JIRSS (2023)
Vol. 22, No. 01, pp 49-66
DOI: 10.22034/jirss.2024.706593

Bayesian Premium Estimators for Pareto Distribution in the Presence of Outliers

Rahele Mollaie [1], **Mehdi Jabbari Nooghabi** [1]

[1] Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran.

Received: 16/04/2022, Accepted: 17/06/2023, Published online: 15/06/2024

Abstract. We assume the Pareto distribution in the presence of outliers based on the Dixit model. We consider the estimation of the Bayesian Premium under squared error loss function (symmetric), linear exponential, and entropy loss functions (asymmetric), using informative and non-informative priors. We use the Lindley approximation and Markov Chain Monte Carlo methods such as the importance sampling procedure for deriving results. Finally, the results are analyzed using simulation studies.

Keywords. Bayesian Premium, Pareto Distribution, Outliers, Non-informative Prior Distribution, Entropy, LINEX and Squared Error Loss.

MSC: 62F15, 62F35, 91B30.

1 Introduction

The study of Premium principles is one of the essential topics in actuarial science, and there are several approaches to determine a Premium, see Young (2004) for a complete review. The net Premium or expected premium is used in financial economics and Risk theory, also one of its main advantages is its simple understanding and easy explanation for the Policyholder. Therefore, we use the concept of net Premium that is determined by using the expectation of X, where X represents the claim size or loss amount of one contract in an insurance company.The credibility theory is a statistical tool for calculating future period Premiums based on a past experiences of insured. This theory was introduced by Mowbray (1914), and Bailey (1950) showed the relation between credibility theory and Bayesian method. Also, Heilman (1985) presented an

Rahele Mollaie (ra.mollaye@mail.um.ac.ir)
Corresponding Author: Mahdi Jabbari Nooghabi (jabbarinm@um.ac.ir).

approach for premium calculation by credibility model based on the percentile principle and Kiapour (2018) obtained E-Bayesian and robust Bayesian Premium estimators by using a class of conjugate prior distributions under the squared log error loss function. An outlier usually refers to some of the observations in a distribution that deviate very much from the other observations. Hawkins (1980), considered that, if the data is very large and deviates from the rest of the data so that it brings to mind the idea that it was produced by another mechanism, we say that the data are in the presence of outliers. Many definitions for outliers are expressed by Grubbs (1950), Kendall and Buckland (1957), Anscombe (1960), Ferguson (1961), Grubbs (1969), Miller (1981), and Barnett and Lewis (1994). Checking the quality of the data is a crucial step in statistical analysis, as the presence of outliers can lead to inaccurate results if not identified and addressed with suitable statistical models. Therefore, one of the first steps to obtain a correct analysis is to identify outlier observations that may carry important information. One of the most important ways to deal with outlier data is its modelling. In this regard, many statistical models in the presence of outliers are proposed, for example, Kale and Sinha (1971), Joshi (1972), and Dixit (1987) introduced a model contaminated by outliers; for more information about this model, we refer readers to Dixit (1989) and Dixit, Moore, and Barnett (1996).

Dixit and Jabbari Nooghabi (2011) obtained the maximum likelihood and uniformly minimum variance unbiased estimator of unknown shape parameter α for the Pareto distribution in the presence of outliers. Also, Okhli and Jabbari Nooghabi (2021) used Bayesian approach for obtaining the estimators of parameters for contaminated exponential distribution for various sample sizes. In this paper by using the concept of net Premium, the Bayesian Premium estimator proposed by Bailey is obtained, where the Pareto distribution in the presence of k outliers under Dixit model is used as the claim size distribution. Several authors have performed application of Pareto distribution in insurance real data. One can refer to Benktander (1963), Dixit and Jabbari Nooghabi (2011a, 2011b), Jeevanand and Nair (1992, 1993, 1996, 1998), Jabbari Nooghabi and Khaleghpanah Nooghabi (2016), Jabbari Nooghabi (2019), etc. For example, Jabbari Nooghabi (2019) presented three examples that the datasets follow a Pareto distribution in the presence of outliers. Data of Norwegian fire claims for the year 1975 are considered and the Pareto distribution is fitted on it and in other example a random sample of size 30 is taken from Danish data. This data involves large fire insurance claims and the Pareto distribution is fitted to this dataset. Also, one example is presented for motor insurance, that is related to the claim amounts greater than 500,000 Rials. In this example some vehicles are very expensive or severely damaged, therefore the claim amounts of these vehicles are much higher than other cases. By using one-sample Kolmogorov-Smirnov goodness of fit test, it has been seen that data follows the Pareto distribution.

The paper is organized as follows. In Section 2, we describe the model and net Premium. The maximum likelihood estimator (MLE) of the net Premium is considered in Section 3. Section 4, is devoted to the approximate Bayesian estimators of net Premium by using the Lindley approximation and importance sampling method. In

Section 5, a Monte Carlo simulation study is used for comparing various estimators and some concluding remarks are made in Section 7. Finally, we present an example to calculate the Premium in Section 6.

2 Model

In insurance for modeling the claims, where the minimum claim is a modal value, we can use the Pareto distribution. The Pareto distribution has the probability density function (pdf)

$$f_X(x) = \frac{\alpha\theta^{\alpha}}{x^{\alpha+1}}, \qquad 0 < \theta \le x, \quad \alpha > 0, \quad \theta > 0,$$

and cumulative distribution function (cdf)

$$F_X(x) = 1 - \left(\frac{\theta}{x}\right)^{\alpha}.$$

Benktander (1963) observed that the Pareto distribution is useful for automobile insurance problems. Dixit and Jabbari Nooghabi (2011) suggested a Pareto distribution contaminated by k outliers for modeling the claim amounts.

Let a set of random variables $(X_1, X_2, \ldots, X_n)$ represent the claim amounts of a motor insurance company. It is assumed that some of these claims are β times higher than the claims of normal vehicles. Therefore the random variables $(X_1, X_2, \ldots, X_n)$ are assumed such that any k (number of outliers) of them are distributed with pdf

$$f_2(x; \alpha, \beta, \theta) = \frac{\alpha(\beta\theta)^{\alpha}}{x^{\alpha+1}}, \quad 0 < \beta\theta \le x, \quad \alpha > 0, \quad \beta > 1, \quad \theta > 0, \tag{2.1}$$

and $n-k$ out of n random variables are distributed with pdf

$$f_1(x; \alpha, \theta) = \frac{\alpha\theta^{\alpha}}{x^{\alpha+1}}, \quad 0 < \theta \le x, \quad \alpha > 0, \tag{2.2}$$

where the shape parameter α is unknown. Also the scale parameter θ, contamination factor β, and k are known. According to Dixit and Jabbari Nooghabi (2011), the joint pdf of $(X_1, X_2, \ldots, X_n)$ in the presence of k outliers is given by

$$f(x_1, x_2, \ldots, x_n; \alpha, \beta, \theta) = \frac{\alpha^n \beta^{k\alpha}\theta^{n\alpha}}{c(n,k)} \left(\prod_{i=1}^{n} x_i\right)^{-(\alpha+1)} \sum_{A_1=1}^{n-k+1} \sum_{A_2=A_1+1}^{n-k+2} \cdots \sum_{A_k=A_{k-1}+1}^{n} \prod_{j=1}^{k} I(x_{A_j} - \beta\theta) I(x_{A_j} - \theta),$$

where $c(n,k) = \frac{n!}{k!(n-k)!}$ and $I(y)$ is the indicator function defined as

$$I(y) = \begin{cases} 1, & y \ge 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

Observe that the indicator function in (2.3) is defined based on (2.1) and (2.2).
The marginal distribution of X_i, $i = 1, 2, \ldots, n$, is

$$f(x_i; \alpha, \beta, \theta) = b\frac{\alpha(\beta\theta)^\alpha}{x_i^{\alpha+1}} I(x_i - \beta\theta) + \bar{b}\frac{\alpha\theta^\alpha}{x_i^{\alpha+1}} I(x_i - \theta), \quad \alpha > 0, \quad \beta > 1, \quad \theta > 0, \tag{2.4}$$

where, $b = \frac{k}{n}$ and $\bar{b} = 1 - b = \frac{n-k}{n}$.

Note: In Dixit model of outliers, the number of outliers is assumed to be known as well as the sample size (a given parameter). But in a real example, number of outliers is always unknown and it should be estimated. One of the methods is evaluating the likelihood for different values of k and selecting each k which maximizes the likelihood.

Also, the net Premium (P) is given by

$$P = \mu(\alpha) = \left(\frac{\alpha}{\alpha - 1}\right)\theta(b\beta + \bar{b}). \tag{2.5}$$

3 Maximum Likelihood Estimator

In this section, the MLE of P from the Pareto distribution contaminated by k outliers in equation (2.4) is derived. Dixit and Jabbari Nooghabi (2011) obtained different estimators of shape parameter for the Pareto distribution with k outliers. The likelihood function is

$$\begin{aligned} L(\alpha|\underline{x}; \beta, \theta) = {} & \frac{\alpha^n \beta^{k\alpha} \theta^{n\alpha}}{c(n,k)} \left(\prod_{i=1}^{n} x_i\right)^{-(\alpha+1)} \\ & \sum_{A_1=1}^{n-k+1} \sum_{A_2=A_1+1}^{n-k+2} \cdots \sum_{A_k=A_{k-1}+1}^{n} \prod_{j=1}^{k} I(x_{A_j} - \beta\theta) I(x_{A_j} - \theta), \end{aligned} \tag{3.1}$$

where $\underline{x} = (x_1, x_2, \ldots, x_n)$ and the MLE of α, denoted by $\hat{\alpha}_{ml}$, is given as

$$\hat{\alpha}_{ml} = \frac{n}{\sum_{i=1}^{n} \log(x_i) - k\log(\beta) - n\log(\theta)}.$$

Also by using the property of MLE and substituting $\hat{\alpha}_{ml}$ in equation (2.5), the MLE of P is obtained as

$$\hat{P}_{ml} = \hat{\mu}(\alpha) = \left(\frac{\hat{\alpha}_{ml}}{\hat{\alpha}_{ml} - 1}\right)\theta(b\beta + \bar{b}). \tag{3.2}$$

4 Bayesian Estimator

In this section, we consider the Bayesian estimator of P under symmetric (squared error) and asymmetric (LINEX and entropy) loss functions, which are defined, respectively, as

$$L_S(\hat{u}(\alpha), u(\alpha)) = (\hat{u}(\alpha) - u(\alpha))^2, \tag{4.1}$$

$$L_L(\hat{u}(\alpha), u(\alpha)) = e^{m(\hat{u}(\alpha)-u(\alpha))} - m(\hat{u}(\alpha) - u(\alpha)) - 1, \quad m \neq 0, \tag{4.2}$$

and

$$L_E(\hat{u}(\alpha) - u(\alpha)) = \left(\frac{\hat{u}(\alpha)}{u(\alpha)}\right)^q - q\log\left(\frac{\hat{u}(\alpha)}{u(\alpha)}\right) - 1, \quad q \neq 0,$$

where $\hat{u}(\alpha)$ is an estimator of $u(\alpha)$. The Bayesian estimator under the loss function L_S is the posterior mean of $u(\alpha)$, which is denoted by $\hat{u}_{BS}$. The corresponding Bayesian estimators under the loss function L_L and L_S are given, respectively by

$$\hat{u}_{BL} = -\frac{1}{m}\log\left(E_\alpha(e^{-mu(\alpha)}|\underline{x})\right),$$

and

$$\hat{u}_{BE} = \left(E_\alpha(u^{-q}(\alpha)|x)\right)^{-\frac{1}{q}}.$$

4.1 Bayesian Premium Estimator based on non-informative Prior Distribution

Assume that β and θ are known and that the shape parameter α is a random variable with Jeffreys prior distribution. The Jeffreys prior is a non-informative prior distribution in the form $\pi(\alpha) = \sqrt{I(\alpha)}$, where $I(\alpha)$ is the Fisher information and obtained by $I(\alpha) = -E\left(\frac{\partial^2\ell}{\partial\alpha^2}\right) = \frac{n}{\alpha^2}$ when $\ell = \log(L(\alpha|\underline{x}; \beta, \theta))$. Hence $\pi(\alpha)$ is given by

$$\pi(\alpha) = \frac{\sqrt{n}}{\alpha}. \tag{4.3}$$

Using equations (4.3) and (3.1), we get the posterior distribution of α as follows:

$$\pi(\alpha|\underline{x}; \beta, \theta) = \frac{\pi(\alpha)L(\alpha|\underline{x}; \beta, \theta)}{\int_0^\infty \pi(\alpha)L(\alpha|\underline{x}; \beta, \theta)d\alpha} = \frac{1}{C}\alpha^{n-1}\beta^{k\alpha}\theta^{n\alpha}\left(\prod_{i=1}^n x_i\right)^{-(\alpha+1)},$$

where $C = \int_0^\infty \alpha^{n-1}\beta^{k\alpha}\theta^{n\alpha}\left(\prod_{i=1}^n x_i\right)^{-(\alpha+1)} d\alpha$.

First, we obtain the Bayesian estimator of P under the loss function L_S, using the posterior distribution $\pi(\alpha|x; \beta, \theta)$.
If we assume that $u(\alpha) = \mu(\alpha)$, then the estimator under L_S is obtained as $\hat{P}_{BS} = E(\mu(\alpha)|\underline{x})$ and

$$\hat{P}_{BS} = \frac{1}{C}\int_0^\infty \left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta + \bar{b})\alpha^{n-1}\beta^{k\alpha}\theta^{n\alpha}\left(\prod_{i=1}^n x_i\right)^{-(\alpha+1)} d\alpha.$$

When the loss function is L_L, we have

$$\hat{P}_{BL} = -\frac{1}{m}\log\left(E(e^{-m\mu(\alpha)}|\underline{x})\right), \qquad m \neq 0,$$

where

$$E(e^{-m\mu(\alpha)}|\underline{x}) = \frac{1}{C}\int_0^\infty e^{-m\left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta+\bar{b})}\alpha^{n-1}\beta^{k\alpha}\theta^{n\alpha}\left(\prod_{i=1}^n x_i\right)^{-(\alpha+1)} d\alpha,$$

and under the loss function L_E, the Bayesian estimator of $\mu(\alpha)$ is obtained to be

$$\hat{P}_{BE} = \left(E\left(\mu^{-q}(\alpha)|\underline{x}\right)\right)^{-\frac{1}{q}},$$

where

$$E\left(\mu^{-q}(\alpha)|\underline{x}\right) = \frac{1}{C}\int_0^\infty \left(\frac{\alpha\theta}{\alpha-1}(b\beta+\bar{b})\right)^{-q}\alpha^{n-1}\beta^{k\alpha}\theta^{n\alpha}\left(\prod_{i=1}^n x_i\right)^{-(\alpha+1)} d\alpha.$$

Note that the estimators described above are of the form of ratio of two integrals, which cannot be calculated into a closed form. In this situation, we need to use the Lindley approximation method that was developed by Lindley (1980) for obtaining the approximate Bayesian estimators in a form containing no integrals.

Let $u(\alpha)$ be a function of α, and we want to obtain Bayesian estimator by using $\pi(\alpha)$ as a prior distribution.

The likelihood function for the Pareto distribution contaminated by k outliers is given by (3.1). Hence, the Bayesian estimator of $u(\alpha)$ using the Lindley approximation is obtained as follows:

$$E(u(\alpha)|\underline{x};\beta,\theta) = \int_0^\infty \left(\frac{u(\alpha)\pi(\alpha)L(\alpha|\underline{x};\beta,\theta)}{\int_0^\infty \pi(\alpha)L(\underline{x};\beta,\theta|\alpha)d\alpha}\right)d\alpha.$$

Let $\varphi(\alpha) = \log(\pi(\alpha))$; then

$$E(u(\alpha)|\underline{x}) \approx \left(u(\alpha) + u_1\varphi_1\tau_{11} + \frac{1}{2}u_{11}\tau_{11} + \frac{1}{2}L_{111}u_1\tau_{11}^2\right)_{\hat{\alpha}_{ml}}, \tag{4.4}$$

where

$$\varphi_1 = \frac{\partial\varphi(\alpha)}{\partial\alpha}, \quad u_1 = \frac{\partial u(\alpha)}{\partial\alpha}, \quad u_{11} = \frac{\partial^2 u(\alpha)}{\partial\alpha^2},$$

$$L_{11} = \frac{\partial^2\ell}{\partial\alpha^2}, \quad L_{111} = \frac{\partial^3\ell}{\partial\alpha^3}, \quad \tau_{11} = -(L_{11})^{-1}.$$

Now, we can obtain the values of the Bayesian estimates of P under different loss functions.

a) Under the squared error loss function:

By substituting $u = \mu(\alpha) = \left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta + \bar{b})$ in equation (4.4), the approximate Bayesian estimator of P is given by

$$\hat{P}_{BS} \simeq \left\{\left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta+\bar{b}) - \frac{1}{\alpha}u_1\tau_{11} + \frac{1}{2}u_{11}\tau_{11} + \frac{1}{2}L_{111}u_1\tau_{11}^2\right\}_{\hat{\alpha}_{ml}}.$$

Thus

$$\hat{P}_{BS} \simeq \frac{\hat{\alpha}_{ml}\theta(b\beta+\bar{b})}{\hat{\alpha}_{ml}-1}\left(1 + \frac{\hat{\alpha}_{ml}}{n(\hat{\alpha}_{ml}-1)^2}\right).$$

b) Under the LINEX loss function:

By replacing $u = e^{-m(\mu(\alpha))} = e^{-m\left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta+\bar{b})}$ in equation (4.4), the approximate Bayesian estimator of P is obtained as

$$\hat{P}_{BL} \simeq -\frac{1}{m}\log\left(e^{-m\left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta+\bar{b})} - \frac{1}{\alpha}u_1\tau_{11} + \frac{1}{2}u_{11}\tau_{11} + \frac{1}{2}L_{111}u_1\tau_{11}^2\right)_{\hat{\alpha}_{ml}},$$

or equivalently as

$$\hat{P}_{BL} \simeq -\frac{1}{m}\log\left(e^{-m\theta\left(\frac{\hat{\alpha}_{ml}}{\hat{\alpha}_{ml}-1}\right)(b\beta+\bar{b})}\left(1 + \frac{\hat{\alpha}_{ml}^2 m^2\theta^2(b\beta+\bar{b})^2}{2n(\hat{\alpha}_{ml}-1)^4} - \frac{\hat{\alpha}_{ml}^2 m\theta(b\theta+\bar{b})}{n(\hat{\alpha}_{ml}-1)^3}\right)\right).$$

c) Under the entropy loss function:

Similar to what has been done in Part a and b, by substituting $u = \mu^{-q}(\alpha) = \left(\left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta+\bar{b})\right)^{-q}$ in equation (4.4), the Bayesian estimator of P is as the following from:

$$\hat{P}_{BE} \simeq \left(\left(\left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta+\bar{b})\right)^{-q} - \frac{1}{\alpha}u_1\tau_{11} + \frac{1}{2}u_{11}\tau_{11} + \frac{1}{2}L_{111}u_1\tau_{11}^2\right)_{\hat{\alpha}_{ml}}^{-\frac{1}{q}}$$

or

$$\hat{P}_{BE} \simeq \left(\left(\frac{\hat{\alpha}_{ml}}{\hat{\alpha}_{ml}-1}\right)\theta(b\beta+\bar{b})\left(1 - \frac{q\hat{\alpha}_{ml}}{n(\hat{\alpha}_{ml}-1)^2} + \frac{q^2+q}{2n(\hat{\alpha}_{ml}-1)^2}\right)\right)^{-\frac{1}{q}}.$$

4.2 Bayesian Premium Estimator based on Informative Prior Distribution

Assume that the informative prior distribution for the shape parameter α is a gamma distribution, such as

$$\pi(\alpha,\eta,\nu) = \frac{\nu^\eta}{\Gamma(\eta)}\alpha^{\eta-1}e^{-\nu\alpha}, \qquad \alpha > 0, \quad (\eta,\nu > 0). \tag{4.5}$$

By using equations (3.1) and (4.5), we get the posterior distribution of α as follows.

$$\pi(\alpha|\underline{x},\beta,\theta)=\frac{1}{C'}\alpha^{n+\eta-1}e^{-\nu\alpha}\beta^{k\alpha}\theta^{n\alpha}\left(\prod_{i=1}^{n}x_i\right)^{-(\alpha+1)},$$

where $C'=\int_0^\infty \alpha^{n+\eta-1}e^{-\nu\alpha}\beta^{k\alpha}\theta^{n\alpha}\left(\prod_{i=1}^{n}x_i\right)^{-(\alpha+1)}d\alpha$.

First we obtain the Bayesian estimator of P under the loss function L_S, using the posterior distribution $\pi(\alpha|x;\beta,\theta)$.

The estimator is obtained as $\hat{P}_{BS}=E(\mu(\alpha)|\underline{x})$ and

$$\hat{P}_{BS}=\frac{1}{C'}\int_0^\infty\left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta+\bar{b})\alpha^{n+\eta-1}e^{-\nu\alpha}\beta^{k\alpha}\theta^{n\alpha}\left(\prod_{i=1}^{n}x_i\right)^{-(\alpha+1)}d\alpha.$$

When the loss function is L_L, we have

$$\hat{P}_{BS}=-\frac{1}{m}\log\left(E(e^{-m\mu(\alpha)}|\underline{x})\right),\quad m\neq 0,$$

where

$$E\left(e^{-m\mu(\alpha)}|\underline{x}\right)=\frac{1}{C'}\int_0^\infty e^{-m\left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta+\bar{b})}\alpha^{n+\eta-1}e^{-\nu\alpha}\beta^{k\alpha}\theta^{n\alpha}\left(\prod_{i=1}^{n}x_i\right)^{-(\alpha+1)}d\alpha.$$

Under the loss function L_E, the Bayesian estimator of $\mu(\alpha)$ is obtained to be

$$\hat{P}_{BE}=\left(E(\mu^{-q}(\alpha)|\underline{x})\right)^{-\frac{1}{q}}$$

and

$$E(\mu^{-q}(\alpha)|\underline{x})=\frac{1}{C'}\int_0^\infty\left(\left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta+\bar{b})\right)^{-q}\alpha^{n+\eta-1}e^{-\nu\alpha}\beta^{k\alpha}\theta^{n\alpha}\left(\prod_{i=1}^{n}x_i\right)^{-(\alpha+1)}d\alpha.$$

The above integrals cannot be simplified into a closed form. By using the Lindley approximation, we obtain the approximate Bayesian estimators.

a) Under the squared error loss function:

By substituting $u=\mu(\alpha)=\left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta+\bar{b})$ in equation (4.4), the approximate Bayesian estimator of P is given by

$$\hat{P}_{BS}\simeq\left(\left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta+\bar{b})+\left(\frac{\eta-1}{\alpha}-\nu\right)u_1\tau_{11}+\frac{1}{2}u_{11}\tau_{11}+\frac{1}{2}L_{111}u_1\tau_{11}^2\right)_{\hat{\alpha}_{ml}}.$$

Then

$$\hat{P}_{BS}\simeq\frac{\hat{\alpha}_{ml}}{\hat{\alpha}_{ml}-1}\theta(b\beta+\bar{b})\left(1+\frac{\nu\hat{\alpha}_{ml}-\eta}{\hat{\alpha}_{ml}-1}+\frac{\hat{\alpha}_{ml}}{n(\hat{\alpha}_{ml}-1)^2}\right).$$

b) Under the LINEX loss function:
Substituting $u = e^{-m\mu(\alpha)} = e^{-m((\frac{\alpha}{\alpha-1})\theta(b\beta+\bar{b}))}$ in equation (4.4), we can get the approximate Bayesian estimator of P as follows:

$$\hat{P}_{BL} \simeq -\frac{1}{m}\log\left(e^{-m(\frac{\alpha}{\alpha-1})\theta(b\beta+\bar{b})} + \left(\frac{\eta-1}{\alpha} - \nu\right)u_1\tau_{11} + \frac{1}{2}u_{11}\tau_{11} + \frac{1}{2}L_{111}u_1\tau_{11}^2\right)_{\hat{\alpha}_{ml}}.$$

Then

$$\hat{P}_{BL} \simeq -\frac{1}{m}\log\Bigg(e^{-m\theta\left(\frac{\hat{\alpha}_{ml}}{\hat{\alpha}_{ml}-1}\right)(b\beta+\bar{b})}\Bigg(1 + \frac{\hat{\alpha}_{ml}^2 c\theta(b\beta+\bar{b})(\eta-\nu\hat{\alpha}_{ml})}{n(\hat{\alpha}_{ml}-1)^2} + \frac{\hat{\alpha}_{ml}^2 c\theta(b\beta+\bar{b})}{n(\hat{\alpha}_{ml}-1)^2} + \frac{\hat{\alpha}_{ml}^2 c^2\theta^2(b\beta+\bar{b})}{2n(\hat{\alpha}_{ml}-1)^4}\Bigg)\Bigg).$$

c) Under the entropy loss function:
By replacing $u = (\mu(\alpha))^{-q} = \left(\left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta+\bar{b})\right)^{-q}$ in equation (4.4), $\hat{P}_{BE}$ is obtained to be

$$\hat{P}_{BE} \simeq \left(\left(\left(\frac{\alpha}{\alpha-1}\right)\theta\left(b\beta+\bar{b}\right)\right)^{-q} + \left(\frac{\eta-1}{\alpha} - \nu\right)u_1\tau_{11} + \frac{1}{2}u_{11}\tau_{11} + \frac{1}{2}L_{111}u_1\tau_{11}^2\right)^{-\frac{1}{q}}$$

or equivalently

$$\hat{P}_{BE} \simeq \left(\left(\frac{\hat{\alpha}_{ml}}{\hat{\alpha}_{ml}-1}\right)\theta(b\beta+\bar{b})\right)\left(1 + \frac{q(\eta-\nu\hat{\alpha}_{ml})}{n(\hat{\alpha}_{ml}-1)} - \frac{q\hat{\alpha}_{ml}}{n(\hat{\alpha}_{ml}-1)^2} + \frac{q^2+q}{2n(\hat{\alpha}_{ml}-1)^2}\right).$$

4.3 Importance Sampling Procedure

In the previous section, we use the Lindley approximation for obtaining the Bayesian estimators under the loss functions, such as squared error loss, LINEX, and entropy. This section applies the importance sampling for approximating the Bayesian Premium estimator. We know that the posterior distribution $\pi(\alpha|x;\beta,\theta)$ is determined as $\pi(\alpha|x;\beta,\theta) = \frac{1}{C}L(\alpha|x;\beta,\theta)\pi(\alpha)$, where $C = \int L(\alpha|x;\beta,\theta)\pi(\alpha)d\alpha$ and also $\pi(\alpha)$ and $L(\alpha|x;\beta,\theta)$ are prior distribution and likelihood function, respectively.

Let $u(\alpha)$ be a function of α. We want to determine the $E(u(\alpha))$, where

$$E(u(\alpha)) = \int u(\alpha)\pi(\alpha|x;\beta,\theta)d\alpha. \tag{4.6}$$

Most of the time, equation (4.6) cannot be solved, so by using the Monte Carlo method, we generate the random sample $\alpha_1, \alpha_2, \ldots, \alpha_N$ from $\pi(\alpha|x;\beta,\theta)$, and $\bar{u}_N$ is defined as an approximate for $E(u(\alpha))$ that is given by

$$\bar{u}_N = \frac{1}{N}\sum_{i=1}^{N} u(\alpha_i).$$

Indeed, in most cases, the posterior distribution cannot be obtained in a closed form and cannot be sampled directly from the posterior distribution, so in this condition, we use the importance sampling. This method is based on generating a random sample of a proposed distribution with density function $g(\alpha)$.

Let $\alpha_1, \alpha_2, \dots, \alpha_N$ be a random sample from $g(\alpha)$. Then $E(u(\alpha))$ can be estimated by

$$\bar{u}(\alpha) = \frac{\sum_{i=1}^{N} u(\alpha_i)h(\alpha_i)}{\sum_{i=1}^{N} h(\alpha_i)},$$

where $h(\alpha_i) = \frac{L(\alpha_i|\underline{x};\beta,\theta)\pi(\alpha_i)}{g(\alpha_i)}$, for more details about this method, one can refer to Rubinstein (1981).

1- **Bayesian Premium Estimator based on non-informative Prior Distribution**

Let $u = \mu(\alpha) = \left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta + \bar{b})$ and $\pi(\alpha) = \frac{\sqrt{n}}{\alpha}$, that is, $\pi(\alpha|x;\beta,\theta)$ is

$$\begin{aligned} \pi(\alpha|x;\beta,\theta) &\approx L(\alpha|\underline{x};\beta,\theta)\pi(\alpha) \\ &\approx \alpha^{n-1}\beta^{k\alpha}\theta^{n\alpha}\left(\prod_{i=1}^{n} x_i\right)^{-(\alpha+1)}. \end{aligned} \tag{4.7}$$

Then

$$\pi(\alpha|x;\beta,\theta) \approx \Gamma_\alpha\left(n, \sum_{i=1}^{n}\log(x_i)\right) \times h(\alpha),$$

where $h(\alpha) = \frac{\beta^{k\alpha}\theta^{n\alpha}}{\prod_{i=1}^{n} x_i}$ and $\Gamma_\alpha\left(n, \sum_{i=1}^{n}\log(x_i)\right)$ is the gamma distribution with parameters n and $\sum_{i=1}^{n}\log(x_i)$.

To obtain the estimators, we use the following procedure:
Step 1: Generate α from the gamma distribution with parameters n and $\sum_{i=1}^{n}\log(x_i)$.
Step 2: Repeat the previous step N times to determine $(\alpha_1, \alpha_2, \dots, \alpha_N)$.
So $\hat{P}_{BS}, \hat{P}_{BL}$, and $\hat{P}_{BE}$ are given, respectively, as

$$\begin{aligned} \hat{P}_{BS} &\approx \frac{\sum_{i=1}^{N} \mu(\alpha_i)h(\alpha_i)}{\sum_{i=1}^{N} h(\alpha_i)}, \\ \hat{P}_{BL} &\approx -\frac{1}{C}\log\left[\frac{\sum_{i=1}^{N} e^{-c\mu(\alpha_i)}h(\alpha_i)}{\sum_{i=1}^{N} h(\alpha_i)}\right], \end{aligned} \tag{4.8}$$

and

$$\hat{P}_{BE} \approx \left[\frac{\sum_{i=1}^{N}(\mu(\alpha_i))^{-q}h(\alpha_i)}{\sum_{i=1}^{N} h(\alpha_i)}\right]^{-\frac{1}{q}}.$$

2- **Bayesian Premium Estimator based on Informative Prior Distribution**

If we consider $u = \mu(\alpha) = \left(\frac{\alpha}{\alpha-1}\right)\theta(b\beta + \bar{b})$ and $\pi(\alpha,\eta,\nu) = \frac{\nu^\eta}{\Gamma(\eta)}\alpha^{\eta-1}e^{-\nu\alpha}$, then $\pi(\alpha|x;\beta,\theta)$ is obtained as

$$\pi(\alpha|x;\beta,\theta) \approx \alpha^{n+\eta-1}e^{-\nu\alpha}\beta^{k\alpha}\theta^{n\alpha}\left(\prod_{i=1}^{n} x_i\right)^{-(\alpha+1)}.$$

Therefore,

$$\pi(\alpha|x;\beta,\theta) \approx \Gamma_\alpha\left(n, \sum_{i=1}^{n}\log(x_i)\right) \times s(\alpha),$$

where $h(\alpha) = \frac{\alpha^\eta e^{-\nu\alpha}\beta^{k\alpha}\theta^{n\alpha}}{\prod_{i=1}^{n} x_i}$.

The rest of calculations are the same as before.

5 Numerical Comparisons

In the previous section, several approximate Bayesian Premium estimators of the Pareto distribution in the presence of outliers are obtained. In this section, the performance of all these estimates is numerically compared in terms of average values and mean squared errors (MSEs). We have generated 1000 samples of size $n = 20, 40, 60, 80, 100$ from the Pareto distribution contaminated by k outliers with $k = 1, 2, 3$, $\alpha = 3, 10$, $\beta = 1.5$, and $\theta = 1$, and hyper parameters are assigned by values as $\nu = 0.1$ and $\eta = 0.3, 1$. In each case, $\hat{P}_{BL}$ is computed for $c = 0.1$, and $\hat{P}_{BE}$ is obtained for $q = 0.5$.

Table 1: Average values and MSEs of Bayesian Premium estimators ($\alpha = 3, \eta = 0.3, \nu = 0.1, k = 1$)

n	P	Prior function	Lindley method			Importance sampling method		
			$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$	$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$
20	1.5375	Gamma	1.631653	1.629634	1.613003	1.689287	1.594239	1.547081
			(0.065727)	(0.06418)	(0.053801)	(0.124505)	(0.035341)	(0.019723)
		Jeffrey	1.632401	1.630405	1.613853	1.696184	1.595364	1.547459
			(0.067566)	(0.066036)	(0.055628)	(0.140755)	(0.03619)	(0.020396)
40	1.51875	Gamma	1.558451	1.557641	1.550613	1.562448	1.561119	1.551823
			(0.018322)	(0.018124)	(0.016654)	(0.019442)	(0.019034)	(0.016899)
		Jeffrey	1.55859	1.557781	1.55076	1.562934	1.561582	1.552173
			(0.018577)	(0.018379)	(0.016906)	(0.019815)	(0.01939)	(0.017195)
60	1.5125	Gamma	1.546861	1.546315	1.541584	1.548533	1.547807	1.542164
			(0.01348)	(0.013372)	(0.012568)	(0.013789)	(0.013633)	(0.012635)
		Jeffrey	1.546967	1.546422	1.541694	1.548778	1.548045	1.542362
			(0.013611)	(0.013504)	(0.012698)	(0.013952)	(0.013793)	(0.012777)
80	1.509375	Gamma	1.52361	1.523239	1.51996	1.524348	1.523897	1.520181
			(0.008279)	(0.008237)	(0.007923)	(0.008344)	(0.00829)	(0.007923)
		Jeffrey	1.523627	1.523256	1.519978	1.524429	1.523975	1.520242
			(0.008339)	(0.008297)	(0.007982)	(0.008412)	(0.008358)	(0.007987)
100	1.5075	Gamma	1.528932	1.528622	1.525897	1.529259	1.528897	1.525889
			(0.007199)	(0.007164)	(0.006892)	(0.007254)	(0.007211)	(0.006905)
		Jeffrey	1.528974	1.528664	1.525939	1.529342	1.528979	1.52596
			(0.007242)	(0.007207)	(0.006935)	(0.007302)	(0.007259)	(0.00695)

Table 2: Average values and MSEs of Bayesian Premium estimators ($\alpha = 3, \eta = 0.3, \nu = 0.1, k = 2$)

n	P	Prior function	Lindley method			Importance sampling method		
			$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$	$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$
20	1.575	Gamma	1.667605	1.665397	1.648047	1.727592	1.614922	1.563853
			(0.083337)	(0.080758)	(0.065556)	(0.212468)	(0.034321)	(0.017663)
		Jeffrey	1.668357	1.666184	1.648928	1.73289	1.614912	1.563881
			(0.085873)	(0.083376)	(0.068124)	(0.24701)	(0.035914)	(0.01823)
40	1.5375	Gamma	1.59034	1.589438	1.581854	1.59446	1.59287	1.581741
			(0.024222)	(0.023938)	(0.021883)	(0.025701)	(0.025022)	(0.021438)
		Jeffrey	1.590584	1.589684	1.58211	1.595095	1.593467	1.5822
			(0.02457)	(0.02428)5	(0.022225)	(0.026217)	(0.025504)	(0.021823)
60	1.525	Gamma	1.549217	1.548692	1.544145	1.550628	1.549936	1.544532
			(0.012007)	(0.011919)	(0.011272)	(0.012265)	(0.012138)	(0.01133)
		Jeffrey	1.549272	1.548748	1.544203	1.550812	1.550114	1.544674
			(0.012123)	(0.012035)	(0.011387)	(0.012408)	(0.012278)	(0.011456)
80	1.51875	Gamma	1.54028	1.539889	1.536481	1.541032	1.540555	1.536686
			(0.009223)	(0.009171)	(0.008781)	(0.009318)	(0.009251)	(0.008796)
		Jeffrey	1.540324	1.539934	1.536527	1.541144	1.540665	1.536777
			(0.00929)	(0.00923)8	(0.008847)	(0.009396)	(0.009328)	(0.008868)
100	1.515	Gamma	1.530608	1.530304	1.527634	1.531127	1.530773	1.527827
			(0.007062)	(0.007031)	(0.006794)	(0.00713)	(0.007092)	(0.006825)
		Jeffrey	1.530632	1.530328	1.527658	1.531192	1.530836	1.527879
			(0.007104)	(0.007073)	(0.006835)	(0.007177)	(0.007139)	(0.006869)

Table 3: Average values and MSEs of Bayesian Premium estimators ($\alpha = 3, \eta = 0.3, \nu = 0.1, k = 3$)

n	P	Prior function	Lindley method			Importance sampling method		
			$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$	$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$
20	1.6125	Gamma	1.702266	1.700172	1.683604	1.737625	1.648582	1.590246
			(0.063201)	(0.061723)	(0.052156)	(0.216315)	(0.030942)	(0.015891)
		Jeffrey	1.702906	1.700831	1.684333	1.741751	1.649396	1.590111
			(0.064927)	(0.063463)	(0.053868)	(0.249763)	(0.031708)	(0.016385)
40	1.55625	Gamma	1.59793	1.597071	1.589841	1.60178	1.600378	1.590855
			(0.020016)	(0.019791)	(0.018174)	(0.021164)	(0.020712)	(0.018391)
		Jeffrey	1.598081	1.597224	1.590001	1.602291	1.600864	1.591224
			(0.020297)	(0.020073)	(0.018451)	(0.021573)	(0.021103)	(0.018715)
60	1.5375	Gamma	1.562019	1.561485	1.556902	1.563606	1.562902	1.557453
			(0.012277)	(0.012186)	(0.011523)	(0.012588)	(0.012456)	(0.011635)
		Jeffrey	1.562075	1.561542	1.55696	1.563789	1.563079	1.557593
			(0.012396)	(0.012305)	(0.011641)	(0.012733)	(0.0126)	(0.011764)
80	1.552812	Gamma	1.547914	1.547524	1.544137	1.548757	1.548283	1.544447
			(0.00852)	(0.00847)2	(0.008111)	(0.00864)	(0.008578)	(0.008155)
		Jeffrey	1.547951	1.547561	1.544175	1.548861	1.548384	1.54453
			(0.008582)	(0.008534)	(0.008173)	(0.008712)	(0.008649)	(0.008222)
100	1.5225	Gamma	1.53385	1.533551	1.530927	1.534252	1.533905	1.53102
			(0.006394)	(0.006368)	(0.006169)	(0.006456)	(0.006423)	(0.0062)
		Jeffrey	1.533861	1.533562	1.530938	1.534303	1.533955	1.53106
			(0.006432)	(0.006405)	(0.006206)	(0.006498)	(0.006465)	(0.006239)

Table 4: Average values and MSEs of Bayesian Premium estimators ($\alpha = 10, \eta = 1, \nu = 0.1, k = 1$)

n	P	Prior function	Lindley method			Importance sampling method		
			$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$	$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$
20	1.138889	Gamma	1.145045	1.145004	1.144511	1.145226	1.145171	1.144542
			(0.000966)	(0.000964)	(0.000941)	(0.000971)	(0.000968)	(0.000939)
		Jeffrey	1.145038	1.144997	1.144506	1.128249	1.128229	1.127987
			(0.001061)	(0.001059)	(0.001035)	(0.000406)	(0.000406)	(0.000401)
40	1.125	Gamma	1.128249	1.128229	1.127987	1.128318	1.128295	1.128021
			(0.000406)	(0.000406)	(0.000401)	(0.000408)	(0.000407)	(0.000402)
		Jeffrey	1.128252	1.128232	1.12799	1.128416	1.128393	1.128111
			(0.000426)	(0.000426)	(0.000421)	(0.000429)	(0.000429)	(0.000423)
60	1.12037	Gamma	1.123974	1.123961	1.123796	1.124007	1.123993	1.123814
			(0.0002670	(0.000267)	(0.000264)	(0.000267)	(0.000267)	(0.000264)
		Jeffrey	1.123999	1.123986	1.123821	1.124075	1.12406	1.123878
			(0.000276)	(0.000276)	(0.000273)	(0.000277)	(0.000276)	(0.000273)
80	1.118056	Gamma	1.119723	1.119713	1.119593	1.119729	1.119719	1.119591
			(0.000192)	(0.0001920	(0.000191)	(0.000193)	(0.000193)	(0.000191)
		Jeffrey	1.119724	1.119714	1.119594	1.119752	1.119742	1.119613
			(0.000197)	(0.000197)	(0.000196)	(0.000198)	(0.000198)	(0.000196)
100	1.116667	Gamma	1.118197	1.118189	1.118093	1.118222	1.118214	1.118113
			(0.000158)	(0.000158)	(0.000157)	(0.000158)	(0.000158)	(0.000157)
		Jeffrey	1.1182	1.118192	1.118096	1.11824	1.118232	1.118129
			(0.000161)	(0.000161)	(0.00016)	(0.000161)	(0.000161)	(0.00016)

Table 5: Average values and MSEs of Bayesian Premium estimators ($\alpha = 10, \eta = 1, \nu = 0.1, k = 2$)

n	P	Prior function	Lindley method			Importance sampling method		
			$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$	$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$
20	1.166667	Gamma	1.174043	1.174	1.17349	1.175489	1.175433	1.174811
			(0.000942)	(0.00094)	(0.000916)	(0.000959)	(0.000956)	(0.000926)
		Jeffrey	1.174085	1.174042	1.173535	1.176185	1.176125	1.175459
			(0.001033)	(0.001031)	(0.001007)	(0.001059)	(0.001056)	(0.00102)
40	1.138889	Gamma	1.142028	1.142008	1.141763	1.14214	1.142118	1.141845
			(0.000423)	(0.000422)	(0.000417)	(0.000425)	(0.000425)	(0.000419)
		Jeffrey	1.142027	1.142007	1.141763	1.142255	1.142231	1.141951
			(0.000444)	(0.000443)	(0.000438)	(0.000447)	(0.000447)	(0.00044)
60	1.12963	Gamma	1.132115	1.132102	1.131939	1.132132	1.132117	1.131941
			(0.000273)	(0.000273)	(0.00027)	(0.000273)	(0.000273)	(0.00027)
		Jeffrey	1.132122	1.132108	1.131946	1.132183	1.132168	1.131988
			(0.000282)	(0.000282)	(0.000279)	(0.000283)	(0.000282)	(0.00028)
80	1.125	Gamma	1.126994	1.126984	1.126862	1.127086	1.127076	1.126947
			(0.000203)	(0.000203)	(0.000202)	(0.000204)	(0.000204)	(0.000203)
		Jeffrey	1.126999	1.126989	1.126868	1.127117	1.127106	1.126976
			(0.000208)	(0.000208)	(0.000207)	(0.00021)	(0.00021)	(0.000208)
100	1.122222	Gamma	1.123566	1.123559	1.123462	1.123601	1.123592	1.123492
			(0.00015)	(0.00015)	(0.000149)	(0.00015)	(0.000149)	(0.000149)
		Jeffrey	1.123567	1.123559	1.123463	1.123616	1.123608	1.123507
			(0.000153)	(0.000153)	(0.000152)	(0.000153)	(0.000152)	(0.000152)

Table 6: Average values and MSEs of Bayesian Premium estimators ($\alpha = 10, \eta = 1, \nu = 0.1, k = 3$)

n	P	Prior function	Lindley method			Importance sampling method		
			$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$	$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$
20	1.194444	Gamma	1.20236	1.202315	1.201792	1.207568	1.207515	1.206493
			(0.000961)	(0.000959)	(0.000934)	(0.001033)	(0.00103)	(0.000997)
		Jeffrey	1.20242	1.202374	1.201854	1.208532	1.208475	1.20786
			(0.001053)	(0.001051)	(0.001026)	(0.001133)	(0.001129)	(0.00109)
40	1.152778	Gamma	1.158617	1.158595	1.158336	1.15976	1.159737	1.159467
			(0.000464)	(0.000464)	(0.000457)	(0.000485)	(0.000484)	(0.000476)
		Jeffrey	1.158682	1.15866	1.158401	1.160006	1.159982	1.159704
			(0.000487)	(0.000486)	(0.000479)	(0.000509)	(0.000508)	(0.000499)
60	1.138889	Gamma	1.14034	1.140327	1.140165	1.140708	1.140694	1.140526
			(0.000292)	(0.000291)	(0.000289)	(0.000295)	(0.000295)	(0.000293)
		Jeffrey	1.140329	1.140316	1.140155	1.140764	1.14075	1.140579
			(0.000301)	(0.000301)	(0.000299)	(0.000305)	(0.000305)	(0.000302)
80	1.131944	Gamma	1.133427	1.133417	1.133296	1.133627	1.133616	1.133491
			(0.000212)	(0.000212)	(0.000211)	(0.000212)	(0.000212)	(0.00021)
		Jeffrey	1.133426	1.133416	1.133295	1.13366	1.13365	1.133522
			(0.000218)	(0.000217)	(0.000216)	(0.000217)	(0.000217)	(0.000215)
100	1.127778	Gamma	1.129294	1.129286	1.129189	1.129358	1.129349	1.129248
			(0.000153)	(0.000153)	(0.000152)	(0.000152)	(0.000152)	(0.000151)
		Jeffrey	1.129296	1.129288	1.129191	1.129379	1.129371	1.129269
			(0.000156)	(0.000156)	(0.000155)	(0.000155)	(0.000155)	(0.000154)

6 Actual Data Analysis

A real example is presented in this section for the purpose of illustration. This data set is related to a random sample of size 20 which is taken from the well-known Danish fire losses for claim of at least 1.5. The sample is as follows.

1.581612, 1.584488, 1.756955, 1.722223, 2.036376, 2.036378, 2.051958, 2.102489, 2.146618, 2.9238653, 3.263154, 3.367496, 4.530015, 4.856098, 5.417277, 5.563852, 6.319914, 7.320644, 9.174312, 56.225426.

Using one-sample Kolmogorov–Smirnov test ($D = 0.13697$ and $p = 0.7992$), the Pareto distribution with $\hat{\alpha} = 1.208316$ and $\theta = 1.5$ is fitted to the data. Histogram, Pareto Q-Q plot, boxplot and empirical distribution are shown in Figure 1. In this data set, $x = 56.225426$ is the largest value and it maybe is an outlier. Therefore, the contaminated factor ($\beta = \frac{min(x_i)}{\theta}$) is obtained as 1.05. Also, the number of outliers (k) and shape parameter (α) are unknown. To obtain the Premium, k must be predetermined. So, we can use the profile likelihood function with respect to k and select k which maximizes the likelihood. Therefore, MLE of α and the values of likelihood function corresponding to k, $L(\hat{\alpha}_{ml}|\underline{x})$, are shown in Table 7.

According to Table 7, the likelihood function is maximized for k=1. So, in this example, $\hat{\alpha}_{ml} = 1.138773$. Therefore, the MLE and different Bayesian estimates of the Premium by using non-informative prior and Lindley approximation are presented in Table 8.

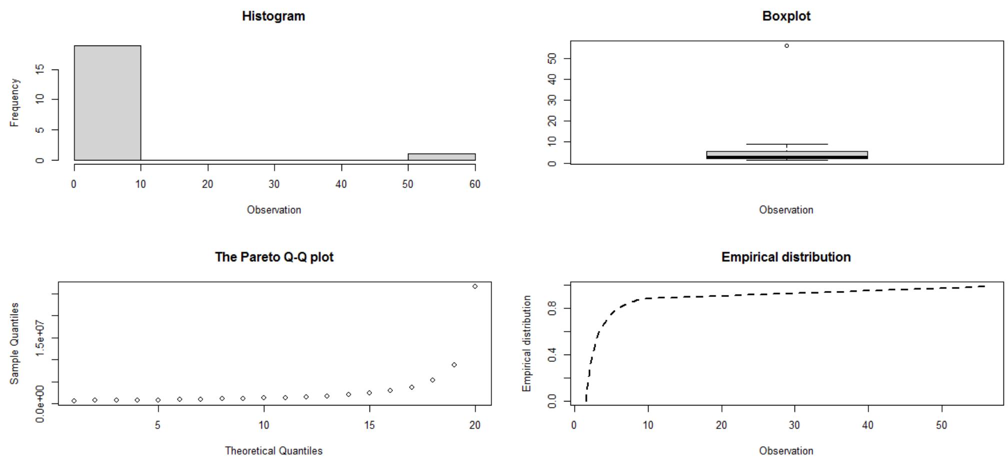

Figure 1: Histogram, Pareto Q-Q plot, boxplot and emprical distribution of data of example.

Table 7: The MLE of α and the likelihood functions corresponding to k for β=1.05 and θ=1.5.

k	$\hat{\alpha}_{ml}$	$L(\hat{\alpha}_{ml}\|\underline{x})$
1	1.138773	9.362670e-21
2	1.141945	1.041932e-21
3	1.145136	1.836196e-22
4	1.148344	4.569079e-23
5	1.151570	1.510238e-23

Table 8: The Maximum likelihood estimate and different Bayesian estimates of Premium.

$\hat{P}_{ml}$	$\hat{P}_{BS}$	$\hat{P}_{BL}$	$\hat{P}_{BE}$
9.02250	22.55625	27.63942	17.46510

7 Discussion and Conclusion

In this study, an attempt has been made to examine the Bayesian estimators for the Pareto distribution in the presence of outliers with insurance applications. The Bayesian estimators of Premium (P) were obtained under squared error, LINEX and entropy loss functions by using gamma and Jeffreys prior through the Lindley approximation and importance sampling procedure. By using the simulation study, in Tables 1–6, the average values and MSEs of the estimates $\hat{P}_{BS}$, $\hat{P}_{BL}$, and $\hat{P}_{BE}$ of P were presented for different choices of n and k. Tables 1–3 is calculated based on $\alpha = 3$, $\eta = 0.3$, $\nu = 0.1$ and Tables 4–6 is obtained for $\alpha = 10$, $\eta = 1$, $\nu = 0.1$. By comparing Tables 1–3 and 4–6 for $k = 1, 2, 3$ and $n = 20, 40, 60, 80, 100$, it can be seen that by increasing α, the value of P decreases. Besides, the MSEs of Bayesian Premium estimator decrease and tend to P. Also, we concluded that by using both Lindley approximation and importance sampling methods, the same result is obtained and that MSEs of Bayesian Premium estimator under entropy loss are smaller than the MSEs of other estimators under LINEX and squared error loss functions. It may be noted that when k increases, P also increases and with the increase in sample size, the Bayesian Premium estimator tends to P.

Also, based on the results of Tables 1–6, it has been seen that under different loss functions, using the gamma prior distribution is more appropriate than the Jeffreys prior distribution for all sample sizes. It may be mentioned that the proposed method can be extended for other positive value distributions.

Acknowledgment

The authors are thankful to the editors and the referees for their valuable comments.

References

Bailey, A. L. (1950), *Credibility Procedures: Laplace's generalization of Bayes' Rule and the combination of collateral knowledge with observed data*, New York State Insurance Department.

Barnett, V. and Lewis, T. (1994), *Outliers in Statistical Data*, 3rd ed., Wiley, New York.

Benktander, G. (1963), A note on the most "dangerous" and skewest class of distribution, *Astin Bulletin*, **2**, 87–390.

Dixit, U. J. (1987), Characterization of the gamma distribution in the presence of k outliers, *Bull. Bombay Mathematical Colloquium*, **4**, 54–59.

Dixit, U. J. (1989), Estimation of parameters of the gamma distribution in the presence of outliers, *Communications in Statistics-Theory and Methods*, **18**(8), 3071–3085.

Dixit, U. J., Moore, K. L. and Barnett, V. (1996), On the estimation of the scale parameter of the exponential distribution in the presence of outliers generated from uniform distribution, *Metron*, **LIV(3-4)**, 201–211.

Dixit, U. J. and Jabbari Nooghabi, M. (2011a), Efficient estimation in the Pareto distribution with the presence of outliers, *Stat Methodol*, **8**(4), 340–355.

Dixit, U. J. and Jabbari Nooghabi, M. (2011b), Efficient Estimation of the Parameters of the Pareto Distribution in the Presence of Outliers, *Communications of the Korean Statistical Society*, **18**(6), 817–835.

Grubbs, F. E. (1950). Sample criteria for testing outlying observations, *Ann. Math. Statist.*, **21**, 27–58.

Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples, *Technometrics*, **11**, 1–21.

Hassan Zadeh, A. and Stanford, D. A. (2016). Bayesian and Bühlmann credibility for phase-type distributions with a univariate risk parameter, *Scandinavian Actuarial Journal*, **2016**(4), 338–355.

Hawkins, D. M. (1980), *Identification of outliers*, Chapman and Hall, London.

Heilmann, W. R. (1985), Tolerance intervals in risk theory, *Insurance: Mathematics and Economics*, **4**(3), 173–177.

Jabbari Nooghabi, M. and Khaleghpanah Nooghabi, E. (2016), On entropy of a Pareto distribution in the presence of outliers, *Comm. Statist. Theory Methods*, **45**(17), 5234–5250.

Jabbari Nooghabi, M. (2019), On detecting outliers in the Pareto distribution, *Journal of Statistical Computation and Simulation*, **89**(8), 1466–1481.

Jeevanand, E. S., Nair, N. U. (1992), Bayes estimates of Pareto parameters in the presence of outliers, *Proceedings of the Symposium on Statistical Inference*, Trivandrum, India, 53–70.

Jeevanand, E. S., Nair, N. U. (1993), Prediction of future observations form Pareto population in the presence of outliers, *Statistica*, **53**, 171–176.

Jeevanand, E. S., Nair, N. U. (1996), Bayes prediction interval for the exponential and Pareto population in the presence of an outlier, *Far East J Math Sci.*, **1**(1), 81–95.

Jeevanand, E. S., Nair, N. U. (1998), On determining the number of outliers in exponential and Pareto samples, *Statist Papers*, **39**, 277–290.

Joshi, P. C. (1972), Efficient estimation of the mean of an exponential distribution when an outlier is present, *Technometrics*, **41**, 137–143.

Kale, B. K. and Sinha, S. K. (1971), Estimation of expected life in the presence of outlier observation, *Technometrics*, **13**(4), 755–759.

Kendall, M. G. and Buckland, W. R. (1957), *A Dictionary of Statistical Terms*, London: Longman.

Kiapour, A. (2018), Bayes, E-Bayes and robust Bayes premium estimation and prediction under the squared log error loss function, *Journal of the Iranian Statistical Society*, **17**(1), 33–47.

Li, C. P. and Hao, H. B. (2016), Likelihood and Bayesian estimation in stress strength model from generalized exponential distribution containing outliers, *IAENG International Journal of Applied Mathematics*, **46**(2), 155–159.

Lindley, D. V. (1980), Approximate Bayesian Methods, *Trabajos de Estadistica Y de Investigacion Operativa*, **31**(1), 223–245.

Miller, R. G. Jr. (1981), *Simultaneous Statistical Inference*, 2nd ed., Springer Verlag, New York.

Mowbray, A. H. (1914), How Extensive a Payroll Exposure is Necessary to Give a Dependable Pure Premium?, *Proceedings of the Casualty Actuarial Society*, **1**, 25–30.

Rastogi, M. K. and Merovci, F. (2018), Bayesian estimation for parameters and reliability characteristic of the Weibull Rayleigh distribution, *Journal of King Saud University-Science*, **30**(4), 472–478.

Rubinstein, R. Y. (1981), *Simulation and the Monte Carlo Method*, Wiley, New York.

Young, V. (2004), *Premium Principles In Encyclopedia of Actuarial Science* , Wiley, New York.