



The Bayes Estimators of the Variance and Scale Parameters of the Normal Model With a Known Mean for the Conjugate and Noninformative Priors Under Stein's Loss

Ying-Ying Zhang^{1,2*}, Teng-Zhong Rong^{1,2} and Man-Man Li^{1,2}

¹Department of Statistics and Actuarial Science, College of Mathematics and Statistics, Chongqing University, Chongqing, China, ²Chongqing Key Laboratory of Analytic Mathematics and Applications, Chongqing University, Chongqing, China

OPEN ACCESS

Edited by:

Niansheng Tang,
Yunnan University, China

Reviewed by:

Guikai Hu,
East China University of Technology,
China

Akio Namba,
Kobe University, Japan

Tabassum Sindhu,
Quaid-i-Azam University, Pakistan

*Correspondence:

Ying-Ying Zhang
robertzhangying@qq.com

Specialty section:

This article was submitted to
Data Science,
a section of the journal
Frontiers in Big Data

Received: 24 August 2021

Accepted: 01 November 2021

Published: 03 January 2022

Citation:

Zhang Y-Y, Rong T-Z and Li M-M
(2022) The Bayes Estimators of the
Variance and Scale Parameters of the
Normal Model With a Known Mean for
the Conjugate and Noninformative
Priors Under Stein's Loss.
Front. Big Data 4:763925.
doi: 10.3389/fdata.2021.763925

For the normal model with a known mean, the Bayes estimation of the variance parameter under the conjugate prior is studied in Lehmann and Casella (1998) and Mao and Tang (2012). However, they only calculate the Bayes estimator with respect to a conjugate prior under the squared error loss function. Zhang (2017) calculates the Bayes estimator of the variance parameter of the normal model with a known mean with respect to the conjugate prior under Stein's loss function which penalizes gross overestimation and gross underestimation equally, and the corresponding Posterior Expected Stein's Loss (PESL). Motivated by their works, we have calculated the Bayes estimators of the variance parameter with respect to the noninformative (Jeffreys's, reference, and matching) priors under Stein's loss function, and the corresponding PESLs. Moreover, we have calculated the Bayes estimators of the scale parameter with respect to the conjugate and noninformative priors under Stein's loss function, and the corresponding PESLs. The quantities (prior, posterior, three posterior expectations, two Bayes estimators, and two PESLs) and expressions of the variance and scale parameters of the model for the conjugate and noninformative priors are summarized in two tables. After that, the numerical simulations are carried out to exemplify the theoretical findings. Finally, we calculate the Bayes estimators and the PESLs of the variance and scale parameters of the S&P 500 monthly simple returns for the conjugate and noninformative priors.

Keywords: Bayes estimator, variance and scale parameters, normal model, conjugate and noninformative priors, Stein's loss

1 INTRODUCTION

There are four basic elements in Bayesian decision theory and specifically in Bayesian point estimation: The data, the model, the prior, and the loss function. In this paper, we are interested in the data from the normal model with a known mean, with respect to the conjugate and noninformative (Jeffreys's, reference, and matching) priors, under Stein's and the squared error loss functions. We will analytically calculate the Bayes estimators of the variance and scale parameters of the normal model with a known mean, with respect to the conjugate and noninformative priors under Stein's and the squared error loss functions.

The squared error loss function has been used by many authors for the problem of estimating the variance, σ^2 , based on a random sample from a normal distribution (see for instance (Maatta and Casella, 1990)). As pointed out by (Casella and Berger, 2002), the squared error loss function penalizes overestimation and underestimation equally, which is fine for the location parameter with parameter space $\Theta = (-\infty, \infty)$. For a variance or scale parameter, the parameter space is $\Theta = (0, \infty)$ where 0 is a natural lower bound and the estimation problem is not symmetric. In these cases, we should not choose the squared error loss function, but choose a loss function which penalizes gross overestimation and gross underestimation equally, that is, an action a will incur an infinite loss when it tends to 0 or ∞ . Stein's loss function has this property, and thus it is recommended to use for the positive restricted parameter space $\Theta = (0, \infty)$ by many authors (see for example (James and Stein, 1961; Petropoulos and Kourouklis, 2005; Oono and Shinozaki, 2006; Bobotas and Kourouklis, 2010; Zhang, 2017; Xie et al., 2018; Zhang et al., 2019; Sun et al., 2021)). In the normal model with a known mean μ , our parameters of interest are $\theta = \sigma^2$ (a variance parameter) and $\theta = \sigma$ (a scale parameter). Therefore, we will select Stein's loss function.

The motivation and contributions of our paper are summarized as follows. For the normal model with a known mean μ , the Bayes estimation of the variance parameter $\theta = \sigma^2$ under the conjugate prior which is an Inverse Gamma distribution is studied in Example 4.2.5 (p.236) of (Lehmann and Casella, 1998) and Example 1.3.5 (p.15) of (Mao and Tang, 2012). However, they only calculate the Bayes estimator with respect to a conjugate prior under the squared error loss. (Zhang, 2017) calculates the Bayes estimator of the variance parameter $\theta = \sigma^2$ of the normal model with a known mean with respect to the conjugate prior under Stein's loss function which penalizes gross overestimation and gross underestimation equally, and the corresponding Posterior Expected Stein's Loss (PESL). Motivated by the works of (Lehmann and Casella, 1998; Mao and Tang, 2012; Zhang, 2017), we want to calculate the Bayes estimators of the variance and scale parameters of the normal model with a known mean for the conjugate and noninformative priors under Stein's loss function. The contributions of our paper are summarized as follows. In this paper, we have calculated the Bayes estimators of the variance parameter $\theta = \sigma^2$ with respect to the noninformative (Jeffreys's, reference, and matching) priors under Stein's loss function, and the corresponding Posterior Expected Stein's Losses (PESLs). Moreover, we have calculated the Bayes estimators of the scale parameter $\theta = \sigma$ with respect to the conjugate and noninformative priors under Stein's loss function, and the corresponding PESLs. For more literature on Bayesian estimation and inference, we refer readers to (Sindhu and Aslam, 2013a; Sindhu and Aslam, 2013b; Sindhu et al., 2013; Sindhu et al., 2016a; Sindhu et al., 2016b; Sindhu et al., 2016c; Sindhu et al., 2017; Sindhu et al., 2018; Sindhu and Hussain, 2018)

The rest of the paper is organized as follows. In the next **Section 2**, we analytically calculate the Bayes estimators of the variance and scale parameters of the normal model with a known mean, with respect to the conjugate and noninformative priors under Stein's loss function, and the corresponding PESLs. We

also analytically calculate the Bayes estimators under the squared error loss function, and the corresponding PESLs. The quantities (prior, posterior, three posterior expectations, two Bayes estimators, and two PESLs) and expressions of the variance and scale parameters for the conjugate and noninformative priors are summarized in two tables. **Section 3** reports vast amount of numerical simulation results of the combination of the noninformative prior and the scale parameter to support the theoretical studies of two inequalities of the Bayes estimators and the PESLs, and that the PESLs depend only on the number of observations, but do not depend on the mean and the sample. In **Section 4**, we calculate the Bayes estimators and the PESLs of the variance and scale parameters of the S&P 500 monthly simple returns for the conjugate and noninformative priors. Some conclusions and discussions are provided in **Section 5**.

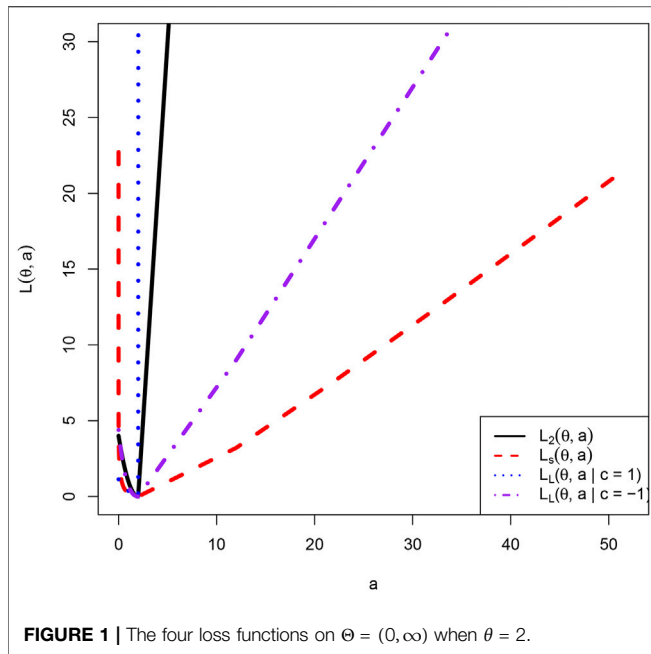
2 BAYES ESTIMATOR, PESL, IRSL, AND BRSL

In this section, we will analytically calculate the Bayes estimator $\delta_s^{\pi, \theta}(\mathbf{x})$ of the variance parameter $\theta = \sigma^2 \in \Theta = (0, \infty)$ under Stein's loss function, the PESL at $\delta_s^{\pi, \theta}(\mathbf{x})$, $PESL_s^{\pi, \theta}(\mathbf{x})$, and the Integrated Risk under Stein's Loss (IRSL) at $\delta_s^{\pi, \theta}$, $IRSL_s^{\pi, \theta} = BRSL^{\pi, \theta}$, which is also the Bayes Risk under Stein's Loss (BRSL) for π, θ . See (Robert, 2007) for the definitions of the posterior expected loss, the integrated risk, and the Bayes risk. We will also analytically calculate the Bayes estimator $\delta_s^{\pi, \sigma}(\mathbf{x})$ of the scale parameter $\sigma \in \Theta = (0, \infty)$ under Stein's loss function, the PESL at $\delta_s^{\pi, \sigma}(\mathbf{x})$, $PESL_s^{\pi, \sigma}(\mathbf{x})$, and the IRSL at $\delta_s^{\pi, \sigma}$, $IRSL_s^{\pi, \sigma} = BRSL^{\pi, \sigma}$, which is also the BRSL for π, σ .

Suppose that we observe X_1, X_2, \dots, X_n from the hierarchical normal model with a mixing variance parameter $\theta = \sigma^2$:

$$\begin{cases} X_i | \theta \stackrel{iid}{\sim} N(\mu, \theta), & i = 1, 2, \dots, n, \\ \theta \sim \pi(\theta), \end{cases} \quad (1)$$

where $-\infty < \mu < \infty$ is a known constant, $N(\mu, \theta)$ is the normal distribution with a known mean μ and an unknown variance θ , and $\pi(\theta)$ is the prior distribution of θ . For the normal model with a known mean μ , the Bayes estimation of the variance parameter $\theta = \sigma^2$ under the conjugate prior which is an Inverse Gamma distribution is studied in Example 4.2.5 (p.236) of (Lehmann and Casella, 1998) and Example 1.3.5 (p.15) of (Mao and Tang, 2012). However, they only calculate the Bayes estimator with respect to a conjugate prior under the squared error loss. (Zhang, 2017) calculates the Bayes estimator of the variance parameter $\theta = \sigma^2$ with respect to the conjugate prior under Stein's loss function, and the corresponding PESL. Motivated by the works of (Lehmann and Casella, 1998; Mao and Tang, 2012; Zhang, 2017), we want to calculate the Bayes estimators of the variance parameter of the normal model with a known mean for the noninformative (Jeffreys's, reference, and matching) priors under Stein's loss function. The usual Bayes estimator with respect to a prior $\pi(\theta)$ is to calculate $\delta_2^{\pi, \theta}(\mathbf{x}) = E(\theta | \mathbf{x})$ under the squared error loss function. As pointed out in the introduction, we should calculate and use the Bayes estimator



of the variance parameter θ with respect to a prior $\pi(\theta)$ under Stein's loss function, that is, $\delta_s^{\pi, \theta}(\mathbf{x})$.

Alternatively, we may be interested in the scale parameter $\theta = \sigma$. Motivated by the works of (Lehmann and Casella, 1998; Mao and Tang, 2012; Zhang, 2017), we also want to calculate the Bayes estimators of the scale parameter $\theta = \sigma$ with respect to the conjugate and noninformative priors under Stein's loss function, and the corresponding PESLs. Suppose that we observe X_1, X_2, \dots, X_n from the hierarchical normal model with a mixing scale parameter $\theta = \sigma$:

$$\begin{cases} X_i | \sigma \stackrel{iid}{\sim} N(\mu, \sigma^2), & i = 1, 2, \dots, n, \\ \sigma \sim \pi(\sigma), \end{cases} \quad (2)$$

where $-\infty < \mu < \infty$ is a known constant, $N(\mu, \sigma^2)$ is the normal distribution with a known mean μ and an unknown variance σ^2 , and $\pi(\sigma)$ is the prior distribution of σ . The usual Bayes estimator with respect to a prior $\pi(\sigma)$ is to calculate $\delta_2^{\pi, \sigma}(\mathbf{x}) = E(\sigma | \mathbf{x})$ under the squared error loss function. As pointed out in the introduction, we should calculate and use the Bayes estimator of the scale parameter σ with respect to a prior $\pi(\sigma)$ under Stein's loss function, that is, $\delta_s^{\pi, \sigma}(\mathbf{x})$.

Now let us explain why we choose Stein's loss function on $\Theta = (0, \infty)$. Stein's loss function is given by

$$L_s(\theta, a) = \frac{a}{\theta} - \log \frac{a}{\theta} - 1, \quad (3)$$

where $\theta > 0$ is the unknown parameter of interest and a is an action or estimator. The squared error loss function is given by

$$L_2(\theta, a) = (a - \theta)^2. \quad (4)$$

The asymmetric Linear Exponential (LINEX) loss function ((Varian et al., 1975; Zellner, 1986; Robert, 2007)) is given by

$$L_L(\theta, a) = e^{c(a-\theta)} - c(a - \theta) - 1, \quad (5)$$

where $c \neq 0$ serving to determine its shape. In particular, when $c > 0$, the LINEX loss function tends to ∞ exponentially, while when $c < 0$, the LINEX loss function tends to ∞ linearly. Note that on the positive restricted parameter space $\Theta = (0, \infty)$, Stein's loss function penalizes gross overestimation and gross underestimation equally, that is, an action a will incur an infinite loss when it tends to 0 or ∞ . Whereas, the squared error loss function does not penalize gross overestimation and gross underestimation equally, as an action a will incur a finite loss (in fact θ^2) when it tends to 0 and incur an infinite loss when it tends to ∞ . Similarly, the LINEX loss functions also do not penalize gross overestimation and gross underestimation equally, as an action a will incur a finite loss (in fact $e^{-c\theta} + c\theta - 1$) when it tends to 0 and incur an infinite loss when it tends to ∞ . **Figure 1** shows the four loss functions on $\Theta = (0, \infty)$ when $\theta = 2$.

As pointed out by (Zhang, 2017), the Bayes estimator

$$\delta_s^{\pi, \theta}(\mathbf{x}) = \frac{1}{E\left(\frac{1}{\theta} | \mathbf{x}\right)}$$

minimizes the PESL, that is,

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

where $\mathcal{A} = \{a(\mathbf{x}) : a(\mathbf{x}) > 0\}$ is an action space, $a = a(\mathbf{x}) > 0$ is an action (estimator), which is a function only of \mathbf{x} , $L_s(\theta, a)$ given by (Eq. 3) is Stein's loss function, and $\theta > 0$ is the unknown parameter of interest. Note that Stein's loss function has a nice property that it penalizes gross overestimation and gross underestimation equally, that is, an action a will incur an infinite loss when it tends to 0 or ∞ . Moreover, note that θ may be the variance parameter σ^2 or the scale parameter σ .

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

$$\delta_s^{\pi, \theta}(\mathbf{x}) \leq \delta_2^{\pi, \theta}(\mathbf{x}), \quad (6)$$

whose proof exploits Jensen's inequality and the proof can be found in (Zhang, 2017). Note that the inequality (Eq. 6) is a special inequality in (Zhang et al., 2018). As calculated in (Zhang, 2017), the PESL at $\delta_s^{\pi, \theta}(\mathbf{x}) = [E(\theta^{-1} | \mathbf{x})]^{-1}$ is

$$PESL_s^{\pi, \theta}(\mathbf{x}) = E[L_s(\theta, a) | \mathbf{x}]_{a=\frac{1}{E\left(\frac{1}{\theta} | \mathbf{x}\right)}} = \log E\left(\frac{1}{\theta} | \mathbf{x}\right) + E(\log \theta | \mathbf{x}),$$

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

$$\begin{aligned} PESL_2^{\pi, \theta}(\mathbf{x}) &= E[L_s(\theta, a) | \mathbf{x}]_{a=E(\theta | \mathbf{x})} \\ &= E(\theta | \mathbf{x}) E\left(\frac{1}{\theta} | \mathbf{x}\right) - \log E(\theta | \mathbf{x}) + E(\log \theta | \mathbf{x}) - 1. \end{aligned}$$

As observed in (Zhang, 2017),

$$PESL_s^{\pi, \theta}(\mathbf{x}) \leq PESL_2^{\pi, \theta}(\mathbf{x}), \quad (7)$$

which is a direct consequence of the general methodology for finding a Bayes estimator or due to $\delta_s^{\pi, \theta}(\mathbf{x})$ minimizes the PESL.

TABLE 1 | The quantities and expressions for the conjugate prior.

Quantities	Expressions for $\theta = \sigma^2$, the variance parameter	Expressions for $\theta = \sigma$, the scale parameter
$\pi_c(\theta)$	$IG(\alpha, \beta)$	$SRIG(\alpha, \beta)$
$\pi_c(\theta \mathbf{x})$	$IG(\alpha^*, \beta^*)$	$SRIG(\alpha^*, \beta^*)$
$E^{\pi_c}(\theta \mathbf{x})$	$\frac{1}{(\alpha^*-1)\beta^*}$, for $\alpha^* > 1$	$\frac{\Gamma(\alpha^*-\frac{1}{2})}{\Gamma(\alpha^*)\beta^{\frac{1}{2}}}$, for $\alpha^* > \frac{1}{2}$
$E^{\pi_c}(\frac{1}{\theta} \mathbf{x})$	$\alpha^*\beta^*$	$\frac{\Gamma(\alpha^*+\frac{1}{2})\beta^{\frac{1}{2}}}{\Gamma(\alpha^*)}$
$E^{\pi_c}(\log \theta \mathbf{x})$	$-\log \beta^* - \psi(\alpha^*)$	$-\frac{1}{2} \log \beta^* - \frac{1}{2} \psi(\alpha^*)$
$\delta_s^{\pi_c, \theta}(\mathbf{x})$	$\frac{1}{\alpha^*\beta^*}$	$\frac{\Gamma(\alpha^*)}{\Gamma(\alpha^*+\frac{1}{2})\beta^{\frac{1}{2}}}$
$\delta_2^{\pi_c, \theta}(\mathbf{x})$	$\frac{1}{(\alpha^*-1)\beta^*}$, for $\alpha^* > 1$	$\frac{\Gamma(\alpha^*-\frac{1}{2})}{\Gamma(\alpha^*)\beta^{\frac{1}{2}}}$, for $\alpha^* > \frac{1}{2}$
$PESL_s^{\pi_c, \theta}(\mathbf{x})$	$\log \alpha^* - \psi(\alpha^*)$	$\log \Gamma(\alpha^* + \frac{1}{2}) - \log \Gamma(\alpha^*) - \frac{1}{2} \psi(\alpha^*)$
$PESL_2^{\pi_c, \theta}(\mathbf{x})$	$\frac{1}{\alpha^*-1} + \log(\alpha^*-1) - \psi(\alpha^*)$, for $\alpha^* > 1$	$\frac{\Gamma(\alpha^* - \frac{1}{2})\Gamma(\alpha^* + \frac{1}{2})}{\Gamma^2(\alpha^*)} - 1 - \log \Gamma(\alpha^* - \frac{1}{2})$ $+ \log \Gamma(\alpha^*) - \frac{1}{2} \psi(\alpha^*)$, for $\alpha^* > \frac{1}{2}$

The numerical simulations will exemplify (Eqs 6, 7) later. Note that the calculations of $\delta_s^{\pi_c, \theta}(\mathbf{x})$, $\delta_2^{\pi_c, \theta}(\mathbf{x})$, $PESL_s^{\pi_c, \theta}(\mathbf{x})$, and $PESL_2^{\pi_c, \theta}(\mathbf{x})$ depend only on the three expectations $E(\theta|\mathbf{x})$, $E(\theta^{-1}|\mathbf{x})$, and $E(\log \theta|\mathbf{x})$.

2.1 Conjugate Prior

The problem of finding the Bayes estimator under a conjugate prior is a standard problem that is treated in almost every text on Mathematical Statistics.

The quantities and expressions of the variance and scale parameters of the normal models (Eqs 1, 2) with a known mean μ for the conjugate prior are summarized in Table 1. In the table, $\alpha > 0$ and $\beta > 0$ are known constants,

$$\alpha^* = \alpha + \frac{n}{2}, \beta^* = \left[\frac{1}{\beta} + \frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2 \right]^{-1}$$

$$\psi(z) = \frac{\Gamma'(z)}{\Gamma(z)} = \frac{d}{dz} \log \Gamma(z) = \text{digamma}(z)$$

is the digamma function, and $\Gamma(z)$ is the gamma function. In R software (R Core Team. R, 2021), the function digamma(z) calculates $\psi(z)$. The quantities and expressions of the variance parameter $\theta = \sigma^2$ for the conjugate prior are calculated in and quoted from (Zhang, 2017). The calculations of the quantities and expressions of the scale parameter $\theta = \sigma$ for the conjugate prior can be found in the **Supplementary Material**. We remark that the calculations of the quantities and expressions in Table 1 are not trivial, especially $E^{\pi_c}(\log \theta|\mathbf{x})$.

2.2 Noninformative Priors

Famous noninformative priors include the Jeffreys's (Jeffreys, 1961), reference (Bernardo, 1979; Berger and Bernardo, 1992), and matching (Tibshirani, 1989; Datta and Mukerjee, 2004) priors. See also (Berger, 2006; Berger et al., 2015) and the references therein.

The Jeffreys's noninformative prior for $\theta = \sigma^2$ is

$$\pi_J(\theta) \propto \frac{1}{\theta} \text{ or } \pi_J(\sigma^2) \propto \frac{1}{\sigma^2}$$

See Part I (p.66) of (Chen, 2014), where μ is assumed known in the normal model $N(\mu, \theta)$. The Jeffreys's noninformative prior for $\theta = \sigma$ is

$$\pi_J(\sigma) \propto \frac{1}{\sigma}$$

See Example 3.5.6 (p.131) of (Robert, 2007), where μ is assumed known in the normal model $N(\mu, \sigma^2)$.

Since μ is assumed known in the normal models, there is only one unknown parameter. Therefore, the reference prior is equal to the Jeffreys's prior, and the matching prior is also equal to the Jeffreys's prior (see pp.130–131 of (Ghosh et al., 2006)). In summary, when μ is assumed known in the normal models, the three noninformative priors equal, that is,

$$\pi_n(\theta) = \pi_J(\theta) = \pi_R(\theta) = \pi_M(\theta) \propto \frac{1}{\theta}$$

and

$$\pi_n(\sigma) = \pi_J(\sigma) = \pi_R(\sigma) = \pi_M(\sigma) \propto \frac{1}{\sigma}$$

where $\pi_n(\cdot)$ stands for the noninformative prior.

Note that as in many statistics textbooks, the probability density function (pdf) of $\theta \sim IG(\alpha, \beta)$ is given by

$$f_\theta(\theta|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \left(\frac{1}{\theta}\right)^{\alpha+1} \exp\left(-\frac{1}{\beta\theta}\right), \theta > 0, \alpha > 0, \beta > 0.$$

The conjugate prior of the scale parameter $\theta = \sigma$ is a Square Root of the Inverse Gamma (SRIG) distribution that we define below.

DEFINITION 1. Let $\theta = \sigma^2 \sim IG(\alpha, \beta)$ with $\alpha > 0$ and $\beta > 0$. Then $\sigma = \sqrt{\theta} \sim SRIG(\alpha, \beta)$ and the pdf of σ is given by

$$f_{\sigma}(\sigma|\alpha, \beta) = \frac{2}{\Gamma(\alpha)\beta^{\alpha}} \left(\frac{1}{\sigma}\right)^{2\alpha+1} \exp\left(-\frac{1}{\beta\sigma^2}\right), \quad \sigma > 0, \alpha > 0, \beta > 0.$$

Definition 1 gives the definition of the SRIG distribution, which is the conjugate prior of the scale parameter $\theta = \sigma$ of the normal distribution. Because the SRIG distribution can not be found in standard textbooks, so we give its definition here. Moreover, Definition 1 is reasonable, since

$$\begin{aligned} f_{\sigma}(\sigma|\alpha, \beta) &= f_{\theta}(\theta|\alpha, \beta)|\theta'(\sigma)| \\ &= \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \left(\frac{1}{\sigma^2}\right)^{\alpha+1} \exp\left(-\frac{1}{\beta\sigma^2}\right) \cdot 2\sigma \\ &= \frac{2}{\Gamma(\alpha)\beta^{\alpha}} \left(\frac{1}{\sigma}\right)^{2\alpha+1} \exp\left(-\frac{1}{\beta\sigma^2}\right). \end{aligned}$$

We have the following proposition which gives the three expectations of the $SRIG(\alpha, \beta)$ distribution. The calculations needed in the proposition can be found in the **Supplementary Material**. We remark that the calculations of $E(\sigma)$ and $E(\sigma^{-1})$ are straightforward by utilizing a simple transformation of $\theta = \sigma^2$ and the integration of an $IG(\alpha, \beta)$ distribution. However, the calculations of $E(\log \sigma)$ is skillful by first a transformation of $y = 1/(\beta\sigma^2)$ and then a change of the order of integration and differentiation.

PROPOSITION 1. Let $\sigma = \sqrt{\theta} \sim SRIG(\alpha, \beta)$ with $\alpha > 0$ and $\beta > 0$. Then

$$\begin{aligned} E(\sigma) &= \frac{\Gamma\left(\alpha - \frac{1}{2}\right)}{\Gamma(\alpha)\beta^{\frac{1}{2}}}, \text{ for } \alpha > \frac{1}{2} \text{ and } \beta > 0, \\ E\left(\frac{1}{\sigma}\right) &= \frac{\Gamma\left(\alpha + \frac{1}{2}\right)\beta^{\frac{1}{2}}}{\Gamma(\alpha)}, \text{ for } \alpha > 0 \text{ and } \beta > 0, \\ E(\log \sigma) &= -\frac{1}{2}\log \beta - \frac{1}{2}\psi(\alpha), \text{ for } \alpha > 0 \text{ and } \beta > 0. \end{aligned}$$

The relationship between the two distributions $IG(\alpha, \beta)$ and $SRIG(\alpha, \beta)$ are given in the following proposition whose proof can be found in the **Supplementary Material**. We remark that the proof of the proposition is straightforward by utilizing monotone transformations $\theta = \sigma^2$ and $\sigma = \sqrt{\theta}$.

PROPOSITION 2. $\theta = \sigma^2 \sim IG(\alpha, \beta)$ if and only if $\sigma = \sqrt{\theta} \sim SRIG(\alpha, \beta)$, where $\alpha > 0$ and $\beta > 0$.

The posterior distributions of θ and σ for the noninformative priors are given in the following theorem whose proof can be found in the **Supplementary Material**.

THEOREM 1. Let $X|\theta \sim N(\mu, \theta)$ and $X|\sigma \sim N(\mu, \sigma^2)$ where μ is known and $\theta = \sigma^2$ is unknown, $\pi(\theta) \propto \frac{1}{\theta}$ and $\pi(\sigma) \propto \frac{1}{\sigma}$. Then

$$\pi(\theta|\mathbf{x}) \sim IG(\tilde{\alpha}, \tilde{\beta}) \text{ and } \pi(\sigma|\mathbf{x}) \sim SRIG(\tilde{\alpha}, \tilde{\beta}),$$

where

$$\tilde{\alpha} = \frac{n}{2} \text{ and } \tilde{\beta} = \frac{2}{\sum_{i=1}^n (x_i - \mu)^2}. \tag{8}$$

We have the following two remarks for Theorem 1.

Remark 1. Let $\theta = \sigma^2$. In the derivation of $\pi(\sigma|\mathbf{x})$, if we derive it in this way,

$$\begin{aligned} f_{\sigma}(\sigma) &= \pi(\sigma|\mathbf{x}) \\ &\propto \left(\frac{1}{\sigma}\right)^{n+1} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right) \\ &= \left(\frac{1}{\sigma^2}\right)^{\frac{n+1}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right) \\ &= \left(\frac{1}{\theta}\right)^{\frac{n+1}{2}} \exp\left(-\frac{1}{2\theta} \sum_{i=1}^n (x_i - \mu)^2\right) = f_{\theta}(\theta) \\ &\sim IG(\tilde{\alpha}_1, \tilde{\beta}), \end{aligned}$$

where

$$\tilde{\alpha}_1 = \frac{n-1}{2} \text{ and } \tilde{\beta} = \frac{2}{\sum_{i=1}^n (x_i - \mu)^2},$$

then by Proposition 2, $f_{\sigma}(\sigma) = \pi(\sigma|\mathbf{x}) \sim SRIG(\tilde{\alpha}_1, \tilde{\beta})$, which is different from $SRIG(\tilde{\alpha}, \tilde{\beta})$. In fact, the above practice is equivalent to the derivation of the pdf of θ in terms of the pdf of σ by $f_{\theta}(\theta) = f_{\sigma}(\sigma)$, ignoring the $|\sigma'(\theta)|$ term, which is obviously wrong. Therefore, the above derivation which is a pitfall for incautious users is wrong. ||

Remark 2. The two posterior distributions in Theorem 1, $\pi(\theta|\mathbf{x}) \sim IG(\tilde{\alpha}, \tilde{\beta})$ and $\pi(\sigma|\mathbf{x}) \sim SRIG(\tilde{\alpha}, \tilde{\beta})$, follow Proposition 2 by accident. We have

$$f_{\theta}(\theta) = \pi(\theta|\mathbf{x}) \propto f(\mathbf{x}|\theta)\pi(\theta) \propto f(\mathbf{x}|\theta) \frac{1}{\theta} \sim IG(\tilde{\alpha}, \tilde{\beta})$$

and

$$f_{\sigma}(\sigma) = \pi(\sigma|\mathbf{x}) \propto f(\mathbf{x}|\sigma)\pi(\sigma) \propto f(\mathbf{x}|\theta) \frac{1}{\sigma} \sim SRIG(\tilde{\alpha}, \tilde{\beta}).$$

Note that $\sigma = \sqrt{\theta}$, and thus

$$\begin{aligned} f_{\sigma}(\sigma)|\sigma'(\theta)| &\propto f(\mathbf{x}|\theta) \frac{1}{\sigma} \frac{1}{2\sqrt{\theta}} = f(\mathbf{x}|\theta) \frac{1}{\sqrt{\theta}} \frac{1}{2\sqrt{\theta}} \\ &= f(\mathbf{x}|\theta) \frac{1}{2\theta} \propto f_{\theta}(\theta), \end{aligned} \tag{9}$$

which is the reason why $\pi(\theta|\mathbf{x}) = f_{\theta}(\theta)$ and $\pi(\sigma|\mathbf{x}) = f_{\sigma}(\sigma)$ follow Proposition 2. Note that the posterior distributions depend on the prior distributions. If the prior distributions $\pi(\theta)$ and $\pi(\sigma)$ are selected different from $\frac{1}{\theta}$ and $\frac{1}{\sigma}$, then the relationship (Eq. 9) may not be satisfied, and thus $\pi(\theta|\mathbf{x})$ and $\pi(\sigma|\mathbf{x})$ may not follow Proposition 2. ||

2.2.1 The Quantities and Expressions of the Variance Parameter

In this subsection, we will calculate the expressions of the quantities (three posterior expectations, two Bayes estimators, and two PESLs) of the variance parameter $\theta = \sigma^2$.

Now we calculate the three expectations $E(\theta|\mathbf{x})$, $E(\theta^{-1}|\mathbf{x})$, and $E(\log \theta|\mathbf{x})$ for the variance parameter $\theta = \sigma^2$. By Theorem 1, $\pi(\theta|\mathbf{x}) \sim IG(\tilde{\alpha}, \tilde{\beta})$, and thus

$$E(\theta|\mathbf{x}) = \frac{1}{(\tilde{\alpha} - 1)\tilde{\beta}}, \quad \tilde{\alpha} > 1 \text{ and } E\left(\frac{1}{\theta}|\mathbf{x}\right) = \tilde{\alpha}\tilde{\beta}.$$

From (Zhang, 2017), we know that

$$E(\log \theta|\mathbf{x}) = -\log \tilde{\beta} - \psi(\tilde{\alpha}).$$

It is easy to see that, for $\tilde{\alpha} > 1$,

$$\delta_s^{\pi,\theta}(\mathbf{x}) = \frac{1}{E\left(\frac{1}{\theta}|\mathbf{x}\right)} = \frac{1}{\tilde{\alpha}\tilde{\beta}} < \frac{1}{(\tilde{\alpha} - 1)\tilde{\beta}} = E(\theta|\mathbf{x}) = \delta_2^{\pi,\theta}(\mathbf{x}),$$

which exemplifies (Eq. 6). From (Zhang, 2017), we find that

$$PESL_s^{\pi,\theta}(\mathbf{x}) = \log \tilde{\alpha} - \psi(\tilde{\alpha}), \text{ for } \tilde{\alpha} > 0,$$

and

$$PESL_2^{\pi,\theta}(\mathbf{x}) = \frac{1}{\tilde{\alpha} - 1} + \log(\tilde{\alpha} - 1) - \psi(\tilde{\alpha}), \text{ for } \tilde{\alpha} > 1.$$

It can be directly proved that $PESL_s^{\pi,\theta}(\mathbf{x}) \leq PESL_2^{\pi,\theta}(\mathbf{x})$ for $\tilde{\alpha} > 1$, which exemplifies (Eq. 7), and its proof which exploits the Taylor series expansion for e^x can be found in the **Supplementary Material**. Note that $PESL_s^{\pi,\theta}(\mathbf{x})$ and $PESL_2^{\pi,\theta}(\mathbf{x})$ depend only on $\tilde{\alpha} = n/2$. Therefore, they depend only on n , but do not depend on μ and x . Numerical simulations will exemplify this result.

The IRSL at $\delta_s^{\pi,\theta}$ or the BRSL for $\theta = \sigma^2$ is (similar to (Robert, 2007))

$$\begin{aligned} IRSL_s^{\pi,\theta} &= BRSL^{\pi,\theta} = r(\pi, \delta_s^{\pi,\theta}) \\ &= E^\pi [R(\theta, \delta_s^{\pi,\theta})] \\ &= \int_{\Theta} R(\theta, \delta_s^{\pi,\theta}) \pi(\theta) d\theta \\ &= \int_{\Theta} \int_{\mathcal{X}} L(\theta, \delta_s^{\pi,\theta}(\mathbf{x})) f(\mathbf{x}|\theta) d\mathbf{x} \pi(\theta) d\theta \\ &= \int_{\mathcal{X}} \int_{\Theta} L(\theta, \delta_s^{\pi,\theta}(\mathbf{x})) f(\mathbf{x}|\theta) \pi(\theta) d\theta d\mathbf{x} \\ &= \int_{\mathcal{X}} \int_{\Theta} L(\theta, \delta_s^{\pi,\theta}(\mathbf{x})) \pi(\theta|\mathbf{x}) d\theta m^{\pi,\theta}(\mathbf{x}) d\mathbf{x} \\ &= \int_{\mathcal{X}} PESL^{\pi,\theta}(a(\mathbf{x})|\mathbf{x})|_{a=\delta_s^{\pi,\theta}} m^{\pi,\theta}(\mathbf{x}) d\mathbf{x} \\ &= \int_{\mathcal{X}} PESL_s^{\pi,\theta}(\mathbf{x}) m^{\pi,\theta}(\mathbf{x}) d\mathbf{x} \\ &= \int_{\mathcal{X}} [\log \tilde{\alpha} - \psi(\tilde{\alpha})] m^{\pi,\theta}(\mathbf{x}) d\mathbf{x} \\ &= \log \tilde{\alpha} - \psi(\tilde{\alpha}) \\ &= PESL_s^{\pi,\theta}(\mathbf{x}), \end{aligned}$$

since $\tilde{\alpha}$ does not depend on x , where

$$m^{\pi,\theta}(\mathbf{x}) = \int_0^\infty f(\mathbf{x}|\theta) \pi(\theta) d\theta$$

is the marginal density of \mathbf{x} with prior $\pi(\theta)$.

2.2.2 The Quantities and Expressions of the Scale Parameter

In this subsection, we will calculate the expressions of the quantities (three posterior expectations, two Bayes estimators, and two PESLs) of the scale parameter $\theta = \sigma$.

Now let us calculate $\delta_s^{\pi,\sigma}(\mathbf{x})$, $\delta_2^{\pi,\sigma}(\mathbf{x})$, $PESL_s^{\pi,\sigma}(\mathbf{x})$, and $PESL_2^{\pi,\sigma}(\mathbf{x})$ for the scale parameter σ . To calculate these quantities, we need to calculate the three expectations $E(\sigma|\mathbf{x})$, $E(\sigma^{-1}|\mathbf{x})$, and $E(\log \sigma|\mathbf{x})$. Since $\pi(\sigma|\mathbf{x}) \sim SRIG(\tilde{\alpha}, \tilde{\beta})$ by Theorem 1, from Proposition 1, we have

$$E(\sigma|\mathbf{x}) = \frac{\Gamma(\tilde{\alpha} - \frac{1}{2})}{\Gamma(\tilde{\alpha})\tilde{\beta}^{\frac{1}{2}}}, \text{ for } \tilde{\alpha} > \frac{1}{2} \text{ and } \tilde{\beta} > 0, \tag{10}$$

$$E\left(\frac{1}{\sigma}|\mathbf{x}\right) = \frac{\Gamma(\tilde{\alpha} + \frac{1}{2})\tilde{\beta}^{\frac{1}{2}}}{\Gamma(\tilde{\alpha})}, \text{ for } \tilde{\alpha} > 0 \text{ and } \tilde{\beta} > 0, \tag{11}$$

$$E(\log \sigma|\mathbf{x}) = -\frac{1}{2} \log \tilde{\beta} - \frac{1}{2} \psi(\tilde{\alpha}), \text{ for } \tilde{\alpha} > 0 \text{ and } \tilde{\beta} > 0. \tag{12}$$

It can be proved that, for $\tilde{\alpha} > \frac{1}{2}$,

$$\delta_s^{\pi,\sigma}(\mathbf{x}) = \frac{1}{E\left(\frac{1}{\sigma}|\mathbf{x}\right)} = \frac{\Gamma(\tilde{\alpha})}{\Gamma(\tilde{\alpha} + \frac{1}{2})\tilde{\beta}^{\frac{1}{2}}} < \frac{\Gamma(\tilde{\alpha} - \frac{1}{2})}{\Gamma(\tilde{\alpha})\tilde{\beta}^{\frac{1}{2}}} = E(\sigma|\mathbf{x}) = \delta_2^{\pi,\sigma}(\mathbf{x}),$$

which exemplifies (Eq. 6), and the proof which exploits the positivity of $\psi'(x)$ can be found in the **Supplementary Material**.

Now we calculate $PESL_s^{\pi,\sigma}(\mathbf{x})$ and $PESL_2^{\pi,\sigma}(\mathbf{x})$ for the scale parameter σ . From (Zhang, 2017), we know that the PESL at $\delta_s^{\pi,\sigma}(\mathbf{x}) = [E(\sigma^{-1}|\mathbf{x})]^{-1}$ is

$$PESL_s^{\pi,\sigma}(\mathbf{x}) = E[L_s(\theta, a)|\mathbf{x}]|_{a=\frac{1}{E\left(\frac{1}{\sigma}|\mathbf{x}\right)}} = \log E\left(\frac{1}{\sigma}|\mathbf{x}\right) + E(\log \sigma|\mathbf{x}),$$

and the PESL at $\delta_2^{\pi,\sigma}(\mathbf{x}) = E(\sigma|\mathbf{x})$ is

$$\begin{aligned} PESL_2^{\pi,\sigma}(\mathbf{x}) &= E[L_s(\theta, a)|\mathbf{x}]|_{a=E(\sigma|\mathbf{x})} \\ &= E(\sigma|\mathbf{x})E\left(\frac{1}{\sigma}|\mathbf{x}\right) - 1 - \log E(\sigma|\mathbf{x}) + E(\log \sigma|\mathbf{x}). \end{aligned}$$

Substituting (Eqs 10, 11, 12), into the above expressions, we obtain

$$\begin{aligned} PESL_s^{\pi,\sigma}(\mathbf{x}) &= \log \frac{\Gamma(\tilde{\alpha} + \frac{1}{2})\tilde{\beta}^{\frac{1}{2}}}{\Gamma(\tilde{\alpha})} - \frac{1}{2} \log \tilde{\beta} - \frac{1}{2} \psi(\tilde{\alpha}) \\ &= \log \Gamma\left(\tilde{\alpha} + \frac{1}{2}\right) - \log \Gamma(\tilde{\alpha}) - \frac{1}{2} \psi(\tilde{\alpha}), \end{aligned}$$

for $\tilde{\alpha} > 0$ and $\tilde{\beta} > 0$, and

$$\begin{aligned} PESL_2^{\pi,\sigma}(\mathbf{x}) &= \frac{\Gamma(\tilde{\alpha} - \frac{1}{2})}{\Gamma(\tilde{\alpha})\tilde{\beta}^{\frac{1}{2}}} \frac{\Gamma(\tilde{\alpha} + \frac{1}{2})\tilde{\beta}^{\frac{1}{2}}}{\Gamma(\tilde{\alpha})} - 1 - \log \frac{\Gamma(\tilde{\alpha} - \frac{1}{2})}{\Gamma(\tilde{\alpha})\tilde{\beta}^{\frac{1}{2}}} - \frac{1}{2} \log \tilde{\beta} - \frac{1}{2} \psi(\tilde{\alpha}) \\ &= \frac{\Gamma(\tilde{\alpha} - \frac{1}{2})\Gamma(\tilde{\alpha} + \frac{1}{2})}{\Gamma^2(\tilde{\alpha})} - 1 - \log \Gamma\left(\tilde{\alpha} - \frac{1}{2}\right) + \log \Gamma(\tilde{\alpha}) - \frac{1}{2} \psi(\tilde{\alpha}), \end{aligned}$$

for $\tilde{\alpha} > \frac{1}{2}$ and $\tilde{\beta} > 0$. It can be directly proved that $PESL_s^{\pi,\sigma}(\mathbf{x}) \leq PESL_2^{\pi,\sigma}(\mathbf{x})$ for $\tilde{\alpha} > \frac{1}{2}$ and $\tilde{\beta} > 0$, which exemplifies (Eq. 7), and its proof which exploits the Taylor series expansion for $\log u$ with u near 1 can be found in the **Supplementary Material**. Note that $PESL_s^{\pi,\sigma}(\mathbf{x})$ and $PESL_2^{\pi,\sigma}(\mathbf{x})$ depend only on $\tilde{\alpha} = n/2$. Therefore, they depend only on n , but do not depend on μ and x . Numerical simulations will exemplify this result.

TABLE 2 | The quantities and expressions for the noninformative priors.

Quantities	Expressions for $\theta = \sigma^2$, the variance parameter	Expressions for $\theta = \sigma$, the scale parameter
$\pi_n(\theta)$	$\frac{1}{\theta}$	$\frac{1}{\sigma}$
$\pi_n(\theta \mathbf{x})$	$IG(\tilde{\alpha}, \tilde{\beta})$	$SRIG(\tilde{\alpha}, \tilde{\beta})$
$E^{\pi_n}(\theta \mathbf{x})$	$\frac{1}{(\tilde{\alpha}-1)\tilde{\beta}}$, for $\tilde{\alpha} > 1$	$\frac{\Gamma(\tilde{\alpha}-\frac{1}{2})}{\Gamma(\tilde{\alpha})\tilde{\beta}^2}$, for $\tilde{\alpha} > \frac{1}{2}$
$E^{\pi_n}(\frac{1}{\theta} \mathbf{x})$	$\tilde{\alpha}\tilde{\beta}$	$\frac{\Gamma(\tilde{\alpha}+\frac{1}{2})\tilde{\beta}^2}{\Gamma(\tilde{\alpha})}$
$E^{\pi_n}(\log \theta \mathbf{x})$	$-\log \tilde{\beta} - \psi(\tilde{\alpha})$	$-\frac{1}{2}\log \tilde{\beta} - \frac{1}{2}\psi(\tilde{\alpha})$
$\delta_s^{\pi_n, \theta}(\mathbf{x})$	$\frac{1}{\tilde{\alpha}\tilde{\beta}}$	$\frac{\Gamma(\tilde{\alpha})}{\Gamma(\tilde{\alpha}+\frac{1}{2})\tilde{\beta}^2}$
$\delta_2^{\pi_n, \theta}(\mathbf{x})$	$\frac{1}{(\tilde{\alpha}-1)\tilde{\beta}}$, for $\tilde{\alpha} > 1$	$\frac{\Gamma(\tilde{\alpha}-\frac{1}{2})}{\Gamma(\tilde{\alpha})\tilde{\beta}^2}$, for $\tilde{\alpha} > \frac{1}{2}$
$PESL_{-s}^{\pi_n, \theta}(\mathbf{x})$	$\log \tilde{\alpha} - \psi(\tilde{\alpha})$	$\log \Gamma(\tilde{\alpha} + \frac{1}{2}) - \log \Gamma(\tilde{\alpha}) - \frac{1}{2}\psi(\tilde{\alpha})$
$PESL_{-2}^{\pi_n, \theta}(\mathbf{x})$	$\frac{1}{\tilde{\alpha}-1} + \log(\tilde{\alpha}-1) - \psi(\tilde{\alpha})$, for $\tilde{\alpha} > 1$	$\frac{\Gamma(\tilde{\alpha}-\frac{1}{2})\Gamma(\tilde{\alpha}+\frac{1}{2})}{\Gamma^2(\tilde{\alpha})} - 1 - \log \Gamma(\tilde{\alpha}-\frac{1}{2})$ $+ \log \Gamma(\tilde{\alpha}) - \frac{1}{2}\psi(\tilde{\alpha})$, for $\tilde{\alpha} > \frac{1}{2}$

The IRSL at $\delta_s^{\pi_n, \sigma}$ or the BRSL for $\theta = \sigma$ is (similar to (Robert, 2007))

$$\begin{aligned}
 IRSL_s^{\pi_n, \sigma} &= BRSL^{\pi_n, \sigma} = r(\pi, \delta_s^{\pi_n, \sigma}) \\
 &= E^{\pi} [R(\sigma, \delta_s^{\pi_n, \sigma})] \\
 &= \int_{\Sigma} R(\sigma, \delta_s^{\pi_n, \sigma}) \pi(\sigma) d\sigma \\
 &= \int_{\Sigma} \int_{\mathcal{X}} L(\sigma, \delta_s^{\pi_n, \sigma}(\mathbf{x})) f(\mathbf{x}|\sigma) d\mathbf{x} \pi(\sigma) d\sigma \\
 &= \int_{\mathcal{X}} \int_{\Sigma} L(\sigma, \delta_s^{\pi_n, \sigma}(\mathbf{x})) f(\mathbf{x}|\sigma) \pi(\sigma) d\sigma d\mathbf{x} \\
 &= \int_{\mathcal{X}} \int_{\Sigma} L(\sigma, \delta_s^{\pi_n, \sigma}(\mathbf{x})) \pi(\sigma|\mathbf{x}) d\sigma m^{\pi_n, \sigma}(\mathbf{x}) d\mathbf{x} \\
 &= \int_{\mathcal{X}} PESL^{\pi_n, \sigma}(a(\mathbf{x})|\mathbf{x})|_{a=\delta_s^{\pi_n, \sigma}} m^{\pi_n, \sigma}(\mathbf{x}) d\mathbf{x} \\
 &= \int_{\mathcal{X}} PESL_s^{\pi_n, \sigma}(\mathbf{x}) m^{\pi_n, \sigma}(\mathbf{x}) d\mathbf{x} \\
 &= \int_{\mathcal{X}} \left[\log \Gamma\left(\tilde{\alpha} + \frac{1}{2}\right) - \log \Gamma(\tilde{\alpha}) - \frac{1}{2}\psi(\tilde{\alpha}) \right] m^{\pi_n, \sigma}(\mathbf{x}) d\mathbf{x} \\
 &= \log \Gamma\left(\tilde{\alpha} + \frac{1}{2}\right) - \log \Gamma(\tilde{\alpha}) - \frac{1}{2}\psi(\tilde{\alpha}) \\
 &= PESL_s^{\pi_n, \sigma}(\mathbf{x}),
 \end{aligned}$$

since $\tilde{\alpha}$ does not depend on x , where

$$m^{\pi_n, \sigma}(\mathbf{x}) = \int_0^{\infty} f(\mathbf{x}|\sigma)\pi(\sigma)d\sigma$$

is the marginal density of \mathbf{x} with prior $\pi(\sigma)$.

The quantities and expressions of the variance and scale parameters for the noninformative priors are summarized in Table 2. In the table, $\tilde{\alpha}$ and $\tilde{\beta}$ are given by (Eq. 8).

From Tables 1, 2, we find that there are four combinations of the expressions of the quantities: conjugate prior and variance

parameter, conjugate prior and scale parameter, noninformative prior and variance parameter, and noninformative prior and scale parameter. The forms of the expressions of the quantities are the same for the variance parameter under the conjugate and noninformative priors, since they have the same Inverse Gamma posterior distributions. Similarly, the forms of the expressions of the quantities are the same for the scale parameter under the conjugate and noninformative priors, since they have the same Square Root of the Inverse Gamma posterior distributions.

The inequalities (Eqs 6, 7) exist in Tables 1, 2. In fact, there are 8 inequalities in Tables 1, 2 and 4 inequalities in each table. Since the forms of the expressions of the quantities are the same in Tables 1, 2, with the only difference of the parameters, there are actually 4 different inequalities which are in Table 2. One inequality of the four inequalities about the Bayes estimators is obvious, and the proofs of the other three inequalities can be found in the Supplementary Material.

3 NUMERICAL SIMULATIONS

In this section, we will numerically exemplify the theoretical studies of (Eqs 6, 7), and that the PESLs depend only on n , but do not depend on μ and x . The numerical simulation results are similar for the four combinations of the expressions of the quantities, and thus we only present the results for the combination of the noninformative prior and the scale parameter.

First, we fix $\mu = 0$ and $n = 10$, and assume that $\sigma = 1$ is drawn from the improper prior distribution. After that, we draw a random sample

$$\mathbf{x} = \text{rnorm}(n = n, \text{mean} = \mu, \text{sd} = \sigma)$$

from $N(\mu, \sigma^2)$.

To generate a random sample $\sigma = (\sigma_1, \dots, \sigma_k)$ with $k = 1000$ from

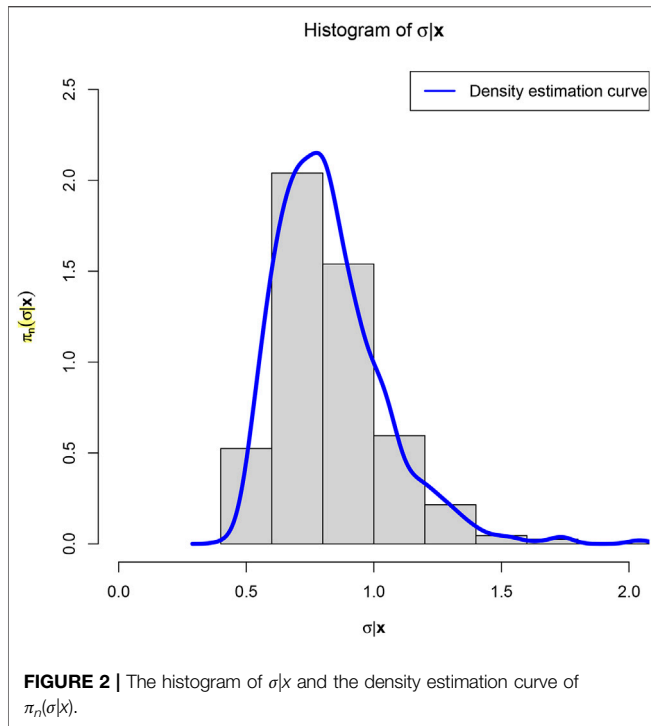


FIGURE 2 | The histogram of $\sigma|x$ and the density estimation curve of $\pi_n(\sigma|x)$.

$$\pi_n(\sigma|x) = SRIG(\tilde{\alpha}, \tilde{\beta}),$$

we will adopt the following algorithm. First, compute $\tilde{\alpha}$ and $\tilde{\beta}$ from (Eq. 8). Second, generate a random sample

$$G = \text{rgamma}\left(n = k, \text{shape} = \tilde{\alpha}, \text{scale} = \tilde{\beta}\right) \sim G(\tilde{\alpha}, \tilde{\beta}).$$

Third, compute

$$IG = \frac{1}{G} \sim IG(\tilde{\alpha}, \tilde{\beta}).$$

Fourth, compute

$$\sigma = \sqrt{IG} \sim SRIG(\tilde{\alpha}, \tilde{\beta}).$$

Hence, σ is a random sample from the $SRIG(\tilde{\alpha}, \tilde{\beta})$ distribution. Figure 2 shows the histogram of $\sigma|x$ and the density estimation curve of $\pi_n(\sigma|x)$. It is $\pi_n(\sigma|x)$ that we find $\delta_s^{\pi_n, \sigma}(\mathbf{x})$ to minimize the PESL. From the figure, we see that the $SRIG(\tilde{\alpha}, \tilde{\beta})$ distribution is left peaked, right skewed, and continuous.

The Bayes estimators ($\delta_s^{\pi_n, \sigma}(\mathbf{x})$ and $\delta_2^{\pi_n, \sigma}(\mathbf{x})$) and the PESLs ($PESL_s^{\pi_n, \sigma}(\mathbf{x})$ and $PESL_2^{\pi_n, \sigma}(\mathbf{x})$) are computed by the following algorithm. First, compute $\tilde{\alpha}$ and $\tilde{\beta}$ from (Eq. 8). Second, compute

$$E_1 = E(\sigma|x) = \frac{\Gamma(\tilde{\alpha} - \frac{1}{2})}{\Gamma(\tilde{\alpha})\tilde{\beta}^{\frac{1}{2}}},$$

$$E_2 = E\left(\frac{1}{\sigma}|x\right) = \frac{\Gamma(\tilde{\alpha} + \frac{1}{2})\tilde{\beta}^{\frac{1}{2}}}{\Gamma(\tilde{\alpha})},$$

$$E_3 = E(\log \sigma|x) = -\frac{1}{2} \log \tilde{\beta} - \frac{1}{2} \psi(\tilde{\alpha}).$$

Third, compute

$$\delta_s^{\pi_n, \sigma}(\mathbf{x}) = \frac{1}{E_2},$$

$$\delta_2^{\pi_n, \sigma}(\mathbf{x}) = E_1,$$

$$PESL_s^{\pi_n, \sigma}(\mathbf{x}) = \log(E_2) + E_3,$$

$$PESL_2^{\pi_n, \sigma}(\mathbf{x}) = E_1 \times E_2 - \log(E_1) + E_3 - 1.$$

Numerical results show that

$$\delta_s^{\pi_n, \sigma}(\mathbf{x}) = 0.7712483 < 0.8152161 = \delta_2^{\pi_n, \sigma}(\mathbf{x})$$

and

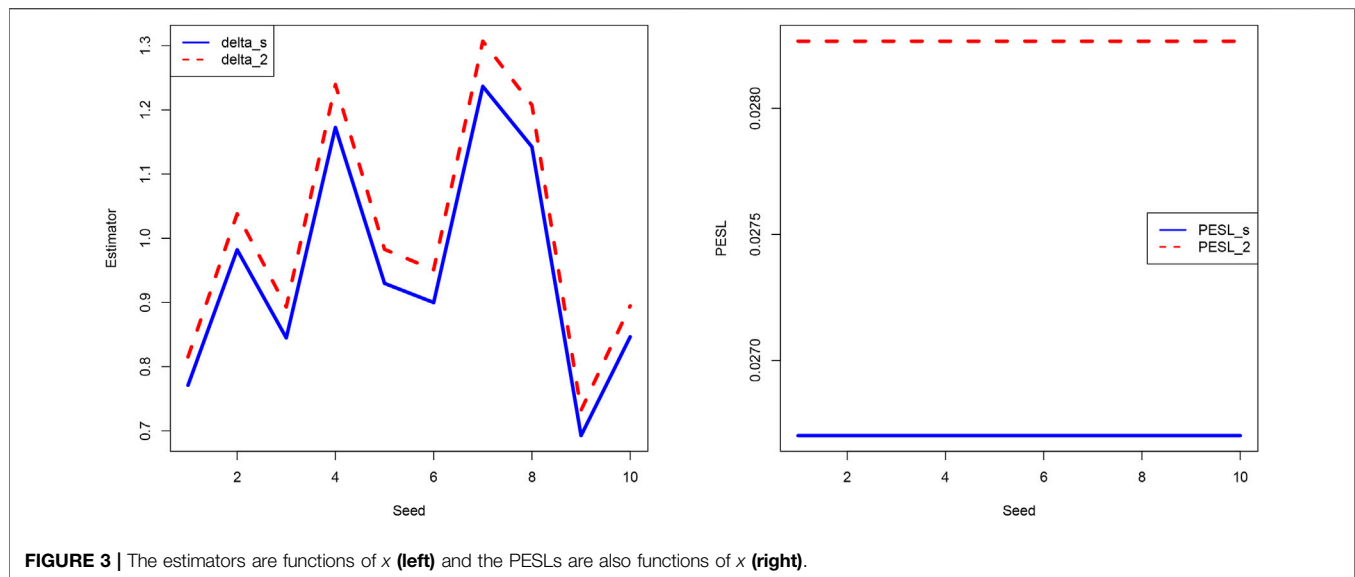
$$PESL_s^{\pi_n, \sigma}(\mathbf{x}) = 0.0267013 < 0.02826706 = PESL_2^{\pi_n, \sigma}(\mathbf{x}),$$

which exemplify the theoretical studies of (6) and (7).

In Figure 3, we fix $\mu = 0$ and $n = 10$, but allow the seed number to change from 1 to 10 (i.e., we change x). From the figure we see that the estimators and PESLs are functions of x . We see from the left plot of the figure that the estimators depend on x in an unpredictable manner, and $\delta_s^{\pi_n, \sigma}(\mathbf{x})$ are unanimously smaller than $\delta_2^{\pi_n, \sigma}(\mathbf{x})$, and thus (Eq. 6) is exemplified. The two Bayes estimators are distinguishable since we fix $n = 10$ to be a small number. The right plot of the figure exhibits that the PESLs do not depend on x , and $PESL_s^{\pi_n, \sigma}(\mathbf{x})$ are unanimously smaller than $PESL_2^{\pi_n, \sigma}(\mathbf{x})$, and thus (Eq. 7) is exemplified.

Now we allow one of the two parameters μ and n to change, holding other parameters fixed. Moreover, we also assume that the sample x is fixed, as it is the case for the real data. Figure 4 shows the estimators and PESLs as functions of μ and n . We see from the left plots of the figure that the estimators depend on μ and n , and (Eq. 6) is exemplified. More specifically, the estimators are first decreasing and then increasing functions of μ , and the estimators attain the minimum when $\mu = 0$. However, the estimators fluctuate around some value when n increases. The right plots of the figure exhibit that the PESLs depend only on n , but do not depend on μ , and (Eq. 7) is exemplified. More specifically, the PESLs are decreasing functions of n . Furthermore, the two PESLs as functions of n are indistinguishable, as the two PESLs are very close. In summary, the results of the figure exemplify the theoretical studies of (Eqs 6, 7).

Since the estimators $\delta_s^{\pi_n, \sigma}(\mathbf{x})$ and $\delta_2^{\pi_n, \sigma}(\mathbf{x})$ and the PESLs $PESL_s^{\pi_n, \sigma}(\mathbf{x})$ and $PESL_2^{\pi_n, \sigma}(\mathbf{x})$ depend on $\tilde{\alpha}$ and $\tilde{\beta}$, where $\tilde{\alpha} > 1/2$ and $\tilde{\beta} > 0$, we can plot the surfaces of the estimators and the PESLs on the domain $(\tilde{\alpha}, \tilde{\beta}) \in (0.5, 10] \times (0, 10] = D$ via the R function `persp3d()` in the R package `rgl` (see (Adler and Murdoch, 2017; Zhang et al., 2017; Zhang et al., 2019; Sun et al., 2021)). We remark that the R function `persp()` in the R package `graphics` can not add another surface to the existing surface, but `persp3d()` can. Moreover, `persp3d()` allows one to rotate the perspective plots of the surface according to one's wishes. Figure 5 plots the surfaces of the estimators and the PESLs, and the surfaces of the difference of the estimators and the difference of the PESLs. From the left two plots of the figure, we see that $\delta_s^{\pi_n, \sigma}(\mathbf{x}) < \delta_2^{\pi_n, \sigma}(\mathbf{x})$ for all $(\tilde{\alpha}, \tilde{\beta})$ on D , which exemplifies (Eq. 6). From the right two plots of the figure, we see that $PESL_s^{\pi_n, \sigma}(\mathbf{x}) < PESL_2^{\pi_n, \sigma}(\mathbf{x})$ for all $(\tilde{\alpha}, \tilde{\beta})$ on D , which



exemplifies (Eq. 7). In summary, the results of the figure exemplify the theoretical studies of (Eqs 6, 7).

4 A REAL DATA EXAMPLE

In this section, we exploit the data from finance. The R package `quantmod` (Ryan and Ulrich, 2017) is exploited to download the data `GSPC` (the S&P 500) during 2020-04-24 and 2021-07-02 from “finance.yahoo.com.” It is commonly believed that the monthly simple returns of the index data or the stock data are normally distributed. It is simple to check that the S&P 500 monthly simple returns follow the normal model. Usually, the data from real examples can be regarded as iid from the normal model with an unknown mean μ . However, the mean μ could be estimated by prior information or historical information. Alternatively, the mean μ could be estimated by the sample mean. Therefore, for simplicity, we assume that the mean μ is known. Assume that

$$\mu = \bar{x}, \alpha = 1, \beta = 1$$

for the S&P 500 monthly simple returns.

The Bayes estimators and the PESLs of the variance and scale parameters of the S&P 500 monthly simple returns for the conjugate and noninformative priors are summarized in Table 3. From the table, we observe the following facts.

- The two inequalities (Eqs 6, 7) are exemplified.
- Given the prior (conjugate or noninformative), the Bayes estimators are similar across different loss functions (Stein’s or squared error).
- Given the loss function, the Bayes estimators are quite different across different priors. Therefore, the prior has a larger influence than the loss function in calculating the Bayes estimators.

More results (the data of the S&P 500 monthly simple returns, the plot of the S&P 500 monthly close prices, the plot of the S&P 500 monthly simple returns, the histogram of the S&P 500 monthly simple returns) for the real data example can be found in the **Supplementary Material** due to space limitations.

5 CONCLUSIONS AND DISCUSSIONS

For the variance ($\theta = \sigma^2$) and scale ($\theta = \sigma$) parameters of the normal model with a known mean μ , we recommend and analytically calculate the Bayes estimators, $\delta_s^{\pi, \theta}(\mathbf{x})$, with respect to the conjugate and noninformative (Jeffreys’s, reference, and matching) priors under Stein’s loss function which penalizes gross overestimation and gross underestimation equally. These estimators minimize the PESLs. We also analytically calculate the Bayes estimators, $\delta_2^{\pi, \theta}(\mathbf{x}) = E(\theta|\mathbf{x})$, with respect to the conjugate and noninformative priors under the squared error loss function, and the corresponding PESLs. The quantities ($\pi(\theta)$, $\pi(\theta|\mathbf{x})$, $E^\pi(\theta|\mathbf{x})$, $E^\pi(\theta^{-1}|\mathbf{x})$, $E^\pi(\log \theta|\mathbf{x})$, $\delta_s^{\pi, \theta}(\mathbf{x})$, $\delta_2^{\pi, \theta}(\mathbf{x})$, $PESL_s^{\pi, \theta}(\mathbf{x})$, $PESL_2^{\pi, \theta}(\mathbf{x})$) and expressions of the variance and scale parameters for the conjugate and noninformative priors are summarized in Tables 1, 2, respectively. Note that $E^\pi(\log \theta|\mathbf{x})$, which is essential for the calculation of $PESL_s^{\pi, \theta}(\mathbf{x})$ and $PESL_2^{\pi, \theta}(\mathbf{x})$, depends on the digamma function.

Proposition 1 gives the three expectations of the $SRIG(\alpha, \beta)$ distribution. Moreover, Proposition 2 gives the relationship between the two distributions $IG(\alpha, \beta)$ and $SRIG(\alpha, \beta)$.

For the conjugate and noninformative priors, the posterior distribution of $\theta = \sigma^2$, $\pi(\theta|\mathbf{x})$, follows an Inverse Gamma distribution, and the posterior distribution of σ , $\pi(\sigma|\mathbf{x})$, follows an SRIG distribution which is defined in Definition 1.

We find that the IRSL at $\delta_s^{\pi, \theta}$ or the BRS� for $\theta = \sigma^2$ is

$$PESL_s^{\pi, \theta}(\mathbf{x}) = \log \tilde{\alpha} - \psi(\tilde{\alpha}).$$

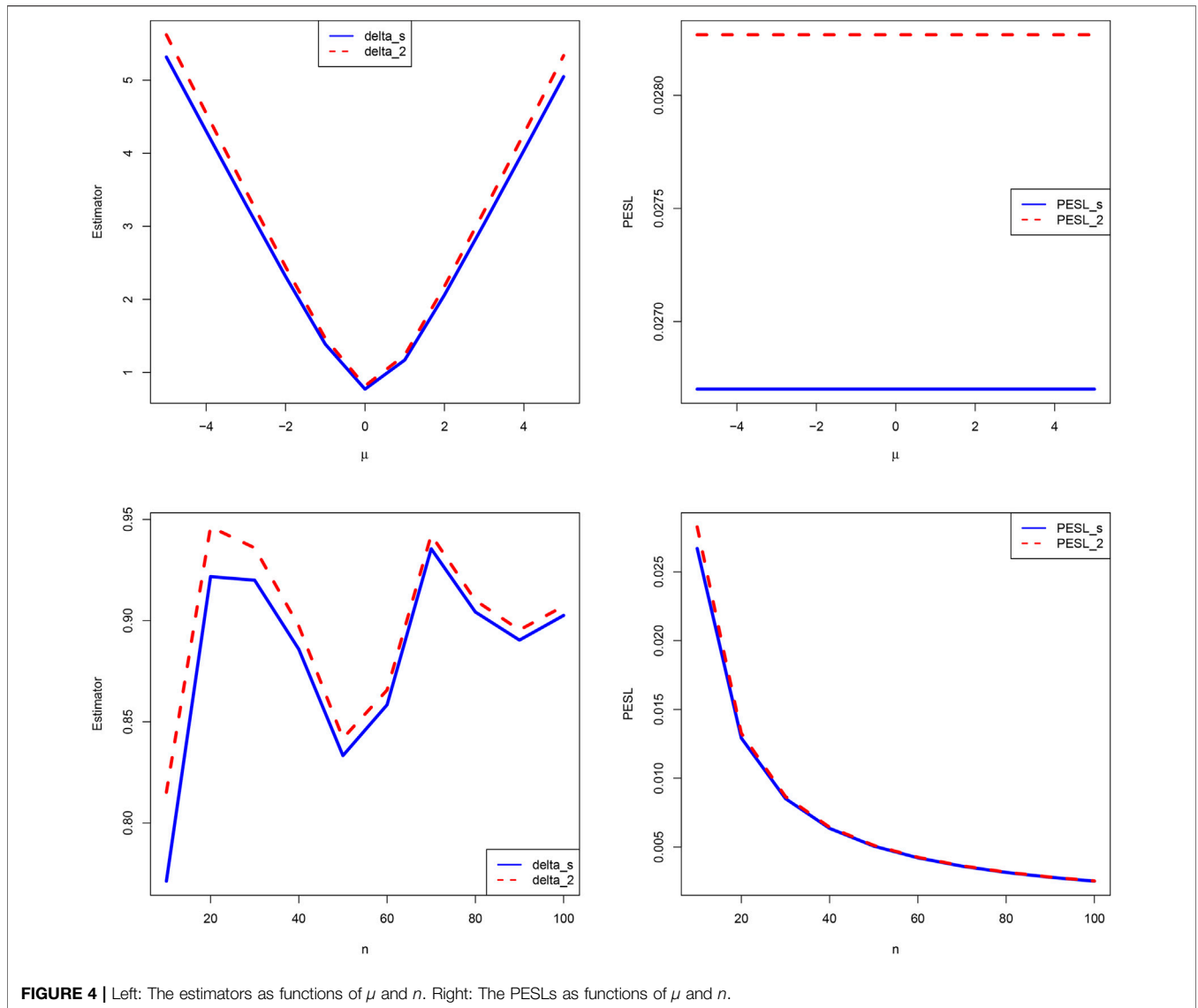


FIGURE 4 | Left: The estimators as functions of μ and n . Right: The PESLs as functions of μ and n .

In addition, the IRSL at $\delta_s^{\pi, \sigma}$ or the BRSL for $\theta = \sigma$ is

$$PESL_s^{\pi, \sigma}(\mathbf{x}) = \log \Gamma\left(\tilde{\alpha} + \frac{1}{2}\right) - \log \Gamma(\tilde{\alpha}) - \frac{1}{2}\psi(\tilde{\alpha}).$$

The numerical simulations of the combination of the noninformative prior and the scale parameter exemplify the theoretical studies of (Eqs 6, 7), and that the PESLs depend only on n , but do not depend on μ and x . Moreover, in the real data example, we have calculated the Bayes estimators and the PESLs of the variance and scale parameters of the S&P 500 monthly simple returns for the conjugate and noninformative priors.

Unlike in frequentist paradigm, if $\hat{\sigma}$ is the Maximum Likelihood Estimator (MLE) of σ , then $\hat{\sigma}^2$ is the MLE of σ^2 . In Bayesian

paradigm, we usually should estimate the variance parameter σ^2 and the scale parameter σ separately. In Table 2, we find that

$$\delta_s^{\pi, \sigma^2}(\mathbf{x}) = \frac{1}{\tilde{\alpha}\tilde{\beta}} \text{ and } \delta_s^{\pi, \sigma}(\mathbf{x}) = \frac{\Gamma(\tilde{\alpha})}{\Gamma(\tilde{\alpha} + \frac{1}{2})\tilde{\beta}^2}.$$

It is easy to see that

$$\delta_s^{\pi, \sigma^2}(\mathbf{x}) \neq [\delta_s^{\pi, \sigma}(\mathbf{x})]^2.$$

Similarly,

$$\delta_2^{\pi, \sigma^2}(\mathbf{x}) \neq [\delta_2^{\pi, \sigma}(\mathbf{x})]^2.$$

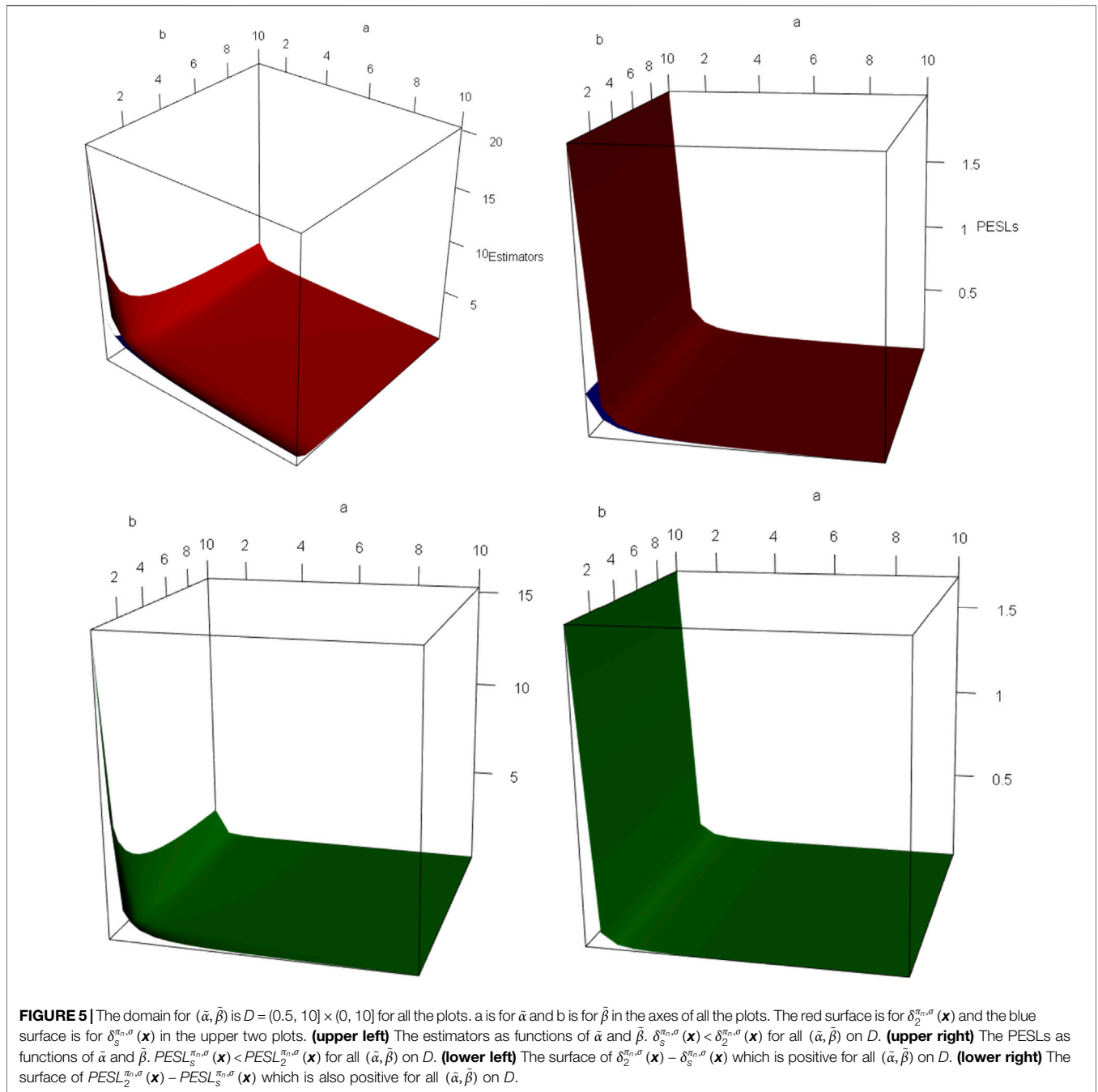


TABLE 3 | The Bayes estimators and the PESLs of the S&P 500 monthly simple returns.

	Conjugate prior		Noninformative prior	
	$\theta = \sigma^2$	$\theta = \sigma$	$\theta = \sigma^2$	$\theta = \sigma$
$\delta_s^{\pi, \theta}(\mathbf{x})$	0.111474	0.338545	0.000408	0.020528
$\delta_2^{\pi, \theta}(\mathbf{x})$	0.125408	0.348644	0.000467	0.021224
$PESL_s^{\pi, \theta}(\mathbf{x})$	0.056583	0.014410	0.063800	0.016285
$PESL_2^{\pi, \theta}(\mathbf{x})$	0.063800	0.014846	0.073126	0.016846

When there is no prior information about the unknown parameter of interest, we prefer the noninformative prior, as the hyperparameters α and β are somewhat arbitrary for the conjugate prior.

We remark that the Bayes estimator under Stein’s loss function is more appropriate than that under the squared error loss function, not because the former is smaller, but because Stein’s loss function which penalizes gross overestimation and gross underestimation equally is more appropriate for the positive restricted parameter.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

This work was carried out in collaboration among all authors. Author YYZ wrote the first draft of the article. Author TZR did literature searches and revised the article. Author MML revised the article. All authors read and approved the final article.

FUNDING

The research was supported by the Ministry of Education (MOE) project of Humanities and Social Sciences on the west and the

REFERENCES

- Adler, D., and Murdoch, D. (2017). *Rgl: 3D Visualization Using OpenGL*. OthersR package version 0.98.1.
- Berger, J. O., and Bernardo, J. M. (1992). *On the Development of the Reference Prior Method*. Bayesian Statistics 4. London: Oxford University Press.
- Berger, J. O., Bernardo, J. M., and Sun, D. C. (2015). Overall Objective Priors. *Bayesian Anal.* 10, 189–221. doi:10.1214/14-ba915
- Berger, J. O. (2006). The Case for Objective Bayesian Analysis. *Bayesian Anal.* 1, 385–402. doi:10.1214/06-ba115
- Bernardo, J. M. (1979). Reference Posterior Distributions for Bayesian Inference. *J. R. Stat. Soc. Ser. B (Methodological)* 41, 113–128. doi:10.1111/j.2517-6161.1979.tb01066.x
- Bobotas, P., and Kourouklis, S. (2010). On the Estimation of a Normal Precision and a Normal Variance Ratio. *Stat. Methodol.* 7, 445–463. doi:10.1016/j.stamet.2010.01.001
- Casella, G., and Berger, R. L. (2002). *Statistical Inference (USA: Duxbury)*. 2nd edition.
- Chen, M. H. (2014). *Bayesian Statistics Lecture*. Changchun, China: Statistics Graduate Summer School/School of Mathematics and Statistics, Northeast Normal University.
- Datta, G. S., and Mukerjee, R. (2004). *Probability Matching Priors: Higher Order Asymptotics*. New York: Springer.
- Ghosh, J. K., Delampady, M., and Samanta, T. (2006). *An Introduction to Bayesian Analysis*. New York: Springer.
- James, W., and Stein, C. (1961). Estimation with Quadratic Loss. *Proc. Fourth Berkeley Symp. Math. Stat. Probab.* 1, 361–380.
- Jeffreys, H. (1961). *Theory of Probability*. 3rd edition. Oxford: Clarendon Press.
- Lehmann, E. L., and Casella, G. (1998). *Theory of Point Estimation*. 2nd edition. New York: Springer.
- Maatta, J. M., and Casella, G. (1990). Developments in Decision-Theoretic Variance Estimation. *Stat. Sci.* 5, 90–120. doi:10.1214/ss/1177012263
- Mao, S. S., and Tang, Y. C. (2012). *Bayesian Statistics*. 2nd edition. Beijing, China: Statistics Press.
- Oono, Y., and Shinozaki, N. (2006). On a Class of Improved Estimators of Variance and Estimation under Order Restriction. *J. Stat. Plann. Inference* 136, 2584–2605. doi:10.1016/j.jspi.2004.10.023
- Petropoulos, C., and Kourouklis, S. (2005). Estimation of a Scale Parameter in Mixture Models with Unknown Location. *J. Stat. Plann. Inference* 128, 191–218. doi:10.1016/j.jspi.2003.09.028
- R Core Team. R (2021). *A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- border area (20XJC910001), the National Social Science Fund of China (21XTJ001), the National Natural Science Foundation of China (12001068; 72071019), and the Fundamental Research Funds for the Central Universities (2020CDJQY-Z001; 2021CDJQY-047).

ACKNOWLEDGMENTS

The authors are extremely grateful to the editor, the guest associate editor, and the reviewers for their insightful comments that led to significant improvement of the article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fdata.2021.763925/full#supplementary-material>

- Robert, C. P. (2007). *The Bayesian Choice: From Decision-Theoretic Motivations to Computational Implementation*. 2nd paperback edition. New York: Springer.
- Ryan, J. A., and Ulrich, J. M. (2017). *R Package Version 0, 4–10*. Quantmod: Quantitative Financial Modelling Framework.
- Sindhu, T. N., and Aslam, M. (2013). Bayesian Estimation on the Proportional Inverse Weibull Distribution under Different Loss Functions. *Adv. Agric. Sci. Eng. Res.* 3, 641–655.
- Sindhu, T. N., Aslam, M., and Hussain, Z. (2016). Bayesian Estimation on the Generalized Logistic Distribution under Left Type-II Censoring. *Thailand Statistician* 14, 181–195.
- Sindhu, T. N., and Aslam, M. (2013). Objective Bayesian Analysis for the Gompertz Distribution under Doudly Type II Cesored Data. *Scientific J. Rev.* 2, 194–208.
- Sindhu, T. N., and Hussain, Z. (2018). Mixture of Two Generalized Inverted Exponential Distributions with Censored Sample: Properties and Estimation. *Stat. Applicata-Italian J. Appl. Stat.* 30, 373–391.
- Sindhu, T. N., Saleem, M., and Aslam, M. (2013). Bayesian Estimation for Topp Leone Distribution under Trimmed Samples. *J. Basic Appl. Scientific Res.* 3, 347–360.
- Sindhu, T. N., Aslam, M., and Hussain, Z. (2016). A Simulation Study of Parameters for the Censored Shifted Gompertz Mixture Distribution: A Bayesian Approach. *J. Stat. Manage. Syst.* 19, 423–450. doi:10.1080/09720510.2015.1103462
- Sindhu, T. N., Feroze, N., and Aslam, M. (2017). A Class of Improved Informative Priors for Bayesian Analysis of Two-Component Mixture of Failure Time Distributions from Doubly Censored Data. *J. Stat. Manage. Syst.* 20, 871–900. doi:10.1080/09720510.2015.1121597
- Sindhu, T. N., Khan, H. M., Hussain, Z., and Al-Zahrani, B. (2018). Bayesian Inference from the Mixture of Half-Normal Distributions under Censoring. *J. Natn. Sci. Found. Sri Lanka* 46, 587–600. doi:10.4038/jnsfr.v46i4.8633
- Sindhu, T. N., Riaz, M., Aslam, M., and Ahmed, Z. (2016). Bayes Estimation of Gumbel Mixture Models with Industrial Applications. *Trans. Inst. Meas. Control.* 38, 201–214. doi:10.1177/0142331215578690
- Sun, J., Zhang, Y.-Y., and Sun, Y. (2021). The Empirical Bayes Estimators of the Rate Parameter of the Inverse Gamma Distribution with a Conjugate Inverse Gamma Prior under Stein's Loss Function. *J. Stat. Comput. Simulation* 91, 1504–1523. doi:10.1080/00949655.2020.1858299
- Tibshirani, R. (1989). Noninformative Priors for One Parameter of Many. *Biometrika* 76, 604–608. doi:10.1093/biomet/76.3.604
- Varian, H. R. (1975). "A Bayesian Approach to Real Estate Assessment," in *Studies in Bayesian Econometrics and Statistics*. Editors S. E. Fienberg and A. Zellner (Amsterdam: North Holland), 195–208.

- Xie, Y.-H., Song, W.-H., Zhou, M.-Q., and Zhang, Y.-Y. (2018). The Bayes Posterior Estimator of the Variance Parameter of the Normal Distribution with a Normal-Inverse-Gamma Prior Under Stein's Loss. *Chin. J. Appl. Probab. Stat.* 34, 551–564.
- Zellner, A. (1986). Bayesian Estimation and Prediction Using Asymmetric Loss Functions. *J. Am. Stat. Assoc.* 81, 446–451. doi:10.1080/01621459.1986.10478289
- Zhang, Y.-Y. (2017). The Bayes Rule of the Variance Parameter of the Hierarchical Normal and Inverse Gamma Model under Stein's Loss. *Commun. Stat. - Theor. Methods* 46, 7125–7133. doi:10.1080/03610926.2016.1148733
- Zhang, Y.-Y., Wang, Z.-Y., Duan, Z.-M., and Mi, W. (2019). The Empirical Bayes Estimators of the Parameter of the Poisson Distribution with a Conjugate Gamma Prior under Stein's Loss Function. *J. Stat. Comput. Simulation* 89, 3061–3074. doi:10.1080/00949655.2019.1652606
- Zhang, Y.-Y., Xie, Y.-H., Song, W.-H., and Zhou, M.-Q. (2018). Three Strings of Inequalities Among Six Bayes Estimators. *Commun. Stat. - Theor. Methods* 47, 1953–1961. doi:10.1080/03610926.2017.1335411
- Zhang, Y.-Y., Zhou, M.-Q., Xie, Y.-H., and Song, W.-H. (2017). The Bayes Rule of the Parameter in (0,1) under the Power-Log Loss Function with an Application to the Beta-Binomial Model. *J. Stat. Comput. Simulation* 87, 2724–2737. doi:10.1080/00949655.2017.1343332
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
- Copyright © 2022 Zhang, Rong and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*