Journal of Sciences, Islamic Republic of Iran 13(3): 267-269 (2002)
National Center For Scientific Research, ISSN 1016-1104

IMPROVED ESTIMATOR OF THE VARIANCE IN THE LINEAR MODEL

N. Sanjari Farsipour*

Department of Statistics, Faculty of Science, Shiraz University, Shiraz, Islamic Republic of Iran

Abstract

The improved estimator of the variance in the general linear model is presented under an asymmetric linex loss function.

Keywords: Equivariant estimator; Normal variance estimator; Improved estimator; Linex loss function

1. Introduction

Consider the canonical form of the general linear model and suppose $X \sim N_P(\mu,\tau I)$ and $U \sim N_n(O,\tau I)$ are to be independently observed. On the basis of these observations, τ is to be estimated, where the loss function is given by

$$L(\tau,\delta) = b\left\{e^{a\left(\frac{\delta}{\tau}-1\right)} - a\left(\frac{\delta}{\tau}-1\right)-1\right\}, \qquad (1.1)$$

where $a \neq 0$ is a shape parameter and $b>0$ is a scale parameter. This loss function which was introduced by Varian [1] and was extensively discussed by Zellner [2], is useful when overestimation is regarded as more serious than underestimation or *vice versa*. In this regard see Parsian and Sanjari Farsipour [3].

A sufficient statistic in this problem is (X,T), where if $\|.\|$ denotes the usual Euclidean norm, $T=\|U\|^2$.

2. MLE and Bayes Estimators

With U unobserved, we can write down the likelihood function, given our normality assumptions, and easily obtain the maximum likelihood estimator. The likelihood function is

$$L(\mu,\tau) =$$

$$(2\pi)^{-\frac{p+n}{2}}(\tau^{-1})\exp\left\{-\frac{1}{2\tau}(X-\mu)'(X-\mu)-\frac{1}{2\tau}U'U\right\}.$$

So we have **X** as an MLE of μ, and $\frac{1}{2}\sum_{i=1}^{n} U_i^2$ as an MLE of τ. Now, we calculate the risk function relative to the loss function in (1.1) of $T = \sum_{i=1}^{n} U_i^2$, we have

$$R(\tau,\hat{\tau}) = e^{-a}(1-a)^{-\frac{n}{2}} - \frac{an}{2} + a - 1 \qquad (2.1)$$

Now, let $\lambda=\tau^{-1}$, and introducing a diffuse prior, as the one cited in the article by Zellner [1], *i.e.*, $\pi(\lambda)=\frac{1}{\lambda}$ we can derive an optimal estimate that minimizes the posterior expected loss of our loss function in (1.1), as a solution of the following equation

* *E-mail: nsf@stat.susc.ac.ir*

$$E_\lambda\left[\lambda e^{a\lambda\delta_B} \mid T=t\right] = e^a E_\lambda\left[\lambda \mid T=t\right]. \quad (2.2)$$

Hence, the Bayes estimator is $\delta_B = \frac{1}{2a}(1-e^{-\frac{2a}{3}})T$. Now we are able to obtain the risk function associated with this estimator as the following equation

$$R(\lambda,\delta_B) = \tfrac{1}{2}\left(1+e^{-\frac{2a}{3}}\right)^{-n} e^{-a} + \frac{n}{2}e^{-\frac{2a}{3}} - \frac{n}{2} + a - 1, \quad (2.3)$$

and we can compare it with that we already derived under the assumption that U is observed. Obviously δ_B works better than T, since it is the best invariant estimator, and T is an invariant estimator.

For the loss function of the form $L(\delta,\lambda) = (\frac{\delta}{\lambda}-1)^2$ the problem was solved by some authors such as Brewster and Zidek [4] as well as Hodges and Lehmann [5].

3. Improved Estimators

The problem remains invariant under the transformation group A under which

$$(\mathbf{X},T) \to (\alpha\Gamma\mathbf{X}+\beta, \alpha^2 T)$$
$$(\mu,\tau) \to (\alpha\Gamma\mu+\beta, \alpha^2\tau) \quad (3.1)$$
$$\delta \to \alpha^2\delta$$

where $\alpha > 0, \beta \in \Re^P$ and Γ is a $p \times p$ orthogonal matrix. It follows that any nonrandomized $\mathcal{A}$-invariant estimator of τ is of the form cT, for some constant c>0. Since $\mathcal{A}$ acts transitively on the parameter space, the risk function of cT,

$$E_{\mu,\tau}\left[\rho\left(\frac{cT}{\tau}\right)\right] = E_{0,1}\left[\rho(cT)\right],$$

is independent of the unknown parameters, where ρ (.) is the scale invariant low function. Then the optimum choice for c is derived from the equation

$$E_{0,1}\left[\frac{\partial}{\partial c^*}\rho(c^*T)T\right] = 0$$

and for the loss function (1.1), c^* is a multiplier of $\sum_{i=1}^n X_i^2$ [3].

Let $\mathcal{H}$ denote the subgroup of $\mathcal{A}$ obtained by requiring in (3.1) that $\beta = 0$ and that Γ be a diagonal orthogonal matrix. Any $\mathcal{H}$-invariant estimator is of the form ϕ (**Z**)T, where **Z**=(Z1, Z2, … , Zp)' and $Z_i = |X_i| T^{-\frac{1}{2}}, i = 1,\ldots,p$. We can see that the risk of such an estimator is

$$R(\mu,\tau;\delta) = E_{\mu,\tau}\left[\rho\left(\frac{\phi(z)T}{\tau}\right)\right]$$
$$= E_{\xi,1}\left[\rho(\phi(z)T)\right]$$
$$= R(\xi;\delta), (say)$$

where $\xi = (\xi_1,\xi_2,\ldots\xi_p)'$ and $\xi_i = |\mu_i|\tau^{-\frac{1}{2}}, i = 1,\ldots,p$. Since we deal only with $\mathcal{H}$–invariant estimators, we may assume without loss of generality that τ =1.

On the other hand, X_i^2 has a chi-squared distribution with 1+2K$_i$ degrees of freedom, where K_i denotes a Poisson random variable with mean $\lambda_i = \frac{1}{2}\xi_i^2$, and the K_i's, i=1,…,p, are independent of each other and of T. Let **K**=($K_1,K_2,\ldots,K_p$), the joint density of T and **Z** conditional on **K**=**k**= ($k_1,k_2,\ldots,k_p$) is

$$f_{T,Z}(t,z \mid k) \propto t^{\frac{1}{2}(n+p)+k_\bullet-1} e^{-\frac{1}{2}t(1+\|z\|^2)} \prod_{i=1}^p z_i^{2k_i},$$

Independent of ξ, where $k_\bullet = \sum_{i=1}^p k_i$.

Now since the loss (1.1) is strictly convex, it uniquely minimized at $\phi_k(z)$ satisfying

$$E\left\{\rho'(\phi_k(\mathbf{Z})T)T \mid \mathbf{Z}=z, \mathbf{K}=k\right\} = 0$$

which is equivalent to

$$E\{Te^{a\phi}k^{(\mathbf{Z})T} \mid \mathbf{Z}=z, \mathbf{K}=k]\} = e^a E[T \mid \mathbf{Z}=z, \mathbf{K}=k].$$

Now, for any estimator ϕ (**Z**)T define $\phi^*(z)$ = $\min\{\phi(z),\phi_o(z)\}$, then let

$$R(\xi;\phi) = E_\xi\left\{E[\rho(\phi(\mathbf{Z})T) \mid \mathbf{Z},\mathbf{K}]\right\}$$
$$= E_\xi\left\{R(\phi(\mathbf{Z}) \mid \mathbf{Z},\mathbf{K})\right\}.$$

Now, either $\phi^*(\mathbf{z}) = \phi(\mathbf{z})$, then $R(\phi^*(\mathbf{z}) \mid \mathbf{z},\mathbf{k}) = R(\phi(\mathbf{z}) \mid \mathbf{z},\mathbf{k})$ or $\phi^*(\mathbf{z}) = \phi_o(\mathbf{z}) < \phi(\mathbf{z})$, then since $R(\phi \mid \mathbf{z},\mathbf{k})$ is strictly convex, and $\phi_{\mathbf{k}}(\mathbf{z}) \le \phi_o(\mathbf{z})$ for all

k, it follows that $R(\phi^*(\mathbf{z})\,|\,\mathbf{z},\mathbf{k}) < R(\phi(\mathbf{z})\,|\,\mathbf{z},\mathbf{k})$, see Figure 2.1, which is also cited in Maatta and Casella [6] in the univariate set up. Therefore, for any $\xi, R(\xi,\phi^*) \le R(\xi,\phi)$ with inequality if $P_\xi(\phi^*(\mathbf{Z}) \ne \phi(\mathbf{z})) > 0$. Now, let $\phi(\mathbf{z}) = c^* = \frac{1}{2a}(1-e^{-\frac{2a}{n+2}})$, then to find $\phi_o(\mathbf{z})$ in this case, note that

$$R(c\,|\,\mathbf{z},\mathbf{O}) \propto \int \rho(ct)t^{\frac{1}{2}(n+p)-1}e^{-\frac{1}{2}t(1+\|z\|^2)}dt.$$

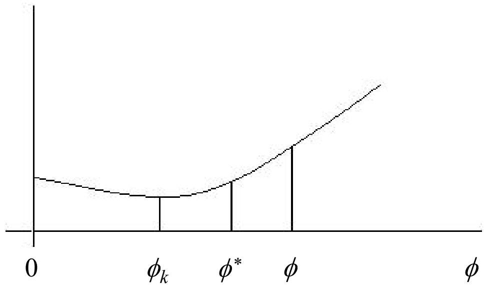

Figure 3.1.

So, using the transformation $t \to t(1+\|\mathbf{z}\|^2)$, we can see that

$$R(c\,|\,\mathbf{z},\mathbf{O}) \propto \int \rho(\tilde{c}t)t^{\frac{p}{2}}t^{\frac{n}{2}-1}e^{-\frac{1}{2}t}dt$$
$$\propto E\left[\rho(\tilde{c}T)T^{\frac{p}{2}}\right] \qquad (3.2)$$

where $\tilde{c} = c/(1+\|\mathbf{z}\|^2)$, so the minimum is attained at $\tilde{c} = \phi_o(\mathbf{z})/(1+\|\mathbf{z}\|^2)$. For finding the value of $\tilde{c}$, using (2.2), $\tilde{c}$ must satisfy the following relation

$$E\left[T^{\frac{p}{2}+1}e^{a\tilde{c}T}\right] = e^a E\left[T^{\frac{p}{2}+1}\right]$$

which is obtained by

$$\tilde{c} = \frac{1}{2a}\left(1-e^{-\frac{2a}{n+p+2}}\right).$$

Hence, $\phi_o(\mathbf{z}) = \frac{1}{2a}(1-e^{-\frac{2a}{n+p+2}})(1+\|\mathbf{z}\|^2)$, and so by the above discussion c^*T is dominated by

$$\delta^* = \min\left\{c^*, \tilde{c}\left(1+\|z\|^2\right)\right\}T. \qquad (3.3)$$

References

1. Varian H.R. A Bayesian approach to real estate assessment, in studies in Bayesian Econometrics and statistics in Honor of Leonard J. Savge, Eds. Fienberg S.E. and Zellner A., Amesterdon, North Holland, 195-208 (1975).
2. Zellner A. Bayesian estimation and prediction using asymmetric loss function. *J. Amer. Statist. Assoc.*, **81**: 446-451 (1986).
3. Parsian A. and Sanjari Farsipour N. On the Admissibility and Inadmissibility of Estimators of Scale Parameter using an Asymmetric Loss Function. *Commun. Statist.-Theory Meth.*, **22**(10): 2877-2901 (1993).
4. Brewster J.F. and Zidek J.V. Improving on equivariant estimators. *Annals of Statistics*, **2**: 21-38 (1974).
5. Hodges J.L. and Lehmann E.L. Some applications of the Cramer-Rao inequality. *Proc. 2nd Berkeley Symp. Math. Statist. Probab*., **1**: 13-22 (1951).
6. Maatta J. and Casella G. Developments in decision theoretic variance estimation. *Statistical Science*, **5**: 90-120 (1990).