应用概率统计　第 34 卷
第 6 期　2018 年 12 月

Chinese Journal of Applied Probability and Statistics
Dec., 2018, Vol. 34, No. 6, pp. 551-564
doi: 10.3969/j.issn.1001-4268.2018.06.001

The Bayes Posterior Estimator of the Variance Parameter of the Normal Distribution with a Normal-Inverse-Gamma Prior under Stein's Loss *

XIE Yuhan　SONG Wenhe　ZHOU Mingqin　ZHANG Yingying*

(*Department of Statistics and Actuarial Science, Chongqing University, Chongqing, 401331, China*)

Abstract: For the variance parameter of the normal distribution with a normal-inverse-gamma prior, we analytically calculate the Bayes posterior estimator with respect to a conjugate normal-inverse-gamma prior distribution under Stein's loss function. This estimator minimizes the Posterior Expected Stein's Loss (PESL). We also analytically calculate the Bayes posterior estimator and the PESL under the squared error loss function. The numerical simulations exemplify our theoretical studies that the PESLs do not depend on the sample, and that the Bayes posterior estimator and the PESL under the squared error loss function are unanimously larger than those under Stein's loss function. Finally, we calculate the Bayes posterior estimators and the PESLs of the monthly simple returns of the SSE Composite Index.

Keywords: Bayes posterior estimator; restricted parameter space $(0, \infty)$; Stein's loss function; posterior expected loss; the normal distribution with normal-inverse-gamma prior

2010 Mathematics Subject Classification: 62F10; 62F15; 62C10

Citation: XIE Y H, SONG W H, ZHOU M Q, et al. The Bayes posterior estimator of the variance parameter of the normal distribution with a normal-inverse-gamma prior under Stein's loss [J]. Chinese J Appl Probab Statist, 2018, 34(6): 551–564.

§1. Introduction

Point estimation is an important class of statistical inference. The study of the performance as well as the optimality of point estimators are usually evaluated through the loss function. In Bayesian analysis, the Bayes risk is frequently computed to assess the performance of an estimator with respect to a given loss function. See [1] for techniques of finding Bayes posterior estimators.

*The project was supported by the Fundamental Research Funds for the Central Universities (Grant Nos. CQDXWL-2012-004; CDJRC10100010; 106112016CDJXY100002), the China Scholarship Council (Grant No. 201606055028), and the MOE Project of Humanities and Social Sciences on the West and the Border Area (Grant No. 14XJC910001).

*Corresponding author, E-mail: robertzhangyying@qq.com.

Received September 12, 2016. Revised August 20, 2017.

As pointed out by [2], the squared error loss function penalizes equally for overestimation and underestimation, which is appropriate for the location case. For the scale or variance case (see [3]) where 0 is a natural lower bound and the estimation problem is not symmetric, Stein's loss function instead of the squared error loss function should be chosen due to Stein's loss function penalizes gross overestimation and gross underestimation equally. Stein's loss function also arises out of the log likelihood function for σ^2 under a normal distribution, and thus ties together good decision theoretic and likelihood properties. See the supplement for a discussion. See [4] for a detailed discussion of the relationship between the log likelihood function and Stein's loss function. Consequently, Stein's loss function is recommended to use for the positive restricted parameter space $\Theta = (0, \infty)$ by many authors (see for instance [5–12]). Moreover, for estimating a covariance matrix which is assumed to be positive definite, many researchers exploit Stein's loss function (see for example [8, 13–21]. For the normal distribution with a normal-inverse-gamma prior, our parameter of interest is $\theta = \sigma^2$ which is a variance parameter. Therefore, we will choose Stein's loss function.

Zhang [12] analytically calculates the Bayes posterior estimator of the normal distribution with a conjugate inverse-gamma prior distribution under Stein's loss function where the mean parameter is assumed known. In this paper, we analytically calculate the Bayes posterior estimator of the normal distribution with a conjugate normal-inverse-gamma prior distribution under Stein's loss function, where the mean parameter is assumed unknown which is a more realistic and more sophisticated situation.

The rest of the paper is organized as follows. In the next Section 2, the Bayes posterior estimator, $\delta_s^\pi(\boldsymbol{x})$, with respect to a joint conjugate normal-inverse-gamma prior under Stein's loss function, as well as the Bayes posterior estimator, $\delta_2^\pi(\boldsymbol{x}) = \mathsf{E}(\theta \mid \boldsymbol{x})$, with respect to the same prior under the squared error loss function are both analytically calculated. In addition, we analytically calculate the Posterior Expected Stein's Loss (PESL) at $\delta_s^\pi(\boldsymbol{x})$ and $\delta_2^\pi(\boldsymbol{x})$. Section 3 reports vast amount of numerical simulation results to support the theoretical studies of (2) and (3), and that the PESLs depend only on v_0 and n, but do not depend on μ_0, κ_0, σ_0, and especially $\boldsymbol{x}$. In Section 4, we calculate the Bayes posterior estimators and the PESLs of the monthly simple returns of the SSE Composite Index. Section 5 concludes.

§2. Bayes Posterior Estimator, PESL, IRSL, and BRSL

In this section, we will analytically calculate the Bayes posterior estimator $\delta_s^\pi(\boldsymbol{x})$ of the parameter $\theta \in \Theta = (0, \infty)$ under Stein's loss function, the PESL at $\delta_s^\pi(\boldsymbol{x})$, $\mathrm{PESL}_s(\pi, \boldsymbol{x})$, and the Integrated Risk under Stein's Loss (IRSL) at $\delta_s^\pi(\boldsymbol{x})$, $\mathrm{IRSL}_s(\pi, \boldsymbol{x}) = \mathrm{BRSL}(\pi, \boldsymbol{x})$, which is also the Bayes Risk under Stein's Loss (BRSL). See [22] for the definitions of the posterior expected loss, the integrated risk, and the Bayes risk.

Suppose that the observations $X_1, X_2, \ldots, X_n$ are from the normal distribution with a normal-inverse-gamma prior:

$$\begin{cases} X_i \,|\, (\mu, \theta) \overset{\text{i.i.d.}}{\sim} \mathrm{N}(\mu, \theta), & i = 1, 2, \ldots, n, \\ \mu \,|\, \theta \sim \mathrm{N}(\mu_0,\, \theta/\kappa_0), & \theta \sim \mathrm{IG}(v_0/2,\, v_0\sigma_0^2/2), \end{cases} \tag{1}$$

where $-\infty < \mu_0 < \infty$, $\kappa_0 > 0$, $v_0 > 0$, and $\sigma_0 > 0$ are known hyper-parameters, $\mathrm{N}(\mu, \theta)$ is a normal distribution with an unknown mean μ and an unknown variance θ, the conditional conjugate prior distribution of μ given θ is $\mathrm{N}(\mu_0,\, \theta/\kappa_0)$ which is a normal distribution with a known mean μ_0 and an unknown variance θ/κ_0, the marginal conjugate prior distribution of θ is $\mathrm{IG}(v_0/2,\, v_0\sigma_0^2/2)$ which is an inverse gamma distribution with a known shape parameter $v_0/2$ and a known rate parameter $v_0\sigma_0^2/2$. Note that the problem of finding the Bayes posterior estimator under a conjugate prior is a standard problem that is treated in almost every text on Mathematical Statistics. Raiffa, and Schlaifer [23] put forward the idea of selecting an appropriate prior from the conjugate family. Specifically, the posterior distribution of θ with a joint conjugate prior $\pi(\mu, \theta) \sim \mathrm{N}\text{-}\mathrm{IG}(\mu_0, \kappa_0, v_0, \sigma_0^2)$ which is the normal-inverse-gamma distribution, was studied in Example 1.5.1 (p. 20) of [24] and Part I (pp. 69–70) of [25]. However, they did not provide any Bayes posterior estimator of θ. Moreover, the normal distribution with a normal-inverse-gamma prior which assumes that μ is unknown is more realistic than the normal distribution with an inverse-gamma prior investigated by [12] which assumes that μ is known.

As pointed out by [12], the Bayes posterior estimator

$$\delta_s^\pi(\boldsymbol{x}) = \frac{1}{\mathsf{E}(1/\theta \,|\, \boldsymbol{x})}$$

minimizes the PESL, that is,

$$\delta_s^\pi(\boldsymbol{x}) = \arg\min_{a \in \mathscr{A}} \mathsf{E}[L_s(\theta, a) \,|\, \boldsymbol{x}],$$

where $\mathscr{A} = \{a(\boldsymbol{x}) : a(\boldsymbol{x}) > 0\}$ is an action space, $a = a(\boldsymbol{x}) > 0$ is an action (estimator), which is a function of $\boldsymbol{x}$ instead of θ,

$$L_s(\theta, a) = \frac{a}{\theta} - \ln \frac{a}{\theta} - 1$$

is Stein's loss function, and $\theta > 0$ is the unknown parameter of interest.

The usual Bayes posterior estimator of θ is $\delta_2^\pi(\boldsymbol{x}) = \mathsf{E}(\theta \,|\, \boldsymbol{x})$ which minimizes the Posterior Expected Squared Error Loss. It is interesting to note that

$$\delta_2^\pi(\boldsymbol{x}) \geqslant \delta_s^\pi(\boldsymbol{x}), \tag{2}$$

whose proof exploits Jensen's inequality and the proof can be found in [12]. As calculated in [12], the PESL at $\delta_s^\pi(\boldsymbol{x}) = [\mathsf{E}(\theta^{-1} \,|\, \boldsymbol{x})]^{-1}$ is

$$\mathrm{PESL}_s(\pi, \boldsymbol{x}) = \mathsf{E}[L_s(\theta, a) \,|\, \boldsymbol{x}]\big|_{a=1/\mathsf{E}(1/\theta|\boldsymbol{x})} = \ln \mathsf{E}\Big(\frac{1}{\theta} \,\Big|\, \boldsymbol{x}\Big) + \mathsf{E}(\ln\theta \,|\, \boldsymbol{x}),$$

and the PESL at $\delta_2^\pi(\boldsymbol{x}) = \mathsf{E}(\theta \,|\, \boldsymbol{x})$ is

$$\begin{aligned}\mathrm{PESL}_2(\pi, \boldsymbol{x}) &= \mathsf{E}[L_s(\theta, a) \,|\, \boldsymbol{x}]\big|_{a=\mathsf{E}(\theta|\boldsymbol{x})} \\ &= \mathsf{E}(\theta \,|\, \boldsymbol{x})\mathsf{E}\Big(\frac{1}{\theta} \,\Big|\, \boldsymbol{x}\Big) - \ln \mathsf{E}(\theta \,|\, \boldsymbol{x}) + \mathsf{E}(\ln\theta \,|\, \boldsymbol{x}) - 1.\end{aligned}$$

Note that

$$\mathrm{PESL}_2(\pi, \boldsymbol{x}) \geqslant \mathrm{PESL}_s(\pi, \boldsymbol{x}), \tag{3}$$

which is a direct consequence of the general methodology for finding a Bayes posterior estimator or due to $\delta_s^\pi(\boldsymbol{x})$ minimizes the PESL. The numerical simulations will exemplify (2) and (3) later. Note that the calculations of $\delta_s^\pi(\boldsymbol{x})$, $\delta_2^\pi(\boldsymbol{x})$, $\mathrm{PESL}_s(\pi, \boldsymbol{x})$, and $\mathrm{PESL}_2(\pi, \boldsymbol{x})$ depend only on $\mathsf{E}(\theta \,|\, \boldsymbol{x})$, $\mathsf{E}(\theta^{-1} \,|\, \boldsymbol{x})$, and $\mathsf{E}(\ln\theta \,|\, \boldsymbol{x})$.

Suppose that $X \sim \mathrm{G}(\alpha, \beta)$ and $Y = 1/X \sim \mathrm{IG}(\alpha, \beta)$. More specifically, the pdfs of X and Y are respectively given by

$$\begin{aligned} f_X(x \,|\, \alpha, \beta) &= \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-x\beta}, \qquad x > 0,\ \alpha, \beta > 0, \\ f_Y(y \,|\, \alpha, \beta) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \Big(\frac{1}{y}\Big)^{\alpha+1} \mathrm{e}^{-\beta/y}, \qquad y > 0,\ \alpha, \beta > 0. \end{aligned}$$

We know that $\mathsf{E}X = \alpha/\beta$. It is easy to calculate

$$\mathsf{E}Y = \frac{\beta}{\alpha - 1}, \qquad \alpha > 1,\ \beta > 0.$$

The calculation of $\mathsf{E}Y$ can be found in the supplement.

The posterior distribution of the variance parameter θ of the normal distribution with a normal-inverse-gamma prior is

$$\pi(\theta \,|\, \boldsymbol{x}) \sim \mathrm{IG}(\alpha^*, \beta^*),$$

where

$$\alpha^* = \frac{v_0 + n}{2}, \qquad \beta^* = \frac{1}{2}\Big[v_0\sigma_0^2 + (n-1)s^2 + \frac{n\kappa_0}{n + \kappa_0}(\overline{x} - \mu_0)^2\Big],$$

$\overline{x}$ is the sample mean, and s^2 is the sample variance. The derivation of $\pi(\theta \mid \boldsymbol{x})$ can be found in the supplement, where it is also shown that the posterior distribution of the mean parameter μ is $\pi(\mu \mid \boldsymbol{x}) \sim \mathrm{t}_{v_n}(\mu_n, \sigma_n^2/\kappa_n)$, and the Bayes posterior estimator of μ is $\delta_2^{\pi,\mu}(\boldsymbol{x}) = \mathsf{E}(\mu \mid \boldsymbol{x}) = \mu_n$. Therefore, we have

$$\theta \mid \boldsymbol{x} \sim \mathrm{IG}(\alpha^*, \beta^*) \qquad \text{and} \qquad \frac{1}{\theta} \,\Big|\, \boldsymbol{x} \sim \mathrm{G}(\alpha^*, \beta^*).$$

Consequently,

$$\mathsf{E}(\theta \mid \boldsymbol{x}) = \frac{\beta^*}{\alpha^* - 1}, \quad \alpha^* > 1 \qquad \text{and} \qquad \mathsf{E}\Big(\frac{1}{\theta} \,\Big|\, \boldsymbol{x}\Big) = \frac{\alpha^*}{\beta^*}.$$

It is easy to obtain that

$$\delta_2^{\pi}(\boldsymbol{x}) = \mathsf{E}(\theta \mid \boldsymbol{x}) = \frac{\beta^*}{\alpha^* - 1} > \frac{\beta^*}{\alpha^*} = \frac{1}{\mathsf{E}(1/\theta \mid \boldsymbol{x})} = \delta_s^{\pi}(\boldsymbol{x}),$$

which exemplifies (2). To calculate the PESL at $\delta_s^{\pi}(\boldsymbol{x})$ and $\delta_2^{\pi}(\boldsymbol{x})$, we have to calculate $\mathsf{E}(\ln\theta \mid \boldsymbol{x})$. From [12], we have

$$\mathsf{E}(\ln\theta \mid \boldsymbol{x}) = \ln\beta^* - \psi(\alpha^*),$$

where $\psi(z) = \Gamma'(z)/\Gamma(z)$ is the digamma function, and $\Gamma(z)$ is the gamma function. In R software [26], the function digamma(z) calculates $\psi(z)$. Note that our forms of the pdfs of the gamma as well as inverse gamma distributions are different from those in [12].

Consequently, the PESL at $\delta_s^{\pi}(\boldsymbol{x})$ and $\delta_2^{\pi}(\boldsymbol{x})$ are respectively given by

$$\mathrm{PESL}_s(\pi, \boldsymbol{x}) = \ln\Big(\frac{\alpha^*}{\beta^*}\Big) + \ln\beta^* - \psi(\alpha^*) = \ln\alpha^* - \psi(\alpha^*)$$

and

$$\begin{aligned}\mathrm{PESL}_2(\pi, \boldsymbol{x}) &= \frac{\beta^*}{\alpha^* - 1}\frac{\alpha^*}{\beta^*} - \ln\frac{\beta^*}{\alpha^* - 1} + \ln\beta^* - \psi(\alpha^*) - 1 \\ &= \frac{1}{\alpha^* - 1} + \ln(\alpha^* - 1) - \psi(\alpha^*), \qquad \text{for } \alpha^* > 1.\end{aligned}$$

It is straightforward to check that

$$\mathrm{PESL}_2(\pi, \boldsymbol{x}) = \frac{1}{\alpha^* - 1} + \ln(\alpha^* - 1) - \psi(\alpha^*) \geqslant \ln\alpha^* - \psi(\alpha^*) = \mathrm{PESL}_s(\pi, \boldsymbol{x}),$$

which exemplifies (3). Note that $\mathrm{PESL}_s(\pi, \boldsymbol{x})$ and $\mathrm{PESL}_2(\pi, \boldsymbol{x})$ of the current paper are the same as those in [12], which is not surprising since the posterior distributions of θ are the same. We also note that $\mathrm{PESL}_s(\pi, \boldsymbol{x})$ and $\mathrm{PESL}_2(\pi, \boldsymbol{x})$ depend only on $\alpha^* = (v_0+n)/2$, and thus on v_0 and n. However, they do not depend on μ_0, κ_0, σ_0, and especially $\boldsymbol{x}$, which is quite interesting. The numerical simulations will exemplify this result later.

The IRSL at $\delta_s^\pi(\boldsymbol{x})$, which is the smallest IRSL, or the BRSL is

$$\begin{aligned}
&\mathrm{IRSL}_s(\pi,\boldsymbol{x})=\mathrm{BRSL}(\pi,\boldsymbol{x})=r_s(\pi,\boldsymbol{x})=\int_{\mathscr{X}}\mathrm{PESL}(\pi,a\,|\,\boldsymbol{x})m(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\Big|_{a=\delta_s^\pi(\boldsymbol{x})}\\
&=\int_{\mathscr{X}}\Big[a\mathsf{E}\Big(\frac{1}{\theta}\,\Big|\,\boldsymbol{x}\Big)-\ln a+\mathsf{E}(\ln\theta\,|\,\boldsymbol{x})-1\Big]m(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\Big|_{a=\delta_s^\pi(\boldsymbol{x})}\\
&=\int_{\mathscr{X}}\Big[a\frac{\alpha^*}{\beta^*}-\ln a+\ln\beta^*-\psi(\alpha^*)-1\Big]\Big|_{\substack{\alpha^*=(v_0+n)/2\\ \beta^*=[v_0\sigma_0^2+(n-1)s^2+n\kappa_0(\overline{x}-\mu_0)^2/(n+\kappa_0)]/2}}m(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\Big|_{a=\delta_s^\pi(\boldsymbol{x})}\\
&\neq\int_{\mathscr{X}}\mathrm{PESL}(\pi,a\,|\,\boldsymbol{x})\big|_{a=\delta_s^\pi(\boldsymbol{x})}m(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\triangleq I,
\end{aligned}$$

since we cannot take the evaluation of $a=\delta_s^\pi(\boldsymbol{x})$ into the integral of $\boldsymbol{x}$. The BRSL is such complicated that we do not intend to calculate it. To be honest, we do not know how to analytically calculate the smallest IRSL, or the BRSL. We also do not know how to numerically compute it. We find that

$$\begin{aligned}
I&=\int_{\mathscr{X}}\mathrm{PESL}_s(\pi,\boldsymbol{x})m(\boldsymbol{x})\mathrm{d}\boldsymbol{x}=\int_{\mathscr{X}}[\ln\alpha^*-\psi(\alpha^*)]\big|_{\alpha^*=(v_0+n)/2}m(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\\
&=[\ln\alpha^*-\psi(\alpha^*)]\big|_{\alpha^*=(v_0+n)/2}\int_{\mathscr{X}}m(\boldsymbol{x})\mathrm{d}\boldsymbol{x}=[\ln\alpha^*-\psi(\alpha^*)]\big|_{\alpha^*=(v_0+n)/2}.
\end{aligned}$$

However, the BRSL is not equal to I.

§3. Numerical Simulations

In this section, we will numerically exemplify the theoretical studies of (2) and (3), and that $\mathrm{PESL}_s(\pi,\boldsymbol{x})$ and $\mathrm{PESL}_2(\pi,\boldsymbol{x})$ depend only on v_0 and n, but do not depend on μ_0, κ_0, σ_0, and especially $\boldsymbol{x}$. Firstly, we fix $\mu_0=0$, $\kappa_0=1$, $\sigma_0=2$, $v_0=2$, and $n=10$. Secondly, we set seed(1) and draw one θ from $\mathrm{IG}(v_0/2,v_0\sigma_0^2/2)$, and then we draw one μ from $\mathrm{N}(\mu_0,\theta/\kappa_0)$. After that, we draw a random sample $\boldsymbol{x}$ from $\mathrm{N}(\mu,\theta)$. The generated sample $\boldsymbol{x}$ is

$$\boldsymbol{x}=(13.21,\ 8.86,\ -1.07,\ 2.04,\ 5.26,\ 6.72,\ 18.96,\ 10.63,\ 2.70,\ 0.92)'.$$

For simplicity, all the numerical results in this section are rounded to 2 or 3 decimal places. Figure 1 shows the histogram of $\theta\,|\,\boldsymbol{x}$ and the density estimation curve of $\pi(\theta\,|\,\boldsymbol{x})$. We then find $\delta_s^\pi(\boldsymbol{x})$ to minimize the PESL under $\pi(\theta\,|\,\boldsymbol{x})$. Numerical results show that

$$\delta_2^\pi(\boldsymbol{x})=39.66>33.05=\delta_s^\pi(\boldsymbol{x})$$

and

$$\mathrm{PESL}_2(\pi,\boldsymbol{x})=0.10>0.09=\mathrm{PESL}_s(\pi,\boldsymbol{x}),$$

which exemplify the theoretical studies of (2) and (3).

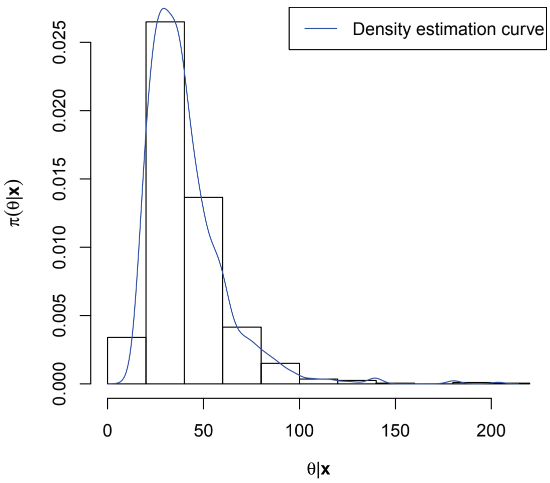

Figure 1 The histogram of $\theta\,|\,\boldsymbol{x}$ and the density estimation curve of $\pi(\theta\,|\,\boldsymbol{x})$

In Figure 2, we fix $\mu_0 = 0$, $\kappa_0 = 1$, $\sigma_0 = 2$, $\upsilon_0 = 2$, and $n = 10$. However, we allow the random seed number to change from 1 to 10 (i.e., $\boldsymbol{x}$ is changed). From Figure 2 we see that the estimators are functions of $\boldsymbol{x}$. Specifically, the left plot of Figure 2 illustrates that the estimators depend on $\boldsymbol{x}$, as well as $\delta_2^\pi(\boldsymbol{x})$ are unanimously larger than $\delta_s^\pi(\boldsymbol{x})$. The right plot of Figure 2 exhibits that the PESLs do not depend on $\boldsymbol{x}$, and $\mathrm{PESL}_2(\pi, \boldsymbol{x})$ are unanimously larger than $\mathrm{PESL}_s(\pi, \boldsymbol{x})$. The numerical values of the Bayes posterior estimators and the PESLs of Figure 2 are summarized in Table 1. The results of Figure 2 and Table 1 exemplify the theoretical studies of (2) and (3), and that the PESLs do not depend on $\boldsymbol{x}$.

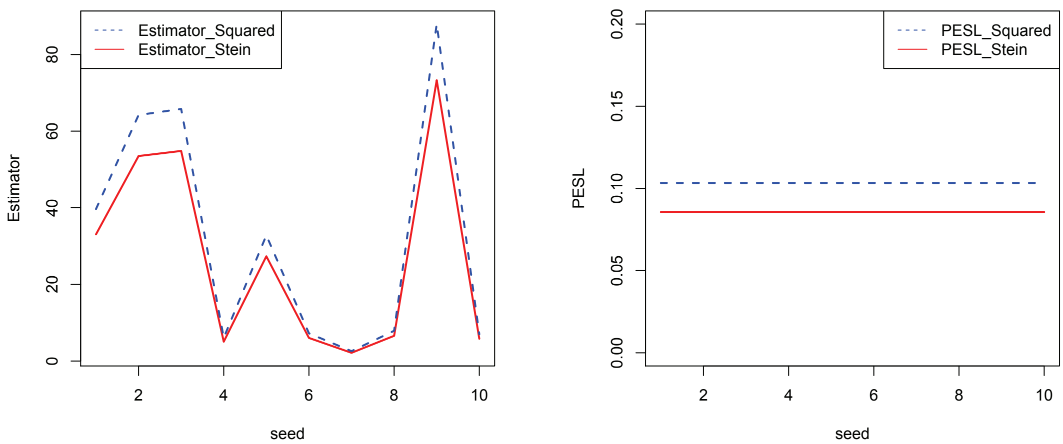

Figure 2 The estimators are functions of x (left) and the PESLs are functions of x (right)

Table 1 The numerical values of the Bayes posterior estimators and the PESLs of Figure 2

seed	1	2	3	4	5	6	7	8	9	10
$\delta_2^\pi(\boldsymbol{x})$	39.66	64.20	65.80	6.06	32.81	7.23	2.58	7.90	87.94	6.99
$\delta_s^\pi(\boldsymbol{x})$	33.05	53.50	54.83	5.05	27.34	6.03	2.15	6.58	73.29	5.82
$\mathrm{PESL}_2(\pi,\boldsymbol{x})$	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
$\mathrm{PESL}_s(\pi,\boldsymbol{x})$	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09

Now we allow one of the five parameters μ_0, κ_0, σ_0, υ_0, and n to change, holding other parameters fixed. Figures 3 and 4 show the estimators and the PESLs as functions of μ_0, κ_0, σ_0, υ_0, and n. We see from the left plots of Figures 3 and 4 that the estimators depend on μ_0, κ_0, σ_0, υ_0, and n, and $\delta_2^\pi(\boldsymbol{x})$ are unanimously larger than $\delta_s^\pi(\boldsymbol{x})$. The right plots of Figures 3 and 4 exhibit that the PESLs do not depend on μ_0, κ_0, and σ_0, only depend on υ_0 and n. In addition, $\mathrm{PESL}_2(\pi,\boldsymbol{x})$ are unanimously larger than $\mathrm{PESL}_s(\pi,\boldsymbol{x})$. The numerical values of the Bayes posterior estimators and the PESLs of Figures 3 and 4 are summarized in Tables 2 and 3 respectively. The results of Figures 3 and 4 and Tables 2 and 3 exemplify the theoretical studies of (2) and (3), and that the PESLs depend only on υ_0 and n, but do not depend on μ_0, κ_0, and σ_0.

Table 2 The numerical values of the Bayes posterior estimators and the PESLs of Figure 3

μ_0	-4	-3	-2	-1	0	1	2	3	4	5
$\delta_2^\pi(x)$	39.66	17.58	27.22	8.32	17.85	25.46	14.97	18.42	17.99	40.87
$\delta_s^\pi(x)$	33.05	14.65	22.69	6.93	14.88	21.21	12.48	15.35	14.99	34.06
$\mathrm{PESL}_2(\pi,x)$	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
$\mathrm{PESL}_s(\pi,x)$	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
κ_0	1	2	3	4	5	6	7	8	9	10
$\delta_2^\pi(x)$	39.66	18.34	27.22	8.31	17.77	27.84	14.28	17.96	18.42	44.13
$\delta_s^\pi(x)$	33.05	15.28	22.68	6.92	14.81	23.20	11.90	14.97	15.35	36.77
$\mathrm{PESL}_2(\pi,x)$	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
$\mathrm{PESL}_s(\pi,x)$	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
σ_0	1	2	3	4	5	6	7	8	9	10
$\delta_2^\pi(x)$	9.92	37.68	17.33	13.07	39.80	2 389.82	61.06	220.40	77.84	7 974.80
$\delta_s^\pi(x)$	8.26	31.40	14.44	10.89	33.17	1 991.52	50.88	183.67	64.87	6 645.66
$\mathrm{PESL}_2(\pi,x)$	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
$\mathrm{PESL}_s(\pi,x)$	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09

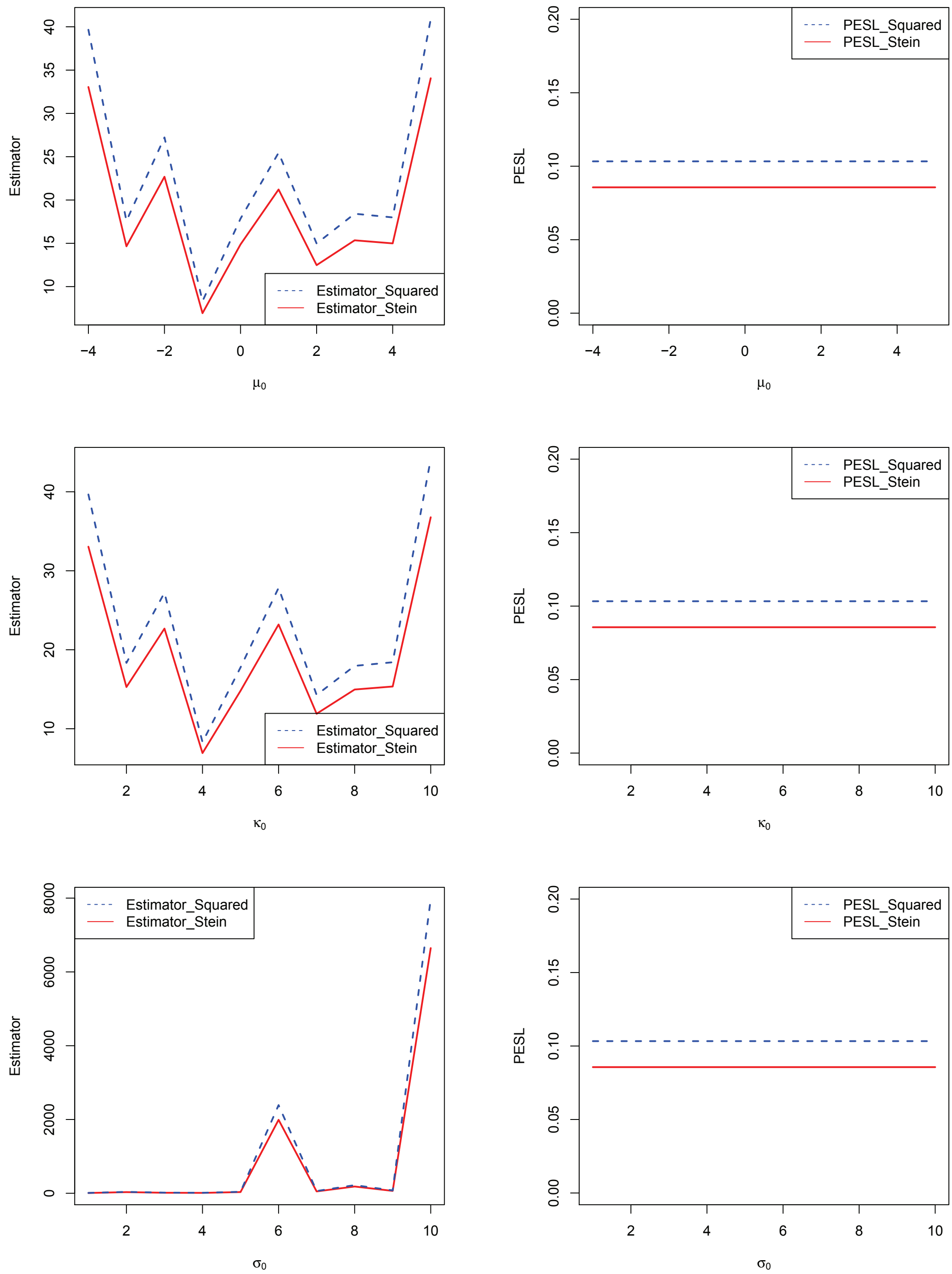

Figure 3 **Left: The estimators as functions of μ_0, κ_0, and σ_0**

Right: The PESLs as functions of μ_0, κ_0, and σ_0

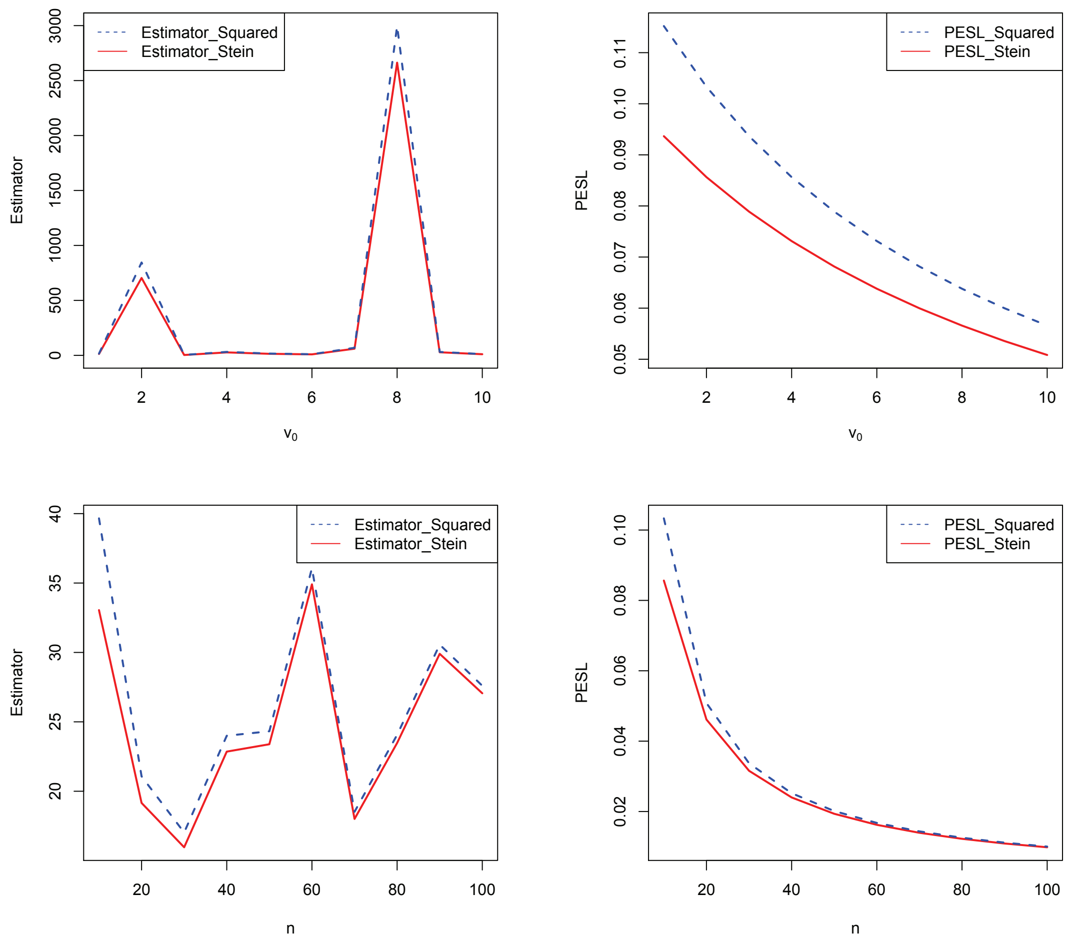

Figure 4 Left: The estimators as functions of v_0 and n
Right: The PESLs as functions of v_0 and n

Table 3 The numerical values of the Bayes posterior estimators and the PESLs of Figure 4

v_0	1	2	3	4	5	6	7	8	9	10
$\delta_2^\pi(x)$	15.289	845.088	4.029	31.491	16.357	10.349	68.704	2 995.897	31.336	11.566
$\delta_s^\pi(x)$	12.509	704.240	3.409	26.992	14.176	9.056	60.621	2 663.020	28.037	10.409
$\mathrm{PESL}_2(\pi,x)$	0.115	0.103	0.094	0.086	0.079	0.073	0.068	0.064	0.060	0.057
$\mathrm{PESL}_s(\pi,x)$	0.094	0.086	0.079	0.073	0.068	0.064	0.060	0.057	0.054	0.051
n	10	20	30	40	50	60	70	80	90	100
$\delta_2^\pi(x)$	39.660	21.073	17.028	23.998	24.322	36.071	18.520	24.063	30.565	27.598
$\delta_s^\pi(x)$	33.050	19.157	15.964	22.855	23.386	34.907	18.005	23.476	29.900	27.057
$\mathrm{PESL}_2(\pi,x)$	0.103	0.051	0.034	0.025	0.020	0.017	0.014	0.013	0.011	0.010
$\mathrm{PESL}_s(\pi,x)$	0.086	0.046	0.032	0.024	0.019	0.016	0.014	0.012	0.011	0.010

§4. Real Data Example

Many real world samples are from normal distributions with unknown mean and variance. In this section, we use a sample from finance. We exploit an R package **quantmod**[27] to download the data 000001.SS (the SSE Composite Index) during 2007-01-08 and 2017-01-13 from "finance.yahoo.com". It is commonly believed that the monthly simple returns of the stock data or the index data are normally distributed. The simple return is calculated as

$$r = \frac{S_1 - S_0}{S_0},$$

where S_0 and S_1 are the close prices of some month and the next month of the data 000001.SS. The monthly simple returns of the data 000001.SS pass the normality test by the function shapiro.test( ) in R. The histogram of the SSE monthly simple returns are given in Figure 5. From the figure we see that the density estimation curve and the normal approximation curve are very close, which further exemplifies that the SSE monthly simple returns are normally distributed.

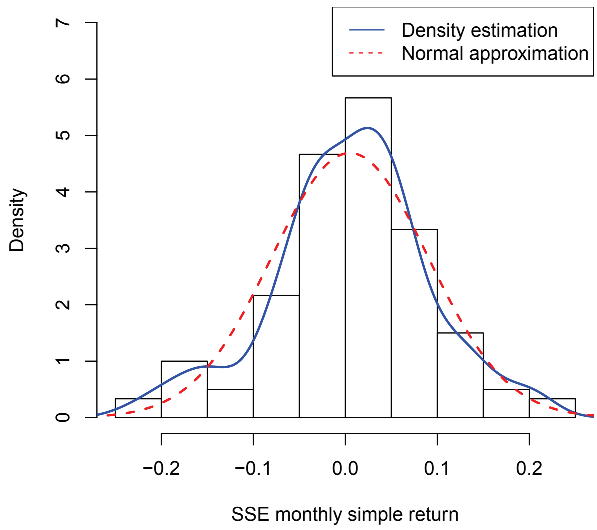

Figure 5 The histogram of the SSE monthly simple returns

Therefore, it is reasonable to assume that the SSE monthly simple returns follow the normal model (1). For simplicity, we assume that

$$\mu_0 = 0, \quad \kappa_0 = 1, \quad \upsilon_0 = 10, \quad \sigma_0 = 1.$$

Alternatively, one could calculate the empirical Bayes estimators of the hyper-parameters of model (1) by the moment method or the MLE method. Numerical results show that

$$\delta_2^\pi(\boldsymbol{x}) = 0.08485648 > 0.083551 = \delta_s^\pi(\boldsymbol{x})$$

and

$$\mathrm{PESL}_2(\pi, \boldsymbol{x}) = 0.007832845 > 0.007712031 = \mathrm{PESL}_s(\pi, \boldsymbol{x}),$$

which exemplify the theoretical studies of (2) and (3). Note that we believe that $\delta_s^\pi(\boldsymbol{x})$ is better than $\delta_2^\pi(\boldsymbol{x})$, not because $\delta_s^\pi(\boldsymbol{x})$ is larger or smaller than $\delta_2^\pi(\boldsymbol{x})$, but because Stein's loss function is more appropriate than the squared error loss function for the variance parameter case.

§5. Conclusion

For the variance parameter θ of the normal distribution with a normal-inverse-gamma prior, we recommend and analytically calculate the Bayes posterior estimator, $\delta_s^\pi(\boldsymbol{x})$, with respect to a conjugate prior $\mu \,|\, \theta \sim \mathrm{N}(\mu_0, \theta/\kappa_0)$ and $\theta \sim \mathrm{IG}(v_0/2, v_0\sigma_0^2/2)$ under Stein's loss function which penalizes gross overestimation and gross underestimation equally. This estimator minimizes the PESL. As comparisons, the Bayes posterior estimator, $\delta_2^\pi(\boldsymbol{x}) = \mathsf{E}(\theta \,|\, \boldsymbol{x})$, with respect to the same conjugate prior under the squared error loss function, and the PESL at $\delta_2^\pi(\boldsymbol{x})$ are calculated. The calculations of $\delta_s^\pi(\boldsymbol{x})$, $\delta_2^\pi(\boldsymbol{x})$, $\mathrm{PESL}_s(\pi, \boldsymbol{x})$, and $\mathrm{PESL}_2(\pi, \boldsymbol{x})$ depend only on $\mathsf{E}(\theta \,|\, \boldsymbol{x})$, $\mathsf{E}(\theta^{-1} \,|\, \boldsymbol{x})$, and $\mathsf{E}(\ln\theta \,|\, \boldsymbol{x})$. The numerical simulations exemplify our theoretical studies that the PESLs depend only on v_0 and n, but do not depend on μ_0, κ_0, σ_0, and especially $\boldsymbol{x}$. The estimators $\delta_2^\pi(\boldsymbol{x})$ are unanimously larger than the estimators $\delta_s^\pi(\boldsymbol{x})$, as well as $\mathrm{PESL}_2(\pi, \boldsymbol{x})$ are unanimously larger than $\mathrm{PESL}_s(\pi, \boldsymbol{x})$. Finally, we calculate the Bayes posterior estimators and the PESLs of the monthly simple returns of the SSE Composite Index, which also exemplify the theoretical studies of (2) and (3).

Acknowledgments The authors gratefully acknowledge the constructive comments offered by the anonymous referee. His/her comments improve the quality of the paper significantly.

Supporting Information: Additional information for this article is available.

Supplement: Some proofs of the article.

R folder: R codes used in the article.

References

[1] BROWN L D, PURVES R. Measurable selections of extrema [J]. *Ann Statist*, 1973, **1(5)**: 902–912.

[2] CASELLA G, BERGER R L. *Statistical Inference* [M]. 2nd ed. Pacific Grove: Duxbury Press, 2001.

[3] LIU Y J, SUN D C. Should we use asis? [J]. *Chinese J Appl Probab Statist*, 2014, **30(1)**: 1–11.

[4] ANDERSON T W. *An Introduction to Multivariate Statistical Analysis* [M]. 2nd ed. New York: Wiley, 1984.

[5] BOBOTAS P, KOUROUKLIS S. On the estimation of a normal precision and a normal variance ratio [J]. *Stat Methodol*, 2010, **7(4)**: 445–463.

[6] BROWN L D. Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters [J]. *Ann Math Statist*, 1968, **39(1)**: 29–48.

[7] BROWN L D. [Developments in decision-theoretic variance estimation]: comment [J]. *Statist Sci*, 1990, **5(1)**: 103–106.

[8] JAMES W, STEIN C. Estimation with quadratic loss [C] // NEYMAN J. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, Berkeley, CA: University of California Press, 1961: 361–379.

[9] OONO Y, SHINOZAKI N. On a class of improved estimators of variance and estimation under order restriction [J]. *J Statist Plann Inference*, 2006, **136(8)**: 2584–2605.

[10] PARSIAN A, NEMATOLLAHI N. Estimation of scale parameter under entropy loss function [J]. *J Statist Plann Inference*, 1996, **52(1)**: 77–91.

[11] PETROPOULOS C, KOUROUKLIS S. Estimation of a scale parameter in mixture models with unknown location [J]. *J Statist Plann Inference*, 2005, **128(1)**: 191–218.

[12] ZHANG Y Y. The Bayes rule of the variance parameter of the hierarchical normal and inverse gamma model under Stein's loss [J]. *Comm Statist Theory Methods*, 2017, **46(14)**: 7125–7133.

[13] DEY D K, SRINIVASAN C. Estimation of a covariance matrix under Stein's loss [J]. *Ann Statist*, 1985, **13(4)**: 1581–1591.

[14] KONNO Y. Estimation of a normal covariance matrix with incomplete data under Stein's loss [J]. *J Multivariate Anal*, 1995, **52(2)**: 308–324.

[15] KONNO Y. Estimation of normal covariance matrices parametrized by irreducible symmetric cones under Stein's loss [J]. *J Multivariate Anal*, 2007, **98(2)**: 295–316.

[16] MA T F, JIA L J, SU Y S. A new estimator of covariance matrix [J]. *J Statist Plann Inference*, 2012, **142(2)**: 529–536.

[17] SHEENA Y, TAKEMURA A. Inadmissibility of non-order-preserving orthogonally invariant estimators of the covariance matrix in the case of Stein's loss [J]. *J Multivariate Anal*, 1992, **41(1)**: 117–131.

[18] SUN X Q, SUN D C, HE Z Q. Bayesian inference on multivariate normal covariance and precision matrices in a star-shaped model with missing data [J]. *Comm Statist Theory Methods*, 2010, **39(4)**: 642–666.

[19] TSUKUMA H. Estimation of a high-dimensional covariance matrix with the Stein loss [J]. *J Multivariate Anal*, 2016,**148**: 1–17.

[20] XU K, HE D J. Further results on estimation of covariance matrix [J]. *Statist Probab Lett*, 2015, **101**: 11–20.

[21] YE R D, WANG S G. Improved estimation of the covariance matrix under Stein's loss [J]. *Statist Probab Lett*, 2009, **79(6)**: 715–721.

[22] ROBERT C P. *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation* [M]. 2nd ed. New York: Springer, 2007.

[23] RAIFFA H, SCHLAIFER R. *Applied Statistical Decision Theory* [M]. Cambridge, MA: Harvard University Press, 1961.

[24] MAO S S, TANG Y C. *Bayesian Statistics* [M]. 2nd ed. Beijing: China Statistics Press, 2012. (in Chinese)

[25] CHEN M H. Bayesian statistics lecture [R]. Statistics Graduate Summer School, School of Mathematics and Statistics, Northeast Normal University, Changchun, China, July 2014.

[26] R Core Team. R: a language and environment for statistical computing [CP]. Vienna, Austria: R Foundation for Statistical Computing, 2017.

[27] RYAN J A. quantmod: quantitative financial modelling framework [CP]. R package version 0.4-7, 2016.

具有正态逆伽玛先验的正态分布中的方差参数在 Stein 损失下的贝叶斯后验估计量

解宇涵　宋文和　周明琴　张应应

(重庆大学统计与精算学系, 重庆, 401331)

摘　要: 对于先验分布为正态逆伽玛分布的正态分布的方差参数, 我们解析地计算了具有共轭的正态逆伽玛先验分布的在 Stein 损失函数下的贝叶斯后验估计量. 这个估计量最小化后验期望 Stein 损失. 我们还解析地计算了在平方误差损失函数下的贝叶斯后验估计量和后验期望 Stein 损失. 数值模拟的结果例证了我们的如下理论研究: 后验期望 Stein 损失不依赖于样本; 在平方误差损失函数下的贝叶斯后验估计量和后验期望 Stein 损失要一致地大于在 Stein 损失函数下的对应的量. 最后, 我们计算了上证综指的月度的简单回报的贝叶斯后验估计量和后验期望 Stein 损失.

关键词: 贝叶斯后验估计量; 限制参数空间 $(0,\infty)$; Stein 的损失函数; 后验期望损失; 具有正态逆伽玛先验的正态分布

中图分类号: O212