International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 14, 2014, pp 37-44

Minimax Estimation of the Variance of a Normal Distribution for an Asymmetric Loss Function

Chandan Kumer Podder
Department of Statistics
University of Chittagong
Chittagong-4331, Bangladesh
Email: *podder_ck@yahoo.com*

[Received September 29, 2013; Revised June 14, 2015; Accepted August 4, 2015]

Abstract

In this paper, minimax estimation of the variance of a normal distribution for an asymmetric loss function has been derived. The asymmetry brings about the shift of location of the loss function.

Keywords and Phrases: Normal distribution, Bayes' and Minimax estimators, Non-informative prior, Risk function, SE and MLINEX loss functions.

AMS Classification: Primary 62J02; Secondary 62J20.

1 Introduction

Let $x_1, x_2, \cdots, x_n$ be a random sample of size n drawn from a normal population whose mean μ and variance θ are both unknown. Several authors have considered the problem of estimation of the variance of a normal distribution under different loss functions. Among them we mention Evan [1], Goodman [2], Pandey and Singh [6], Pal and Ling [7], Prakash and Pandey [8], Solomon [9] and Strawderman [10] etc. Hodges and Lehmann [4] investigated some problems in minimax point estimation.

In the present paper, a minimax estimator of the variance of a normal distribution with unknown mean has been derived by assuming an asymmetric loss function named modified linear-exponential (MLINEX). The asymmetry brings about the shift of location of the loss function. There are many real life data where the use of symmetric loss functions may be inappropriate. In some cases a given positive error may be more serious than a given negative error or vice-versa. In this cases asymmetric loss functions may be used.

2 Preliminaries

Let X be a random variable whose distribution depends on k parameters $\theta_1, \theta_2, \cdots, \theta_k$ and let Ω denotes the parameter space of values of θ, the $k-$ dimensional vector $(\theta_1, \theta_2, \cdots, \theta_k)$. Now consider the general problem of estimating the unknown parameter θ, from the results of a random sample of n observations by the method of Bayesian point estimation.

Denoting the sample results $x_1, x_2, \cdots, x_n$ by x, let $\hat{\theta}$ be an estimator of θ and also let $L(\hat{\theta}, \theta)$ be a loss function, the loss incurred by taking the value of θ to be $\hat{\theta}$. The risk function $R(\hat{\theta}, \theta)$ is the expected value of the loss function with respect to the sample observations.

If $l(\theta|x)$ is the likelihood function of θ given the sample x, and $\pi(\theta)$ is the prior density of θ, then combining $l(\theta|x)$ and $\pi(\theta)$, the Bayes' estimator $\hat{\theta}$ of θ will be a solution of the equation

$$\int_\Omega \frac{\delta L}{\delta \hat{\theta}} l(\theta|x)\pi(\theta)d\theta = 0, \tag{1}$$

where L stands for loss function and assuming that the sufficient regularity conditions prevail to permit differentiation under the sign of integral.

Let us consider an asymmetric loss function as

$$L(\hat{\theta}, \theta) = \varpi \left[\left(\frac{\hat{\theta}}{\theta} \right)^\gamma - \gamma \ln \left(\frac{\hat{\theta}}{\theta} \right) - 1 \right]; \quad \gamma \neq 0, \quad \varpi > 0, \tag{2}$$

where ϖ serving to change the scale of the loss function and γ serving to determine its shape.

When $\hat{\theta} = \theta$ i.e., $\frac{\hat{\theta}}{\theta} = 1$, then $L(\hat{\theta}, \theta) = 0$, writing $R = \frac{\hat{\theta}}{\theta}$, the relative error $L(R)$ is minimum at $R = 1$. If we write $D = \ln R = \ln \hat{\theta} - \ln \theta$, where $D = \hat{\theta} - \theta$ (say) represent the estimation error in estimating θ by $\hat{\theta}$, then $L(R)$ can be expressed as the same form of LINEX loss function,

$$L(\hat{\theta}, \theta) = \kappa \left[e^{\lambda(\hat{\theta}-\theta)} - \lambda(\hat{\theta} - \theta) - 1 \right]; \quad \lambda \neq 0, \quad \kappa > 0, \tag{3}$$

where κ and λ are scale and shape characteristics of the loss function. The loss function defined in (3) is also an asymmetric one, was first introduced by Varian [11] and developed by Zellner [12] in such a form. For this reason (2) is called modified linear-exponential (MLINEX) loss function.

However, the Bayes' estimator of the parameter θ for MLINEX loss function is $\hat{\theta} = [E_\theta(\theta^{-\gamma}|x)]^{-\frac{1}{\gamma}}$; provided that such expectation exists.

It is evident from (1) that, for squared-error (SE) loss function

$$L(\hat{\theta}, \theta) = c(\hat{\theta} - \theta)^2; \quad c > 0, \tag{4}$$

where c serving to change the scale of the loss function, the Bayes' estimator of the parameter θ is simply the mean of the posterior distribution.

The derivation depends primarily on a theorem which is due to Lehmann [5] and can be stated as follows.

Theorem 2.1. *Let $\tau = \{F_\theta; \theta \in \Theta\}$ be a family of distribution functions and D be a class of estimators of θ. Suppose that $\delta^* \in D$ is a Bayes' estimator concerning to a prior distribution $\pi(\theta)$ on the parameter space Θ. If the risk function $R(\delta^*, \theta) =$ constant on Θ, then δ^* is a minimax estimator for θ.*

3 Main Results

Consider the case of estimating the variance θ of a normal distribution of unknown mean μ. Here, $\Omega = (\mu, \theta)$, Ω is the half-plane; $-\infty < \mu < \infty,\ 0 < \theta < \infty$, and

$$\begin{aligned} l(\mu, \theta|x) &= (2\pi\theta)^{-n/2} \exp\{-\textstyle\sum(x_i - \mu)^2/2\theta\} \\ &= (2\pi\theta)^{-n/2} \exp\left[-\{S + n(\mu - \bar{x})^2\}/2\theta\right], \end{aligned} \tag{5}$$

where $S = \sum(x_i - \bar{x})^2$. A mathematically convenient and widely applicable joint prior density for the problem under consideration suggested by Evan [1] is the class of natural conjugates

$$\pi(\mu, \theta) \propto \theta^{-(1+\nu/2)} \exp\left[-\{\eta + \zeta(\mu - \xi)^2\}/2\theta\right], \tag{6}$$

where $\nu,\ \eta,\ \zeta \geq 0$ and $-\infty < \xi < \infty$. This is obtained by generalizing the likelihood (5) regarded as a function of the unknown parameters, and here is seen to be equivalent to assuming that the prior marginal density of θ is such that η/θ is distributed as chi-square with $(\nu - 1)$ degrees of freedom and that the prior conditional density of μ

given θ is normal with mean ξ and variance θ/ζ.

The advantage of taking the prior distribution to be natural conjugate lies in the fact that the likelihood $l(\mu, \theta|x)$, the prior density $\pi(\mu, \theta)$ and the posterior density $\pi(\mu, \theta|x)$ are all of the same functional form, thus ensuring mathematically tractability.

For the limiting case, when $\eta = \zeta = 0$ we have the subclass of prior densities given by

$$\pi(\mu, \theta) \propto \theta^{-(1+\nu/2)}, \tag{7}$$

which is equivalent to assuming the prior distributions of μ and θ to be independent, that of μ being uniform and that of θ being proportional to $\theta^{-(1+\nu/2)}$. The particular form of (7) corresponding to $\nu = 0$ is precisely the prior distribution advocated by Jeffreys' [3] when one is completely ignorant to the value of μ and θ apart, of course, from their admissible ranges.

Substitution from (5) and (6) in (1), the Bayes' estimator $\hat{\theta}$ of θ is a solution of

$$\iint \frac{\delta L(\hat{\theta}, \theta)}{\delta \hat{\theta}} \theta^{-(n+\nu+2)/2} \exp\left[-\{S + n(\mu - \bar{x})^2 + \zeta(\mu - \xi)^2 + \eta\}/2\theta\right] d\mu d\theta = 0,$$

the integration being over $\Omega : -\infty < \mu < \infty,\ 0 < \theta < \infty$. On noting that $L(\hat{\theta}, \theta)$ is independent of μ and that

$$\int_{-\infty}^{\infty} \exp\left[-\{n(\mu - \bar{x})^2 + \zeta(\mu - \xi)^2\}/2\theta\right] d\mu = \{2\pi\theta/(n+\zeta)\}^{1/2} \exp\{-n\zeta(\xi - \bar{x})^2/2(n+\zeta)\theta\},$$

we find that $\hat{\theta}$ is a solution of

$$\int_{0}^{\infty} \frac{\delta L(\hat{\theta}, \theta)}{\delta \hat{\theta}} \theta^{-(n+\nu+1)/2} \exp\left(-K/2\theta\right) d\theta = 0, \tag{8}$$

where $K = S + \eta + \zeta(\xi - \bar{x})^2/(n + \zeta)$.

For the loss function given by (2), it follows from (8) that the estimator $\hat{\theta}$ is given by

$$\hat{\theta}^{\gamma} = \frac{\int_0^{\infty} \theta^{-(n+\nu+1)/2} \exp(-K/2\theta) d\theta}{\int_0^{\infty} \theta^{-(n+\nu+2\gamma+1)/2} \exp(-K/2\theta) d\theta}, \tag{9}$$

using the transformation $K/2\theta = y$, then (9) becomes

$$\begin{aligned}\hat{\theta}^{\gamma} &= \left(\frac{K}{2}\right)^{\gamma} \frac{\int_0^\infty y^{(n+\nu-1)/2-1}\exp(-y)dy}{\int_0^\infty y^{(n+\nu+2\gamma-1)/2-1}\exp(-y)dy} \\ &= \left(\frac{K}{2}\right)^{\gamma} \frac{\Gamma\left(\frac{n+\nu-1}{2}\right)}{\Gamma\left(\frac{n+\nu+2\gamma-1}{2}\right)}.\end{aligned}$$

Hence,

$$\hat{\theta} = CK, \tag{10}$$

where $C = \frac{1}{2}\left[\frac{\Gamma\left(\frac{n+\nu-1}{2}\right)}{\Gamma\left(\frac{n+\nu+2\gamma-1}{2}\right)}\right]^{\frac{1}{\gamma}}$.

For the limiting case $\eta = \zeta = 0$, we have the Bayes' estimator $\hat{\theta} = CS$.

As $x \sim N(\mu, \theta)$ then $u = \frac{S}{\theta}$ is distributed as chi-square with $(n-1)$ degrees of freedom. The probability density function of u is

$$f(u) = \frac{1}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)} u^{\frac{n-1}{2}-1}\exp(-\frac{1}{2}u); \quad u \geq 0,$$

and hence the probability density function of S is given by

$$f(S) = \frac{1}{(2\theta)^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)} S^{\frac{n-1}{2}-1}\exp(-\frac{1}{2\theta}S); \quad S \geq 0. \tag{11}$$

Therefore, the risk function of the estimator $\hat{\theta}$ for the MLINEX loss function (2) is

$$\begin{aligned}R_{ML}(\hat{\theta}, \theta) &= E\left[L(\hat{\theta}, \theta)\right] \\ &= \varpi\left[\frac{1}{\theta^{\gamma}}E(\hat{\theta}^{\gamma}) - \gamma E(\ln\hat{\theta}) + \gamma\ln\theta - 1\right].\end{aligned} \tag{12}$$

For simplicity,

$$\begin{aligned}E(\hat{\theta}^{\gamma}) &= E(CS)^{\gamma} \\ &= C^{\gamma}E(S^{\gamma}) \\ &= C^{\gamma}\int_0^\infty S^{\gamma}f(S)dS \\ &= C^{\gamma}\frac{(2\theta)^{\gamma}\Gamma(\frac{n+2\gamma-1}{2})}{\Gamma(\frac{n-1}{2})},\end{aligned}$$

and

$$E(\ln\hat{\theta}) = E(\ln CS)$$
$$= \ln C + E(\ln S),$$

where

$$E(\ln S) = \int_0^\infty \ln S f(S) dS$$
$$= \frac{1}{(2\theta)^{\frac{n-1}{2}}\Gamma(\frac{n-1}{2})}\int_0^\infty \ln S \; S^{\frac{n-1}{2}-1}\exp(-\frac{1}{2\theta}S)dS,$$

let us also make a transformation, $\frac{S}{2\theta} = z$.

Then

$$E(\ln S) = \frac{1}{\Gamma(\frac{n-1}{2})}\int_0^\infty \{\ln(2\theta) + \ln z\} \; z^{\frac{n-1}{2}-1}\exp(-z)dz$$
$$= \ln(2\theta) + \frac{1}{\Gamma(\frac{n-1}{2})}\int_0^\infty \ln z \; z^{\frac{n-1}{2}-1}\exp(-z)dz$$
$$= \ln(2\theta) + \frac{\Gamma'(\frac{n-1}{2})}{\Gamma(\frac{n-1}{2})},$$

where $\Gamma'(\frac{n-1}{2}) = \int_0^\infty \ln z \; z^{\frac{n-1}{2}-1}\exp(-z)dz$ is the first derivative of $\Gamma(\frac{n-1}{2})$ with respect to n.

Therefore,

$$E(\ln\hat{\theta}) = \ln C + \ln(2\theta) + \frac{\Gamma'(\frac{n-1}{2})}{\Gamma(\frac{n-1}{2})}.$$

Using these above results, the risk function of the estimator $\hat{\theta}$ for the MLINEX loss function becomes

$$R_{ML}(\hat{\theta},\theta) = \varpi\left[\frac{\Gamma(\frac{n+\nu-1}{2})}{\Gamma(\frac{n+\nu+2\gamma-1}{2})}\frac{\Gamma(\frac{n+2\gamma-1}{2})}{\Gamma(\frac{n-1}{2})} - \ln\frac{\Gamma(\frac{n+\nu-1}{2})}{\Gamma(\frac{n+\nu+2\gamma-1}{2})} - \gamma\frac{\Gamma'(\frac{n-1}{2})}{\Gamma(\frac{n-1}{2})} - 1\right],$$

which is a constant with respect to θ, as n, ν, and γ are known and independent of θ.

For the limiting case $\eta = \zeta = 0$ then according to the Lehmann [5], $\hat{\theta} = CS$ is a minimax estimator of θ, where $C = \frac{1}{2}\left[\frac{\Gamma(\frac{n+\nu-1}{2})}{\Gamma(\frac{n+\nu+2\gamma-1}{2})}\right]^{\frac{1}{\gamma}}$ and $S = \sum(x_i - \bar{x})^2$.

As $\gamma = -1$, then

$$\begin{aligned}\hat{\theta} &= \frac{1}{2}\frac{\Gamma\left(\frac{n+\nu-3}{2}\right)}{\Gamma\left(\frac{n+\nu-1}{2}\right)}S\\ &= \frac{S}{n+\nu-3}\\ &= \frac{S}{d(n)},\end{aligned}$$

where $d(n) = n+\nu-3$ is same as the Bayesian estimation of the variance of a normal distribution for squared-error loss function derived by Evans [1].

When $\nu = 0$, then $\hat{\theta} = \frac{1}{2}\left[\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n+2\gamma-1}{2})}\right]^{\frac{1}{\gamma}} \sum(x_i - \bar{x})^2$ is a minimax estimator of θ for MLINEX loss function using the Jeffreys' prior density and the risk function of the estimator as

$$R_{ML}(\hat{\theta}, \theta) = \varpi\left[-\ln\frac{\Gamma'\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n+2\gamma-1}{2}\right)} - \gamma\frac{\Gamma'\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right],$$

which is a constant with respect to θ, as n and γ are known and independent of θ too.

It is seen that when $\nu = 0$ and $\gamma = 1$, then $\hat{\theta} = \frac{1}{n-1}\sum(x_i - \bar{x})^2$ is an unbiased estimator of the variance of a normal distribution.

Acknowledgement

The author is thankful to the referees for their valuable suggestions.

References

[1] Evans, I.G. (1964). Bayesian estimation of the variance of a Normal Distribution, *JRSS*, **26**, 63-68.

[2] Goodman, L.A. (1960). "A note on the estimation of variance", *Sankhya*, **22**, 221-228.

[3] Jeffreys', H. (1961). Theory of probability, *3rd ed. Oxford: Clarendon Press.*

[4] Hodges, J.I. and Lehmann, E. L. (1950). Some problems in minimax point estimation, *Ann. of Math. Stat.*, **21**, 182-197.

[5] Lehmann E.L. (1983). Theory of Point Estimation, New York, John Wiley.

[6] Pandey, B.N. and Singh, J. (1977). Estimation of variance of Normal distribution using apriori information. *Journal of Indian Statistical Association*, **15**, 141-150.

[7] Pal, N. and Ling, C.J. (1996). Estimation of a Normal scale parameter under balanced loss function. *Journal of Indian Statistical Association*, **34**, 21-38.

[8] Prakash, G. and Pandey, B.N. (2007). Shrinkage testimation for the variance of normal distribution under asymmetric loss function. *Journal of Statistical Research*, **41(7)**, 17-35.

[9] Solomon, D.L. (1972). $\bigwedge -$ Minimax estimation of a scale parameter, *Journal of American Statistical Association*, **67**, 647-649.

[10] Strawderman, W.E. (1974). Minimax estimation of power of the variance of a normal population under squared error loss. *Annals of Statistics*, **2(1)**, 190-198.

[11] Varian, H.R. (1975). "A Bayesian Approach to Real Estate Assessment" in studies in Bayesian Econometrics and Statistics in Honour of L. J. Savage, eds, S. E. Fienberg and A. Zellner, Amsterdam, *North-Holland Publishing Co.*, 195-208.

[12] Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss function, *Journal of American Statistical Association*, **81**, 446-451.