SOME IMPROVED SHRINKAGE TESTIMATORS FOR VARIANCE OF NORMAL DISTRIBUTION UNDER ASYMMETRIC LOSS FUNCTION

1Rakesh Srivastava and 2Tejal Shah
1Department of Statistics, The M.S. University of Baroda, Vadodara, India
2Centre for Management Studies Ganpat University
E Mail: 1rakeshsrivastava30@yahoo.co.in; 2tts01@ganpatuniversity.ac.in

Received March 21, 2016
Modified October 06, 2016
Accepted November 12, 2016

Abstract
The present paper proposes some improved shrinkage testimator(s) for the variance of a normal distribution in presence of a guess value on it. Risk properties of these have been studied under an asymmetric loss function. It has been observed that proposed testimators perform better than the conventional estimators over a fairly large range of guess value to parameter ratio. The arbitrariness in the choice of shrinkage factor has been removed by making it dependent on test statistic. We have also considered the square of shrinkage factor as a different choice of shrinkage factor and it is observed that estimator proposed using this choice performs better than the other estimators. Recommendation on the degree(s) of asymmetry and level of significance has been made.

Key Words: Normal Distribution, Variance, Shrinkage Testimator(s), Asymmetric Loss Function, Degrees of Asymmetry, Level of Significance, Relative Risk.

1. Introduction
Normal distribution occupies an important place in the study of statistics. Estimation and testimation of its parameters has drawn attention of many research workers since long. Several authors have proposed different estimators for variance of a Normal distribution such as Pandey et. al (1988) considered some shrinkage testimators for the variance under Mean Square Error (MSE) criterion.

Parsian and Farsipour (1999), Singh et. al (2002), Mishra and Meulen (2003), Pandey et. al (2004), Ahmadi et. al (2005), Prakash and Singh (2006), Singh et. al (2007) among others. All the mentioned authors have considered the estimation of variance under the LINEX loss function.

The choice of an appropriate loss function is guided by the consequences of over estimation or under estimation which may be more serious than the other in a given context. As the Squared Error Loss Function (SELF) gives equal importance to both over and under estimation, it has been established in many research papers that an asymmetric loss function is more appropriate in such situations.
(i) We have proposed two testimators for the variance of a normal distribution $\tilde{\sigma}^2_{ST1}$ when the shrinkage factor (S.F) is dependent on test statistics $k_1 = \frac{\bar{x}s^2}{\sigma_0^2x^2}$ used to test the available guess value of the variance, and

(ii) $\tilde{\sigma}^2_{ST2}$ square of the shrinkage factor (S.F) $k_2 = k_1^2$ the choice is motivated by the fact that testimators proposed by taking square of the S.F. perform better.

1.1 Utilizing guess

When a guess of the parameter of interest $\theta$ is available to the experimenter either due to past studies or his familiarity with the behavior of the population then this guess may improve the estimation procedure. Several authors have suggested so many ways of utilizing available guess value commencing from the pioneering work by Bancroft (1944) and an extensive bibliography in this context is provided by Han et. al (1988).

1.2 Asymmetric loss functions

The loss function $L(\tilde{\theta}^2, \sigma^2)$ provides a measure of financial consequences arising from a wrong estimate of the unknown quantity $\sigma^2$. As in many real life situations, particularly in insurance claims, estimating any health statistics parameter, the over-estimation and under-estimation are having different impacts. So giving ‘equal’ importance to these as the squared error loss function (SELF) does, may not be useful. Several authors such as Canfield (1970), Zellner (1986), Basu and Ebrahimi (1991), Srivastava (1996), Srivastava and Tanna (2001), Srivastava and Shah (2010) and others have demonstrated the superiority of the asymmetric loss functions, over squared error loss functions in several contexts.

A useful asymmetric loss function known as LINEX loss function was introduced by Varian (1975) which was extended by Zellner (1986). If the parameter of interest is scale parameter then Basu and Ebrahimi (1991) have proposed the following loss function:

$$L(\Delta) = b[e^{a\Delta} - a\Delta - 1] , a \neq 0, b > 0$$

where $\Delta = \left(\frac{\tilde{\sigma}^2}{\sigma^2} - 1\right)$

(1.2.1)

The sign and magnitude of ‘$a$’ represents the direction and degree of asymmetry respectively. Positive values of ‘$a$’ are suggested for situations where overestimation is more serious than the under estimation, while negative values of ‘$a$’ are recommended in reverse situations. ‘$b$’ is constant of proportionality. $L(\Delta)$ rises exponentially when $\Delta < 0$ and almost linearly when $\Delta > 0$. Hence, the loss function defined by (1.2.1) is known as LINEAR EXPONENTIAL (LINEX) loss function.

In section-2, we have defined the testimators, section-3 is devoted to the derivation of risk(s) of the proposed testimator(s), section-4 deals with the relative risk(s) of these testimators with respect to the conventional estimator(s). The paper concludes with recommendations and suggestions in section-5.

2. The shrinkage testimator(s)

Let $x_1, x_2, \ldots, x_n$ be a random sample of size $n$ from a Normal population with mean $\mu$ and variance $\sigma^2$. It is assumed that some initial (point) guess value for $\sigma^2$
is known and let it be $\sigma_0^2$, available from the past experience or some other reliable sources. It is well known that the unbiased estimate of $\sigma^2$ is $s^2$ with variance of $s^2$ as 
\[
\frac{2s^4}{v}
\] where $v = (n - 1)$ before incorporating this guess value we test the null hypothesis $H_0: \sigma^2 = \sigma_0^2$ against the alternative $H_1: \sigma^2 \neq \sigma_0^2$ using the test statistic 
\[
\frac{\frac{v s^2}{\sigma_0^2}}{\frac{a^2}{s^2}}
\] where $v = n - 1$ and $s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$ which follows $\chi^2$ distribution with $v$ degrees of freedom. Now $H_0$ may be accepted at a% level of significance if, $x_1^2 < \frac{v s^2}{\sigma_0^2} < x_2^2$ where $x_1^2$ and $x_2^2$ are lower and upper points of the uniformly most powerful unbiased test of $H_0$, then a shrinkage testimator may be proposed with shrinkage factor 
\[
k = \frac{\frac{v s^2}{\sigma_0^2}}{s^2}
\] which is inversely proportional to $x_2^2$ where $x_2^2 = x_2^2 - x_1^2$. If the data does not support $H_0$ it may be rejected and in this case it is recommended to use $s^2$, the best available (UMVUE) estimator of $\sigma^2$.

Thus, the proposed shrinkage testimator $\hat{\sigma}_2^{2 ST_1}$ of $\sigma^2$ is defined as
\[
\hat{\sigma}_2^{2 ST_1} = \begin{cases} 
  k_0 s^2 + (1 - k_0)\sigma_0^2 & \text{if } x_1^2 \leq \frac{v s^2}{\sigma_0^2} \leq x_2^2 \\
  s^2 & \text{otherwise}
\end{cases}
\]

Where $k_0 = k$

Estimators of this type with an arbitrary 'k' ($0 \leq k \leq 1$) have been proposed by Pandey and Singh (1976, 77), Srivastava (1976) and others. In all of these studies it has been shown that the shrinkage testimators perform better than the conventional estimators if 'k' is near zero, $n$ is small and $|\sigma^2 - \sigma_0^2|$ is also small. Hence, we should select the shrinkage factor which approaches to zero rapidly, with this motivation we have taken square of the S.F. and proposed another shrinkage testimator $\hat{\sigma}_2^{2 ST_2}$ of $\sigma^2$ as:
\[
\hat{\sigma}_2^{2 ST_2} = \left(\frac{\frac{v s^2}{\sigma_0^2}}{\sigma_0^2} \right)^2 s^2 + \left[1 - \left(\frac{\frac{v s^2}{\sigma_0^2}}{\sigma_0^2} \right)^2\right] \sigma_0^2 \quad \text{if } H_0 \text{ is accepted}
\]
otherwise
\]

\[
R(\hat{\sigma}_2^{2 ST_1}) = E[\hat{\sigma}_2^{2 ST_1} | L(\Lambda)]
\]
\[
= E\left[k s^2 + (1 - k)\sigma_0^2 \bigg| \begin{array}{c}
  x_1^2 < \frac{v s^2}{\sigma_0^2} < x_2^2 \\
  x_1^2 < \frac{v s^2}{\sigma_0^2} < x_2^2
\end{array}\right] \cdot P\left[\begin{array}{c}
  x_1^2 < \frac{v s^2}{\sigma_0^2} < x_2^2 \\
  x_1^2 < \frac{v s^2}{\sigma_0^2} < x_2^2
\end{array}\right]
\]
\[
+ E\left[s^2 \bigg| \begin{array}{c}
  \frac{v s^2}{\sigma_0^2} < x_1^2 \\
  \frac{v s^2}{\sigma_0^2} > x_2^2
\end{array}\right] \cdot P\left[\begin{array}{c}
  \frac{v s^2}{\sigma_0^2} < x_1^2 \\
  \frac{v s^2}{\sigma_0^2} > x_2^2
\end{array}\right]
\]

(3.1)
\[R(\theta) = \begin{cases} 1 - \frac{xa}{\lambda x} \left( \frac{y}{2} + 1 \right) \left\{ I \left( \chi \lambda \frac{y}{2} + 2 \right) - I \left( \chi \lambda \frac{y}{2} \right) \right\} \\
+ a \left\{ I \left( \chi \lambda \frac{y}{2} + 1 \right) - I \left( \chi \lambda \frac{y}{2} + 1 \right) \left( \frac{y}{2} + 1 \right) \right\} \\
- a \left\{ I \left( \chi \lambda \frac{y}{2} \right) - I \left( \chi \lambda \frac{y}{2} + \frac{y}{2} + 1 \right) \right\} \\
\left[ 1 - I \left( \chi \lambda \frac{y}{2} \right) - I \left( \chi \lambda \frac{y}{2} + \frac{y}{2} + 1 \right) + 1 \right] \end{cases}\]

(3.3)
Where

\[ I^* = \frac{e^{\alpha (3-\lambda)}}{2^{\nu/2} \Gamma(\frac{\nu}{2})} \int_{x_0^2} \frac{e^{-\frac{\nu s^2}{2\sigma^2}}}{\sigma^\nu} \, e^{-\frac{1}{2} t \nu^{-1} \sigma_{0}^2} \, dt \quad \text{and} \quad \lambda = \frac{\sigma^2}{\sigma_0^2}. \]

Again, we obtain the risk of \( \hat{\sigma}^2_{ST_2} \) under \( L(\Delta) \) with respect to \( s^2 \), given by

\[
R(\hat{\sigma}^2_{ST_2}) = E[ \hat{\sigma}^2_{ST_2} \mid L(\Delta)] = \frac{\sigma^2}{\sigma_0^2} \int_{x_0^2} \frac{e^{-\frac{\nu s^2}{2\sigma^2}}}{\sigma^\nu} \, e^{-\frac{1}{2} t \nu^{-1} \sigma_{0}^2} \, dt \quad \text{and} \quad \lambda = \frac{\sigma^2}{\sigma_0^2}. \]

\[
= e^{-a} \int_{\chi_{1/\sigma_0^2}^2} e^{\frac{\nu s^2}{\sigma^2}} (s^2 - \sigma_0^2) f(s^2) \, ds^2
\]

\[
- a \int_{\chi_{1/\sigma_0^2}^2} \left( \frac{\nu s^2}{\sigma_0^2 \chi^2} \right) \frac{(s^2 - \sigma_0^2)}{\sigma_0^2} f(s^2) \, ds^2
\]

\[
- \int_{\chi_{1/\sigma_0^2}^2} f(s^2) \, ds^2 + e^{-a} \int_{0}^{\chi_{1/\sigma_0^2}^2} e^{\frac{\nu s^2}{\sigma^2}} f(s^2) \, ds^2
\]

\[
+ e^{-a} \int_{\chi_{1/\sigma_0^2}^2} e^{\frac{\nu s^2}{\sigma^2}} f(s^2) \, ds^2 - a \int_{0}^{\chi_{1/\sigma_0^2}^2} \left( \frac{s^2}{\sigma_0^2} - 1 \right) f(s^2) \, ds^2
\]

\[
- a \int_{\chi_{1/\sigma_0^2}^2} \left( \frac{s^2}{\sigma_0^2} - 1 \right) f(s^2) \, ds^2 - \int_{0}^{\chi_{1/\sigma_0^2}^2} f(s^2) \, ds^2 - \int_{\chi_{1/\sigma_0^2}^2} f(s^2) \, ds^2
\]

\[(3.5)\]
Where \( f(s^2) = \frac{1}{2^{\nu/2} \Gamma(\nu/2)} (s^2)^{\nu - 1} e^{-\frac{1}{2} \frac{s^2}{\sigma^2}} \)

Straight forward integration of (3.5) gives

\[
R(\hat{\sigma}_{ST_2}) = \left( \frac{\sigma^2}{\nu} \right)^{\nu/2} \begin{pmatrix}
I^* - \frac{4a}{\lambda^2(\nu^2)} \left( \frac{\nu}{2} + 1 \right) \left( \frac{\nu}{2} + 2 \right) \\
\left\{ I \left( x^2 \Lambda, \frac{\nu}{2} + 3 \right) - I \left( x^2 \Lambda, \frac{\nu}{2} + 2 \right) \right\} \\
\left\{ -a \left\{ I \left( x^2 \Lambda, \frac{\nu}{2} + 1 \right) - I \left( x^2 \Lambda, \frac{\nu}{2} + 2 \right) \right\} \right\} \\
+ \frac{e^{-a}}{2^{\nu/2} (1 - \frac{a}{\nu})} \left[ 1 - I \left( x^2 \Lambda, \frac{\nu}{2} \right) - I \left( x^2 \Lambda, \frac{\nu}{2} \right) + 1 \right]
\end{pmatrix}
\]

Where \( I^* = \frac{e^{a(\Lambda + 1)}}{2^{\nu/2} \Gamma^{(\nu/2)}(\frac{\nu}{2})} \int x^2 \Lambda e^{\frac{a^2}{\Lambda^2}} \frac{a^2}{\Lambda^2} \left( e^{\frac{a^2}{\Lambda^2}} - \frac{a^2}{\Lambda^2} \right) e^{-\frac{a^2}{\nu} t \frac{e}{t} \frac{e}{t} dt} \)

Where \( I(\chi; p) = \left( \Gamma(\frac{\nu}{2}) \right) \int_0^\chi e^{-x^2} x^{\nu-1} dx \) refers to the standard incomplete gamma function.

**4. Relative Risk(s)**

A natural way of comparing the risk of the proposed testimators, is to study its performance with respect to the best available estimator \( \hat{\sigma}^2 \) in this case. For this purpose, we obtain the risk of \( s^2 \) under \( L_0(\hat{\sigma}^2, \sigma^2) \) as:

\[
R_\nu(s^2) = E[s^2 | L_0(\hat{\sigma}^2, \sigma^2)]
\]

\[
= e^{-x} \int_0^\infty \int_0^\nu f(s^2) ds^2 - a \int_0^\nu \left[ \frac{s^2}{\sigma^2} - 1 \right] f(s^2) ds^2 - \int_0^\nu f(s^2) ds^2
\]

(4.1)

Where \( f(s^2) = \frac{1}{2^{\nu/2} \Gamma(\nu/2)} (s^2)^{\nu - 1} e^{-\frac{1}{2} \frac{s^2}{\sigma^2}} \)

A straight forward integration of (4.1) gives
Some improved shrinkage testimators for variance …

\[ R_e(s^2) = \left( \frac{\sigma^2}{\nu} \right)^{\nu/2} \left[ \frac{e^{-a}}{2^{\nu/2} \left( \frac{1}{2} - \frac{a}{\nu} \right)^{\nu/2}} - 1 \right] \]  

(4.2)

Now, we define the Relative Risk of \( \hat{\sigma}^2 \), with respect to \( s^2 \) under \( L_e(\hat{\sigma}^2, \sigma^2) \) as follows:

\[ RR_1 = \frac{R_e(s^2)}{R(\hat{\sigma}^2_{ST_1})} \]  

(4.3)

Using (4.2) and (3.3) the expression for \( RR_1 \) given in (4.3) can be obtained; it is observed that \( RR_1 \) is a function of \( \lambda, \nu, \alpha \) and ‘a’. To observe the performance of \( \hat{\sigma}^2 \), we have taken several values of these parameters viz. \( \alpha = 1\%, 5\%, 10\%, \nu = 5, 8, 10, 12 \text{ and } a = -2.0, -1.0, 1.0, 1.5, 1.75 \) i.e. positive as well as negative values, as ‘a’ is the prime important factor and decides about the seriousness of over/under estimation in the real life situation further \( \lambda = 0.2(0.2)2.0 \). Similarly, we define the Relative Risk of \( \hat{\sigma}^2_2 \) with respect to \( s^2 \) under \( L_e(\hat{\sigma}^2, \sigma^2) \) as follows:

\[ RR_2 = \frac{R_e(s^2)}{R(\hat{\sigma}^2_{ST_2})} \]  

(4.4)

Using (4.2) and (3.6) the expression for \( RR_2 \) given in (4.4) can be obtained, it is observed that \( RR_2 \) is a function of \( \lambda, \nu, \alpha \) and ‘a’. To observe the behavior of \( \hat{\sigma}^2_2 \), we have taken several values of these, same as in the case of \( \hat{\sigma}^2_1 \).

Some of the graphs of \( RR_1 \) and \( RR_2 \) for the values considered above are provided in the appendix. However, our conclusions based on all the graphs are given in the next section.

5. Conclusion

We wish to compare the performance of \( \hat{\sigma}^2_{ST_1} \) and \( \hat{\sigma}^2_{ST_2} \) with respect to the best available (unbiased) estimator of \( \sigma^2 \). For this purpose we define the Relative Risks (RR) as

\[ RR_1 = \frac{\text{Risk}(\sigma^2/L(\Delta))}{\text{Risk}(\hat{\sigma}^2_{ST_1}/L(\Delta))} \]  

(5.1)

and \[ RR_2 = \frac{\text{Risk}(\sigma^2/L(\Delta))}{\text{Risk}(\hat{\sigma}^2_{ST_2}/L(\Delta))} \]  

(5.2)
It is observed that the expressions of \( RR_1 \) and \( RR_2 \) are functions of \( \nu, \alpha, \lambda \) and the degrees of asymmetry "a". For the comparison of the proposed testimators with the best available estimator we have considered the values of these as mentioned in the previous section. There will be several tables and graphs for Relative Risk(RR) values for both the testimators. We have assembled some of graphs in the appendix of the paper. However our recommendations based on all these computations are as follows:

(i) \( \hat{\sigma}^2_{ST} \) performs better than \( \hat{\sigma}^2 \) for a considerably large range of \( \lambda \) for different degrees of asymmetry. For \( a = -2 \) the range is \( 0.6 \leq \lambda \leq 1.8 \), which changes slightly for \( a = -1 \) and becomes \( 0.6 \leq \lambda \leq 1.6 \). For the positive values of 'a' we have observed that when \( 0.8 \leq \lambda \leq 1.4 \), the performance of \( \hat{\sigma}^2_{ST} \) is better than \( \hat{\sigma}^2 \). Similar pattern is observed for the other two positive values of 'a' i.e. \( a = 1.5 \) and \( a = 1.75 \). However these values of RR values are smaller in magnitude for positive values of 'a' as compared to those for the negative values.

(ii) For higher values of \( \nu \) i.e. 5% and 10% a similar kind of behaviour of RR values is observed but the range of \( \lambda \) changes, it is \( 0.6 \leq \lambda \leq 2.0 \) for \( a = -2 \) and \( a = 5\% \) and this becomes \( 0.8 \leq \lambda \leq 2.0 \) for \( a = -2 \) and \( a = 10\% \). Similarly for other values of negative 'a' the range of \( \lambda \) changes these are the maximum ranges reported here. Again, when 'a' is positive the range of \( \lambda \) is \( 0.8 \leq \lambda \leq 1.8 \) for \( a = 1.75 \) which is the best obtained range for positive values of 'a' at higher levels of significance.

As the value of '\( \nu \)' increases there is a decrease in the RR\( _1 \) values for different values of levels of significance and degrees of asymmetries. However the best performance of \( \hat{\sigma}^2_{ST} \) is observed at \( \nu = 1\% \) for \( a = -2 \) and \( \nu = 1\% \) for \( a = 1.75 \).

It is recommended therefore to consider a smaller level of significance (preferably \( 1\% \)) and smaller sample size \( \nu = 5 \) or 8 for positive / negative values of 'a' at particular \( a = 1.75 \) and \( a = -2 \).

Next we have considered another testimator \( \hat{\sigma}^2_{ST} \) which is obtained by squaring the shrinkage factor, we have evaluated the expression RR\( _2 \) for the same set of values as considered for RR\( _1 \) and our recommendations are as follows:

(i) \( \hat{\sigma}^2_{ST} \) performs better than the usual estimator \( \hat{\sigma}^2 \) for different ranges of \( \lambda \) i.e. for \( a = -2 \), it is \( 0.6 \leq \lambda \leq 1.8 \), however for \( a = -1 \) it becomes \( 0.6 \leq \lambda \leq 1.6 \) i.e. almost the same whole range as observed for \( \hat{\sigma}^2_{ST} \) but the magnitude of RR\( _2 \) values are higher than the magnitude of RR\( _1 \) values indicating a 'better' control over the risk of \( \hat{\sigma}^2_{ST} \) as compared to \( \hat{\sigma}^2_{ST} \). This pattern is observed when \( \alpha = 1\%, \nu = 5 \) and \( a = -1.0 \). Similarly for the positive values of 'a' the magnitude of RR\( _2 \) values are higher for \( a = 1.75 \).

(ii) A similar kind of pattern for the performance of \( \hat{\sigma}^2_{ST} \) is observed for higher levels of significance i.e. \( \alpha = 5\% and \alpha = 10\% \) for different the ranges of \( \lambda \).
however it decreases slightly as now it becomes $0.6 \leq \lambda \leq 1.4$

(iii) From the above discussion It is observed that the values of $RR_2$ are more than unity for some positive and negative values of ‘a’. So, it is concluded that in both the situations i.e. over/under estimation the proposed estimators behaves nicely in the sense of having smaller risks.

(iv) The maximum values of $RR_2$ are observed for $\alpha = 1\%$, $a = -2.0$ and $\nu = 5$. Similarly the highest values of these are observed for $a = 1.75$, $\alpha = 1\%$ and $\nu = 5$.

(v) It is observed that the RR values decrease for higher values of ‘$\nu$’ and ‘$\alpha$’.

(vi) So, it is recommended to consider smaller level of significance along with a smaller sample size with proper choice of ‘$a$’.

To conclude, we recommend that: A shrinkage testimator $\hat{\sigma}^2_{ST}$ (i.e. ‘square’ of shrinkage factor) should be considered with lower level of significance $\alpha = 1\%$, smaller sample size $\nu = 5$ or $8$ and the degree of asymmetry as positive $a = 1.75$ (for situations where overestimation is more serious) and a negative $a = -2.0$ (for situations where under estimation is more serious).

Acknowledgements

The authors are grateful to the referees for their comments which led to improvement of the paper in present form.

References


Appendix

Graphs of Relative Risk for $\hat{\sigma}_{ST_1}$ and $\hat{\sigma}_{ST_2}$ with respect to conventional estimator.

Graphs of Relative Risk for $\hat{\sigma}_{ST_1}$

![Graph of Relative Risk for $\hat{\sigma}_{ST_1}$](image-url)
Some improved shrinkage estimators for variance …

\begin{align*}
\alpha &= 1\%, \nu = 8 \\
\alpha &= 5\%, \nu = 5 \\
\alpha &= 5\%, \nu = 8
\end{align*}

\begin{align*}
\lambda \\
\lambda \\
\lambda
\end{align*}
$a = 1.75, \nu = 5$

- $\alpha = 1\%$
- $\alpha = 5\%$
- $\alpha = 10\%$

$a = -2.0, \nu = 5$

- $\alpha = 1\%$
- $\alpha = 5\%$
- $\alpha = 10\%$
Graphs of Relative Risk for $\hat{\delta}_T$.

**$a = 1\%, \nu = 5$**

- $a=-2$
- $a=-1$
- $a=1$
- $a=1.5$
- $a=1.75$

**$a = 5\%, \nu = 5$**

- $a=1.75$
- $a=2$
- $a=1$
- $a=1.5$
\[ a = 1.75, \nu = 5 \]

\[ a = -2, \nu = 5 \]