

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Publications, Department of Statistics

Statistics, Department of

10-2-2022

Bayesian analysis for the Lomax model using noninformative priors

Daojiang He

Dongchu Sun

Qing Zhu

Follow this and additional works at: <https://digitalcommons.unl.edu/statisticsfacpub>



Part of the [Other Statistics and Probability Commons](#)

This Article is brought to you for free and open access by the Statistics, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, Department of Statistics by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Bayesian analysis for the Lomax model using noninformative priors

Daojiang He^a, Dongchu Sun^b and Qing Zhu^a

^aDepartment of Statistics, Anhui Normal University, Wuhu, People's Republic of China; ^bDepartment of Statistics, University of Nebraska-Lincoln, Lincoln, NE, USA

ABSTRACT

The Lomax distribution is an important member in the distribution family. In this paper, we systematically develop an objective Bayesian analysis of data from a Lomax distribution. Non-informative priors, including probability matching priors, the maximal data information (MDI) prior, Jeffreys prior and reference priors, are derived. The propriety of the posterior under each prior is subsequently validated. It is revealed that the MDI prior and one of the reference priors yield improper posteriors, and the other reference prior is a second-order probability matching prior. A simulation study is conducted to assess the frequentist performance of the proposed Bayesian approach. Finally, this approach along with the bootstrap method is applied to a real data set.

ARTICLE HISTORY

Received 9 November 2021
Revised 6 August 2022
Accepted 2 October 2022

KEYWORDS

Lomax model; probability matching priors; MDI prior; Jeffreys prior; reference priors; posterior propriety

2000 MATHEMATICS

SUBJECT CLASSIFICATIONS
62F15; 62N05

1. Motivation

A random variable is said to be distributed as a Lomax distribution if its density function has the following form:

$$f(x; \alpha, \beta) = \frac{\alpha}{\beta} \left(1 + \frac{x}{\beta}\right)^{-(\alpha+1)}, \quad x > 0, \quad (1)$$

where $\alpha > 0$ is the shape parameter and $\beta > 0$ is the scale parameter. The distribution is originally introduced in Lomax (1954) for the analysis of business failure data. Since then, the Lomax model has been widely applied in many other fields. For example, Atkinson and Harrison (1978) utilized the Lomax distribution to model personal wealth data; Bain and Engelhardt (1992) found that the Lomax distribution provided a good model for biomedical problems, such as survival time following a heart transplant; Holland et al. (2006) applied the Lomax distribution to model the distribution of the sizes of computer files on servers; and Marshall and Olkin (2007) showed that the Lomax distribution can be applied as a lifetime distribution. For some extensions of the Lomax distribution, one is referred to Nayak (1987), Roy and Gupta (1996), Nadarajah (2005), Lemonte and Cordeiro (2013), Kang et al. (2021) among others.

For the Lomax model (1), there is no closed-form expression for the classical maximum likelihood estimator (MLE). More importantly, as pointed out by Deville (2016), the MLE does not exist if the sample coefficient of variation $CV_n < 1$, which was also analyzed in detail in Chakraborty (2019). In addition, a simulation study indicates that the probability of $CV_n < 1$ is not negligible. For example, when $(\alpha, \beta) = (3, 1)$ and $n = 20$, the empirical probability of $CV_n < 1$ is as high as 0.25. In this sense, the MLE method is not applicable under these cases.

Thus, it is natural to consider Bayesian estimation for the parameters (α, β) in model (1). In a Bayesian paradigm, the specification of a prior distribution is one of the most important problems. We first consider the following vague prior for $\theta = (\alpha, \beta)$, that is,

$$\pi_G(\theta) \propto \frac{1}{\beta} \alpha^{\tau-1} e^{-\alpha}, \quad (2)$$

where $\tau > 0$ is a hyperparameter. That is to say, the 'marginal' prior for α is a gamma distribution. It is shown that the posterior distribution of $\pi_G(\theta | \mathbf{X})$ is proper for any $n \geq 1$ (see the Appendix for proofs), where $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is the sample. To assess the sensitivity of the corresponding Bayesian estimation with respect to the hyperparameter τ , we conduct a sensitivity analysis here.

The values of (α, β) are set as (2, 1.5), and τ is set as 1, 5 and 9, respectively. The empirical square root of the mean squared error ($\sqrt{\text{MSE}}$) and coverage probability (CP) of the Bayesian estimation with respect to the sample

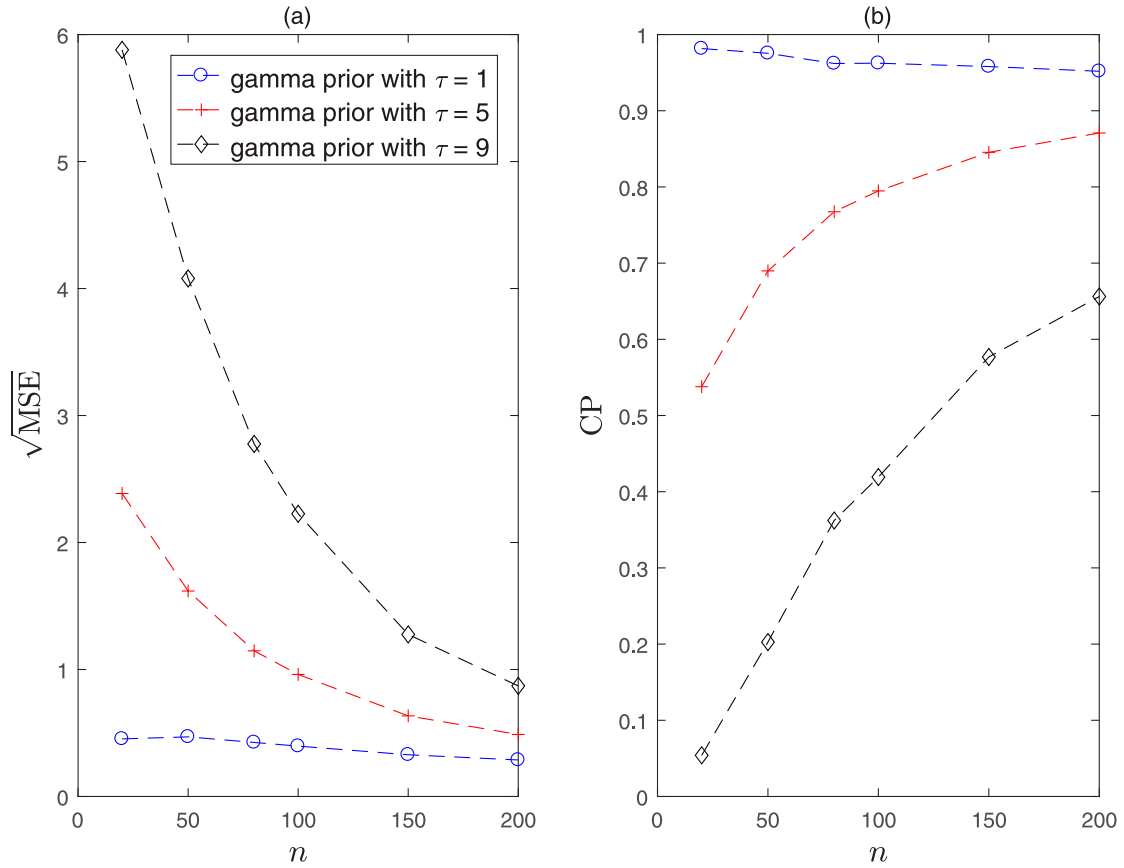


Figure 1. Square root of the mean squared error and coverage probability of the Bayesian estimation of α based on the gamma priors π_G with $\tau = 1$ (circle), $\tau = 5$ (cross), and $\tau = 9$ (diamond). Panel (a) is for the square root of mean squared error, and panel (b) for is the coverage probability.

size n are shown in Figure 1. It can be observed that the performance of the MSE and CP based on the prior π_G is very sensitive to the choice of the hyperparameter τ , which makes it difficult to specify π_G in applications. Obviously, if one is interested in making an optimal decision based on his/her beliefs, such a (subjective) prior could be appropriate.

Based on the above motivation, we propose an objective Bayesian analysis for the Lomax model in this paper, particularly when there is little prior information on the parameters. As we know, one of the most appealing features of the objective Bayesian analysis is to use noninformative priors. In the work of Ferreira et al. (2016, 2020), Jeffreys prior and independent Jeffreys prior (i.e., one of the reference priors) were considered. In this paper, we propose a more systematic and deeper analysis based on an extensive class of objective priors including probability matching priors, the maximal data information (MDI) prior, Jeffreys prior and reference priors.

The remainder of this paper is organized as follows. In Section 2, noninformative priors, including probability matching priors, the MDI prior, Jeffreys prior and reference priors are derived. Moreover, the posterior propriety under each prior is validated. In Section 3, a simulation study is conducted to evaluate the frequentist properties of Bayesian estimates based on the noninformative priors. In Section 4, the proposed Bayesian approach is applied to analyze a real data set. Some concluding remarks are given in Section 5.

2. Noninformative priors and their properties

In this section, we derive some important noninformative priors for the parameters (α, β) , which contain probability matching priors, the MDI prior, Jeffreys prior and reference priors.

2.1. Probability matching priors

The rationale behind a probability matching prior is that a noninformative prior should provide inferences that are similar to those obtained from a frequentist perspective, such as in terms of credible versus confidence intervals. In this perspective, a probability matching prior is a prior such that the posterior coverage probability of Bayesian credible interval matches the corresponding frequentist coverage probability (Consonni et al., 2018).

Given a prior $\pi(\cdot)$ for the parameters (ϕ, φ) , suppose that ϕ is the parameter of interest, and $\phi^{(1-\gamma)}(\pi(\cdot), \mathbf{X})$ is the $(1 - \gamma)$ th percentile of the marginal posterior distribution of ϕ . Then, $\pi(\cdot)$ is called a second-order probability matching prior if

$$P\left(\phi \leq \phi^{(1-\gamma)}(\pi(\cdot), \mathbf{X})\right) = 1 - \gamma + o(n^{-1})$$

holds for all $\gamma \in (0, 1)$, see Datta and Mukerjee (2004) for more details.

For the parameters α and β in the Lomax model (1), we have the following theorem.

Theorem 2.1: (a) *When α is the parameter of interest and β is the nuisance parameter, the second-order probability matching prior has the form of*

$$\pi_{M_1}(\boldsymbol{\theta}) \propto F_1(\alpha) \cdot G_1(\beta), \quad (3)$$

where $F_1(\alpha) \propto \exp\{-\int \frac{(c_1+3)\alpha+2c_1+3}{\alpha(\alpha+1)} d\alpha\}$, $G_1(\beta) \propto \beta^{c_1}$, and c_1 is an arbitrary constant.

(b) *When β is the parameter of interest and α is the nuisance parameter, the second-order probability matching prior is given by*

$$\pi_{M_2}(\boldsymbol{\theta}) \propto F_2(\alpha) \cdot G_2(\beta), \quad (4)$$

where $F_2(\alpha) \propto \exp\{-\int \frac{(c_2+3)\alpha^2+3(c_2+2)\alpha+2c_2+2}{\alpha^2(\alpha+2)} d\alpha\}$, $G_2(\beta) \propto \beta^{c_2}$, and c_2 is an arbitrary constant.

2.2. The MDI prior

Lindley (1956) applied the Shannon entropy to develop an information theoretic analysis of the structure of Bayesian modelling. This prompted the works on the definition of the least informative prior distribution based on some definitions of the amount of information. Zellner (1977) proposed an important noninformative prior, which is called the MDI prior. Zellner (1977) proved that using this prior could emphasize the information in the likelihood function. Therefore, the information in the prior is weak compared with that in the data (Ramos et al., 2018).

For the Lomax model (1), we have the following result, the proofs of which are deferred to the Appendix.

Theorem 2.2: (a) *The MDI prior for the parameters $\boldsymbol{\theta} = (\alpha, \beta)$ is given by*

$$\pi_M(\boldsymbol{\theta}) \propto \frac{\alpha}{\beta e^{\frac{1}{\alpha}}}. \quad (5)$$

(b) *For any $n \geq 1$, the posterior distribution under $\pi_M(\boldsymbol{\theta})$ is improper.*

2.3. Jeffreys prior

Jeffreys prior is probably the most popular noninformative prior method among practitioners. According to Jeffreys (1961), Jeffreys prior is proportional to the square root of the determinant of the Fisher information matrix. Besides being parametrization invariant, Jeffreys prior enjoys many optimality properties in the absence of nuisance parameters. It maximizes the asymptotic divergence between the prior and the posterior under several different metrics. However, Jeffreys prior also has some potential drawbacks. Particularly, in the multidimensional case, its use may lead to incoherence and paradoxes. See Consonni et al. (2018) for more discussions.

For the Lomax model (1), Jeffreys prior for the parameters $\boldsymbol{\theta} = (\alpha, \beta)$ has the following form:

$$\pi_J(\boldsymbol{\theta}) \propto \frac{1}{\beta(\alpha+1)\sqrt{\alpha(\alpha+2)}}. \quad (6)$$

And it was shown in Ferreira et al. (2020) that, for any $n \geq 1$, the posterior distribution under $\pi_J(\boldsymbol{\theta})$ is proper.

By Theorem 2.1, we have

Theorem 2.3: *Regardless of whether α is the parameter of interest or β is the parameter of interest, $\pi_J(\boldsymbol{\theta})$ is always not a second-order probability matching prior.*

2.4. Reference priors

Reference analysis uses information-theoretical concepts to precisely define the objective prior, which should be maximally dominated by the data, in the sense of maximizing the missing information on the parameters (Berger et al., 2009). The original formulation of reference priors was introduced in Bernardo (1979), which was largely informal. Berger and Bernardo (1992) gave more precise definitions of the sequential reference process in continuous multiparameter problems. In addition, a rigorous general definition of reference priors was formally given in Berger et al. (2009) for one block of parameters.

As we know, reference priors separate the parameters into different ordering groups of interest. For the ordering group $\{\beta, \alpha\}$, it was shown in Ferreira et al. (2020) that the reference prior is

$$\pi_{R_1}(\boldsymbol{\theta}) \propto \frac{1}{\alpha\beta}. \quad (7)$$

Furthermore, for any $n \geq 1$, the posterior distribution under the reference prior $\pi_{R_1}(\boldsymbol{\theta})$ is improper.

For the ordering group (α, β) , we have

Theorem 2.4: (a) *The reference prior under the ordering group $\{\alpha, \beta\}$ is given by*

$$\pi_{R_2}(\boldsymbol{\theta}) \propto \frac{1}{\beta\alpha(\alpha + 1)}. \quad (8)$$

- (b) *For $n = 1$, the posterior distribution under the reference prior $\pi_{R_2}(\boldsymbol{\theta})$ is improper; while for $n \geq 2$, the posterior under $\pi_{R_2}(\boldsymbol{\theta})$ is proper.*
 (c) *The prior $\pi_{R_2}(\boldsymbol{\theta})$ is a second-order probability matching prior while $\pi_{R_1}(\boldsymbol{\theta})$ is not.*

The proofs of Theorem 2.4 are also deferred to the Appendix. It follows from Theorems 2.2–2.4 that that only Jeffreys prior π_J and the reference prior π_{R_2} enable posterior inferences. However, π_{R_2} is a second-order probability matching prior while π_J is not. In this sense, π_{R_2} is recommended for potential users. In fact, this is also verified in the following numerical studies.

3. Simulation study

To evaluate the frequentist performance of the Bayesian estimation based on π_J and π_{R_2} , we simulate data from the Lomax model (1) with different true values of the parameters α and β and different sample sizes n . Then, posterior samples are drawn from the joint posterior distribution of α and β by using the random-walk Metropolis algorithm in Roberts et al. (1997). For each chain, the sample size is 50,000 after 5000 burn-in samples. By choosing samples with jump of 10, a final chain of 5000 values is obtained. In order to make the estimation more robust, we take the posterior median as the Bayesian estimator for each parameter. The process is replicated 5000 times. Thus, we can obtain estimated mean squared errors and coverage probabilities of credible intervals (CIs).

The empirical results of the MSE and CP for the 95% CIs are listed in Table 1, where the estimated probabilities that the sample coefficient of variation CV_n is less than 1 are also associated. From Table 1, the following observations can be found:

- As is expected, the MSEs of the Bayesian estimators decrease as the sample size increases. Meanwhile, the CPs of the 95% CIs approach the nominal level of 0.95.
- The larger the value of α is, the higher the probability $P(CV_n < 1)$. For each parameter, the larger the true value is, the larger the corresponding MSE.
- According to both the MSE and CP, the performance of the Bayesian estimators under the reference prior π_{R_2} is much better than that under Jeffreys prior π_J . In fact, this is because π_{R_2} is a second-order probability matching prior while π_J is not.

4. Real data analysis

Now we apply the proposed Bayesian approach to analyze a sample of computer file sizes (in bytes) for 269 files with the *.ini extension on a Windows-based personal computer. The data are available on the website <http://web.uvic.ca/~dgiles/downloads/data>. The data were also analyzed by Holland et al. (2006) and Ferreira et al. (2016), where the Lomax distribution was shown to appropriately fit this data.

Table 1. Empirical MSEs and CPs (within parentheses) of Bayesian estimators based on the priors π_J and π_{R_2} .

n		20	50	80	100	150
			$\alpha = 3, \beta = 1$			
	$P(CV_n < 1)$	0.250	0.076	0.025	0.013	0.002
π_J	α	4.445 (0.983)	7.601 (0.974)	6.743 (0.961)	5.532 (0.959)	3.414 (0.948)
	β	1.021 (0.984)	1.442 (0.972)	1.254 (0.958)	1.007 (0.955)	0.654 (0.950)
π_{R_2}	α	3.023 (0.977)	5.957 (0.972)	5.528 (0.959)	4.646 (0.957)	2.965 (0.952)
	β	0.685 (0.980)	1.128 (0.969)	1.027 (0.958)	0.844 (0.956)	0.567 (0.951)
			$\alpha = 2, \beta = 1.5$			
	$P(CV_n < 1)$	0.146	0.021	0.003	0.001	0.000
π_J	α	3.864 (0.978)	3.312 (0.958)	2.085 (0.938)	1.495 (0.942)	0.565 (0.941)
	β	5.006 (0.979)	3.930 (0.956)	2.412 (0.939)	1.689 (0.938)	0.653 (0.943)
π_{R_2}	α	2.538 (0.971)	2.516 (0.960)	1.733 (0.942)	1.272 (0.946)	0.505 (0.944)
	β	3.277 (0.975)	3.080 (0.959)	2.007 (0.943)	1.431 (0.942)	0.582 (0.947)
			$\alpha = 1.5, \beta = 2$			
	$P(CV_n < 1)$	0.082	0.005	0.000	0.000	0.000
π_J	α	2.898 (0.975)	1.405 (0.947)	0.660 (0.939)	0.346 (0.946)	0.154 (0.946)
	β	14.218 (0.966)	6.280 (0.947)	3.015 (0.943)	1.456 (0.944)	0.639 (0.946)
π_{R_2}	α	1.927 (0.972)	1.113 (0.951)	0.572 (0.942)	0.304 (0.947)	0.143 (0.947)
	β	9.361 (0.971)	4.926 (0.952)	2.593 (0.942)	1.280 (0.948)	0.590 (0.949)

Table 2. Summary of the parametric bootstrap and the Bayesian estimates.

Method	Parameter	Estimator	SD	95% CI
Bootstrap	α	0.504	0.047	(0.422, 0.606)
	β	132.654	27.649	(86.717, 194.908)
π_J	α	0.497	0.043	(0.420, 0.587)
	β	128.406	24.125	(89.679, 182.416)
π_{R_2}	α	0.495	0.043	(0.419, 0.586)
	β	127.067	23.545	(89.755, 180.693)

For comparison purposes, the parametric bootstrap and the Bayesian approaches based on the Jeffreys prior π_J and the reference prior π_{R_2} are included here. The parametric bootstrap is based on the MLE since the sample coefficient of variation CV_n is greater than 1 here. The estimator along with the corresponding standard deviation (SD) and 95% confidence/credible interval (CI) for α and β are listed in Table 2. It can be seen from Table 2 that the Bayesian estimates of β are much more accurate than those of the parametric bootstrap according to the SD and the width of the CI. In addition, the performances of the two Bayesian estimates are close to each other, although the prior π_{R_2} behaves slightly better than π_J . To conclude, it is noted that our results are close to those in Ferreira et al. (2016) with respect to Jeffreys prior.

5. Concluding remarks

In this paper, objective Bayesian methods are developed to make inferences on the parameters of a Lomax distribution. Compared with the work in the literature, our contribution lies in the following points. First, we consider a larger class of noninformative priors, which includes probability matching priors, the MDI prior and both of reference priors. Second, it is revealed that one of reference priors is a second-order probability matching prior while Jeffreys prior is not. Third, we clarify that the MLE does not exist if the sample coefficient of variation $CV_n < 1$, and also consider the probability of such phenomenon in the simulation study. As a result, it is feasible to use objective Bayesian analysis for the Lomax distribution in practice.

Acknowledgments

The authors would like to thank the Editor, the Associate Editor and the anonymous Reviewers for their valuable comments and suggestions on earlier versions of this paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work is supported by the National Social Science Foundation of China Grant No. 21BTJ034.

References

- Atkinson, A. B., & Harrison, A. J. (1978). *Distribution of personal wealth in Britain*. Cambridge University Press.
- Bain, L. J., & Engelhardt, M. (1992). *Introduction to probability and mathematical statistics*. PWSKENT Publishing Company.
- Berger, J. O., & Bernardo, J. M. (1992). Ordered group reference priors with application to the multinomial problem. *Biometrika*, 79(1), 25–37. <https://doi.org/10.1093/biomet/79.1.25>
- Berger, J. O., Bernardo, J. M., & Sun, D. C. (2009). The formal definition of reference priors. *The Annals of Statistics*, 37(2), 905–938. <https://doi.org/10.1214/07-AOS587>
- Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference (with discussion). *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2), 113–147. <https://doi.org/10.1111/j.2517-6161.1979.tb01066.x>
- Chakraborty, T. (2019). An analysis of the maximum likelihood estimates for the Lomax distribution. *ArXiv: 1911.12612v2*, 1–14.
- Consonni, G., Fouskakis, D., Liseo, B., & Ntzoufras, I. (2018). Prior distributions for objective Bayesian analysis. *Bayesian Analysis*, 13(2), 627–679. <https://doi.org/10.1214/18-BA1103>
- Datta, G. S., & Mukerjee, R. (2004). *Probability matching priors: Higher order asymptotics*. Lecture notes in statistics. Springer.
- Deville, Y. (2016). *Renext: Renewal method for extreme values extrapolation* (p. 25). <https://cran.r-project.org/web/packages/Renext/Renext.pdf>.
- Ferreira, P. H., Gonzales, J. F. B., Tomazella, V. L. D., Ehlers, R. S., Louzada, F., & Silva, E. B. (2016). Objective Bayesian analysis for the Lomax distributionms. *ArXiv: 1602.08450v1*, 1–19.
- Ferreira, P. H., Ramos, E., Ramos, P. L., Gonzales, J. F. B., Tomazella, V. L. D., R. S. Ehlers, Silva, E. B., & Louzad, F. (2020). Objective Bayesian analysis for the Lomax distribution. *Statistics & Probability Letters*, 159, Article 108677. <https://doi.org/10.1016/j.spl.2019.108677>
- Holland, O., Golaup, A., & Aghvami, A. H. (2006). Traffic characteristics of aggregated module downloads for mobile terminal reconfiguration. *IEEE Proceedings: Communications*, 153(5), 683–690. <https://doi.org/10.1049/ip-com:20045155>
- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.
- Kang, S. G., Lee, W. D., & Kim, Y. (2021). Posterior propriety of bivariate lomax distribution under objective priors. *Communication in Statistics – Theory and Methods*, 50(9), 2201–2209. <https://doi.org/10.1080/03610926.2019.1662049>
- Lemonte, A. J., & Cordeiro, G. M. (2013). An extended Lomax distribution. *Statistics*, 47(4), 800–816. <https://doi.org/10.1080/02331888.2011.568119>
- Lindley, D. V. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 27(4), 986–1005. <https://doi.org/10.1214/aoms/1177728069>
- Lomax, K. (1954). Business failures: Another example of the analysis of failure data. *Journal of the American Statistical Association*, 49(268), 847–852. <https://doi.org/10.1080/01621459.1954.10501239>
- Marshall, A. W., & Olkin, I. (2007). *Life distributions: Structure of nonparametric, semiparametric, and parametric families*. Springer.
- Nadarajah, S. (2005). Sums, products, and ratios for the bivariate Lomax distribution. *Computational Statistics & Data Analysis*, 49(1), 109–129. <https://doi.org/10.1016/j.csda.2004.05.003>
- Nayak, T. K. (1987). Multivariate Lomax distribution: Properties and usefulness in reliability theory. *Journal of Applied Probability*, 24(1), 170–177. <https://doi.org/10.2307/3214068>
- Peers, H. W. (1965). On confidence sets and Bayesian probability points in the case of several parameters. *Journal of the Royal Statistical Society: Series B (Methodological)*, 27(1), 9–16. <https://doi.org/10.1111/j.2517-6161.1965.tb00581.x>
- Ramos, P. L., Louzada, F., & Ramos, E. (2018). Posterior properties of the Nakagami-*m* distribution using non-informative priors and applications in reliability. *IEEE Transactions on Reliability*, 67(1), 105–117. <https://doi.org/10.1109/TR.24>
- Roberts, G. O., Gelman, A., & Gilks, W. R. (1997). Weak convergence and optimal scaling of random walk metropolis algorithms. *Annals of Applied Probability*, 7(1), 110–120. <http://doi.org/10.1214/aoap/1034625254>
- Roy, D., & Gupta, R. P. (1996). Bivariate extension of Lomax and finite range distributions through characterization approach. *Journal of Multivariate Analysis*, 59(1), 22–33. <https://doi.org/10.1006/jmva.1996.0052>
- Zellner, A. (1977). Maximal data information prior distributions. In A. Aykac & C. Brumat (Eds.), *New developments in the applications of Bayesian methods* (pp. 211–232). North-Holland.

Appendix

Proof of the posterior propriety of π_G : The joint posterior density of (α, β) based on the prior π_G is given by

$$\pi_G(\alpha, \beta \mid x_1, \dots, x_n) \propto \pi_G(\theta) \cdot L(x_1, \dots, x_n \mid \alpha, \beta) = \beta^{-1} \alpha^{\tau-1} e^{-\alpha} \left(\frac{\alpha}{\beta}\right)^n \prod_{i=1}^n \left(1 + \frac{x_i}{\beta}\right)^{-(\alpha+1)}.$$

Denote $x_m = \min\{x_1, x_2, \dots, x_n\}$, then we have

$$\begin{aligned} \int_0^{+\infty} \int_0^{+\infty} \pi_G(\alpha, \beta \mid x_1, \dots, x_n) d\beta d\alpha &\leq \int_0^{+\infty} \alpha^{n+\tau-1} e^{-\alpha} \int_0^{+\infty} \beta^{-(n+1)} \left(1 + \frac{x_m}{\beta}\right)^{-n(\alpha+1)} d\beta d\alpha \\ &\propto \int_0^{+\infty} \alpha^{n+\tau-1} e^{-\alpha} B(n\alpha, n) d\alpha \\ &\propto \int_0^{+\infty} \alpha^{n+\tau-1} e^{-\alpha} \frac{1}{n\alpha(n\alpha+1) \cdots (n\alpha+n-1)} d\alpha, \end{aligned}$$

where $B(a, b)$ is the Beta function. Let

$$g_1(\alpha) = \frac{\alpha^{n+\tau-1}e^{-\alpha}}{n\alpha(n\alpha+1)\cdots(n\alpha+n-1)}.$$

Then, it can be seen that

$$\begin{aligned} g_1(\alpha) &= O(\alpha^{n+\tau-2}), \quad \alpha \rightarrow 0, \\ g_1(\alpha) &= O(\alpha^{\tau-1}e^{-\alpha}), \quad \alpha \rightarrow +\infty. \end{aligned}$$

Thus, $\int_0^{+\infty} g_1(\alpha) d\alpha < \infty$ for any $n \geq 1$. Consequently, the posterior distribution of π_G is proper. ■

Proof of Theorem 2.1: According to Peers (1965), the second-order probability matching prior $\pi_{M_1}(\theta)$ satisfies the following partial differential equation:

$$\frac{\partial}{\partial \alpha} \left\{ \alpha(\alpha+1)\pi_{M_1}(\alpha, \beta) \right\} + \frac{\partial}{\partial \beta} \left\{ \beta(\alpha+2)\pi_{M_1}(\alpha, \beta) \right\} = 0,$$

the solution of which is given by formula (3). Similarly, the second-order probability matching prior $\pi_{M_2}(\theta)$ is such that

$$\frac{\partial}{\partial \alpha} \left\{ \alpha^{\frac{3}{2}}(\alpha+2)^{\frac{1}{2}}\pi_{M_2}(\alpha, \beta) \right\} + \frac{\partial}{\partial \beta} \left\{ \alpha^{-\frac{1}{2}}\beta(\alpha+1)(\alpha+2)^{\frac{1}{2}}\pi_{M_2}(\alpha, \beta) \right\} = 0,$$

and the solution of this equation is given by formula (4). ■

Proof of Theorem 2.2: (a) In the light of Zellner (1977), the MDI prior for $\theta = (\alpha, \beta)$ has the following form:

$$\pi_M(\alpha, \beta) \propto \exp \{H(\alpha, \beta)\},$$

where $H(\alpha, \beta) = E[\log f(X)]$, and $f(x)$ is the density function of the Lomax distribution. Note that

$$H(\alpha, \beta) = \int_0^{+\infty} \log \left\{ \frac{\alpha}{\beta} \left(1 + \frac{x}{\beta}\right)^{-(\alpha+1)} \right\} \cdot \frac{\alpha}{\beta} \left(1 + \frac{x}{\beta}\right)^{-(\alpha+1)} dx = \log \left(\frac{\alpha}{\beta}\right) - \frac{\alpha+1}{\alpha}.$$

It follows that the MDI prior is

$$\pi_M(\theta) \propto \exp \left\{ \log \left(\frac{\alpha}{\beta}\right) - \frac{\alpha+1}{\alpha} \right\} \propto \frac{\alpha}{\beta e^{\frac{1}{\alpha}}}.$$

(b) The joint posterior density of (α, β) based on π_M is

$$\begin{aligned} \pi_M(\alpha, \beta | x_1, \dots, x_n) &\propto \pi_M(\theta) \cdot L(x_1, \dots, x_n | \alpha, \beta) \\ &= \frac{1}{\beta e^{\frac{1}{\alpha}}} \left(\frac{\alpha}{\beta}\right)^n \prod_{i=1}^n \left(1 + \frac{x_i}{\beta}\right)^{-(\alpha+1)}. \end{aligned}$$

Denote $x_M = \max\{x_1, x_2, \dots, x_n\}$. Then, we have

$$\begin{aligned} &\int_0^{+\infty} \int_0^{+\infty} \pi_M(\alpha, \beta | x_1, \dots, x_n) d\beta d\alpha \\ &\geq \int_0^{+\infty} \alpha^n e^{-\frac{1}{\alpha}} \int_0^{+\infty} \beta^{-(n+1)} \left(1 + \frac{x_M}{\beta}\right)^{-n(\alpha+1)} d\beta d\alpha \\ &\propto \int_0^{+\infty} \alpha^n e^{-\frac{1}{\alpha}} B(n\alpha, n) d\alpha \\ &\propto \int_0^{+\infty} \alpha^n e^{-\frac{1}{\alpha}} \frac{1}{n\alpha(n\alpha+1)\cdots(n\alpha+n-1)} d\alpha. \end{aligned}$$

Let

$$g_2(\alpha) = \alpha^n e^{-\frac{1}{\alpha}} \frac{1}{n\alpha(n\alpha+1)\cdots(n\alpha+n-1)}.$$

Then, it can be seen that $g_2(\alpha) \rightarrow \frac{1}{n^n}$ as $\alpha \rightarrow +\infty$. Thus, $\int_0^{+\infty} g_2(\alpha) d\alpha = \infty$ for any $n \geq 1$. Consequently, the posterior distribution $\pi_M(\alpha, \beta | x_1, \dots, x_n)$ is improper. ■

Proof of Theorem 2.4: (a) Let S be the inverse of the Fisher information matrix I . Then, up to a constant,

$$S = \begin{pmatrix} \alpha^2(\alpha+1)^2 & \alpha\beta(\alpha+1)(\alpha+2) \\ \alpha\beta(\alpha+1)(\alpha+2) & \beta^2\alpha^{-1}(\alpha+1)^2(\alpha+2) \end{pmatrix}.$$

Following the notations in Bernardo (1979), it holds that

$$h_1 = \frac{1}{\alpha^2(\alpha+1)^2}, \quad h_2 = \frac{\alpha}{\beta^2(\alpha+2)}.$$

Now we select compact set series $\Omega_l = [c_{1l}, d_{1l}] \times [c_{2l}, d_{2l}]$ for (α, β) , $l = 1, 2, \dots$, such that $c_{1l}, c_{2l} \rightarrow 0$, and $d_{1l}, d_{2l} \rightarrow +\infty$ as $l \rightarrow \infty$. Then,

$$\pi_2^l(\beta | \alpha) = \frac{|h_2|^{1/2} \mathbf{1}_{[c_{2l}, d_{2l}]}(\beta)}{\int_{c_{2l}}^{d_{2l}} |h_2|^{1/2} d\beta} := \frac{k_1}{\beta} \mathbf{1}_{[c_{2l}, d_{2l}]}(\beta),$$

where $\mathbf{1}_{[a,b]}(\cdot)$ refers to the indicator function on the interval $[a, b]$, and $k_1 = \log(d_{2l}) - \log(c_{2l})$ is a constant. Note that h_1 is independent of β . It follows that

$$E_1^l(\log |h_1| | \alpha) = \int_{c_{2l}}^{d_{2l}} \log |h_1| \cdot \frac{k_1}{\beta} d\beta = \log |h_1|.$$

Subsequently,

$$\begin{aligned} \pi_1^l(\alpha, \beta) &= \frac{\pi_2^l(\beta | \alpha) \cdot \exp\left\{\frac{1}{2} E_1^l(\log |h_1| | \alpha)\right\} \mathbf{1}_{[c_{1l}, d_{1l}]}(\alpha)}{\int_{c_{1l}}^{d_{1l}} \exp\left\{\frac{1}{2} E_1^l(\log |h_1| | \alpha)\right\} d\alpha} \\ &= \frac{\pi_2^l(\beta | \alpha) \cdot |h_1|^{\frac{1}{2}} \cdot \mathbf{1}_{[c_{1l}, d_{1l}]}(\alpha)}{\int_{c_{1l}}^{d_{1l}} |h_1|^{\frac{1}{2}} d\alpha} \\ &:= \frac{k_2}{\beta \alpha (\alpha + 1)} \mathbf{1}_{[c_{1l}, d_{1l}]}(\alpha) \mathbf{1}_{[c_{2l}, d_{2l}]}(\beta), \end{aligned}$$

where $k_2 = k_1 \cdot (\log \frac{d_{1l}}{d_{1l}+1} - \log \frac{c_{1l}}{c_{1l}+1})^{-1}$ is a constant.

Let (α^*, β^*) be an inner point of Ω_l . Then, the reference prior under the ordering group $\{\alpha, \beta\}$ is given by

$$\pi_{R_2}(\theta) = \lim_{l \rightarrow \infty} \frac{\pi_1^l(\alpha, \beta)}{\pi_1^l(\alpha^*, \beta^*)} \propto \frac{1}{\beta \alpha (\alpha + 1)}.$$

(b) Let $\pi_{R_2}(\alpha, \beta | x_1, \dots, x_n)$ be the posterior density based on the prior $\pi_{R_2}(\theta)$. Then,

$$\pi_{R_2}(\alpha, \beta | x_1, \dots, x_n) \propto \frac{1}{\beta \alpha (\alpha + 1)} \left(\frac{\alpha}{\beta}\right)^n \prod_{i=1}^n \left(1 + \frac{x_i}{\beta}\right)^{-(\alpha+1)}. \quad (\text{A1})$$

When $n \geq 2$, we have

$$\begin{aligned} &\int_0^{+\infty} \int_0^{+\infty} \pi_{R_2}(\alpha, \beta | x_1, \dots, x_n) d\beta d\alpha \\ &\propto \int_0^{+\infty} \int_0^{+\infty} \frac{1}{\beta \alpha (\alpha + 1)} \left(\frac{\alpha}{\beta}\right)^n \prod_{i=1}^n \left(1 + \frac{x_i}{\beta}\right)^{-(\alpha+1)} d\beta d\alpha \\ &\leq \int_0^{+\infty} \frac{\alpha^{n-1}}{\alpha + 1} B(n\alpha, n) d\alpha \\ &\propto \int_0^{+\infty} \frac{\alpha^{n-1}}{\alpha + 1} \cdot \frac{1}{n\alpha(n\alpha + 1) \cdots (n\alpha + n - 1)} d\alpha. \end{aligned}$$

Denote

$$g_3(\alpha) = \frac{\alpha^{n-1}}{\alpha + 1} \cdot \frac{1}{n\alpha(n\alpha + 1) \cdots (n\alpha + n - 1)}.$$

Then, $g_3(\alpha) = O(\alpha^{n-2})$ as $\alpha \rightarrow 0$, and $g_3(\alpha) = O(\alpha^{-2})$ as $\alpha \rightarrow +\infty$. It follows that $\int_0^{+\infty} g_3(\alpha) d\alpha < \infty$ for $n \geq 2$, which implies that the posterior is proper for $n \geq 2$.

When $n = 1$, it follows from (A1) that

$$\pi_{R_2}(\alpha, \beta | x_1, \dots, x_n) \propto \frac{1}{\beta^2(\alpha + 1)} \left(1 + \frac{x_1}{\beta}\right)^{-(\alpha+1)}.$$

Note that

$$\int_0^{+\infty} \int_0^{+\infty} \frac{1}{\beta^2(\alpha + 1)} \left(1 + \frac{x_1}{\beta}\right)^{-(\alpha+1)} d\beta d\alpha \propto \int_0^{+\infty} \frac{1}{\alpha(\alpha + 1)} d\alpha = \infty.$$

Thus, we have

$$\int_0^{+\infty} \int_0^{+\infty} \pi_{R_2}(\alpha, \beta | x_1, \dots, x_n) d\beta d\alpha = \infty,$$

which shows that $\pi_{R_2}(\alpha, \beta | x_1, \dots, x_n)$ is improper for $n = 1$.

(c) The result follows from Theorem 2.1. ■