MOAD'2024
The Sixth International Colloquium on
Methods and Tools for Decision Support
October 20th–22nd 2024, Tizi-Ouzou, Algeria

Bayesian Loss Robustness of Exponential Models

Lyasmine Harrouche
Laboratory of Pure and Applied Mathematics. Mouloud Mammeri University of Tizi-Ouzou. Algeria
Hocine Fellag
Laboratory of Pure and Applied Mathematics. Mouloud Mammeri University of Tizi-Ouzou. Algeria

Lynda Atil
Laboratory of Pure and Applied Mathematics. Mouloud Mammeri University of Tizi-Ouzou. Algeria

Abstract

Bayesian loss robustness of estimators for exponential samples is considered. The posterior expected loss variations are studied under general class of LINEX functions. To study the robustness, the ranges of the posterior expected loss of the Bayesian estimator is calculated. Using the range of posterior expected loss under class of LINEX functions, the robustness of the Bayesian estimator for exponential samples is studied.

Keywords and phrases Bayesian robustness; Exponential sample; LINEX loss function; Posterior loss-robustness; Prior distribution.

Received before September 15, 2024. **Accepted** October 3, 2024.

1 Introduction

The problem of robustness has always been an important element of the foundation of statistical analysis and has enjoyed the attention of authors in statistics for over half a century.
Frequent's criticism of the Bayesian approach often concentrates on what, for them, is an arbitrary choice of a prior distribution and/or loss function, a choice that greatly influences the optimal decision taken. Bayesians too have been aware of the sensitivity of their analysis to possible misspecification of the prior distribution, the loss function, or the model, to quantify and interpret the uncertainty induced by the partial knowledge of one (or more) of the three elements in Bayesian analysis (prior, likelihood and loss). An awareness that led to a series of papers on robust Bayesian analysis, also called Bayesian sensitivity analysis see Berger [4] and Bhatia et al [6].

In a Bayesian framework, most of the research efforts in the study of Bayesian robustness have been directed towards prior-robustness, see for example, Berger [3], Betró and Guglielmi [5], Carota and Ruggeri [7], Joshi et al [11], Perone-Pacifico et al [15]. According to Berger [4], p. 251, the robustness problems for loss functions are often not as severe as those for priors. As a simple example, quadratic and absolute error estimation loss are quite different, and yet they result in estimating θ by the posterior mean and posterior median, respectively, these two estimates are typically reasonably similar.

In a decision-theoretic framework, a specification of a loss function can be even more severe than that specifying a prior.
For example, in the context of a skewed posterior distribution covering a range of severe losses, as is typically the case in actuarial losses, the seemingly small distance between these two estimators, mentioned in the previous example can be crucial. In an extensive review of the decision-theoretic approach to credibility, see Heilmann [10], eight different loss functions were listed together with the kind of Bayesian estimators, they produce for a variety of models and prior distributions.

Several authors have paid attention to loss robustness, the most important papers included Arias et al [2] and Micheas and Dey [14] who introduced the maximum posterior estimate range for measuring loss robustness. Abraham and Daurès [1] derived an analytic approximation to a set of Bayes actions associated with the class of loss functions. Dey et al [8] discussed variations of the posterior expected loss concerning the asymmetric parameter when the loss functions belong to a certain class. Ranges of the posterior expected loss, influence function and minimax regret principles are adopted to measure the robustness and optimum choice of the loss functions, with applications on the continuous exponential family and discrete power series distributions under the class of LINEX loss functions. Moreover, Dey and Micheas [9] discussed applications including the class of weighted squared error and Hellinger distance loss. Makov [13] suggested several measures of Bayesian loss-robustness and discussed the influence approach to robustness and used it for the study of families of loss functions.

In this paper, we introduce a robustness measure of posterior expected loss, using the LINEX loss functions belonging to a certain class and using the range as robustness measure of the Bayes estimator of exponential sample.

The structure of the paper is as follows : we introduce in Section 2, the concept of the range of posterior expected loss as a measure of assessing loss robustness under the class of LINEX loss function. In Section 3, we study loss robustness for the exponential sample under the LINEX loss function and determine the parameters of the prior function, which minimize the range for Bayes estimate. The work ends with conclusion.

2 Models and methods

Let us define a random variable X on a probability space $(\mathbb{R}^+, \mathbb{F}, \mathbf{P})$ and assume that, for a given $x \in \mathcal{X}$, the sample space $\mathcal{X}$ is fixed.

Let an action a and parameter θ belong to an action space $\mathcal{A}$ and a parameter space Θ respectively. Also, we will assume that the posterior distribution of θ given x; defined by $\pi(\theta|x) = \frac{\pi(\theta)f(x|\theta)}{m(x)}$, where $m(x) = \int_\Theta \pi(\theta)f(x|\theta)d\theta$ is the marginal distribution of x, $\pi(\theta)$ is the prior distribution of parameter θ and $f(x|\theta)$ the likelihood function.

▶ **Definition 1.** *Let a loss function denoted by $L_\theta(\theta, a) \in \mathcal{L}$ a class of loss functions. We define the LINEX loss function which is an asymmetric loss introduced by Klebanov [12] as follows*

$$L_c(\theta, a) = \exp[c(a - \theta)] - c(a - \theta) - 1$$

where the constant $c \neq 0$ determines the shape of the loss function, its behaviour is defined by Both Zellner [17] and Varian [16].

▶ **Definition 2.** *For $L \in \mathcal{L}$, we define the posterior expected loss in using action a by*

$$\rho_L(\pi, x, a) = \mathbb{E}_{\pi(.|x)}[L_c(\theta, a)] = \int_\Theta L_c(\theta, a)\pi(\theta|x)d\theta \tag{1}$$

For all a, the posterior expected loss under LINEX loss function is given by

$$\rho_L(\pi, x, a) = \exp(ac)E_{\theta|x}(\exp(-c\theta)) - c(a - E_{\theta|x}(\theta)) - 1 \tag{2}$$

▶ **Definition 3.** *The Bayesian estimator denoted by $\hat{\theta}_B$, is defined as the value of $\theta \in \Theta$ which minimizes posterior expected loss function $\rho_L(\pi, x, a)$.*

$$\hat{\theta}_B = -\frac{1}{c} \ln E_{\theta|x}(\exp(-c\theta)) \tag{3}$$

▶ **Definition 4.** *In the case of Bayesian robustness analysis, the range is used as measure of robustness of the quantity of interest (prior, likelihood and loss functions).*
For assessment of the sensitivity of the class $\mathcal{L}$, the range of the posterior expected loss is then defined as follows

$$R(a, x) = \sup_{L\in\mathcal{L}} \rho_L(\pi, x, a) - \inf_{L\in\mathcal{L}} \rho_L(\pi, x, a) \tag{4}$$

The class $\mathcal{L}$ is robust if $R(a,x)$ is sufficiently small.

▶ **Example 5.** A plot of $R(a)$ against a indicates the degree of robustness for different values of a. An example is illustrated by Makov [13].

To derive the range, in the LINEX class loss function, the steps are as follows:

1. Modelling uncertainty in the loss function by specifying a class of possible loss functions. This class of LINEX loss function $\mathcal{L}$ is given by

$$\mathcal{L} = \{L_\theta(\theta, a) : c_L < c < c_U\}$$

   where, $c \neq 0$. Overall, it seems reasonable to consider a class of LINEX loss functions where c_L and c_U have the same sign see Dey et al[8].
2. Determining the loss posterior ranges over the classes given above.
   Here, in order to calculate the range of the posterior expected loss $\rho_c(\pi, x, a)$, we need to know the behavior of $\rho_c(\pi, x, a)$.

In exponential family case, $\rho_c(\pi, x, a)$ is a convex function in c having minimum at c tends to 0. If c_L and c_U have the same sign, we calculate the range from (4) (see Dey et al[8]) as follows

$$R(a,x) = \begin{cases} \rho_{L_{c_U}}(\pi, x, a) - \rho_{L_{c_L}}(\pi, x, a) & \text{if } 0 < c_L < c_U) \\ \rho_{L_{c_L}}(\pi, x, a) - \rho_{L_{c_U}}(\pi, x, a) & \text{if } c_L < c_U < 0) \end{cases}$$

In the following, we assume that $0 < c_L < c_U$.

▶ **Proposition 6.** *Let us calculate the range of the posterior loss under the Bayesian estimator $a = \hat{\theta}$ for a given $x \in \mathcal{X}$*
We replace $a = \hat{\theta}$ by the Bayesian estimator in (2) and obtain

$$\rho_L(\pi, x, \hat{\theta}_B) = \ln\left[E_{\theta|x}(\exp(-c\theta))\right] + cE_{\theta|x}(\theta) \tag{5}$$

The corresponding range is then equal to

$$R(\hat{\theta}_B, x) = \ln\left[E_{\theta|x}(\exp(-c_U\theta))\right] - \ln\left[E_{\theta|x}(\exp(-c_L\theta))\right] + E_{\theta|x}(\theta)(c_U - c_L) \tag{6}$$

3 Loss robustness for exponential samples under LINEX loss function

Let us consider a sample of an exponential random variable $X \sim Ex(\theta)$ with density $f(x) = \theta \exp(-\theta x)$ where x and $\theta \in \mathbb{R}_+^*$. The likelihood function is given by

$$L(x|\theta) = \theta^n \exp(-\theta s)$$

where $s = \Sigma_{i=1}^n x_i$
Let $\pi(\theta)$ be the prior distribution defined by $\Gamma(\alpha, \beta)$ distribution:

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}$$

with α, β, $\theta \in \mathbb{R}_+^*$ and $\Gamma(\alpha) = \int_0^{+\infty} \theta^{\alpha-1} e^{-\theta} d\theta$
The posterior distribution $\pi(\theta|x)$ is then $\theta|x \sim \Gamma(n+\alpha, s+\beta)$. From (1), the posterior expected loss is given by

$$\rho_L(\pi, x, a) = \exp(ac)\left(1 + \frac{c}{\beta + s}\right)^{-(\alpha+n)} - c\left(a - \frac{\alpha+n}{\beta+s}\right) - 1 \tag{7}$$

Then, the corresponding Bayesian estimator is

$$\hat{\theta}_B = \frac{\alpha+n}{c} \ln\left(1 + \frac{c}{\beta+s}\right) \tag{8}$$

3.1 Computation of the range

Since $\rho(\pi, x, a)$ is a convex function in c and has a minimum when c tends to 0 the range is given by

$$\begin{aligned} R(a,x) &= \left(\exp(ac_U)\left(\frac{\beta+s}{\beta+s+c_U}\right)^{\alpha+n} - c_U\left(a - \frac{\alpha+n}{\beta+s}\right)\right) \\ &\quad - \left(\exp(ac_L)\left(\frac{\beta+s}{\beta+s+c_L}\right)^{\alpha+n} - c_L\left(a - \frac{\alpha+n}{\beta+s}\right)\right) \end{aligned} \tag{9}$$

With $0 < c_L < c_U$

▶ Remark 7. Following (9), the range depends on the sample size n, the sum s of the observations, the parameters of asymmetry c_U and c_L, the parameters (α, β) of the prior distribution $\Gamma(\alpha, \beta)$, and the action a.
To obtain the robustness of the action, we need to derive the parameters (α, β) of the prior which minimize the range $R(a, x)$. In the following, we set $R(a, x) = R(\alpha, \beta)$ when a is the Bayesian estimator.

3.2 Minimisation of the range

The minimisation of the range is performed with respect to the parameters of the prior distribution in the two cases, Bayesian estimator, from (7) and (8) we obtain the posterior loss as follows

$$\rho(\pi, x, \hat{\theta}_B) = (\alpha+n)\left[\frac{c}{\beta+s} - \ln(1 + \frac{c}{\beta+s})\right]$$

From (4) we have $R(a,x) = \sup_{L\in\mathcal{L}} \rho_L(\pi, x, a) - \inf_{L\in\mathcal{L}} \rho_L(\pi, x, a)$. Since we are interested by the couple (α, β) which minimizes the range $R(a, x)$, we denote it by $R(\alpha, \beta)$ as a function of these parameters.
The range of posterior expected loss function is then

$$R(\alpha,\beta) = (\alpha+n)\left[\left(\left(\frac{c_U}{\beta+s}\right) - \ln\left(1+\frac{c_U}{\beta+s}\right)\right) - \left(\left(\frac{c_L}{\beta+s}\right) - \ln\left(1+\frac{c_L}{\beta+s}\right)\right)\right] \tag{10}$$

▶ **Proposition 8.**
1. *Notice that α is in the first terms of the formula (10) only. Then, the range is small if α tends to 0.*
2. *Concerning β, we know that if x tends to 0, $\ln(1 + x)$ tends to 0. This means that if $\frac{c_U}{\beta+s}$ and $\frac{c_L}{\beta+s}$ are close to zero for big values of β, then the second term of (10) is again close to zero.*

$$\tfrac{c_U}{\beta+s} - \ln\left(1+\tfrac{c_U}{\beta+s}\right) \simeq 0 \textit{ for } \tfrac{c_L}{\beta+s} + \ln\left(1+\tfrac{c_L}{\beta+s}\right) \simeq 0$$

For the range to be reduced, β must be large and α small, hence the robustness of the class of loss functions.

3.3 Variation of the range with respect to α and β

1. Variation of the range with α
   In this case, following Proposition 8, we have α tends to 0 and $\frac{c_U}{\beta+s}$ tends to 0 then in neighborhood to 0, we can say that $\frac{c_U}{\beta+s} = \alpha$. This allows us to set $\beta = \frac{c_U}{\alpha} - s$ with $0 < \alpha < \frac{c_U}{s}$ in formula (10).
   We obtain the range denoted by $R(\alpha)$ defined by :

$$R(\alpha) = (n+\alpha)\left[\alpha - \ln(1+\alpha) - \frac{\alpha c_L}{c_U} + \ln(1 + \frac{\alpha c_L}{c_U})\right] \tag{11}$$

2. Variation of the range with β
   The range is obtained after replacing α by $\frac{c_U}{\beta+s}$ in (2) and then,

$$R(\beta) = \left[\frac{c_U}{\beta+s} + n\right]\left[\left(\frac{c_U}{\beta+s}\right) - \ln\left(1+\frac{c_U}{\beta+s}\right) - \left(\frac{c_L}{\beta+s}\right) + \ln\left(1+\frac{c_L}{\beta+s}\right)\right] \quad (12)$$

To illustrate the behavior of the range with respect to α and β, let us consider the following example.

▶ **Example 9.** We define a random variable from $Ex(1)$ sample of size $n = 50$, fixed beforehand. The sum s of the sample $Ex(1)$ observations is given by the mean of the simulation sums of $10,000$ samples $Ex(1)$ of size 50, then we obtain $s = 49.97$, c_U and c_L are considered constants, with $c_U = 5$ and $c_L = 1$, using formulas (11) and (12) for the variation of α and β respectively, we obtain the variation of the range illustrated by Figure 1 and Figure 2

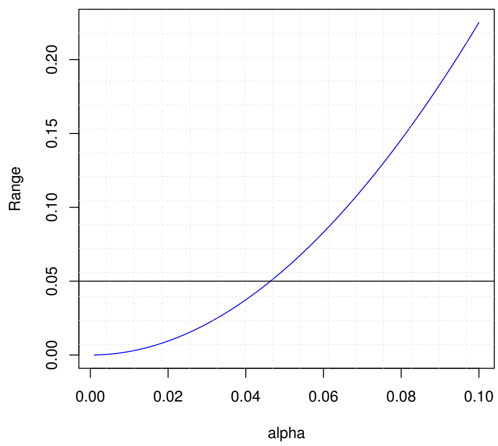

Figure 1 Variation of the range with α

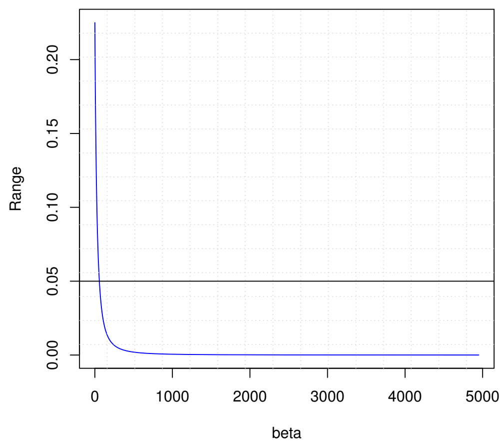

Figure 2 Variation of the range with β

Figure 1: The graph represents the range $R(\alpha)$. We notice that for small values of α, the range is reduced and the robustness is verified than, in the case where the range is equal to 0.05, we calculated α and deduced β as follows

$$\alpha = 0.04634277$$
$$\beta = 57.9217$$

Figure 2: From the graph representing the range $R(\beta)$, we notice that $R(\beta)$ is small and tends towards 0 for large values of β. In the case where the range takes the value 0.05, α and β are

$$\alpha = 0.0463428$$
$$\beta = 57.92162$$

According to both figures, one can say that, to reduce the value of the range, it is sufficient to take large values of β and small values of α. As we demonstrated in Proposition 8.

4 Conclusion

In this work, we have discussed the problem of Bayesian robustness of the class of LINEX loss function, for the exponential sample with the conjugate prior gamma distribution $\Gamma(\alpha, \beta)$. The robustness is obtained when the range of the expected loss function is minimised. In our case, using the Bayesian estimator, we demonstrated that the robustness is verified when the β parameter is high and the α parameter is small. In future work, simulated studies can be extended to more general models.

References

1 Christophe Abraham and Jean-Pierre Daurès. Analytic approximation of the interval of Bayes actions derived from a class of loss functions. *Test*, 8(1):129–145, June 1999. URL: `http://link.springer.com/10.1007/BF02595866`, `doi:10.1007/BF02595866`.

2 J.P. Arias-Nicolás, J. Martín, F. Ruggeri, and A. Suárez-Llorens. Optimal actions in problems with convex loss functions. *International Journal of Approximate Reasoning*, 50(2):303–314, February 2009. URL: `https://linkinghub.elsevier.com/retrieve/pii/S0888613X08000558`, `doi:10.1016/j.ijar.2008.03.014`.

3 James O. Berger. Robust Bayesian analysis: sensitivity to the prior. *Journal of Statistical Planning and Inference*, 25(3):303–328, July 1990. URL: `https://linkinghub.elsevier.com/retrieve/pii/037837589090079A`, `doi:10.1016/0378-3758(90)90079-A`.

4 James O. Berger. *Statistical decision theory and Bayesian analysis*. Springer Series in Statistics. Springer New York, New York, NY, softcover reprint of the hardcover 2. ed edition, 2010.

5 Bruno Betrò and Alessandra Guglielmi. Methods for Global Prior Robustness under Generalized Moment Conditions. In P. Bickel, P. Diggle, S. Fienberg, K. Krickeberg, I. Olkin, N. Wermuth, S. Zeger, David Ríos Insua, and Fabrizio Ruggeri, editors, *Robust Bayesian Analysis*, volume 152, pages 273–293. Springer New York, New York, NY, 2000. Series Title: Lecture Notes in Statistics. URL: `http://link.springer.com/10.1007/978-1-4612-1306-2_15`, `doi:10.1007/978-1-4612-1306-2_15`.

6 Kush Bhatia, Yi-An Ma, Anca D. Dragan, Peter L. Bartlett, and Michael I. Jordan. Bayesian Robustness: A Nonasymptotic Viewpoint. *Journal of the American Statistical Association*, 119(546):1112–1123, April 2024. URL: `https://www.tandfonline.com/doi/full/10.1080/01621459.2023.2174121`, `doi:10.1080/01621459.2023.2174121`.

7 Cinzia Carota and Fabrizio Ruggeri. Robust Bayesian analysis given priors on partition sets. *Test*, 3(2):73–86, December 1994. URL: `http://link.springer.com/10.1007/BF02562694`, `doi:10.1007/BF02562694`.

8 Dipak K. Dey, Kuo-ren Lou, and Sudip Bose. A BAYESIAN APPROACH TO LOSS ROBUSTNESS. *Statistics & Risk Modeling*, 16(1), January 1998. URL: `https://www.degruyter.com/document/doi/10.1524/strm.1998.16.1.65/html`, `doi:10.1524/strm.1998.16.1.65`.

9 Dipak K. Dey and Athanasios C. Micheas. Ranges of Posterior Expected Losses and ϵ-Robust Actions. In P. Bickel, P. Diggle, S. Fienberg, K. Krickeberg, I. Olkin, N. Wermuth, S. Zeger, David Ríos Insua, and Fabrizio Ruggeri, editors, *Robust Bayesian Analysis*, volume 152, pages 145–159. Springer New York, New York, NY, 2000. Series Title: Lecture Notes in Statistics. URL: `http://link.springer.com/10.1007/978-1-4612-1306-2_8`, `doi:10.1007/978-1-4612-1306-2_8`.

10 Wolf-Rüdiger Heilmann. Decision theoretic foundations of credibility theory. *Insurance: Mathematics and Economics*, 8(1):77–95, March 1989. URL: `https://linkinghub.elsevier.com/retrieve/pii/0167668789900504`, `doi:10.1016/0167-6687(89)90050-4`.

11 Chaitanya Joshi, Fabrizio Ruggeri, and Simon P. Wilson. Prior Robustness for Bayesian Implementation of the Fault Tree Analysis. *IEEE Transactions on Reliability*, 67(1):170–183, March 2018. URL: `http://ieeexplore.ieee.org/document/8240697/`, `doi:10.1109/TR.2017.2778241`.

12 Lev (Leo) Borisovich Klebanov. "universal" loss functions and unbiased estimates. In *Doklady Akademii Nauk*, volume 203, pages 1249–1251. Russian Academy of Sciences, 1972.

13 Udi E. Makov. Some aspects of Bayesian loss-robustness. *Journal of Statistical Planning and Inference*, 38(3):359–370, March 1994. URL: `https://linkinghub.elsevier.com/retrieve/pii/0378375894900167`, `doi:10.1016/0378-3758(94)90016-7`.

14 Athanasios C. Micheas and Dipak K. Dey. On Measuring Loss Robustness Using Maximum A Posteriori Estimate. *Communications in Statistics - Theory and Methods*, 33(5):1069–1085, January 2004. URL: `http://www.tandfonline.com/doi/abs/10.1081/STA-120029826`, `doi:10.1081/STA-120029826`.

15 M. Perone-Pacifico, G. Salinetti, and L. Tardella. Bayesian robustness on constrained density band classes. *Test*, 5(2):395–409, December 1996. URL: `http://link.springer.com/10.1007/BF02562625`, `doi:10.1007/BF02562625`.

16 Hal R Varian. A bayesian approach to real estate assessment. *Studies in Bayesian econometrics and statistics in honor of Leonard J. Savage*, 1975.

17 Arnold Zellner. Bayesian Estimation and Prediction Using Asymmetric Loss Functions. *Journal of the American Statistical Association*, 81(394):446–451, June 1986. URL: `http://www.tandfonline.com/doi/abs/10.1080/01621459.1986.10478289`, `doi:10.1080/01621459.1986.10478289`.