



**HAL**  
open science

# Parameter estimation for stable distributions and their mixture

Omar Hajjaji, Solym Manou-Abi, Yousri Slaoui

► **To cite this version:**

Omar Hajjaji, Solym Manou-Abi, Yousri Slaoui. Parameter estimation for stable distributions and their mixture. 2023. hal-04218721

**HAL Id: hal-04218721**

**<https://hal.science/hal-04218721v1>**

Preprint submitted on 26 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

# Parameter estimation for stable distributions and their mixture

Omar Hajjaji<sup>1†</sup>, Solym M. Manou-Abi<sup>1,2,3\*†</sup> and Yousri  
Slaoui<sup>1†</sup>

<sup>1</sup>\*Laboratoire de Mathématiques et Applications, UMR CNRS  
7031, Futuroscope Chasseneuil, Poitiers, 86000, France.

<sup>2</sup>nstitut Montpelliérain Alexander Grothendieck, UMR CNRS  
5149, Place Eugene Bataillon, Montpellier, 34090, France.

<sup>3</sup>Centre Universitaire de Formation et de Recherche, Rue de  
l'Université, 97660, Dembeni, 97660, Mayotte, France.

\*Corresponding author(s). E-mail(s):

[solym-mawaki.manou-abi@umontpellier.fr](mailto:solym-mawaki.manou-abi@umontpellier.fr);

Contributing authors: [omar.hajjaji@univ-poitiers.fr](mailto:omar.hajjaji@univ-poitiers.fr);

[yousri.slaoui@univ-poitiers.fr](mailto:yousri.slaoui@univ-poitiers.fr);

<sup>†</sup>These authors contributed equally to this work.

## Abstract

In this paper, we consider parameter estimation of univariate  $\alpha$ -stable distributions and their mixture. We propose firstly an estimation method based on the characteristic function by the use of a Gaussian kernel estimator of the density distribution. The choice of the optimal bandwidth parameter was done using a plug-in method. We highlight another estimation procedure for the maximum likelihood framework based on the false-position algorithm to find a numerical root of the log-likelihood through the score functions. In the case of a mixture  $\alpha$ -stable distributions, the EM algorithm and the Bayesian estimation method were modified in order to propose an efficient tool for parameter estimation. Although we have limited the mixture study to two components, the proposed methods can be generalized to several components of mixture. A simulation study is conducted to evaluate the performance of our methods which are then applied to real data. Our results seems accurately estimating mixture of  $\alpha$ -stable distributions. The application concerned the estimation of reproduction

number of the Covid-19 in Mayotte and the Enzyme dataset density distribution. We compare the proposed methods along with a detailed discussion, and we conclude with some other forthcoming works.

**Keywords:** Stable distribution; Parametric estimation; Mixture model; EM; MCMC method.

## 1 Introduction

In recent decades, many researchers have shown interest in studying  $\alpha$ -stable distributions due to their ability to generalize widely used laws such as Gaussian, Lévy and Cauchy in order to handle impulsive and skewed data, which is particularly important in the financial field. In 1925, Levy discovered that  $\alpha$ -stable distributions arise as the limit of normalized sums of independent and identically distributed random variables. The family of stable distributions possesses skewness and tail thickness. Unfortunately, there is no closed-form expression for the cumulative distribution function and probability density function except some few cases such that the Gaussian, Lévy and Cauchy distribution. Let  $X$  be a stable random variable and  $\phi(t) = E(\exp(itX))$  it's characteristic function. Note that this family laws has multiple parameterizations. We follows the presentation in [1] and consider two type of representations. We write  $X \sim S(\alpha, \beta, \delta, \omega; 0)$  for the so called parametrization 0 and  $X \sim S(\alpha, \beta, \delta, \omega; 1)$  for the parametrization 1. A random variable  $X \sim S(\alpha, \beta, \delta, \omega; 0)$  if it's characteristic function is expressed as follows.

$$\begin{aligned} \phi(t) = & \exp\left(-\delta^\alpha t^\alpha \left[1 + i\beta \tan\left(\frac{\pi\alpha}{2}\right) \text{sign}(t)(\delta t^{1-\alpha} - 1)\right] + i\omega t\right) \mathbb{I}_{\{\alpha \neq 1\}} \\ & + \exp\left(-\delta t \left(1 + i\beta \frac{2}{\pi} \text{sign}(t) \log(\delta t)\right) + i\omega t\right) \mathbb{I}_{\{\alpha=1\}}, \end{aligned} \quad (1)$$

with  $\alpha \in (0, 2]$  is the index of stability,  $\beta \in [-1, 1]$  the skewness parameter;  $\delta > 0$  the scale parameter,  $\omega \in \mathbb{R}$  the location or shift parameter and  $\text{sign}(\cdot)$  is the signe function. The formula in (1) can be rewrite as follows.

$$\phi(t) = g(t) + ih(t) = \exp(-\delta^\alpha [t^\alpha + i\beta\eta(\delta t^\alpha; 0)] + i\omega t)$$

where the expression of  $\eta$  function is given by

$$\eta(\delta t|\alpha; 0) = \frac{2}{\pi} t \log(\delta t) \mathbb{I}_{\{\alpha=1\}} + \tan\left(\frac{\pi\alpha}{2}\right) \text{sign}(t) \delta^{1-\alpha} (t - t^\alpha) \mathbb{I}_{\{\alpha \neq 1\}}. \quad (2)$$

A random variable  $X$  is  $S(\alpha, \beta, \delta, \omega; 1)$  if,

$$\begin{aligned} \phi(t) = & \exp\left(-\delta t^\alpha \left[1 - i\beta \left(\tan\left(\frac{\pi\alpha}{2}\right)\right) \text{sign}(t)\right] + i\omega t\right) \mathbb{I}_{\{\alpha \neq 1\}} \\ & + \exp(-\delta t [1 + i\beta \frac{2}{\pi} \text{sign}(t) \log(t)] + i\omega t) \mathbb{I}_{\{\alpha = 1\}}. \end{aligned}$$

These parametrizations are commonly characterised by four parameters:  $\alpha$  the tail index or the index of stability that governs the heaviness of the tail bounded in  $(0, 2]$ , the skewness parameter  $\beta$ , which ranges in  $[-1, 1]$  and when  $\beta = 0$  we have a symmetric stable distribution. The parameter  $\delta$  denote the scale, and  $\omega$  the location parameter. When  $\delta = 1$  and  $\omega = 0$ , one says that  $X$  is a standard stable random variable. By specifying some of this parameters, for example: In the case of  $\alpha = 2$  and  $\beta = 0$ , we get the Gaussian distribution, same for  $\alpha = 1$  and  $\beta = 0$ , it becomes Cauchy distribution, and Levy when  $\alpha = \frac{1}{2}$  and  $\beta = 1$ .

The parametrization 0 is very important because the characteristic function, density and cumulative distribution function are continuous with respect to the four parameters. It is therefore well conditioned numerically for computing. While the parametrization 1 is suitable to study algebraic properties. Note that for  $\beta = 0$ , the two parametrizations are identical. It's worth mentioning that  $\alpha$ -stable distributions have an infinite variance when  $\alpha < 1$ , which complicates the estimation of the parameters in this case. There are many approaches to estimate parameters of such distributions, namely the Fractional Lower Order Moments (FLOM) method, as described in [7]. This method considers the lower-order fractional moments of the data to estimate the parameters of the distribution. It is also less sensitive to outliers and doesn't require any iterative optimization procedures. However, one limitation of the FLOM method is that it may not work well for data with heavy tails, as it relies on the estimation of fractional moments, which may not exist for  $\alpha$ -stable distributions with  $\alpha < 1$ . Additionally, the FLOM method assumes that the data are independently and identically distributed, which may not be the case in some real-world applications. Furthermore, as shown in [2], it doesn't guarantee a good estimation of the parameter of asymmetry  $\beta$ . Another approach is the quantile method improved by McCulloch [3] which is based on the relationship between the  $q$ -th quantile and the distribution parameters for any distribution. In the case of the  $\alpha$ -stable distribution, the  $q$ -th quantile is related to the scale and location parameters of the distribution, as well as the  $\alpha$  and  $\beta$  parameters, it's also robust to outliers and can handle missing data, but it may not perform well for distributions with slowly varying tails or non-smooth density functions. In the same context a new estimation algorithm for the tail index was proposed in [4] by considering a quantile conditional variance ratio. Based on the explicit formula (1) of the characteristic function, some authors established an estimation method (ECF) using regression, [1]. Even though the lack of closed-form expression for the probability density is theoretically a disadvantage, note that in practise they are computed numerically using an integral transform. It is in

this way that many authors performs the maximum likelihood (ML) method for parameter estimation, [1]. Finite mixture models are becoming increasingly popular and play a crucial role in density estimation. In particular, the mixture of  $\alpha$ -stable distributions is a popular tool for modeling skewed and impulsive data, making it applicable in various fields. When dealing with such mixture method, one need to initialize and choose the components of the mixture. The k-means clustering [5] allows to better estimate the components. A well known method of mixture estimation is the Expectation-Maximization algorithm (EM). We shall consider ECF (2.1) and ML (2.2) methods to estimate the four parameters of this distributions, in order to choose the most efficient of them to apply it in the EM algorithm. This algorithm is particularly useful in cases where the data is incomplete or partially observed. Many work has been done in this direction; we quoted the following papers of [6] and references therein. Let us mentioned the work of [7], [8] which presents a methodology for the EM algorithm applied to symmetric  $\alpha$ -stable distributions. In [9] the authors established some formula that involves symmetric and non symmetric  $\alpha$ -stable distributions in order to estimate parameters with the use of EM algorithm but the paper is still not published and we remark their algorithm does not converge for non symmetric  $\alpha$ -stable distributions. Since the Bayesian approach is also useful and being an efficient tool for parameter estimation in mixture models, we proposed a modified framework that involves updating the posterior distribution until convergence to the stationary distribution by combining the Gibbs sampling and Metropolis-Hastings algorithm, as described in [10]. The novelty of this latest approach compared to [10], is essentially the choice of the rejection zones in Metropolis-Hastings algorithm, that can significantly affect estimates additionally to the updating of parameters.

Mixture models, in general, could be used in the understanding and development of an epidemic through the estimation of the the generation time and reproduction number; see for instance the work in [11], where the authors consider mixture of weibull, lognormal and other distributions of mixture models to estimate the reproduction number during the period of Covid-19 outbreak in Mayotte island. It was announced in this paper that mixture estimation considering negative serial interval can enhance the estimates by treating  $\alpha$ -stable mixtures with the various Expectation- Maximization algorithm and Bayesian methods. Therefore, this work can be consider as an extension of this study.

In this paper, we have chosen to deal with the two parameterizations of  $\alpha$ -stable distributions in order to gain in performance, which can be affected according to the method used. Our contribution on this paper regarding the literature is the introduction of Kernel in the estimation method of characteristic function (ECF) which reveals to perform better in a simulation study when  $\alpha \geq 1$ . Additionally, we propose a new methodology that employs score function and False-position algorithm to perform the ML estimation. Furthermore, we have improved the EM algorithm of [6] and the Bayesian approach of [10], which also demonstrated efficient performance in a simulation study.

All these methods were used to estimate the reproduction number during the Covid-19 outbreak data of Mayotte as in [11]. We also give another application with the Enzyme data, see [12]. The use of K-means method on the above data allows to choose a two components in the mixture estimation methods of this paper. We also aims to establish new mixture properties results that involved non symmetric stable distributions as tried in the non published paper [9].

This paper is organized as follows. In Section 2, we first recall some basic quantiles method originally developed by McCulloch and introduce the Kernel estimation method of the characteristic function and the estimation procedure for  $\alpha$ -stable distributions. We also consider a new framework consideration of the ML estimation of  $\alpha$ -stable distributions based on score function and False-position algorithm. Next, in Section 3, we outline an adapted Bayesian and EM algorithm approaches to estimate the parameters of a two mixture of  $\alpha$ -stable distributions than can be generalized to more than two components mixtures. In Section 4, we apply the proposed methods with the above mentioned Covid-19 outbreak data of Mayotte together with the Enzyme data, a conclusion and future developments. The last section of this paper is an Appendix section, where we established some modified mixture properties in the case of non-symmetric  $\alpha$ -stable in order to extend the work of [9] for practical issues as a forthcoming work.

## 2 Parameter estimation of stable distributions

The aim of this section is to present the relevant methods that we consider in order to improve our mains results. We recall firstly, the quantiles method originally developed by McCulloch, as described in [3] provided by the `libstableR` package. We'll use such method to initialize our algorithms. The method has been simplified in recent years, and the version we use is based on the work presented in [1] that can be described as follows. Let  $x_p$  be the  $p^{\text{th}}$  quantile of  $X \sim S(\alpha, \beta, \delta, \omega; 0)$  distribution and define the following quantities:

$$\begin{aligned} v_\alpha(\alpha, \beta, \delta, \omega) &= \frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}}, \\ v_\beta(\alpha, \beta, \delta, \omega) &= \frac{x_{0.05} + x_{0.95} - 2x_{0.5}}{x_{0.95} - x_{0.05}}, \\ v_\delta(\alpha, \beta, \delta, \omega) &= x_{0.75} - x_{0.25}, \\ v_\omega(\alpha, \beta, \delta, \omega) &= -x_{0.5}. \end{aligned}$$

McCulloch in [3] provides tables, where one could deduce the value of the above parameters based on this quantities considering a standard stable distribution  $Z \sim S(\alpha, \beta, 1, 0; 0)$ . This is not restrictive since the scaling property, as outlined in [13] allow to have  $x_p = \delta z_p + \omega$ , where  $z_p$  is  $p^{\text{th}}$  quantile of  $Z$ . Then, one could deduce the value of the above parameters based on this quantities

based on the following relations.

$$\begin{aligned}v_{\alpha}(\alpha, \beta, \delta, \omega) &= v_{\alpha}(\alpha, \beta, 1, 0), \\v_{\beta}(\alpha, \beta, \delta, \omega) &= v_{\beta}(\alpha, \beta, 1, 0), \\v_{\delta}(\alpha, \beta, \delta, \omega) &= \delta v_{\delta}(\alpha, \beta, 1, 0), \\v_{\omega}(\alpha, \beta, \delta, \omega) &= \delta v_{\omega}(\alpha, \beta, 1, 0) - \omega.\end{aligned}$$

As we can see  $v_{\alpha}$  and  $v_{\beta}$  are independent of the scale and location parameters and the above relations allow to obtain a reliable estimation of the four parameters, when the sample set is large, see [1]. In the sequel the vector parameter  $\Theta_0 = (\alpha_0, \beta_0, \delta_0, \omega_0)$  will refer to the initial vector value parameter according to the McCulloch method. Now, we will introduce the other methods.

## 2.1 Estimation method based on characteristic function

Let  $X \sim S(\alpha, \beta, \delta, \omega; 0)$  and consider  $X_1, \dots, X_n$  a sample of size  $n$  of  $X$  and  $F_n$  the empirical cumulative distribution function:

$$\hat{\phi}_{n_1}(t) = \int_{\mathbf{R}} \exp(itx) dF_n(x) = \frac{1}{n} \sum_{j=1}^n \exp(itX_j).$$

Note that  $\hat{\phi}_{n_1}(t)$  is a consistent estimator of  $\phi(\cdot)$  for large value of  $n$  (by a simple application of the law of large number). We introduce an alternative estimation of characteristic function using Kernel method. We shall be concerned with the Gaussian kernel for the case when  $\alpha > 1$  expressed as  $K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right)$  for all  $x \in \mathbb{R}$  and satisfies  $\int_{\mathbf{R}} K(z) dz = 1$ . This alternative estimation of the characteristic function is defined as follow.

$$\hat{\phi}_{n_2}(t) = \frac{1}{nh_n} \int_{\mathbf{R}} \exp(itx) \sum_{j=1}^n K\left(\frac{x - X_j}{h_n}\right) dx,$$

where the bandwidth satisfies  $\lim_{n \rightarrow \infty} h_n = 0$  in order to have the consistency of  $\hat{\phi}_{n_2}$ . For practical setting and in the case  $\alpha > 1$ , we choose the optimal bandwidth  $h_n$  by considering the plug-in method proposed by Sheather and Jones [14] since it is based on square mean error. On the other hand (specifically for  $\alpha < 1$ ), we consider the plug-in method of second generation proposed by Slaoui in [15] (see also [16]). Now, we define the following sample functions  $g_n(t)$  and  $h_n(t)$  to compute the estimated characteristic function on sample size  $n$ :

$$\begin{cases} g_n(t) = \Re(\hat{\phi}_n(t)), \\ h_n(t) = \Im(\hat{\phi}_n(t)), \end{cases}$$

where  $\hat{\phi}_n(t)$  could be  $\hat{\phi}_{n_1}(t)$  or  $\hat{\phi}_{n_2}(t)$ . We have:

$$\begin{cases} \sqrt{g_n(t)^2 + h_n(t)^2} = |\hat{\phi}_n(t)| = \exp(-\delta^\alpha t^\alpha), \\ \arctan\left(\frac{h_n(t)}{g_n(t)}\right) = \arg(\hat{\phi}_n(t)) = -\delta^\alpha \beta \eta(\delta t | \alpha; 0) + \omega t. \end{cases}$$

Moreover, we define the following quantities:

$$\begin{cases} y(t) &= \log(-\log(|\phi(t)|)) = \log(\delta^\alpha) + \alpha \log(|t|), \\ z_k &= \arg(\hat{\phi}_n(t_k)), \\ B_k &= \hat{\delta}^{\hat{\alpha}} \eta(\hat{\delta} t_k | \hat{\alpha}; 0). \end{cases}$$

The relationship between  $y(t)$  and  $\log(|t|)$  can be modeled linearly with a slope of  $\alpha$  and an intercept of  $a = \log(\delta^\alpha)$ . Set

$$y_k = \log(-\log(|\hat{\phi}_n(t_k)|)) = \log\left(-\log\left(\sqrt{g_n(t_k)^2 + h_n(t_k)^2}\right)\right)$$

where  $t_k$  is choose on a grid of equally spaced points  $t_1, \dots, t_m$  using the sample data. We make use of the following weighted least squares method to minimize

$$\begin{aligned} S(a, \alpha) &= \sum_{k=1}^m W_k (y_k - a - \alpha \log(|t_k|))^2 \\ (\hat{a}, \hat{\alpha}) &= \arg \min_{(a, \alpha)} S(a, \alpha) \end{aligned}$$

where the weights  $W_k = \frac{1}{\sigma_k^2}$  with  $\sigma_k^2$  is the variance of the  $k$ -th observation. We obtain the following estimators:

$$\begin{aligned} \hat{\alpha}_m &= \left( \sum_{k=1}^m W_k \log(t_k) y_k - \frac{\sum_{k=1}^m W_k y_k}{\sum_{k=1}^m W_k} \times \sum_{k=1}^m W_k \log(t_k) \right) \\ &\quad \times \left( \sum_{k=1}^m W_k \log(t_k)^2 - \frac{\sum_{k=1}^m W_k \log(t_k)}{\sum_{k=1}^m W_k} \times \sum_{k=1}^m W_k \log(t_k) \right)^{-1}, \\ \hat{a}_m &= \frac{\sum_{k=1}^m W_j y_k - \hat{\alpha}_m \sum_{k=1}^m W_k \log(t_k)}{\sum_{k=1}^m W_j}, \\ \hat{\delta}_m &= \delta_0 \exp\left(\frac{\hat{a}_m}{\hat{\alpha}_m}\right). \end{aligned}$$

Then, from the following formula:

$$z_k = \beta B_k + \omega t_k,$$



we use again the above weighted least squares method to obtain

$$\begin{aligned}\hat{\omega}_m &= \delta_0 \left( \sum_{k=1}^m W_k B_k z_k - \frac{\sum_{k=1}^m W_k t_k z_k}{\sum_{k=1}^m W_k t_k B_k} \times \sum_{k=1}^m W_k B_k^2 \right) \\ &\quad \times \left( \sum_{k=1}^m W_k B_k t_k - \frac{\sum_{k=1}^m W_k t_k^2}{\sum_{k=1}^m W_k t_k B_k} \times \sum_{k=1}^m W_k B_k^2 \right)^{-1} + \omega_0, \\ \hat{\beta}_m &= \frac{\sum_{k=1}^m W_k t_k z_k - \hat{\omega}_m \sum_{k=1}^m W_k t_k^2}{\sum_{k=1}^m W_k t_k B_k}.\end{aligned}$$

Since  $\hat{\phi}_m$  is consistent, one could obtain the consistency of the estimated parameters  $\hat{\alpha}_m, \hat{\beta}_m, \hat{\delta}_m$  using the classical regression estimation method. A simulation study is conducted to evaluate the performance of the proposed estimation. The choice of  $m$  was done from [17], which suggest selecting points  $t_k$  in the interval  $[0.1, 1]$ .

## 2.2 Maximum likelihood method

In the following lines, we present a method for estimating stable distributions within the framework of maximum likelihood (ML) method. Since probability density function does not have a closed-form expression, the classical ML method does not apply in this context because the likelihood ratio doesn't exist explicitly. For this reason, we first make use of numerical approximation of the density function of stable distributions, which provides accurate estimates. We introduce the following method based on score functions and the False-Position algorithm, which is an intuitive way to estimate the underline parameters of the stable distributions. Assume that we are in parametrization 1 and  $\alpha > 1$ . Set

$$\begin{aligned}g_d(x|\alpha, \beta) &= \int_0^\infty \cos(xr + \beta\eta(r, \alpha; 1))r^{d-1} \exp(-r^\alpha)dr 1_{\{0 < d < \infty\}} \\ &\quad + \int_0^\infty [\cos(xr + \beta\eta(r, \alpha; 1)) - 1]r^{d-1} \exp(-r^\alpha)dr 1_{\{-2 \min(1, \alpha) < d \leq 0\}},\end{aligned}$$

$$\begin{aligned}\tilde{g}_d(x|\alpha, \beta) &= \int_0^\infty \sin(xr + \beta\eta(r, \alpha; 1))r^{d-1} \exp(-r^\alpha)dr 1_{\{-\min(1, \alpha) < d < \infty\}} \\ &\quad + \int_0^\infty [\sin(xr + \beta\eta(r, \alpha; 1)) - xr]r^{d-1} \exp(-r^\alpha)dr 1_{\{\alpha > 1, -\alpha < d \leq -1\}},\end{aligned}$$

$$h_d(x|\alpha, \beta) = \int_0^\infty \cos(xr + \beta\eta(r, \alpha; 1)) \log(r)r^{d-1} \exp(-r^\alpha)dr.$$

$$\tilde{h}_d(x|\alpha, \beta) = \int_0^\infty \sin(xr + \beta\eta(r, \alpha; 1)) \log(r)r^{d-1} \exp(-r^\alpha) dr,$$

For  $x \in \mathbb{R}$  and  $d \in \mathbb{N}$ . Where  $\eta(r, \alpha; 1) = -\tan(\frac{\pi\alpha}{2}r^\alpha)$  should not be confused with the previous one in (2) from parameterization 0. Let recall the following theorem from [1].

**Theorem 1** (Stable score function) *Let  $\alpha \neq 1$ . The univariate stable density in the 1-parameterization is given by:*

$$f(x|\alpha, \beta, \delta, \omega; 1) = \frac{1}{\pi\delta} g_1\left(\frac{x-\omega}{\delta}|\alpha, \beta\right).$$

*Then, the score functions which is the derivative of this density on each parameter are:*

$$\begin{aligned} \frac{\partial f}{\partial \alpha}(x|\alpha, \beta, \delta, \omega; 1) &= \frac{1}{\pi\delta} \left[ \frac{\pi\beta}{2\cos(\frac{\pi\alpha}{2})^2} \tilde{g}_{1+\alpha}\left(\frac{x-\omega}{\delta}|\alpha, \beta\right) \right. \\ &\quad \left. + \beta \tan\left(\frac{\pi\alpha}{2}\right) \tilde{h}_{1+\alpha}\left(\frac{x-\omega}{\delta}|\alpha, \beta\right) - h_{1+\alpha}\left(\frac{x-\omega}{\delta}|\alpha, \beta\right) \right], \\ \frac{\partial f}{\partial \beta}(x|\alpha, \beta, \delta, \omega; 1) &= \frac{\tan(\frac{\pi\alpha}{2})}{\pi\delta} \tilde{g}_{1+\alpha}\left(\frac{x-\omega}{\delta}|\alpha, \beta\right), \\ \frac{\partial f}{\partial \delta}(x|\alpha, \beta, \delta, \omega; 1) &= -\frac{1}{\pi\delta^2} g_1\left(\frac{x-\omega}{\delta}|\alpha, \beta\right) + \frac{x-\omega}{\pi\delta^3} \tilde{g}_2\left(\frac{x-\omega}{\delta}|\alpha, \beta\right), \\ \frac{\partial f}{\partial \omega}(x|\alpha, \beta, \delta, \omega; 1) &= \frac{1}{\pi\delta^2} \tilde{g}_2\left(\frac{x-\omega}{\delta}|\alpha, \beta\right). \end{aligned}$$

Many equations, including most of the more complicated ones, can be solved only by iterative numerical approximation. The method of false position is one of them. This consists of trial and error, in which various values of the unknown quantity are tried. That trial-and-error may be guided by calculating, at each step of the procedure, a new estimate for the solution. Let  $c$  be the root of a given continuous function  $g$  defined on the interval  $[a_0, a_1]$  such that  $g(a_0)$  and  $g(a_1)$  are of opposite signs. Then there exists a root  $c$  in the interval  $[a_0, a_1]$  by the intermediate value theorem. There are many root-finding algorithms that can be used to obtain approximations to such a root. One of the most common is Newton's method or secant method. but it can fail to find a root under certain circumstances and it may be computationally costly since it requires a computation of the function derivatives. Other methods are needed and one general class of methods are the two-point bracketing methods. These methods proceed by producing a sequence of shrinking intervals  $[a_k, b_k]$  at the  $k$ -th step, such that  $(a_k, b_k)$  contains a root of  $g$ . The simplest variation, called the bisection method, calculates the solution estimate as the midpoint of the bracketing interval:

$$c = \frac{g(a_1)a_0 - a_1g(a_0)}{g(a_1) - g(a_0)}.$$

The convergence rate of the bisection method could possibly be improved by using a different solution estimate. The false-position method calculates

the new solution estimate as the  $x$ -intercept of the line segment joining the endpoints of the function on the current bracketing interval. This method runs as the first iteration of the bisection method [18]. Essentially, the root is being approximated by replacing the actual function by a line segment on the bracketing interval and then using the classical double false position formula on that line segment, see [18]. Note that the rate of convergence can drop below that of the bisection method. We'll use this methodology to establish our algorithm by finding the roots of the derivative of the above log-likelihood function, which is equal to:

$$\sum_{i=1}^n \frac{\partial \log(f(x_i|\alpha, \beta, \delta, \omega; 1))}{\partial \alpha} = \sum_{i=1}^n \frac{\partial f(x_i|\alpha, \beta, \delta, \omega; 1)}{\partial \alpha} / f(x_i|\alpha, \beta, \delta, \omega; 1).$$

Obviously, since this method requires a good range for each parameter. For this reason we consider the intervals  $[\alpha_0, \alpha_1]$ ,  $[\beta_0, \beta_1]$ ,  $[\delta_0, \delta_1]$ , and  $[\omega_0, \omega_1]$  defined by adding a given margin to the initial values  $\Theta_0 = (\alpha_0, \beta_0, \delta_0, \omega_0)$  based on the Quantile method. The same procedure is used for the other three parameters. In the sequel we refer to this method as ML-Second. However, at this stage, one can be disappointed by not giving a theoretical study about the existence of the root solutions of the above score functions under the false position method. This will be addressed in a forthcoming work. But we're going to look at convergence numerically.

Note that, one can also make use of the `optim` command in R to directly maximize the log-likelihood function and obtain the estimated parameters. This will be referred to as ML-First in the sequel. The performance and consistency of the estimation will be illustrated in a simulation study.

### 3 Parameter estimation for mixture of $\alpha$ -stable distributions

The mixture of  $\alpha$ -stable distributions is a popular tool for modeling skewed and impulsive data, making it applicable in various fields. Two common methods for inferring parameters in mixture models are: the Expectation-Maximization method algorithm and the Bayesian approach, [6, 10]. Bayesian estimation is a framework for the formulation of statistical inference problems. In the prediction or estimation of a random variable or process, the Bayesian philosophy is based on combining the evidence contained in the random variable with prior knowledge of the probability distribution of the random variable. Bayesian estimation method uses prior data to estimate the value of the unknown parameters. This reduces the difference between the estimator and the actual value of that parameter. In Bayesian modeling, the selection of priors plays then a crucial role in determining the posterior inference. The Expectation-Maximization (EM) algorithm is a widely used computational method for estimating the parameters of statistical models with latent or missing variables. This algorithm is particularly useful in cases where the data

is incomplete or partially observed. The EM algorithm works by iteratively alternating between the E-step, where we estimate the expected value of the unobserved or latent variables given the current parameter estimates and the M-step, where we maximize the likelihood of the observed data based on the expected values obtained in the E-step. This alternating process continues until convergence is achieved, resulting in the optimal parameter estimates for the model. Let  $n$  be the number of observations and  $z_i$  the latent observations,  $i = 1, \dots, n$  and denote by  $\lambda_1 = \mathbb{P}(z_i = 1)$  the weight. To obtain the mixture component parameter, we use in this paper the k-means clustering method [5]. We assume in the rest of this paper that we are in a two component mixture. Of course this work can be generalized to more than two component.

### 3.1 The proposed Expectation-Maximization algorithm

We introduce in this paper an adapted Expectation-Maximization method by including the above parameter estimation tools namely the ML estimation method and the Estimation through the the characteristic function ECF (Empirical or Kernel) both for the vector parameter  $\Theta$  and for updating the E-step in the EM algorithm. Such adaptation are important when selecting the appropriate parameter estimation method in the EM method. Our proposed Expectation-Maximization algorithm is given as follows.

---

**Algorithm 1** Algorithm EM for mixture of  $\alpha$ -stable distributions

---

- 1: Initialization of the model, with choosing a specific tolerance  $\epsilon$ .
- 2: **repeat**
- 3:   **Step-E:**
- 4:   We compute the posterior probability
- 5:   **for**  $i = 1, \dots, n$  **do**
- 6:     **for**  $j = 1, 2$  **do**
- 7:       The observation  $i$  belongs to component  $j$  with probability

$$P_{i,j}^{(t)} = \frac{\lambda_j^{(t)} f_j(x_i | \alpha_1^{(t)}, \beta_1^{(t)}, \delta_1^{(t)}, \omega_1^{(t)}; 0)}{\lambda_1^{(t)} f_1(x_i | \alpha_1^{(t)}, \beta_1^{(t)}, \delta_1^{(t)}, \omega_1^{(t)}; 0) + \lambda_2^{(t)} f_2(x_i | \alpha_2^{(t)}, \beta_2^{(t)}, \delta_2^{(t)}, \omega_2^{(t)}; 0)}$$

- 8:     **end for**
  - 9:   **end for**
  - 10:   **Step-M:**
  - 11:   **for**  $j = 1, 2$  **do**
  - 12:      $\lambda_j^{(t+1)} = \frac{1}{n} \{\#z_i = j\}$ .
  - 13:     Then, we use ML or ECF to obtain  $\Theta_j^{(t+1)} = (\alpha_j^{(t+1)}, \beta_j^{(t+1)}, \delta_j^{(t+1)}, \omega_j^{(t+1)})$
  - 14:     We compute the log-likelihood at iteration  $t + 1$ , we note it:  $Q^{(t+1)}$
  - 15:   **end for**
  - 16: **until**  $\frac{|Q^{(t+1)} - Q^{(t)}|}{|Q^{(t+1)}|} < \epsilon$ .
-

Now, let us turn into another interesting method that is known for its flexibility namely the Bayesian estimation method.

## 3.2 The proposed Bayesian algorithm

As state above, the central idea behind Bayesian estimation is that before we have seen any data, we already have some prior knowledge about the distribution it came from. Such prior knowledge usually comes from experience or past experiments. In this section we consider uniform priors for the parameters  $\alpha$  and  $\beta$ ; an inverse gamma prior for the parameter  $\delta$ , and a normal prior for the parameter  $\omega$ . Moreover we shall consider a symmetric Dirichlet prior for the weights, see [10]. Each of these priors has been deliberately chosen to be independent from one another to ensure that our model remains free of any unwanted biases. It is imperative to note that the choice of priors can significantly influence the overall results of the model and should be chosen with care. The Bayesian approach involves updating the posterior distribution of the parameters using Baye's theorem. Note that computing the joint posterior distribution of the above parameters given the data and the priors is often analytically intractable because of the lack of closed-form expression for  $\alpha$ -stable densities. To overcome this issue, we'll use Markov chain Monte Carlo (MCMC) methods more precisely a combination of Gibbs sampling and Metropolis-Hastings algorithms (see [10]), as illustrated in the following steps.

### 3.2.1 The distribution of the weights:

We assume that the prior distribution of  $\lambda$  follows a symmetric Dirichlet distribution with parameters  $\zeta$ . Since  $\mathbb{P}(z_i = j)$  is equal to  $\lambda_j$  for  $j = 1, 2$  and  $i = 1, \dots, N$ , where  $N$  is the number of observations. The full conditional distribution for  $\lambda$  is also a Dirichlet distribution, with updated parameters  $\zeta + n_j$ , where  $n_j$  is the number of observations assigned to component  $j$ . Thus, the updated distribution for weights is  $\lambda | \Theta \sim D(\zeta + n_1, \zeta + n_2)$ .

### 3.2.2 Updating vector parameter $\Theta$ using MCMC:

In this step, we consider the Metropolis-Hastings sampling method. We generate a candidate parameter  $\Theta_j^{new} = (\alpha_j^{new}, \beta_j^{new}, \delta_j^{new}, \omega_j^{new})$ , from a proposal distribution  $q(\cdot)$ , and it's accepted with probability  $A_{\Theta_j^{new}}$ , defined by:

$$A_{\Theta_j^{new}} = \min \left( 1, \prod_{i=1, z_i=j}^N \frac{f(x_i | \alpha_j^{new}, \beta_j^{new}, \delta_j^{new}, \omega_j^{new}; 0)}{f(x_i | \alpha_j^{old}, \beta_j^{old}, \delta_j^{old}, \omega_j^{old}; 0)} \times \frac{p(\Theta_j^{new})q(\Theta_j^{old} | \Theta_j^{new})}{p(\Theta_j^{old})q(\Theta_j^{new} | \Theta_j^{old})} \right).$$

We also notice that the priors are independent. Then, we get:

$$p(\Theta) = p(\alpha)p(\beta)p(\delta)p(\omega).$$

In this work, we choose a normal distribution for  $q(\cdot|\cdot)$ . By symmetry, we conclude that

$$q(\Theta_j^{new}|\Theta_j^{old}) = q(\Theta_j^{old}|\Theta_j^{new}).$$

Then  $A_{\Theta_j^{new}}$  become:

$$\min \left( 1, \prod_{i=1, z_i=j}^N \frac{f(x_i|\alpha_j^{new}, \beta_j^{new}, \delta_j^{new}, \omega_j^{new}; 0)}{f(x_i|\alpha_j^{old}, \beta_j^{old}, \delta_j^{old}, \omega_j^{old}; 0)} \times \frac{IG(\delta_j^{new}|\alpha_0, \beta_0)N(\omega_j^{new}|\epsilon, k)}{IG(\delta_j^{old}|\alpha_0, \beta_0)N(\omega_j^{old}|\epsilon, k)} \right). \quad (3)$$

Then, we sample a uniform variable  $u$  in  $[0, 1]$ . If  $A_{\Theta_j^{new}} > u$ , we accept the new candidate variables, otherwise we keep those of the previous iteration. The fact that we consider a single rejection zone associated with the vector parameter  $\Theta$  is possible since the priors are assumed to be independent. Thus the Markov chain  $\tilde{\Theta}_n = (\tilde{\alpha}_n, \tilde{\beta}_n, \tilde{\delta}_n, \tilde{\omega}_n)$  where  $n$  is the index of iteration is stationary. Unlike [10] where the authors consider a multiple Markov chain for each parameter without taking advantage of the independence.

### 3.2.3 Updating the allocation parameter

It's necessary at each iteration to predict which subpopulation that each observation belongs. We do that, by computing the conditional probability,

$$\mathbb{P}(z_i = j|\Theta) = \frac{\lambda_j f_j(x_i|\alpha_j, \beta_j, \delta_j, \omega_j; 0)}{\lambda_1 f_1(x_i|\alpha_1, \beta_1, \delta_1, \omega_1; 0) + \lambda_2 f_2(x_i|\alpha_2, \beta_2, \delta_2, \omega_2; 0)}. \quad (4)$$

which is the probability that the observation  $x_i$  belongs to the component  $j$ . Note that, this method requires ordered steps to converge to the correct distribution, similar to the approach detailed in [10], where the reversible jump Markov chain Monte Carlo was used to determine the number of components of the mixture model. However, in contrast to the approach in [10], we consider equation (3) as the rejection zone in the Metropolis-Hastings step for all the parameters, which seems to be numerically more accurate than considering rejection zones for each parameter separately (considering the bias generated). We hence summarize our adapted method in the following algorithm:

**Algorithm 2** Bayesian Approach for a Mixture of  $\alpha$ -Stable Distributions**Require:** Initialization of the parameters as explained by K-means.**Require:** Number of iterations  $N$  and burn-in  $M$ .

- 1: **for**  $t = 1, \dots, N$  **do**
- 2:   Obtain weights  $\lambda = (\lambda_1, \lambda_2)$  by drawing samples from a symmetric Dirichlet distribution  $\lambda \sim D(\zeta + n_1, \zeta + n_2)$  with  $n_1$  as the number of observations assigned to the first component, and  $n_2$  for the second.
- 3:   Update parameters of the proposal distribution  $q(\cdot|\cdot) = N(\cdot|\theta, \sigma)$ , setting  $\theta$  to the value of the previous iteration and choosing a small value for  $\sigma$  (the standard deviation).
- 4:   Sample new candidates  $\Theta_j^{new} = (\alpha_j^{new}, \beta_j^{new}, \delta_j^{new}, \omega_j^{new})$  from the proposal distribution  $q(\cdot|\cdot) = N(\cdot|\theta, \sigma)$  for each component.
- 5:   Accept  $\Theta_j^{new}$  according to equation (3) and set  $\Theta_j^t = \Theta_j^{new}$ , otherwise set  $\Theta_j^t = \Theta_j^{t-1}$ .
- 6:   **for** each observation  $x_i$  **do**
- 7:     Obtain the allocation variable  $z_i$  using equation (4).
- 8:   **end for**
- 9: **end for**
- 10: Calculate the mean parameters:  $\Theta_j = \frac{1}{N-M} \sum_{k=M}^N \Theta_j^{(k)}$ .

Now, since we have fulfilled the main methodology of interest in this paper, the upcoming section will focus on assessing the efficacy of each approach with simulated data and then applied to real data.

## 4 Applications

This section deals firstly with the estimation method based on characteristic function and the maximum likelihood method for estimating the parameters of  $\alpha$ -stable distributions. We first evaluate the performance through a simulation study of these methods and assess the effect of observation size  $n$  with varying parameter values. The results of these simulations are presented in Tables 1 to 4. We consider the Gaussian Kernel in the ECF method (denoted by ECF-Kernel) and the ECF method with empirical characteristic function is denoted by ECF-Empirical as presented in Section 2.1. Additionally, we used the ML method (First and Second). It's worth mentioning that using the ML-Second method, one observe that numerical convergence holds for the case  $\alpha > 1$  maybe because  $\alpha$ -stable distributions have an infinite variance when  $\alpha < 1$ . Consequently, we didn't provide an estimation of the parameters using the ML-Second method in that case, as shown below.

**Table 1** Comparison of Methods for Parameter Estimation-Configuration 1

Parameter	True value	n	ECF-Kernel	ECF-Empirical	ML-First	ML-Second
$\alpha$	1.6	500	1.6356	<b>1.6023</b>	1.6078	1.6151
		750	<b>1.6191</b>	1.5484	1.5697	1.7641
		1000	<b>1.6075</b>	1.5519	1.5443	1.6894
$\beta$	-0.8	500	0.7541	<b>-0.7566</b>	-0.6692	-0.3714
		750	-0.8375	-0.7668	-0.7389	<b>-0.8081</b>
		1000	-0.7696	-0.7337	<b>-0.7781</b>	-0.7756
$\delta$	5	500	<b>5.0124</b>	4.8583	4.8791	5.0498
		750	5.2009	<b>4.9608</b>	4.9314	5.4934
		1000	5.1666	<b>4.9888</b>	4.9418	5.2904
$\omega$	12	500	11.8914	<b>11.9563</b>	11.8727	12.8558
		750	<b>12.2187</b>	12.2930	12.2241	12.8333
		1000	<b>12.1193</b>	12.1905	12.2968	12.5753

**Table 2** Comparison of Methods for Parameter Estimation - Configuration 2

Parameter	True value	n	ECF-Kernel	ECF-Empirical	ML-First	ML-Second
$\alpha$	1.4	500	<b>1.3848</b>	1.3243	1.3588	1.3605
		750	<b>1.3973</b>	1.3613	1.3649	1.4420
		1000	1.3329	1.2985	1.3210	<b>1.3924</b>
$\beta$	0.5	500	0.4304	0.4464	<b>0.5150</b>	0.5800
		750	<b>0.5173</b>	0.5651	0.5523	0.6135
		1000	0.5381	0.5542	<b>0.5280</b>	0.6106
$\delta$	2	500	<b>1.9819</b>	1.8854	1.9495	1.9278
		750	2.1368	<b>2.0573</b>	2.0668	2.2048
		1000	2.0910	<b>2.0162</b>	2.0397	2.2395
$\omega$	-10	500	-9.9746	-10.0041	<b>-10.0032</b>	-8.4292
		750	-9.9876	-10.0300	<b>-10.0021</b>	-9.2277
		1000	-10.0807	-10.1048	<b>-10.0705</b>	-10.8141

**Table 3** Comparison of Methods for Parameter Estimation - Configuration 3

Parameter	True value	n	ECF-Kernel	ECF-Empirical	ML-First
$\alpha$	0.8	500	0.9831	0.8523	<b>0.7717</b>
		750	0.8601	<b>0.7891</b>	0.7730
		1000	0.8614	0.7502	<b>0.7510</b>
$\beta$	0.8	500	<b>0.7755</b>	0.7249	0.8436
		750	0.8903	0.9398	<b>0.8166</b>
		1000	0.6330	0.7373	<b>0.8056</b>
$\delta$	3	500	3.3389	3.1960	<b>3.0887</b>
		750	3.2231	<b>3.0899</b>	3.0994
		1000	3.0905	2.8824	<b>2.9579</b>
$\omega$	-12	500	-11.5866	-11.6107	<b>-11.9937</b>
		750	<b>-12.0331</b>	-12.1348	-12.0439
		1000	<b>-11.9651</b>	-12.0894	-12.1991

**Table 4** Comparison of Methods for Parameter Estimation - Configuration 4

Parameter	True value	n	ECF-Kernel	ECF-Empirical	ML-First
$\alpha$	0.6	500	0.7295	0.6118	<b>0.5932</b>
		750	0.7136	<b>0.5949</b>	0.5852
		1000	0.6826	0.6447	<b>0.5801</b>
$\beta$	-0.5	500	-0.3862	<b>-0.5280</b>	-0.5449
		750	-0.3669	-0.4413	<b>-0.4831</b>
		1000	-0.4074	<b>-0.4941</b>	-0.4436
$\delta$	4	500	4.7518	<b>4.0385</b>	4.0950
		750	4.3565	3.7984	<b>3.8458</b>
		1000	4.4246	<b>4.2515</b>	3.7425
$\omega$	5	500	4.1715	4.4790	<b>4.8560</b>
		750	4.5660	4.7277	<b>5.1534</b>
		1000	4.7741	<b>4.9808</b>	5.1215



These results seems to confirm that the ECF-Kernel estimation method is better than the ECF-Empirical method, which in our case whenever  $\alpha > 1$ .

We also consider in a second way, mixture estimation framework of two  $\alpha$ -stable distributions with distinct parameters. The objective is to evaluate the precision and efficacy of the methods described earlier namely our adapted Expectation-Maximization method (including the use of ECF-Kernel, ECF-Empirical and ML-First to update the parameters in the **M**-step of EM algorithm) and Bayesian method. Let  $\Theta_1 = (\alpha_1, \beta_1, \delta_1, \omega_1)$  and  $\Theta_2 = (\alpha_2, \beta_2, \delta_2, \omega_2)$ , the density of the mixture model is giving by:

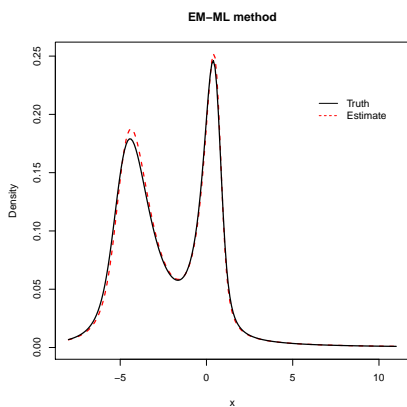
$$f(x, \lambda_1, \Theta_1, \Theta_2; 0) = \lambda_1 \times f(x, \Theta_1; 0) + (1 - \lambda_1) \times f(x, \Theta_2; 0).$$

Upon implementing the four methods, we obtain the following table of estimated parameters:

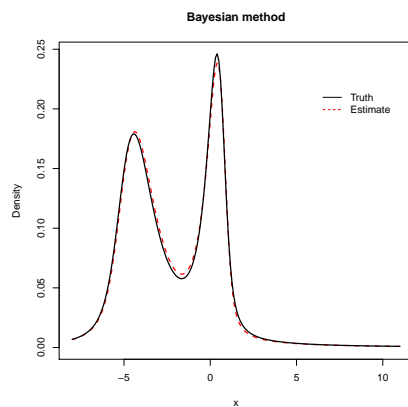
**Table 5** Comparison of Methods for Parameter Estimation of the Mixture Model

Parameter	True value	Bayesian	EM-ECF-Kernel	EM-ECF-Empirical	EM-ML
$\alpha_1$	1.2	<b>1.2032</b>	1.2099	1.2886	1.1689
$\beta_1$	0.5	<b>0.5015</b>	0.2977	0.4095	0.4906
$\delta_1$	1	<b>0.9988</b>	1.0338	1.0253	0.9638
$\omega_1$	-4.25	<b>-4.1939</b>	-4.1146	-4.1474	-4.1906
$\lambda_1$	0.6	0.6037	0.622	<b>0.599</b>	<b>0.601</b>
$\alpha_2$	1.2	<b>1.1812</b>	1.4926	1.1579	1.1434
$\beta_2$	-0.5	-0.6354	-0.6280	-0.3839	<b>-0.5497</b>
$\delta_2$	0.5	<b>0.5062</b>	0.5155	0.5103	0.4887
$\omega_2$	0.3	<b>0.3043</b>	0.2694	0.2808	0.3157

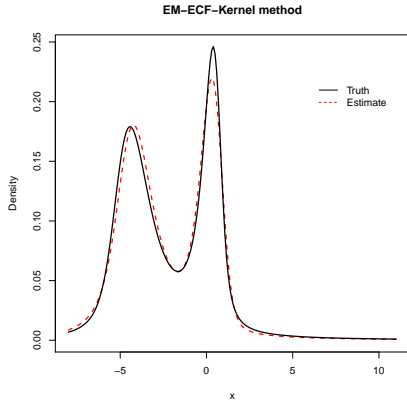
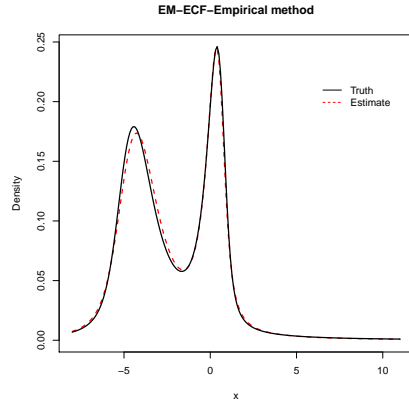
The corresponding graphs to visualize the plots is presented from Figure 1 to 4.



**Fig. 1** EM-ML method



**Fig. 2** Bayesian method

**Fig. 3** EM-ECF-Kernel method**Fig. 4** EM-ECF-Empirical method

The above simulation performance show numerically accurate parameter estimation for the mixture of  $\alpha$ -stable distributions with a given two components, which can be extended to more than two components. Now let us propose an application of the above described methods to real data, namely the serial interval (SI) and enzyme dataset distribution.

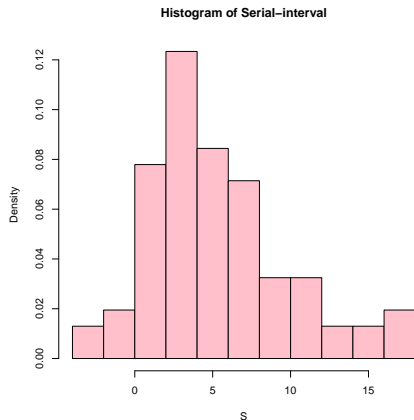
To enhance the understanding of the estimated parameters, we will provide tables and compare AIC and BIC criteria. The COVID-19 pandemic has caused extensive damage across the globe, disrupting productivity and triggering widespread panic. In the French region of Mayotte, the Regional Health Agency has undertaken significant efforts to collect and monitor the propagation, as documented in [11] through the estimation of the time varying reproduction number which is a non-pharmaceutical monitoring tool. But this need to determine firstly the transmission dynamics based on the serial interval described as follow using the data involved in [11]:

$$\text{serial\_interval} = y - \left( \frac{x.\text{ub} + x.\text{lb}}{2} \right).$$

with:

- $x.\text{lb}$ : Lower limit dates of onset of symptoms of infectors.
- $x.\text{ub}$ : Upper limit dates of onset of symptoms of infectors.
- $y$ : Dates of onset of symptoms of the infectee.

The associated histogram of this interval is presented in Figure 5, which will later be used to adjust the curves obtained by each approach:

**Fig. 5** Serial-interval

Upon examining the histogram of the distribution, we noticed that it exhibits left-skewness and is not symmetric. Consequently, we propose applying a mixture of two  $\alpha$ -stable distributions not necessarily symmetric. The K-means algorithm was utilized to select two components. We observed that the EM-ML method performed exceptionally well during our analysis, owing to the small dataset size, the peaked distribution, and the fact that the allocation vector changes at each iteration, which may impact model parameter estimation. We chose EM-ML due to its effectiveness, as demonstrated in [6]. To validate our decision, we calculated the AIC and BIC criteria for all four methods, which led to this optimal choice. While the AIC and BIC values were relatively close to each other, we visually confirmed this selection in Figure 4. The tables associated with these methods are presented below.

**Table 6** Estimated Parameters for Serial-Interval dataset

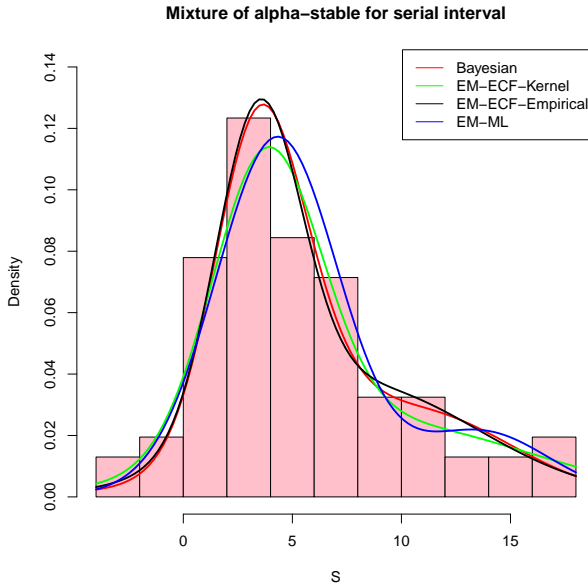
Parameter	Bayesian	EM-ECF-Kernel	EM-ECF-Empirical	EM-ML
$\alpha_1$	1.7140	2	2	1.9614
$\beta_1$	0.4010	1	1	-0.8793
$\delta_1$	1.6734	1.7524	1.3736	2.0038
$\omega_1$	3.6323	3.8096	3.4368	4.2990
$\lambda_1$	0.7111	0.5974	0.5064	0.8311
$\alpha_2$	1.7625	2	2	2
$\beta_2$	-0.6113	-1	-1	0
$\delta_2$	3.2759	4.7511	3.8474	2.2267
$\omega_2$	10.8648	9.0474	8.0447	13.5769

The estimated curves obtained by combining the values of each method on the histogram are displayed in the figure 6. It is evident that the curves generated by the EM-ML method closely resemble the shape of the histogram, indicating a good fit. However, the curves generated by the other four methods show deviation from the actual shape. Therefore, overall, the mixture of two

**Table 7** Table of comparison between criterion of selection for Serial-interval dataset

Methods	EM-ML	EM-ECF-Kernel	EM-ECF-Empirical	Bayesian
AIC	<b>455.2616</b>	457.9891	455.935	456.9538
BIC	<b>476.3558</b>	479.0833	477.0293	478.048

$\alpha$ -stable distributions using the EM-ML method provides the best fit of the serial interval.

**Fig. 6** Mixture methods applied on Serial-interval

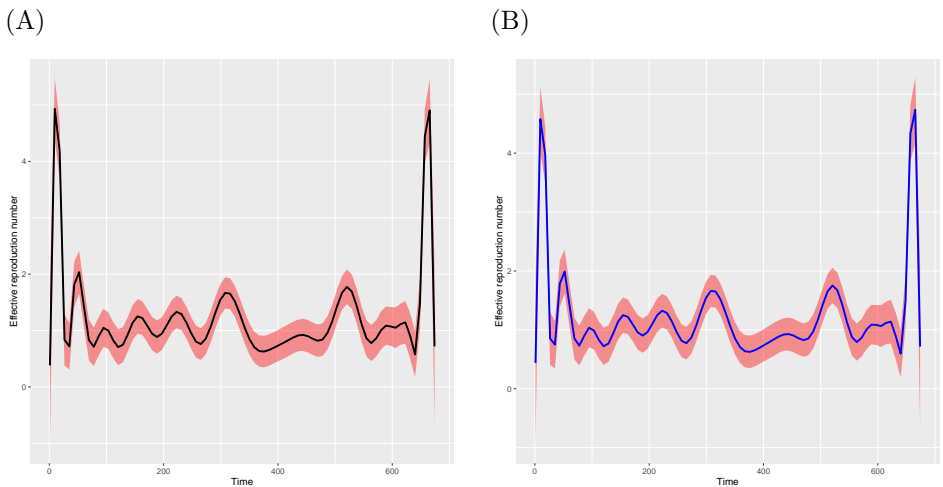
### *Time varying Reproduction Number*

The basic reproduction number  $R_0$  at the start of an epidemic and the time varying reproduction numbers during a development of an epidemic is an important tool. It is historically defined as the average number of new cases of infection generated by an individual during a period of infectivity (see [19]). There are several methods for computing this parameter. We consider a non-parametric approach [20] based on the generation time function linked to the serial interval distribution; (see [21, 22]). We give an estimation of this generation time distribution taking into account an  $\alpha$ -stable mixture modeling framework. Indeed, the choices of serial interval estimation models in the R0 package (see [21, 22]) are fixed family distributions, which are Gamma, Log-normal and Weibull. A mixture consideration of such distributions was treated in [11] but do not allow mixture of  $\alpha$ -stable distributions, which consider also the negative serial interval.

Let  $\mu$  be the probability distribution of the transmissibility of an infectious individual at the age of infection  $a$ , assuming that the entire population is susceptible. Let  $\Gamma(t)$  be the number of new infections during the time interval  $]t; t + dt[$  and  $s(t)$  the proportion of susceptibles in the population. For discrete time  $t \neq 0$ , we have the following non parametric formula of the time varying reproduction number

$$R_0(t) = \frac{\Gamma(t)}{\sum_{\tau < t} \mu(\tau) \Gamma(t - \tau)}. \quad (5)$$

Other improved estimation methods of the reproduction number also exist in the literature, see [23, 24] and references therein. In this work, using the epidemic incidence curve in Mayotte (between 13 March 2020 to 11 January 2022), we derive a generation time distribution in order to estimate the time varying reproduction number  $R_0(t)$  using the non parametric formula in equation (5) and the best-fit of the SI in an  $\alpha$ -stable mixture modeling framework. The following pictures in Figure 7 represent the evolution of the time varying reproduction number  $R_0(t)$ ; we make the curve smooth using estimated values.



**Fig. 7** Evolution of the time varying reproductive number in Mayotte from 13 March 2020 to 11 January 2022 with a mixture of  $\alpha$ -stable distribution on Serial-interval estimated by EM-ML (A), and estimated by EM-ECF-Empirical (B).

Looking at the graphs of figure 7, we notice a similarity between them and for both, we can see that the reproduction number starts at a value of around 2.5, indicating that each infected individual is, on average, infecting 2-3 other people. As time progresses, we see a sharp increase in the reproduction number, peaking at around 4.5, indicating that the disease is spreading rapidly. This could be due to a variety of factors, such as increased travel,

relaxed social distancing measures, or a new variant of the disease that is more transmissible. After the peak, we see a decline in the reproduction number, indicating that the disease is spreading more slowly. This could be due to interventions such as increased vaccination rates, stricter social distancing measures, or natural immunity acquired by those who have recovered from the disease. The reproduction number eventually drops below 1, indicating that the disease is no longer spreading and may be considered under control. But, at the end the reproduction number suddenly increase due to the effect of the new variant omicron. This task builds on previous work presented in [11], which focused exclusively on the non-negative portion of the serial interval. In this paper, we will examine a second dataset that contains a significantly greater number of observations than the serial interval. The results of the four approaches outlined in Section 3 will be presented as follows.

Finally let us consider here a dataset relates to the distribution of enzymatic activity in the blood of 245 unrelated individuals for an enzyme involved in the metabolism of carcinogenic substances, which is available in [12]. Mixture models have been applied to this dataset distribution as it contains two subpopulations of slow and fast metabolizers within the population.

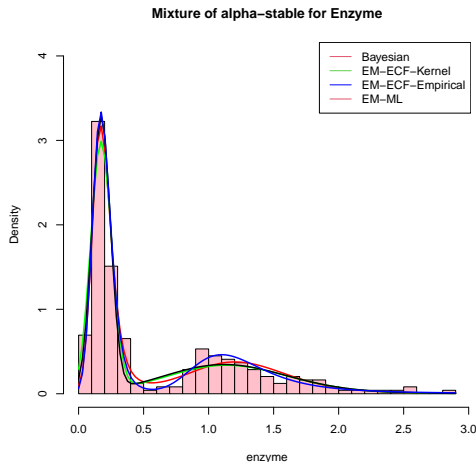
**Table 8** Table of estimated parameters for enzyme dataset

Parameter	Bayesian	EM-ECF-Kernel	EM-ECF-Empirical	EM-ML
$\alpha_1$	1.5936	1.9857	2	1.6939
$\beta_1$	0.8501	1	1	0.9999
$\delta_1$	0.0549	0.0565	0.0499	0.0525
$\omega_1$	0.1756	0.1726	0.1741	0.1752
$\lambda_1$	0.6188	0.5918	0.5755	0.6244
$\alpha_2$	1.7175	1.8921	1.8864	1.3883
$\beta_2$	0.6134	1	1	0.9964
$\delta_2$	0.2885	0.3409	0.3458	0.2273
$\omega_2$	1.2204	1.1447	1.1292	1.1502

**Table 9** Table of comparison between criterion of selection for enzyme dataset

Methods	EM-ML	EM-ECF-Kernel	EM-ECF-Empirical	Bayesian
AIC	<b>109.2265</b>	129.2364	129.3528	116.7802
BIC	<b>140.7378</b>	160.7478	160.8641	148.2916

We observe that almost all of the estimated parameters are similar, except for the index of stability  $\alpha$ . As previously mentioned, the EM-ML method is the most effective, as demonstrated in Table 9, with the smallest AIC and BIC values. Additionally, the associated figure for the four methods is presented in Figure 8 :



**Fig. 8** Mixture methods applied on Enzyme

As expected, the EM-ML method provides a better fit to the enzyme histogram compared to the other approaches. This is because the other methods, particularly EM-ECF, are sensitive to changes in the allocation vector at each iteration. Furthermore, the Bayesian approach has the advantage of utilizing prior knowledge of each parameter, which explains its efficiency on simulated data. However, it also requires certain conditions, such as the length of the observations, to ensure accurate results.

## 4.1 Conclusion and perspectives

In this paper, we consider the parameter estimation of uni-variate  $\alpha$ -stable distributions and their mixture. We introduced some new techniques, such as the Gaussian kernel estimator in the characteristic function for the case  $\alpha \geq 1$ , which has shown more efficient performance than the empirical characteristic function in the simulation study. We also perform another estimation procedure in the maximum likelihood framework based on the false-position algorithm method to find the root of the log-likelihood through the score functions established in [1]. In the case of estimation for the mixture of  $\alpha$ -stable distributions, although we limited our analysis to two components, the proposed methods can be generalized to several components. The EM algorithm was adapted to estimate the parameters of each sub-population until convergence by combining the ECF and ML method in the M-step. The Bayesian method, which is more flexible but requires many steps to perform the estimation was adapted also in the parameter estimation of given mixture model of  $\alpha$ -stable distributions. Finally, we consider two type of applications of our estimation methods on real data, namely the estimation of reproduction number of the Covid-19 in Mayotte and the Enzyme dataset. In terms of forthcoming

study, we are currently interested in practical developments based on the mixture properties of  $\alpha$ -stable distributions given in the Appendix section by an adaptation of the suitable Expectation- Maximization algorithm and Bayesian methods. Another practical perspective study is to implement an R package for parameter estimation of a mixture of  $\alpha$ -stable distributions.

## 5 Appendix

We aims here to establish some slightly different formulas related to mixture properties of  $\alpha$ -stable distributions, compared to [9] in order to provide some forthcoming practical work based on the estimations methods developed in this paper. Firstly, note that it is straightforward to show that if  $X \sim S(\alpha, \beta, \sigma, 0; 1)$  with  $\alpha < 2$ ; then, there exist two independent and identically distributed random variables  $Y_1$  and  $Y_2$  with the same distribution  $S(\alpha, 1, \sigma, 0; 1)$ , such that for  $\alpha \neq 1$ :

$$X \stackrel{d}{=} \left(\frac{1+\beta}{2}\right)^{\frac{1}{\alpha}} Y_1 - \left(\frac{1-\beta}{2}\right)^{\frac{1}{\alpha}} Y_2.$$

**Theorem 2** Let  $Y \sim S(\alpha, \beta, \sigma, \mu; 1)$ . For  $\alpha \in (0, 2)$  and  $\beta \in (0, 1)$ , we have:

$$Y = \left(\cos\left(\frac{\pi\alpha}{2}\right)\right)^\alpha \times \sigma \times \left(\frac{E_1}{W_1'} - \frac{E_2}{W_2'}\right) + \mu,$$

where  $E_1, E_2 \sim \mathcal{E}(1)$  are i.i.d., and  $W_1' \sim \text{Weibull}(0, \frac{1}{\lambda_1}, \alpha)$  and  $W_2' \sim \text{Weibull}(0, \frac{1}{\lambda_2}, \alpha)$  are also i.i.d., with  $\lambda_1 = \left(\frac{1+\beta}{2}\right)^{\frac{1}{\alpha}}$  and  $\lambda_2 = \left(\frac{1-\beta}{2}\right)^{\frac{1}{\alpha}}$ .

In view of the proof, let us introduce the following preliminaries.

**Lemma 1** Let  $0 < \alpha' < 2$  and  $0 < \alpha < \alpha'$ . Consider  $X_1$  and  $X_2$  as two independent copies with a (positive stable) distribution  $S(\frac{\alpha}{\alpha'}, 1, (\cos(\frac{\pi\alpha}{2\alpha'}))^{\frac{\alpha}{\alpha'}}, 0)$ , and  $V_1, V_2$  as two independent copies with a distribution  $S(\alpha', 1, (\cos(\frac{\pi\alpha'}{2}))^{\frac{1}{\alpha'}}, 0)$ . Then:

$$Z = \cos\left(\frac{\pi\alpha}{2}\right)^\alpha \times \left[\left(\frac{1+\beta}{2}\right)^{\frac{1}{\alpha}} X_1^{\frac{1}{\alpha'}} V_1 - \left(\frac{1-\beta}{2}\right)^{\frac{1}{\alpha}} X_2^{\frac{1}{\alpha'}} V_2\right]$$

is a random variable with a stable distribution  $S(\alpha, \beta, 1, 0)$ .

*Proof* Let  $S_1 = X_1^{\frac{1}{\alpha'}} V_1$ . We use the Laplace transform, as follows:

$$E(\exp(-uS_1)) = E\left(\exp(-uX_1^{\frac{1}{\alpha'}} V_1)\right)$$



$$\begin{aligned}
&= E\left(\exp(-u^{\alpha'} X_1)\right) \\
&= \exp\left(-\left(u^{\alpha'}\right)^{\frac{\alpha'}{\alpha'}} \times \frac{\left(\cos\left(\frac{\pi\alpha'}{2\alpha'}\right)\right)^{\frac{\alpha'}{\alpha'}}}{\cos\left(\frac{\pi\alpha'}{2\alpha'}\right)}\right) \\
&= \exp\left(-u^\alpha \times \frac{\left(\cos\left(\frac{\pi\alpha}{2}\right)\right)^{\frac{1}{\alpha}}}{\cos\left(\frac{\pi\alpha}{2}\right)}\right).
\end{aligned}$$

Therefore,  $S_1 \sim S(\alpha, 1, \cos(\frac{\pi\alpha}{2})^{\frac{1}{\alpha}}, 0; 1)$ . The same applies to  $S_2 = X_2^{\frac{1}{\alpha'}} V_2$ . Now, by using the property of multiple by a scalar in [13], we get:

$$\cos\left(\frac{\pi\alpha}{2}\right)^\alpha \times S_1 \sim S(\alpha, 1, 1, 0; 1).$$

Then, we obtain the desired result by using the algebraic property on sums of  $\alpha$ -stable distributions, see [13].  $\square$

**Lemma 2** *Let  $A$  and  $X$  be independent random variables, with  $A \sim S(\rho, 1, (\cos(\frac{\rho\pi}{2}))^{\frac{1}{\rho}}, 0; 1)$  and  $X \sim S(\alpha, \beta, 1, 0; 1)$ , such that  $0 < \rho < 1$ ,  $0 < \alpha \leq 2$ , and  $\alpha \neq 1 \neq \rho\alpha$ . Then  $A^{\frac{1}{\alpha}}X \sim S(\rho\alpha, \hat{\beta}, \hat{\sigma}, 0)$ . If  $\alpha < 1$ , and  $\beta = -1, 0$ , or  $1$ , then  $\hat{\beta} = \beta$ .*

*Proof* Since, we have  $A \sim S(\rho, 1, (\cos(\frac{\rho\pi}{2}))^{\frac{1}{\rho}}, 0)$ , it comes that  $E(\exp(-wA)) = \exp(-w^\rho)$ . Taking  $w > 0$ , we can deduce that  $A > 0$ . Now, let's compute the characteristic function of  $A^{\frac{1}{\alpha}}X$ . For  $\alpha \neq 1$ ,

$$\begin{aligned}
&E\left(\exp(itA^{\frac{1}{\alpha}}X)\right) \\
&= E\left(\exp(-tA^{\frac{1}{\alpha}}(1 - i\beta \tan(\frac{\alpha\pi}{2})\text{sign}(t)))\right) \\
&= E\left(\exp\left(-\left[t^\alpha(1 - i\beta \tan(\frac{\alpha\pi}{2})\text{sign}(t))\right]A\right)\right) \\
&= \exp\left(-\left[t^\alpha(1 - i\beta \tan(\frac{\alpha\pi}{2})\text{sign}(t))\right]^\rho\right) \\
&= \exp\left(-t^{\rho\alpha}\left(1 + \beta^2 \tan^2\left(\frac{\alpha\pi}{2}\right)\right)^{\frac{\rho}{2}}[\cos(\rho\theta) - i\sin(\rho\theta)\text{sign}(t)]\right) \\
&= \exp\left(-\cos(\rho\theta)(1 + \beta^2 \tan^2\left(\frac{\alpha\pi}{2}\right))^{\frac{\rho}{2}}t^{\rho\alpha}\left[1 - i\hat{\beta} \tan\left(\frac{\rho\alpha\pi}{2}\right)\text{sign}(t)\right]\right),
\end{aligned}$$

where  $\theta = \arctan(\beta \tan(\frac{\alpha\pi}{2}))$ ,  $\hat{\beta} = \tan(\rho\theta) / \tan(\frac{\rho\alpha\pi}{2})$ , et  $\hat{\sigma} = \cos(\rho\theta)^{\frac{1}{\rho\alpha}}(1 + \beta^2 \tan^2(\frac{\alpha\pi}{2}))^{\frac{1}{2\alpha}}$ .  $\square$

## Proof of Theorem 2

Let  $X_1 \sim S(\frac{\alpha}{\alpha'}, 1, (\cos(\frac{\pi\alpha}{2\alpha'}))^{\frac{\alpha'}{\alpha}}, 0; 1)$ , and  $V_1 \sim S(\alpha', 1, (\cos(\frac{\pi\alpha'}{2}))^{\frac{1}{\alpha'}}, 0; 1)$ , we define  $S_1 = X_1^{\frac{1}{\alpha'}} V_1$ . Thus,

$$S_1 \sim S(\alpha, 1, (\cos(\frac{\pi\alpha}{2}))^{\frac{1}{\alpha}}, 0; 1).$$

Moreover, we have:

$$\frac{E_1}{S_1} \sim \text{Weibull}(0, 1, \alpha).$$

We now return to the proved theorem, where  $S_2 = X_2^{\frac{1}{\alpha}} V_2$ , where  $X_2$  and  $V_2$  follow the same distributions as  $X_1$  and  $V_1$ , respectively, and are i.i.d. Thus:

$$X = \lambda_1 S_1 - \lambda_2 S_2.$$

With  $X \sim S(\alpha, \beta, (\cos(\frac{\pi\alpha}{2}))^{\frac{1}{\alpha}}, 0)$  and:

$$\begin{aligned} \lambda_1 S_1 &\stackrel{d}{=} \frac{E_1}{W_1}, \\ \lambda_2 S_2 &\stackrel{d}{=} \frac{E_2}{W_2}. \end{aligned}$$

By independence, we infer that:

$$X \stackrel{d}{=} \frac{E_1}{W_1} - \frac{E_2}{W_2}.$$

Hence the result follows.

## Supplementary information

Additional supporting information of the original contributions presented in the study and the R codes including datasets are available from the corresponding author on reasonable request.

## Declarations

The authors declare that they have no competing interests. The co-authors have seen and agree with the contents of the manuscript. We certify that the submission is an original work and is not under review at any other publication.

## References

- [1] Nolan, J.P.: Univariate Stable Distributions. Springer, ??? (2020)
- [2] Kuruoglu, E.E.: Density parameter estimation of skewed/spl alpha/-stable distributions. IEEE Transactions on signal processing **49**(10), 2192–2201 (2001)

- [3] McCulloch, J.H.: Simple consistent estimators of stable distribution parameters. *Communications in statistics-simulation and computation* **15**(4), 1109–1136 (1986)
- [4] Paczek, K., Jelito, D., Pitera, M., Wyłomańska, A.: Estimation of stability index for symmetric  $\{\alpha\}$ -stable distribution using quantile conditional variance ratios. *arXiv preprint arXiv:2212.13502* (2022)
- [5] Sinaga, K.P., Yang, M.-S.: Unsupervised k-means clustering algorithm. *IEEE access* **8**, 80716–80727 (2020)
- [6] Castillo-Barnes, D., Martínez-Murcia, F.J., Ramírez, J., Górriz, J., Salas-Gonzalez, D.: Expectation–maximization algorithm for finite mixture of  $\alpha$ -stable distributions. *Neurocomputing* **413**, 210–216 (2020)
- [7] Teimouri, M., Rezakhah, S., Mohammadpour, A.: Em algorithm for symmetric stable mixture model. *Communications in Statistics-Simulation and Computation* **47**(2), 582–604 (2018)
- [8] Teimouri, M.: Finite mixture of skewed sub-gaussian stable distributions. *arXiv preprint arXiv:2205.14067* (2022)
- [9] Teimouri, M.: Statistical inference for stable distribution using em algorithm. *arXiv preprint arXiv:1811.04565* (2018)
- [10] Salas-Gonzalez, D., Kuruoglu, E.E., Ruiz, D.P.: Finite mixture of  $\alpha$ -stable distributions. *Digital Signal Processing* **19**(2), 250–264 (2009)
- [11] Manou-Abi, S.M., Slaoui, Y., Balicchi, J.: Estimation of some epidemiological parameters with the covid-19 data of mayotte. *Frontiers in Applied Mathematics and Statistics* **8**, 870080 (2022)
- [12] Bechtel, Y.C., Bonaiti-Pellie, C., Poisson, N., Magnette, J., Bechtel, P.R.: A population and family study n-acetyltransferase using caffeine urinary metabolites. *Clinical Pharmacology & Therapeutics* **54**(2), 134–141 (1993)
- [13] Samorodnitsky, G., Taqqu, M.S., Linde, R.: Stable non-gaussian random processes: stochastic models with infinite variance. *Bulletin of the London Mathematical Society* **28**(134), 554–555 (1996)
- [14] Sheather, S.J., Jones, M.C.: A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B (Methodological)* **53**(3), 683–690 (1991)
- [15] Slaoui, Y.: Bandwidth selection for recursive kernel density estimators defined by stochastic approximation method. *Journal of Probability and*

Statistics **2014** (2014)

- [16] Slaoui, Y.: The stochastic approximation method for the estimation of a distribution function. *Mathematical Methods of Statistics* **23**(4), 306–325 (2014)
- [17] Kogon, S.M., Williams, D.B.: Characteristic function based estimation of stable distribution parameters. *A practical guide to heavy tails: statistical techniques and applications*, 311–338 (1998)
- [18] Thakur, G., Saini, J.K.: Comparative study of iterative methods for solving non-linear equations. *journal of university of shanghai for science and technology*. *Journal of University of Shanghai for Science and Technology* (2021)
- [19] Diekmann, O., Heesterbeek, J.A.P., Metz, J.A.: On the definition and the computation of the basic reproduction ratio  $r_0$  in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology* **28**, 365–382 (1990)
- [20] Wallinga, J., Lipsitch, M.: How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences* **274**(1609), 599–604 (2007)
- [21] Obadia, T., Haneef, R., Boëlle, P.-Y.: The r0 package: a toolbox to estimate reproduction numbers for epidemic outbreaks. *BMC medical informatics and decision making* **12**(1), 1–9 (2012)
- [22] Boelle, P.-Y., Obadia, T., Obadia, M.T.: Package ‘r0’ (2015)
- [23] Demongeot, J., Oshinubi, K., Rachdi, M., Seligmann, H., Thuderoz, F., Waku, J.: Estimation of daily reproduction numbers during the covid-19 outbreak. *Computation* **9**(10), 109 (2021)
- [24] Waku, J., Oshinubi, K., Demongeot, J.: Maximal reproduction number estimation and identification of transmission rate from the first inflection point of new infectious cases waves: Covid-19 outbreak example. *Mathematics and Computers in Simulation* **198**, 47–64 (2022)