

ORIGINAL RESEARCH PAPER

The Fréchet distribution: Estimation and Application an Overview

Pedro L. Ramos¹, Francisco Louzada¹, Eduardo Ramos¹ and Sanku Dey²

ARTICLE HISTORY

Compiled July 11, 2019

ABSTRACT

In this article we consider the problem of estimating the parameters of the Fréchet distribution from both frequentist and Bayesian points of view. First we briefly describe different frequentist approaches, namely, maximum likelihood, method of moments, percentile estimators, L-moments, ordinary and weighted least squares, maximum product of spacings, maximum goodness-of-fit estimators and compare them using extensive numerical simulations. Next, we consider the Bayesian inference approach using reference priors. Five real data sets related to the minimum flow of water on Piracicaba river in Brazil are used to illustrate the applicability of the discussed procedures.

2010 AMS : 62F10; 62P12

KEYWORDS : Bayesian Inference; Fréchet distribution; Hydrological applications; Maximum product of spacings; Reference prior.

1. Introduction

Extreme value theory plays an important role in statistical analysis. The most used distribution to describe extreme data is the generalized extreme value (GEV) distribution [21]. Its cumulative density function (CDF) is given by

$$F(t|\sigma, \mu, \xi) = \begin{cases} \exp \left\{ -[1 + \xi(t - \mu)/\sigma]_+^{-1/\xi} \right\}, & \text{for } \xi \neq 0 \\ \exp \left\{ -\exp [-(t - \mu)/\sigma] \right\}, & \text{for } \xi = 0 \end{cases} \quad (1)$$

where $\sigma > 0$, $\mu, \xi \in \mathbb{R}$ and $x_+ = \max(x, 0)$. Gumbel, Weibull and Fréchet distributions are special cases of the so-called generalized extreme value (GEV) distribution. The Fréchet distribution is named after French mathematician Maurice René Fréchet, who developed it in the 1920s as a maximum value distribution (which is also known as the extreme value distribution of type II). Kotz and Nadarajah [25] describe this distribution and discussed its wide applicability in different spheres such as accelerated life testing, natural calamities, horse racing, rainfall, queues in supermarkets, sea currents, wind speeds, track race records and so on.

Let the random variable T follows Fréchet distribution then its probability density

¹ Institute of Mathematical and Computer Sciences, University of São Paulo, Avenida Trabalhador São-Carlense, 400, 13566-590, São Carlos, SP, Brazil

² Department of Statistics, St. Anthony's College, Shillong, Meghalaya, India

function (PDF) and cdf are given by

$$f(t|\lambda, \alpha) = \lambda \alpha t^{-(\alpha+1)} e^{-\lambda t^{-\alpha}} \quad \text{and} \quad F(t|\lambda, \alpha) = e^{-\lambda t^{-\alpha}}, \quad (2)$$

for all $t > 0$ and the quantities $\alpha > 0$ and $\lambda > 0$ are the shape and the scale parameters respectively. This distribution is also referred as Inverse Weibull distribution. The PDF can be unimodal or decreasing depending on the choice of the shape parameter while its hazard function is always unimodal.

Several researchers have studied different aspects of inferential procedures for the Fréchet distribution. From the classical perspective, Calabria and Pulcini [7] and Erto [14] discussed the properties of the maximum likelihood estimators (MLE) and the ordinary least-square estimators (LSE) respectively. Ramos et al. [37] presented the MLE for the the Fréchet distribution in the presence of cure fraction. Loganathan and Uma [28] compared the MLE, LSE, weighted LSE (WLSE) and the method of moments (MME) and outlined that the WLSE provided similar results. Salman et al. [41] and Maswadah [31] studied the Fréchet distribution in the context of order statistics and generalized order statistics respectively. The Bayes estimators were discussed by Calabria and Pulcini [8] and Kundu and Howlader [26] using informative or subjective priors such as Gamma priors (also known as flat priors). However, Bernardo [5] argued that the use of simple proper flat priors presumed to be non-informative often hide important unwarranted assumptions which may easily dominate, or even invalidate the statistical analysis and should be strongly discouraged. Recently, Abbas and Tang [1] studied Fréchet distribution based on Jeffreys and reference priors.

Parameter estimation is significant for any probability distribution and therefore various estimation methods are frequently studied in the statistical literature. Traditional estimation methods such as the MLE, MME, LSE and WLSE are often opted for parameter estimation. Each has its own merits and demerits but the most popular method of estimation is the maximum likelihood estimation method. Besides the above cited methods, we consider five additional methods to estimate the parameters of Fréchet distribution. These additional methods are the maximum product spacing estimator (MPS), percentile estimator (PE), Cramér-von-Mises estimator (CME), Anderson-Darling estimator (ADE) and L-moment (LME) estimator. Several authors have compared these different estimation methods for different distributions [2, 11–13, 27, 32, 36, 40]. From the Bayesian point of view different Bayes estimators are discussed using objective priors and different loss functions. Also, the coverage probability with a confidence level equal to 95% for the estimates are obtained. To evaluate the performance of the estimators, a simulation study is carried out. Finally, five real life data sets have been analyzed for illustrative purposes.

The objective of this paper is to estimate the parameters of the model from both frequentist and Bayesian perspective and to develop a guideline for choosing the best estimation method for the Fréchet distribution, which we would be of profound interest to applied statisticians. The present study is unique because of the fact that thus far, no attempt has been made to compare all these aforementioned estimators for the two-parameter Fréchet distribution. At the end, we present the better estimation procedure for the Fréchet distribution.

The paper is organized as follows. Section 2 describes nine frequentist methods of estimation. The Bayes estimators are presented in section 3. In Section 4, we present the Monte Carlo simulation results. In Section 5, the usefulness of the Fréchet distribution is illustrated by using five real data sets. Finally, concluding remarks are provided in Section 6.

2. Classical parameter estimation methods

In this section, we describe nine methods for estimating the parameters λ and α of the Fréchet distribution.

2.1. Maximum Likelihood Estimation

Among the statistical inference methods, the maximum likelihood method is widely used due to its desirable properties including consistency, asymptotic efficiency and invariance. Under the maximum likelihood method, the estimators are obtained by maximizing the log-likelihood function. Let T_1, \dots, T_n be a random sample such that $T \sim \text{Fréchet}(\lambda, \alpha)$. Then, the likelihood function from (2) is given by

$$L(\lambda, \alpha | t) = \prod_{i=1}^n f(t_i, \lambda, \alpha) = \lambda^n \alpha^n \left(\prod_{i=1}^n t_i^{-(\alpha+1)} \right) \exp \left(-\lambda \sum_{i=1}^n t_i^{-\alpha} \right). \quad (3)$$

The log-likelihood function (3) is given by

$$l(\lambda, \alpha | t) = n \log(\lambda) + n \log(\alpha) - (\alpha + 1) \sum_{i=1}^n \log(t_i) - \lambda \sum_{i=1}^n t_i^{-\alpha}.$$

From $\partial l(\lambda, \alpha | t) / \partial \lambda = 0$ and $\partial l(\lambda, \alpha | t) / \partial \alpha = 0$, we get the likelihood equations

$$\frac{n}{\lambda} - \sum_{i=1}^n t_i^{-\alpha} = 0 \quad (4)$$

and

$$\frac{n}{\alpha} - \sum_{i=1}^n \log(t_i) + \lambda \sum_{i=1}^n t_i^{-\alpha} \log(t_i) = 0, \quad (5)$$

whose solutions provide $\hat{\lambda}_{MLE}$ and $\hat{\alpha}_{MLE}$. After some algebraic manipulations, the estimate $\hat{\alpha}_{MLE}$ can be obtained by solving the following non-linear equation

$$\frac{n}{\alpha} - \sum_{i=1}^n \log(t_i) + \frac{n \sum_{i=1}^n t_i^{-\alpha} \log(t_i)}{\sum_{i=1}^n t_i^{-\alpha}} = 0. \quad (6)$$

The estimate $\hat{\lambda}_{MLE}$ can be obtained by substituting $\hat{\alpha}_{MLE}$ in

$$\hat{\lambda}_{MLE} = \frac{n}{\sum_{i=1}^n t_i^{-\alpha}}. \quad (7)$$

The obtained ML estimates are asymptotically normally distributed with a joint bivariate normal distribution given by

$$(\hat{\lambda}_{MLE}, \hat{\alpha}_{MLE}) \sim N_2[(\lambda, \alpha), I^{-1}(\lambda, \alpha)] \text{ for } n \rightarrow \infty,$$

where $I(\lambda, \alpha)$ is the Fisher information matrix given by

$$I(\lambda, \alpha) = \begin{bmatrix} \frac{n}{\lambda^2} & \frac{n(1 - \gamma - \log(\lambda))}{\lambda\alpha} \\ \frac{n(1 - \gamma - \log(\lambda))}{\lambda\alpha} & \frac{n}{\alpha^2} \left(\frac{\pi^2}{6} + (1 - \gamma - \log(\lambda))^2 \right) \end{bmatrix}, \quad (8)$$

and $\gamma \approx 0.5772156649$ is known as Euler-Mascheroni constant.

In the following we prove the existence and uniqueness of MLEs.

Theorem 2.1. *Let t_1, \dots, t_n be not all equal. Then the MLEs of the parameters α and λ are unique and are given by $\hat{\alpha}$ and*

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^n t_i^{-\hat{\alpha}}}, \quad (9)$$

where $\hat{\alpha}$ is the only solution of non-linear equation

$$G(\alpha) = \frac{n}{\alpha} - \sum_{i=1}^n \log t_i - \frac{n}{\sum_{i=1}^n t_i^{-\alpha}} \sum_{i=1}^n t_i^{-\alpha} \log t_i.$$

Proof. See Appendix A. □

2.2. Moments Estimators

The method of moments is one of the oldest method used for estimating the parameters of the statistical models. The raw moments of T for the Fréchet distribution is

$$E(T^r | \lambda, \alpha) = \lambda^{\frac{r}{\alpha}} \Gamma \left(1 - \frac{r}{\alpha} \right), \quad (10)$$

where $r \in N$ and $\Gamma(\lambda) = \int_0^\infty e^{-x} x^{\lambda-1} dx$ is the gamma function. Note that $E(T^r | \gamma, \alpha)$ does not have a finite value for $\alpha > r$. The moment estimators (MEs) for the Fréchet distribution can be obtained by equating the first two theoretical moments with the sample moments. However, instead of equating the first two theoretical moments, we consider that

$$E(T | \lambda, \alpha) = \lambda^{\frac{1}{\alpha}} \Gamma \left(1 - \frac{1}{\alpha} \right) \quad \text{and} \quad \text{Var}(T | \lambda, \alpha) = \lambda^{\frac{2}{\alpha}} \left(\Gamma \left(1 - \frac{2}{\alpha} \right) - \Gamma^2 \left(1 - \frac{1}{\alpha} \right) \right). \quad (11)$$

Therefore, the population coefficient of variation is given by

$$CV(X | \lambda, \alpha) = \frac{\sqrt{\text{Var}(T | \lambda, \alpha)}}{E(T | \lambda, \alpha)} = \sqrt{\frac{\Gamma(1 - 2\alpha^{-1})}{\Gamma^2(1 - \alpha^{-1})} - 1},$$

which is independent of the scale parameter λ . So, the estimator for $\hat{\alpha}_{ME}$ can be obtained by solving the following non-linear equation

$$\sqrt{\frac{\Gamma(1-2\alpha^{-1})}{\Gamma^2(1-\alpha^{-1})}} - 1 - \frac{s}{\bar{t}} = 0.$$

Substituting $\hat{\alpha}_{MME}$ in (11) the estimate $\hat{\lambda}_{MME}$ can be obtained by solving

$$\hat{\lambda}_{ME} = \frac{\bar{t}^\alpha}{\Gamma^\alpha(1-\alpha^{-1})}. \quad (12)$$

However, this estimator can only be computed for $\alpha > 2$ which is undesirable.

2.3. Percentile Estimator

The percentile estimator is a statistical method used to estimate the parameters by comparing the sample points with the theoretical points. This method was originally suggested by Kao [22, 23] and has been widely used for distributions that has the quantile function in closed form expression, such as the Weibull and the Generalized Exponential distribution. The quantile function of the Fréchet distribution has the closed form and is given by

$$Q(p|\lambda, \alpha) = \left(\frac{1}{\lambda} \log \left(\frac{1}{p} \right) \right)^{-\frac{1}{\alpha}}. \quad (13)$$

Therefore, the percentile estimates (PCEs), $\hat{\lambda}_{PCE}$ and $\hat{\alpha}_{PCE}$, can be obtained by minimizing

$$\sum_{i=1}^n \left(t_{(i)} - \left(\frac{1}{\lambda} \log \left(\frac{1}{p_i} \right) \right)^{-\frac{1}{\alpha}} \right)^2,$$

with respect to λ and α , where p_i denotes an estimate of $F(t_{(i)}; \lambda, \alpha)$ and $t_{(i)}$ is the i th order statistics (we assume the same notation for the next sections). The estimates of λ and α can be also be obtained by solving the following non-linear equations:

$$\sum_{i=1}^n \left[t_{(i)} - \left(\frac{1}{\lambda} \log \left(\frac{1}{p_i} \right) \right)^{-\frac{1}{\alpha}} \right] \left(\frac{1}{\lambda} \log \left(\frac{1}{p_i} \right) \right)^{-\frac{1}{\alpha}} = 0,$$

$$\sum_{i=1}^n \left[t_{(i)} - \left(\frac{1}{\lambda} \log \left(\frac{1}{p_i} \right) \right)^{-\frac{1}{\alpha}} \right] \log \left(\frac{1}{\lambda} \log \left(\frac{1}{p_i} \right) \right) \left(\frac{1}{\lambda} \log \left(\frac{1}{p_i} \right) \right)^{-\frac{1}{\alpha}} = 0,$$

respectively. In this paper, we consider $p_i = \frac{i}{n+1}$. However, several estimators of p_i can be used instead [30].

2.4. L-Moments Estimators

Hosking [19] proposed an alternative method of estimation analogous to conventional moments, namely L-moments estimators. L-moments estimators can be obtained by equating the sample with the population L-moments. Hosking [19] stated that the L-moment estimators are more robust than the usual moment estimators, and also relatively robust to the effects of outliers and reasonably efficient when compared to the MLE for some distributions.

For the Fréchet distribution, the L-moments estimators can be obtained by equating the first two sample L-moments with the corresponding population L-moments. The first two sample L-moments are

$$l_1 = \frac{1}{n} \sum_{i=1}^n t_{(i)}, \quad l_2 = \frac{2}{n(n-1)} \sum_{i=1}^n (i-1)t_{(i)} - l_1,$$

where $t_{(1)}, t_{(2)}, \dots, t_{(n)}$ denotes the order statistics of a random sample from a distribution function $F(t|\lambda, \alpha)$. The first two population L-moments are

$$\mu_1(\lambda, \alpha) = \int_0^1 Q(p|\lambda, \alpha) dp = E(X|\lambda, \alpha) = \lambda^{\frac{1}{\alpha}} \Gamma\left(1 - \frac{1}{\alpha}\right) \quad (14)$$

and

$$\mu_2(\lambda, \alpha) = \int_0^1 Q(p|\lambda, \alpha)(2p-1) dp = \lambda^{\frac{1}{\alpha}} \left(2\frac{1}{\alpha} - 1\right) \Gamma\left(1 - \frac{1}{\alpha}\right), \quad (15)$$

where $Q(p|\lambda, \alpha)$ is given in (13). After some algebraic manipulation, the estimator for $\hat{\alpha}_{LME}$ can be obtained as

$$\hat{\alpha}_{LME} = \frac{\log(2)}{\log(2) + \log\left(\sum_{i=1}^n (i-1)t_{(i)}\right) - \log(n(n-1)\bar{t})}. \quad (16)$$

Note that, substituting $\hat{\alpha}_{LME}$ in (14) the estimator for $\hat{\lambda}_{LME}$ can be obtained by solving

$$\hat{\lambda}_{LME} = \frac{\bar{t}^{\alpha}}{\Gamma^{\alpha}(1 - \alpha^{-1})}, \quad \alpha > 1. \quad (17)$$

It is worth noting that, among the chosen methods, the L-moments estimator was the only that has closed-form solution for both parameters.

2.5. Ordinary and Weighted Least-Square Estimate

The least square estimators $\hat{\lambda}_{LSE}$ and $\hat{\alpha}_{LSE}$, can be obtained by minimizing

$$S(\lambda, \alpha) = \sum_{i=1}^n \left[F(t_{(i)} | \theta, \lambda) - \frac{i}{n+1} \right]^2,$$

with respect to λ and α . Similarly, they can also be obtained by solving the following non-linear equations (see Erto [14] for more details):

$$\begin{aligned}\sum_{i=1}^n \left[F(t_{(i)} | \lambda, \alpha) - \frac{i}{n+1} \right] \eta_1(t_{(i)} | \lambda, \alpha) &= 0, \\ \sum_{i=1}^n \left[F(t_{(i)} | \lambda, \alpha) - \frac{i}{n+1} \right] \eta_2(t_{(i)} | \lambda, \alpha) &= 0,\end{aligned}$$

where

$$\eta_1(t_{(i)} | \lambda, \alpha) = t_{(i)}^{-\alpha} e^{-\lambda t_{(i)}^{-\alpha}} \text{ and } \eta_2(t_{(i)} | \lambda, \alpha) = \lambda t_{(i)}^{-\alpha} \log(t_{(i)}) e^{-\lambda t_{(i)}^{-\alpha}}. \quad (18)$$

The weighted least-squares estimators (WLSEs), $\hat{\lambda}_{WLSE}$ and $\hat{\alpha}_{WLSE}$, can be obtained by minimizing

$$W(\lambda, \alpha) = \sum_{i=1}^n \frac{(n+1)^2 (n+2)}{i(n-i+1)} \left[F(t_{(i)} | \lambda, \alpha) - \frac{i}{n+1} \right]^2.$$

The estimators can also be obtained by solving the following non-linear equations,

$$\begin{aligned}\sum_{i=1}^n \frac{(n+1)^2 (n+2)}{i(n-i+1)} \left[F(t_{(i)} | \lambda, \alpha) - \frac{i}{n+1} \right] \eta_1(t_{(i)} | \lambda, \alpha) &= 0, \\ \sum_{i=1}^n \frac{(n+1)^2 (n+2)}{i(n-i+1)} \left[F(t_{(i)} | \lambda, \alpha) - \frac{i}{n+1} \right] \eta_2(t_{(i)} | \lambda, \alpha) &= 0.\end{aligned}$$

2.6. Method of Maximum Product of Spacings

The maximum product of spacings (MPS) method is a powerful alternative to ML method for estimating the unknown parameters of continuous univariate distributions. This method was proposed by Cheng and Amin [9, 10], and later independently developed by Ranneby [38] as an approximation to the Kullback-Leibler measure of information.

Let $D_i(\lambda, \alpha) = F(t_{(i)} | \lambda, \alpha) - F(t_{(i-1)} | \lambda, \alpha)$, for $i = 1, 2, \dots, n+1$, be the uniform spacings of a random sample from the Fréchet distribution, where $F(t_{(0)} | \lambda, \alpha) = 0$, $F(t_{(n+1)} | \lambda, \alpha) = 1$ and $\sum_{i=1}^{n+1} D_i(\lambda, \alpha) = 1$. The maximum product of spacings estimators $\hat{\lambda}_{MPS}$ and $\hat{\alpha}_{MPS}$ are obtained by maximizing the geometric mean of the spacings

$$G(\lambda, \alpha) = \left[\prod_{i=1}^{n+1} D_i(\lambda, \alpha) \right]^{\frac{1}{n+1}},$$

with respect to λ and α , or, equivalently, by maximizing the logarithm of the geometric

mode of sample spacings

$$H(\lambda, \alpha) = \frac{1}{n+1} \sum_{i=1}^{n+1} \log D_i(\lambda, \alpha).$$

The estimators $\hat{\lambda}_{MPS}$ and $\hat{\alpha}_{MPS}$ of the parameters λ and α can be obtained by solving the following nonlinear equations

$$\begin{aligned} \frac{\partial H(\lambda, \alpha)}{\partial \lambda} &= \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{D_i(\lambda, \alpha)} [\eta_1(t_{(i)}|\lambda, \alpha) - \eta_1(t_{(i-1)}|\lambda, \alpha)] = 0, \\ \frac{\partial H(\lambda, \alpha)}{\partial \alpha} &= \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{D_i(\lambda, \alpha)} [\eta_2(t_{(i)}|\lambda, \alpha) - \eta_2(t_{(i-1)}|\lambda, \alpha)] = 0. \end{aligned} \quad (19)$$

Note that if $t_{(i+k)} = t_{(i+k-1)}$ then $D_{i+k}(\lambda, \alpha) = D_{i+k-1}(\lambda, \alpha) = 0$ for some i . Therefore, the MPS estimators are sensitive to closely spaced observations, especially ties. When ties are due to multiple observations, $D_i(\lambda, \alpha)$ should be replaced by the corresponding likelihood $f(t_{(i)}, \lambda, \alpha)$ since $t_{(i)} = t_{(i-1)}$.

Cheng and Amin [10] presented desirable properties of the MPS such as asymptotic efficiency and invariance. They also proved that the consistency of maximum product of spacing estimators holds under much more general conditions than for maximum likelihood estimators. Therefore, the MPS estimators are asymptotically normally distributed with a joint bivariate normal distribution given by

$$(\hat{\lambda}_{MPS}, \hat{\alpha}_{MPS}) \sim N_2[(\lambda, \alpha), I^{-1}(\lambda, \alpha)], \quad \text{for } n \rightarrow \infty.$$

2.7. The Cramér-von Mises maximum goodness-of-fit estimators

The Cramér-von Mises is a type of maximum goodness-of-fit estimators (also called minimum distance estimators) and is based on the difference between the estimate of the cumulative distribution function and the empirical distribution function [6].

Macdonald [29] motivated the choice of the Cramér-von Mises statistic and provided empirical evidence that the bias of the estimator is smaller than the other goodness-of-fit estimators. The proposed estimator is based on the Cramér-von Mises statistics given by

$$W_n^2 = n \int_{-\infty}^{\infty} (F(t_{(i)}) - E_n(t_{(i)}))^2 dF(t_{(i)}),$$

where $E_n(\cdot)$ is the empirical density function. Boos [6] discussed its asymptotic properties and presented its computational form which is given by

$$C(\lambda, \alpha) = \frac{1}{12n} + \sum_{i=1}^n \left(F(t_{(i)} | \theta, \lambda) - \frac{2i-1}{2n} \right)^2. \quad (20)$$

The Cramér-von Mises estimators $\hat{\lambda}_{CME}$ and $\hat{\alpha}_{CME}$ of the parameters λ and α are

obtained by minimizing

$$\begin{aligned}\sum_{i=1}^n \left(F(t_{(i)} | \lambda, \alpha) - \frac{2i-1}{2n} \right) \eta_1(t_{(i)} | \lambda, \alpha) &= 0, \\ \sum_{i=1}^n \left(F(t_{(i)} | \lambda, \alpha) - \frac{2i-1}{2n} \right) \eta_2(t_{(i)} | \lambda, \alpha) &= 0,\end{aligned}$$

where $\eta_1(\cdot | \lambda, \alpha)$ and $\eta_2(\cdot | \lambda, \alpha)$ are given respectively in (18).

2.8. Method of Anderson-Darling

Other type of maximum goodness-of-fit estimators is based on an Anderson-Darling statistic and is known as the Anderson-Darling estimator. The Anderson-Darling statistic is given by

$$\text{ADS}_n^2 = n \int_{-\infty}^{\infty} \frac{(F(t_{(i)}) - E_n(t_{(i)}))^2}{F(t)(1-F(t))} dF(t_{(i)}).$$

Boos [6] also discussed the properties of the AD estimators and presented its computational form which the Anderson-Darling estimators $\hat{\lambda}_{ADE}$ and $\hat{\alpha}_{ADE}$ are obtained by minimizing, with respect to λ and α , the function

$$A(\lambda, \alpha) = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \left[-\lambda t_{(i)}^{-\alpha} + \log \left(1 - e^{-\lambda t_{(n+1-i)}^{-\alpha}} \right) \right].$$

These estimators can also be obtained by solving the following non-linear equations:

$$\begin{aligned}\sum_{i=1}^n (2i-1) \left[\frac{\eta_1(t_{(i)} | \lambda, \alpha)}{F(t_{(i)} | \lambda, \alpha)} - \frac{\eta_1(t_{n+1-i:n} | \lambda, \alpha)}{1 - F(t_{n+1-i:n} | \lambda, \alpha)} \right] &= 0, \\ \sum_{i=1}^n (2i-1) \left[\frac{\eta_2(t_{(i)} | \lambda, \alpha)}{F(t_{(i)} | \lambda, \alpha)} - \frac{\eta_2(t_{n+1-i:n} | \lambda, \alpha)}{1 - F(t_{n+1-i:n} | \lambda, \alpha)} \right] &= 0,\end{aligned}$$

where $\eta_1(\cdot | \lambda, \alpha)$ and $\eta_2(\cdot | \lambda, \alpha)$ are given respectively in (18).

3. Bayesian Analysis

In the previous sections, we have presented different estimation procedures using the frequentist approach. In this section, we consider the Bayesian inference approach for estimating the the unknown parameters of the Fréchet distribution. Bayesian analysis is an attractive framework in practical problems and has grown popularity in recent years. The prior distribution is a key part of the Bayesian inference and there are different types of priors distribution available in the literature (see, for instance, Ramos et al. [35]). Usually, non-informative priors are preferable because if prior information on study parameters is unavailable or does not exist for a device, then initial

uncertainty about the parameters can be quantified with a non-informative prior distribution. An important non-informative reference prior was introduced by Bernardo [4], with further developments (see Bernardo [5] and references therein). The proposed idea was to maximize the expected Kullback-Leibler divergence between the posterior distribution and the prior. The reference prior provides posterior distribution with interesting properties, such as invariance under one-to-one transformations, consistent marginalization and consistent sampling properties. Recently, Abbas and Tang [1] derived two reference priors as well as the Jeffreys' prior for the Fréchet distribution. Here, we derive such priors by using different approach, the following proposition is used to obtain the reference priors.

Proposition 3.1. [Bernardo [5], pg 40, Theorem 14]. *Let $\boldsymbol{\theta} = (\theta_1, \theta_2)$ be a vector of the ordered parameters of interest and $I(\theta_1, \theta_2)$ is the Fisher information matrix. If the parameter space of θ_1 does not depend of θ_2 and $I_{j,j}(\boldsymbol{\theta}), j = 1, 2$ are factorized in the form*

$$S_{1,1}^{-\frac{1}{2}}(\boldsymbol{\theta}) = f_1(\theta_1)g_1(\theta_2) \quad \text{and} \quad I_{2,2}^{\frac{1}{2}}(\boldsymbol{\theta}) = f_2(\theta_1)g_2(\theta_2).$$

Then the reference prior for the ordered parameters $\boldsymbol{\theta}$ is given by $\pi_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = f_1(\theta_1)g_2(\theta_2)$ and there is no need for compact approximations, even if the conditional priors are not proper.

Theorem 3.2. *Let $\omega_1 = (\lambda, \alpha)$ and $\omega_2 = (\alpha, \lambda)$ be the vectors of the ordered parameters then the reference priors for the ordered parameters are respectively given by*

$$\pi_{\omega_1}(\lambda, \alpha) \propto \frac{1}{\lambda\alpha} \quad \text{and} \quad \pi_{\omega_2}(\alpha, \lambda) \propto \frac{1}{\lambda\alpha}. \quad (21)$$

Proof. For $\omega_1 = (\lambda, \alpha)$, we have $f_1(\lambda) = \lambda^{-1}$, $g_1(\alpha) = 1$, $f_2(\lambda) = \sqrt{\frac{\pi^2}{6} + (1 - \gamma - \log(\lambda))^2}$, $g_2(\alpha) = \alpha^{-1}$. Therefore, $\pi_{\omega_1}(\lambda, \alpha) \propto f_1(\lambda)g_2(\alpha) \propto \lambda^{-1}\alpha^{-1}$. Analogously, we obtain $\pi_{\omega_2}(\alpha, \lambda)$. \square

Moreover, Berger et al. [3] suggested by starting with a collection of reference priors and then taking the arithmetic mean or the geometric mean to obtain an overall reference prior. Therefore, an overall reference prior is the same as (21). Another well-known non-informative prior was introduced by Jeffreys [20] and can be obtained through the square root of the determinant of Fisher information matrix (8), such prior is widely used due to its invariance property under one-to-one transformations. After some algebraic manipulations we have

$$\pi_J(\lambda, \alpha) \propto \frac{1}{\lambda\alpha}, \quad (22)$$

which is equal to (21). It is worth noting that such prior also arises considering the Jeffreys rule (see, Kass and Wasserman [24]). Therefore, even considering different methods to obtain non-informative priors, we have the same prior for the Fréchet distribution.

Combining the likelihood (3) and the prior (22), the posterior distribution is

$$\pi(\lambda, \alpha | \mathbf{t}) = \frac{(\lambda\alpha)^{n-1}}{c(\mathbf{t})} \prod_{i=1}^n t_i^{-\alpha} \exp \left\{ -\lambda \sum_{i=1}^n t_i^{-\alpha} \right\}, \quad (23)$$

where

$$c(\mathbf{t}) = \int_{\mathcal{A}} (\lambda\alpha)^{n-1} \prod_{i=1}^n t_i^{-\alpha} \exp \left\{ -\lambda \sum_{i=1}^n t_i^{-\alpha} \right\} d\boldsymbol{\theta} \quad (24)$$

and $\mathcal{A} = \{(0, \infty) \times (0, \infty)\}$ is the parameter space of $\boldsymbol{\theta}$. To get reliable inference, first we have to check whether the posterior distribution (23) is a proper posterior, i.e., $c(\mathbf{t}) < \infty$.

Theorem 3.3. *The posterior distribution (23) is proper if $n \geq 2$.*

Proof. Since $(\lambda\alpha)^{n-1} \prod_{i=1}^n t_i^{-\alpha} \exp \left\{ -\lambda \sum_{i=1}^n t_i^{-\alpha} \right\} \geq 0$, by Tonelli theorem (see Folland [15]) we have

$$\begin{aligned} c(\mathbf{t}) &= \int_{\mathcal{A}} (\lambda\alpha)^{n-1} \prod_{i=1}^n t_i^{-\alpha} \exp \left\{ -\lambda \sum_{i=1}^n t_i^{-\alpha} \right\} d\boldsymbol{\theta} \\ &= \int_0^\infty \int_0^\infty (\lambda\alpha)^{n-1} \prod_{i=1}^n t_i^{-\alpha} \exp \left\{ -\lambda \sum_{i=1}^n t_i^{-\alpha} \right\} d\lambda d\alpha \\ &= \Gamma(n) \int_0^\infty \alpha^{n-1} \prod_{i=1}^n t_i^{-\alpha} \left\{ \sum_{i=1}^n t_i^{-\alpha} \right\}^{-n} d\alpha \\ &\leq \int_0^\infty \alpha^{n-1} \prod_{i=1}^n \left(\frac{t_i}{\min(t_1, \dots, t_n)} \right)^{-\alpha} d\alpha < \infty, \end{aligned}$$

the last inequality holds if $n \geq 2$. □

The marginal posterior distribution for α is given by

$$\pi(\alpha | \mathbf{t}) \propto \alpha^{n-2} \prod_{i=1}^n t_i^{-\alpha} \left\{ \sum_{i=1}^n t_i^{-\alpha} \right\}^{-n}. \quad (25)$$

The conditional posterior distribution for λ is

$$\pi(\lambda | \alpha, \mathbf{t}) \sim \text{Gamma} \left(n, \sum_{i=1}^n t_i^{-\alpha} \right), \quad (26)$$

where $\text{Gamma}(a, c)$ is the Gamma distribution with PDF, $f(x, a, c) = c^a x^{a-1} \exp(-cx)/\Gamma(a)$. By considering marginal and conditional posterior distributions (25) and (26), the convergence of Markov Chain Monte Carlo (MCMC) method can be easily achieved.

Abbas and Tang [1] derived the same priors (21) and proved that the obtained posterior is proper. However, the authors did not prove that the obtained posterior means for α and λ are finite. These proofs are important in order to obtain reliable results. The following theorem prove that the posterior mean of λ is improper depending on the data.

Theorem 3.4. *The posterior mean for λ is improper in case $\prod_{i=1}^n \left(\frac{t_i}{\min(t_1, \dots, t_n)} \right) \leq \min(t_1, \dots, t_n)$ and $n \geq 2$.*

Proof. The posterior mean for λ is given by

$$\hat{\lambda}(\mathbf{t}) = \int_{\mathcal{A}} \lambda (\lambda \alpha)^{n-1} \prod_{i=1}^n t_i^{-\alpha} \exp \left\{ -\lambda \sum_{i=1}^n t_i^{-\alpha} \right\} d\boldsymbol{\theta}.$$

Notice that, since $\prod_{i=1}^n t_i / \min(t_1, \dots, t_n)^{n+1} \leq 1$ by hypothesis it follows that $(\prod_{i=1}^n t_i / \min(t_1, \dots, t_n)^{n+1})^{-\alpha} \geq 1^{-\alpha} = 1$.

Now, since $\lambda (\lambda \alpha)^{n-1} \prod_{i=1}^n t_i^{-\alpha} \exp \left\{ -\lambda \sum_{i=1}^n t_i^{-\alpha} \right\} \geq 0$, by Tonelli theorem we have

$$\begin{aligned} \hat{\lambda}(\mathbf{t}) &= \int_{\mathcal{A}} \lambda (\lambda \alpha)^{n-1} \prod_{i=1}^n t_i^{-\alpha} \exp \left\{ -\lambda \sum_{i=1}^n t_i^{-\alpha} \right\} d\boldsymbol{\theta} \\ &= \int_0^\infty \prod_{i=1}^n t_i^{-\alpha} \int_0^\infty \alpha^{n-1} \lambda^n \exp \left\{ -\lambda \sum_{i=1}^n t_i^{-\alpha} \right\} d\lambda d\alpha \\ &= \Gamma(n+1) \int_0^\infty \alpha^{n-1} \prod_{i=1}^n t_i^{-\alpha} \left\{ \sum_{i=1}^n t_i^{-\alpha} \right\}^{-n} d\alpha \\ &\geq \Gamma(n+1) n^{-n} \int_0^\infty \alpha^{n-1} \left(\frac{\prod_{i=1}^n t_i}{\min(t_1, \dots, t_n)^{n+1}} \right)^{-\alpha} d\alpha \\ &\geq \Gamma(n+1) n^{-n} \int_0^\infty \alpha^{n-1} d\alpha = \infty, \end{aligned}$$

where the last inequality holds for $n \geq 2$. □

From the above theorem, we can see that the posterior mean of λ may be improper depending on the data, which is undesirable. The use of Monte Carlo methods in improper posterior was discussed by Hobert and Casella [18]. The authors argued that “one can not expect the Gibbs output to provide a “red flag”, informing the user that the posterior is improper. The user must demonstrate propriety before a Markov Chain Monte Carlo technique is used.” In our case, as the properness of the posterior mean of λ depend on the data, other alternative is needed to be considered as Bayes estimator. A discussion about other alternatives will be presented in the next section.

4. Simulation Study

In this section, we present some experimental results to evaluate the performance of the different methods of estimation discussed in the previous sections.

4.1. Classical approach

In this subsection, we compare the efficiency of the different estimation methods considering classical approach. The following procedures are adopted:

- (1) Generate N samples from the $\text{Fr\'echet}(\lambda, \alpha)$ distribution with size n and compute $\hat{\boldsymbol{\theta}} = (\hat{\lambda}, \hat{\alpha})$ via MLE, ME, LME, LSE, WLSE, PCE, MPS, CME and ADE.
- (2) Using $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}$, compute the mean relative estimates (MRE) = $\sum_{i=1}^N (\hat{\theta}_i / \theta_j) / N$ and the mean squared errors (MSE) = $\sum_{i=1}^N (\hat{\theta}_i - \theta_j)^2 / N$, for $j = 1, 2$.

Considering the above approach, the most efficient estimator will be the one whose MREs is closer to one with smaller MSEs. These results are computed using the software R (see, R Core Team [34]). The seed used to generate the pseudo-random samples from the Fr\'echet distribution was 2018. We have chosen the values to perform this procedure are $N = 500,000$ and $n = (15, 20, 25, \dots, 140)$. We presented results only for $\boldsymbol{\theta} = (2, 4)$ due to space constraint. However, results are similar for other choices of λ and α . Figure 1 shows the MREs, MSEs for the estimates of $\boldsymbol{\theta}$. The horizontal lines in the Figure correspond to MREs and MSEs being one and zero respectively.

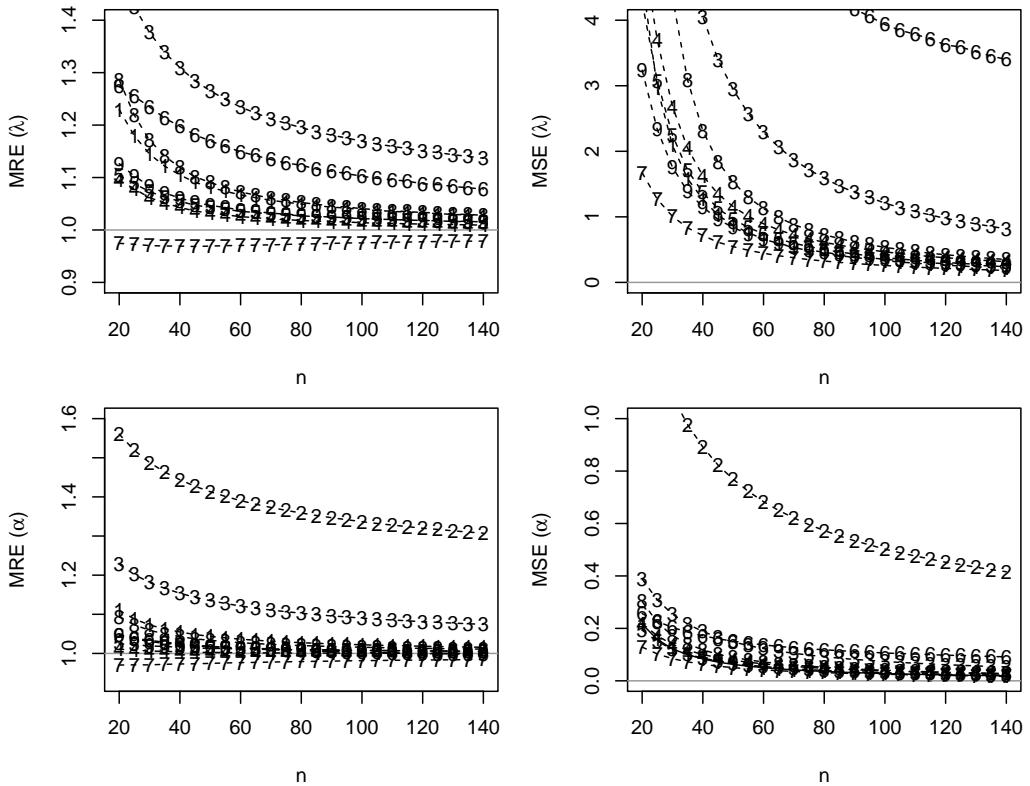


Figure 1. MREs, MSEs for the estimates of $\lambda = 2$ and $\alpha = 4$ for $N = 500,000$ simulated samples, considering different values of n using the following estimation method 1-MLE, 2-ME, 3-LME, 4-LSE, 5-WLSE, 6-PCE, 7-MPS, 8-CME, 9-ADE.

From Figure 1, we observe that the MSEs of all estimators of the parameters tend to zero for large n and also, the values of MREs tend to one, i.e. the estimators are asymptotically unbiased and consistent for the parameters. For both parameters, we observe that the moment estimators has the largest MREs and MSEs respectively among all the considered estimators. Further, we also observe that the MPS, ADE, LSE and WLSE performs better than the MLEs for small and moderate sample sizes in terms of MREs and MSEs. Moreover, the MPS estimators have the smallest MSEs among all the considered estimators.

Combining all results with the good properties of the MPS method such as consistency, asymptotic efficiency, normality and invariance, we suggest to use MPS estimators of the parameters of Fréchet distribution in all practical purposes.

4.2. Bayesian approach

In this subsection, we obtain the Bayes estimator under the same assumptions of section 4.1. The 95% coverage probability of the asymptotic confidence intervals under the classical set-up and the credible intervals ($CI_{95\%}$) under the Bayesian set-up are also evaluated. For large number of experiments and considering confidence level of 95%, the frequencies of intervals that covered the true values of θ should be closer to 95%.

4.2.1. Metropolis-Hastings (M-H) algorithm

Since the marginal posterior distribution of α does not have closed form, the Metropolis-Hastings (M-H) algorithm is applied to generate samples from this marginal distribution. In this case, we have used the Gamma distribution as transition kernel $q(\alpha^{(j)}|\alpha^{(*)}, b)$ for sampling values of α , where b is a known hyperparameter that controls the acceptance rate of the algorithm. It is worth mentioning that, the choice of the transition kernel is arbitrary, and other non-negative random variables could also be used instead. The M-H algorithm operates as follows:

- (1) Start with an initial value $\alpha^{(1)}$ and set the iteration counter $j = 1$;
- (2) Generate a random value $\alpha^{(*)}$ from the proposal $\text{Gamma}(\alpha^{(j)}, b)$;
- (3) Evaluate the acceptance probability

$$h(\alpha^{(j)}, \alpha^{(*)}) = \min \left(1, \frac{\pi(\alpha^{(*)}|\mathbf{t})}{\pi(\alpha^{(j)}|\mathbf{t})} \frac{q(\alpha^{(j)}, \alpha^{(*)}, b)}{q(\alpha^{(*)}, \alpha^{(j)}, b)} \right),$$

where $\pi(\cdot)$ is given in (25). Generate a random value u from an independent uniform in $(0, 1)$;

- (4) If $h(\alpha^{(j)}, \alpha^{(*)}) \geq u(0, 1)$ then $\alpha^{(j+1)} = \alpha^{(*)}$, otherwise $\alpha^{(j+1)} = \alpha^{(j)}$;
- (5) Change the counter from j to $j + 1$ and return to step 2 until convergence is reached.

In this case, we choose b to be equal to one. However, other values can also be considered. To decrease the necessary time taken for M-H method to reach the convergence, we can use (16) as a good initial value for $\alpha^{(1)}$. Considering the conditional posterior distribution (26), the Bayes estimator for λ can be obtained directly from the Gamma distribution with $\left(n, \sum_{i=1}^n t_i^{-\hat{\alpha}_{Bayes}}\right)$ as well as its respective credible inter-

val is evaluated by the quantile function. The decision rule used to obtain the Bayes estimators will be presented in the following.

4.2.2. Bayes estimator

The selection of a decision rule to obtain posterior estimates is of fundamental problem in Bayesian statistics. Usually, this problem is over looked by many authors. The most common risk function used to obtain the Bayes estimates is the mean squared error, by considering this risk function, the obtained Bayes estimates are the posterior means. Other alternative functions can be considered, for instance, the posterior mode, also known as maximum a posteriori probability (MAP) is obtained by assuming a 0-1 loss function, while the posterior median is obtained considering a linear loss function.

For the Fréchet distribution, Abbas and Tang [1] presented two Bayes estimators, the first is the MAP estimators obtained from Laplace's approximation. Despite of the fact that reference priors are invariant under-one-to-one transformation, the MAP is not [33], and therefore, the obtained Bayes estimator is not invariant. Although the authors considered other Bayes estimator, but they did not specify how the proposed estimator was obtained. This lack of information makes such results non-reproducible, which is undesirable for any application. As the authors used Monte Carlo methods, the most common Bayesian estimator is the posterior mean. However, from Theorem 3.4 we have proved that the posterior mean may be improper depending on the data, which is undesirable. In fact, consider the example analyzed by Abbas and Tang [1] related to fatigue lifetime data. The data is given by: 152.7, 172.0, 172.5, 173.3, 193.0, 204.7, 216.5, 234.9, 262.6, 422.6. From the proposed theorem, the posterior mean of λ is improper if

$$\prod_{i=1}^n \left(\frac{t_i}{\min(t_1, \dots, t_n)} \right) \leq \min(t_1, \dots, t_n),$$

since $25.4 \leq 152.7$, then the posterior mean of λ is improper and can not be used. As it was discussed earlier, for the proposed dataset we can easily draw a sample for the marginal distribution and compute the posterior mean without any "red flag", but such estimate is meaningless. Therefore, for the Fréchet distribution the posterior median is a reasonable choice as it is invariant under-one-to-one transformation and finite for $n \geq 2$ almost surely.

4.2.3. Results

For each simulated data set under the Bayesian approach, 5,500 iterations are performed using the MCMC methods. As a burning sample, we discarded the first 500 initial values taking jumps of size 5 to reduce the auto-correlation values among the chain, getting at the end one chain of size 1,000. To validate the convergence of the obtained chain, we used the Geweke criterion [16] with a 95% confidence level. At the end, 10,000 posterior medians for α and λ were computed.

The results are presented only for $\theta = (2, 4)$ due to space constraint. However, the following results are also similar for other choices of λ and α . Figure 2 presents the MREs, MSEs and the coverage probability with a confidence level equal to 95% for the estimates obtained for the MPS and posterior medians. It is to be noted that, since MPS performs better than their counterparts, we compare Bayes estimates with MPS.

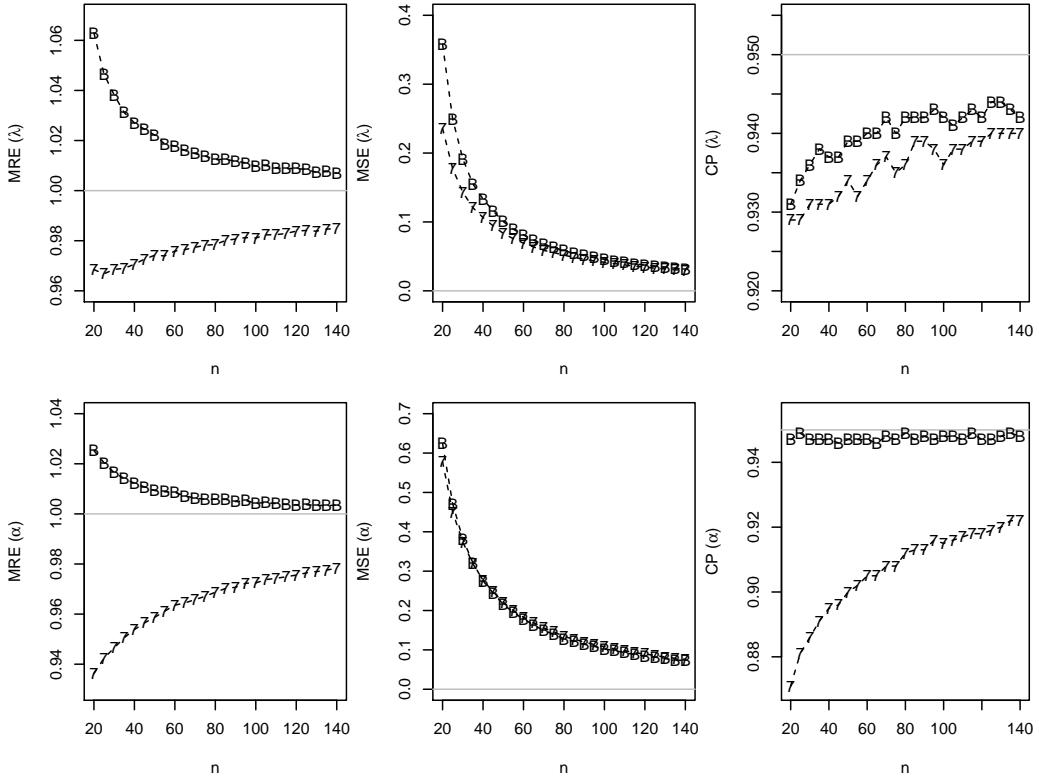


Figure 2. MREs, MSEs and the coverage probability with a 95% confidence level for the estimates of $\lambda = 2$ and $\alpha = 4$ for $N = 500,000$ simulated samples, considering different values of n using the following estimation method, 7 - MPS, B - Bayes estimator.

From Figure 2 we observe that the Bayes estimates have the MRE nearest to one while both approaches (MPS and Bayes estimates) have approximately the same MSE. Moreover, considering the Bayesian approach, we have obtained accurate coverage probability through the highest posterior density intervals. Therefore, we conclude that the Bayesian approach is preferred in order to make inference on unknown parameters of the Fréchet distribution.

5. Applications

In this section, we analyze five real data sets related to minimum monthly flows of water (m^3/s) on the Piracicaba River, located in São Paulo state, Brazil. This study can be useful to protect and maintain aquatic resources for the state [39, 42]. The data sets (see in Supplemental material for more details) are obtained from the Department of Water Resources and Power agency manager of water resources of the State of São Paulo from 1960 to 2014.

Table 1 presents the posterior median (Bayes estimators) and 95% credible intervals for λ and α of the Fréchet distribution for the data sets related to the total monthly rainfall during May, June, July, August and September at Piracicaba River.

The results obtained using the Fréchet distribution are compared with the Weibull, Gamma, Lognormal (LN), Gumbel and Generalized Exponential (GE) distribution [17]. We consider certain discrimination criteria such as BIC (Bayesian Information

Table 1. Bayes estimates, 95% credible intervals for λ and α for the data sets related to the total monthly rainfall during May, June, July, August and September at Piracicaba River.

Month	$\boldsymbol{\theta}$	$\hat{\boldsymbol{\theta}}_{Bayes}$	CI _{95%} ($\boldsymbol{\theta}$)
May	λ	309.890	(223.248; 416.505)
	α	1.817	(1.401; 2.293)
June	λ	89.758	(64.376; 121.075)
	α	1.585	(1.194; 2.033)
July	λ	204.493	(146.666; 275.840)
	α	2.048	(1.549; 2.637)
August	λ	401.656	(290.594; 537.969)
	α	2.4585	(1.876; 3.119)
September	λ	55.128	(39.539; 74.362)
	α	1.529	(1.163; 1.939)

Criteria), AIC (Akaike Information Criteria) and AICc (Corrected Akaike information criterion) and computed respectively by $BIC = -2l(\hat{\boldsymbol{\theta}}; \mathbf{t}) + k \log(n)$, $AIC = -2l(\hat{\boldsymbol{\theta}}; \mathbf{t}) + 2k$ and $AICc = AIC + [2k(k+1)]/[(n-k-1)]$, where k is the number of parameters to be fitted and $\hat{\boldsymbol{\theta}}$ is estimation of $\boldsymbol{\theta}$. Given a set of candidate models for \mathbf{t} , the preferred model is the one which provides the minimum values of the aforementioned statistics.

Table 2 presents the results of BIC, AIC and AICc, for different probability distributions. The goodness of fit can also be checked through the over plot of the survival function adjusted by the proposed theoretical models onto the empirical survival function as shown in Figure 3.

Table 2. Results of the BIC, AIC and the AICc for different probability distributions for the data sets related to the minimum flows of water (m³/s) during May, June, July and August at Piracicaba River.

Month	Test	Fréchet	Weibull	Gamma	LN	Gumbel	GE
May	BIC	361.43	390.91	386.90	369.93	394.23	384.04
	AIC	358.06	387.53	383.52	366.55	390.85	380.66
	AICc	358.38	387.86	383.84	366.88	391.18	380.98
June	BIC	346.92	379.72	381.02	359.70	403.55	380.81
	AIC	343.60	376.39	377.69	356.37	400.22	377.48
	AICc	343.93	376.73	378.03	356.70	400.55	377.81
July	BIC	302.86	336.50	332.78	316.75	341.31	330.30
	AIC	299.54	333.17	329.45	313.42	337.98	326.97
	AICc	299.87	333.50	329.79	313.75	338.32	327.30
August	BIC	283.92	310.33	303.41	294.30	303.68	299.35
	AIC	280.49	306.90	299.98	290.87	300.25	295.92
	AICc	280.81	307.22	300.30	291.19	300.57	296.24
September	BIC	329.06	344.21	341.77	332.96	351.45	340.68
	AIC	325.73	340.89	338.44	329.63	348.12	337.35
	AICc	326.06	341.22	338.77	329.96	348.45	337.69

Comparing the empirical survival function with the adjusted distributions, we observe that Fréchet distribution fits better than the chosen models. These results are confirmed from the BIC, AIC and AICc values, since Fréchet distribution has the minimum values for the proposed data sets. Therefore, our proposed methodology can be used successfully to analyze the minimum flow of water during May, June, July,

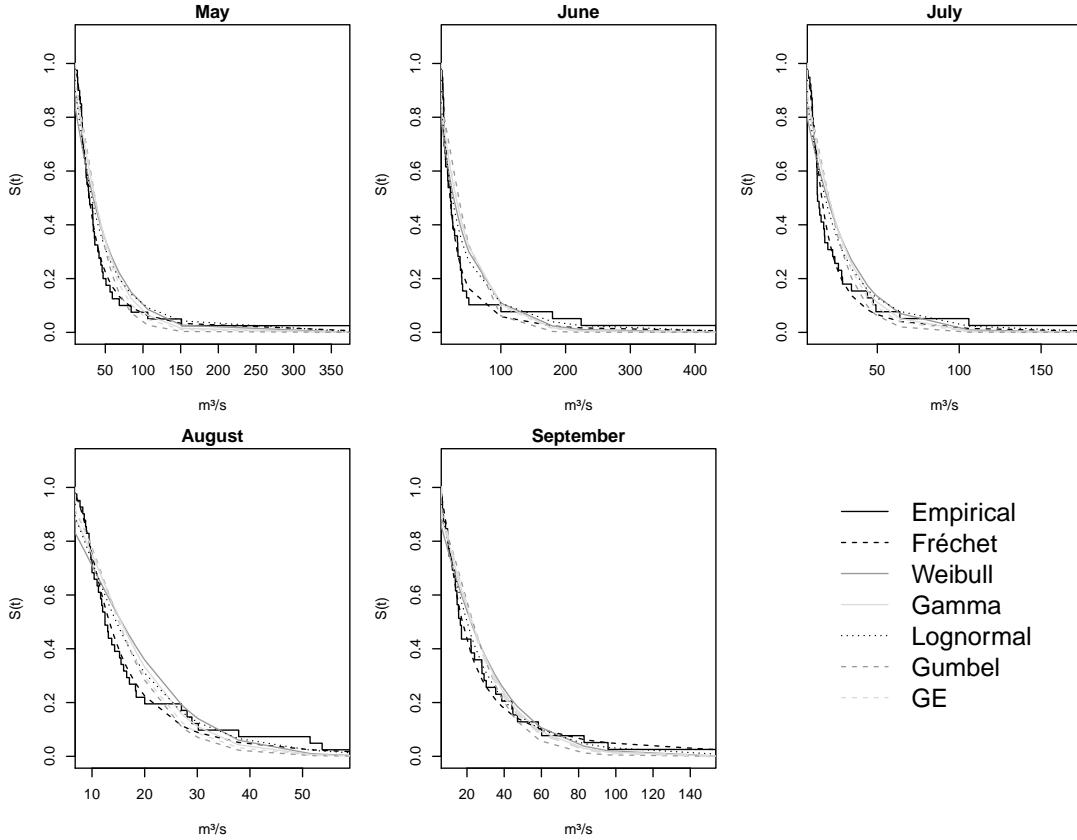


Figure 3. Survival function adjusted by the empirical and the different distributions for the data sets related to the minimum flows of water (m^3/s) during May till September at Piracicaba River.

August and September at Piracicaba River using the Fréchet distribution with the Bayesian approach.

6. Conclusions

In this paper, we considered different methods of estimation of the unknown parameters both from frequentist and Bayesian viewpoint of Fréchet distribution. We considered the MLEs, MMEs, LME, PCEs, LSEs, WLSE, MPS, CME and ADE as frequentist estimators. As it is not feasible to compare these methods of estimation theoretically, we have performed an extensive simulation study to compare these nine methods of estimation. We compared these frequentist estimators mainly with respect to MREs and MSEs. The results show that among the classical estimators, MPS performs better than their counterpart in terms of MSEs in the simulation study. Additionally, we have obtained Bayes estimates through posterior median and compared with respect to MREs and MSEs. We have obtained accurate coverage probability through the credible intervals. Therefore, combining all these results, we conclude that the Bayesian approach is preferred in order to make inference on unknown parameters of the Fréchet distribution. We also compared this model with some existing competing models and the Fréchet distribution performed reasonably well.

Acknowledgements

The authors are grateful to the principal editor and the anonymous reviewers for their valuable suggestions and comments which have improved the paper substantially. Pedro L. Ramos is grateful to the São Paulo State Research Foundation (FAPESP Proc. 2017/25971-0). The data used in the application and the codes are available in the supplemental material.

References

- [1] K. Abbas and Y. Tang, *Analysis of frechet distribution using reference priors*, Communications in Statistics-Theory and Methods 44 (2015), pp. 2945–2956.
- [2] M.R. Alkasasbeh and M.Z. Raqab, *Estimation of the generalized logistic distribution parameters: Comparative study*, Statistical Methodology 6 (2009), pp. 262–279.
- [3] J.O. Berger, J.M. Bernardo, D. Sun, *et al.*, *Overall objective priors*, Bayesian Analysis 10 (2015), pp. 189–221.
- [4] J.M. Bernardo, *Reference posterior distributions for bayesian inference*, Journal of the Royal Statistical Society. Series B (Methodological) (1979), pp. 113–147.
- [5] J.M. Bernardo, *Reference analysis*, Handbook of statistics 25 (2005), pp. 17–90.
- [6] D.D. Boos, *Minimum distance estimators for location and goodness of fit*, Journal of the American Statistical association 76 (1981), pp. 663–670.
- [7] R. Calabria and G. Pulcini, *Confidence limits for reliability and tolerance limits in the inverse weibull distribution*, Reliability Engineering & System Safety 24 (1989), pp. 77–85.
- [8] R. Calabria and G. Pulcini, *Bayes 2-sample prediction for the inverse weibull distribution*, Communications in Statistics-Theory and Methods 23 (1994), pp. 1811–1824.
- [9] R. Cheng and N. Amin, *Maximum product of spacings estimation with application to the lognormal distribution*, Math. Report (1979), pp. 79–1.
- [10] R. Cheng and N. Amin, *Estimating parameters in continuous univariate distributions with a shifted origin*, Journal of the Royal Statistical Society. Series B (Methodological) (1983), pp. 394–403.
- [11] S. Dey, T. Dey, S. Ali, and M.S. Mulekar, *Two-parameter maxwell distribution: Properties and different methods of estimation*, Journal of Statistical Theory and Practice 10 (2016), pp. 291–310.
- [12] S. Dey, T. Dey, and D. Kundu, *Two-parameter rayleigh distribution: different methods of estimation*, American Journal of Mathematical and Management Sciences 33 (2014), pp. 55–74.
- [13] S. Dey, D. Kumar, P.L. Ramos, and F. Louzada, *Exponentiated chen distribution: Properties and estimation*, Communications in Statistics-Simulation and Computation (2016).
- [14] P. Erto, *Genesis, properties and identification of the inverse weibull lifetime model*, Statistica Applicata 1 (1989), pp. 117–128.
- [15] G.B. Folland, *Real analysis: modern techniques and their applications*, 2nd ed., Wiley, New York, 1999.
- [16] J. Geweke, *et al.*, *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*, Vol. 196, Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA, 1991.
- [17] R.D. Gupta and D. Kundu, *Generalized exponential distribution: different method of estimations*, Journal of Statistical Computation and Simulation 69 (2001), pp. 315–337.
- [18] J.P. Hobert and G. Casella, *The effect of improper priors on gibbs sampling in hierarchical linear mixed models*, Journal of the American Statistical Association 91 (1996), pp. 1461–1473.
- [19] J.R. Hosking, *L-moments: analysis and estimation of distributions using linear combinations of order statistics*, Journal of the Royal Statistical Society. Series B (Methodological)

(1990), pp. 105–124.

[20] H. Jeffreys, *An invariant form for the prior probability in estimation problems*, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 186. The Royal Society, 1946, pp. 453–461.

[21] A.F. Jenkinson, *The frequency distribution of the annual maximum (or minimum) values of meteorological elements*, Quarterly Journal of the Royal Meteorological Society 81 (1955), pp. 158–171.

[22] J.H. Kao, *Computer methods for estimating weibull parameters in reliability studies*, IRE Transactions on Reliability and Quality Control (1958), pp. 15–22.

[23] J.H. Kao, *A graphical estimation of mixed weibull parameters in life-testing of electron tubes*, Technometrics 1 (1959), pp. 389–407.

[24] R.E. Kass and L. Wasserman, *The selection of prior distributions by formal rules*, Journal of the American Statistical Association 91 (1996), pp. 1343–1370.

[25] S. Kotz and S. Nadarajah, *Extreme value distributions: theory and applications*, World Scientific, 2000.

[26] D. Kundu and H. Howlader, *Bayesian inference and prediction of the inverse weibull distribution for type-ii censored data*, Computational Statistics & Data Analysis 54 (2010), pp. 1547–1558.

[27] D. Kundu and M.Z. Raqab, *Generalized rayleigh distribution: different methods of estimations*, Computational statistics & data analysis 49 (2005), pp. 187–200.

[28] A. Loganathan and A. Uma, *Comparison of estimation methods for inverse weibull parameters*, Global and Stochastic Analysis 4 (2017), pp. 83–93.

[29] P. Macdonald, *An estimation procedure for mixtures of distribution*, J. Royal Statist. Soc. Ser 33 (1971), pp. 326–329.

[30] N.R. Mann, N.D. Singpurwalla, and R.E. Schafer, *Methods for statistical analysis of reliability and life data* (1974).

[31] M. Maswadah, *Conditional confidence interval estimation for the inverse weibull distribution based on censored generalized order statistics*, journal of Statistical Computation and Simulation 73 (2003), pp. 887–898.

[32] J. Mazucheli, F. Louzada, and M. Ghitany, *Comparison of estimation methods for the parameters of the weighted lindley distribution*, Applied Mathematics and Computation 220 (2013), pp. 463–471.

[33] K.P. Murphy, *Machine learning: a probabilistic perspective*, MIT press, 2012.

[34] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2017). Available at <https://www.R-project.org/>.

[35] P.L. Ramos, J.A. Achcar, F.A. Moala, E. Ramos, and F. Louzada, *Bayesian analysis of the generalized gamma distribution using non-informative priors*, Statistics 51 (2017), pp. 824–843.

[36] P. Ramos and F. Louzada, *The generalized weighted lindley distribution: Properties, estimation and applications*, Cogent Mathematics 3 (2016), p. 1256022.

[37] P. Ramos, D. Nascimento, and F. Louzada, *The long term fréchet distribution: Estimation, properties and its application*, Biom Biostat Int J 6 (2017), p. 00170.

[38] B. Ranneby, *The maximum spacing method. an estimation method related to the maximum likelihood method*, Scandinavian Journal of Statistics (1984), pp. 93–112.

[39] D.W. Reiser, T.A. Wesche, and C. Estes, *Status of instream flow legislation and practices in north america*, Fisheries 14 (1989), pp. 22–29.

[40] G.C. Rodrigues, F. Louzada, and P.L. Ramos, *Poissonexponential distribution: different methods of estimation*, Journal of Applied Statistics 45 (2018), pp. 128–144.

[41] M.A.W.M.K. SALMAN and S.S.M. AMER, *Order statistics from inverse weibull distribution and characterizations*, Metron 61 (2003), pp. 389–401.

[42] D.L. Tennant, *Instream flow regimens for fish, wildlife, recreation and related environmental resources*, Fisheries 1 (1976), pp. 6–10.

Appendix A: Proof of Theorem 2.1.

Proof. It is important to notice that the solutions to the likelihood equation (4) and (5) may fail to be the global maximum of $l(\lambda, \alpha|t)$. However, the reasoning that follows justify why in this case these solution will in fact be the global maximum. Notice that

$$\max_{\lambda \in \mathbb{R}^+, \alpha \in \mathbb{R}^+} l(\lambda, \alpha|t) = \max_{\alpha \in \mathbb{R}^+} \left(\max_{\lambda \in \mathbb{R}^+} l(\lambda, \alpha|t) \right)$$

To find $\max_{\lambda \in \mathbb{R}^+} l(\lambda, \alpha|t)$ we notice its derivative in relation to λ , for α fixed, is given by $\frac{\partial l(\lambda, \alpha|t)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n t_i^{-\alpha}$ which is positive for $\lambda < \hat{\lambda}(\alpha) = \frac{n}{\sum_{i=1}^n t_i^{-\alpha}}$ and negative in case $\lambda > \hat{\lambda}(\alpha)$. Therefore $\hat{\lambda}(\alpha)$ is the unique value that provides $\max_{\lambda \in \mathbb{R}^+} l(\lambda, \alpha|t)$ for α fixed and

$$\max_{\lambda \in \mathbb{R}^+} l(\lambda, \alpha|t) = l\left(\hat{\lambda}(\alpha), \alpha\right)$$

Now, to find $\max l\left(\hat{\lambda}(\alpha), \alpha|t\right)$ notice, by the chain rule, that $l\left(\hat{\lambda}(\alpha), \alpha|t\right)$ is a differentiable function in relation to α with derivative given by $\frac{dl(\hat{\lambda}(\alpha), \alpha|t)}{d\alpha} = \frac{\partial l(\lambda(\alpha), \alpha|t)}{\partial \lambda} \lambda'(\alpha) + \frac{\partial l(\lambda(\alpha), \alpha|t)}{\partial \alpha} = \frac{\partial l(\lambda(\alpha), \alpha|t)}{\partial \alpha}$ because $\frac{\partial l(\lambda(\alpha), \alpha|t)}{\partial \lambda} = 0$. Let $G(\alpha) = \frac{dl(\lambda(\alpha), \alpha|t)}{d\alpha}$

Let us prove that

$$G(\alpha) = \frac{n}{\alpha} - \sum_{i=1}^n \log t_i + \frac{n}{\sum_{i=1}^n t_i^{-\alpha}} \sum_{i=1}^n t_i^{-\alpha} \log t_i$$

admits a unique solution. It is straightforward to see that $\lim_{\alpha \rightarrow 0^+} G(\alpha) = \infty$. We now prove that $\lim_{\alpha \rightarrow \infty} G(\alpha)$ exists and is negative. Let $u_i = \frac{t_i}{\sqrt[n]{\prod_{i=1}^n t_i}}$ for $i = 1 \dots, n$. Then $t_i = u_i \sqrt[n]{\prod_{i=1}^n t_i}$ for $i = 1, \dots, n$ and therefore

$$\begin{aligned} G(\alpha) &= \frac{n}{\alpha} - \sum_{i=1}^n \log t_i + \frac{n}{\sum_{i=1}^n u_i^{-\alpha}} \sum_{i=1}^n u_i^{-\alpha} \log \left(u_i \sqrt[n]{\prod_{i=1}^n t_i} \right) \\ &= \frac{n}{\alpha} - \sum_{i=1}^n \log t_i + \log \prod_{i=1}^n t_i + \frac{n}{\sum_{i=1}^n u_i^{-\alpha}} \sum_{i=1}^n u_i^{-\alpha} \log(u_i) = \frac{n}{\alpha} + \frac{n}{\sum_{i=1}^n u_i^{-\alpha}} \sum_{i=1}^n u_i^{-\alpha} \log(u_i). \end{aligned}$$

Now let $u_{\min} = \min u_1, \dots, u_n$, and r be the number of times u_{\min} appears in u_1, \dots, u_m . Then $\lim_{\alpha \rightarrow \infty} \frac{\sum_{i=1}^n u_i^{-\alpha}}{u_{\min}^{-\alpha}} = r$, $\lim_{\alpha \rightarrow \infty} \frac{\sum_{i=1}^n u_i^{-\alpha} \log(u_i)}{u_{\min}^{-\alpha}} = r \log(u_{\min})$ and therefore

$$\lim_{\alpha \rightarrow \infty} G(\alpha) = 0 + \frac{n}{r} r \log(u_{\min}) = n \log(u_{\min}) < 0,$$

where the last inequality follows since $u_{\min} < 1$, which is true by consequence of the hypothesis that not all t_i are equal. It follows that, because of the intermediate value

theorem, $G(\alpha)$ has at least one root in the interval $[0, \infty)$.

We now prove that $G'(\alpha)$ is negative, which in turn will imply that $G(\alpha)$ cannot have more than one root in $[0, \infty)$, where $G'(\alpha)$ is given by

$$G'(\alpha) = -\frac{n}{\alpha^2} - n \frac{\sum_{i=1}^n t_i^{-\alpha} (\log t_i)^2 \left(\sum_{i=1}^n t_i^{-\alpha} \right) - \left(\sum_{i=1}^n t_i^{-\alpha} \log t_i \right)^2}{\left(\sum_{i=1}^n t_i^{-\alpha} \right)^2}. \quad (27)$$

To show that $G'(\alpha)$ is negative, it is enough to show that

$$\sum_{i=1}^n t_i^{-\beta} (\log t_i)^2 \left(\sum_{i=1}^n t_i^{-\beta} \right) - \left(\sum_{i=1}^n t_i^{-\beta} \log t_i \right)^2 > 0. \quad (28)$$

To prove this, one can use the Cauchy-Schwartz inequality (a special case of Holder's inequality with $p = q = 2$) stated as follows

$$\left(\sum_{i=1}^n a_i^2 \right) \left(\sum_{i=1}^n b_i^2 \right) \geq \left(\sum_{i=1}^n a_i b_i \right)^2, \quad (29)$$

where equality holds if $\frac{a_i}{b_i} = \text{constant}$.

To prove the equation (28), take $a_i^2 = t_i^{-\alpha} (\log t_i)^2$ and $b_i^2 = t_i^{-\alpha}$. Then clearly $a_i b_i = t_i^{-\alpha} \log t_i$, and hence the inequality (28) follows easily from the application of the Cauchy-Schwartz inequality.

It follows that $G(\alpha) = \frac{dl(\lambda(\alpha), \alpha|t)}{d\alpha}$ has only one root $\hat{\alpha}$ such that $G(\alpha) > 0$ for $\alpha < \hat{\alpha}$ and $G(\alpha) < 0$ for $\alpha > \hat{\alpha}$ and therefore $\hat{\alpha}$ is the only value that attains the maximum of $l(\lambda(\alpha), \alpha)$. Then

$$\max_{\lambda \in \mathbb{R}^+, \alpha \in \mathbb{R}^+} l(\lambda, \alpha|t) = l(\hat{\lambda}(\hat{\alpha}), \hat{\alpha}) = l(\hat{\lambda}, \hat{\alpha})$$

and $(\hat{\lambda}, \hat{\alpha})$ is the only pair that attains the maximum of $l(\lambda, \alpha)$. □

SUPPLEMENTAL MATERIAL

Data set

The datasets used in Section 5 is given as follow:

- May: 29.19, 18.47, 12.86, 151.11, 19.46, 19.46, 84.30, 19.30, 18.47, 34.12, 374.54, 19.72, 25.58, 45.74, 68.53, 36.04, 15.92, 21.89, 40.00, 44.10, 33.35, 35.49, 56.25, 24.29, 23.56, 50.85, 24.53, 13.74, 27.99, 59.27, 13.31, 41.63, 10.00, 33.62, 32.90, 27.55, 16.76, 47.00, 106.33, 21.03.
- June: 13.64, 39.32, 10.66, 224.07, 40.90, 22.22, 14.44, 23.59, 47.02, 37.01, 432.11, 10.63, 28.51, 11.77, 25.35, 25.80, 39.73, 9.21, 22.36, 11.63, 33.35, 18.00, 18.62, 17.71, 100.10, 23.32, 11.63, 10.20, 12.04, 11.63, 50.57, 11.63, 33.72, 14.69, 12.30, 32.90, 179.75, 37.57, 7.95.
- July: 12.98, 15.66, 13.18, 174.94, 10.35, 47.52, 13.28, 24.03, 11.40, 22.71, 43.96, 9.38, 11.40, 13.28, 14.84, 14.44, 63.74, 12.04, 17.26, 28.74, 12.25, 10.22, 26.25, 13.31, 28.24, 12.88, 17.71, 8.82, 10.40, 7.67, 49.15, 17.93, 9.80, 105.88, 10.77, 13.49, 19.77, 34.22, 7.26.
- August: 16.00, 9.52, 9.43, 53.72, 17.10, 8.52, 10.00, 15.23, 8.78, 28.97, 28.06, 18.26, 9.69, 51.43, 10.96, 13.74, 20.01, 10.00, 12.46, 10.40, 26.99, 7.72, 11.84, 18.39, 11.22, 13.10, 16.58, 12.46, 58.98, 7.11, 11.63, 8.24, 9.80, 15.51, 37.86, 30.20, 8.93, 14.29, 12.98, 12.01, 6.80.
- September: 29.19, 8.49, 7.37, 82.93, 44.18, 13.82, 22.28, 28.06, 6.84, 12.14, 153.78, 17.04, 13.47, 15.43, 30.36, 6.91, 22.12, 35.45, 44.66, 95.81, 6.18, 10.00, 58.39, 24.05, 17.03, 38.65, 47.17, 27.99, 11.84, 9.60, 6.72, 13.74, 14.60, 9.65, 10.39, 60.14, 15.51, 14.69, 16.44

Code of the Metropolis-Hasting algorithm within Gibbs

```

library(coda)
#####
### Gibbs with Metropolis-Hasting algorithm #####
### R: Iteration Number; burn: Burn in; #####
### jump: Jump size; b= Control generation values #####
### log_posteriori: logarithm of posteriori density #####
#####

MCMC<-function(t,R,burn,jump,b=1) {
log_posteriori <- function (alfa) {
posterior= (n-2)*log(alfa)-(alfa*sum(log(t)))-n*log(sum(t^(-alfa)))
return(posterior) } ##logarithm of the marginal posterior of alpha
valpha<-length(R+1) ; n<-length(t)
valpha[1]<-max(log(2)/(log(((2/(n*(n-1))*sum(seq(0,n-1,1)*sort(t)))-mean(t))/mean(t))+1)),1
##Set the initial value of alpha based on the L-moments
c1<-rep(0,times=R)
###Starts the M-H algorithm described in Section 4.2.1
for(i in 1:R){
prop1<-rgamma(1,shape=b*valpha[i],rate=b)
ratio1<-log_posteriori(prop1)+dgamma(valpha[i],shape=b*prop1,rate=b,log=TRUE)-dgamma(
prop1,shape=b*valpha[i],rate=b,log=TRUE)-log_posteriori(valpha[i])
h<-min(1,exp(ratio1)); u1<-runif(1)
if (u1<h & is.finite(h)) {valpha[i+1]<-prop1 ; c1[i]<-0} else
{valpha[i+1]<-valpha[i] ; c1[i]<-1}
###Ends the M-H algorithm
valpha2<-valpha[seq(burn,R,jump)]  ###Remove the burn-in and takes jump
}

```

```

ge1<-abs(geweke.diag(valpha2)$z[1]) ### Compute the Geweke diagnostic
alpha<-median(valpha2);    ### Compute the median of alpha
### Compute the median of lambda
lambda<-qgamma(0.5, shape=n, rate = sum(t^(-alpha)), lower.tail = TRUE)
prai<-quantile(valpha2, probs = 0.025)  ## Compute the Lower credibility interval of alpha
pras<-quantile(valpha2, probs = 0.975)  ## Compute the Upper credibility interval of alpha
## Compute the Lower credibility interval of lambda
prli<-qgamma(0.025, shape=n, rate = sum(t^(-alpha)), lower.tail = TRUE)
## Compute the Upper credibility interval of lambda
prls<-qgamma(0.975, shape=n, rate = sum(t^(-alpha)), lower.tail = TRUE)
return(list(acep=(1-sum(c1)/length(c1)),lambda=lambda,alpha=alpha, LCI_alpha=prai,
UCI_alpha=pras, LCI_Lambda=prli, UCI_Lambda=prls, Geweke.statistics=ge1))
}

#####
## Example ## t: data vector ##
#####

rIW<-function(n,lambda,alpha) {
  U<-runif(n,0,1)
  t<-(((1/lambda)*(log(1/U)))^(-1/alpha))
  return(t) }

set.seed(2018)
t<-rIW(n=30,lambda=4,alpha=2)

MCMC(t,R=15000,burn=500,jump=5)

$acep      ##Acceptance rate
[1] 0.2490667

$lambda     ##Posterior median of lambda
[1] 4.880441

$alpha      ##Posterior median of alpha
[1] 2.236982

$LCI_alpha  ##Lower credibility interval of alpha
2.5%
1.618934

$UCI_alpha  ##Upper credibility interval of alpha
97.5%
2.917118

$LCI_Lambda ##Lower credibility interval of lambda
[1] 3.329736

$UCI_Lambda ##Upper credibility interval of lambda
[1] 6.851465

$Geweke.statistics ## Geweke Statistics
1.835165

```