Thailand Statistician
April 2025; 23(2): 355-367
http://statassoc.or.th
Contributed paper

A Pseudo Estimation of Variance using Prior Information with Unknown Shape Parameter: A Study in Normal Case

Shreya Bhunia*, **Babulal Seal and Proloy Banerjee**
Department of Mathematics and Statistics, Aliah University, Kolkata, West Bengal, India.
*Corresponding author; e-mail: shreyabhunia.stat@gmail.com

Received: 18 August 2022
Revised: 10 October 2022
Accepted: 31 December 2022

Abstract

In this article, a variant of Bayes estimate for variance parameter from normal distribution is investigated under weighted squared error loss function. For this purpose, inverse gamma conjugate prior distribution with an unknown hyper-parameter is considered. Further, empirical Bayes approach is used to estimate that unknown hyper-parameter from the marginal likelihood equation. We consider some numerical iterative procedure to approximate the hyper-parameter and in this context, the Bayes estimator may be called a pseudo empirical Bayes estimator. To study the performance of the estimator, the integrated risk and bias performance are carried out through an extensive simulation. It is seen that the estimator performs well in accordance with the integrated risk values. To reduce the bias of the estimator, jackknife resampling technique is used further. Finally, the efficiency of the proposed estimator is studied using two real-life datasets and it is found to be satisfactory as they produce lower posterior risk values.

Keywords: Jackknife resampling method, empirical Bayes, risk values.

1. Introduction

In Bayesian analysis, the procedure using the estimates of the hierarchical parameters of the prior from the data, is known as empirical Bayes method or hierarchical Bayes method. The difference between these two methods is that hierarchical involves hyperpriors for estimating the unknown parameter whereas empirical method estimates that parameter directly from the data. Works on empirical and hierarchical Bayes procedure were done in different areas like epidemiology (Greenland, 2009; Maclehose and Hamra, 2014), biopharmaceutical research (Louis, 1991), count data analysis (Tunaru, 2002) etc. A Quasi-Empirical Bayes method for small area estimation was developed by Raghunathan (1993). The canonical empirical Bayes and a general maximum likelihood empirical Bayes method for the problem of estimating Normal means was discussed by Zhang (1997) and Jiang and Zhang (2009) respectively. For large scale prediction problem, empirical Bayes estimates were used by Efron (2009). Also, the empirical Bayes estimation procedure has been worked out by Seal and Hossain (2015) for the transition probability matrix whether the states seems to be equal or not. Hierarchical empirical Bayes and Benchmarked hierarchical empirical Bayes estimators of positive small area means under multiplicative models were developed by Ghosh et al. (2015). When no proper information regarding the joint parameters of the model of the variable is available, Banerjee and Seal (2021) introduced Partial Bayes estimation process to estimate only the parameter of interest in presence of another nuisance parameter. Raweesawat et al. (2017) considered the empirical Bayes estimator of odds ratio in rare data. Zhang et al. (2019) derived empirical Bayes estimators of the mean and variance parameters of the normal distribution using a conjugate normal-inverse-gamma prior by the method of moment and the maximum likelihood estimation. Grabski and Załęska-Fornal (2010) applied bootstrap resampling methods in an empirical Bayes estimation of the reliability parameters for censored data. Based on upper record values, entropy Bayesian estimation for Lomax distribution has been done by Hassan and Zaky (2021).

In this study, we obtain the Bayes estimate of the unknown variance parameter of the normal distribution in presence of unknown shape parameter of the prior distribution. We consider an empirical Bayes approach to estimate that shape parameter from marginal likelihood equation. It is observed that the estimator cannot be obtained in closed form. Therefore, some numerical iteration procedure is taken into consideration and then after substituting it, the empirical Bayes estimate of variance parameter is obtained. Finally, to study the performance of this estimator, integrated risk is calculated through an extensive simulation process. Furthermore, jackknife resampling technique is used to modify the bias performance of the estimator as an approximation technique is used. Also, to demonstrate the performance of the proposed estimator in real-life scenario, two carbon fiber datasets are considered.

We consider $N(\theta, \sigma^2)$ as the baseline distribution with θ being known and σ^2 is our parameter of interest. A conjugate prior is taken as $\sigma^2 \sim IG(\alpha, p)$; where α is the unknown parameter and p being known. In Section 2, the Bayes estimate of σ^2 is obtained under weighted squared error loss function but due to the presence of hyper-parameter α, the expression of $\hat{\sigma}^2$ is to be modified. In Section 3, the unknown hyper-parameter α is estimated through empirical Bayes approach. But due to intractability of equation, some iterative procedure is considered and thus the approximate Bayes estimator $\hat{\sigma}^2$ is obtained. We call it pseudo empirical Bayes estimate. In this section the method is described fully. In Section 4, the integrated risk of this estimator is calculated numerically. In Section 5, the bias performance of $\hat{\sigma}^2$ is taken by adjusting the bias through jackknife resampling technique for different choices of parameter values. Real data applications are presented and discussed in Section 6. Finally we conclude the paper in Section 7.

2. Bayes Estimation under Weighted Squared Error Loss Function

Let $x_1, x_2, \cdots, x_n$ be a random sample drawn from $N(\theta, \sigma^2)$, with parameters θ and σ^2. Then the likelihood function becomes

$$f(\mathbf{X} \mid \sigma^2) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\theta)^2}; -\infty < \theta < +\infty \text{ and } \sigma^2 > 0. \tag{1}$$

Since the selection of prior information is an essential part in Bayesian context, in this study the inverse gamma distribution is chosen as the prior distribution for the scale parameter σ^2 i.e. $\sigma^2 \sim IG(\alpha, p)$, with shape parameter α being unknown and scale parameter p is known. The pdf of the prior distribution of σ^2 is

$$\pi(\sigma^2) = \frac{1}{p^\alpha \Gamma(\alpha)} (\sigma^2)^{-\alpha-1} e^{-\frac{1}{p\sigma^2}}; \text{ where } p > 0 \text{ and } \alpha > 0. \tag{2}$$

Combining the likelihood function (1) and the inverse gamma prior (2), the posterior distribution of σ^2 given the data $\mathbf{X} = (x_1, x_2, \cdots, x_n)$ becomes

$$h(\sigma^2 \mid \mathbf{X}) = \frac{(\sigma^2)^{-(\alpha+\frac{n}{2})-1} e^{-1/\left\{\frac{2p}{p\sum_{i=1}^n (x_i-\theta)^2+2}\right\}\sigma^2}}{\left\{\frac{2p}{p\sum_{i=1}^n (x_i-\theta)^2+2}\right\}^{(\alpha+\frac{n}{2})} \Gamma(\alpha+\frac{n}{2})}, \tag{3}$$

and it follows that the posterior density also belongs to the inverse gamma distribution.

Therefore,

$$\sigma^2 \mid \mathbf{X} \sim IG\left(\alpha + \frac{n}{2}, \frac{2p}{p\sum_{i=1}^{n}(x_i-\theta)^2+2}\right).$$

For scale parameter estimation, the choice for the loss function is usually weighted squared error loss (WSEL) (Ferguson, 2014)

$$L(\sigma^2, \hat{\sigma^2}) = \left(\frac{\hat{\sigma^2}}{\sigma^2} - 1\right)^2 \tag{4}$$

$$\text{i.e;} \quad L(\sigma^2, \hat{\sigma^2}) = (\hat{\sigma^2} - \sigma^2)^2 \omega(\sigma^2)$$

$$\text{and} \quad w(\sigma^2) = \frac{1}{(\sigma^2)^2}.$$

So the Bayes estimate of the parameter of interest σ^2 based on the WSEL function is given by

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{E\{\sigma^2.w(\sigma^2)|X=x\}}{E\{w(\sigma^2)|X=x\}} = \frac{\int_0^\infty \sigma^2 \frac{1}{(\sigma^2)^2} h(\sigma^2|\mathbf{X})d\sigma^2}{\int_0^\infty \frac{1}{(\sigma^2)^2} h(\sigma^2|\mathbf{X})d\sigma^2} \\
&= \frac{\left\{\frac{2p}{p\sum_{i=1}^{n}(x_i-\theta)^2+2}\right\}^{(\alpha+\frac{n}{2}+1)} \Gamma(\alpha+\frac{n}{2}+1)}{\left\{\frac{2p}{p\sum_{i=1}^{n}(x_i-\theta)^2+2}\right\}^{(\alpha+\frac{n}{2}+2)} \Gamma(\alpha+\frac{n}{2}+2)} \\
&= \frac{p\sum_{i=1}^{n}(x_i-\theta)^2+2}{2p(\alpha+\frac{n}{2}+1)}.
\end{aligned} \tag{5}$$

As in Equation (5), θ was assumed to be known, we may take $\theta = 0$, for simplicity.

So, the expression for Bayes estimator becomes

$$\hat{\sigma}^2 = \frac{p\sum_{i=1}^{n}x_i^2+2}{2p(\alpha+\frac{n}{2}+1)}, \text{ provided } \alpha \text{ is unknown.} \tag{6}$$

In Equation (6), $\hat{\sigma}^2$ is not fully known due to the presence of unknown hyper-parameter α and it will be estimated by using following approach.

3. Empirical Bayes Approach

The empirical Bayes method follows the Bayes model except using a fixed pre-specified value for α. Instead of that, the method suggests to estimate α directly from the observed values. The marginal distribution of $\boldsymbol{X}$ contains all the information about α and is obtained as

$$\begin{aligned}
m(\mathbf{X}) &= \int_0^\infty f(\mathbf{X} \mid \sigma^2)\pi(\sigma^2)d\sigma^2 \\
&= \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{p^\alpha \Gamma\alpha} \left\{\frac{2p}{p\sum_{i=1}^{n}x_i^2+2}\right\}^{(\alpha+\frac{n}{2})} \Gamma(\alpha+\frac{n}{2}).
\end{aligned} \tag{7}$$

Now, taking logarithm we get,

$$\begin{aligned}
log\big(m(\mathbf{X})\big) = -\frac{n}{2}log(2\pi) - \alpha\, log\, p - log\Gamma\alpha + \left(\alpha+\frac{n}{2}\right)log\left(\frac{2p}{p\sum_{i=1}^{n}x_i^2+2}\right) \\
+log\left(\Gamma\left(\alpha+\frac{n}{2}\right)\right).
\end{aligned} \tag{8}$$

Now, differentiating (8) w.r.t. α and equating to 0, we get

$$-\,log\, p - \frac{\delta}{\delta\alpha}\, log\Gamma\alpha + log\left(\frac{2p}{p\sum_{i=1}^{n}x_i^2+2}\right) + \frac{\delta}{\delta\alpha} log\left(\Gamma(\alpha+\frac{n}{2})\right) = 0$$

$$or, \frac{\delta}{\delta\alpha} log\Gamma\left(\alpha+\frac{n}{2}\right)-\frac{\delta}{\delta\alpha} log\Gamma\alpha = log\, p - log\left(\frac{2p}{p\sum_{i=1}^{n} x_i^2+2}\right). \tag{9}$$

As it is difficult to obtain the estimate of α explicitly, we calculate $\hat{\alpha}$ numerically by using Newton-Raphson method. After substituting $\hat{\alpha}$ in Equation (6), the Bayes estimate of σ^2 becomes,

$$\hat{\sigma}^2 = \frac{p\sum_{i=1}^{n} x_i^2+2}{2\,p(\hat{\alpha}+\frac{n}{2}+1)}. \tag{10}$$

It is to be noted that the estimate (10) is not exactly empirical Bayes as we estimate $\hat{\alpha}$ by using some iterative procedure approximately from the marginal likelihood. In this context, we call it pseudo empirical Bayes estimate (PEBE).

4. Integrated Risk using Simulation

In Bayesian inference, risk is one of the most important factors as it quantifies the deviation between the estimate and the unknown parameter through the loss function. The behaviour of PEBE has been studied on the basis of integrated risk to extract the estimator's goodness and it is calculated by integrating the risk function $R\left(\sigma^2, \hat{\sigma}^2\right)$ over the prior distribution. So, the integrated risk becomes,

$$\begin{aligned} E_{\sigma^2}[R(\sigma^2,\hat{\sigma^2})] =& E_{\sigma^2}E_{x|\sigma^2}\left[L(\sigma^2,\hat{\sigma^2})\right] \\ =& \int_{-\infty}^{\infty}\int_{0}^{\infty} L(\sigma^2,\hat{\sigma^2})f(x|\sigma^2)\pi(\sigma^2)d\sigma^2dx \\ =& \int_{-\infty}^{\infty}\int_{0}^{\infty}\left(\frac{\hat{\sigma^2}}{\sigma^2}-1\right)^2\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2} \\ & \frac{1}{p^{\alpha}\Gamma(\alpha)}(\sigma^2)^{-\alpha-1}e^{-\frac{1}{p\sigma^2}}d\sigma^2dx. \end{aligned}$$

The above expression is difficult to obtain in an explicit form. So we perform an extensive simulation to calculate the integrated risk of the estimator by using R Core Team (2019) **(Version 3.6.1)**. In order to generate σ^2 from $IG(\alpha, p)$, the shape parameter α being unknown and from (9) it is seen that, the estimate of α contains $X_i's$ which are generated from $N(0,\sigma^2)$. Hence choosing some initial values of α will be the first step of our simulation process. We construct the following algorithm to calculate the integrated risk of $\hat{\sigma}^2$.

Step 1: Initialize some values of α.

Step 2: Fix the first value of α as α_1 and generate σ_1^2, σ_2^2,..., σ_m^2 from IG(α_1,p) for some known values of p.

Step 3: Fix σ_1^2 and generate $\mathbf{X_1}$ of size n from $N(0,\sigma_1^2)$.

Step 4: Repeat Step 3 for K times, i.e. K many observations from $N(0,\sigma_1^2)$.

Step 5: Calculate $\hat{\alpha}$ for K times from equation (9) by some iterative method, e.g. Newton - Raphson etc.

Step 6: Calculate $\hat{\sigma_1^2}$ for K times using the equation (10).

Step 7: Calculate K replicated loss value L(σ_1^2,$\hat{\sigma_1^2}$) using the equation (4).

Step 8: Calculate the risk value $E_{X|\sigma_1^2}$[L(σ_1^2,$\hat{\sigma_1^2}$)] by taking an average of K replicates of loss values and this can be done according to the strong law of large numbers (SLLN).

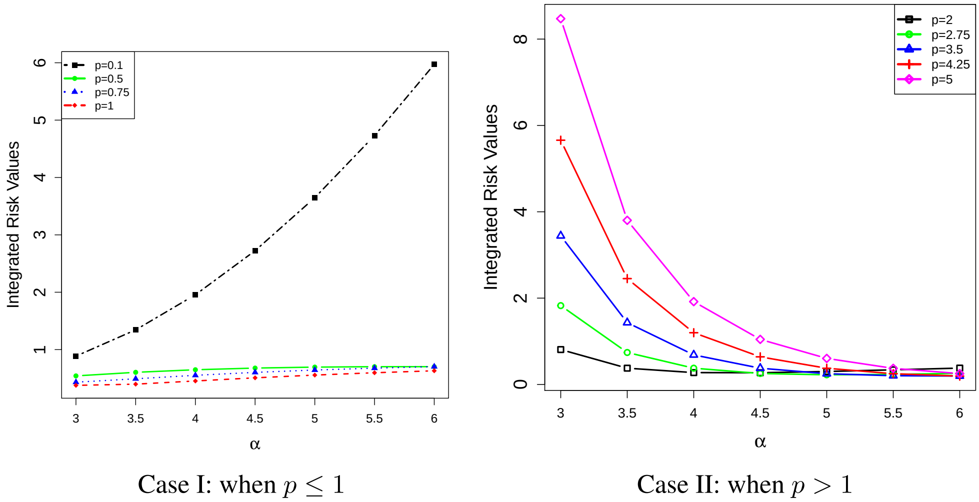

Figure 1 Integrated risk of pseudo empirical Bayes estimator

Table 1 Integrated risk of pseudo empirical Bayes estimate (PEBE)

Sample sizes	Parameter choices	$\alpha = 3$	$\alpha = 3.5$	$\alpha = 4$	$\alpha = 4.5$	$\alpha = 5$	$\alpha = 5.5$	$\alpha = 6$
	p=0.1	0.8861	1.3413	1.9513	2.7176	3.6412	4.7228	5.9630
	p=0.5	0.5701	0.6225	0.6607	0.6862	0.7015	0.7088	0.7099
$n = 10$	p=0.75	0.4802	0.5309	0.5820	0.6250	0.6601	0.6887	0.7118
	p=1.0	0.4350	0.4562	0.5027	0.5480	0.5878	0.6219	0.6511
	p=0.1	0.8862	1.3423	1.9530	2.7198	3.6438	4.7257	5.9661
	p=0.5	0.5522	0.6117	0.6549	0.6833	0.7000	0.7076	0.7086
$n = 25$	p=0.75	0.4475	0.5048	0.5628	0.6112	0.6504	0.6819	0.7073
	p=1.0	0.3946	0.4169	0.4700	0.5220	0.5673	0.6058	0.6384
	p=0.1	0.8859	1.3440	1.9559	2.7236	3.6482	4.7305	5.9713
	p=0.5	0.5415	0.6037	0.6487	0.6778	0.6946	0.7018	0.7022
$n = 50$	p=0.75	0.4315	0.4910	0.5518	0.6027	0.6438	0.6766	0.7029
	p=1.0	0.3764	0.3976	0.4530	0.5078	0.5556	0.5963	0.6306
	p=0.1	0.8857	1.3457	1.9589	2.7274	3.6526	4.7354	5.9766
	p=0.5	0.5355	0.5984	0.6435	0.6722	0.6883	0.6948	0.6950
$n = 75$	p=0.75	0.4245	0.4845	0.5463	0.5981	0.6398	0.6729	0.6994
	p=1.0	0.3691	0.3893	0.4454	0.5012	0.5500	0.5914	0.6265
	p=0.1	0.8856	1.3474	1.9618	2.7312	3.6570	4.7403	5.9819
	p=0.5	0.5314	0.5942	0.6385	0.6665	0.6816	0.6875	0.6875
$n = 100$	p=0.75	0.4208	0.4809	0.5432	0.5953	0.6370	0.6701	0.6965
	p=1.0	0.3652	0.3849	0.4414	0.4976	0.5469	0.5888	0.6241

Step 9: Repeat **Step 3** - **Step 7** for remaining σ_i^2 ; i = 2, 3, ..., m.

Step 10: Calculate $E_{\sigma^2} E_{X|\sigma^2}$[L($\sigma^2$,$\hat{\sigma^2}$)] by taking average of such m replicated risk values.

Step 11: Repeat **Step 2** - **Step 9** to obtain the integrated risk of $\hat{\sigma}^2$ for remaining values of α.

Table 2 Integrated risk of pseudo empirical Bayes estimate (PEBE)

Sample sizes	Parameter choices	$\alpha=3$	$\alpha=3.5$	$\alpha=4$	$\alpha=4.5$	$\alpha=5$	$\alpha=5.5$	$\alpha=6$
$n=10$	p=2.0	0.7735	0.4370	0.3618	0.3632	0.3889	0.4212	0.4539
	p=2.75	1.5873	0.7203	0.4371	0.3455	0.3264	0.3368	0.3591
	p=3.5	2.8947	1.2698	0.6755	0.4359	0.3399	0.3081	0.3065
	p=4.25	4.7012	2.0909	1.0821	0.6393	0.4336	0.3385	0.2990
	p=5.0	7.0095	3.1861	1.6594	0.9583	0.6100	0.4307	0.3389
$n=25$	p=2.0	0.7914	0.3927	0.3006	0.3000	0.3288	0.3660	0.4042
	p=2.75	1.7408	0.7274	0.3913	0.2798	0.2545	0.2646	0.2894
	p=3.5	3.2560	1.3709	0.6743	0.3897	0.2733	0.2326	0.2282
	p=4.25	5.3392	2.3255	1.1518	0.6319	0.3873	0.2722	0.2225
	p=5.0	7.9914	3.5923	1.8248	1.0073	0.5977	0.3844	0.2734
$n=50$	p=2.0	0.8088	0.3776	0.2752	0.2719	0.3009	0.3396	0.3795
	p=2.75	1.8267	0.7411	0.3776	0.2547	0.2247	0.2334	0.2583
	p=3.5	3.4429	1.4324	0.6851	0.3772	0.2493	0.2030	0.1961
	p=4.25	5.6582	2.4526	1.1988	0.6403	0.3756	0.2495	0.1938
	p=5.0	8.4731	3.8020	1.9190	1.0446	0.6042	0.3734	0.2519
$n=75$	p=2.0	0.8187	0.3730	0.2658	0.2609	0.2897	0.3287	0.3692
	p=2.75	1.8668	0.7499	0.3742	0.2460	0.2137	0.2214	0.2461
	p=3.5	3.5266	1.4622	0.6929	0.3746	0.2415	0.1925	0.1842
	p=4.25	5.7986	2.5108	1.2224	0.6472	0.3736	0.2425	0.1839
	p=5.0	8.6832	3.8959	1.9633	1.0642	0.6104	0.3717	0.2455
$n=100$	p=2.0	0.8236	0.3707	0.2611	0.2554	0.2840	0.3231	0.3639
	p=2.75	1.8867	0.7542	0.3725	0.2417	0.2082	0.2154	0.2400
	p=3.5	3.5680	1.4770	0.6967	0.3733	0.2375	0.1872	0.1782
	p=4.25	5.8680	2.5396	1.2342	0.6506	0.3726	0.2390	0.1789
	p=5.0	8.7868	3.9423	1.9852	1.0739	0.6135	0.3709	0.2424

It is to be noted that, σ^2 is an original variance parameter of the baseline distribution. So, we choose such combinations of the hyper-parameters α and p for which small σ^2 values are generated. Also, the shape parameter α should not be too small as it generates large value of σ^2. So the initial choices of α are taken as $3,\ 3.5,\ 4,\ 4.5,\ 5,\ 5.5$ and 6.

In this study, the scale parameter of the prior distribution is known. But larger value of the scale parameter indicates higher value of the variance. So, in practical situation to find out a prior distribution having large scale or variance is absolutely insignificant. Therefore, as a rational case, we consider p to be small. Moreover, the choices of scale parameter p is divided into two cases: (i) $p = 0.1, 0.5, 0.75, 1$ i.e. p belongs to the range 0 to 1 and (ii) $p > 1$, in particular, we choose $p = 2, 2.75, 3.5, 4.25, 5$.

By considering each pair of (α, p), we generate 50 random σ^2 and for each of the σ^2, random samples are simulated from $N(0, \sigma^2)$ of different sizes $n = 10, 25, 50, 75, 100$. The sample sizes are selected in an increasing order to study the effect of small, moderate and large sample behaviour on the estimator $\hat{\sigma}^2$. The number of replication in the simulation process is taken as $K = 5000$ times and based on that simulated samples, the integrated risk of $\hat{\sigma}^2$ has been obtained by following the above mentioned algorithm. The results are presented in Table 1 $(0 < p < 1)$ and Table 2 $(p > 1)$ respectively.

From Table 1, it is observed that integrated risk of $\hat{\sigma}^2$ increases when the choices of α become larger and for any fixed choice of α, the integrated risk of $\hat{\sigma}^2$ decreases with p increases. In Table 2, integrated risk of $\hat{\sigma}^2$ decreases as α increases for all the chosen values of p. Also from Figure 1, it is clear that the pseudo empirical Bayes estimator $\hat{\sigma}^2$ performs well for most of the parameter choices. The variation in integrated risk values are due to the combined effect of the prior parameters. Thus we slightly modify it through the bias performance of $\hat{\sigma}^2$ with the help of jackknife resampling technique in the next section.

5. Bias Performance using Jackknife Method

In statistical inference, unbiasedness is an important characteristic to measure the goodness of an estimator. But, it is well known that Bayes estimator is always a biased estimator under squared error loss function. As a biased estimator is undesirable, so here we try to reduce the bias with the help of resampling technique since we have obtained approximately in previous section. Jackknife method is one of the resampling techniques introduced by Quenouille (1956) and named by Tukey (1958). It is conducted based on aggregation of the obtained estimates of each $(n-1)$ sized sub sample. The jackknife bias is defined as,

$$\widehat{Bias}_{Jack} = (n-1)(\bar{\theta} - \hat{\theta}); \;\; \bar{\theta} = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_i,$$

where, $\hat{\theta}$ is the estimate of θ considering the entire sample and $\hat{\theta}_i$ be the estimate of θ obtained by deleting the i^{th} observation.

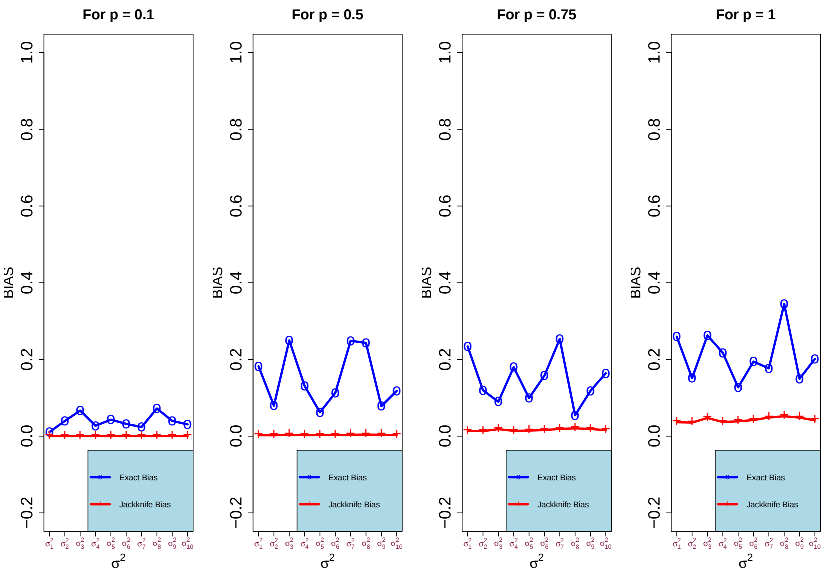

Figure 2 The graphical overview of exact bias and jackknife bias when α=3

To study the bias performance of $\hat{\sigma}^2$, we calculate the exact bias and jackknife bias numerically by considering the prior parameters as $\alpha = 3, 3.5, 4, 4.5, 5$ and $p = 0.1, 0.5, 0.75, 1$. For these choices of (α, p), σ^2 are simulated from inverse gamma distribution and we use moderate sample size to generate samples from $N(0, \sigma^2)$. Both the exact bias and jackknife bias have been obtained based on 5000 simulated samples and are displayed in Table 3.

To understand the exact and jackknife bias from Table 3, we represent it in a graphical way for some of the chosen values of α. From Figure 2 and 3, it is seen that for the choice $p = 0.1$, both the exact and jackknife biases are close to zero. As p increases, both the the bias of $\hat{\sigma}^2$ increases but in case of exact bias, the rate of change of the bias is higher as compared to the jackknife bias. That is why, we adopt the jackknife resampling technique to reduce the bias of the estimate. It illustrates that the jackknife bias of $\hat{\sigma}^2$ is almost close to zero for the chosen values of α and p.

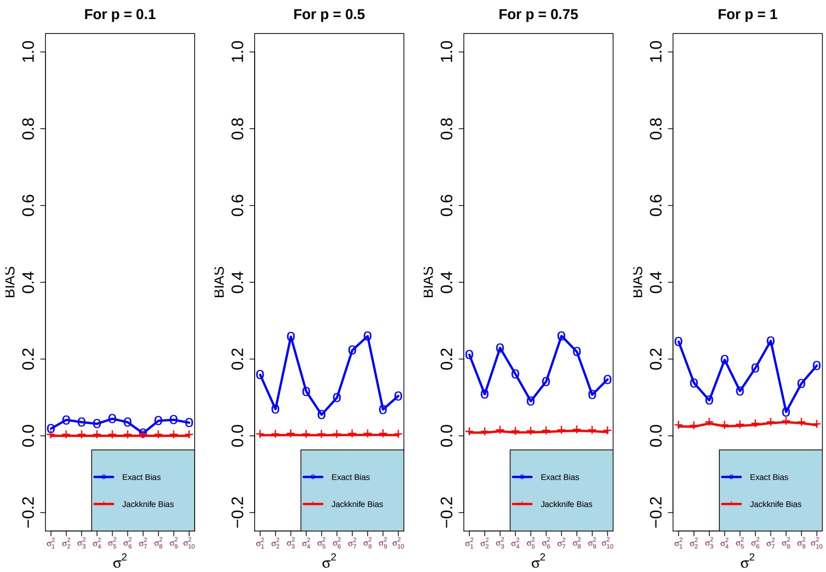

Figure 3 The graphical overview of exact bias and jackknife bias when α=4

Table 3 Bias performance of pseudo empirical Bayes estimator (PEBE)

	p=0.1		p=0.5		p=0.75		p=1	
Parameter choices	Exact Bias	Jackknife Bias	Exact Bias	Jackknife Bias	Exact Bias	Jackknife Bias	Exact Bias	Jackknife Bias
	0.0006	0.0001	0.2105	0.0045	0.2550	0.0225	0.2583	0.0565
	0.0369	0.0001	0.0922	0.0044	0.1337	0.0218	0.1678	0.0551
	0.1215	0.0002	0.1127	0.0064	0.4344	0.0301	1.3895	0.0727
	0.0212	0.0001	0.1510	0.0046	0.2026	0.0229	0.2385	0.0573
	0.0412	0.0001	0.0715	0.0049	0.1104	0.0240	0.1405	0.0599
$\alpha = 3$	0.0276	0.0001	0.1294	0.0055	0.1776	0.0264	0.2153	0.0648
	0.0514	0.0002	0.2599	0.0066	0.1727	0.0309	0.0619	0.0744
	0.1308	0.0002	0.0737	0.0072	0.5595	0.0332	1.6455	0.0792
	0.0372	0.0002	0.0907	0.0066	0.1320	0.0307	0.1658	0.0740
	0.0258	0.0001	0.1355	0.0053	0.1850	0.0258	0.2225	0.0636
	0.0108	0.0001	0.1821	0.0025	0.2339	0.0135	0.2595	0.0364
	0.0395	0.0001	0.0795	0.0024	0.1196	0.0131	0.1514	0.0354
	0.0667	0.0001	0.2496	0.0036	0.0899	0.0184	0.2627	0.0475
	0.0271	0.0001	0.1311	0.0026	0.1798	0.0138	0.2174	0.0370
	0.0431	0.0001	0.0622	0.0027	0.0994	0.0145	0.1273	0.0387
$\alpha = 3.5$	0.0321	0.0001	0.1129	0.0030	0.1576	0.0160	0.1945	0.0421
	0.0238	0.0001	0.2487	0.0037	0.2534	0.0189	0.1768	0.0486
	0.0721	0.0001	0.2431	0.0041	0.0545	0.0204	0.3445	0.0520
	0.0398	0.0001	0.0782	0.0037	0.1181	0.0188	0.1496	0.0483

Table 3 (continued)

	0.0307	0.0001	0.1181	0.0030	0.1640	0.0156	0.2013	0.0413
	0.0186	0.0001	0.1592	0.0017	0.2115	0.0086	0.2456	0.0245
	0.0415	0.0001	0.0697	0.0017	0.1083	0.0083	0.1380	0.0238
	0.0357	0.0001	0.2582	0.0022	0.2291	0.0118	0.0925	0.0324
	0.0313	0.0001	0.1157	0.0017	0.1610	0.0088	0.1981	0.0249
	0.0445	0.0001	0.0549	0.0018	0.0904	0.0092	0.1166	0.0261
$\alpha = 4$	0.0354	0.0001	0.0992	0.0019	0.1415	0.0102	0.1767	0.0286
	0.0067	0.0001	0.2229	0.0022	0.2602	0.0122	0.2473	0.0332
	0.0392	0.0001	0.2596	0.0024	0.2189	0.0132	0.0625	0.0356
	0.0417	0.0001	0.0686	0.0022	0.1070	0.0121	0.1364	0.0330
	0.0342	0.0001	0.1041	0.0019	0.1471	0.0100	0.1830	0.0280
	0.0243	0.0001	0.1408	0.0013	0.1911	0.0057	0.2283	0.0171
	0.0431	0.0001	0.0619	0.0013	0.0990	0.0055	0.1268	0.0166
	0.0162	0.0001	0.2390	0.0016	0.2604	0.0079	0.2153	0.0229
	0.0346	0.0001	0.1027	0.0013	0.1456	0.0058	0.1812	0.0174
	0.0457	0.0001	0.0491	0.0014	0.0826	0.0061	0.1077	0.0183
$\alpha = 4.5$	0.0379	0.0001	0.0874	0.0015	0.1284	0.0068	0.1617	0.0201
	0.0049	0.0001	0.1976	0.0016	0.2466	0.0082	0.2620	0.0235
	0.0186	0.0001	0.2424	0.0017	0.2589	0.0089	0.2042	0.0253
	0.0433	0.0001	0.0609	0.0016	0.0979	0.0081	0.1255	0.0233
	0.0370	0.0001	0.0919	0.0014	0.1333	0.0067	0.1674	0.0196
	0.0286	0.0000	0.1259	0.0011	0.1735	0.0039	0.2111	0.0123
	0.0444	0.0000	0.0555	0.0010	0.0911	0.0038	0.1174	0.0119
	0.0029	0.0001	0.2154	0.0013	0.2574	0.0055	0.2550	0.0166
	0.0371	0.0000	0.0913	0.0011	0.1327	0.0040	0.1667	0.0125
	0.0467	0.0000	0.0443	0.0011	0.0757	0.0042	0.1001	0.0132
$\alpha = 5$	0.0399	0.0000	0.0776	0.0012	0.1174	0.0047	0.1489	0.0145
	0.0131	0.0001	0.1757	0.0013	0.2280	0.0057	0.2566	0.0171
	0.0047	0.0001	0.2191	0.0014	0.2590	0.0062	0.2517	0.0184
	0.0446	0.0001	0.0547	0.0013	0.0901	0.0056	0.1162	0.0169
	0.0391	0.0000	0.0815	0.0012	0.1218	0.0046	0.1540	0.0142

6. Real Data Application

In this section, the pseudo empirical Bayes estimation procedure is applied to two real datasets. The first dataset is a strength data for single carbon fibers and impregnated 1000-cardon fiber tows, measured in GPa and it is obtained from Ristic and Kundu (2015) whereas, the second data is taken from Nichols and Padgett (2006) representing the breaking stress of carbon fibers of 50 mm in length. Both the datasets are provided below:

Dataset I : Strength data for single carbon fibers

2.247	2.640	2.908	3.099	3.126	3.245	3.328	3.355	3.383	3.572	3.581	3.681
3.726	3.727	3.728	3.783	3.785	3.786	3.896	3.912	3.964	4.050	4.063	4.082
4.111	4.118	4.141	4.246	4.251	4.262	4.326	4.402	4.457	4.466	4.519	4.542
4.555	4.614	4.632	4.634	4.636	4.678	4.698	4.738	4.832	4.924	5.043	5.099
5.134	5.359	5.473	5.571	5.684	5.721	5.998	6.06				

A transformation is taken into consideration to incorporate the real data application with the theoretical results by subtracting the mean from the original datasets to make the baseline distribution as $N(0, \sigma^2)$. The sample mean and sample standard deviation of the first dataset is 0 and 0.814, and for the second dataset these are 0 and 1.009 respectively. The respective kurtosis values for both the datasets are 2.9 and 3.1. To determine the model's goodness-of-fit, the Kolmogorov-Smirnov test (K-S) statistics are calculated (Bhunia and Banerjee , 2022). It measures the differences between the theoretical cdf and the empirical cdf from the data. The K-S statistic and the corresponding p-values are 0.064 (0.962) and 0.053 (0.938) respectively for dataset I and II.

Dataset II : Breaking stress of carbon fibers data

0.39	0.81	0.85	0.98	1.08	1.12	1.17	1.18	1.22	1.25	1.36	1.41	1.47	1.57
1.59	1.61	1.69	1.71	1.73	1.80	1.84	1.87	1.89	1.92	2.00	2.03	2.05	2.12
2.17	2.35	2.38	2.41	2.43	2.48	2.50	2.53	2.55	2.56	2.59	2.67	2.73	2.74
2.76	2.77	2.79	2.81	2.82	2.83	2.85	2.87	2.88	2.93	2.95	2.96	2.97	3.09
3.11	3.15	3.19	3.22	3.27	3.28	3.31	3.33	3.39	3.51	3.56	3.60	3.65	3.68
3.70	3.75	4.20	4.38	4.42	4.70	4.90	4.91	5.08	5.56	1.58	1.60	1.61	1.70
1.85	2.04	2.17	2.17	2.48	2.55	2.82	2.98	3.11	3.16	3.19	3.23	3.31	3.39
3.68	3.69												

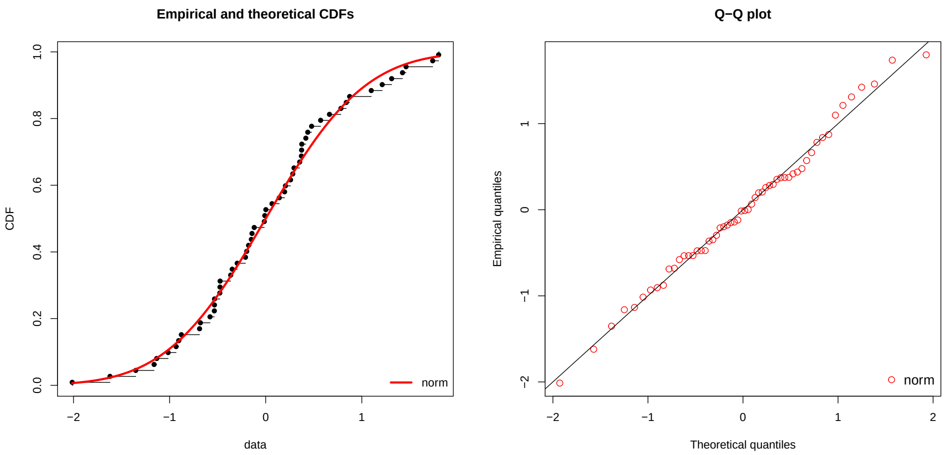

Figure 4 Empirical vs fitted cdf plot and Q-Q plot for dataset I

Apart from the theoretical criteria of model fitting, we also represent it graphically by plotting the empirical cdf with the fitted cdf and the Q-Q plot for each of the datasets in Figure 4 and 5 respectively. It illustrates that in the Q-Q plot, the data points approximately fall on the straight line and also the fitted cdf is close to the empirical cdf for both the datasets.

To obtain the empirical Bayes estimate of σ^2 from the given datasets, we compute $\hat{\alpha}$ by maximizing the marginal likelihood and this maximization procedure is carried out by using Newton Raphson method. In this context, the Bayes estimate may be termed as pseudo empirical Bayes estimate. For both the real datasets, the Bayes estimate, posterior risk and respective 95% and 99% credible intervals have been obtained for different choices of scale parameter and are presented in Table 4 and 5 respectively.

From the above tables, it is observed that the posterior risk values are very small for different choices of scale parameter p. Additionally, as it is expected, the posterior risk increases with the larger choices of the scale parameter. Overall, the pseudo empirical Bayes estimator has proven to be effective for both the datasets. Consequently, these two real-life datasets demonstrate the usefulness

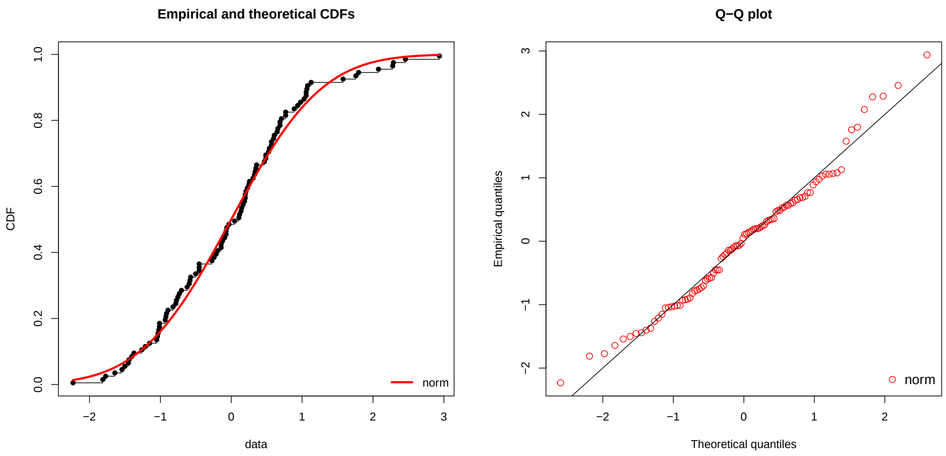

Figure 5 Empirical vs fitted cdf plot and Q-Q plot for dataset II

Table 4 PEBE and posterior risk with associated credible intervals for real dataset I

Choices of Scale Parameter	PEBE	Posterior Risk	Credible Interval 95%	99%
0.1	0.6400878	0.02242504	(0.4968110, 0.9025924)	(0.4569143, 1.0028145)
0.5	0.6321641	0.03077207	(0.4732282, 0.9567252)	(0.4295906, 1.0859900)
1	0.6307794	0.03227576	(0.4694693, 0.9662091)	(0.4252748, 1.1007757)
3	0.630592	0.03340571	(0.4673518, 0.9746058)	(0.4226976, 1.1133630)
5	0.6309434	0.0336621	(0.4671709, 0.9771192)	(0.4223864, 1.1169185)

Table 5 PEBE and posterior risk with associated credible intervals for real dataset II

Choices of Scale Parameter	PEBE	Posterior Risk	Credible Interval 95%	99%
0.1	0.9926128	0.01630763	(0.7960003, 1.321301)	(0.7405010, 1.442592)
0.5	0.9892781	0.01871222	(0.7825748, 1.348082)	(0.7245637, 1.482325)
1	0.9889982	0.01906759	(0.7808530, 1.352250)	(0.7224841, 1.488427)
3	0.9896466	0.01932853	(0.7802756, 1.356469)	(0.7215973, 1.494185)
5	0.9901507	0.01938886	(0.7804227, 1.357930)	(0.7216521, 1.496051)

of the estimation procedure.

7. Conclusion

In this paper, a variant of Bayes estimation process has been described by considering normal distribution as a baseline model with an unknown variance parameter σ^2. It is estimated in the presence of inverse gamma prior distribution, where the hyper-parameter α is unknown. The Bayes estimate of σ^2 is considered to be known when we estimate α through the empirical Bayes approach and substitute it into the original Bayes estimator $\hat{\sigma}^2$. During this process, due to intractability of the marginal likelihood equation, some iteration procedure is taken into consideration to approximate the MLE value and hence the estimator $\hat{\sigma}^2$ may be called as pseudo empirical Bayes estimator (PEBE). We carried out Monte-Carlo simulation technique to study the behaviour of the estimator in accordance with the integrated risk values. Also, we use jackknife resampling technique on the pseudo

Bayes estimate to reduce the bias of the estimator. It is found that the estimator performs quite well for the chosen values of hyper-parameters of the inverse gamma distribution.

Also the real-life example depicts that the pseudo empirical Bayes estimator has very small posterior risk values for the chosen scale parameter of the prior distribution. Furthermore, the estimator PEBE belongs to its corresponding credible interval for both the considered datasets. As a result, the potential of the pseudo empirical Bayes estimator has been illustrated through the real-life datasets. Also, this work may eventually be expanded by comparing this estimator with some variant of Bayes estimators such as partial Bayes, E-Bayesian estimator available in literature.

Acknowledgements

We would like to thank the editor and reviewers for their comments and suggestions on the manuscript.

References

Efron B. Empirical Bayes estimates for large-scale prediction problems. J Am Stat Assoc. 2009; 104(487):1015–1028.

Ferguson TS. Mathematical statistics: A decision theoretic approach. New York and London: Academic press; 2014.

Banerjee P, Seal B. Partial Bayes Estimation of Two Parameter Gamma Distribution under Non-Informative Prior. Stat Optim Inf Comput. 2021; 10(4):1110-1125.

Bhunia S, Banerjee P. Some Properties and Different Estimation Methods for Inverse A (α) Distribution with an Application to Tongue Cancer Data. Reliab.: Theory Appl. 2022; 17(1 (67)):251–266.

Ghosh M, Kubokawa T, Kawakubo Y. Benchmarked empirical Bayes methods in multiplicative area-level models with risk evaluation. Biometrika. 2015; 102(3):647–659.

Greenland S. Bayesian perspectives for epidemiologic research: III. Bias analysis via missing-data methods. Int J Epidemiol. 2009; 38(6): 1662–1673, doi:10.1093/ije/dyp278.

Grabski F, Załęska-Fornal A. Bootstrap methods for the censored data in empirical Bayes estimation of the reliability parameters. Reliab.: Theory Appl. 2010; 5(2 (17)):115–121.

Hassan AS, Zaky AN. Entropy Bayesian estimation for Lomax distribution based on record. Thail Stat. 2021; 19(1):95–114.

Jiang W, Zhang CH. General maximum likelihood empirical Bayes estimation of normal means. Ann Stat. 2009; 37(4):1647–1684.

MacLehose RF, Hamra GB. Applications of Bayesian methods to epidemiologic research. Curr Epidemiol Rep. 2014; 1(3):103-9.

Nichols MD and Padgett WJ. A bootstrap control chart for Weibull percentiles. Qual Reliab Eng Int. 2006; 22(2): 141–151.

Louis TA. Using empirical Bayes methods in biopharmaceutical research. Stat Med. 1991; 10(6):811–829.

Quenouille MH. Notes on bias in estimation. Biometrika. 1956; 43(3/4):353–360.

Raghunathan, TE. A quasi-empirical Bayes method for small area estimation. J Am Stat Assoc. 1993; 88(424):1444–1448.

Ristić MM, Kundu D. Marshall-Olkin generalized exponential distribution. Metron. 2015; 73(3):317–333.

Raweesawat K, Areepong Y, Sukparungsee S, Jampachaisri K. Odds Ratio Estimation in Rare Data by Empirical Bayes Method. Thail Stat. 2017; 15(2):149–156.

R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2019. http://www.R-project.org/.

Seal B, Hossain SJ. Empirical Bayes estimation of parameters in Markov transition probability matrix with computational methods. J Appl Stat. 2015; 42(3):508–519.

Tukey J. Bias and confidence in not quite large samples. Ann Math Statist. 1958; 29: 614.

Tunaru R. Hierarchical Bayesian models for multiple count data. Austrian J Stat. 2002; 31(2&3):221–229, doi:https://doi.org/10.17713/ajs.v31i23.484.

Zhang CH. Empirical Bayes and compound estimation of normal means. Stat Sin. 1997; 7(1):181–193.

Zhang YY, Rong TZ, Li MM. The empirical Bayes estimators of the mean and variance parameters of the normal distribution with a conjugate normal-inverse-gamma prior by the moment method and the MLE method. Commun Stat Theory. 2019; 48(9):2286–2304.