应用概率统计　第 36 卷
第 3 期　2020 年 6 月

Chinese Journal of Applied Probability and Statistics
Jun., 2020, Vol. 36, No. 3, pp. 249-260
doi: 10.3969/j.issn.1001-4268.2020.03.003

Linear Approximate Bayes Estimator for Uniform Distribution *

ZHANG Fengyue　　WANG Lichun⋆

(*School of Science, Beijing Jiaotong University, Beijing, 100044, China*)

Abstract: We employ a linear Bayes procedure to estimate the unknown parameter of the uniform distribution $R(-\theta,\theta)$ and propose a linear approximate Bayes estimator (LABE) for θ, which has a closed analytic solution form and is convenient to use. Numerical simulations indicate that the proposed LABE is close to the ordinary Bayes estimator (BE), which is calculated by numerical integration and the so-called brute-force method as well. Furthermore, we compare the proposed LABE with the Lindley's approximation. The superiorities of the LABE over the classical estimators are also established in terms of the mean squared error (MSE) criterion.

Keywords: uniform distribution; linear approximate Bayes estimator; Lindley's approximation; mean squared error (MSE) criterion

2010 Mathematics Subject Classification: 62C10

Citation: ZHANG F Y, WANG L C. Linear approximate Bayes estimator for uniform distribution [J]. Chinese J Appl Probab Statist, 2020, 36(3): 249–260.

§1. Introduction

Perhaps one of the most important distributions is the uniform distribution for continuous random variables and it has been applied in many fields including natural and social sciences. The theory of uniform distribution first appeared in Hermann's paper published in 1916 and since then many scholars have paid attention to it. Chen and Ni [1] gave some conventional estimators for the parameter of uniform distribution including the maximum likelihood estimator (MLE) and the uniformly minimum variance unbiased estimator (UMVUE) and so on, which are commonly used in classical statistics.

Hartigan [2] considered the linear regression from a Bayesian point of view and proposed a method of linear prediction, which uses only the first two moments of the distribution of parameters and observations, rather than the complete probability distribution

*The project was partly supported by the National Natural Science Foundation of China (Grant No. 11371051).

⋆Corresponding author, E-mail: wlc@amss.ac.cn.

Received November 22, 2018. Revised October 19, 2019.

model. Subsequently, Rao [3] proposed the Linear Bayes procedure from a linear optimization viewpoint. Recently, Wei and Zhang [4] discussed the linear Bayesian estimation and its superiorities for linear model. Wang and Singh [5] studied the linear Bayesian estimation of the two-parameter exponential distribution and exhibited the superiorities of the linear Bayes procedure over some classical estimators. In application fields, Kuo [6] explored the linear Bayes estimators on the class of potency curves in quantal bioassay under the integrated squared error loss. In Bayesian statistics, the posterior distribution offers a sensible compromise between the prior and the observed data, and combined strength of the two sources of information leads to increased precision in the understanding of the parameter, which improves the effect of statistical inference.

We consider the following uniform distribution:

$$X \sim R(-\theta, \theta), \qquad \theta > 0.$$

Let $X_1, X_2, \cdots, X_n$ be independently drawn from the distribution and denote $M = \max_i X_i$, $m = \min_i X_i$.

Then, the joint probability density function of the samples $X = (X_1, X_2, \cdots, X_n)$ is

$$f(x \,|\, \theta) = f(x_1, x_2, \cdots, x_n \,|\, \theta) = \begin{cases} \dfrac{1}{(2\theta)^n}, & -\theta < m < M < \theta; \\ 0, & \text{otherwise}. \end{cases}$$

If we assume that the prior distribution of θ is $\pi(\theta)$, then the posterior distribution of θ, say $\pi(\theta \,|\, x)$, can be obtained by

$$\mathrm{d}\pi(\theta \,|\, x) \propto f(x \,|\, \theta)\mathrm{d}\pi(\theta).$$

Thus, under the squared loss function

$$L(\widehat{\theta}, \theta) = (\widehat{\theta} - \theta)^2, \tag{1}$$

the BE would be the posterior expectation of $\pi(\theta \,|\, x)$. However, the BE is somewhat complicated and inconvenient to use when the integrals appear in the posterior density are not tractable, which causes the BE has no explicit expression and thus must be evaluated numerically. Therefore, in this case we generally use simulation-based methods to obtain approximate Bayesian estimations which can escape this unpleasant situation. Lindley [7] developed asymptotic expansions for the ratios of integrals that occur in Bayesian analysis. However, calculating the third derivative of the posterior density function is often an arduous task. Also, Tierney and Kadane [8] proposed an approximation procedure to the posterior moments and marginal densities, but this approximation process needs to satisfy a main condition that the product of the likelihood and the prior is unimodal, which more or less limits its application.

This paper finds that when the prior distribution is specified as Pareto distribution, the posterior expectation has an explicit form, but integral difficulties often occur with other prior distributions specified. Therefore, a linear Bayes procedure is employed to estimate the unknown parameter, and numerical simulations show that the proposed linear approximate Bayes estimator is easy to use, which not only has an explicit form but also performs well.

The paper is organized as follows. In Section 2, we first illustrate some properties of the uniform distribution and its sufficient and complete statistic. Then, we employ the linear Bayes procedure to construct the linear approximate Bayes estimator (LABE) for the parameter. We also investigate the superiorities of the LABE over some classical estimators in terms of the mean squared error (MSE) criterion. Numerical simulations are carried out in Section 3 to make comparisons between the LABE and the BE, the latter is obtained via numerical integration and validated by the brute-force method. Also, we compare the proposed LABE with the Lindley's approximation. Section 4 is devoted to conclusions.

§2. LABE and Its Superiorities

As indicated in the previous section, by denoting $M = \max\limits_i X_i$, $m = \min\limits_i X_i$, we can get the sufficient and complete statistic $T = \max\{M, -m\}$ and its probability density function $f(t \,|\, \theta) = nt^{n-1}/\theta^n$, where $0 < t < \theta$. We omit the proof processes since they are relatively simple. Note that information in samples is all derived from the statistic T. Then we can employ it to compute the posterior density function of θ.

First, we consider the case that the prior distribution is specified as Pareto distribution, which is the conjugate prior distribution of the uniform distribution. We state the following Theorem.

Theorem 1 Let $X_1, X_2, \cdots, X_n$ be independently drawn from the distribution $R(-\theta, \theta)$ and the prior of θ is $\mathrm{Pareto}(\theta_0, \alpha)$, where $\theta_0 > 0$, $\alpha > 0$, then under the squared loss function the Bayes estimator of θ is $\widehat{\theta}_{\mathrm{BE}} = (\alpha + n)(\alpha + n - 1)^{-1}\theta_1$, where $\theta_1 = \max(t, \theta_0)$.

Proof The density function of $\mathrm{Pareto}(\theta_0, \alpha)$ is

$$\pi(\theta) = \frac{\alpha\theta_0^\alpha}{\theta^{\alpha+1}}, \qquad \theta > \theta_0.$$

Denote $\theta_1 = \max(t, \theta_0)$, together with the density function of T, we have

$$\pi(\theta \,|\, t) \propto f(t \,|\, \theta)\pi(\theta) \propto \frac{1}{\theta^{\alpha+n+1}}, \qquad \theta > \theta_1,$$

which exactly follows the distribution of Pareto$(\theta_1, \alpha + n)$.

Then, we can get the posterior density of θ is

$$\pi(\theta \,|\, t) = \frac{(\alpha + n)\theta_1^{\alpha+n}}{\theta^{\alpha+n+1}}, \qquad \theta > \theta_1.$$

Obviously, the posterior expectation is $\mathsf{E}(\theta \,|\, t) = (\alpha + n)(\alpha + n - 1)^{-1}\theta_1$.

Hence, under the loss function (1), the Bayes estimator of θ is

$$\widehat{\theta}_{\mathrm{BE}} = (\alpha + n)(\alpha + n - 1)^{-1}\theta_1.$$

We obtain the conclusion of Theorem 1. $\square$

If other prior distributions are adopted, there may arise some troubles in integration and the Bayesian estimation obtained by numerical simulations may not have an explicit expression, which imposes certain limitations on the application of Bayesian estimation.

Enlightened by Theorem 1, we employ T to construct a linear estimator of θ. Define the LABE $\widehat{\theta}_{\mathrm{LB}}$ be of the form $\widehat{\theta}_{\mathrm{LB}} = aT + b$ satisfying

$$R(\widehat{\theta}_{\mathrm{LB}}, \theta) = \min_{a,b} \mathsf{E}_{(T,\theta)}(\widehat{\theta}_{\mathrm{LB}} - \theta)^2, \tag{2}$$

$$\mathsf{E}_{(T,\theta)}(\widehat{\theta}_{\mathrm{LB}} - \theta) = 0, \tag{3}$$

where a and b are unknown scalars, $\mathsf{E}_{(T,\theta)}$ indicates that the expectation is with respect to the joint distribution of T and θ.

Assume that the prior distribution $\pi(\theta)$ of the parameter θ belongs to the following prior family:

$$\zeta = \{\pi(\theta) : \mathsf{E}\theta^2 < \infty\}. \tag{4}$$

Thus we can get the following Theorem.

Theorem 2 Under the prior assumption (4), the expression of $\widehat{\theta}_{\mathrm{LB}}$ satisfying the conditions (3) and (2) is given as follows

$$\widehat{\theta}_{\mathrm{LB}} = c\mathsf{Var}\,(\theta)MT + \mathsf{E}\theta - c^2\mathsf{Var}\,(\theta)M\mathsf{E}\theta, \tag{5}$$

where $\mathsf{E}(T \,|\, \theta) = c\theta$, $W = \mathsf{E}_\theta[\mathsf{Var}\,(T \,|\, \theta)]$, and $M = [W + c^2\mathsf{Var}\,(\theta)]^{-1}$ with $c = n/(n + 1)$.

Proof From $\mathsf{E}_{(T,\theta)}(aT+b-\theta)=0$, we know

$$b=\mathsf{E}_{(T,\theta)}\theta-a\mathsf{E}_{(T,\theta)}T=\mathsf{E}\theta-a\mathsf{E}[\mathsf{E}(T\,|\,\theta)]=\mathsf{E}\theta-ac\mathsf{E}\theta.$$

Hence, $R(\widehat{\theta},\theta)$ can be written as

$$\begin{aligned}R(\widehat{\theta},\theta)&=\mathsf{E}_{(T,\theta)}(aT+b-\theta)^2\\&=\mathsf{E}_{(T,\theta)}(aT+\mathsf{E}\theta-ac\mathsf{E}\theta-\theta)^2\\&=\mathsf{E}_{(T,\theta)}[a(T-c\mathsf{E}\theta)+(\mathsf{E}\theta-\theta)]^2\\&=a^2\mathsf{E}_{(T,\theta)}(T-c\mathsf{E}\theta)^2+\mathsf{E}_{(T,\theta)}(\mathsf{E}\theta-\theta)^2+2a\mathsf{E}_{(T,\theta)}(T-c\mathsf{E}\theta)(\mathsf{E}\theta-\theta)\\&=a^2\mathsf{E}_\theta\mathsf{E}_{T|\theta}(T-c\theta+c\theta-c\mathsf{E}\theta)^2+\mathsf{Var}\,(\theta)+2a\mathsf{E}_\theta\mathsf{E}_{T|\theta}(T-c\mathsf{E}\theta)(\mathsf{E}\theta-\theta)\\&=a^2\mathsf{E}_\theta\mathsf{Var}\,(T\,|\,\theta)+a^2c^2\mathsf{Var}\,(\theta)+\mathsf{Var}\,(\theta)-2ac\mathsf{Var}\,(\theta).\end{aligned}$$

Let $\partial R(\widehat{\theta},\theta)/\partial a=0$, then

$$\frac{\partial R(\widehat{\theta},\theta)}{\partial a}=2a\mathsf{E}_\theta\mathsf{Var}\,(T\,|\,\theta)+2ac^2\mathsf{Var}\,(\theta)-2c\mathsf{Var}\,(\theta)=0,$$

which yields

$$a=c\mathsf{Var}\,(\theta)[\mathsf{E}_\theta\mathsf{Var}\,(T\,|\,\theta)+c^2\mathsf{Var}\,(\theta)]^{-1}=c\mathsf{Var}\,(\theta)M.$$

Together with

$$b=\mathsf{E}\theta-c^2\mathsf{Var}\,(\theta)M\mathsf{E}\theta,$$

we come to the conclusion of Theorem 2. The proof Theorem 2 is complete. □

Remark 3 There are some unknown elements in the equation (5) related to the prior distribution of θ, such as $\mathsf{E}\theta$, $\mathsf{Var}\,(\theta)$ and M, which can be figured out for a specific prior distribution.

In what follows, we discuss the superiorities of the proposed LABE over some classical estimators. Note that the UMVUE (the uniformly minimum variance unbiased estimator) of the parameter θ is $[(n+1)/n]T$, which can be represented as

$$\widehat{\theta}_{\mathrm{U}}=\frac{n+1}{n}T=c^{-1}T. \tag{6}$$

And it is easy to see that the MLE (maximum likelihood estimator) of θ is T, that is

$$\widehat{\theta}_{\mathrm{ML}}=T.$$

Theorem 4 Let $\widehat{\theta}_{\mathrm{LB}}$ and $\widehat{\theta}_{\mathrm{U}}$ be given by (5) and (6) respectively, then $\widehat{\theta}_{\mathrm{LB}}$ is superior to $\widehat{\theta}_{\mathrm{U}}$ in terms of MSE criterion, i.e. $\mathsf{MSE}\,(\widehat{\theta}_{\mathrm{LB}})\leqslant\mathsf{MSE}\,(\widehat{\theta}_{\mathrm{U}})$.

Proof Note that the MSE of the $\widehat{\theta}_{\rm U}$ is

$$\begin{aligned}\mathrm{MSE}\,(\widehat{\theta}_{\rm U}) &= \mathsf{E}_{(T,\theta)}(\widehat{\theta}_{\rm U}-\theta)^2 = \mathsf{E}_\theta\mathsf{E}_{T|\theta}(c^{-1}T-\theta)^2\\ &= \mathsf{E}_\theta\{\mathsf{Var}_{T|\theta}(c^{-1}T)+[\mathsf{E}_{T|\theta}(c^{-1}T-\theta)]^2\}\\ &= c^{-2}\mathsf{E}_\theta\mathsf{Var}\,(T\mid\theta)=c^{-2}W.\end{aligned}\tag{7}$$

On the other hand, it is readily seen that

$$\begin{aligned}\mathrm{MSE}\,(\widehat{\theta}_{\rm LB}) &= \mathsf{E}_{(T,\theta)}(\widehat{\theta}_{\rm LB}-\theta)^2 = \mathsf{Var}\,[\mathsf{E}(\widehat{\theta}_{\rm LB}-\theta|\theta)]+\mathsf{E}[\mathsf{Var}\,(\widehat{\theta}_{\rm LB}-\theta|\theta)]\\ &= \mathsf{Var}\,[c^2\mathsf{Var}\,(\theta)M\theta+\mathsf{E}\theta-c^2\mathsf{Var}\,(\theta)M\mathsf{E}\theta-\theta]\\ &\quad+\mathsf{E}\{\mathsf{Var}\,[c\mathsf{Var}\,(\theta)MT+\mathsf{E}\theta-c^2\mathsf{Var}\,(\theta)M\mathsf{E}\theta]\}\\ &= \mathsf{Var}\,[c^2\mathsf{Var}\,(\theta)M(\theta-\mathsf{E}\theta)-(\theta-\mathsf{E}\theta)]+\mathsf{E}\{\mathsf{Var}\,[c\mathsf{Var}\,(\theta)MT]\}\\ &= \mathsf{Var}\,\{[c^2\mathsf{Var}\,(\theta)M-1](\theta-\mathsf{E}\theta)\}+[c\mathsf{Var}\,(\theta)M]^2\mathsf{E}_\theta\mathsf{Var}\,(T\mid\theta)\\ &= [c^2\mathsf{Var}\,(\theta)M-1]^2\mathsf{Var}\,(\theta)+[c\mathsf{Var}\,(\theta)M]^2W\\ &= [c\mathsf{Var}\,(\theta)M]^2[c^2\mathsf{Var}\,(\theta)+W]-2[c\mathsf{Var}\,(\theta)]^2M+\mathsf{Var}\,(\theta)\\ &= [c\mathsf{Var}\,(\theta)]^2M-2[c\mathsf{Var}\,(\theta)]^2M+\mathsf{Var}\,(\theta)\\ &= -c^{-2}(M^{-1}-W)^2M+c^{-2}(M^{-1}-W)\\ &= c^{-2}W-c^{-2}W^2M.\end{aligned}\tag{8}$$

Comparing (7) with (8), we conclude that $\mathrm{MSE}\,(\widehat{\theta}_{\rm LB})\leqslant\mathrm{MSE}\,(\widehat{\theta}_{\rm U})$. The proof of Theorem 4 is complete. □

Similarly, mimicking the proof processes of the Theorem 4, we can prove that $\widehat{\theta}_{\rm LB}$ is superior to $\widehat{\theta}_{\rm ML}$. Hence, we can get the conclusion that $\mathrm{MSE}\,(\widehat{\theta}_{\rm LB})$ is less than both $\mathrm{MSE}\,(\widehat{\theta}_{\rm ML})$ and $\mathrm{MSE}\,(\widehat{\theta}_{\rm U})$, which indicates the superiorities of the proposed LABE over the classical estimators under the MSE criterion.

§3. Numerical Comparisons

At the beginning of this section, we present the numerical comparisons between the $\widehat{\theta}_{\rm LB}$ and $\widehat{\theta}_{\rm BE}$ by $|\widehat{\theta}_{\rm LB}-\widehat{\theta}_{\rm BE}|$ under the specific prior distribution Pareto(α,θ_0) in Table 1, in this case the ordinary Bayes estimator has an explicit form. It should be noted that those simulation results are based on 1 000 replications and the true value of θ is specified as 5. Also, we compute the MSE of the LABE to assess the accuracy of the linear Bayes method, which is defined as

$$\mathrm{MSE}\,(\widehat{\theta}_{\rm LB})=\frac{1}{1\,000}\sum_{i=1}^{1\,000}(\widehat{\theta}_{{\rm LB},i}-\theta)^2.$$

Moreover, we use the Lindley's approximation as a contrast. We will give the expression of the Lindley's approximation $\widehat{\theta}_{\text{Lindley}}$ later.

Table 1 $\theta \sim \text{Pareto}(\alpha, \theta_0)$

n	α	θ_0	$\lvert\widehat{\theta}_{\text{LB}} - \widehat{\theta}_{\text{BE}}\rvert$	$\lvert\widehat{\theta}_{\text{Lindley}} - \widehat{\theta}_{\text{BE}}\rvert$	MSE ($\widehat{\theta}_{\text{LB}}$)
30	5	3	0	0.6645	4.526×10^{-6}
	3	3	0	0.3328	5.267×10^{-7}
	3	4	0	0.3325	1.653×10^{-6}
50	5	3	0	0.3995	8.019×10^{-7}
	3	3	0	0.1999	1.833×10^{-6}
	3	4	0	0.1998	3.481×10^{-7}
100	5	3	0	0.1999	6.017×10^{-6}
	3	3	0	0.1000	1.399×10^{-6}
	3	4	0	0.0999	1.960×10^{-6}

It is easily seen from Table 1 that $\widehat{\theta}_{\text{LB}}$ is equal to the $\widehat{\theta}_{\text{BE}}$, which means in this situation the BE is the linear form of the statistic T and $\widehat{\theta}_{\text{LB}}$ is better than the Lindley's approximation obviously. Also, we find that $\widehat{\theta}_{\text{Lindley}}$ also approximates well and tends to be closer to the $\widehat{\theta}_{BE}$ with the increasing of sample size.

However, $\widehat{\theta}_{\text{BE}}$ may have no explicit form when we select other prior distributions. In what follows, for some other priors, we make some numerical comparisons between the proposed LABE, the BE obtained by numerical integration and the Lindley's approximation method as well. The prior distributions are listed in Table 2, where we use beta and gamma prior distributions and assign two different values to the hyperparameters. They obviously belong to the prior family (4). Moreover, the variance of the prior distributions is used to measure the variation of the prior information, that is, the smaller variance represents the more prior information.

Table 2 Priors of θ

Prior distributions	Variance of the prior
Pr1: $\theta \sim \text{B}(2, 2)$	0.05
Pr2: $\theta \sim \text{B}(10, 10)$	0.0119
Pr3: $\theta \sim \Gamma(5, 2)$	1.25
Pr4: $\theta \sim \Gamma(5, 10)$	0.05

In order to solve the complex integral problem appears in the expectation of the posterior distribution, Lindley [7] proposed a kind of approximate method, which has been used in many Bayes computations. For the parameter θ, the Lindley's approximation $\widehat{\theta}_{\text{Lindley}}$ is defined as follows:

$$\widehat{\theta}_{\text{Lindley}} = \widehat{\theta}_{\text{ML}} + \rho_1\sigma^2 + \frac{1}{2}L_3\sigma^4,$$

where $\rho_1 = \mathrm{d}\ln\pi(\theta)/\mathrm{d}\theta$ with $\pi(\theta)$ being a specific prior distribution, $L_2 = \mathrm{d}\ln^2 f(t\,|\,\theta)/\mathrm{d}\theta^2$, $L_3 = \mathrm{d}\ln^3 f(t\,|\,\theta)/\mathrm{d}\theta^3$, $\sigma^2 = -L_2^{-1}$ and $\sigma^4 = (\sigma^2)^2$, where $\ln f(t\,|\,\theta)$ is the log-likelihood function.

According to the prior distributions of θ and computing the above expressions, we can obtain the Lindley's approximation $\widehat{\theta}_{\text{Lindley}}$. Furthermore, we also exhibit the distances $|\widehat{\theta}_{\text{BE}} - \theta|$ and $|\widehat{\theta}_{\text{LB}} - \theta|$, from which we can see the differences between various of estimators and the true values of θ. Here we adopt two different prior distributions, the true values of θ are specified as 0.5 and 1, respectively.

Table 3 $\theta \sim \mathrm{B}(a, b)$

n	Prior	$\lvert\widehat{\theta}_{\text{LB}} - \widehat{\theta}_{\text{BE}}\rvert$	$\lvert\widehat{\theta}_{\text{Lindley}} - \widehat{\theta}_{\text{BE}}\rvert$	$\lvert\widehat{\theta}_{\text{BE}} - \theta\rvert$	$\lvert\widehat{\theta}_{\text{LB}} - \theta\rvert$	MSE $(\widehat{\theta}_{\text{LB}})$
30	Pr1	1.100×10^{-3}	3.431×10^{-2}	5.104×10^{-5}	5.181×10^{-5}	2.685×10^{-6}
	Pr2	7.047×10^{-4}	4.162×10^{-2}	5.487×10^{-5}	5.534×10^{-5}	3.063×10^{-6}
50	Pr1	3.977×10^{-4}	2.037×10^{-2}	8.256×10^{-6}	7.840×10^{-6}	6.147×10^{-8}
	Pr2	3.201×10^{-4}	2.312×10^{-2}	1.370×10^{-7}	1.847×10^{-7}	3.408×10^{-11}
100	Pr1	9.990×10^{-5}	1.010×10^{-2}	4.130×10^{-6}	4.028×10^{-6}	1.623×10^{-8}
	Pr2	9.088×10^{-5}	1.083×10^{-2}	7.249×10^{-7}	7.336×10^{-7}	1.530×10^{-8}

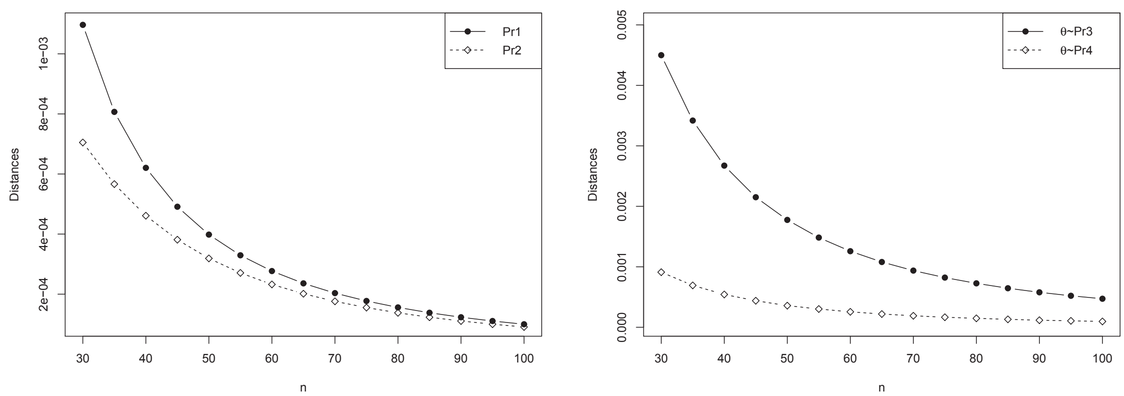
Table 4 $\theta \sim \Gamma(\alpha, \lambda)$

n	Prior	$\lvert\widehat{\theta}_{\text{LB}} - \widehat{\theta}_{\text{BE}}\rvert$	$\lvert\widehat{\theta}_{\text{Lindley}} - \widehat{\theta}_{\text{BE}}\rvert$	$\lvert\widehat{\theta}_{\text{BE}} - \theta\rvert$	$\lvert\widehat{\theta}_{\text{LB}} - \theta\rvert$	MSE $(\widehat{\theta}_{\text{LB}})$
30	Pr3	4.486×10^{-3}	1.358×10^{-1}	9.058×10^{-6}	1.351×10^{-5}	1.824×10^{-7}
	Pr4	9.129×10^{-4}	1.234×10^{-1}	1.750×10^{-5}	1.854×10^{-5}	3.439×10^{-7}
50	Pr3	1.773×10^{-3}	8.084×10^{-2}	1.483×10^{-5}	1.656×10^{-5}	2.743×10^{-7}
	Pr4	3.591×10^{-4}	7.611×10^{-2}	1.660×10^{-5}	1.701×10^{-5}	2.893×10^{-7}
100	Pr3	4.718×10^{-4}	4.021×10^{-2}	7.197×10^{-6}	7.663×10^{-6}	5.873×10^{-8}
	Pr4	9.590×10^{-5}	3.899×10^{-2}	1.297×10^{-5}	1.289×10^{-5}	1.661×10^{-7}

From Table 3 and Table 4, we can see that the distance $|\widehat{\theta}_{\text{LB}} - \widehat{\theta}_{\text{BE}}|$ tends to decrease as the prior information gets more concentrated for the two kinds of prior distributions, respectively, which indicates that a more informative prior may lead to the LABE $\widehat{\theta}_{\text{LB}}$ be closer to the BE $\widehat{\theta}_{\rm BE}$. In addition, it is readily seen that when the prior distribution of θ is changed from the beta distributions to the gamma distributions, the LABE $\widehat{\theta}_{\rm LB}$ has a little change, but it is not so sensitive to the choice of the prior distributions, which reflects the robustness of the linear Bayes estimator.

Moreover, we easily see that all $|\widehat{\theta}_{\rm LB}-\widehat{\theta}_{\rm BE}|$s are relatively small and $|\widehat{\theta}_{\rm LB}-\widehat{\theta}_{\rm BE}|\leqslant|\widehat{\theta}_{\rm Lindley}-\widehat{\theta}_{\rm BE}|$ can be found from all the numerical comparison results, which implies that as an approximation to the BE the LABE outperforms the Lindley's approximation for the above cases. When comparing the distances $|\widehat{\theta}_{\rm LB}-\theta|$ with $|\widehat{\theta}_{\rm BE}-\theta|$, we can find that the LABE is close to the true value of θ and sometimes even performs better than the BE. Note that the proposed LABE possesses an explicit form and is easy to use. Hence, it is suitable to employ the linear Bayes method to estimate θ.

Furthermore, for the two kinds of prior distributions, we display the distances $|\widehat{\theta}_{\rm LB}-\widehat{\theta}_{\rm BE}|$ varying with the sample size n in Figure 1 (a) and (b), respectively. From Figure 1, we find that the distance $|\widehat{\theta}_{\rm LB}-\widehat{\theta}_{\rm BE}|$ has a decreasing trend with the increasing of sample size, and the choice of the prior distribution has little effect on the approximate performance. Also, we can find that a more informative prior distribution helps obtain a closer LABE to the BE.



(a) (b)

Figure 1 The distances $|\widehat{\theta}_{\rm LB}-\widehat{\theta}_{\rm BE}|$ vary with the sample size under the beta prior distributions in (a) and under the gamma prior distributions in (b)

Finally, in order to further illustrate the robustness of the linear Bayes estimator to the choice of the prior distributions, we use the inverse gamma distribution as prior. For $n=50$ and $\theta=1.5$, we compute $|\widehat{\theta}_{\rm LB}-\widehat{\theta}_{\rm BE}|$, $|\widehat{\theta}_{\rm Lindley}-\widehat{\theta}_{\rm BE}|$, $|\widehat{\theta}_{\rm LB}-\theta|$, $|\widehat{\theta}_{\rm Lindley}-\theta|$ and the results are shown in Table 5 and Table 6. From Table 6, we see that $|\widehat{\theta}_{\rm LB}-\widehat{\theta}_{\rm BE}|$ and $|\widehat{\theta}_{\rm LB}-\theta|$ are both smaller, which means the LABE is both closer to the $\widehat{\theta}_{\rm BE}$ and the true value of θ when compared with the Lindley's approximation $\widehat{\theta}_{\text{Lindley}}$.

Table 5 $\theta \sim \Gamma^{-1}(\alpha, \lambda)$

Prior	$\widehat{\theta}_{\text{BE}}$	$\widehat{\theta}_{\text{LB}}$	$\widehat{\theta}_{\text{Lindley}}$
Pr5: $\theta \sim \Gamma^{-1}(5, 20)$	1.5066	1.5056	1.2179

Table 6 Distances under the Pr5

Prior	$\lvert\widehat{\theta}_{\text{LB}} - \widehat{\theta}_{\text{BE}}\rvert$	$\lvert\widehat{\theta}_{\text{Lindley}} - \widehat{\theta}_{\text{BE}}\rvert$	$\lvert\widehat{\theta}_{\text{BE}} - \theta\rvert$	$\lvert\widehat{\theta}_{\text{LB}} - \theta\rvert$	MSE $(\widehat{\theta}_{\text{LB}})$
Pr5	0.0011	0.2888	0.0056	0.2821	1.212×10^{-7}

Also, we present the frequency histograms of the four distances $|\widehat{\theta}_{\text{LB}} - \widehat{\theta}_{\text{BE}}|$, $|\widehat{\theta}_{\text{Lindley}} - \widehat{\theta}_{\text{BE}}|$, $|\widehat{\theta}_{\text{LB}} - \theta|$ and $|\widehat{\theta}_{\text{Lindley}} - \theta|$ in Figure 2 under the inverse gamma distribution $\Gamma^{-1}(5, 20)$. They are obtained by repeating the simulations 1 000 times, which further show that the linear Bayes procedure performs well.

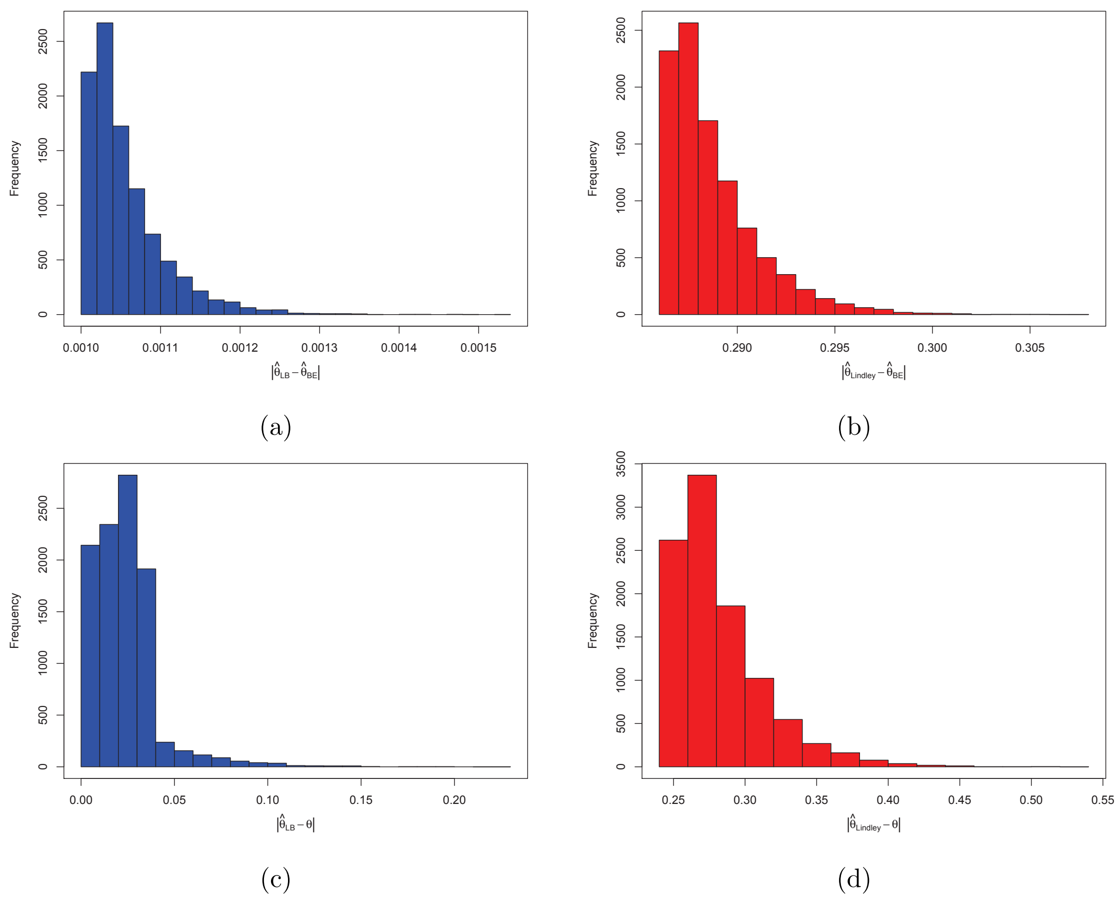
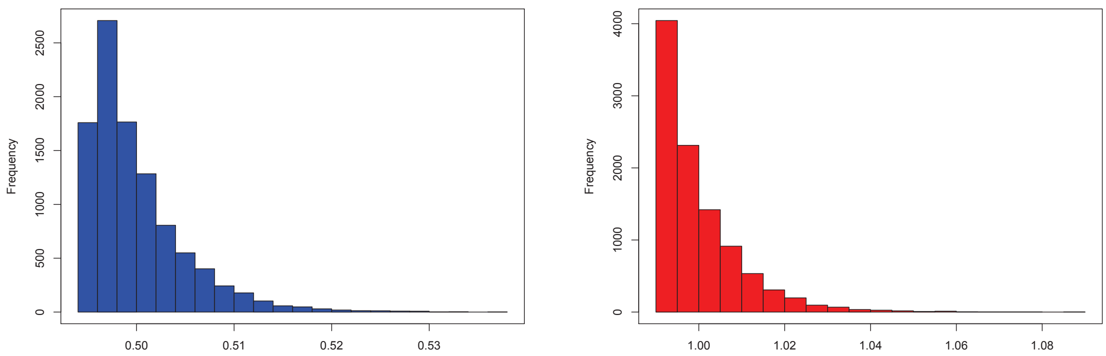


Figure 2 The histograms of the distances for $n = 50$ and $\theta = 1.5$ under the Pr5

Besides the numerical integration method to obtain the BE, there are some other methods, such as the brute-force method (see [9], etc.), the simulation-based MCMC methods (see [10,11], etc.) including Metropolis-Hastings procedure and Gibbs procedure, which are suggested to obtain the approximate Bayes estimator. Here, we produce the simulation results by the brute-force method under the Pr1 and the Pr3 and present them in Figure 3.



(a) (b)

Figure 3 Histograms of the samples from the posterior distributions under the Pr1 in (a) and under the Pr3 in (b)

For the prior distributions Pr1 and Pr3, it is easily seen that the samples from the posterior distributions are concentrated at the value 0.5 and 1 (the true values of θ that we specified) respectively, which means that the outcomes are very similar to the numerical integration method.

§4. Conclusions

This article employs the linear Bayes method to estimate the parameter θ of the uniform distribution $R(-\theta, \theta)$ and proposes a linear approximate Bayes estimator (LABE). Under several different prior distributions, we compare the LABE with the BE and the Lindley's approximation numerically. Moreover, the superiorities of the LABE over some classical estimators in terms of the MSE criterion are also exhibited.

Compared with the BE and the Lindley's approximation, we find that the proposed LABE is not only simple and easy to calculate but also is a good approximation, and in numerical simulations the LABE is very close to the BE regardless of the choice of priors and the change of the hyperparameters. Also, we find that the LABE outperforms the Lindley's approximation for our cases, which furthermore reveals the efficiency of the proposed LABE. The procedure used in this article can be extended easily to many other useful distributions such as normal, log-normal, exponential family, etc.

References

[1] CHEN X R, NI G X. *Course of Mathematical Statistics* [M]. Hefei: Publishing House of USTC, 2009. (in Chinese)

[2] HARTIGAN J A. Linear Bayesian Methods [J]. *J Roy Statist Soc Ser B*, 1969, **31(3)**: 446–454.

[3] RAO C R. *Linear Statistical Inference and Its Applications* [M]. 2nd ed. New York: Wiley, 1973.

[4] WEI L S, ZHANG W P. The superiorities of Bayes linear minimum risk estimation in linear model [J]. *Comm Statist Theory Methods*, 2007, **36(5-8)**: 917–926.

[5] WANG L C, SINGH R S. Linear Bayes estimator for the two-parameter exponential family under type II censoring [J]. *Comput Statist Data Anal*, 2014, **71**: 633–642.

[6] KUO L. Linear Bayes estimators of the potency curve in bioassay [J]. *Biometrika*, 1988, **75(1)**: 91–96.

[7] LINDLEY D V. Approximate Bayesian methods [J]. *Trabajos Estadíst Investigación Oper*, 1980, **31(1)**: 223–245.

[8] TIERNEY L, KADANE J B. Accurate approximations for posterior moments and marginal densities [J]. *J Amer Statist Assoc*, 1986, **81(393)**: 82–86.

[9] ALBERT J. *Bayesian Computation with R* [M]. New York: Springer-Verlag, 2008.

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

[11] RIZZO M L. *Statistical Computing with R* [M]. New York: Chapman and Hall/CRC, 2008.

[12] WEI L S, ZHANG W P. *Bayesian Analysis* [M]. Hefei: Publishing House of USTC, 2013. (in Chinese)

均匀分布的线性近似贝叶斯估计

张凤月　　王立春

(北京交通大学理学院, 北京, 100044)

摘　要: 本文利用线性贝叶斯方法估计均匀分布 $R(-\theta,\theta)$ 的未知参数, 提出了 θ 的线性近似贝叶斯估计 (LABE), LABE 具有封闭解析解的形式且便于使用. 数值模拟表明本文提出的 LABE 与普通的贝叶斯估计 (BE) 很接近, 其中 BE 由数值积分得到, 我们也使用了所谓的强力算法来获得 BE. 进一步, 我们比较了 LABE 与 Lindley 近似. 在均方误差准则下, LABE 相对于经典估计量的优越性也得到证明.

关键词: 均匀分布; 线性近似贝叶斯估计; Lindley 近似; 均方误差 (MSE) 准则

中图分类号: O212.8