Finding Posterior Moments for a Normal Mean Parameter and its Application in Bayesian Estimation Using LINEX Loss Function

Danni Shi

Stat 544 - Final Project

Abstract

This project reproduces works done in Rodrigues (1994) by reviewing how to find and approximate posterior moments for a normal mean parameter, and how to apply such methods in finding the Bayes estimator of the normal mean parameter using the LINEX loss function. Since the LINEX loss function is asymmetric, we need to consider the robustness of different priors with respect to the corresponding Bayes estimators. This project will show that the Bayes estimator under a double-exponential or Student's *t* prior is robust when replacing the squared error loss function with the LINEX one by showing the exact and approximation of the posterior moments under both priors.

Introduction & Motivation

The loss function along with the assigned prior distribution plays a vital role in Bayes estimation. In a study of real estate assessment, Varian (1975) introduced the LINEX loss function, which is asymmetric and is a useful method for estimation and prediction problem which rises exponentially on one side of zero and almost linearly on the other side of zero. It is defined as

$$L(\Delta) = b\left(e^{a\Delta} - a\Delta - 1\right), \quad a \neq 0, b > 0 \tag{1}$$

where $\Delta = \hat{\theta} - \theta$ is the prediction error.

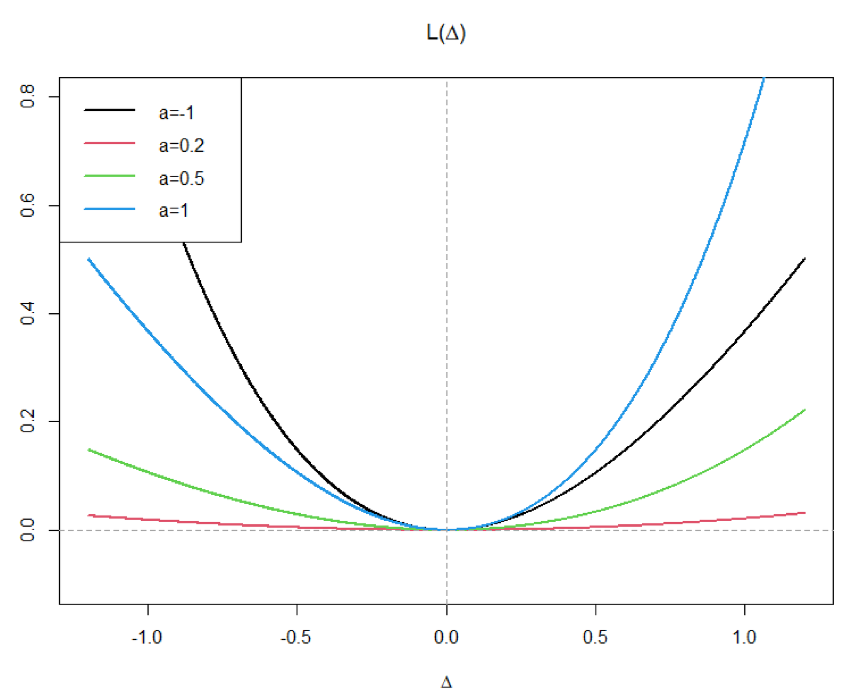

Figure 1. Plot of $L(\Delta)$ when $a = -1, 0.2, 0.5, 1$

For $|a| \geq 1$, $L(\Delta)$ is quite asymmetric. For small $|a|$, $L(\Delta)$ is almost symmetric and close to a squared error loss function. Let $\pi(\theta)$ and $p(\theta|D)$ be the prior and posterior respectively. The Bayes estimator under the LINEX is

$$\hat{\theta}_B = -\frac{1}{a}\log\left[E\left(e^{-a\theta}|D\right)\right] \tag{2}$$

which involves moment generation function evaluation and approximation for the posterior.

Consider n IID observations with normal likelihood with unknown mean θ and known variance σ^2. Since the LINEX loss function is asymmetric when $|a|$ is large, we need to consider the sensitivity of the Bayes estimator to the prior when replacing the squared error loss function with the LINEX loss function. Let $\hat{\theta}_C$ denote the Bayes estimator under squared error loss, so $\hat{\theta}_C = E(\theta|D)$. When the prior is normal, by Zellner (1986), $\hat{\theta}_B = \hat{\theta}_C - \frac{a}{2}\text{Var}(\theta|D)$. If $\frac{a}{2}\text{Var}(\theta|D)$ is small, $\hat{\theta}_B \to \hat{\theta}_C$. Since this is not always the case, $\hat{\theta}_B$ lacks robustness.

I choose to reproduce Rodrigues (1994) as it applies techniques of evaluating posterior moments to provide a link between the robust Bayesian analysis for the normal mean parmameter when adopting the LINEX and squared error loss function and the next two sections will show that the sensitivity of $\hat{\theta}_B$ to the double-exponential prior and Student's *t* prior is not affected under the LINEX.

Key Theoretical Ideas

Let the density of $Y \equiv \frac{1}{n}(X_1 + \cdots + X_n) \sim N(\theta, \sigma^2/n)$ be $p(y-\theta)$. Let $m(y) = \int p(y-\theta)\pi(\theta)\mathrm{d}\theta$ and

$$s(y) = -\frac{\partial \log m(y)}{\partial y}, S(y) = \frac{\partial s(y)}{\partial y}, H(y) = \frac{1}{a}\log\left[\frac{m(y - a\sigma^2/n)}{m(y)}\right] \tag{3}$$

$$\Rightarrow H(y) = \frac{1}{a}\int_{y-a\sigma^2/n}^{y} s(t)\mathrm{d}t \tag{4}$$

$$\hat{\theta}_C = y - \frac{\sigma^2}{n} - s(y), \hat{\theta}_B = y - \frac{a\sigma^2}{2n} - H(y) \tag{5}$$

Similar to the conclusion in Pericchi & Smith (1992) of the properties of $s(y)$ and the influence of the prior on the posterior mean, if $\pi(\theta)$ is such that $H(y)$ is bounded as a function of $\hat{\theta} - \mu$, then $|\hat{\theta}_B - \hat{\theta}_C|$ is bounded and $\hat{\theta}_B$ is robust. This does not hold for conjugate (normal) prior.

Note that by (4) and the Mean Value Theorem, without loss of generality, if $s(t)$ is continuous on $J = [y - a\sigma^2/n, y]$, $(a > 0)$, $\exists c \in J$ such that

$$H(y) = \frac{\sigma^2}{n}s(c) \text{ and } \hat{\theta}_B = \hat{\theta}_C - \frac{\sigma^2}{n}s(c) \tag{6}$$

Results: Double Exponential Prior

Consider $\pi(\theta) = \frac{1}{\sqrt{2\pi\nu}}\exp\left(-\frac{\sqrt{2}}{\nu}|\theta - \mu|\right)$. Let $a^* = \exp\left(\frac{\sigma^2}{n\nu^2}\right), b^* = \frac{\sqrt{2}\sigma^2}{n\nu}, c(y) = \frac{\sqrt{2}}{\nu}(y-\mu)$ and $\Phi(\cdot)$ be the cumulative distribution for standard normal. Here,

$$m(y) = \frac{a^*}{\sqrt{2}\nu}[F(y) + G(y)] \tag{7}$$

$$F(y) = e^{c(y)}\Phi\left[\frac{\sqrt{n}}{\sigma}(\mu - y - b)\right], G(y) = e^{-c(y)}\Phi\left[-\frac{\sqrt{n}}{\sigma}(\mu - y + b)\right] \tag{8}$$

From Pericchi & Smith (1992),

$$E(\theta|y) = w(y)(y+b) + [1 - w(y)](y - b) \tag{9}$$

where $w(y) = F(y)/[F(y) + G(y)]$, which is proportional to $s(y)$.

Meanwhile, since Pericchi and Smith derived that $|E(\theta|y) - y| \leq b$, then $s(y)$ is bounded (for finite σ^2) and based on (6), so is $H(y)$. Since $H(y)$ is bounded, then so is $\left|\hat{\theta}_B - \hat{\theta}_C\right|$ and $\hat{\theta}_B$ is robust when replacing the LINEX loss function with the squared error one.

Results: Student's *t* Prior

Consider $\theta \sim t(\theta|\alpha, \mu, \tau)$ where $\pi(\theta) \propto \left[1 + \frac{(\theta-\mu)^2}{\alpha\tau^2}\right]^{-(\alpha+1)/2}$. Lindley (1968) derived that $E(\theta|y) = \hat{E}(\theta|y) + O\left(n^{-2}\right)$ where

$$\hat{E}(\theta|y) = y - \frac{\sigma^2}{n}\cdot\frac{(\alpha+1)(y-\mu)}{\alpha\tau^2 + (y-\mu)^2} = y - \frac{\sigma^2}{n}\cdot s^*(y) \tag{10}$$

where $s^*(y)$ is an approximation for $s(y)$. Then by (5) and (10),

$$\hat{\theta}_b \doteq y - \frac{a\sigma^2}{2n} - \frac{(\alpha+1)}{2a}\log\left[1 - \frac{\sigma^2}{n}\cdot\frac{2a(y - a\sigma^2/2n - \mu)}{\alpha\tau^2 + (y-\mu)^2}\right] \tag{11}$$

Based on my reproduction of works in Rodrigues (1994) and Lindley (1968), Figure 1 shows the influence of the dependence of the difference between $\hat{\theta}_B$ and $\hat{\theta}_C$ on the discrepancy between the prior location parameter and the estimated posterior mean. Same as in Rodrigues (1994), I choose $n = 10, \alpha = 9, \sigma = \tau = 1$. Here I extended Rodrigues' choices of μs to $\pm\{0, 0.25, 0.5, 1, 2.5, 5, 7.5, 10, 12, 15, 20\}$ and plot the corresponding cases when $a = -4, -2, 2$ and 4.

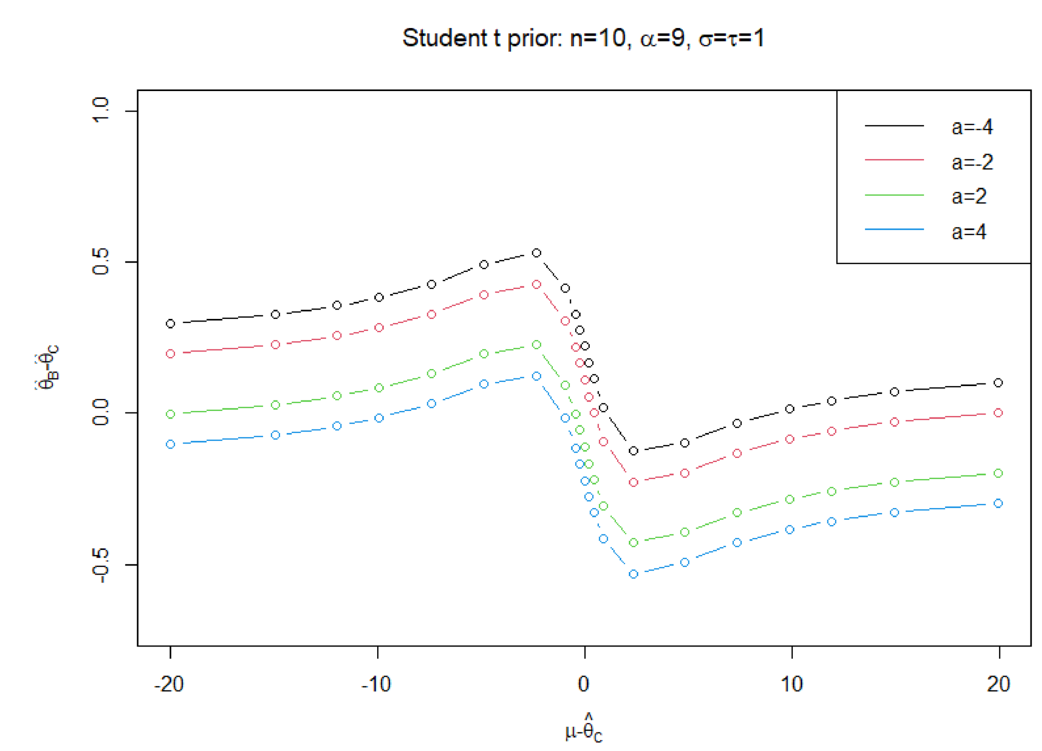

Figure 2. Student's *t* prior: $n = 10, \alpha = 9, \sigma = \tau = 1$

The shapes of the lines and their general convergence as $\mu - \hat{\theta}_C \to \pm\infty$ are the same for different values of a. In general $\hat{\theta}_B - \hat{\theta}_C$ decreases as a increases, but $|\hat{\theta}_B - \hat{\theta}_C|$ is always bounded. In general, my reproduction is consistent with Rodrigues' work. So the sensitivity to Student's *t* prior for the normal mean parameter is not affected when replacing the squared error loss function by the LINEX.

Discussion

When a LINEX loss function is necessary in practice, one should be careful in the choice of prior distribution. This project shows that the sensitivity of the Bayes estimator to double exponential and Student's *t* prior is not affected when replacing the squared error loss function by the LINEX one, while it is affected under normal prior.

One can extend this project to broader cases, for instance, when the likelihood is of a general exponential family form, or when an arbitrary prior is adopted when a LINEX loss function is chosen. Or one can extend this project to the robustness consideration of Bayes estimators under all highly asymmetric loss functions.

References

1. Lindley, D. V. (1968) The choice of variables in multiple regression (with discussion). *J. R. Statist. Soc.* B, 30, 31-66.
2. Miyasawa, K. (1960) An empirical Bayes estimator of the mean of a normal population. *Bull. Int. Statist. Inst.*, 38, 181.
3. Pericchi, L. R. and Smith, A. E M. (1992). Exact and approximate posterior moments for a normal location parameter. *J. Roy. Statist. Soc.* B 54, 793-804.
4. Rodrigues. (1994). Bayesian estimation of a normal mean parameter using the Linex loss function and robustness considerations. *Test (Madrid, Spain)*, 3(2), 237–246. `https://doi.org/10.1007/BF02562704`
5. Varian, tt. R. (1975). A Bayesian approach to real estate assessment. Studies in Bayesian Econometrics and Statistics in Honor of L .L Savage. (S. E. Feinberg and A. Zellner, eds,), Amsterdam: North Holland, 195-208.
6. Zellner, A. (1986). Bayes estimation and prediction using asymmetric loss functions. *J. Amer. Statist. Assoc.* 81,446-451.