



HAL
open science

On parameter estimation for Amoroso family of distributions

Catherine Combes, Hon Keung Tony Ng

► **To cite this version:**

Catherine Combes, Hon Keung Tony Ng. On parameter estimation for Amoroso family of distributions. *Mathematics and Computers in Simulation*, 2022, 191, pp.309-327. 10.1016/j.matcom.2021.07.004 . hal-03332015

HAL Id: hal-03332015

<https://hal.science/hal-03332015>

Submitted on 16 Oct 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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

On Parameter Estimation for Amoroso Family of Distributions

Catherine Combes^{a,*}, Hon Keung Tony Ng^b

^a*Univ Lyon, UJM-Saint-Étienne, CNRS, Institut d'Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, Saint-Etienne, France*

^b*Department of Statistical Science, Southern Methodist University, Dallas, TX 75275-0332, USA*

Abstract

The four-parameter generalized gamma (GF) distribution, also known as the Amoroso family of distributions, is a flexible and versatile statistical distribution that encapsulates many well-known lifetime distributions, including the exponential, Weibull, lognormal, and gamma distributions as special instances. The four-parameter GF distribution is shown to be appropriate for fitting skewed and heavy-tailed data sets. However, even though the GF distribution is very useful and flexible, it remains less studied than its counterparts, probably due to the difficulty in estimating the parameters of the distribution. In this paper, we explore several novel iterative parameter estimation approaches for the four-parameter GF distribution, which includes the maximum likelihood estimation and minimum distance estimation approaches.

Standard error and confidence interval of a function of the parameter estimates based on bootstrap method are also discussed.

An R package is developed based on the proposed estimation methods. Numerical examples and Monte Carlo simulations are used to illustrate the usefulness of the proposed approaches for fitting the four-parameter GF distribution.

Keywords: Maximum likelihood estimation, minimum distance estimation, optimization, generalized gamma distribution.

*Corresponding author.

Email address: `catherine.combes@univ-st-etienne.fr` (Catherine Combes)

1. Introduction

The generalized gamma (GF) distribution is a statistical distribution originally introduced by Amoroso [1] and D'Addario [2] as a special case of a mathematical function for analyzing and modeling economic income data. The GF distribution is also known as the Amoroso family of distributions. The functional form commonly used in practice was suggested by Stacy [3], which corresponds to a three-parameter exponentiated gamma distribution. The four-parameter GF distribution is a flexible distribution that has been a great interest in many practical applications. For instance, since the four-parameter GF distribution provides different shapes of hazard function, it can be used to model duration or time-to-event data and to describe a wide variety of life-cycle phenomena. Moreover, in finance and insurance, the GF distribution is commonly used as a model for financial losses, stock exchange prices, or insurance claim sizes, mainly because it is a fairly flexible positive-skewed distribution. The four-parameter GF distribution includes many well-known distributions as special cases. Furthermore, since the GF distribution can be considered as a sum of independent exponential random variables, it is also a natural model for individual waiting times. However, as pointed out by Song [4], despite this wide range of potential applications, the GF distribution is less used in practice, while the major reason may due to the difficulty of estimating the model parameters. The difficulty lies in the issue that different sets of model parameters can provide similar fits of the GF distribution. Therefore, in this paper, we aim to address the problem of estimating the parameters of the four-parameter GF distribution by providing some feasible estimation methods.

For the parameter estimation of the three-parameter GF distribution proposed in [3], which is a member of the Amoroso family of distributions, it was first discussed by Parr and Webster [5] and Stacy and Mihram [6]. Different iterative schemes to obtain the maximum likelihood estimates of the model parameters have been studied, see, for example, [7, 8, 9, 10, 11, 12, 13]. However, those iterative algorithms, such as the Newton-Raphson method, to obtain the

maximum likelihood estimates may not work well unless an appropriate and reliable set of initial values are provided. Nevertheless, as pointed out by Prentice [9], the asymptotic properties of the maximum likelihood estimates may not hold even when sample size is large. Hence, Hirose [14, 15] proposed a reparameterization of the three-parameter GF distribution function and applied the predictor-corrector method with the continuation method [16]. For a comprehensive literature review on the maximum likelihood estimation of the three-parameter GF distribution and the iterative approaches to obtain the solution of the likelihood score equations, one may refer to [12].

For some interesting applications of the three-parameter GF distribution in the context of time-to-event or lifetime data analysis, the readers can refer to some recent work by Balakrishnan and Pal [17], Ling [18], Ramos et al. [19] and Pal et al. [20].

In addition to the maximum likelihood estimation method, moment-based estimation approaches have been studied in the literature; see, for example, [6, 21, 22, 23, 24, 25, 26, 27, 28]. Though, those moment-based methods are difficult to use due to the fact that different sets of parameters conduce to similar probability density functions (PDFs). Furthermore, those moment-based methods are essentially proposed for non-negative data. In the last decades, Bayesian estimation methods for the three-parameter GF distribution have attracted a great deal of interest. The main issue of using Bayesian methods resides in the choice of appropriate priors for the parameters, which is an important prerequisite to the success of the Bayesian approach. Different authors, including [29] [30] [31] [32] [33] [34], discussed the strengths and weaknesses inaccessible to theoretical analysis such as different prior distributions can have a great effect on the subsequent estimates.

In summary, even for the special case of the four-parameter GF distribution, the moment-based method, the maximum likelihood estimation method, and the Bayesian method have difficulties and weaknesses. Therefore, for a more general four-parameter GF distribution, special attention is needed for the estimation of the model parameters.

In this article, we study the maximum likelihood estimation method and propose the minimum distance estimation (MDE) method [35]. The proposed MDE approach is based on the minimization of various distance metrics between the theoretical distribution and the observed data. The major contribution of this paper is threefold:

- Novel estimation procedures are developed to estimate the model parameters for the four-parameter GF distribution in order to fit data with positive and/or negative values with the four-parameter GF distribution;
- 70 • R package that implements the proposed estimation methods and functions related to the four-parameter GF distribution is developed;
- The proposed estimation methods are shown to be feasible and effective.

The remainder of this article is organized as follows. Section 2 provides the preliminary knowledge about the Amoroso family of distributions, including the functional form of the distribution, special cases, and some basic properties. In Section 3, different parameter estimation methods for the model parameters in a four-parameter GF distribution are discussed. Specifically, a method to obtain initial estimates, the maximum likelihood estimation method, and the minimum distance estimation method are considered. We also discuss a bias-reduction estimate for the parameter μ and use the bootstrap method to obtain the standard error of each estimate and the confidence interval of each parameter in Section 3. In Section 4, the features of the R package created for the Amoroso family of distributions are discussed. In Section 5, two numerical examples based on real data are used to illustrate the estimation methods considered in this paper. In Section 6, Monte Carlo simulation studies are used to evaluate the performance of the proposed estimation methods. Finally, some concluding remarks are provided in Section 7.

2. Preliminary Knowledge of Amoroso Family of Distributions

In this section, we first introduce the mathematical notations used and the
 90 four-parameter Γ distribution considered in this paper. Some properties of
 the four-parameter Γ distribution will be discussed.

2.1. Notations

The notations used in this paper are the following:

- Upper case letters denote cumulative distribution function (CDF) (e.g.,
 95 G) and lower case letters denote the corresponding PDF (e.g., g).
- The true distribution is referred to as G , while the empirical distribution
 corresponding to $n \in \mathbb{N}$ independent and identically distributed (i.i.d.)
 observations $\mathbf{x} = (x_1, \dots, x_n)$, is referred to as G_n .
- $\boldsymbol{\theta}$ is a m -dimensional vector of unknown parameters.
- 100 • $\hat{\boldsymbol{\theta}}$ is a m -dimensional vector of estimated parameters.
- Θ is an open subset of a separable Hubert space.
- \mathbf{Y} is a vector of random variables following the four-parameter Γ with
 $\mathbf{Y} = \{Y_1, Y_2, \dots, Y_n\}$.
- \mathbf{y} is the vector of observations with $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$.

105 2.2. Amoroso family of distributions and four-parameter Generalized Gamma distribution

A random variable Y follows a four-parameter Γ distribution with PDF

$$g(y; a, \ell, c, \mu) = \frac{1}{\Gamma(\ell)} \left| \frac{c}{a} \right| \left(\frac{y - \mu}{a} \right)^{\ell c - 1} \exp \left[- \left(\frac{y - \mu}{a} \right)^c \right], \quad (1)$$

where $a \in \mathbb{R} \setminus \{0\}$ is the scale parameter, $\mu \in (-\infty, \infty)$ is the location parameter,
 $\ell \in \mathbb{R}^+$ and $c \in \mathbb{R} \setminus \{0\}$ are the shape parameters, and the support of the

110 distribution is

$$\begin{cases} y \geq \mu & \text{if } a > 0, \\ y \leq \mu & \text{if } a < 0. \end{cases}$$

We denote this distribution as $\text{GF4}(a, \ell, c, \mu)$ distribution. Note that the PDF in Eq. (1) is used in Crocks [36], while the probability distribution is defined for $a > 0$ and the support of the distribution is $\mu < y < \infty$ in the original paper by Amoroso [1].

115 The CDF of random variable Y that follows a $\text{GF4}(a, \ell, c, \mu)$ distribution can be expressed as

$$G(y; a, \ell, c, \mu) = \frac{\gamma\left(\ell, \left(\frac{y-\mu}{a}\right)^c\right)}{\Gamma(\ell)}, \quad (2)$$

where $\gamma(s, x) = \int_0^x t^{s-1} \exp(-t) dt$ is the lower incomplete gamma function. In this paper, we are interested in estimating the parameter vector $\boldsymbol{\theta} = (a, \ell, c, \mu)$ based on the observed sample of size n , $\mathbf{y} = (y_1, y_2, \dots, y_n)$, from the $\text{GF4}(a, \ell, c, \mu)$ 120 distribution. The properties of the $\text{GF4}(a, \ell, c, \mu)$ distribution that are used in the development of the parameter estimation methods are described here.

Moments. Suppose Y follows the $\text{GF4}(a, \ell, c, \mu)$ distribution, the r^{th} moment of Y ($r = 1, 2, \dots$) is

$$E[Y^r] = \begin{cases} a^r \frac{\Gamma\left(\frac{c\ell+r}{c}\right)}{\Gamma(\ell)}, & \text{if } r/c > -\ell \\ \infty, & \text{otherwise.} \end{cases}$$

Then, the mean and variance of the random variable Y are

$$\begin{aligned} E[Y] &= \frac{a\Gamma\left(\ell + \frac{1}{c}\right)}{\Gamma(\ell)} + \mu \\ \text{and } Var[Y] &= a^2 \left\{ \frac{\Gamma\left(\ell + \frac{2}{c}\right)\Gamma(\ell) - [\Gamma\left(\ell + \frac{1}{c}\right)]^2}{\Gamma(\ell)^2} \right\}, \end{aligned} \quad (3)$$

respectively.

Distribution of transformed random variables. Suppose Y follows the $\text{GF4}(a, \ell, c, \mu =$ 125 $0)$ distribution, then we have

- (1) $X = Y^s \sim \text{GF}(a^s, \ell, c/s, 0)$;
- (2) $Z = \omega Y \sim \text{GF}(\omega a, \ell, c, 0)$;
- (3) $W = Y^c \sim \text{GF}(a^c, \ell, 1, 0)$ (i.e., a two-parameter gamma distribution with scale parameter a^c and shape parameter ℓ);
- 130 (4) If V follows a two-parameter gamma distribution with scale parameter 1 and shape parameter ℓ , then $Y = aV^{\frac{1}{c}} + \mu \sim \text{GF}(a, \ell, c, \mu)$.

Property (4) provides a simple way to generate a random sample of the four-parameter GF distribution in using the quantile function of the two-parameter gamma distribution.

135 *Special cases.* As mentioned in the Section 1, the four-parameter GF distribution encapsulates a large number of well-known and widely used distributions. These connections are presented in Figure 1. Crocks [36] also presented a review of important properties of the Amoroso family of distributions and pointed out that many commonly used probability distributions occur as special cases or
 140 limits (such as the power law, the log-gamma, the log-normal, and the normal distributions).

The four-parameter GF model is important generalizations of other models and the inferential procedures of these special cases have been studied in the literature. For instance, for statistical estimation and inferential procedures of
 145 the generalized Fisher-Tippett distribution (also known as the generalized extreme value distribution), the generalized Weibull distribution, and the Pearson Type III distribution, one can refer to [37], [38], and [39], respectively, and the references therein.

3. Parameter Estimation Methods

150 3.1. A method to obtain initial estimates

Based on the observed sample $\mathbf{y} = (y_1, y_2, \dots, y_n)$, we denote the i -th order statistic of the observed sample as $y_{i:n}$. Based on the observed data, the parameter space can be written as two disjoint sets $\{a > 0, \mu \in (y_{1:n}, \infty), \ell \in \mathbb{R}^+,$

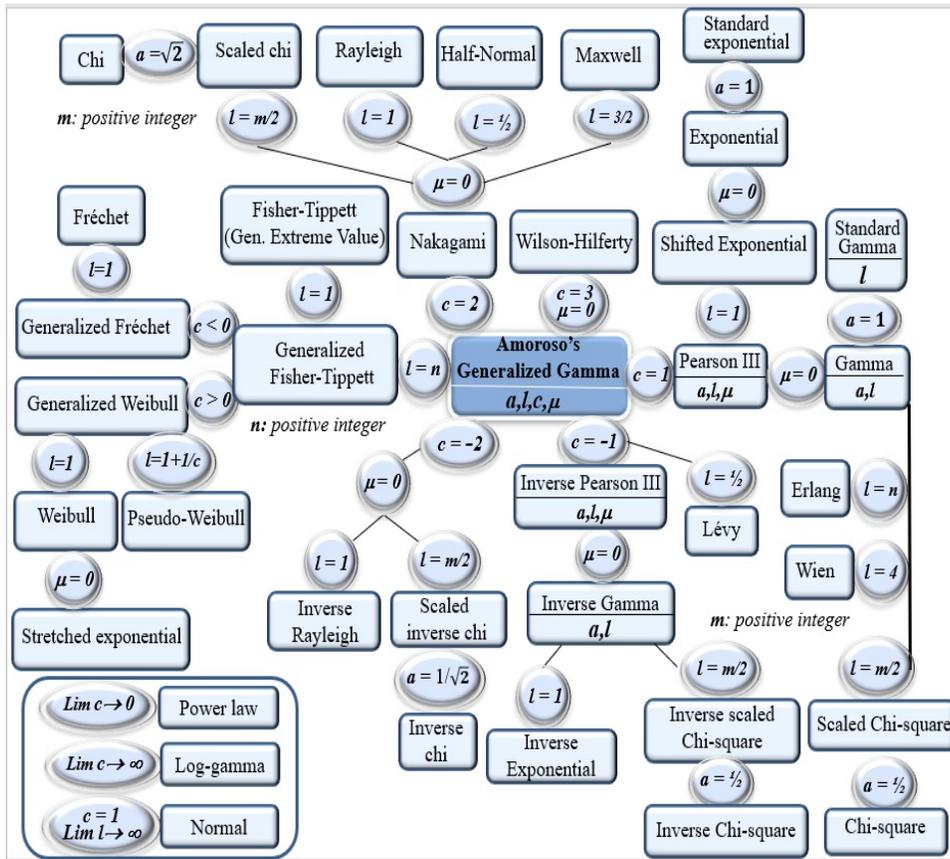


Figure 1: Amoroso family of distributions and their connections

$c \in \mathbb{R} \setminus \{0\}$ and $\{a < 0, \mu \in (-\infty, y_{n:n}), \ell \in \mathbb{R}^+, c \in \mathbb{R} \setminus \{0\}\}$, where $y_{i:n}$ is
 155 the i -th order statistic of the observed sample $\mathbf{y} = (y_1, y_2, \dots, y_n)$. Therefore,
 the estimation in each of the two disjoint parameter spaces can be considered
 separately and then the two sets of estimates are compared.

Usually, due to the form of the four-parameter $\text{G}\Gamma$ distribution, close-formed
 solution of the parameter estimates are not available and iterative procedures
 160 are needed to obtain the estimates. In the past years, iterative procedures
 coupled with Monte Carlo simulation have been largely studied for estimating
 the four-parameter $\text{G}\Gamma$ distribution [13, 27, 40]. Since those iterative procedures
 require a reliable starting value of the estimates, we propose here a method to
 obtain the initial estimates.

165 For $a > 0$, since the support of the distribution is $y > \mu$, we set the initial
 estimate of μ as $\hat{\mu}_I = py_{1:n}$, where $p \leq 1$ when $y_{1:n} > 0$ and $\hat{\mu}_I = qy_{1:n}$, where
 $q \geq 1$ when $y_{1:n} < 0$, and transformed the observed data by $x_{i:n} = y_{i:n} - \hat{\mu}_I$,
 $i = 1, 2, \dots, n$. Note that when $p = 1$, $x_{1:n} = 0$ which needs to be discarded in
 the process of obtaining the initial estimates of parameters a , l and c to avoid
 170 computational error (e.g., taking logarithm of $x_{1:n}$ results in $-\infty$ if $x_{1:n} = 0$).
 Similarly, for $a < 0$, since the support of the distribution is $y < \mu$, we set the
 initial estimate of μ as $\hat{\mu}_I = qy_{n:n}$, where $q \geq 1$ when $y_{n:n} > 0$ and $\hat{\mu}_I = py_{n:n}$,
 where $p \leq 1$ when $y_{n:n} < 0$, and transformed the observed data by $x_{i:n} =$
 $y_{i:n} - \hat{\mu}_I$, $i = 1, 2, \dots, n$. When $q = 1$, $x_{n:n} = 0$ which needs to be discarded in
 175 the process of obtaining the initial estimates of parameters a , l and c to avoid
 computational error. Here, we suggest to use the value $p = 0.99$ and $q = 1.01$.

Using the property that $W = (Y - \mu)^c$ follows a two-parameter gamma
 distribution with scale parameter a^c and shape parameter ℓ if Y follows the
 $\text{G}\Gamma(a, \ell, c, \mu)$ distribution, and the method of moment estimates of the scale and
 180 shape parameters of two-parameter gamma distribution, the proposed procedure
 to obtain the initial parameter estimates are described as follows:

Step A1. Obtain an initial estimate of μ as

$$\hat{\mu}_I = \begin{cases} py_{1:n}, & \text{when } y_{1:n} > 0, \text{ for } a > 0; \\ qy_{1:n}, & \text{when } y_{1:n} < 0, \text{ for } a > 0; \\ qy_{n:n}, & \text{when } y_{n:n} > 0, \text{ for } a < 0; \\ py_{n:n}, & \text{when } y_{n:n} < 0, \text{ for } a < 0, \end{cases}$$

with $p \leq 1$ and $q \geq 1$.

Step A2. Compute $x_{i:n} = y_{i:n} - \mu_I$ ($i = 1, 2, \dots, n$) and discard the values $x_{i:n} = 0$;

Step A3. Consider a sequence of values for parameter c , $c_j, j = 1, 2, \dots, J$;

Step A4. Fixed the value of $c = c_j$, compute $w_{i:n} = x_{i:n}^{c_j}$ ($i = 1, 2, \dots, n$);

Step A5. Obtain the estimates of a and ℓ as

$$a_j = \left(\frac{s_w^2}{\bar{w}} \right)^{1/c_j} \quad \text{and} \quad \ell_j = \left(\frac{\bar{w}^2}{s_w^2} \right),$$

respectively, where

$$\bar{w} = \frac{1}{n} \sum_{i=1}^n w_i \quad \text{and} \quad s_w^2 = \frac{1}{n-1} \sum_{i=1}^n (w_i - \bar{w})^2;$$

Step A6. Compute the likelihood of the observed data based on parameter vector

$$\boldsymbol{\theta} = (a_j, \ell_j, c_j, \mu_I):$$

$$L_j = L(\mathbf{y}; \boldsymbol{\theta} = (a_j, \ell_j, c_j, \mu_I)) = \prod_{i=1}^n g(y_i; a_j, \ell_j, c_j, \mu_I);$$

Step A7. Repeat Steps A3–A6 for different values of c_j to obtain $L_j, j = 1, 2, \dots, J$;

Step A8. Compare the values of the likelihood $L_j, j = 1, 2, \dots, J$ to obtain j^* in which

$$L_{j^*} = \max_j L_j;$$

Step A9. Set the initial estimate of $\boldsymbol{\theta}$ as $\hat{\boldsymbol{\theta}}_I = (a_{j^*}, \ell_{j^*}, c_{j^*}, \mu_I)$.

The above algorithm can be used to obtain initial estimates of $\boldsymbol{\theta}$ for $a > 0$ and $a < 0$.

3.2. Maximum likelihood estimation

The maximum likelihood estimate (MLE) of the parameter vector $\boldsymbol{\theta} = (a, \ell, c, \mu)$, denoted as $\hat{\boldsymbol{\theta}}_{ML} = (\hat{a}_{ML}, \hat{\ell}_{ML}, \hat{c}_{ML}, \hat{\mu}_{ML})$ can be obtained by maximizing the likelihood function

$$L(\mathbf{y}; \boldsymbol{\theta} = (a, \ell, c, \mu)) = \prod_{i=1}^n g(y_i; a, \ell, c, \mu).$$

We can first obtain the MLE of $\boldsymbol{\theta}$ in the parameter space $\{a > 0, \mu \in (y_{1:n}, \infty), \ell \in \mathbb{R}^+, c \in \mathbb{R} \setminus \{0\}\}$, denoted as $\hat{\boldsymbol{\theta}}_{ML+}$, and then obtain the MLE of $\boldsymbol{\theta}$ in the parameter space $\{a < 0, \mu \in (-\infty, y_{n:n}), \ell \in \mathbb{R}^+, c \in \mathbb{R} \setminus \{0\}\}$, denoted as $\hat{\boldsymbol{\theta}}_{ML-}$. The MLE of $\boldsymbol{\theta}$ can be obtained as

$$\hat{\boldsymbol{\theta}}_{ML} = \begin{cases} \hat{\boldsymbol{\theta}}_{ML+}, & \text{if } L(\mathbf{y}; \hat{\boldsymbol{\theta}}_{ML+}) > L(\mathbf{y}; \hat{\boldsymbol{\theta}}_{ML-}), \\ \hat{\boldsymbol{\theta}}_{ML-}, & \text{if } L(\mathbf{y}; \hat{\boldsymbol{\theta}}_{ML+}) < L(\mathbf{y}; \hat{\boldsymbol{\theta}}_{ML-}). \end{cases} \quad (4)$$

3.3. Minimum distance estimation based on nonparametric density estimation

The MDE method [35, 41] based on nonparametric density estimation refers to a very general technique that formalizes the parameters inference problem as a problem of minimizing a distance, over the set of parameter $\boldsymbol{\theta} \in \Theta$, between the model PDF $g(x; \boldsymbol{\theta})$ and some empirical density estimate, \hat{f}_n obtained from a sample of size n . Formally, the main objective of the MDE is to minimize the distance between $g(x; \boldsymbol{\theta})$ and $\hat{f}_n(x)$ with respect to $\boldsymbol{\theta} = (a, \ell, c, \mu)$, i.e.,

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmin}} d(g(x; \boldsymbol{\theta}) || \hat{f}_n(x)),$$

where $d(g||f)$ denotes any proper distance or divergence function that evaluates the closeness of two density functions g and f . Wolfowitz [35] showed that MDE is more robust to departures from underlying assumptions than maximum likelihood estimation. Millar [42] explored the asymptotic behavior of the MDE within a general framework. The MDE approach is different from the moment-based or maximum likelihood approach and is widely studied in the literature [41]. The classic examples of this method are the least-square and the minimum chi-square estimators.

Based on the observed sample $\mathbf{y} = (y_1, y_2, \dots, y_n)$, we consider the empirical density estimate based on the Gaussian kernel density estimate \hat{f}_n . The Gaussian kernel density estimate, \hat{f}_n , of a univariate density f is (see, for example, [43])

$$\hat{f}_n(x) = \frac{1}{n} \sum_{i=1}^n \frac{K(x - y_i; h)}{h},$$

where $K(x; h) \propto \exp[-x^2/(2h^2)]$ is the Gaussian kernel function, h is the smoothing parameter (bandwidth) chosen to be 0.9 times the minimum of the standard deviation and the interquartile range of the sample \mathbf{y} divided by $1.34 \times n^{-1/5}$ [44]. Specifically, we discretize the values of x into N points (x_1, x_2, \dots, x_N) and minimizing the function

$$\sum_{k=1}^N d(g(x_k; a, \ell, c, \mu) || \hat{f}_n(x_k)).$$

The following distance or divergence measures are considered.

Kullback-Leibler divergence. The Kullback-Leibler divergence between two densities f and g is defined as

$$d_{KL}(g(x)||f(x)) = \int g(x) \ln \left[\frac{g(x)}{f(x)} \right] dx.$$

The estimator of θ , denoted as $\hat{\theta}_{KL}$, based on minimizing Kullback-Leibler divergence can be obtained as

$$\hat{\theta}_{KL} = \underset{\theta \in \Theta}{\operatorname{argmin}} \left\{ \sum_{k=1}^N \left| g(x_k; \theta) \ln \left[\frac{g(x_k; \theta)}{\hat{f}_n(x_k)} \right] \right| \right\}. \quad (5)$$

Jensen-Shanon Bregman divergence. Since the Kullback-Leibler divergence defined above is not symmetric (i.e., $D_{KL}(g(x)||f(x)) \neq (f(x)||g(x))$), we consider the average of the two divergence measures

$$\frac{1}{2} \int \left\{ g(x) \ln \left[\frac{g(x)}{f(x)} \right] + f(x) \ln \left[\frac{f(x)}{g(x)} \right] \right\} dx,$$

which is equivalent to consider the Jensen-Shanon Bregman divergence defined as

$$d_{JSB}(g(x)||f(x)) = \int [g(x) - f(x)] \ln \left[\frac{g(x)}{f(x)} \right] dx.$$

220 The estimator of $\boldsymbol{\theta}$, denoted as $\hat{\boldsymbol{\theta}}_{JSB}$, based on minimizing the Jensen-Shanon Bregman divergence can be obtained as

$$\hat{\boldsymbol{\theta}}_{JSB} = \operatorname{argmin}_{\boldsymbol{\theta} \in \Theta} \left\{ \sum_{k=1}^N \left| g(x_k; \boldsymbol{\theta}) \ln \left[\frac{g(x_k; \boldsymbol{\theta})}{\hat{f}_n(x_k)} \right] \right| + \left| \hat{f}_n(x_k) \ln \left[\frac{\hat{f}_n(x_k)}{g(x_k; \boldsymbol{\theta})} \right] \right| \right\}. \quad (6)$$

Hellinger distance. The Hellinger distance between two densities f and g is defined as

$$d_{HD}(g(x)||f(x)) = \int \left[\sqrt{g(x)} - \sqrt{f(x)} \right]^2.$$

The estimator of $\boldsymbol{\theta}$, denoted as $\hat{\boldsymbol{\theta}}_{HD}$, based on minimizing Hellinger distance can be obtained as

$$\hat{\boldsymbol{\theta}}_{HD} = \operatorname{argmin}_{\boldsymbol{\theta} \in \Theta} \left\{ \sum_{k=1}^N \left[\sqrt{g(x_k; \boldsymbol{\theta})} - \sqrt{\hat{f}_n(x_k)} \right]^2 \right\}. \quad (7)$$

Wassertein distance. The Wassertein distance (also known as Kantorovich distance) is an interesting metric on probability distributions and popular in many areas of statistics and machine learning. An estimate of $\boldsymbol{\theta}$ based on Wassertein distance can be obtained by minimizing the function

$$d_{WD}(g(x)||f(x)) = \left[\int |g(x) - f(x)|^\delta dx \right]^{\frac{1}{\delta}}.$$

225 Based on the N discretized points, the estimator of $\boldsymbol{\theta}$, denoted as $\hat{\boldsymbol{\theta}}_{WD}$, based on minimizing Wassertein distance can be obtained as

$$\hat{\boldsymbol{\theta}}_{WD} = \operatorname{argmin}_{\boldsymbol{\theta} \in \Theta} \left\{ \left[\frac{1}{N} \sum_{k=1}^N |g(x_k; \boldsymbol{\theta}) - \hat{f}_n(x_k)|^\delta \right]^{\frac{1}{\delta}} \right\}. \quad (8)$$

Here, we consider $\delta = 4$. Minimum Wasserstein distance estimators are particular instances of minimum distance estimators [41] and appear to be practical and robust alternatives to likelihood-based estimation in the setting of generative models. The study asymptotic properties of such minimum Wasserstein distance estimators, complementing results derived by Bassetti, Bodini
230 and Regazzini [45], are presented in [46].

Squared Euclidean distance. The squared Euclidean distance between two densities f and g is defined as

$$d_{SE}(g(x)||f(x)) = \int [g(x) - f(x)]^2.$$

The estimator of θ , denoted as $\hat{\theta}_{SE}$, based on minimizing squared Euclidean distance can be obtained as

$$\hat{\theta}_{SE} = \operatorname{argmin}_{\theta \in \Theta} \left\{ \left[\frac{1}{N} \sum_{k=1}^N |g(x_k; \theta) - \hat{f}_n(x_k)|^2 \right] \right\}. \quad (9)$$

Note that the estimator obtained based minimizing the squared Euclidean distance can be viewed as a special case of the minimum Wasserstein distance estimator when $\delta = 2$.

3.4. Minimum distance estimation based on empirical CDF

Based on the observed sample $\mathbf{y} = (y_1, y_2, \dots, y_n)$, we consider the empirical CDF \hat{F}_n as

$$\hat{F}_n(y_{i:n}) = \frac{i - 0.5}{n},$$

for $i = 1, 2, \dots, n$. Note that other kinds of nonparametric estimate of CDF, such as $\hat{F}_n(x_{i:n}) = i/(n + 1)$ can be considered here. Then, the minimum distance estimate based on empirical CDF can be obtained by minimizing the function

$$\sum_{i=1}^n \mathcal{D}(G(y_{i:n}; a, \ell, c, \mu) || \hat{F}_n(y_{i:n})).$$

The following distance or divergence measures are considered.

Hellinger distance. The Hellinger distance between two CDFs F and G is defined as

$$\mathcal{D}_{HD}(G(y)||F(y)) = \int [\sqrt{G(y)} - \sqrt{F(y)}]^2.$$

The estimator of θ , denoted as $\hat{\theta}_{HDC}$, based on minimizing Hellinger distance can be obtained as

$$\hat{\theta}_{HDC} = \operatorname{argmin}_{\theta \in \Theta} \left\{ \sum_{i=1}^n \left[\sqrt{G(y_{i:n}; \theta)} - \sqrt{\hat{F}_n(y_{i:n})} \right]^2 \right\}. \quad (10)$$

Wassertein distance. The Wassertein distance between two CDFs F and G is defined as

$$\mathcal{D}_{WD}(G(y), F(y)) = \left[\int |G(y) - F(y)|^\delta dy \right]^{1/\delta}.$$

The estimator of $\boldsymbol{\theta}$, denoted as $\hat{\boldsymbol{\theta}}_{WDC}$, based on minimizing Wassertein distance can be obtained as

$$\hat{\boldsymbol{\theta}}_{WDC} = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmin}} \left\{ \left[\frac{1}{n} \sum_{i=1}^n |G(y_{i:n}; \boldsymbol{\theta}) - \hat{F}_n(y_{i:n})|^\delta \right]^{\frac{1}{\delta}} \right\}. \quad (11)$$

245 Here, we consider $\delta = 4$.

Squared Euclidean distance. The squared Euclidean distance between two CDFs F and G is defined as

$$\mathcal{D}_{SE}(G(y), F(y)) = \int [G(y) - F(y)]^2.$$

The estimator of $\boldsymbol{\theta}$, denoted as $\hat{\boldsymbol{\theta}}_{SEC}$, based on minimizing squared Euclidean distance can be obtained as

$$\hat{\boldsymbol{\theta}}_{SEC} = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmin}} \left\{ \frac{1}{n} \sum_{i=1}^n [G(y_{i:n}; \boldsymbol{\theta}) - \hat{F}_n(y_{i:n})]^2 \right\}. \quad (12)$$

For the MDE based on nonparametric density estimation and the MDE based on empirical CDF, similar to the MLE, we can first obtain the MDE of $\boldsymbol{\theta}$ in the parameter space $\{a > 0, \mu \in (y_{1:n}, \infty), \ell \in \mathbb{R}^+, c \in \mathbb{R} \setminus \{0\}\}$, denoted as $\hat{\boldsymbol{\theta}}_{MD+}$, and then obtain the MDE of $\boldsymbol{\theta}$ in the parameter space $\{a < 0, \mu \in (-\infty, y_{n:n}), \ell \in \mathbb{R}^+, c \in \mathbb{R} \setminus \{0\}\}$, denoted as $\hat{\boldsymbol{\theta}}_{MD-}$. The MDE of $\boldsymbol{\theta}$ can be obtained as

$$\hat{\boldsymbol{\theta}}_{MD} = \begin{cases} \hat{\boldsymbol{\theta}}_{MD+}, & \text{if } d(g(x; \hat{\boldsymbol{\theta}}_{MD+}) || f_n(x)) < d(g(x; \hat{\boldsymbol{\theta}}_{MD-}) || f_n(x)) \\ \hat{\boldsymbol{\theta}}_{MD-}, & \text{if } d(g(x; \hat{\boldsymbol{\theta}}_{MD+}) || f_n(x)) > d(g(x; \hat{\boldsymbol{\theta}}_{MD-}) || f_n(x)). \end{cases}$$

3.5. Bias-reduced estimator for parameter μ

In Sections 3.2, 3.3 and 3.4, we studied different methods to obtain the four parameters in the four-parameter $G\Gamma$ distribution. After obtaining the estimates

of a, ℓ, c (denoted as $\hat{a}, \hat{\ell}, \hat{c}$), from Eq. (3), a bias-reduced estimator of μ can be obtained based on the first sample moment as

$$\hat{\mu}_{BR} = \begin{cases} \min \left\{ y_{1:n}, \frac{1}{n} \sum_{i=1}^n y_i - \frac{\hat{a}\Gamma(\hat{\ell}+1/\hat{c})}{\Gamma(\hat{\ell})} \right\}, & \text{for } \hat{a} > 0, \\ \max \left\{ y_{n:n}, \frac{1}{n} \sum_{i=1}^n y_i - \frac{\hat{a}\Gamma(\hat{\ell}+1/\hat{c})}{\Gamma(\hat{\ell})} \right\}, & \text{for } \hat{a} < 0. \end{cases} \quad (13)$$

3.6. Standard error estimation and confidence interval based on bootstrap procedure

260 In statistical practice, in addition to reporting point estimates of the model parameters, it is desirable to give an estimate of the variation of each parameter estimate or an interval estimate with a certain degree of confidence for each parameter or a function of the parameters (e.g., the percentiles and survival probabilities) in general. Since the analytic asymptotic variances of the estimators proposed in Section 3.2, 3.3, and 3.4 are intractable due to different reasons 265 such as the parameter spaces are depending on the parameters themselves, we propose using the bootstrap method (see, for example, [47]) to estimate the standard error of the estimate and the confidence interval of a function of model parameters, $\lambda(a, \ell, c, \mu)$ in the four-parameter GF distribution, in which the individual model parameters can be considered as special cases of the function λ . 270 Suppose the estimate of the parameter vector $\boldsymbol{\theta} = (a, \ell, c, \mu)$ obtained by one of the methods presented in Sections 3.2, 3.3, and 3.4 is denoted as $\hat{\boldsymbol{\theta}}_* = (\hat{a}, \hat{\ell}, \hat{c}, \hat{\mu})$, where $*$ = *ML, KL, JSB, HD, WD, SE, HDC, WDC*, or *SEC*, the following bootstrap procedure can be used to obtain the standard error estimate and the 275 confidence interval of a function of the model parameters $\lambda(a, \ell, c, \mu)$:

Step B1. Generate a bootstrap sample, $\mathbf{y}^{(b)} = (y_1^{(b)}, y_2^{(b)}, \dots, y_n^{(b)})$, from the four-parameter GF distribution with parameter vector $\hat{\boldsymbol{\theta}}_* = (\hat{a}, \hat{\ell}, \hat{c}, \hat{\mu})$ using Property (4) presented in Section 2.2.

Step B2. Based on the bootstrap sample $\mathbf{y}^{(b)}$, obtain the estimates of the four parameters using the same method to obtain $\hat{\boldsymbol{\theta}}_*$, and denote the estimate of 280 $\boldsymbol{\theta}$ based on $\mathbf{y}^{(b)}$ as $\hat{\boldsymbol{\theta}}_*^{(b)} = (\hat{a}^{(b)}, \hat{\ell}^{(b)}, \hat{c}^{(b)}, \hat{\mu}^{(b)})$.

Step B3. Repeat Steps B1–B2 for B times to obtain $\hat{\theta}_*^{(1)}, \hat{\theta}_*^{(2)}, \dots, \hat{\theta}_*^{(B)}$.

Step B4. The bootstrap estimate of standard error of $\lambda(\hat{a}, \hat{\ell}, \hat{c}, \hat{\mu})$ can be obtained as

$$\sqrt{\frac{1}{B} \sum_{b=1}^B [\lambda(\hat{a}^{(b)}, \hat{\ell}^{(b)}, \hat{c}^{(b)}, \hat{\mu}^{(b)}) - \bar{\lambda}]^2},$$

285 where

$$\bar{\lambda} = \frac{1}{B} \sum_{b=1}^B \lambda(\hat{a}^{(b)}, \hat{\ell}^{(b)}, \hat{c}^{(b)}, \hat{\mu}^{(b)}).$$

Step B5. Order $\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(B)}$, where $\lambda^{(b)} = \lambda(\hat{a}^{(b)}, \hat{\ell}^{(b)}, \hat{c}^{(b)}, \hat{\mu}^{(b)})$, to obtain $\lambda^{[1]} < \lambda^{[2]} < \dots < \lambda^{[B]}$. Then, a $100(1 - \pi)\%$ percentile bootstrap confidence interval of $\lambda(a, \ell, c, \mu)$ can be obtained as

$$\left\{ \lambda^{[int(B(\pi/2))]}, \lambda^{[int(B(1-\pi/2))]} \right\},$$

where $int(x)$ is the integer part of x .

290 Note that if one wants to estimate the standard error and/or construct confidence interval for a particular parameter, for example the parameter a , we can set $\lambda(a, \ell, c, \mu) = a$.

4. R Package

There are some existing R packages that have the capability to estimate the
 295 parameter of different statistical distributions using maximum likelihood estimation method in which suitable modifications of the R programs are needed for applying to the four-parameter GF distribution. For example, the **fitdistrplus** package [48], the **MASS** package [49], and the **survival** package [50]. There are also some R packages for fitting the three-parameter generalized gamma distribution, a special case of the four-parameter GF distribution with $\mu = 0$ such as
 300 the **fit.GenGamma** function in the R package **temporal** (Parametric Time to Event Analysis) [51]. However, these existing R functions can only obtain the

maximum likelihood estimates of the model parameters for some special cases of the four-parameter GG distribution with $c > 0$ (power shape parameter) and $a > 0$ (positive scale parameter) or require substantial R programming to modify the code.

Here, we aim to develop an R package to estimate the parameters of the four-parameter GG distribution based on different estimation methods proposed in Section 3. The package called **AmoRosoDistrib** (<https://github.com/ccombesGG4/AmoRosoDistrib/>) written in R [52] is a general package that aims at fitting univariate parametric distributions in the Amoroso family of distributions with the scale parameter $a \in \mathbb{R} \setminus \{0\}$, the location parameter $\mu \in (-\infty, \infty)$ and the shape parameters $\ell \in \mathbb{R}^+$ and $c \in \mathbb{R} \setminus \{0\}$.

The **AmoRosoDistrib** package also provides the PDF, CDF, quantile function, hazard function for the four-parameter GG distribution using the parametrization originating from Amoroso [1]. The package also includes a random number generation for the four-parameter GG distribution based on the property that if $Z \sim \text{Gamma}(1, \ell)$, then $Y = aZ^{\frac{1}{c}} + \mu \sim GG(a, \ell, c, \mu)$.

The **AmoRosoDistrib** package includes different parameter estimation methods based on optimization approach to fit the univariate population. By default, direct optimization is performed using the general-purpose constraint optimization function `constrOptim` with Nelder-Mead method available in the R package **stats** [52] which minimizes a function subject to linear inequality constraints using an adaptive barrier algorithm. In summary, the nine estimation methods that are provided in the package are:

1. Maximum Likelihood Estimate: `fit.mle()` (see, Eq. (4));
2. Minimum Kullback-Leibler divergence estimate based on PDF: `fit.mkle()` (see, Eq. (5));
3. Minimum Jensen-Shanon divergence estimate based on PDF: `fit.mjse()` (see, Eq. (6));
4. Minimum Hellinger distance estimate based on PDF: `fit.mhe()` (see, Eq. (7));

5. Minimum Wasserstein distance estimate based on PDF: `fit.mwe()` (see, Eq. (8));
- 335 6. Minimum squared distance estimate based on PDF: `fit.msqe()` (see, Eq. (9));
7. Minimum Hellinger distance estimate based on CDF: `fit.mhdfe()` (see, Eq. (10));
8. Minimum Wasserstein distance estimate based based on CDF: `fit.mwdfe()` (see, Eq. (11));
- 340 9. Minimum squared distance estimate based on CDF: `fit.msqdfc()` (see, Eq. (12)).

For all these estimation methods, the users have the flexibility to provide the initial values of the parameters. Since the initial values are essential for the optimization algorithm and they may have a strong impact on the quality of the estimates due to the problem of local optimum, we have also provide the function called `init.theta()` in the **AmoRosoDistrib** package to obtain reasonable initial values based on the method presented in Section 3.1. Moreover, we also provide the function called `mu.bias_reduced_estimate` to compute the bias-reduced estimator for parameter μ based on Eq. (13).

In addition, the **AmoRosoDistrib** package also provides functions to compute different theoretical characteristics of the GF4 distributions (i.e., the values of the parameters are input variables) and different statistics computed from a sample (i.e., the data are the input variables):

- 355 • Theoretical mean: `mean_theo`;
- Theoretical variance: `var_theo`;
- Theoretical moments: `moment_theo`;
- Sample moments: `moment`;
- Sample skewness: `skewness`;
- 360 • Sample kurtosis: `kurtosis`.

5. Numerical Examples

In this section, we use two real data examples to illustrate the proposed parameter estimation procedures for four-parameter GF distribution.

5.1. Neurotransmission in Guinea Pig Brains (`paulsen` data)

365 The `paulsen` data frame in R package `boot` [53, 47] has 346 observations of the spontaneous current flowing into individual brain cells in the brain of adult guinea pigs. The currents are measured in pico-amperes. These observations are obtained from an experiment to see if the current flow was quantal in nature [54].

370 We fit the `paulsen` data by using the four-parameter GF distribution. First, we use the method proposed in Section 3.1 with 1,000 equal-spaced points in $(-20, 20)$ as the sequence of values for parameter c to obtain the initial value of θ in the parameter space $\Omega_+ = \{a > 0, \mu \in (y_{1:n}, \infty), \ell \in \mathbb{R}^+, c \in \mathbb{R} \setminus \{0\}\}$, and the initial value of θ in the parameter space $\Omega_- = \{a < 0, \mu \in (-\infty, y_{n:n}),$
375 $\ell \in \mathbb{R}^+, c \in \mathbb{R} \setminus \{0\}\}$. Then, based on these initial estimates, we obtain the parameter estimates using the estimation methods described in Section 3 in parameter spaces Ω_+ and Ω_- . The parameter estimates are presented in Table 1 and the parameter estimates that give a better value of objective function are indicated by an asterisk. The histogram and the fitted PDFs for the `paulsen`
380 data with the two initial estimates, and the estimates obtained from different estimation methods (i.e., the set of estimates with an asterisk in Table 1) are presented in Figure 2. From Figure 2, we observe that all the proposed methods, including the initial estimates, provide a reasonable fit to the data. Moreover, the fitted PDFs based on different parameter estimation methods are similar
385 except for the initial estimate in parameter spaces Ω_- and the MDE-HD.

To provide an estimate of the variation of each parameter estimate and an interval estimate with a certain degree of confidence for each parameter, we compute the estimated standard errors of the selected point estimates (indicated with asterisk in Table 1) and the 95% confidence intervals the four parameters

Table 1: Parameter estimates of the four-parameter GF distribution based on different estimation methods for the `paulsen` data set.

Methods	Parameter	Parameter estimates				Value of objective function
	space	a	ℓ	c	μ	
Initial	Ω_+	2.56418	2.65481	0.98098	2.87100	-949.2714*
	Ω_-	-24.25277	0.21465	20.00000	29.08800	-950.3481
MLE	Ω_+	0.00243	20.94687	0.38253	2.23448	-946.1885
	Ω_-	-33.63409	0.19869	30.07615	38.39699	-941.8875*
MDE-KL	Ω_+	2.46675	2.74086	0.94761	2.90000	3.577317
	Ω_-	-34.21816	0.30807	19.84622	39.15368	2.43993*
MDE-JSB	Ω_+	1.35520	3.75067	0.81873	2.88080	7.25734
	Ω_-	-35.00385	0.31365	21.09198	39.99006	4.86410*
MDE-HD	Ω_+	4.36420	1.79173	1.16580	2.90000	0.32401
	Ω_-	-351.85447	0.61295	141.59483	357.89817	0.02574*
MDE-WD	Ω_+	0.00005	27.45406	0.28308	2.75488	0.01190
	Ω_-	-113.95793	0.25992	85.50569	118.82425	0.00792*
MDE-SE	Ω_+	0.00011	27.18671	0.29909	2.55018	0.04358
	Ω_-	-213.03378	0.31778	137.54679	218.10832	0.00505*
MDE-HDC	Ω_+	2.19365	2.96368	0.98964	2.90000	0.18574*
	Ω_-	-41.89894	0.08910	81.44715	46.52610	0.21051
MDE-WDC	Ω_+	0.76866	4.48282	0.7036	2.90000	0.03549
	Ω_-	-104.08400	0.12016	167.1149	108.88144	0.02908*
MDE-SEC	Ω_+	0.21302	6.98729	0.57329	2.90000	0.28877
	Ω_-	-154.48710	0.16138	192.87155	159.35486	0.19522*

* indicates parameter estimates that give a better value of objective function.

Figure 2: Histogram and fitted PDFs of the `paulsen` data set.

390 based on the bootstrap procedure presented in Section 3.6 with $B = 1,000$
bootstrap samples. The results are presented in Table 2.

5.2. Anorexia Data on Weight Change (`anorexia` data)

The `anorexia` data set in the `MASS` package [49] contains the weight change
data for young female anorexia patients. Here, the studied variable is the Weight
395 of the patient after the study period (in lbs.): `Postwt`. There are 72 observations
in the data set. The objective of this example is to study the behavior of the

Table 2: Bootstrap standard error (SE) estimates and 95% confidence interval (CI) based
on 1,000 bootstrap data sets according to the four-parameter Γ distribution based on the
different parameter estimates (indicated with an asterisk) in Table 1.

Methods	\hat{a}		$\hat{\ell}$		\hat{c}		$\hat{\mu}$	
	SE	95% CI	SE	95% CI	SE	95% CI	SE	95% CI
MLE	16.389	(-38.577, -22.297)	0.062	(0.150, 0.332)	12.256	(25.631, 66.959)	16.545	(31.856, 78.683)
MDE-KL	3.386	(-32.832, -22.27)	0.111	(0.342, 0.707)	2.501	(12.064, 18.916)	3.456	(34.426, 41.243)
MDE-JSB	3.180	(-32.077, -22.433)	0.120	(0.350, 0.749)	2.657	(12.161, 19.952)	3.220	(33.773, 40.975)
MDE-HD	80.815	(-120.11, -23.148)	0.254	(0.400, 1.164)	22.790	(12.283, 86.72)	81.621	(33.259, 276.734)
MDE-WD	18.445	(-70.527, -29.073)	0.107	(0.305, 0.609)	12.183	(23.873, 62.426)	18.466	(52.859, 107.98)
MDE-SE	26.517	(-97.982, -36.992)	0.125	(0.370, 0.723)	14.570	(29.125, 72.250)	26.598	(72.134, 150.519)
MDE-HDC	2.658	(0.011, 8.143)	12.832	(1.525, 42.002)	0.530	(0.455, 2.052)	0.690	(2.311, 3.644)
MDE-WDC	25.040	(-74.122, -26.8)	0.064	(0.053, 0.244)	100.981	(92.511, 436.78)	25.146	(43.385, 119.962)
MDE-SEC	45.799	(-122.874, -27.732)	0.071	(0.095, 0.316)	64.087	(87.123, 261.553)	45.929	(55.326, 195.493)

Table 3: Parameter estimates of the four-parameter GF distribution based on different estimation methods for the **anorexia** data set.

Methods	Parameter	Parameter estimates				Value of
	space	a	ℓ	c	μ	objective function
Initial	Ω_+	25.92861	0.32126	4.38438	70.58700	-247.7085
	Ω_-	-32.07908	0.08792	18.91892	104.63600	-246.4883*
MLE	Ω_+	26.38376	0.26558	4.46531	71.20407	247.0739*
	Ω_-	-31.70010	0.10629	15.44827	104.47086	246.4163
MDE-KL	Ω_+	19.95212	0.83938	2.02986	70.38661	0.9072
	Ω_-	-56.07346	0.31234	12.66618	130.71678	0.72463*
MDE-JSB	Ω_+	0.018040	51.72513	0.52424	51.02371	1.66871
	Ω_-	-33.86042	0.28699	7.24023	107.65803	1.1343*
MDE-HD	Ω_+	24.75166	0.42261	3.26804	70.83969	0.0531
	Ω_-	-34.07224	0.27576	7.45461	107.65995	0.0121*
MDE-WD	Ω_+	19.05471	0.90292	1.86721	70.37946	0.0028
	Ω_-	-36.82913	0.25279	8.99935	110.51578	0.0021*
MDE-SE	Ω_+	18.72147	1.00642	1.89545	69.24949	0.0029
	Ω_-	-35.27432	0.27184	8.00305	108.99058	0.0014*
MDE-HDC	Ω_+	15.90206	1.13124	1.71907	70.00529	0.0156
	Ω_-	-32.52647	0.13146	13.79075	105.64742	0.0153*
MDE-WDC	Ω_+	14.18711	1.34909	1.59308	69.73335	0.0236
	Ω_-	-40.37422	0.29562	10.11036	115.24716	0.0228*
MDE-SEC	Ω_+	16.87272	0.99972	1.79398	70.42269	0.0263
	Ω_-	-35.81867	0.21207	10.84299	109.69528	0.0232*

* indicates parameter estimates that give a better value of objective function.

proposed estimation methods based on a relatively small sample.

Following the same approach in the previous numerical example in Section 5.1, the parameter estimates are presented in Table 3 and the parameter estimates that give a better value of objective function are indicated by an asterisk. The histogram and the fitted PDFs for the variable `Postwt` in **anorexia** data set with the two initial estimates, and the estimates obtained from different estimation methods (i.e., the set of estimates with an asterisk in Table 3) are presented in Figure 3. From Figure 3, we observe that all the proposed methods, provide a reasonable fit to the data, however, the variation between the fitted PDFs based on different parameter estimation methods are larger compared to the example in Section 5.1 probably due to the small sample size.

To provide an estimate of the variation of each parameter estimate and

Figure 3: Histogram and fitted PDFs of the `anorexia` data set.

an interval estimate with a certain degree of confidence for each parameter, we
 410 compute the estimated standard errors of the selected point estimates (indicated
 with asterisk in Table 3) and the 95% confidence intervals the four parameters
 based on the bootstrap procedure presented in Section 3.6 with $B = 1,000$
 bootstrap samples. The results are presented in Table 4.

Table 4: Bootstrap standard error (SE) estimates and 95% confidence interval (CI) based
 on 1,000 bootstrap data sets according to the four-parameter GF distribution based on the
 different parameter estimates in Table 3.

Methods	\hat{a}		$\hat{\ell}$		\hat{c}		$\hat{\mu}$	
	SE	95% CI	SE	95% CI	SE	95% CI	SE	95% CI
MLE	3.010	(25.431, 31.467)	0.496	(0.109, 0.573)	456.541	(4.053, 59.019)	0.629	(71.367, 73.094)
MDE-KL	14.052	(-68.267, -34.654)	0.722	(0.305, 2.921)	5.261	(6.835, 24.407)	15.444	(117.357, 156.425)
MDE-JSB	8.608	(-42.544, -25.217)	1.521	(0.267, 4.251)	4.348	(4.567, 20.874)	12.231	(105.939, 136.417)
MDE-HD	53.009	(-38.002, -26.799)	11.201	(0.278, 8.543)	3.077	(5.105, 12.173)	67.908	(105.403, 404.756)
MDE-WD	25.822	(-56.633, -26.289)	4.557	(0.255, 10.020)	6.193	(5.518, 22.031)	32.704	(108.004, 217.270)
MDE-SE	47.222	(-76.828, -25.871)	10.960	(0.271, 34.013)	7.617	(5.359, 27.562)	63.083	(107.004, 296.130)
MDE-HDC	21.764	(-35.483, -22.882)	2.432	(0.049, 4.646)	57.176	(7.466, 220.913)	25.370	(104.486, 198.399)
MDE-WDC	49.836	(-102.387, -25.894)	3.952	(0.158, 4.892)	38.825	(8.351, 117.528)	52.500	(109.416, 272.711)
MDE-SEC	48.788	(-61.370, -26.048)	3.799	(0.112, 4.463)	42.775	(8.347, 169.406)	52.058	(106.776, 281.766)

6. Monte Carlo Simulation Studies

415 In order to give a comprehensive evaluation of the proposed estimation procedures, we conduct extensive experimentation on synthetic univariate populations drawn from a specific Amoroso distribution, and compare results obtained with the parameter estimation methods presented in Section 3.

All the Monte Carlo simulations are performed with the R package
420 **AmoRosoDistrib** described in Section 4 and the code to reproduce the results can be obtained from the authors upon request. In our preliminary simulation studies, we observed that the variations of the parameter estimates, especially for parameter a , can be large (with standard deviations in the order of 10^{13} in some cases). Moreover, we observe that some of the optimization procedures
425 do not converge and the parameter estimates cannot be obtained. These results agree with the issues in estimating the parameters of the four-parameter GF distribution discussed in the literature (see, Section 1), which is one of the reasons we propose different estimation methods in this paper.

Since the variations of the parameter estimates can be large, in order to provide
430 reasonable and meaningful comparisons of different estimation methods, instead of using the conventional measures for comparing estimation methods using biases and variances, we consider some robust measures including the medians and interquartile ranges (IQRs). In addition, we also present the proportion of simulations that the parameter estimates cannot be obtained.

435 For the initial estimates, 1,000 equal-spaced points in $(-20, 20)$ as the sequence of values for parameter c are used to obtain the initial value of θ in the parameter space $\Omega_+ = \{a > 0, \mu \in (y_{1:n}, \infty), \ell \in \mathbb{R}^+, c \in \mathbb{R} \setminus \{0\}\}$, and the initial value of θ in the parameter space $\Omega_- = \{a < 0, \mu \in (-\infty, y_{n:n}), \ell \in \mathbb{R}^+, c \in \mathbb{R} \setminus \{0\}\}$.

440 6.1. Simulation setting based on the `paulsen` data

In this setting, we simulate the data sets with sample size $n = 346$ from the four-parameter GF distribution with parameters $a = -33.6$, $\ell = 0.2$, $c = 30.00$,

$\mu = 38.4$, which are the maximum likelihood estimates based on the `paulsen` data presented in Section 5.1. The simulated medians and IQRs, and the proportion of the set of estimates in the correct sample space (i.e., Ω_+ in this setting) is selected for each estimation method, except for the method to obtain the initial estimates, based on 2,000 simulations (denoted as % choose -ve) are presented in Table 5. The simulated medians closest to the true values of the parameters and the smallest IQR among all the methods, except for the method to obtain the initial estimates, are highlighted in bold.

Table 5: Simulated medians and IQRs based on 2000 simulated data sets from the four-parameter G Γ distribution with $(a = -33.6, \ell = 0.2, c = 30.0, \mu = 38.4)$ based on different estimation methods proposed in this paper.

Methods	Parameter space	% choose -ve	a (True value = -33.6)			ℓ (True value = 0.2)			c (True value = 30)			μ (True value = 38.4)		
			Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR		
Initial	Ω_+	-	4.602	2.324	1.626	0.778	1.221	0.320	3.330	0.422				
	Ω_-		-22.080	2.859	0.205	0.065	20.000	0.000	26.892	3.023				
MLE	Ω_+	0.750	2.099	3.632	2.899	4.246	0.927	0.481	3.054	0.724				
	Ω_-		-31.584	10.854	0.185	0.080	30.804	11.510	36.356	10.995				
MDE-KL	Ω_+	0.884	3.111	3.455	2.204	2.393	0.996	0.432	3.238	0.774				
	Ω_-		-30.616	3.750	0.306	0.078	17.620	3.953	35.649	3.670				
MDE-JSB	Ω_+	0.938	0.761	3.600	4.730	23.257	0.704	0.647	2.642	2.526				
	Ω_-		-30.709	3.647	0.315	0.091	17.569	4.279	35.792	3.604				
MDE-HD	Ω_+	0.885	5.688	2.200	1.336	0.555	1.327	0.296	3.340	0.486				
	Ω_-		-33.341	65.438	0.362	0.240	17.903	23.186	38.341	66.627				
MDE-WD	Ω_+	0.948	0.001	0.026	26.101	11.876	0.379	0.152	1.536	0.757				
	Ω_-		-42.468	23.441	0.329	0.125	23.581	13.250	47.562	23.792				
MDE-SE	Ω_+	0.990	0.001	0.010	25.743	11.183	0.377	0.117	1.406	0.843				
	Ω_-		-61.057	34.952	0.378	0.124	31.040	19.343	66.643	35.242				
MDE-HDC	Ω_+	0.631	0.192	1.666	7.087	21.621	0.569	0.508	2.967	0.640				
	Ω_-		-35.159	22.461	0.205	0.151	35.618	25.076	40.104	22.609				
MDE-WDC	Ω_+	0.779	0.941	2.715	4.071	5.298	0.737	0.438	3.237	0.563				
	Ω_-		-36.267	46.832	0.204	0.138	38.853	36.116	41.109	47.315				
MDE-SEC	Ω_+	0.789	1.170	2.739	3.701	4.065	0.786	0.411	3.241	0.553				
	Ω_-		-36.644	48.746	0.205	0.137	38.405	37.223	41.490	49.231				

From Table 5, we observe that the estimates in parameter space Ω_+ have larger variations compared to the estimates in parameter space Ω_- as the true value of parameter a is negative. We can see that the MDE-HD method perform well in terms of the closeness of the medians to the true values, however, the variations (IQRs) of this method are much larger than the other methods. When we take both IQRs and the closeness of the medians to the true values into account, the performance of the MDE-KL and MDE-JSB methods are better than the other methods. In this case, the MDE-JSB method is preferred since the proportion of times that the parameter estimates in the right sample space are identified for MDE-JSB (93.8%) is higher than that for the MDE-KL method (88.4%). It is clear that there is not a single estimation method outperforms the other methods in estimating the four parameters in the model.

6.2. Simulation setting based on the `anorexia` data

In this setting, we simulate the data sets with sample size $n = 72$ from a four-parameter Γ distribution with parameters $a = 26.4$, $\ell = 0.27$, $c = 4.47$, $\mu = 71.2$, which are the maximum likelihood estimates based on the `anorexia` data presented in Section 5.2. The simulated medians and IQRs, and the proportion of the set of estimates in the correct sample space (i.e., Ω_+ in this setting) is selected for each estimation method, except for the method to obtain the initial estimates, based on 2,000 simulations (denoted as % choose +ve) are presented in Table 6. The simulated medians closest to the true values of the parameters and the smallest IQR among all the methods, except for the method to obtain the initial estimates, are highlighted in bold.

Table 6: Simulated medians and IQRs based on 2000 simulated data sets from the four-parameter GT distribution with $(a = 26.4, \ell = 0.27, c = 4.47, \mu = 71.2)$ based on different estimation methods proposed in this paper.

Methods	Parameter space	% choose +ve	a (True value = 26.4)		ℓ (True value = 0.27)		c (True value = 4.47)		μ (True value = 71.2)	
			Median	IQR	Median	IQR	Median	IQR	Median	IQR
Initial	Ω_+	-	26.365	4.056	0.255	0.211	4.985	3.724	70.997	0.622
	Ω_-	-	-30.511	4.379	0.163	0.190	13.253	12.492	104.726	3.693
MLE	Ω_+	0.910	27.420	3.485	0.173	0.146	5.607	4.021	71.641	0.667
	Ω_-	-	-31.798	4.582	0.045	0.151	32.013	6863.828	104.756	4.882
MDE-KL	Ω_+	0.621	25.831	6.340	0.469	0.511	3.477	2.432	70.568	3.057
	Ω_-	-	-33.065	10.221	0.481	2.271	5.441	5.204	124.337	26.491
MDE-JSB	Ω_+	0.667	26.133	5.830	0.425	0.577	3.676	2.466	69.756	4.100
	Ω_-	-	-32.949	8.384	0.403	2.052	5.943	6.047	111.732	25.924
MDE-HD	Ω_+	0.469	26.432	4.343	0.377	0.299	3.895	1.814	69.913	3.073
	Ω_-	-	-32.329	6.351	0.296	0.202	6.743	2.257	106.203	6.959
MDE-WD	Ω_+	0.516	26.073	6.418	0.479	0.459	3.328	2.334	69.651	3.725
	Ω_-	-	-31.854	6.850	0.356	0.384	5.841	3.868	106.308	7.659
MDE-SE	Ω_+	0.389	26.050	6.529	0.476	0.528	3.413	2.462	69.624	3.745
	Ω_-	-	-33.387	16.428	0.370	1.065	6.301	4.049	107.325	74.431
MDE-HDC	Ω_+	0.542	25.745	5.648	0.290	0.418	4.367	5.352	71.032	1.290
	Ω_-	-	-31.141	8.574	0.467	1.584	5.829	7.198	108.116	12.605
MDE-WDC	Ω_+	0.623	26.143	6.203	0.315	0.501	4.330	4.966	70.716	3.546
	Ω_-	-	-31.185	6.861	0.288	0.509	7.534	9.602	105.914	7.525
MDE-SEC	Ω_+	0.620	26.155	6.117	0.314	0.495	4.406	5.083	70.788	3.080
	Ω_-	-	-31.374	6.723	0.284	0.556	7.651	10.862	106.005	7.855

475 From Table 6, once again, we observe that there is not a single estimation
method that outperforms the other methods in estimating the four parameters
in the model. The MLE and MDE-HD methods perform reasonably well for
all the four parameters in terms of both the closeness of the medians to the
true values and the IQRs. In this case, the MLE method is preferred since
480 the proportion of times that the parameter estimates in the right sample space
are identified for MLE (91.0%) is higher than that for the MDE-HD method
(46.9%).

6.3. Simulation setting based on the bladder cancer data

In this subsection, we consider the simulation setting based on the model
485 fitting of a data set that represents the remission times (in months) of a random
sample of 128 bladder cancer patients reported in Lee and Wang [55] based on
the three-parameter GF distribution (i.e., with $\mu = 0$ in the four-parameter
GF distribution).

Shanker and Shukla [56] modeled this bladder cancer data by using the
490 three-parameter GF distribution and reported the parameter estimates (the
formulation in Eq. (1) with $\mu = 0$) as $\hat{a} = 0.52813$, $\ell = 3.8869$, $c = 0.5139$ (see,
Table 1 of [56]). Here, we simulate the data sets with sample size $n = 128$ from
a four-parameter GF distribution with parameters $a = 0.52813$, $\ell = 3.8869$,
 $c = 0.5139$, and $\mu = 0$. The simulated medians and IQRs, and the proportion
495 of the set of estimates in the correct sample space (i.e., Ω_+ in this setting) is
selected for each estimation method, except for the method to obtain the initial
estimates, based on 2,000 simulations (denoted as % choose +ve) are presented
in Table 7. The simulated medians closest to the true values of the parameters
and the smallest IQR among all the methods, except for the method to obtain
500 the initial estimates, are highlighted in bold.

Table 7: Simulated medians and IQRs based on 2000 simulated data sets from the four-parameter GT distribution with $(a = 0.52813, \ell = 3.8869, c = 0.5139, \mu = 0)$ based on different estimation methods in this paper.

Methods	Parameter space	% choose +ve	a (True value = 0.52813)		ℓ (True value = 3.8869)		c (True value = 0.5139)		μ (True value = 0)	
			Median	IQR	Median	IQR	Median	IQR	Median	IQR
Initial	Ω_+	-	6.489	5.076	1.242	0.628	0.861	0.320	0.225	0.197
	Ω_-		-47.385	21.836	0.817	0.976	20.000	7.467	53.782	20.257
MLE	Ω_+	0.966	4.399	8.632	1.623	2.755	0.780	0.516	0.171	0.244
	Ω_-		-462.761	851.280	0.023	0.054	3263.908	41454.460	462.850	840.460
MDE-KL	Ω_+	0.836	3.660	3.121	2.051	0.795	0.749	0.214	0.199	0.215
	Ω_-		-55.863	22.700	0.246	0.116	23.953	12.237	61.516	18.922
MDE-JSB	Ω_+	0.655	2.025	3.554	2.784	2.793	0.675	0.317	0.056	2.246
	Ω_-		-56.852	21.247	0.274	0.122	23.976	10.963	61.681	17.465
MDE-HD	Ω_+	0.498	6.010	4.390	1.483	0.692	0.904	0.296	0.189	0.190
	Ω_-		-61.045	100.716	0.285	0.135	24.199	18.934	68.091	185.830
MDE-WD	Ω_+	0.933	0.015	11.045	17.615	23.697	0.444	0.857	-2.941	4.403
	Ω_-		-114.570	61.085	0.303	0.108	45.034	32.613	119.090	60.045
MDE-SE	Ω_+	0.879	0.004	4.348	19.013	19.956	0.380	0.396	-3.319	4.172
	Ω_-		-166.282	91.514	0.330	0.113	62.191	43.183	171.255	83.701
MDE-HDC	Ω_+	0.892	0.257	5.144	4.690	13.940	0.464	0.584	0.036	0.321
	Ω_-		-146.334	93.396	0.020	0.075	648.114	1012.150	151.689	84.631
MDE-WDC	Ω_+	0.944	0.852	5.063	3.285	9.043	0.544	0.506	0.115	0.389
	Ω_-		-77.388	45.221	0.023	0.094	352.543	522.557	84.236	51.0312
MDE-SEC	Ω_+	0.933	0.756	5.189	3.343	10.360	0.540	0.518	0.108	0.395
	Ω_-		-94.478	84.023	0.037	0.103	284.195	541.808	104.299	85.928

From the simulation results in Table 7, we observe that the MDE-SEC methods perform reasonably well for all the four parameters in terms of both the closeness of the medians to the true values, the IQRs and the proportion of times that the parameter estimates in the right sample space.

505 From the three settings presented in Sections 6.1, 6.2, and 6.3, the performance of the estimation methods proposed in this paper can be varied and there is not a single method better than the others in all the situations. Therefore, as a practical recommendation, researchers can consider using Monte Carlo simulation methods to evaluate the performance of the estimation procedures under
510 some particular settings specified based on the context of the problem.

7. Concluding Remarks

The Amoroso family of distributions are widely used in many areas for modeling a real phenomenon. The main advantage of the four-parameter GF distribution distribution is its flexibility regarding the varieties of shapes and hazard
515 functions for modeling real-life data. However, the flexibility of GF distribution comes at the price of having difficulties in parameter estimation since different values of parameters may appear almost identical fit of PDF, as discussed in Lawless [10].

In this paper, we present a unified framework for fitting a univariate data set
520 by using the four-parameter GF distribution with different parameter estimation methods. Our proposed approaches have taken the negative shape parameter and negative scale parameter into account. Since the initial values have an impact on the efficiency of the iterative numerical algorithms to obtain the parameter estimates, we also present a feasible and reliable method to obtain
525 initial estimates for those iterative numerical algorithms. The computational algorithms for the methods presented in this paper are programmed in R and made available as an R package called **AmoRosoDistrib**.

From the numerical and simulation studies, we show that the proposed methods are feasible and efficient in fitting the four-parameter GF distribution to

530 data. However, we found that there is no one single method presented in this paper that outperforms the other methods in all situations. Therefore, overall speaking, we propose applying all the estimation methods and graphically display the fitted PDFs of the $G\Gamma$ distribution in order to assess the suitability of using a specific set of parameter estimates.

535 For future research, exploring different methods for fitting the four-parameter $G\Gamma$ distribution to incomplete data, such as censored and truncated data in reliability/survival analysis, is an interesting topic. For instance, the study concerning a three-parameter $G\Gamma$ distribution with $a > 0$ and $c > 0$ by Shang and Ng [57] based on left-truncated and right-censored data can be extended
540 to the four-parameter $G\Gamma$ distribution with $a \in \mathbb{R} \setminus \{0\}$ for the scale parameter, $\mu \in (-\infty, \infty)$ for the location parameter and $c \in \mathbb{R} \setminus \{0\}$. Research in this direction is in progress and we hope to report the results in a future paper.

Acknowledgment

545 The authors are grateful for the referees and editor for their helpful and constructive suggestions that improved the manuscript. H. K. T. Ng's work was supported by a grant from the Simons Foundation (#709773 to Tony Ng).

References

- [1] L. Amoroso, Ricerche intorno alla curva dei redditi, *Annali de Mathematica*
550 2 (1) (1925) 123–159.
- [2] R. D'Addario, Intorno alla curva dei redditi di amoroso, *Riv Italiana Statist Econ Finanza* 4 (1) (1932) 723–729.
- [3] E. W. Stacy, A generalization of the gamma distribution, *The Annals of Mathematical Statistics* 33 (3) (1962) 1187–1192.
- 555 [4] K.-S. Song, Globally convergent algorithms for estimating generalized gamma distributions in fast signal and image processing, *IEEE Transactions on Image Processing* 17 (2008) 1233–1250.

- [5] V. Parr, J. Webster, A method for discriminating between failure density functions used in reliability predictions, *Technometrics* 7 (1) (1965) 1–10.
- 560 [6] E. W. Stacy, G. A. Mihram, Parameter estimation for a generalized gamma distribution, *Technometrics* 7 (3) (1965) 349–358.
- [7] W. H. Hager, B. L. J., Inferential procedures for the generalized gamma distribution, *Journal of the American Statistical Association* 65 (332) (1970) 1601–1609. doi:10.1080/01621459.1970.10481190.
- 565 [8] E. W. Stacy, Quasimaximum likelihood estimators for two-parameter gamma distributions, *IBM Journal of Research and Development* 17 (2) (1973) 115–124.
- [9] R. L. Prentice, A log gamma model and its maximum likelihood estimation, *Biometrika* 61 (3) (1974) 539–544.
- 570 [10] J. F. Lawless, Inference in the generalized gamma and log gamma distributions, *Technometrics* 22 (3) (1980) 409–419.
- [11] J. R. Hobbs, A. H. Moore, W. James, Minimum distance estimation of the three parameters of the gamma distribution, *IEEE Transactions on Reliability* 33 (3) (1984) 237–240.
- 575 [12] A. Noufaily, M. Jones, On maximization of the likelihood for the generalized gamma distribution, *Computational Statistics* 28 (2) (2013) 505–517.
- [13] H. Yilmaz, H. S. Sazak, Double-looped maximum likelihood estimation for the parameters of the generalized gamma distribution, *Mathematics and Computers in Simulation* 98 (2014) 18–30.
- 580 [14] H. Hirose, Maximum likelihood parameter estimation in the three-parameter gamma distribution, *Computational Statistics & Data Analysis* 20 (4) (1995) 343–354.
- [15] H. Hirose, Maximum likelihood parameter estimation by model augmentation with applications to the extended four-parameter generalized gamma

- 585 distribution, *Mathematics and Computers in Simulation* 54 (1) (2000) 81–97.
- [16] E. L. Allgower, K. Georg, *Introduction to Numerical Continuation Methods*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2003.
- 590 [17] N. Balakrishnan, S. Pal, An em algorithm for the estimation of parameters of a flexible cure rate model with generalized gamma lifetime and model discrimination using likelihood- and information-based methods, *Computational Statistics* 30 (03 2014). doi:10.1007/s00180-014-0527-9.
- [18] M. H. Ling, A comparison of estimation methods for generalized gamma
595 distribution with one-shot device testing data, *International Journal of Applied & Experimental Mathematics* 3 (2018) 1–7. doi:10.15344/2456-8155/2018/124.
- [19] P. Ramos, D. Nascimento, P. H. Ferreira da Silva, K. Weber, T. Santos, F. Louzada, Modeling traumatic brain injury lifetime data: Improved estimators for the generalized gamma distribution under small samples, *PLoS ONE* (08 2019). doi:10.1371/journal.pone.0221332.
600
- [20] S. Pal, H. Yu, Z. Loucks, I. Harris, Illustration of the flexibility of generalized gamma distribution in modeling right censored survival data: Analysis of two cancer datasets, *Annals of Data Science* 7 (03 2020).
605 doi:10.1007/s40745-019-00224-5.
- [21] H. N. Phien, T. V. Van Nguyen, J.-H. Kuo, Estimating the parameters of the generalized gamma distribution by mixed moments, *Hydrologic Frequency Modeling* (1984) 407–418.
- [22] D. R. Wingo, Computing maximum-likelihood parameter estimates of the
610 generalized gamma distribution by numerical root isolation, *IEEE Transactions on Reliability* 36 (5) (1987) 586–590.

- [23] F. Ashkar, B. Bobée, D. Leroux, D. Morissette, The generalized method of moments as applied to the generalized gamma distribution, *Stochastic Hydrology and Hydraulics* 2 (3) (1988).
- 615 [24] A. N. A. Maur, Statistical Tools for Drop Size Distributions: Moments and Generalized Gamma, *Journal of the Atmospheric Sciences* 58 (4) (2001) 407–418.
- [25] P.-H. Huang, T.-Y. Hwang, On new moment estimation of parameters of the generalized gamma distribution using it’s characterization, *Taiwanese Journal of Mathematics* 10 (4) (2006) 1083–1093.
- 620 [26] O. Gomes, C. Combes, A. Dussauchoy, Four-parameter generalized gamma distribution used for stock return modelling, in: *The Proceedings of the Multiconference on “Computational Engineering in Systems Applications”*, Vol. 1, 2006, pp. 380–386.
- 625 [27] O. Gomes, C. Combes, A. Dussauchoy, Parameter estimation of the generalized gamma distribution, *Mathematics and Computers in Simulation* 79 (4) (2008) 955–963.
- [28] D. Kurniasari, W. Widiarti, Y. Antonio, M. Azram, M. Usman, Estimation of generalized gamma distribution parameters with probability weighted moment method, *Sci.Int.(Lahore)* 30 (1) (2018) 1–6.
- 630 [29] M. Maswadah, A. M. Seham, M. Ahsanullah, Bayesian inference on the generalized gamma distribution based on generalized order statistics, *Journal of Statistical Theory and Applications* 12 (2013) 356–377.
- [30] R. Kishan, Comparison between MLE and Bayes estimators of scale parameter of generalized gamma distribution with known shape parameter under squared error loss function, *Journal of Reliability and Statistical Studies* 7 (1) (2014) 43–50.
- 635

- [31] K. Ahmad, S. P. Ahmad, A. Ahmed, J. A. Reshi, Bayesian analysis of generalized gamma distribution using R software, *Journal of Applied Probability and Statistics* 4 (2) (2015) 323–335.
- [32] S. Naqash, S. Ahmad, A. Ahmed, Bayesian analysis of generalized gamma distribution, *Journal of Applied Probability and Statistics* 4 (3) (2015) 499–512.
- [33] J. A. Achcar, P. L. Ramos, E. Z. Martinez, Some computational aspects to find accurate estimates for the parameters of the generalized gamma distribution, *Pesquisa Operacional* 37 (2) (2017) 365–385.
- [34] P. L. Ramos, F. Louzada, Bayesian reference analysis for the generalized gamma distribution, *IEEE Communications Letters* 22 (9) (2018) 1950–1953.
- [35] J. Wolfowitz, The minimum distance method, *The Annals of Mathematical Statistics* 28 (1) (1957) 75–88.
- [36] G. E. Crooks, The Amoroso distribution, arXiv e-prints (2015) arXiv:1005.3274v2arXiv:1005.3274v2.
- [37] A. Bücher, J. Segers, On the maximum likelihood estimator for the generalized extreme-value distribution, *Extremes* 20 (4) (2017) 839–872.
- [38] C.-D. Lai, *Generalized Weibull Distributions*, Springer-Verlag, Berlin Heidelberg, 2004.
- [39] V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*, Springer Netherlands, Dordrecht, 1998. doi:10.1007/978-94-017-1431-0_14.
URL https://doi.org/10.1007/978-94-017-1431-0_14
- [40] C. Cohen, A., T. Whitten, B., *Parameter Estimation in Reliability and Life Span Models*, *Statistics: A Series of Textbooks and Monographs*, 1988.
- [41] A. Basu, H. Shioya, C. Park, *Statistical inference: The minimum distance approach*, CRC Press, 2011.

- 665 [42] P. W. Millar, A general approach to the optimality of minimum distance estimators, *Transactions of the American Mathematical Society* 286 (1984) 377–418.
- [43] S. J. Sheather, M. C. Jones, A reliable data-based bandwidth selection method for kernel density estimation, *Journal of the Royal Statistical Society. Series B (Methodological)* 53 (3) (1991) 683–690.
670 URL <http://www.jstor.org/stable/2345597>
- [44] B. W. Silverman, *Density Estimation*, Chapman and Hall, London, 1986.
- [45] F. Bassetti, A. Bodini, E. Regazzini, On minimum kantorovich distance estimators, *Statistics & Probability Letters* 76 (12) (2006) 1298–1302.
675 doi:<https://doi.org/10.1016/j.spl.2006.02.001>.
URL <http://www.sciencedirect.com/science/article/pii/S0167715206000381>
- [46] E. Bernton, P. E. Jacob, M. Gerber, C. P. Robert, On parameter estimation with the Wasserstein distance, *Information and Inference: A Journal of the IMA* 8 (4) (2019) 657–676. arXiv:<https://academic.oup.com/imaiai/article-pdf/8/4/657/31772180/iaz003.pdf>,
680 doi:10.1093/imaiai/iaz003.
URL <https://doi.org/10.1093/imaiai/iaz003>
- [47] A. C. Davison, D. V. Hinkley, *Bootstrap Methods and Their Applications*,
685 Cambridge University Press, Cambridge, 1997.
URL <http://statwww.epfl.ch/davison/BMA/>
- [48] M. L. Delignette-Muller, C. Dutang, fitdistrplus: An R package for fitting distributions, *Journal of Statistical Software* 64 (4) (2015) 1–34.
URL <http://www.jstatsoft.org/v64/i04/>
- 690 [49] W. N. Venables, B. D. Ripley, *Modern Applied Statistics with S*, 4th Edition, Springer, New York, 2002, iSBN 0-387-95457-0.
URL <http://www.stats.ox.ac.uk/pub/MASS4/>

- [50] T. M. Therneau, A Package for Survival Analysis in R, r package version 3.2-3 (2020).
695 URL <https://CRAN.R-project.org/package=survival>
- [51] Z. McCaw, Temporal: Parametric Time to Event Analysis, R package version 0.2.0 (2020).
URL <https://CRAN.R-project.org/package=Temporal>
- [52] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2020).
700 URL <https://www.R-project.org/>
- [53] A. Canty, B. D. Ripley, boot: Bootstrap R (S-Plus) Functions, R package version 1.3-25 (2020).
- [54] O. Paulsen, P. Heggelund, The quantal size at retinogeniculate synapses determined from spontaneous and evoked epscs in guinea-pig thalamic slices.,
705 *The Journal of Physiology* 480 (3) (1994) 505–511.
- [55] E. T. Lee, J. W. Wang, *Statistical methods for survival data analysis*, 3rd ed. John Wiley and Sons: New York, USA, 2003.
- [56] R. Shanker, K. Shukla, On modeling of lifetime data using three-parameter generalized Lindley and generalized gamma distributions, *Biometrics & Biostatistics International Journal* 4 (2016) 283–288.
710
- [57] X. Shang, H. K. T. Ng, On parameter estimation for the generalized gamma distribution based on left-truncated and right-censored data, *Computational and Mathematical Methods* (2020) e1091doi:10.1002/cmm4.1091.