilde{\sigma}^2=\begin{cases}\varphi_1(\hat{\sigma}^2-\sigma_o^2)+\sigma_o^2 & ,\text{if } \hat{\sigma}^2\in R\\ \varphi_2(\hat{\sigma}^2-\sigma_o^2)+\sigma_o^2 & ,\text{if } \hat{\sigma}^2\notin R\end{cases} \qquad \ldots\ldots.(4)$$

where $\varphi_i(\hat{\sigma}^2)$, i =1,2; $0\le\varphi_i(\hat{\sigma}^2)\le1$ represents as shrinkage weight factors which may be a functions of $\hat{\sigma}^2$ or may be constants. The resulting estimator (4) is so called pre-test single stage shrinkage estimator (PTSSSE). Noted that, we assume the prior information regarding due the following reasons, [12]:

1. we believe that (σ_0^2) is close to the true value of σ^2, or
2. we fear that, (σ_0^2) may be near the true value of σ^2, i.e.; something bad happens if (σ^2)

Manuscript Received on May 2014.

Abbas N. Salman is a Professor in the Department of Mathematics, College of Education for pure Sciences / Ibn Al-Haithm - Baghdad University - Baghdad - Iraq.

Bayda Atiya Khalaf is Lecturer in the Department of Mathematics, College of Education for Pure Sciences / Ibn Al-Haithm - Baghdad University – Baghdad - Iraq.

Haydar Sabeeh Kalash is a Master of Electrical - Communication Engineer from University Technology Malaysia.

approximately equal to (σ_0^2) and we do not known about it

Several authors had studied the estimator defined in (4) for special distribution for different parameters and suitable regions (R) as well as for estimate the parameters of linear regression model. For example see [1], [2], [3], [4],[5] ,[6],[7],[8],[9]and [12].

II. STEIN-TYPE ESTIMATOR

As shown in Stein [11], the Stein-type estimator $(\hat{\sigma}_s^2)$ for the variance of normal distribution is defined below:

$$\hat{\sigma}_s^2 = \mathrm{Min}[S^2, S^{*2}], \qquad \text{...............(5)}$$

where,

$$S^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2/(n+1), \qquad \text{.........(6)}$$

and,

$$S^{*2} = \sum_{i=1}^{n}(x_i - \mu_o)^2/(n+2), \qquad \text{...........(7)}$$

In this paper, we considered pre- test single stage shrinkage estimator (PTSSSE) which is defined in (4) using Stein-type estimator $\hat{\sigma}_s^2$ instead of $\hat{\sigma}^2$ like follows:

$$\tilde{\sigma}^2 = \psi(\hat{\sigma}_s^2)\hat{\sigma}_s^2 + (1-\psi(\hat{\sigma}_s^2))\sigma_o^2, \qquad \text{..........(8)}$$

the shrinkage weight factor that we are considered in this context is as follows :

$$\psi(\hat{\sigma}_s^2) = \begin{cases} \varphi_1(\hat{\sigma}_s^2) & , \quad \text{if } \hat{\sigma}_s^2 \in R \\ \varphi_2(\hat{\sigma}_s^2) & , \quad \text{if } \hat{\sigma}_s^2 \notin R \end{cases} \qquad \text{.......(9)}$$

where

$$\varphi_1(\hat{\sigma}_s^2) = 0 \text{ and } \varphi_2(\hat{\sigma}_s^2) = k = e^{-\frac{10}{N}} \quad ,0 \le \varphi_i(.) \le 1$$

,i=1,2 and R is the pre-test region for testing the hypothesis H_o: $\sigma^2 = \sigma_o^2$ against the hypothesis H_1: $\sigma^2 \neq \sigma_o^2$ with level of significance (α) using the test statistic $T = N\hat{\sigma}_s^2/\sigma_o^2$.

i. e.; $R = \left[x^2_{1-\alpha/2,\,N-2}\left(\frac{\sigma_o^2}{N}\right),\ x^2_{\alpha/2,\,N-2}\left(\frac{\sigma_o^2}{N}\right)\right]$

Where N may be (n+1) or (n+2) depend upon $\hat{\sigma}_s^2$, $x^2_{1-\alpha/2,\,N-2}$ and $x^2_{\alpha/2,\,N-2}$ are the lower and upper 100 (α/2) percentile of Chi-Square distribution with (N-2) degree of freedom

Refer to the region R=[a,b] for simple computation.

Now, using the shrinkage weight factor above, the shrinkage estimator which is defined in (4), would be:

$$\tilde{\sigma}^2 = \begin{cases} \sigma_o^2 & , \text{ if } \hat{\sigma}_s^2 \in R \\ k(\hat{\sigma}_s^2 - \sigma_o^2) + \sigma_o^2 & , \text{ if } \hat{\sigma}_s^2 \notin R \end{cases} \qquad \text{.........(10)}$$

The shrinkage estimator mentioned above was used for estimating the variance of normal distribution $N(\mu, \sigma^2)$. Expressions for Bias , Mean Squared Error and Relative Efficiency of the proposed estimator are derived. Numerical computations have been made and presented for the Relative Efficiency and Bias Ratio. These results are compared with the existing results.

III. EXPRESSIONS FOR BIAS AND MSE OF $\tilde{\sigma}^2$

In this section, recall (PTSSSE) defined in (4) which has the following form:

$$\tilde{\sigma}^2 = \begin{cases} \sigma_0^2 & ,\text{if } \hat{\sigma}_s^2 \in R, \\ k[\hat{\sigma}_s^2 - \sigma_0^2] + \sigma_0^2 & ,\text{if } \hat{\sigma}_s^2 \notin R, \end{cases}$$

The expressions for Bias and Mean Squared Error [MSE(·)] of $\tilde{\sigma}_{SS}^2$ are represented respectively as follows:

$$\mathrm{Bias}(\tilde{\sigma}^2 \,|\, \sigma^2, R) = E(\tilde{\sigma}^2) - \sigma^2$$

$$= \int_{\hat{\sigma}_s^2 \in R} [\sigma_0^2 - \sigma^2] f(\hat{\sigma}_s^2) d\hat{\sigma}_s^2 + \int_{\hat{\sigma}_s^2 \in \bar{R}} \left[k(\hat{\sigma}_s^2 - \sigma_0^2) + (\sigma_0^2 - \sigma^2) \right] f(\hat{\sigma}_s^2) d\hat{\sigma}_s^2$$

where $\bar{R}$ is the complement region of R, and $f(\hat{\sigma}_s^2/\sigma^2)$ defined as follows:

$$f(\hat{\sigma}_s^2/\sigma^2) = \begin{cases} \dfrac{\left[N/\sigma^2\right]^{\frac{n-1}{2}} \left[\hat{\sigma}_s^2\right]^{\left(\frac{n-3}{2}\right)} \exp\left[\dfrac{-N}{2\sigma^2}\hat{\sigma}_s^2\right]}{2^{\frac{(n-1)}{2}}\,\Gamma\left[\dfrac{n-1}{2}\right]} & , \text{for } \hat{\sigma}_s^2 \ge 0, \sigma^2 \ge 0 \\ 0 & ,\text{otherwise} \end{cases} \qquad \text{....(11)}$$

The previous expression will result

$$B(\tilde{\sigma}^2 \,|\, \sigma^2, R) = \frac{\sigma^2}{N}\left\{\frac{1}{2}[N\lambda\, j_0(a_1,b_1) + j_1(a_1,b_1)] + \frac{1}{2}N[\lambda - 1]\right\} \qquad \text{...(12)}$$

Where

$$j_\ell(a_1,b_1) = \int_{a_1}^{b_1} y^\ell f(y)dy, \ell = 0,1,2. \qquad \text{...(13)}$$

$$\lambda = \frac{\sigma_0^2}{\sigma^2},\ y = \frac{N\sigma_s^2}{\sigma^2} \sim X^2_{n-1}, \qquad \text{...(14)}$$

$a_1 = \lambda X^2_{1-\alpha/2,n-1}$, $b_1 = \lambda X^2_{\alpha/2,n-1}$, ...(15)

and,

$$\mathrm{MSE}(\tilde{\sigma}^2 \mid \sigma^2, \mathrm{R}) = \mathrm{E}(\tilde{\sigma}^2 - \sigma^2)^2$$
$$= \int_{\hat{\sigma}_s^2 \in \mathrm{R}} [\sigma_0^2 - \sigma^2] \mathrm{f}(\hat{\sigma}_s^2) d\hat{\sigma}_s^2 +$$
$$\int_{\hat{\sigma}_s^2 \in \bar{\mathrm{R}}} \left[\mathrm{k}(\hat{\sigma}_s^2 - \sigma_0^2) + (\sigma_0^2 - \sigma^2) \right]^2 \mathrm{f}(\hat{\sigma}_s^2) d\hat{\sigma}_s^2$$

we conclude,

$$\mathrm{MSE}(\tilde{\sigma}^2 \mid \sigma^2, \mathrm{R}) = \frac{\sigma^4}{\mathrm{N}^2} \{ \; k^2 \{ 2\mathrm{N} + \mathrm{N}^2(1-\lambda) + 2\mathrm{N}(1-\lambda) + 3 \} +$$
$$2k\,(n-1)(\lambda-1)\{ \mathrm{N}(1-\lambda)+1 \} + \mathrm{N}^2(\lambda-1)^2 -$$
$$k^2 \{ \mathrm{J}_2(a_1,b_1) - 2\mathrm{N}\lambda \mathrm{J}_1(a_1,b_1) + \mathrm{N}^2\lambda^2 \mathrm{J}_0(a_1,b_1) +$$
$$2(\mathrm{J}_1(a_1,b_1) - \mathrm{N}\lambda \mathrm{J}_0(a_1,b_1)) + 3\mathrm{J}_0(a_1,b_1) \} -$$
$$2k\mathrm{N}(\lambda-1)\{ \mathrm{J}_1(a_1,b_1) - \mathrm{N}\lambda\, \mathrm{J}_0(a_1,b_1) + \mathrm{J}_0(a_1,b_1) \} \quad ....(16)$$

The Efficiency of $\tilde{\sigma}^2$ relative to $\hat{\sigma}_s^2$ is given by

$$\mathrm{R.Eff}\,(\tilde{\sigma}^2 \mid \sigma^2, \mathrm{R}) = \frac{\mathrm{MSE}(\hat{\sigma}^2)}{\mathrm{MSE}(\tilde{\sigma}^2 \mid \sigma^2, \mathrm{R})}$$

See [1] ,[2],[9],and [12].

IV. CONCLUSIONS AND NUMERICAL RESULTS

1. From the expressions for bias and MSE of the considered estimator $\tilde{\sigma}^2$ we can see the following:

(i) *Bias* $(\tilde{\sigma}^2 / \sigma^2, R)$ is an odd function of λ.

(ii) *MSE* $(\tilde{\sigma}^2 / \sigma^2, R)$ is an even function of λ.

(iii) The considered estimator $\tilde{\sigma}^2$ is a consistent estimator of σ^2,

$$\text{i.e.:} \lim_{n\to\infty} \mathrm{MSE}(\tilde{\sigma}^2 / \sigma^2, \mathrm{R}) = 0$$

(iv) The considered estimator $\tilde{\sigma}^2$ dominates $(\hat{\sigma}_s^2)$ With large sample size (n) in the term of *MSE*,

i.e., $\lim_{n\to\infty} [MSE(\tilde{\sigma}^2) - MSE(\hat{\sigma}_s^2)] \le 0$.

2. The computation of relative efficiency [*R.Eff*($\tilde{\sigma}^2$)=*MSE* $(\hat{\sigma}_s^2)$ /*MSE*($\tilde{\sigma}^2$)] and *Bias* ratio [*NB*($\tilde{\sigma}^2$ /σ^2)/ σ^2] of considered estimator $\tilde{\sigma}^2$ were made for different constant involved in it, some of these computation are given in the annexed table for some samples of these constant e. g. α = 0.0, 0.05, 0.1 and N = 10(10) 30 and λ = 0.25(0.25) 2.0.The following numerical results from the mentioned table were made:

(i) $\tilde{\sigma}^2$ has a maximum *R. Eff* when $\sigma^2 \approx \sigma_0^2$.

(ii) *R. Eff* of $\tilde{\sigma}^2$ is a maximum with small values of α when $\sigma^2 \approx \sigma_0^2$.

(iii) The *Bias* ratio of $\tilde{\sigma}^2$ as reasonably small when $\sigma^2 \approx \sigma_0^2$.

(iv) The *Bias* ratio of $\tilde{\sigma}^2$ increases when α increases especially when $\sigma^2 \approx \sigma_0^2$.

(v) The Effective interval [the values of λ which make the value of *R. Eff* greater than one] of the proposed estimator $\tilde{\sigma}^2$ is [0.25, 2].

3. The considered estimator $\tilde{\sigma}^2$ is better than stein-type estimator[11] and also better than some existing estimator, for example (Al-Hemyari and AL-Joboori[1], AL-Joboori[4]. Davis and Arnold[6], Hirano[7] , Pandey[8] and Pandy and Singh[9]).

we indicate that, α = 0.001 get the optimal *R. Eff*. and *Bias* ratio specially at $\sigma^2 \approx \sigma_0^2$.Also, considered estimator $\tilde{\sigma}^2$ is more efficient than the existing estimators in the terms of Minimum Mean Squared Error(MSE) specially around σ_0^2.

ACKNOWLEDGMENT

The authors wanted to provide thanks to the referees and to the editor for constructive suggestions and valuable comments which resulted in the improvement of this article.

REFERENCES

1. Al-Hemyari, Z. A. and Al-Joboori, A. N. “Pre-Test Estimator for Variance of Normal Distribution”. Pro. 7th, Conf. of Iraqi Statistic. Assoc. [1995].
2. Al-Juboori,A.N., (2002), "On Shrunken Estimators for the Parameters of Simple Linear Regression Model", Ibn-Al-Haitham J. for Pure and Applied Sciences, 15,(4A), 60-67[2002].
3. Al-Juboori,A.N., “Pre-Test Single and Double Stage Shrunken Estimators for the Mean of Normal Distribution With Known Variance”, Baghdad J. for Science, 7 (4), 1432-1442[2010].
4. Al-Juboori,A.N." Single and Double Stage Shrunken Estimators for the Variance of Normal Distribution". Journal of Education, AL-Mustansiryia. University, 2 , 597-608[2011].
5. Al-Juboori,A.N., " Estimate the Reliability Function of Exponential Distribution Via Pre-Test Single Stage Shrinkage Estimator". Accepted to publish ,Tikrit of Mathematics and Computer Journal.[2012].
6. Davis, R. L. and Arnold, J. C. ,“An Efficient Preliminary Test Estimator For The Variance Of Normal Population When The Mean is Unknown”. Biometrika, 57, 674-677 [1970].
7. Hirano, N., “Some Properties of An Estimator for the Variance of a Normal Distribution ”. Ann. Inst. Statistic. Math., 25: 479-492 [1974].
8. Pandey, B. N. “On Shrinkage Estimation of Normal Population Variance”. Comm. Statis. Theor. Meth., 4,A8, 359-345 [1979].
9. Pandey, B. N. and Singh, J. “Estimation of Variance of Normal Population Using Prior Information”. J. Ind. Statist. Assoc., 15, 141-150 [1977].
10. Singh, H. P. and Saxena, S." A class of Shrinkage Estimators for Variance of a Normal Population". Brazilian Journal of Probability and Statistics, 17, 41–56[2003].
11. Stein, C., “In Admissibility of The Usual Estimator For The Variance of A Normal Distribution With Unknown Mean”. Ann. Inst. Statist. Math. 16, 155-160. [1964].
12. Thompson, J. R., “Some Shrinkage Techniques For Estimating The Mean”. J. Amer. Statist. Assoc., 63, 113-122 [1968].

AUTHOR PROFILE

Abbas N. Salman is a Professor in the Department of Mathematics, College of Education for pure Sciences / Ibn Al-Haithm - Baghdad University - Baghdad - Iraq. Received his M.sc. in Mathematical Statistics in 1995, his B.Sc. in Mathematics Sciences from Department of Mathematics - College of Education / Ibn Al-Haithm - Baghdad University respectively .P.O. Box :4150 - AL-Adhamiyah - Anter Square - Baghdad- Iraq - Street: 20.

Bayda Atiya Khalaf is Lecturer in the Department of Mathematics, College of Education for Pure Sciences / Ibn Al-Haithm - Baghdad University – Baghdad - Iraq. Received her M.Sc. in Mathematical Statistics in 2008, her B.Sc. in Mathematical Sciences from Department of Mathematics - College of Education/Ibn Al-Haithm - Baghdad University. Currently she is studying Ph.D. in the department of Mathematics - Faculty of Sciences in university Putra Malaysia - Kuala Lumpur.

Haydar Sabeeh Kalash is a Master of Electrical - Communication Engineer from University Technology Malaysia. He has bachelor with owner of Hardware and Network Engineer from Coventry University - UK. He has a long industrial technical support experience about seven years and teaching experience more than five years in the area of Communication and Network Engineering. Currently he is a Lecturer in the Computing Faculty in Gulf College, Sultan of Oman, Muscat.

Table [1]: Shown Bias Ratio B(.) and Relative Efficiency R.ff(.) of $\tilde{\sigma}^2$ w.r.t. α , N and λ .

α	N	λ	0.25	0.5	0.75	1	1.25	1.5	1.75	2
0.01	10	B(·)	+7.5	+5	+2.5	$2.2512e^{-9}$	2.5	5	7.5	10
		R.Eff	0.35556	0.8	3.2	$2.0752e^{+15}$	3.2	0.8	0.35556	0.2
	20	B(·)	14.966	-9.9848	4.9972	$7.6796e^{-5}$	4.9991	9.9971	14.993	19.986
		R.Eff	0.17859	0.40121	1.6018	$1.7825e^{+6}$	1.6006	0.40023	0.17795	0.10014
	30	B(·)	22.158	14.823	7.4682	0.000488	7.4889	14.956	22.381	29.758
		R.Eff	0.1222	0.27305	1.0756	44172	1.0698	0.26823	0.11977	0.067747
0.05	10	B(·)	7.4981	4.9991	2.4997	$2.1188e^{-5}$	2.4998	4.9996	7.4994	9.999
		R.Eff	0.35573	0.8003	3.2008	$3.5371e^{+7}$	3.2004	0.80012	0.35562	0.20004
	20	B(·)	14.719	9.8466	4.9544	0.0011055	4.9733	9.9263	14.848	19.74
		R.Eff	0.18463	0.41247	1.6288	8383.9	1.6166	0.40585	0.18138	0.10263
	30	B(·)	21.377	14.336	7.3048	0.002843	7.3802	14.628	21.708	28.68
		R.Eff	0.13125	0.29162	1.1213	1273	1.099	0.2799	0.12712	0.072865
0.1	10	B(·)	7.4813	4.99	2.4962	$7.6927e^{-5}$	2.4972	4.9939	7.4895	9.9839
		R.Eff	0.35733	0.80319	3.2096	$2.4416e^{+5}$	3.2071	0.80195	0.35655	0.20065
	20	B(·)	14.402	9.6532	4.8775	0.0009479	4.9062	9.7626	14.552	19.289
		R.Eff	0.19278	0.42885	1.6771	1197.3	1.6573	0.41902	0.18866	0.1074
	30	B(·)	20.72	13.909	7.1143	0.0031478	7.2029	14.182	20.937	27.6
		R.Eff	0.13962	0.30935	1.1763	324.55	1.1467	0.29674	0.13638	0.07859