Article

A New Generalized *t* Distribution Based on a Distribution Construction Method

Ruijie Guan [ID]**, Xu Zhao, Weihu Cheng * and Yaohua Rong**

Faculty of Science, Beijing University of Technology, Beijing 100124, China; guanruijie@emails.bjut.edu.cn (R.G.); zhaox@bjut.edu.cn (X.Z.); rongyaohua@bjut.edu.cn (Y.R.)
* Correspondence: chengweihu@bjut.edu.cn

Abstract: In this paper, a new generalized *t* (new Gt) distribution based on a distribution construction approach is proposed and proved to be suitable for fitting both the data with high kurtosis and heavy tail. The main innovation of this article consists of four parts. First of all, the main characteristics and properties of this new distribution are outined. Secondly, we derive the explicit expression for the moments of order statistics as well as its corresponding variance–covariance matrix. Thirdly, we focus on the parameter estimation of this new Gt distribution and introduce several estimation methods, such as a modified method of moments (MMOM), a maximum likelihood estimation (MLE) using the EM algorithm, a novel iterative algorithm to acquire MLE, and improved probability weighted moments (IPWM). Through simulation studies, it can be concluded that the IPWM estimation performs better than the MLE using the EM algorithm and the MMOM in general. The newly-proposed iterative algorithm has better performance than the EM algorithm when the sample kurtosis is greater than 2.7. For four parameters of the new Gt distribution, a profile maximum likelihood approach using the EM algorithm is developed to deal with the estimation problem and obtain acceptable.

Keywords: generalized *t* distribution; order statistics; EM algorithm; improved PWM; parameter estimation

1. Introduction

The generalized *t* (Gt) distribution was first proposed by McDonald and Newey (1988) [1] to implement a partially adaptive regression models with the following probability density function (pdf)

$$f(x; p, q) = \frac{p}{2q^{1/p}B(1/p, q)} \left(1 + \frac{|x|^p}{q}\right)^{-(q+1/p)}, \quad p, q > 0, \tag{1}$$

where $B(a, b)$ denotes the beta function. With the addition of two shape parameters, p, q, the shape of the distribution becomes more flexibe so that it can fit a wider range of data. Nadarajah (2008) [2] obtained the analytic expression of the cumulative distribution function (cdf) through the incomplete beta function presented below

$$F(x; p, q) = \frac{1}{2}\left[1 + \text{sign}(x)I_{1-(1+|x|^p/q)^{-1}}\left(\frac{1}{p}, q\right)\right], \tag{2}$$

where $I_x(a, b) = \frac{1}{B(a,b)}\int_0^x t^{a-1}(1-t)^{b-1}dt$ denotes the incomplete beta function.

With the passage of time, the generalized *t* distribution warrants special attention by scholars and has been widely applied in several research areas. Galbraith and Zhu (2010) [3] introduced a new class of asymmetric Student *t* distributions and illustrated their applications in financial econometrics. Harvey and Lange (2017) [4] applied the generalized *t* distribution in autoregressive conditional heteroscedasticity models. Furthermore, the

generalized *t* distribution and its extension were involved in extensive research in the fields of robust estimation and robust statistical models, such as [5–8].

As the types of data become more and more widespread, many different generalized *t* distributions have been proposed and the related properties have been studied in depth. Arslan and Genç (2009) [9] obtained the skew-generalized *t* distribution through the scale mixture of the skew-exponential power and generalized gamma distributions. Venegas et al. (2012) [10] constructed a more flexible distribution by introducing skew coefficients *ε* to fit a wider range of data. Papastathopoulos and and Tawn (2013) [11] coined an extension of the Student's *t* distribution that allows for a negative shape parameter. Acitas et al. (2015) [12] introduced an alpha-skew-generalized *t* distribution based on a new skewing procedure, and presented the corresponding properties as well as the estimation methods. Lak et al. (2019) [13] proposed a distribution with a variety of shape variations, namely, the alpha-beta-skew-generalized *t* distribution, which contains several well-known distributions and illustrated its usefulness by means of a real data analysis. For Student's *t* Birnbaum-Saunders distribution, we recommend [14,15].

Although there is plenty of literature on the generalized *t* distribution, two aspects still need to be improved: (1) the type of data that previous distributions can be fitted to is not wide enough; (2) for the flexibility of the shape, the distribution function is often very complex, imposing limitations on the parameter estimation. To solve these two problems, the new Gt distribution we propose will fit both the data with a heavier tail and with higher kurtosis based on different shape parameters, with a more extensive application scope than the generalized *t* distributions introduced before. In addition, our new Gt distribution will make many classical estimation methods suitable for the parameter estimation problem of this distribution.

In this paper, we coined a new Gt distribution based on the distribution construction approach and denote it by $NGt(\mu, \sigma, \alpha, \beta)$. The main innovation of this work consists of four parts. Firstly, we investigate the main properties of the new Gt distribution, including moments, skewness coefficients, kurtosis coefficients and random number generation. Secondly, we derive explicit expressions for moments of order statistics as well as its corresponding variance–covariance matrix through recurrence relations and distribution transformation techniques when the shape parameter $\frac{1}{\beta}$ is either a real non-integer or an integer. Thirdly, we established an EM-type algorithm and a novel iterative algorithm to acquire the maximum likelihood estimation (MLE) of parameters for $NGt(0, \sigma, \alpha, \beta)$. The explicit analytical expressions for the asymptotic variance–covariance matrix of MLE were also presented based on the EM algorithm. Finally, we introduce improved probability-weighted moments (IPWM) based on order statistics transformation and obtain the corresponding algorithm to implement this novel method. For four parameters, $NGt(\mu, \sigma, \alpha, \beta)$, we propose a profile maximum likelihood approach (PLA) using the EM algorithm to deal with the estimation problem.

The rest of this article is organized as follows. In Section 2, we investigate the basic properties of this new distribution and make a comparison with the Gt distribution introduced by McDonald and Newey (1988) [1]. In Section 3, we derive explicit expressions for moments of order statistics from this distribution. In Section 4, we focus on the variance–covariance matrix of order statistics. In Section 5, several estimation methods are considered. In Section 6, we conduct several simulation studies to illustrate the performance of the proposed estimation method and compare these simulation results under different parameter values and different sample sizes. In Section 7, a real data analysis is performed to illustrate the feasibility of the new distribution we proposed. Discussions and conclusions are presented in Section 8 and proofs are given in the Appendix A.

2. Basic Theorems and Properties

In this section, some basic definitions, theorems and properties for the new Gt distribution are elucidated.

The new Gt distribution contains many common distributions, such as the normal distribution, the Laplace distribution, and the Pareto distribution. With the addition of two shape parameters, the shape of the distribution becomes more flexible, which can fit data having both heavy tails and high kurtosis.

Definition 1. *Let the random variable X with the following pdf*

$$f_X(x; \mu, \sigma, \beta) = \frac{\beta}{2^{1+1/\beta}\Gamma(1/\beta)\sigma} \exp\left\{ -\frac{|x-\mu|^\beta}{2\sigma^\beta} \right\}, \qquad x \in R,$$

where $\Gamma(\cdot)$ denotes the gamma function, $\mu \in R$ is a location parameter, $\sigma > 0$ is a scale parameter, and $\beta > 0$ is a shape parameter. Then the random variable X is said to have a generalized normal distribution and denoted by $X \sim GN(\mu, \sigma, \beta)$.

The definition of the new Gt distribution (denoted by $NGt(\mu, \sigma, \alpha, \beta)$) is stated as follows:

Definition 2. *Let the random variable $X - \mu \sim GN(0, \sigma, \beta)$ and $Y \sim Gamma(\alpha, \beta)$. Supposing X and Y are independent, we shall define $T = X/Y^{\frac{1}{\beta}} \sim NGt(\mu, \sigma, \alpha, \beta)$, where $\mu \in R$ is a location parameter, $\sigma > 0$ is a scale parameter, and $\alpha, \beta > 0$ are two shape parameters.*

Theorem 1. *The pdf of $NGt(\mu, \sigma, \alpha, \beta)$ can be expressed as follows*

$$g_T(t; \mu, \sigma, \alpha, \beta) = \frac{1}{\left(\frac{2}{\beta}\right)^{1+\frac{1}{\beta}} B(\alpha, \frac{1}{\beta})\sigma} \left[1 + \frac{\beta|t-\mu|^\beta}{2\sigma^\beta}\right]^{-(\alpha+\frac{1}{\beta})}, \qquad t \in R.$$

Proof. See Appendix A.1 (Part I). □

As is shown in the following two pictures (Figures 1 and 2), two shape parameters (α and β) have a decisive effect on the shape of the density function. The density function has a heavier tail with a smaller value of α and β, but it is the opposite when the parameters α and β are larger. Furthermore, the shape parameter β plays a decisive role in the shape of the distribution and the change of parameter α can control the tail behavior.
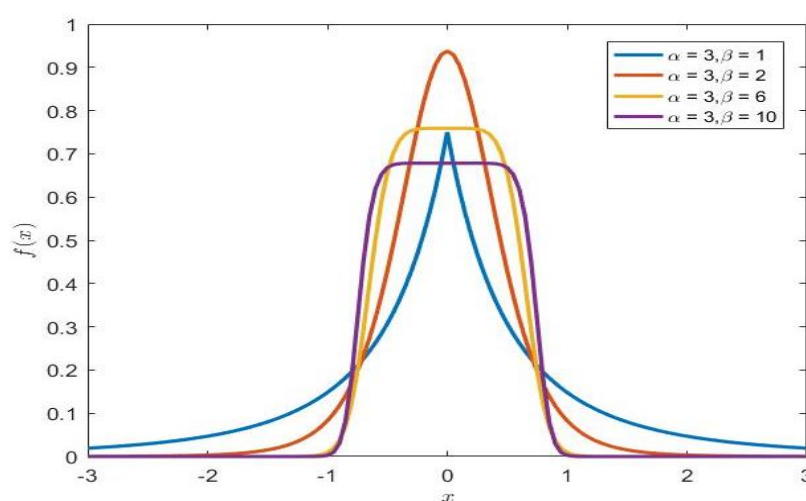

Figure 1. The pdf of the new Gt distribution for different values of β with fixed $\alpha = 3$.
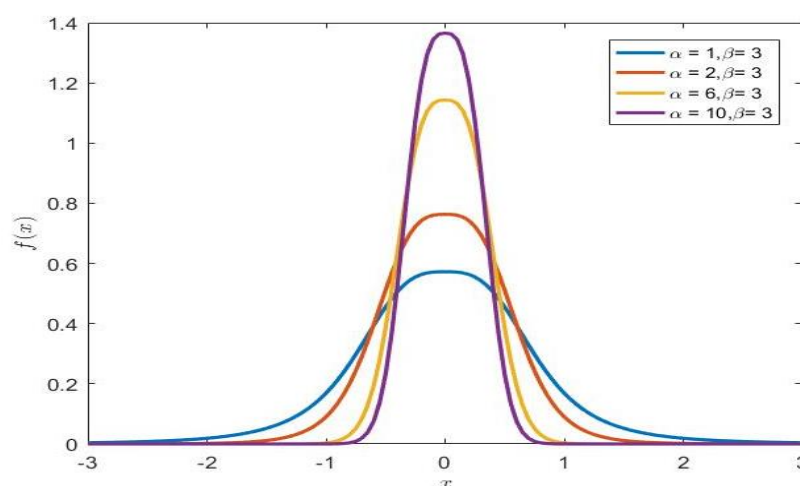
Figure 2. The pdf of the new Gt distribution for different values of α with fixed $\beta = 3$.

Figures 3 and 4 vividly describe the feature that the new Gt distribution is more suitable for fitting both the data with high kurtosis and with heavier tail than the Gt distribution coined by McDonald and Newey (1988) [1] under appropriate parameter range from two perspectives respectively.

It is well-known that the pdf with a steeper shape is suitable for fitting data with high kurtosis, while a pdf with a smoother shape can be used for fitting heavy-tailed data. Since these two distributions we compare are symmetric, we use the maximum value of the pdf to characterize the steepness of its shape. Let $K_1(\sigma, \alpha, \beta)$ and $K_2(\sigma, p, q)$ denote the maximum value of the pdf of Gt distribution and new Gt distribution, respectively.

Since the value of σ has little effect on the value of $K_1(\sigma, \alpha, \beta)$ and $K_2(\sigma, p, q)$, we may assume that $\sigma = 1$. On the one hand, it is not very difficult to prove that, for a fixed β, $K_1(\sigma, \alpha, \beta)$ decreases monotonically with respect to α. Then we have: $\lim\limits_{\alpha \longrightarrow \infty} K_1(\sigma, \alpha, \beta) \longrightarrow +\infty$, $\lim\limits_{\alpha \longrightarrow 0^+} K_1(\sigma, \alpha, \beta) \longrightarrow 0$. For $K_2(\sigma, p, q)$, the interval of the shape parameters p, q are (1, 10) and (0.1, 50), respectively. The reason for choosing these two ranges is that they include possible values for parameters in practical applications. From Figure 3, we can obviously find that the range of function $K_2(\sigma, p, q)$ is never greater than 1. This means that we can only adjust the parameter σ to make the pdf of the Gt distribution have a steeper shape. On the other hand, for fixed β and p, when α, $q \longrightarrow 0^+$, both $K_1(\sigma, \alpha, \beta)$ and $K_2(\sigma, p, q)$ tend to 0 but $K_1(\sigma, \alpha, \beta)$ has a faster convergence. This means that the pdf of the new Gt distribution proposed by us has a heavier tail when the parameters are within a reasonable range. Thus, the conclusion stated in the preceding paragraph is therefore self-evident.

The pdf generated under different shape parameters can be divided into four categories according to different shapes and the tail behavior, i.e., steep in shape and with a light tail, flat in shape and with a light tail, steep in shape and with a heavy tail, flat in shape and with a heavy tail. As is depicted in Figure 4, the solid line denotes the new Gt distribution, and the dashed line represents the previous Gt distribution defined by McDonald and Newey (1988) [1]. The subfigure Figure 4a–d depicts the pdf of the Gt distribution and the new Gt distribution under the above four categories with different shape parameters, respectively. It can be seen from the figure that, if the parameter value is in a reasonable range, the shape of the new Gt distribution is more flexible and more suitable for fitting a wider range of data.
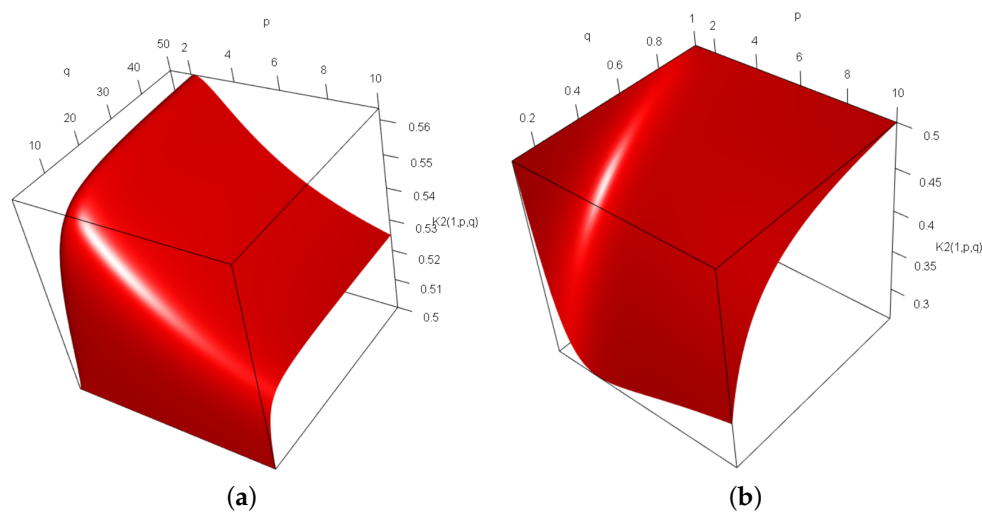
Figure 3. 3D plot for $K_2(1, p, q)$. (**a**) $1 \leq p \leq 10$, $1 \leq q \leq 50$. (**b**) $1 \leq p \leq 10$, $0.1 \leq q \leq 1$.
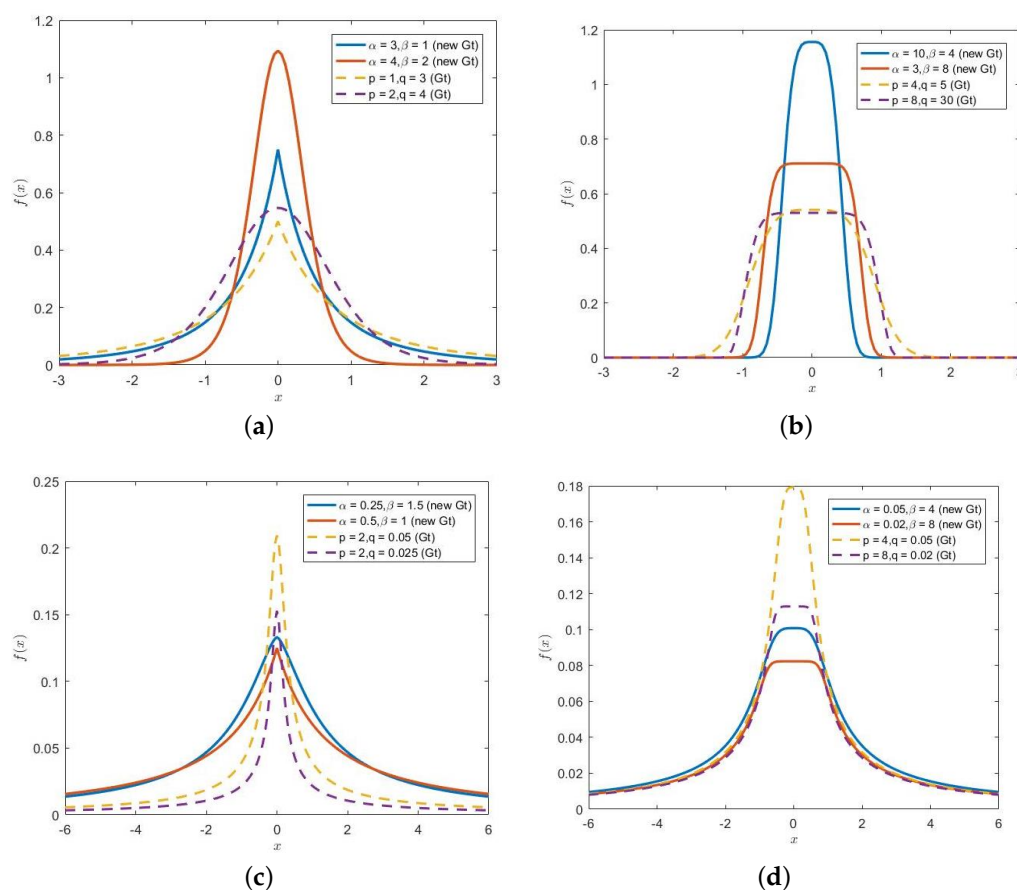
Figure 4. Comparison with the Gt distribution defined by McDonald and Newey.

Some important properties of the new Gt distribution are stated as follows:

Proposition 1. *If $Y|U = u \sim GN(\mu, \sigma u^{-\frac{1}{\beta}}, \beta)$, and $U \sim Gamma(\alpha, \beta)$, then $Y \sim NGt(\mu, \sigma, \alpha, \beta)$. $U|Y = y \sim Gamma(\alpha + \frac{1}{\beta}, \frac{1}{\beta} + \frac{|y-\mu|^\beta}{2\sigma^\beta})$.*

Proof. See Appendix A.1 (Part I). □

Proposition 2. *If $Y|U = u \sim GN(\mu, \sigma(\frac{\beta}{2}u)^{-\frac{1}{\beta}}, \beta)$ and $U \sim \chi^2(2\alpha)$, then $Y \sim NGt(\mu, \sigma, \alpha, \beta)$. $U|Y = y \sim Gamma(\alpha + \frac{1}{\beta}, \frac{1}{2} + \frac{\beta|y-\mu|^{\beta}}{4\sigma^{\beta}})$.*

Proposition 3. *If $X \sim Beta(\frac{1}{\beta}, \alpha)$, let $Y = \mu + \sigma \cdot W \cdot (2Z)^{\frac{1}{\beta}}$ and $Z = \frac{X}{\beta(1-X)}$, then $Y \sim NGt(\mu, \sigma, \alpha, \beta)$, where W is a random variable that only takes ± 1 and satisfies $P(W = 1) = P(W = -1) = \frac{1}{2}$.*

From the above proposition, we can follow the following steps to generate random samples from $NGt(\mu, \sigma, \alpha, \beta)$:

1. Generate $X \sim Beta(1/\beta, \alpha)$ and set $Z = \frac{X}{\beta(1-X)}$.
2. Generate $U \sim U(0, 1)$.
   if $U \leq \frac{1}{2}$, set $W = 1$.
   else set $W = -1$.
3. Set $Y = \mu + \sigma W(2Z)^{\frac{1}{\beta}}$ and return Y.

Proposition 4. *If $T \sim NGt(\sigma, \alpha, \beta)$, we have*
1. the 2k-th order moment of T is given by

$$E(X^{2k}) = \left(\frac{2}{\beta}\right)^{\frac{2k}{\beta}} \frac{\Gamma(\alpha - 2k/\beta)\Gamma[(2k+1)/\beta]}{\Gamma(\alpha)\Gamma(1/\beta)} \sigma^{2k}, \qquad k = 1, 2, \cdots,$$

and the above equation exists for $\alpha\beta > 2k$. Furthermore, from the symmetry of the distribution, we can conclude that $E(X^{2k-1}) = 0, k = 1, 2, \ldots,$
2. the kurtosis coefficient of T is given by

$$G_2(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\alpha - 4/\beta)\Gamma(1 + 1/\beta)\Gamma(5/\beta)}{\Gamma^2(\alpha - 2/\beta)\Gamma^2(3/\beta)} - 3 \qquad \alpha\beta > 4,$$

3. the mean deviation of T is given by

$$\delta = \frac{2\sigma B(2/\beta, \alpha - 1/\beta)}{(2/\beta)^{1-1/\beta} B(\alpha, 1/\beta)\beta}, \qquad \alpha\beta > 1,$$

4. the shannon entropy is given by

$$E[-\log(t)] = \left(1 + \frac{1}{\beta}\right)\log\left(\frac{2}{\beta}\right) + \log B\left(\alpha, \frac{1}{\beta}\right) + \log\sigma - \left(\alpha + \frac{1}{\beta}\right)\left[\varphi(\alpha) - \varphi(\alpha + 1/\beta)\right],$$

5. the Rényi entropy is given by

$$Y_R(\eta) = \frac{2B(1/\beta, \eta\alpha + (\eta - 1)/\beta)}{(1 - \eta)(2/\beta)^{\eta + (\eta-1)/\beta}[B(\alpha, 1/\beta)]^{\eta}\sigma^{\eta-1}\beta},$$

where $\varphi(\cdot)$ denote the digamma function.

3. Moments of Order Statistics

In this section, we derive the explicit expression for moments of order statistics from this distribution under the independent identically-distributed (IID) case and independent not identically-distributed (INID) case when the shape parameters $\frac{1}{\beta}$ and α are both real non-integer or $\frac{1}{\beta}$ is an integer.

It is well-known that order statistics has always been a significant field in statistical research, especially in the study of distribution theory. In addition, it is necessary to study the related properties of order statistics, since many classical parameter estimation methods are based on the application of order statistics.

3.1. The IID Case

In this subsection, we obtain explicit expression for moments of order statistics from $NGt(\mu, \sigma, \alpha, \beta)$ under IID case based on distribution transformation approach and recurrence relations when the shape parameter $\frac{1}{\beta}$ is either an integer or a real non-integer.

3.1.1. Shape Parameter $\frac{1}{\beta}$ Integer Case

Govindarajulu (1963) [16] pointed out that the k-th order moment of order statistics from a symmetric distribution can be linearly expressed by the moments of the order statistics from the corresponding folded distribution

$$2^n \delta_{r:n}^{(k)} = \sum_{m=0}^{r-1} \binom{n}{m} \gamma_{r-m:n-m}^{(k)} + (-1)^k \sum_{m=r}^{n} \binom{n}{m} \gamma_{m-r+1:m}^{(k)} \, ,$$

where $\delta_{r:n}^{(k)}$ denotes the k-th moments of order statistics from symmetric distribution, while $\gamma_{r:n}^{(k)}$ denotes the k-th moments of order statistics from the corresponding folded distribution.

Without loss of generality, let $\mu = 0$, $\sigma = 1$. Making transformation $\frac{|T|^\beta}{2} = Z$, we can derive the pdf of the random variable Z as follows

$$f_Z(z; \alpha, \beta) \;=\; \frac{1}{\left(\frac{1}{\beta}\right)^{\frac{1}{\beta}} B\left(\alpha, \frac{1}{\beta}\right)} z^{\frac{1}{\beta}-1} (1 + \beta z)^{-\left(\alpha + \frac{1}{\beta}\right)} \, , \qquad z > 0 \qquad (3)$$

Under the above transformation, the recurrence relations introduced by Govindarajulu (1963) [16] becomes

Lemma 1.

$$2^{n-\frac{k}{\beta}} \mu_{r:n}^{(k)} = \sum_{m=0}^{r-1} \binom{n}{m} \nu_{r-m:n-m}^{(k)} + (-1)^k \sum_{m=r}^{n} \binom{n}{m} \nu_{m-r+1:m}^{(k)} \, , \qquad (4)$$

where $\mu_{r:n}^{(k)}$ denotes the k-th moments of order statistics from $NGt(0, 1, \alpha, \beta)$ and $\nu_{r:n}^{(k)}$ denotes the k-th moments of order statistics from the corresponding folded distribution $F_Z(z)$.

Proof. See Appendix A.2 (Part II). $\square$

Let $\alpha_{r:n}^{(k)}$ denote $E(Z_{r:n}^k)$, then $\nu_{r:n}^{(k)} = \alpha_{r:n}^{(k/\beta)}$. The cdf of the random variable Z can be written as

$$\begin{aligned}
F_Z(z; \alpha, \beta) &\;=\; \frac{1}{\left(\frac{1}{\beta}\right)^{\frac{1}{\beta}} B\left(\alpha, \frac{1}{\beta}\right)} \int_0^z y^{\frac{1}{\beta}-1} (1 + \beta y)^{-\left(\alpha + \frac{1}{\beta}\right)} dy \\
&\;=\; \frac{1}{B\left(\alpha, \frac{1}{\beta}\right)} \int_0^{1 - \frac{1}{1+\beta z}} y^{\frac{1}{\beta}-1} (1 - y)^{\alpha - 1} dy \, ,
\end{aligned}$$

by binomial expansion we have

$$1 - F_Z(z) = \frac{1}{B\left(\alpha, \frac{1}{\beta}\right)} \sum_{i=0}^{\frac{1}{\beta}-1} C_i \frac{1}{(1 + \beta z)^{\alpha + i}} \, , \qquad (5)$$

where $C_i = (-1)^i \binom{1/\beta - 1}{i} \frac{1}{\alpha + i}$.

Let $H_v(\frac{1}{\beta}, p)$ denote the coefficient in the expansion of $\left[\sum\limits_{i=0}^{\frac{1}{\beta}-1} C_i \frac{1}{(1+\beta y)^i} \right]^p$. Obviously, the coefficients of this polynomial satisfy the following recursive relation

$$H_v\left(\frac{1}{\beta}, p\right) = \sum_{i=0}^{v} C_i H_{v-i}\left(\frac{1}{\beta}, p-1\right),$$

and with the following constraints

$$H_v\left(\frac{1}{\beta}, 1\right) = C_v, \qquad H_v\left(\frac{1}{\beta}, 0\right) = 1,$$

we can cauculate every coefficient striaghtly. Then $[1 - F_Z(z)]^p$ can be expressed as

$$[1 - F_Z(z)]^p = \frac{1}{\left[B\left(\alpha, \frac{1}{\beta}\right) \right]^p} \sum_{v=0}^{(\frac{1}{\beta}-1)p} H_v\left(\frac{1}{\beta}, p\right) \frac{1}{(1+\beta z)^{\alpha p + v}}. \tag{6}$$

It is noted that Equations (5) and (6) have similar forms, so the distribution $F_Z(z)$ is said to have a minimum domains of attraction. With this conclusion, $E(Z_{1:n}^k)$ is derived first. Through the traditional way, we obtain

$$
\begin{aligned}
E(Z_{1:n}^k) &= n \int_0^{+\infty} z^k [1 - F_Z(z)]^{n-1} f(z) \mathrm{d}z \\
&= \frac{n(\frac{1}{\beta})^k}{\left[B\left(\alpha, \frac{1}{\beta}\right) \right]^n} \sum_{v=0}^{(\frac{1}{\beta}-1)(n-1)} H_v\left(\frac{1}{\beta}, n-1\right) \times B\left(\frac{1}{\beta} + k, \alpha n + v - k\right),
\end{aligned}
$$

and the above expression exists for $k < \alpha n$.

After that, using the well-known recurrence relations from David and Nagaraja (2003) [17] presented below

$$(n-r)v_{r:n}^{(k)} + r v_{r+1:n}^{(k)} = n v_{r:n-1}^{(k)}, \tag{7}$$

where $r = 1, \cdots, n-1$, and $k = 1, 2, \cdots$.

We acquire the k-th order moment of order statistics from the distribution $F_Z(z; \alpha, \beta)$. When the sample size is n, the method of obtaining all the k-th moments of order statistics from $NGt(0, 1, \alpha, \beta)$ when $\frac{1}{\beta}$ is an integer can be briefly described as follows

1. Derive $v_{i,j}^{(k)}$ ($1 \le i \le n$, $1 \le j \le n$) by using the above approach.
2. For $1 \le r \le [\frac{n}{2}]$, derive $\mu_{r:n}^{(k)}$ by using the recurrence relations in Equation (4).
3. By using the symmetry of Gt distribution, the k-th moments of the remaining order statistics can be calculated through $\mu_{n-r+1:n}^{(k)} = -\mu_{r:n}^{(k)}$.

Since this method makes full use of the characteristics of the distribution function and recursive relations, it is very efficient and can avoid a lot of unnecessary calculations.

3.1.2. Shape Parameter $\frac{1}{\beta}$ or α Non-Integer Case

The generalized Kampe de Feriet function from Exton (1978) [18] is defined by

$$F_{C:D}^{A:B}\left((a):(b_1);\cdots;(b_n);(c):(d_1);\cdots,(d_n);x_1,\cdots,x_n\right) = \sum_{m_1=0}^{\infty}\cdots\sum_{m_n=0}^{\infty}\frac{((a))_{m_1+\cdots+m_n}((b_1))_{m_1}\cdots((b_n))_{m_n}}{((c))_{m_1+\cdots+m_n}((d_1))_{m_1}\cdots((d_n))_{m_n}}$$
$$\times\frac{x_1^{m_1}\cdots x_n^{m_n}}{m_1!\cdots m_n!},$$

where $a = (a_1, a_2, \cdots, a_A)$, $b_i = (b_{i,1}, b_{i,2}, \cdots, b_{i,B})$ for $i = 1, 2, \cdots, n$, $c = (c_1, \cdots, c_C)$, $d_i = (d_{i,1}, \cdots, d_{i,D})$, or $i = 1, 2, \cdots, n$ and $((f))_k = ((f_1, f_2, \cdots, f_p))_k = (f_1)_k \cdots (f_p)_k$, $(f_i)_k = f_i(f_i + 1) \cdots (f_i + k - 1)$.

By using the above special function, we derive the following theorem:

Theorem 2. *If the shape parameter α is a real non-integer, the k-th order moment of order statistics $X_{r:n}$ from $NGt(0, 1, \alpha, \beta)$ can be calculated by the following convergent expression*

$$E(X_{r:n}^k) = C_{r,n}\left[I_1(k, r-1, n-r) + I_2(k, r-1, n-r)\right],$$

where

$$I_1(k, r-1, n-r) = \frac{2^{\frac{k}{\beta}-n}(\frac{1}{\beta})^{\frac{k}{\beta}}}{B\left(\frac{1}{\beta}, \alpha\right)}\sum_{i=0}^{r-1}\sum_{j=0}^{n-r}(-1)^j\binom{r-1}{i}\binom{n-r}{j}A\left(\frac{k+1}{\beta}, \alpha - \frac{k}{\beta} - 1, i + j\right),$$

$$I_2(k, r-1, n-r) = (-1)^k\frac{2^{\frac{k}{\beta}-n}(\frac{1}{\beta})^{\frac{k}{\beta}}}{B\left(\frac{1}{\beta}, \alpha\right)}\sum_{i=0}^{n-r}\sum_{j=0}^{r-1}(-1)^j\binom{n-r}{i}\binom{r-1}{j}A\left(\frac{k+1}{\beta}, \alpha - \frac{k}{\beta} - 1, i + j\right),$$

$$A\left(\frac{k+1}{\beta}, \alpha - \frac{k}{\beta} - 1, i + j\right) = \frac{\beta^{i+j}B\left(\frac{k+i+j+1}{\beta}, \alpha - \frac{k}{\beta}\right)}{\left[B\left(\alpha, \frac{1}{\beta}\right)\right]^{i+j}} \times F_{1:1}^{1:2}\left(\left(\frac{k+i+j+1}{\beta}\right):(1-\alpha, \frac{1}{\beta});\cdots;\right.$$
$$\left.(1-\alpha, \frac{1}{\beta}):(\alpha + \frac{i+j+1}{\beta}):(\frac{1}{\beta}+1);\cdots;(\frac{1}{\beta}+1):1;\cdots;1)\right),$$

$$C_{r:n} = \frac{n!}{(r-1)!(n-r)!}.$$

Proof. See Appendix A.2 (Part II). □

Theorem 3. *If the shape parameter $\frac{1}{\beta}$ is a real non-integer, $A\left(\frac{k+1}{\beta}, \alpha - \frac{k}{\beta} - 1, i + j\right)$ in Theorem 2 can be calculated by*

$$A(\frac{k+1}{\beta}, \alpha - \frac{k}{\beta} - 1, i + j) = \frac{(\frac{1}{\beta})^{\frac{k+1}{\beta}}\alpha^{-(i+j)}}{\left[B\left(\alpha, \frac{1}{\beta}\right)\right]^{i+j}}B\left(\frac{k+1}{\beta}, \alpha(i+j+1) - \frac{k}{\beta}\right) \times F_{1:1}^{1:2}\left(\left(\alpha(i+j+1) - \frac{k}{\beta}\right):\right.$$
$$(1 - \frac{1}{\beta}, \alpha);\cdots;(1 - \frac{1}{\beta}, \alpha):(\alpha(i+j+1) + \frac{1}{\beta}):(\alpha + 1);\cdots;(\alpha + 1):$$
$$\left.1;\cdots1\right),$$

Proof. See Appendix A.2 (Part II). □

3.2. The INID Case

Suppose that $X_1, \cdots, X_n$ are independent but not identical random samples from the following pdf with the same shape parameter β but the rest of the parameters are different

$$g_i(x) = \frac{1}{\left(\frac{2}{\beta}\right)^{1+\frac{1}{\beta}} B\left(\alpha_i, \frac{1}{\beta}\right)\sigma_i} \left[1 + \frac{\beta|t - \mu_i|^\beta}{2\sigma_i^\beta}\right]^{-\left(\alpha_i + \frac{1}{\beta}\right)}.$$

Theorem 4. *If the shape parameter $\frac{1}{\beta}$ is a real non-integer, the m-th order moment of order statistics $X_{r:n}$ from $NGt(0, 1, \alpha_i, \beta)$ can be calculated by the following convergent expression*

$$\mu_{r:n}^m = \left[\sum_{j=n-r+1}^n (-1)^{j-(n-r+1)} \binom{j-1}{n-r} + \sum_{j=r}^n (-1)^{m+j-r} \binom{j-1}{r-1}\right] I_j^+(m),$$

where

$$
\begin{aligned}
I_j^+(m) &= \frac{m 2^{\frac{m}{\beta}-1-j}}{\prod\limits_{t=1}^{j} \alpha_{i_t} B(\alpha_{i_t}, \frac{1}{\beta})} \left(\frac{1}{\beta}\right)^{\frac{m}{\beta}-1} B\left(\frac{m}{\beta}, \sum_{t=1}^{j} \alpha_{i_t} - \frac{m}{\beta}\right) \times F_{1:1}^{1:2}\left((\sum_{t=1}^{j}\alpha_{i_t} - \frac{m}{\beta}) : (1-\frac{1}{\beta}, \alpha_{i_1}); \cdots \right. \\
&\left. ; (1-\frac{1}{\beta}, \alpha_{i_j}) : (\sum_{t=1}^{j}\alpha_{i_t}) : (\alpha_{i_1}+1); \cdots ; (\alpha_{i_j}+1) : 1; \cdots ; 1\right), \qquad \alpha_{i_t} > \frac{m}{\beta}.
\end{aligned}
\tag{8}
$$

Proof. See Appendix A.2 (Part II). □

4. Product Moments and Variance–Covariance Matrix of Order Statistics

In this section, we obtain explicit expressions for product moments and the corresponding variance–covariance matrix of order statistics from this distribution based on recurrence relations and distribution transformation when the shape parameter $\frac{1}{\beta}$ is an integer.

The covariance matrix of order statistics plays a big role in a number of estimation methods based on order statistics, such as the least square estimation and the BLUE for location-scale parameters. In addition, obtaining a covariance matrix of order statistics is usually not an easy task and requires some related skills.

Under transformation $\frac{|T|^\beta}{2} = Z$ introduced in the previous section, the recursive relations for the product moment of order statistics from a symmetric distribution between its corresponding folded distribution becomes

$$
\begin{aligned}
2^{n-\frac{k_1+k_2}{\beta}} \mu_{r,s:n}^{(k_1,k_2)} &= \sum_{m=0}^{r-1} \binom{n}{m} v_{r-m,s-m:n-m}^{(k_1,k_2)} + (-1)^{k_1} \sum_{m=r}^{s-1} \binom{n}{m} v_{m-r+1:m}^{(k_1)} v_{s-m:n-m}^{(k_2)} \\
&\quad + (-1)^{k_1+k_2} \sum_{m=s}^{n} \binom{n}{m} v_{m+1-s,m+1-r:m}^{(k_1,k_2)}.
\end{aligned}
\tag{9}
$$

The cdf of the random variable Z is denoted by $F_Z(z)$ and suppose $U, V \sim F_Z(z)$, then let $\alpha_{r,s:n}^{(k_1,k_2)}$ denote $E(U_{r:n}^{k_1} V_{s:n}^{k_2})$ $(r \le s)$. Obsiviously, we can write that $v_{r,s:n}^{(k_1,k_2)} = \alpha_{r,s:n}^{(k_1/\beta, k_2/\beta)}$.

First of all, we decide to derive $\alpha_{r,s:n}^{(k_1,k_2)}$. Taking the well-known expression of the product moment of order statistics and then using the binomial expansion, we obtain

$$
\begin{aligned}
\alpha_{r,r+1:s}^{(k_1,k_2)} &= C_{r,r+1:s} \iint\limits_{0<x<y<+\infty} x^{k_1} y^{k_2} [F(x)]^{r-1} [1-F(y)]^{s-r-1} f(x) f(y) \, dx \, dy \\
&= C_{r,r+1:s} \sum_{i=0}^{r-1} (-1)^i \binom{r-1}{i} \int_0^{+\infty} x^{k_1} [1-F(x)]^i f(x) \left(\int_x^{+\infty} y^{k_2} [1-F(y)]^{s-r-1} f(y) \, dy\right) dx \\
&= C_{r,r+1:s} \sum_{i=0}^{r-1} (-1)^i \binom{r-1}{i} I(k_1, k_2, i, s-r-1),
\end{aligned}
\tag{10}
$$

where

$$C_{r,s:n} = \frac{s!}{(r-1)!(s-r-1)!} \, .$$

After that, we attempt to derive the double integral $I(k_1, k_2, i, s-r-1)$. By the multinomial theorem we have

$$[1-F(y)]^{s-r-1} = \frac{1}{\left[B\left(\alpha, \frac{1}{\beta}\right)\right]^{s-r-1}} \sum_{v=0}^{(s-r-1)(\frac{1}{\beta}-1)} H_v\left(\frac{1}{\beta}, s-r-1\right) \times (1+\beta y)^{-[\alpha(s-r-1)+v]} \, .$$

We denote the inner integral $\int_x^{+\infty} y^{k_2} [1-F(y)]^{s-r-1} f(y) dy$ of the above double integral $I(k_1, k_2, i, s-r-1)$ by $T_x(k_2, s-r-1)$ and it can be calculated as follows

$$T_x(k_2, s-r-1) = \frac{\left(\frac{1}{\beta}\right)^{-\frac{1}{\beta}}}{\left[B\left(\alpha, \frac{1}{\beta}\right)\right]^{s-r}} \sum_{v=0}^{(s-r-1)(\frac{1}{\beta}-1)} H_v\left(\frac{1}{\beta}-1, s-r-1\right) \times \int_x^{+\infty} y^{k_2+\frac{1}{\beta}-1}(1+\beta y)^{-[\alpha(s-r)+v+\frac{1}{\beta}]} dy$$

$$= \frac{\left(\frac{1}{\beta}\right)^{k_2}}{\left[B\left(\alpha, \frac{1}{\beta}\right)\right]^{s-r}} \sum_{v=0}^{(s-r-1)(\frac{1}{\beta}-1)} H_v\left(\frac{1}{\beta}, s-r-1\right) \times \int_{\frac{\beta x}{1+\beta x}}^{1} w^{k_2+\frac{1}{\beta}-1}(1-w)^{\alpha(s-r)+v-k_2-1} dw \, .$$

Notice that $(k_2 + \frac{1}{\beta} - 1)$ is still a positive integer, then expanding the integrand on the right hand side of the above expression in terms of $(1-w)$, we acquire

$$\begin{aligned} T_x(k_2, s-r-1) &= \frac{\left(\frac{1}{\beta}\right)^{k_2}}{\left[B\left(\alpha, \frac{1}{\beta}\right)\right]^{s-r}} \sum_{v=0}^{(s-r-1)(\frac{1}{\beta}-1)} H_v\left(\frac{1}{\beta}, s-r-1\right) \sum_{m=0}^{k_2+\frac{1}{\beta}-1} (-1)^m \\ &\times \binom{k_2+\frac{1}{\beta}-1}{m} \frac{(1+\beta x)^{-[\alpha(s-r)+m+v-k_2]}}{\alpha(s-r)+m+v-k_2} \, . \end{aligned} \tag{11}$$

Using the fact that

$$[1-F(x)]^i = \frac{1}{\left[B\left(\alpha, \frac{1}{\beta}\right)\right]^i} \sum_{u=0}^{i(\frac{1}{\beta}-1)} H_u\left(\frac{1}{\beta}, i\right) \times (1+\beta y)^{-(\alpha i + u)} \, , \tag{12}$$

by putting Equations (11) and (12) into $I(k_1, k_2, i, s-r-1)$, we obtain

$$\begin{aligned} I(k_1, k_2, i, s-r-1) &= \frac{\left(\frac{1}{\beta}\right)^{k_1+k_2}}{\left[B\left(\alpha, \frac{1}{\beta}\right)\right]^{s-r+i+1}} \sum_{u=0}^{(\frac{1}{\beta}-1)i} \sum_{v=0}^{(s-r-1)(\frac{1}{\beta}-1)} H_u\left(\frac{1}{\beta}, i\right) H_v\left(\frac{1}{\beta}, s-r-1\right) \\ &\times \sum_{m=0}^{k_2+\frac{1}{\beta}-1} \binom{k_2+\frac{1}{\beta}-1}{m} \frac{(-1)^m}{\alpha(s-r)+m+v-k_2} \\ &\times B\left[k_1+\frac{1}{\beta}, \alpha(s-r+i+1)+m+u+v-k_1-k_2\right] \, . \end{aligned} \tag{13}$$

The above equation will exists for $\alpha > (k_1+k_2)/2$. Finally, using recurrence relations from Arnold et al. (2008) [19]:

For any arbitray distribution and $1 \le p < q \le n$

$$\alpha_{p,q:n}^{(k_1,k_2)} = \sum_{r=p}^{q-1} \sum_{s=n-q+r+1}^{n} (-1)^{s+n-p-q+1} \binom{r-1}{p-1} \binom{s-r-1}{n-q} \binom{n}{s} \alpha_{r,r+1:s}^{(k_1,k_2)} \, , \tag{14}$$

we can derive a closed-form expression for every product moment and the corresponding variance–covariance matrix of order statistics from $F_Z(z)$.

When the sample size is n, the method of obtaining all the product moments and variance–covariance matrix for order statistics from $NGt(\mu, \sigma, \alpha, \beta)$ can be briefly described as:

1. Derive $\nu_{i,j:n}^{(k_1,k_2)}$ $(1 \leq i \leq j \leq n)$ by using the above approach.
2. For $1 \leq r \leq [\frac{n}{2}]$, $r \leq s \leq n - r + 1$, obtain $\mu_{r,s:n}^{(k_1,k_2)}$ by using recurrence relations in Equation (9) and obtain the covariance $\sigma_{r,s:n}$ by $\sigma_{r,s:n} = \mu_{r,s:n} - \mu_{r:n}\mu_{s:n}$.
3. By using the symmetry of Gt distribution and the variance–covariance matrix, the remaining covariance of order statistics can be calculated by $\sigma_{r,s:n} = \sigma_{n-s+1,n-r+1:n}$ and $\sigma_{r,s:n} = \sigma_{s,r:n}$.

5. Parameter Estimation

In this section, we propose several parameter estimation methods for the new Gt distribution.

5.1. Modfied Method of Moments

In this subsection, a modified method of moments (MMOM) for the parameters of $NGt(0, \sigma, \alpha, \beta)$ is proposed.

Let $X_1, \ldots, X_n$ be a random sample from $NGt(\sigma, \alpha, \beta)$. Making transformation $Y = |X|^2$, the k-th theoretical central moments of Y are given by

$$E(Y^k) = \left(\frac{2}{\beta}\right)^{\frac{2k}{\beta}} \frac{\Gamma(\alpha - 2k/\beta)\Gamma[(2k+1)/\beta]}{\Gamma(\alpha)\Gamma(1/\beta)} \sigma^{2k}, \qquad k = 1, 2, \cdots,$$

The MMEs of the new distribution after transformation can be obtained by equating the first three theoretical central moments of Y with the corresponding central sample moments. Eliminating the scale parameter σ, we acquire

$$\frac{[\Gamma(\alpha)]\Gamma(\alpha - 4/\beta)}{[\Gamma(\alpha - 2/\beta)]^2} \times \frac{[\Gamma(1/\beta)]\Gamma(5/\beta)}{[\Gamma(3/\beta)]^2} = \frac{n\sum_{i=1}^n Y_i^2}{(\sum_{i=1}^n Y_i)^2}, \tag{15}$$

$$\frac{[\Gamma(\alpha)]^2\Gamma(\alpha - 6/\beta)}{[\Gamma(\alpha - 2/\beta)]^3} \times \frac{[\Gamma(1/\beta)]^2\Gamma(7/\beta)}{[\Gamma(3/\beta)]^3} = \frac{n^2\sum_{i=1}^n Y_i^3}{(\sum_{i=1}^n Y_i)^3}. \tag{16}$$

In order to establish an effective algorithm to obtain the numerical solution of the above equations, we propose the following theorem:

Theorem 5. *Let*

$$G_k(\alpha, \beta) = \frac{[\Gamma(\alpha)]^{k-1}\Gamma(\alpha - 2k/\beta)}{[\Gamma(\alpha - 2/\beta)]^k} \times \frac{[\Gamma(1/\beta)]^{k-1}\Gamma[(2k+1)/\beta]}{[\Gamma(3/\beta)]^k} \qquad k > 2,$$

then we assert that $G_k(\alpha, \beta)$ is a monotonically-decreasing function of the parameters α and β. In addition, the function has the following limiting properties

$$\text{For fixed } \alpha: \quad \lim_{\beta \longrightarrow 2k/\alpha^+} G_k(\alpha, \beta) = \infty, \qquad \lim_{\beta \longrightarrow +\infty} G_k(\alpha, \beta) = \frac{3^{k/2}}{\sqrt{2k+1}} \times \frac{\prod_{i=1}^k \Gamma[i/(2k+1)]}{[\Gamma(1/3)]^k[\Gamma(2/3)]^k}.$$

$$\text{For fixed } \beta: \quad \lim_{\alpha \longrightarrow 2k/\beta^+} G_k(\alpha, \beta) = \infty, \qquad \lim_{\alpha \longrightarrow +\infty} G_k(\alpha, \beta) = \frac{[\Gamma(1/\beta)]^{k-1}\Gamma[(2k+1)/\beta]}{[\Gamma(3/\beta)]^k}.$$

Proof. See Appendix A.3 (Part III). □

With the above theorem, the steps of the algorithm are presented by the following forms

1. Assign an initial values $\alpha_{(1)}$, and then substitute $\alpha_{(1)}$ into Equation (15). Since $G_k(\alpha, \beta)$ is monotonically-decreasing with respect to parameter β, *uniroot* in the R function can be used to derive $\hat{\beta}_{(1)}$.
2. Substitute $\hat{\beta}_{(1)}$ into Equation (16), since $G_k(\alpha, \beta)$ is monotonically-decreasing with respect to parameter α, using *uniroot* in R function then we obtain $\hat{\alpha}_{(2)}$.
3. If the norm of vector $(\hat{\alpha}_{(i+1)} - \hat{\alpha}_{(i)}, \hat{\beta}_{(i+1)} - \hat{\beta}_{(i)})$ exceed the prescribed value $\varepsilon > 0$, then 2–3 are repeated. Otherwise, the iteration is broken up and the values generated in the last iteration are determined to be estimates of the shape parameters α and β.

5.2. MLE Using the EM Algorithm

In this subsection, an EM-type algorithm to determine the MLEs for the parameters of $NGt(0, \sigma, \alpha, \beta)$ is established.

Through the method of distribution construction, the pdf of the new Gt distribution is treated as a marginal pdf of a bivariate distribution and this bivariate distribution function can be expressed as the product of two common distributions. After that, we treat one random variable in the bivariate distribution as an observed variable, while the other as an unobserved variable, so that the EM algorithm can be applied to obtain the MLE. At the same time, this idea is also widely used to derive the MLE of other distributions. See [15,20,21] for more details about recent works in this topic.

First proposed by Dempster (1977) [22], the EM algorithm is a powerful tool to deal with estimation problem under missing data situation. In this subsection, we are going to use the EM algorithm to acquire the MLEs of $NGt(\sigma, \alpha, \beta)$ based on Proposition 2. Let $Y = (y_1, \ldots, y_n)^T$ denote the observed data, $U = (u_1, \ldots, u_n)^T$ on behalf of the unobserved data. Combining Y and U together we derive the complete data denoted by $Z = (Y, U)$. Besides that, let $\theta = (\sigma, \alpha, \beta)^T$ represent the vector of parameters. Invoking Proposition 2, the complete log-likelihood function can be expressed as follows

$$
\begin{aligned}
\ln L(\theta|Z) &= n\left[\left(1 + \frac{1}{\beta}\right)\ln\beta - \left(1 + \frac{2}{\beta} + \alpha\right)\ln 2 - \ln\Gamma(\frac{1}{\beta}) - \ln\Gamma(\alpha) - \ln\sigma\right] \\
&+ \left(\alpha + \frac{1}{\beta} - 1\right)\sum_{i=1}^{n}\ln u_i - \frac{\beta}{4\sigma^\beta}\sum_{i=1}^{n}|y_i|^\beta u_i - \frac{1}{2}\sum_{i=1}^{n}u_i .
\end{aligned}
\tag{17}
$$

It is well-known that the EM framework is an iterative algorithm consisting of two steps; namely, the expectation step (E-step) and the maximization step (M-step).

E-step: Given the observed data set $Y = (y_1, \ldots, y_n)$ and parameter estimation values $\theta_{(h)}$ for the h-th iteration, the conditional expectation of the complete log-likelihood function can be calculated as follows

$$
\begin{aligned}
Q(\theta|\theta_{(h)}) = E[\ln L(\theta|Z)|Y, \theta_{(h)}] &= n\left[\left(1 + \frac{1}{\beta}\right)\ln\beta - \left(1 + \frac{2}{\beta} + \alpha\right)\ln 2 - \ln\Gamma(\frac{1}{\beta}) - \ln\Gamma(\alpha) - \ln\sigma\right] \\
&+ \left(\alpha + \frac{1}{\beta} - 1\right)\sum_{i=1}^{n}E[\ln u_i|y_i, \theta_{(h)}] - \frac{\beta}{4\sigma^\beta}\sum_{i=1}^{n}|y_i|^\beta E[u_i|y_i, \theta_{(h)}] \\
&- \frac{1}{2}\sum_{i=1}^{n}E[u_i|y_i, \theta_{(h)}] ,
\end{aligned}
$$

where

$$
E[u_i|y_i, \theta_{(h)}] = \frac{4\sigma_{(h)}^{\beta_{(h)}}(\alpha_{(h)} + 1/\beta_{(h)})}{2\sigma_{(h)}^{\beta_{(h)}} + \beta_{(h)}|y_i|^{\beta_{(h)}}} ,
$$

$$
E[\ln u_i|y_i, \theta_{(h)}] = \varphi\left(\alpha_{(h)} + \frac{1}{\beta_{(h)}}\right) - \ln\left(\frac{1}{2} + \frac{\beta_{(h)}|y_i|^{\beta_{(h)}}}{4\sigma_{(h)}^{\beta_{(h)}}}\right) .
$$

In the EM algorithm, the M-step needs to maximize the conditional expectation obtained by E-step and the suggested framework can be briefly described as follows

1. Update $\alpha_{(h+1)}$ through the following nonlinear equation and the corresponding numerical procedures is recommended for *uniroot* in *R* program.

$$\frac{\partial Q(\theta|\theta_{(h)})}{\partial \alpha} = n[\,\varphi(\alpha) + \ln 2\,] - \sum_{i=1}^{n} \mathrm{E}[\ln u_i|y_i, \theta_{(h)}]\,.$$

2. Update $\sigma_{(h+1)}$ through the following explicit expression

$$\sigma_{(h+1)} = \left(\frac{\beta_{(h)}^2}{4n}\right)^{\frac{1}{\beta_{(h)}}} \left(\sum_{i=1}^{n} |y_i|^{\beta_{(h)}} \mathrm{E}[u_i|y_i, \sigma_{(h)}, \alpha_{(h+1)}, \beta_{(h)}]\right)^{\frac{1}{\beta_{(h)}}}\,.$$

3. Update $\beta_{(h+1)}$ by maximizing the observed data log-likelihood function

$$\beta_{(h+1)} = \arg\max_{\beta} \sum_{i=1}^{n} Q(\theta|\sigma_{(h+1)}, \alpha_{(h+1)}, \beta\,)\,.$$

For the initial values $\theta_0 = (\sigma_0, \alpha_0, \beta_0)$, we recommend conducting numerical procedures such as *optim* in *R* program for maximizing the corresponding likelihood function or using the estimation values from the method of moments.

5.3. Explicit Expressions for Asymptotic Variances and Covariances of the Estimates under the EM Algorithm

In this subsection, we obtain closed-form expression for the asymptotic variances and covariances of the estimates under the EM algorithm. Louis (1982) [23] coined the missing information principle to acquire an information matrix in a situation containing latent variables. This principle is given by

Observed information = Complete information − Missing information .

In the following paper, we will denote the complete, observed and missing information by $I^Z(\theta)$, $I^Y(\theta)$ and $I^{(U|Y)}(\theta)$, respectively. Then the observed information $I^Y(\theta)$ is calculated by:

$$I^Y(\theta) = I^Z(\theta) - I^{(U|Y)}(\theta)\,.$$

5.3.1. Complete Information Matrix $I^Z(\theta)$

Taking the second-order partial derivatives of the complete log-likelihood function in Equation (17) and applying the definition of the fisher information, then after some calculation, we obtain

$$I_{11}^Z = \frac{n\beta}{\sigma^2}\,, \qquad I_{22}^Z = n\varphi'(\alpha)\,, \qquad I_{12}^Z = I_{21}^Z = I_{23}^Z = I_{32}^Z = 0\,,$$

$$I_{13}^Z = I_{31}^Z = -\frac{n}{\sigma\beta} + \frac{\beta^2}{4\sigma}nT_1(\alpha, \beta) + \frac{n\beta^{\frac{1}{\beta}}}{2^{\frac{1}{\beta}}\mathrm{B}(\alpha, 1/\beta)}\,,$$

$$I_{33}^Z = \frac{2n}{\beta^3}\left[-\ln\beta + (1+\beta) + 2\ln 2 + \varphi(\tfrac{1}{\beta})\right] + \frac{n}{\beta^2}\left[\frac{1}{\beta} - 1 + \frac{\varphi(\tfrac{1}{\beta})}{\beta^2}\right] - \frac{2}{\beta^3}n\left[\varphi(\alpha) + \frac{\ln 2}{2^\alpha}\right]$$
$$+ \frac{n}{2}T_1(\alpha, \beta) + \frac{\beta n}{4}T_2(\alpha, \beta)\,,$$

where

$$T_1(\alpha, \beta) = \frac{4}{\beta^3}\left[\varphi(1 + \tfrac{1}{\beta}) - \ln\frac{\beta}{2} - \varphi(\alpha)\right]\,,$$

$$T_2(\alpha,\beta) = \frac{2}{\beta^3}\left\{\left[\varphi\left(1+\frac{1}{\beta}\right)\right]^2 + \frac{1}{\beta^2}\varphi'\left(1+\frac{1}{\beta}\right)\right\} + \frac{8}{\beta^3}\left[\varphi\left(1+\frac{1}{\beta}\right)\ln\frac{\beta}{2} + \varphi\left(1+\frac{1}{\beta}\right)\varphi(\alpha)\right]\right\}$$
$$+ \frac{4}{\beta^3}\left\{(\frac{\beta}{2})^2 + [\varphi(\alpha)]^2 + \varphi'(\alpha) + 2\varphi(\alpha)\ln\frac{\beta}{2}\right\}.$$

The second-order partial derivatives of the complete likelihood function and the calculation method of the complete information are shown in the Appendix A.3 (Part III).

5.3.2. Missing Information Matrix $I^{(U|Y)}(\theta)$

From Proposition 2, the logarithm of the joint density function of latent variable U given Y can be written as

$$\ln f_{U|Y}(u|y) = \sum_{i=1}^{n}\left(\alpha+\frac{1}{\beta}\right)\ln\left(\frac{1}{2}+\frac{\beta|y_i|}{4\sigma^\beta}\right) - n\Gamma(\alpha+1/\beta) + \left(\alpha+\frac{1}{\beta}-1\right)\sum_{i=1}^{n}\ln U_i - \sum_{i=1}^{n}\left(\frac{1}{2}+\frac{\beta|y_i|}{4\sigma^\beta}\right)U_i.$$

After taking the second-order partial derivatives with respect to σ, α and β, and utilizing the results in Section 5.2, the elements of the missing information matrix $I^{(U|Y)}(\theta)$ are presented as follows

$$I_{11}^{(U|Y)} = -\beta(\alpha+1)\sum_{i=1}^{n}\left\{\frac{|y_i|[2(\beta+1)\sigma^\beta+\beta|y_i|^2]}{2\sigma^{\beta+1}+\beta\sigma|y_i|^2} - \frac{|y_i|(\beta+1)}{\sigma^2(2\sigma^\beta+\beta|y_i|^\beta)}\right\},$$

$$I_{12}^{(U|Y)} = I_{21}^{(U|Y)} = -\sum_{i=1}^{n}\frac{\beta^2|y_i|}{2\sigma^{\beta+1}+\beta\sigma|y_i|},$$

$$I_{13}^{(U|Y)} = I_{31}^{(U|Y)} = -\sum_{i=1}^{n}\frac{|y_i|}{2\sigma^{\beta+1}+\beta\sigma|y_i|} + (\alpha+\frac{1}{\beta})\sum_{i=1}^{n}\frac{2\beta|y_i|(2\sigma^{\beta+1}+\beta\sigma|y_i|) - \beta^2|y_i|(2\sigma^{\beta+1}\ln\sigma+\sigma|y_i|)}{(2\sigma^{\beta+1}+\beta\sigma|y_i|)^2}$$
$$+ \frac{2\beta|y_i|-\beta^2\sigma|y_i|\ln\sigma}{4\sigma^{\beta+1}},$$

$$I_{22}^{(U|Y)} = n\frac{\partial^2\Gamma(\alpha+1/\beta)}{\partial\alpha^2},$$

$$I_{23}^{(U|Y)} = I_{32}^{(U|Y)} = -\frac{|y_i|(1-\beta\ln\sigma)}{2\sigma^\beta+\beta|y_i|} + \frac{n}{\beta^4}\varphi'(\alpha+1/\beta),$$

$$I_{33}^{(U|Y)} = -\frac{2}{\beta^2}\sum_{i=1}^{n}\ln\left(\frac{1}{2}+\frac{\beta|y_i|}{4\sigma^\beta}\right) + \sum_{i=1}^{n}\frac{|y_i|}{2\sigma^{\beta+1}+\beta\sigma|y_i|} + \frac{1}{\beta^2}\sum_{i=1}^{n}\frac{|y_i|(1-\beta\ln\sigma)}{2\sigma^\beta+\beta|y_i|} + \left(\frac{2n}{\beta^3}+\frac{n}{\beta^4}\right)\frac{\partial\Gamma(\alpha+1/\beta)}{\partial\beta}$$
$$+ \left(\alpha+\frac{1}{\beta}\right)\sum_{i=1}^{n}\left[\frac{|y_i|[(2\sigma^\beta+\beta|y_i|)\ln\sigma+(2\sigma^\beta\ln\sigma+|y_i|)|y_i|(1-\beta\ln\sigma)]}{(2\sigma^\beta+\beta|y_i|)^2} + \frac{|y_i|(2-\beta\ln\sigma)}{2\sigma^\beta+\beta|y_i|^\beta}\right]$$
$$- \frac{2n}{\beta^3}\left[\varphi\left(\alpha+\frac{1}{\beta}\right) - \ln\left(\frac{1}{2}+\frac{\beta|y_i|}{4\sigma^\beta}\right)\right].$$

5.4. MLE via a New Iterative Algorithm

In this subsection, we propose a new iterative algorithm to obtain the MLEs for the parameters of $NGt(0,\sigma,\alpha,\beta)$ because of the difficulty of sloving the estimation equations directly.

Let $X_1,\cdots,X_n$ be a random sample from $NGt(0,\sigma,\alpha,\beta)$, and the log-likelihood function is given by

$$\ln L(X;\sigma,\alpha,\beta) = -n(1+1/\beta)\ln(2/\beta) - n\ln B(\alpha,1/\beta) - n\ln\sigma - \left(\alpha+\frac{1}{\beta}\right)\sum_{i=1}^{n}\ln\left[1+\frac{\beta|x_i|^\beta}{2\sigma^\beta}\right].$$

With partial derivatives respect to (σ,α,σ), we obtain

$$
\begin{cases}
\dfrac{\partial \ln L(X;\sigma,\alpha,\beta)}{\partial \sigma} &= n - (\alpha\beta+1)\displaystyle\sum_{i=1}^{n} \dfrac{\beta|x_i|^\beta}{2\sigma^\beta + \beta|x_i|^\beta} \,, \\[4mm]
\dfrac{\partial \ln L(X;\sigma,\alpha,\beta)}{\partial \alpha} &= n\left[\varphi(\alpha+1/\beta) - \varphi(\alpha)\right] - \displaystyle\sum_{i=1}^{n} \ln\left[1 + \dfrac{\beta|x_i|^\beta}{2\sigma^\beta}\right], \\[4mm]
\dfrac{\partial \ln L(X;\sigma,\alpha,\beta)}{\partial \beta} &= n\ln(2/\beta) + n\beta(1+1/\beta) - n\left[\varphi(\alpha+1/\beta) - \varphi(1/\beta)\right] + \displaystyle\sum_{i=1}^{n}\ln\left[1 + \dfrac{\beta|x_i|^\beta}{2\sigma^\beta}\right] \\[4mm]
&\quad - \beta(\alpha\beta+1)\displaystyle\sum_{i=1}^{n}\dfrac{\left(|x_i|^\beta + \beta|x_i|^\beta\ln(|x_i|/\sigma)\right)}{2\sigma^\beta + \beta|x_i|^\beta} \,.
\end{cases}
\tag{18}
$$

Instead of sloving all three preceding equations simultaneously, we propose a new algorithm because of the complexity of the third equation of Equation (18). In order to establish the new approach, the following lemma is necessary

Lemma 2. *For any fixed β, the following equations*

$$
\begin{cases}
n - (\alpha\beta+1)\displaystyle\sum_{i=1}^{n}\dfrac{\beta|x_i|^\beta}{2\sigma^\beta + \beta|x_i|^\beta} = 0 \,, \\[4mm]
n\left[\varphi(\alpha+1/\beta) - \varphi(\alpha)\right] - \displaystyle\sum_{i=1}^{n}\ln\left[1 + \dfrac{\beta|x_i|^\beta}{2\sigma^\beta}\right] = 0 \,,
\end{cases}
\tag{19}
$$

always possess a convergent solution of parameter (σ,α) and is independent of the initial values (σ_0,α_0).

Proof. See Appendix A.3 (Part III). □

Moreover, the algorithm for obtaining the solution of the equation in Lemma 2 is shown below

1. Given arbitrary initial values $(\sigma_{(0)},\alpha_{(0)})$ and $\sigma_{(0)} > 0$, $\alpha_{(0)} > 0$.
2. Substitute $\alpha_{(0)}$ into the first equation of Equation (19), and we obtain $\sigma_{(1)}$.
3. Substitute $\sigma_{(1)}$ into the second equation of Equation (19), and we obtain $\alpha_{(1)}$.
4. If the norm of vector $(\sigma_{(h+1)} - \sigma_{(h)}, \alpha_{(h+1)} - \alpha_{(h)})$ exceed the prescribed value $\varepsilon > 0$, then 2–3 are repeated. Otherwise, the iteration is broken up and the values generated in the last iteration are determined to be solution of this equation.

In what follows, the new algorithm to acquire the solution of Equation (18) can be briefly described as

1. Give an appropriate interval for the shape parameter β, and then generate a set of values of this parameter at an appropriate distance within the interval and represent the j-th element of this vector by $\beta^{(j)}$.
2. Assigning $\beta^{(j)}$ into the first two equations of Equation (18) and applying Lemma 2, we can obtain their convergence solution denoted by $(\sigma^{(j)}, \alpha^{(j)})$.
3. Repeat this procedure and observe the sign-changing in the left-hand side of the third equation of Equation (18). Then, by interpolating on β values, we derive the final estimate $\hat{\beta}_{MLE}$, which makes the third equation equals to 0 or sufficiently close to 0. Finally, by substituting $\hat{\beta}_{MLE}$ into the first two equations of Equation (18), we derive the final estimates denoted by $(\hat{\sigma}_{MLE}, \hat{\alpha}_{MLE})$.

It is noteworthy that this algorithm is applicable when the sample kurtosis is greater than 2.7. Comparing with other numerical methods, our new algorithm requires less precision of initial values (σ_0,α_0) and can obtain convergent solutions.

5.5. An Improved PWM Estimation

In this subsection, we propose an improved PWM (IPWM) estimation based on order statistics transformation for $NGt(0, \sigma, \alpha, \beta)$ and develop the corresponding algorithm to implement this novel method.

PWM estimation is a very effective and commonly used method for parameter estimation of distributions with heavy tails. Among many parameter estimation methods proposed so far, PWM and MLE are the two most commonly-used methods that have similar priority.

Let $X_1, \cdots, X_n$ be a random sample from $NGt(0, \sigma, \alpha, \beta)$ and $X_{1:n} \leq \cdots \leq X_{n:n}$ behalf of the corresponding order statistics. The traditional theoretical PWM was introduced by Greenwood et al. [24] as an alternative to the method of moments given in the following equation

$$M_{p,r,s}(x) = \mathrm{E}\{X^p [G(x)]^r [1 - G(x)]^s\}, \quad r = 0, 1, \cdots, \quad s = 0, 1, \ldots.$$

Taking $p = 1$, $s = 0$, the moments become $M_{1,r,0}(x)$ denoted by β_r. Wang (1990) [25] obtained an unbiased estimator of β_r through linear combination of order statistics as follows

$$b_r = \frac{1}{n} \binom{n-1}{r}^{-1} \sum_{i=1}^{n} \binom{i-1}{r} x_{i:n}, \quad r = 0, 1, \cdots.$$

However, applying the traditional PWM directly can cause many problems due to the complexity of the distribution function. For example, intricacy expressions of the theoretical PWM will make the estimation equation tedious and difficult to be numerically optimized, thus impeding the use of this estimation method. As a result, we decide to make isotone transformation $Y = \sigma X$. After this monotone transformation, the estimation equation becomes simple and easy to solve by a numerical approach. The improved theoretical PWM (denoted by $M_{1,r,0}^Y$) under this transformation is given by

$$\beta_r^Y = M_{1,r,0}^Y = \mathrm{E}\{Y [G(y)]^r\}, \quad r = 0, 1, \cdots. \tag{20}$$

The corresponding improved sample PWM (denoted by b_r^Y) is defined as follows

$$b_r^Y = \frac{1}{n} \binom{n-1}{r}^{-1} \sum_{i=1}^{n} \binom{i-1}{r} y_{i:n}, \quad r = 0, 1, \cdots, \tag{21}$$

where $y_{i:n} = \tilde{\sigma} x_{i:n}$ and $\tilde{\sigma}$ is an initial guess of the scale parameter σ.

From Equation (20), we can derive the 2–5th improved theoretical PWM as follows

$$\beta_1^Y = \frac{\left(\frac{2}{\beta}\right)^{\frac{1}{\beta}}}{2\mathrm{B}(\alpha, \frac{1}{\beta})} \int_0^1 z^{\frac{2}{\beta}-1} (1-z)^{\alpha-\frac{1}{\beta}-1} I_z\left(\frac{1}{\beta}, \alpha\right) \mathrm{d}z,$$

$$\beta_2^Y = \beta_1^Y,$$

$$\beta_3^Y = \frac{2^{1/\beta-3}\left(\frac{1}{\beta}\right)^{\frac{1}{\beta}}}{\mathrm{B}(\alpha, \frac{1}{\beta})} \left\{ \int_0^1 z^{\frac{2}{\beta}-1} (1-z)^{\alpha-\frac{1}{\beta}-1} I_z\left(\frac{1}{\beta}, \alpha\right) \mathrm{d}z + 3 \int_0^1 z^{\frac{2}{\beta}-1} (1-z)^{\alpha-\frac{1}{\beta}-1} \left[I_z\left(\frac{1}{\beta}, \alpha\right) \right]^3 \mathrm{d}z \right\},$$

$$\beta_4^Y = \frac{2^{1/\beta-2}\left(\frac{1}{\beta}\right)^{\frac{1}{\beta}}}{\mathrm{B}(\alpha, \frac{1}{\beta})} \left\{ \int_0^1 z^{\frac{2}{\beta}-1} (1-z)^{\alpha-\frac{1}{\beta}-1} I_z\left(\frac{1}{\beta}, \alpha\right) \mathrm{d}z + \int_0^1 z^{\frac{2}{\beta}-1} (1-z)^{\alpha-\frac{1}{\beta}-1} \left[I_z\left(\frac{1}{\beta}, \alpha\right) \right]^3 \mathrm{d}z \right\}.$$

According to the traditional idea of the method of moments, let the improved theoretical PWM β_3^Y, β_4^Y equal to the corresponding improved sample PWM. Then we can obtain an equation set whose solution is the estimation of the unknown parameters. Due to the complexity of these equations, we transform the search for the solution of this equation

set into an optimization problem and acquire the simultaneous estimation of the shape parameters denoted by $\hat{\alpha}_{PWM}$, $\hat{\beta}_{PWM}$ through minimizing the following expression to avoid failures during optimization

$$(\hat{\alpha}_{PWM}, \hat{\beta}_{PWM}) = \underset{(\alpha, \beta)}{\arg\min} \left(\beta_3^Y + \beta_4^Y - b_3^Y - b_4^Y \right)^2 . \tag{22}$$

Numerical techniques are recommended, such as the *optim* function in *R* program and the method will be chosen for $L - BFGS - B$.

Finally, the estimation of the scale parameter $\hat{\sigma}_{PWM}$ can be obtained through the first original sample PWM of the random variable X with analytical expression given by the following equation

$$\hat{\sigma}_{PWM} = \frac{2b_1 B(\hat{\alpha}_{PWM}, 1/\hat{\beta}_{PWM})}{(\frac{2}{\hat{\beta}_{PWM}})^{1/\hat{\beta}_{PWM}} \int_0^1 u^{2/\hat{\beta}_{PWM}-1} (1-u)^{\hat{\alpha}_{PWM}-\hat{\beta}_{PWM}-1} I_u(1/\hat{\beta}_{PWM}, \hat{\alpha}_{PWM}) \mathrm{d}u} . \tag{23}$$

In the study, we found that the estimation results are sensitive to the choice of the initial guess $\tilde{\sigma}$. Therefore, we introduce the following two methods to discuss the choice of this parameter.

Method1 (A two step estimation method)

1. Take a consistent estimate of parameter σ as the initial guess of it, such as MLE.
2. Substitute this estimate value $\hat{\sigma}$ into Equation (22) then we derive the improved PWM-estimation of the two shape parameters. The estimation for the scale paramter σ can be obtained through Equation (23).

Through this method, considering that it is relatively fixed to select the value of parameter $\tilde{\sigma}$, and the final estimations are largely dependent on this parameter, sometimes it is inevitable that the fixed parameter $\tilde{\sigma}$ does not apply to all samples. As a result, we establish the following algorithm, which can flexibly select the parameter $\tilde{\sigma}$ and the suggested framework can be briefly described as:

Method2 (A selection algorithm)

1. Give an appropriate interval for the parameter $\tilde{\sigma}$, generate a set of $\tilde{\sigma}$ at appropriate distance within the interval and represent the h-th element of the vector by $\tilde{\sigma}_{(h)}$.
2. Substitute each $\tilde{\sigma}_{(h)}$ into Equation (22) to acquire simultaneously estimation for two shape parameters denoted by $\hat{\alpha}_{PWM}^{(h)}$ and $\hat{\beta}_{PWM}^{(h)}$. After that, cauculate the corresponding original 6th theroetical PWM-kuritosis through the following equation

$$T_k(\hat{\alpha}_{PWM}^{(h)}, \hat{\beta}_{PWM}^{(h)}) = \int_0^1 u^{\frac{2}{\hat{\beta}_{PWM}^{(h)}}-1} (1-u)^{\hat{\alpha}_{PWM}^{(h)} - \frac{1}{\hat{\beta}_{PWM}^{(h)}}-1} \left[I_u \left(\frac{1}{\hat{\beta}_{PWM}^{(h)}}, \hat{\alpha}_{PWM}^{(h)} \right) \right]^k \mathrm{d}u ,$$

$$\tau_6^{(h)PWM} = \frac{T_1(\hat{\alpha}_{PWM}^{(h)}, \hat{\beta}_{PWM}^{(h)}) + 2T_3(\hat{\alpha}_{PWM}^{(h)}, \hat{\beta}_{PWM}^{(h)}) + T_5(\hat{\alpha}_{PWM}^{(h)}, \hat{\beta}_{PWM}^{(h)})}{T_1(\hat{\alpha}_{PWM}^{(h)}, \hat{\beta}_{PWM}^{(h)})} .$$

3. Repeat 2 and cauculate $q^{PWM}(\tilde{\sigma})_{(h)} = |\tau_6^{(h)PWM} - \hat{M}_{1,0,5}^X / \hat{M}_{1,0,1}^X|$ ($\hat{M}_{1,0,5}^X$ and $\hat{M}_{1,0,1}^X$ denote the original 6th and 2th sample PWM of the random variable X, respectively). Then, interpolate on $\tilde{\sigma}$ values and derive $\check{\sigma}$ that minimize $q^{PWM}(\tilde{\sigma})$.
4. Substituting $\check{\sigma}$ into Equation (22) we derive the final estimation for two shape parameters and final estimation for σ can be obtained form Equation (23).

5.6. Parameter Estimation of $NGt(\mu, \sigma, \alpha, \beta)$ $(\mu \neq 0)$

In this subsection, we propose a profile maximum likelihood approach (PLA) using the EM algorithm to cope with parameter estimation for the new Gt distribution with four parameters.

In fitting problems with the distribution of data with better symmetry, three factors are usually comprehensively scrutinized, namely location information, scale parameters, and shape parameters. Therefore, with the addition of the location parameter, we focus on parameter estimation problem of $NGt(\mu, \sigma, \alpha, \beta)$.

Balakrishnan (2019) [20] applied this approach to determine the MLEs for the parameters of Student's t Birnbaum-Saunders distribution. As Balakrishnan points out, this method is an example of an exhaustive search. Due to the addition of location parameter μ, the estimation equation becomes very complicated. To overcome this difficulty, μ can be assumed as a natural number through profile likelihood approach, reducing the dimension of the problem with four parameters to that with three. Then the EM algorithm for the three-parameter case described in the preceding chapter comes in handy here.

Let $Y_1, \ldots, Y_n$ be a random sample from $NGt(\mu, \sigma, \alpha, \beta)$. In this case, the log-likelihood function is given by

$$\ln L(y|\mu, \sigma, \alpha, \beta) = -n(1 + 1/\beta) \ln(2/\beta) - n \ln B(\alpha, 1/\beta) - n \ln \sigma - \left(\alpha + \frac{1}{\beta}\right) \sum_{i=1}^{n} \ln \left[1 + \frac{\beta|y_i - \mu|^\beta}{2\sigma^\beta}\right].$$

Since $\frac{\beta|y_i - \mu|^\beta}{2\sigma^\beta + \beta|y_i - \mu|^\beta} < \ln(1 + y_i) < \frac{\beta|y_i - \mu|^\beta}{2\sigma^\beta}$, we can assert that the likelihood function $L(\mu)$ is bounded when the parameter (σ, α, β) are fixed.

A suggested framework for the PLA using the EM algorithm is briefly described as follows:

1. Give an appropriate interval for the location parameter μ. Then generate a set of μ at appropriate distance within the interval and represent the j-th element of the vector by $\mu^{(j)}$.
2. Substitute each $\mu^{(j)}$ in the log-likelihood function and implement the EM algorithm to derive the MLEs of (σ, α, β).
3. Repeating this procedure, and then interpolating on μ values, we obtain a curve of values of the log-likelihood function and choose the value of μ that maximizes the curve as $\hat{\mu}_{MLE}$.
4. Substitute $\hat{\mu}_{MLE}$ in the log-likelihood function and implement the EM algorithm to determine the final MLEs denoted by $(\hat{\sigma}_{MLE}, \hat{\alpha}_{MLE}, \hat{\beta}_{MLE})$.

The implementation for the EM algorithm is similar to Section 5.2, just replace all the $|y_i|$ in the partial derivative of σ, α, β and the integral with $|y_i - \hat{\mu}_{MLE}|$. The trace of the log-likelihood function is depicted in Figure 5. Through simulation research, we find that this algorithm performs well for $\alpha > \beta$ and is more accurate and stable than using the sample mean as the estimation of the location parameter μ directly under a small sample size.
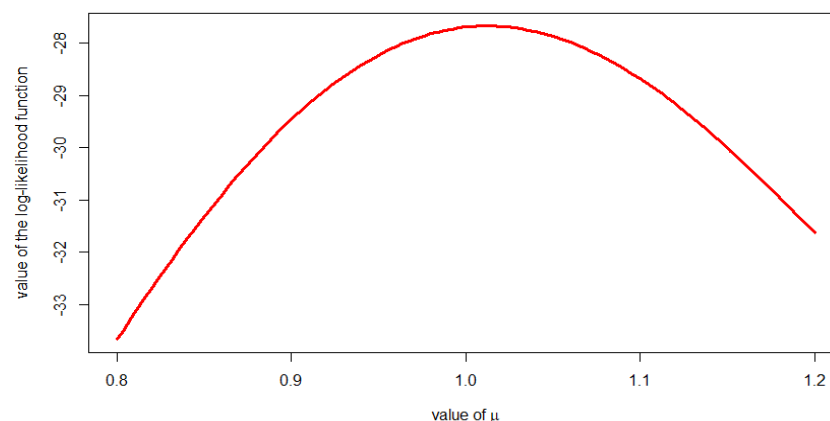

Figure 5. Curve of values of the log-likelihood function under sample size $n = 50$.

6. Simulation Study

In this section, we conduct a simulation study to illustrate the performance of the proposed estimation methods and compare these simulation results under different parameter values and different sample sizes.

According to the actual demand, the sample size will be chosen for 25, 50, 100 and 200. We perform Monte Carlo simulations and evaluate the performance of the considered methods through relative bias (RB) and relative mean-square errors (RMSE). These two indicators will be computed with 5000 replications at the same time and they are defined as follows

$$\text{RB}(\hat{\theta}) = \frac{1}{N} \sum_{i=1}^{N} \frac{\hat{\theta}_i}{\theta_i},$$

$$\text{RMSE}(\hat{\theta}) = \frac{1}{N} \sum_{i=1}^{N} \left[\frac{\hat{\theta}_i}{\theta_i} - \text{RB}(\hat{\theta}) \right].$$

This section will be detailed in three parts. First of all, we conduct a comparative study for three different estimation methods (MMOM, MLE using the EM algorithm and IPWM) of $NGt(0, \sigma, \alpha, \beta)$. We mainly study the influence of shape parameter β on the performance of these methods. As a result, we fix $\sigma = 1$, $\alpha = 3$, and β will be chosen for $2, 3, 4$, respectively. To show the comparison more vividly, we draw the following pictures (Figures 6–8) and the estimated values of these three parameters are shown in Appendix A.4 (Part IV).

It can be seen that the two indicators RB and RMSE of these three estimation methods gradually decrease with the increase of sample size. As the value of the shape parameter, β, increases, the estimation accuracy of the IPWM and the MLE using the EM algorithm for shape parameter α gradually becomes better. Meanwhile, we can conclude that the IPWM performs best for both accuracy and stability, especially for the distribution has a heavier tail, while the MLE using the EM algorithm is the second on the whole. The performance of these two methods is comparable under a large sample size or high theoretical kurtosis. Furthermore, the IPWM is suitable for distributions with heavy tails and the MLE using the EM algorithm can become an alternative estimation method for steep-shaped distributions or under large sample size.

Second, a comparative study of the MLE using the EM algorithm and the MLE via a new iterative algorithm with sample kurtosis more than 2.7 is considered. The numerical results are presented in Table 1 for fixed $\mu = 0$, $\sigma = 1$, $\alpha = 3$, $\beta = 3$.

Table 1. RB, RMSE and estimated values (in parentheses) of (σ, α, β).

Sample Size	MLE Using the EM Algorithm			MLE via a New Iterative Algorithm		
	RB			RB		
	σ	α	β	σ	α	β
$n = 25$	0.8788 (0.879)	1.3800 (4.14)	1.2524 (3.757)	0.7235 (0.724)	0.7337 (2.201)	1.1098 (3.323)
$n = 50$	0.9371 (0.937)	1.1867 (3.56)	1.2078 (3.624)	0.8101 (0.81)	0.8392 (2.518)	1.0586 (3.176)
$n = 100$	0.9482 (0.948)	1.1041 (3.312)	1.1861 (3.558)	0.9144 (0.914)	0.9572 (2.872)	0.9989 (2.997)
$n = 200$	0.9678 (0.968)	0.9555 (2.867)	1.1188 (3.356)	0.9373 (0.937)	1.0052 (3.016)	0.9837(2.951)
	RMSE			RMSE		
	σ	α	β	σ	α	β
$n = 25$	0.2348	1.2119	0.6866	0.2199	0.9791	0.1478
$n = 50$	0.1632	0.9105	0.8034	0.1783	0.8141	0.09916
$n = 100$	0.0874	0.7222	0.7311	0.1269	0.6223	0.0509
$n = 200$	0.0401	0.2789	0.2175	0.0753	0.3971	0.0211

Table 1 suggests that the new iterative algorithm performs better than the EM algorithm under sample kurtosis that exceeds 2.7 in general, especially for the estimation accuracy and stability of shape parameter β. For shape parameter α, the iterative algorithm provides higher estimation accuracy while the EM algorithm is more stable.

At last, a simulation study for the PLA using the EM algorithm for $NGt(\mu, \sigma, \alpha, \beta)$ ($\mu \neq 0$) considered in Section 5.6 will be performed. We mainly focus on the influence of shape parameter α and sample size on the estimation performance of this method. Therefore, we fix $\mu = 1$, $\sigma = 1$, $\beta = 2$ and α will be chosen for $2, 2.5, 3$, respectively. Line graphs of the relationship between RB, RMSE and sample sizes with different value of shape parameter α were shown below (Figure 9) and the estimated value for these parameters are given in Table A4.

Figure 9 reveals that as the value of the shape parameter α increases, the estimation results become better and better. It is concluded that with the increase of the theoretical kurtosis, the accuracy and stability of the MLEs are gradually improved. In addition, the estimation effect of location parameter, μ, and shape parameter, β, is less affected by the change of shape parameter α, while that two indicators for the scale parameter, σ, and shape parameter, α, are more sensitive to the change of it. This means that it is not very difficult to estimate the location parameter for a distribution with good symmetry, while the shape parameters of a heavy tail distribution are more difficult to estimate.
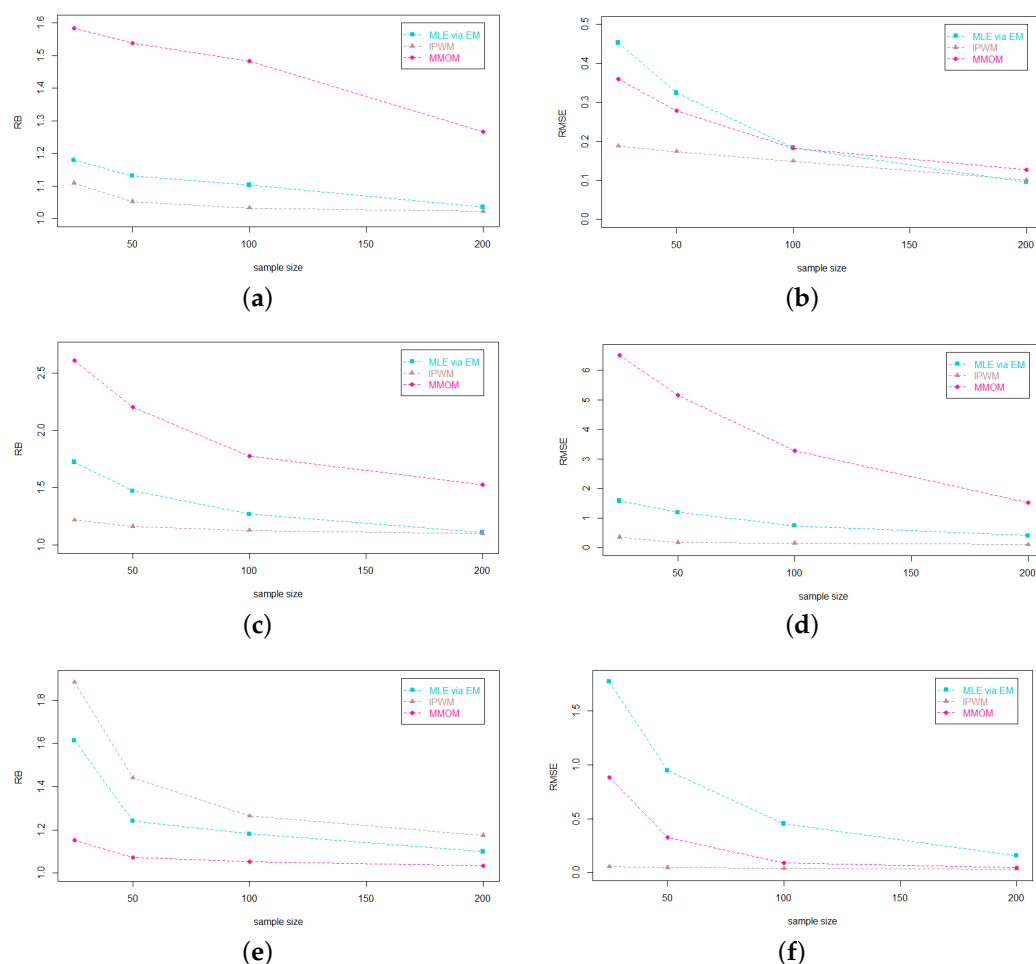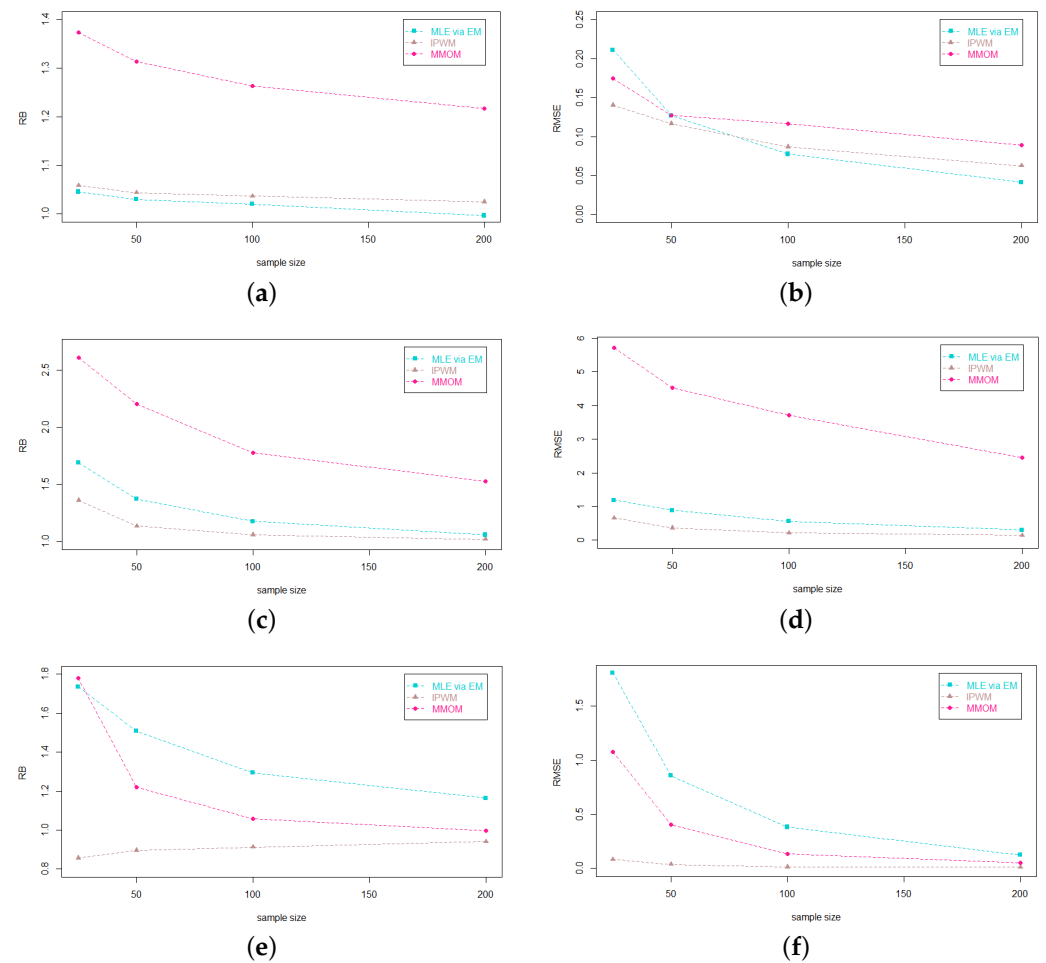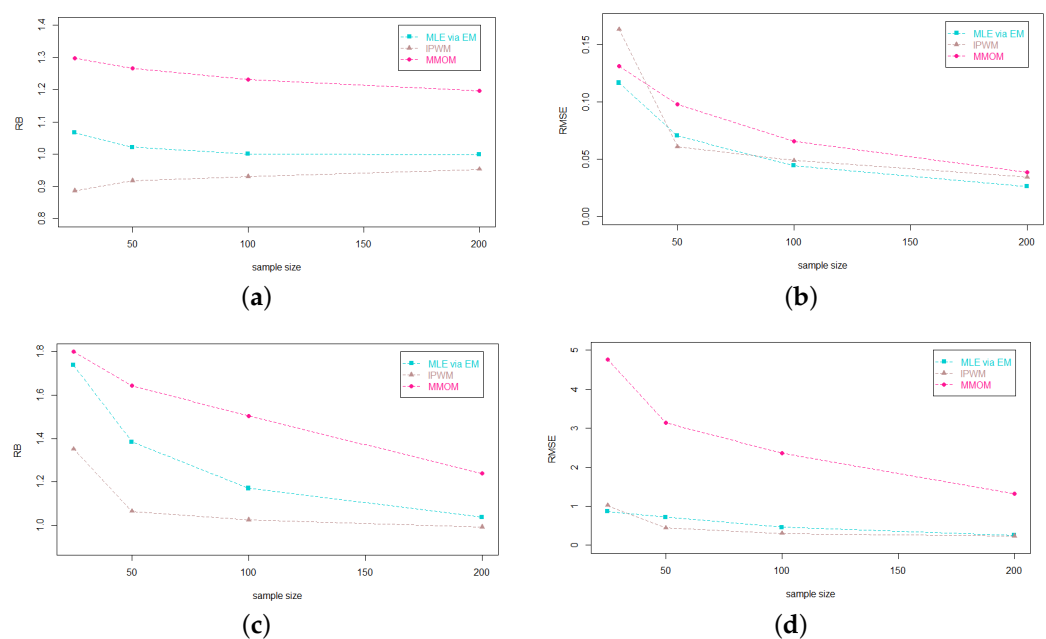

Figure 6. RB and RMSE of $(0, \sigma, \alpha, \beta)$ ($\sigma = 1, \alpha = 3, \beta = 2$). (**a**) RB of scale parameter σ. (**b**) RMSE of scale parameter σ. (**c**) RB of shape parameter α. (**d**) RMSE of shape parameter α. (**e**) RB of shape parameter β. (**f**) RMSE of shape parameter β.
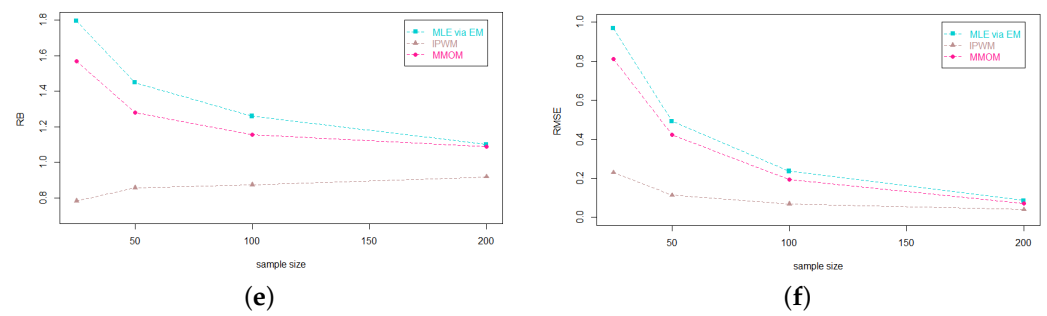
Figure 7. RB and RMSE of $(0, \sigma, \alpha, \beta)$ $(\sigma = 1, \alpha = 3, \beta = 3)$. (**a**) RB of scale parameter σ. (**b**) RMSE of scale parameter σ. (**c**) RB of shape parameter α. (**d**) RMSE of shape parameter α. (**e**) RB of shape parameter β. (**f**) RMSE of shape parameter β.

Figure 8. *Cont.*

Figure 8. RB and RMSE of $(0, \sigma, \alpha, \beta)$ $(\sigma = 1, \alpha = 3, \beta = 4)$. (**a**) RB of scale parameter σ. (**b**) RMSE of scale parameter σ. (**c**) RB of shape parameter α. (**d**) RMSE of shape parameter α. (**e**) RB of shape parameter β. (**f**) RMSE of shape parameter β.
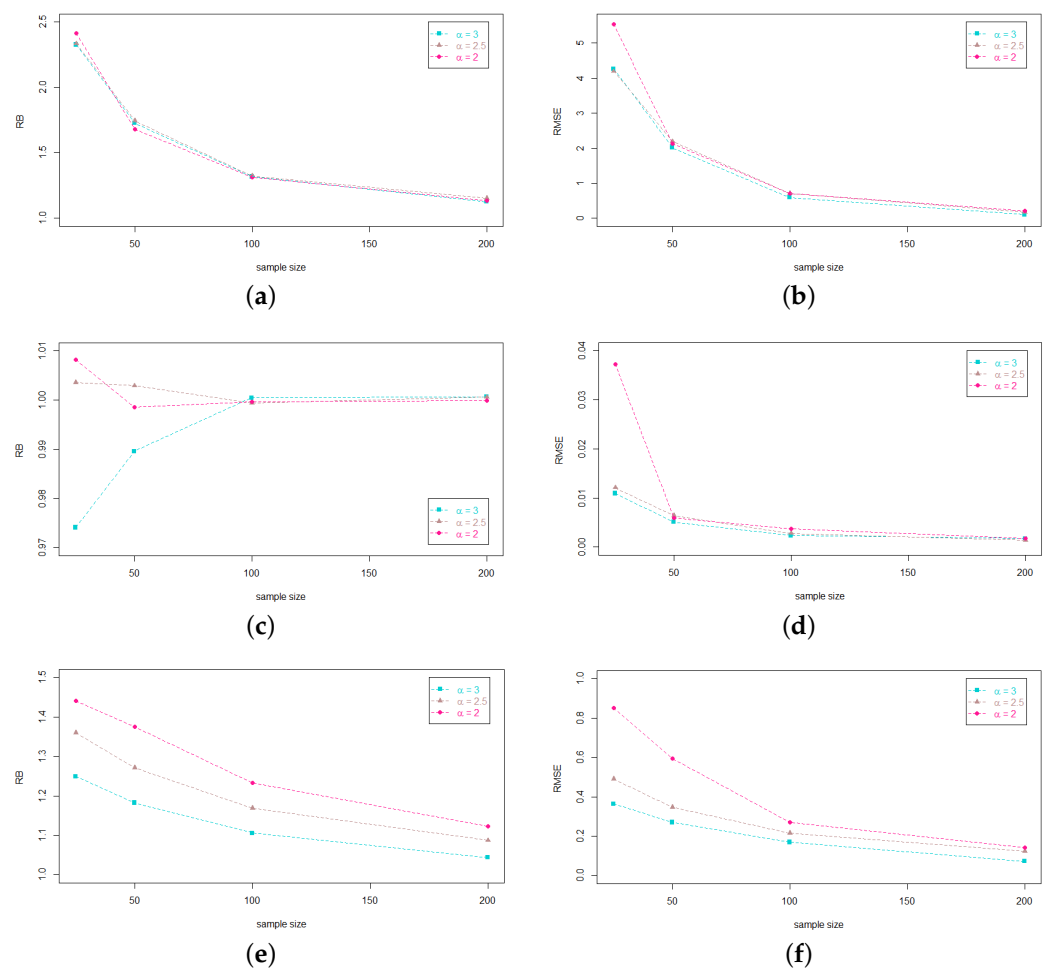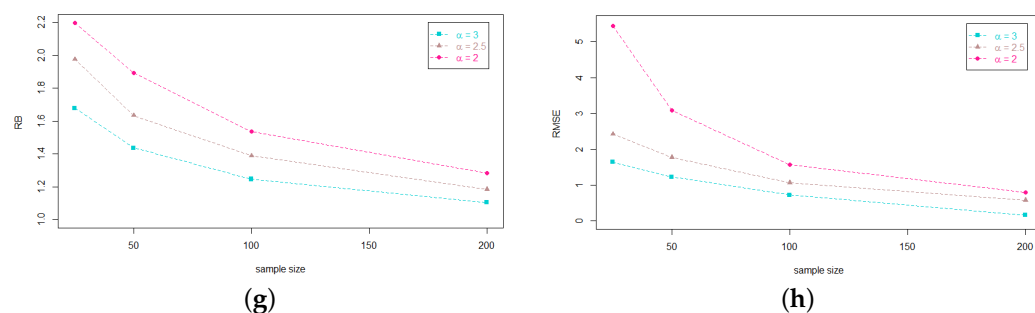

Figure 9. *Cont.*

Figure 9. RB and RMSE of $(\mu, \sigma, \alpha, \beta)$ under different shape parameter α based on the PLA using the EM algorithm. (**a**) RB of location parameter μ. (**b**) RMSE of location parameter μ. (**c**) RB of scale parameter σ. (**d**) RMSE of scale parameter σ. (**e**) RB of shape parameter α. (**f**) RMSE of shape parameter α. (**g**) RB of shape parameter β. (**h**) RMSE of shape parameter β.

7. Real Data Analysis

In this section, a real data analysis is performed to illustrate the feasibility of the new Gt distribution and a comparison with the Gt distribution is also presented.

These data concerned U.S. stock market and were introduced by Shen et al. (2020) [26]. The basic description of the statistical characteristics of the data set are shown in Table 2, which containing the sample size n, the mean value $\bar{Y}$, the standard deviation S, the asymmetry coefficient $\sqrt{b_1}$, the kurtosis coefficient b_2, the minimum value $\min(Y)$, and the maximum value $\max(Y)$. It can be seen from the table that the data set has good symmetry and higher kurtosis.

Table 2. Descriptive statistics for stock data.

n	$\bar{Y}$	S	$\sqrt{b_1}$	b_2	$\min(Y)$	$\max(Y)$
4278	-2.295×10^{-3}	0.01143	0.09466	5.28822	-0.1039	0.08113

To obtain the MLEs of the unknown parameters, the PLA depicted in Figure 10 is recommended to avoid possible numerical optimization problems. The fitting results of these two distributions and the estimated parameter values are presented in Table 3. Based on the AIC criterion, we can infer from it that the new Gt distribution is slightly more suitable for fitting this data set than the Gt distribution. The histogram of the considered data with the two density functions we fitted is drawn in Figure 11. It is clearly seen that the pdf of the new Gt distribution is steeper in shape with a lighter tail, while the Gt distribution is the opposite.

Table 3. Estimation value and the fitting index.

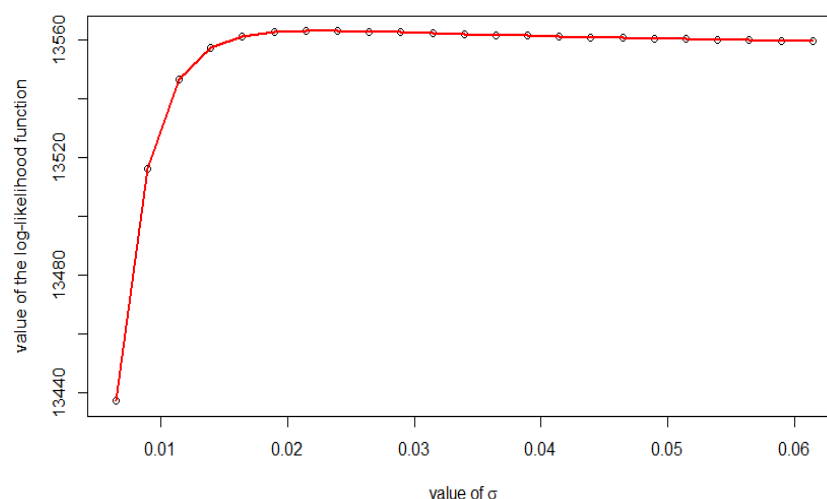| Parameter Estimates	New Gt Distribution	Gt Distribution
σ	0.02143	0.01025
α	5.46063	—
β	1.22367	—
p	—	1.92009
q	—	1.72641
log-likelihood value	13,563.13	13,543.77
AIC	$-27,118.26$	$-27,079.54$

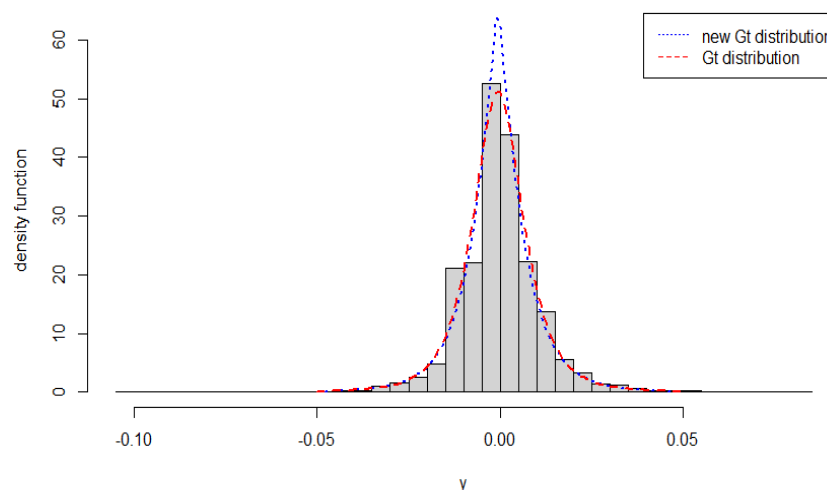**Figure 10.** Trace of the PLA based on the new Gt distribution.

Figure 11. Histogram of the stock data and the estimated pdf.

8. Conclusions

In this paper, we coin a new Gt distribution that is suitable for fitting both the data with high kurtosis and a heavy tail based on a construction approach. Firstly, we have investigated the main properties of the new distribution including moments, skewness coefficients, kurtosis coefficients and random number generation. Secondly, we have derived the explicit expression for the moments of order statistics as well as its corresponding variance–covariance matrix through recurrence relations and the distribution transformation technique. The so obtained method has high efficiency and can greatly reduce the computation time. After that, we have focused on the parameter estimation of this new Gt distribution. Several estimation methods including MMOM, MLE using the EM algorithm, MLE using a new iterative algorithm and IPWM have been introduced. Among all these estimation methods, the IPWM performs best on the whole and this novel method makes the parameter estimation method of this distribution not limited to MLE and MMOM. Furthermore, the new iterative algorithm to acquire MLE is more suitable than the EM algorithm when the sample kurtosis is more than 2.7. For four-parameter new Gt distribution, we have established an EM-type algorithm through the profile maximum likelihood approach and discovered that the variation of the shape parameter α has a significant effect on the estimation performance of scale parameter σ and shape parameter α. However, there are still some limitations and areas to be improved in our research. The parameter estimation method is limited to PLA using the EM algorithm when the distribution has

a heavier tail. Therefore, proposing more efficient and accurate estimation methods will be the focus of our future research. Besides, the new Gt distribution is suitable for fitting data with good symmetry. In future work, it can be applied for the asymmetric situation by adding a skew parameter.

Appendix A

Appendix A.1. Part I

Proof. Proof of Theorem 1
According to the pdf of the quotient of two random variables, we have

$$
\begin{aligned}
g_T(t; \sigma, \alpha, \beta) &= \frac{\beta}{2^{1+1/\beta}\Gamma(1/\beta)\sigma} \times \frac{1}{\Gamma(\alpha)\beta^{\alpha-1}} \int_0^{+\infty} e^{-|t-\mu|^\beta z^\beta/(2\sigma^\beta)} z^{\alpha\beta} e^{-z^\beta/\beta} dz \\
&= \frac{1}{2^{1+\frac{1}{\beta}}\Gamma(\alpha)\Gamma(\frac{1}{\beta})\sigma} \int_0^{+\infty} z^{\alpha\beta} e^{[|t-\mu|^\beta z^\beta/(2\sigma^\beta)+1/\beta]z^\beta} dz .
\end{aligned} \tag{A1}
$$

Let $u = [|t-\mu|^\beta z^\beta/(2\sigma^\beta) + 1/\beta]z^\beta$, we obtain

$$
\begin{aligned}
\int_0^{+\infty} z^{\alpha\beta} e^{[|t-\mu|^\beta z^\beta/(2\sigma^\beta)+1/\beta]z^\beta} dz\} &= \left[\frac{1}{\beta} + \frac{|t-\mu|^\beta}{2\sigma^\beta}\right]^{-(\alpha+\frac{1}{\beta})} \frac{1}{\beta} \int_0^{+\infty} u^{\alpha+\frac{1}{\beta}-1} e^{-u} du \\
&= \beta^{\alpha+\frac{1}{\beta}-1}\Gamma\left(\alpha + \frac{1}{\beta}\right)\left[1 + \frac{|t-\mu|^\beta}{2\sigma^\beta}\right]^{-(\alpha+\frac{1}{\beta})} .
\end{aligned}
$$

by substituting the above expression into (A1) we can derive the pdf of the random variable T in Theorem 1. □

Proof. Proof of Proposition 1

$$
\begin{aligned}
f_Y(y) &= \int_0^{+\infty} g(y|u)g(u) du \\
&= \frac{\beta^{1-\alpha}}{2^{1+\frac{1}{\beta}}\Gamma(\frac{1}{\beta})\Gamma(\alpha)\sigma} \int_0^{+\infty} u^{\alpha+\frac{1}{\beta}-1} exp\left\{-\frac{u|y-\mu|^\beta}{2\sigma^\beta} - \frac{u}{\beta}\right\} du \\
&= \frac{1}{(\frac{2}{\beta})^{1+\frac{1}{\beta}} B(\alpha, \frac{1}{\beta})\sigma}\left[1 + \frac{\beta|y-\mu|^\beta}{2\sigma^\beta}\right]^{-(\alpha+\frac{1}{\beta})} .
\end{aligned}
$$

□

Appendix A.2. Part II

Proof. Proof of Lemma 1

Notice that

$$F_Z(z) = 2G(2^{\frac{1}{\beta}}z) - 1 \quad \text{and} \quad f_Z(z) = 2^{1+\frac{1}{\beta}}g(2^{\frac{1}{\beta}}z),$$

the follwing steps are similar to Govindarajulu (1963) [16]. □

Proof. Proof of Theorem 2

Through the traditional way for cauculating the *k*-th moments of order statistics, we have

$$
\begin{aligned}
E(X_{r:n}^k) &= C_{r,n} \int_{-\infty}^{+\infty} x^k [G(x)]^{r-1} [1 - G(x)]^{n-r} g(x) dx \\
&= C_{r,n} \left\{ \int_0^{+\infty} x^k [G(x)]^{r-1} [1 - G(x)]^{n-r} g(x) dx + \int_{-\infty}^0 x^k [G(x)]^{r-1} [1 - G(x)]^{n-r} g(x) dx \right\} \\
&= C_{r,n} \left[I_1(k, r-1, n-r) + I_2(k, r-1, n-r) \right].
\end{aligned}
$$

From the symetric properties of this distribution, let $u = -x$, we have

$$I_2(k, r-1, n-r) = C_{r,n}(-1)^k \int_0^{+\infty} u^k [1 - G(u)]^{r-1} [G(u)]^{n-r} g(u) du.$$

We notice that

$$G[(2u)^{\frac{1}{\beta}}] = \frac{1}{2} \left[1 + I_{\beta u/(1+\beta u)}\left(\frac{1}{\beta}, \alpha\right) \right] \quad u > 0,$$

where $I_x(a,b) = \frac{1}{B(a,b)} \int_0^x t^{a-1}(1-t)^{b-1} dt$ denote the incomplete beta function.

$$I_1(k, r-1, n-r) = \frac{C_{r,n}}{2(\frac{1}{\beta})^{\frac{1}{\beta}} B(\frac{1}{\beta}, \alpha)} \int_0^{+\infty} u^{\frac{k+1}{\beta}-1} (1+\beta u)^{-(\alpha+\frac{1}{\beta})} \{G[(2u)^{\frac{1}{\beta}}]\}^{r-1} \{1 - G[(2u)^{\frac{1}{\beta}}]\}^{n-r} du$$

let $\quad z = \frac{\beta u}{1 + \beta u}$

$$= \frac{2^{\frac{k}{\beta}-1}(\frac{1}{\beta})^{\frac{k}{\beta}}}{B(\frac{1}{\beta}, \alpha)} \int_0^1 z^{\frac{k+1}{\beta}-1} (1-z)^{\alpha+\frac{k+2}{\beta}-1} \left[1 + I_z\left(\frac{1}{\beta}, \alpha\right) \right]^{r-1} \left[1 - I_z\left(\frac{1}{\beta}, \alpha\right) \right]^{n-r} dz$$

$$= \frac{2^{\frac{k}{\beta}-n}(\frac{1}{\beta})^{\frac{k}{\beta}}}{B(\frac{1}{\beta}, \alpha)} \sum_{i=0}^{r-1} \sum_{j=0}^{n-r} (-1)^j \binom{r-1}{i} \binom{n-r}{j} \int_0^1 z^{\frac{k+1}{\beta}-1} (1-z)^{\alpha-\frac{k}{\beta}-1} \left[I_z\left(\frac{1}{\beta}, \alpha\right) \right]^{i+j} dz$$

$$= \frac{2^{\frac{k}{\beta}-n}(\frac{1}{\beta})^{\frac{k}{\beta}}}{B(\frac{1}{\beta}, \alpha)} \sum_{i=0}^{r-1} \sum_{j=0}^{n-r} (-1)^j \binom{r-1}{i} \binom{n-r}{j} A\left(\frac{k+1}{\beta}, \alpha - \frac{k}{\beta} - 1, i+j\right).$$

When α is a real non-integer, by using generalized multinomial theorem, we obtain

$$I_z\left(\frac{1}{\beta}, \alpha\right) = \frac{z^{\frac{1}{\beta}}}{B(\frac{1}{\beta}, \alpha)} \sum_{k=0}^{\infty} \frac{(1-\alpha)_k z^k}{(\frac{1}{\beta} + k)k!},$$

$$\left[I_z\left(\frac{1}{\beta}, \alpha\right) \right]^{i+j} = \frac{z^{\frac{i+j}{\beta}}}{[B(\frac{1}{\beta}, \alpha)]^{i+j}} \sum_{m_1+\cdots+m_{i+j}=0}^{\infty} \frac{(1-\alpha)_{m_1} \cdots (1-\alpha)_{m_{i+j}}}{(\frac{1}{\beta} + m_1) \cdots (\frac{1}{\beta} + m_{i+j})} \times \frac{z^{m_1+\cdots+m_{i+j}}}{m_1! \cdots m_{i+j}!}$$

With the above equations $A(\frac{k+1}{\beta}, \alpha - \frac{k}{\beta} - 1, i+j)$ can be rewritten as

$$
\begin{aligned}
A\left(\frac{k+1}{\beta}, \alpha - \frac{k}{\beta} - 1, i+j\right) &= \left[B\left(\frac{1}{\beta}, \alpha\right)\right]^{-(i+j)} \sum_{m_1+\cdots+m_{i+j}=0}^{\infty} \frac{(1-\alpha)_{m_1} \cdots (1-\alpha)_{m_{i+j}}}{(\frac{1}{\beta}+m_1)\cdots(\frac{1}{\beta}+m_{i+j})} \\
&= \frac{\beta^{i+j}B(\frac{k+i+j+1}{\beta}, \alpha - \frac{k}{\beta})}{[B(\alpha, \frac{1}{\beta})]^{i+j}} \times F_{1:1}^{1:2}\left(\left(\frac{k+i+j+1}{\beta}\right):(1-\alpha, \frac{1}{\beta});\cdots;\right. \\
&\qquad \left. (1-\alpha, \frac{1}{\beta}):(\alpha+\frac{i+j+1}{\beta}):(\frac{1}{\beta}+1);\cdots;(\frac{1}{\beta}+1):1;\cdots;1\right),
\end{aligned}
$$

and the above expression exists for $\alpha\beta > k$.

For a suifficent large N, we can bound that

$$
\begin{aligned}
&\frac{B(\frac{k+i+j+1}{\beta}, \alpha - \frac{k}{\beta})}{[B(\frac{1}{\beta}, \alpha)]^{i+j}} \sum_{max(m_1,\cdots,m_{n-i+j})>N} \left| \frac{(1-\alpha)_{m_1}\cdots(1-\alpha)_{m_{i+j}}}{(\frac{1}{\beta}+m_1)\cdots(b+m_{i+j})} \times \frac{1}{(m_1!\cdots m_{i+j}!)} \right| \\
&< \frac{\Gamma(\frac{k+i+j+1}{\beta})}{[B(\frac{1}{\beta}, \alpha)]^{i+j}} \sum_{max(m_1,\cdots,m_{i+j})>N} \left| \frac{(1-\alpha)_{m_1}\cdots(1-\alpha)_{m_{n-i+j}}}{(\frac{1}{\beta}+m_1)\cdots(\frac{1}{\beta}+m_{i+j})} \times \frac{1}{m_1!\cdots m_{i+j}!} \right| \\
&= \frac{\Gamma(\frac{k+i+j+1}{\beta})}{[B(\frac{1}{\beta}, \alpha)]^{i+j}} \left[\sum_{m_1=0}^{\infty}\cdots\sum_{m_{i+j}=0}^{\infty} \frac{|(1-\alpha)_{m_1}\cdots(1-a)_{m_{i+j}}|}{(\frac{1}{\beta}+m_1)\cdots(\frac{1}{\beta}+m_{i+j})m_1!\cdots m_{i+j}!} \right. \\
&\qquad \left. - \sum_{m_1=0}^{N}\cdots\sum_{m_{n-i+j}=0}^{N} \frac{|(1-\alpha)_{m_1}\cdots(1-\alpha)_{m_{i+j}}|}{(\frac{1}{\beta}+m_1)\cdots(\frac{1}{\beta}+m_{i+j})m_1!\cdots m_{i+j}!} \right] \\
&= \frac{\Gamma(\frac{k+i+j+1}{\beta})}{[B(\frac{1}{\beta}, \alpha)]^{i+j}} \left\{ \left[\sum_{m=0}^{\infty} \frac{|(1-\alpha)_m|}{(b+m)m!} \right]^{i+j} - \left[\sum_{m=0}^{N} \frac{|(1-\alpha)_m|}{(b+m)m!} \right]^{i+j} \right\} < \infty.
\end{aligned}
$$

$\square$

Proof. Proof of Theorem 3

By using the genralized multinomial theorem, we acquire

$$
I_{\beta u/(1+\beta u)}\left(\frac{1}{\beta}, \alpha\right) = \frac{1}{B(\alpha, \frac{1}{\beta})} \sum_{k=0}^{\infty} \frac{(1-\frac{1}{\beta})_k(1+\beta u)^{-(\alpha+k)}}{(\alpha+k)k!},
$$

and

$$
\left[I_{\beta u/(1+\beta u)}\left(\frac{1}{\beta}, \alpha\right)\right]^{i+j} = \left[B\left(\alpha, \frac{1}{\beta}\right)\right]^{-(i+j)} \sum_{m_1\cdots m_{i+j}=0}^{\infty} \frac{(1-\frac{1}{\beta})_{m_1}\cdots(1-\frac{1}{\beta})_{m_{i+j}}}{(\alpha+m_1)\cdots(\alpha+m_{i+j})}(1+\beta u)^{-[\alpha(i+j)+\sum_{l=1}^{i+j}m_l]}.
$$

By substituting the above two expressions into $A(\frac{k+1}{\beta} - 1, \alpha + \frac{1}{\beta}, i+j)$, we obtain

$$
\begin{aligned}
A(\frac{k+1}{\beta}, \alpha + \frac{1}{\beta}, i+j) &= \frac{(\frac{1}{\beta})^{\frac{k+1}{\beta}}\alpha^{-(i+j)}}{\left[B(\alpha, \frac{1}{\beta})\right]^{i+j}} B\left(\frac{k+1}{\beta}, \alpha(i+j+1) - \frac{k}{\beta}\right) \times F_{1:1}^{1:2}\left(\left(\alpha(i+j+1) - \frac{k}{\beta}\right):\right. \\
&\qquad (1 - \frac{1}{\beta}, \alpha);\cdots;(1 - \frac{1}{\beta}, \alpha):(\alpha(i+j+1) + \frac{1}{\beta}):(\alpha+1);\cdots;(\alpha+1): \\
&\qquad \left. 1;\cdots 1\right),
\end{aligned}
$$

this expression exists for $\alpha > \frac{k}{\beta}$. $\square$

Proof. Proof of Theorem 4

Let $G_i(x)$ denote the cdf with parameters $(\mu_i, \sigma_i, \beta, \alpha_i)$. Barakat and Abdelkader (2004) [27] proposed the following expression to cauculate the moments of order statistics for the INID case

$$\mu_{r:n}^{(m)+} = \sum_{j=n-r+1}^{n} (-1)^{j-(n-r+1)} \binom{j-1}{n-r} I_j^+(m),$$

$$\mu_{r:n}^{(m)-} = \sum_{j=r}^{n} (-1)^{j-r} \binom{j-1}{r-1} I_j^-(m),$$

where

$$I_j^+(m) = \sum_{1 \le i_1 < i_2 \cdots < i_j \le n} m \int_0^{+\infty} x^{m-1} \prod_{t=1}^{j} [1 - G_{i_t}(x)] \mathrm{d}x,$$

$$I_j^-(m) = \sum_{1 \le i_1 < i_2 \cdots < i_j \le n} m \int_{-\infty}^{0} x^{m-1} \prod_{t=1}^{j} G_{i_t}(x) \mathrm{d}x.$$

Let $u = -x$, from the symetric properties of this distribution we can reexpress $I_j^-(m)$ as

$$I_j^-(m) = \sum_{1 \le i_1 < i_2 \cdots < i_j \le n} m(-1)^m \int_0^{+\infty} x^{m-1} \prod_{t=1}^{j} [1 - G_{i_t}(x)] \mathrm{d}x.$$

From Section 2 we know that $G_{i_t}[(2u)^{\frac{1}{\beta}}] = \frac{1}{2}\left[1 + I_{\beta u/(1+\beta u)}(\frac{1}{\beta}, \alpha_{i_t})\right]$. Using series expansion, we immediately acquire the expressions as follows

$$1 - G_{i_t}((2u)^{\frac{1}{\beta}}) = \frac{1}{2B(\alpha_{i_t}, \frac{1}{\beta})} \sum_{k=0}^{\infty} \frac{(1 - \frac{1}{\beta})_k}{(\alpha_{i_t} + k)k!} (1 + \beta u)^{-(\alpha_{i_t} + k)},$$

$$\prod_{t=1}^{j} \left[1 - G_{i_t}((2u)^{\frac{1}{\beta}})\right] = \frac{2^{-j}}{\prod_{t=1}^{j} B(\alpha_{i_t}, \frac{1}{\beta})} \sum_{m_1=0}^{\infty} \cdots \sum_{m_j=0}^{\infty} \frac{(1 - \frac{1}{\beta})_{m_1} \cdots (1 - \frac{1}{\beta})_{m_j}}{(\alpha_{i_1} + m_1) \cdots (\alpha_{i_j} + m_j)}$$
$$\times (m_1! \cdots m_j!)^{-1} \times (1 + \beta u)^{-(\sum_{t=1}^{j} \alpha_{i_t} + m_1 + \cdots m_j)}.$$

By subsituting the above equation into the expression of $I_j^+(m)$, we obtain

$$I_j^+(m) = \sum_{1 \le i_1 < i_2 \cdots < i_j \le n} \frac{m}{\beta} \left[2^j \prod_{t=1}^{j} B(\alpha_{i_t}, \frac{1}{\beta})\right]^{-1} \int_0^{+\infty} (2u)^{\frac{m}{\beta}-1} \sum_{m_1=0}^{\infty} \cdots \sum_{m_j=0}^{\infty}$$
$$\frac{(1 - \frac{1}{\beta})_{m_1} \cdots (1 - \frac{1}{\beta})_{m_j}}{(\alpha_{i_1} + m_1) \cdots (\alpha_{i_j} + m_j)} \times \frac{(1 + \beta u)^{-(\sum_{t=1}^{j} \alpha_{i_t} + m_1 + \cdots m_j)}}{m_1! \cdots m_j!} \mathrm{d}u$$

$$= \sum_{1 \le i_1 < i_2 \cdots < i_j \le n} \frac{m 2^{\frac{m}{\beta}-1-j}}{\prod_{t=1}^{j} B(\alpha_{i_t}, \frac{1}{\beta})} (\frac{1}{\beta})^{\frac{m}{\beta}-1} \sum_{m_1=0}^{\infty} \cdots \sum_{m_j=0}^{\infty} \frac{(1 - \frac{1}{\beta})_{m_1} \cdots (1 - \frac{1}{\beta})_{m_j}}{(\alpha_{i_1} + m_1) \cdots (\alpha_{i_j} + m_j)}$$
$$\times \frac{1}{m_1! \cdots m_j!} B\left(\frac{m}{\beta}, \sum_{t=1}^{j} \alpha_{i_t} + m_1 + \cdots m_j - \frac{m}{\beta}\right).$$

By using the defination of the generalized Kampe de Feriet function, the above equations can be reexpressed as

$$I_j^+(m) = \frac{m2^{\frac{m}{\beta}-1-j}}{\prod\limits_{t=1}^{j} \alpha_{i_t} B(\alpha_{i_t}, \frac{1}{\beta})} \left(\frac{1}{\beta}\right)^{\frac{m}{\beta}-1} B\left(\frac{m}{\beta}, \sum_{t=1}^{j} \alpha_{i_t} - \frac{m}{\beta}\right) \times F_{1:1}^{1:2}\left((\sum_{t=1}^{j} \alpha_{i_t} - \frac{m}{\beta}) : (1 - \frac{1}{\beta}, \alpha_{i_1}); \cdots \right.$$

$$\left. ;(1 - \frac{1}{\beta}, \alpha_{i_j}) : (\sum_{t=1}^{j} \alpha_{i_t}) : (\alpha_{i_1} + 1); \cdots ; (\alpha_{i_j} + 1) : 1; \cdots ; 1\right),$$

and the above expression exists for $\alpha_{i_t} > \frac{m}{\beta}$. □

Appendix A.3. Part III

Proof. Proof of Theorem 5

Let

$$h_k(\beta) = \frac{[\Gamma(1/\beta)]^{k-1}\Gamma[(2k+1)/\beta]}{[\Gamma(3/\beta)]^k}, \qquad I_k(\alpha,\beta) = \frac{[\Gamma(\alpha)]^{k-1}\Gamma(\alpha - 2k/\beta)}{[\Gamma(\alpha - 2/\beta)]^k} \ (k > 2).$$

In order to prove the monotony, we take the logarithms of the above two functions, respectively. Differentiating with respect to parameter β, we obtain

$$\frac{\partial \ln h_k(\beta)}{\partial \beta} = -\frac{1}{\beta^2}\left[(k+1)\varphi\left(\frac{1}{\beta}\right) + (2k+1)\varphi\left(\frac{2k+1}{\beta}\right) - 3k\varphi\left(\frac{3}{\beta}\right)\right]$$

$$\text{using the fact that } \varphi(z) = \sum_{k=0}^{\infty} \frac{1}{k+z}$$

$$= -\frac{4k^2 - 4k}{\beta^3} \times \frac{n}{(n+1/\beta)(n + (2k+1)/\beta)(n+3/\beta)} < 0,$$

as a result $h_k(\beta)$ is monotonically decreasing with respect to the parameter β.

For $I_k(\alpha,\beta)$, taking logarithm, we have

$$\ln I_k(\alpha,\beta) = (k-1)\ln\Gamma(\alpha) + \ln\Gamma(\alpha - 2k/\beta) - k\ln\Gamma(\alpha - 2/\beta).$$

Taking partial derivative with respect to parameter α, we have

$$\frac{\partial \ln I_k(\beta)}{\partial \alpha} = (k-1)\varphi(\alpha) + \varphi\left(\alpha - \frac{2k}{\beta}\right) - k\varphi\left(\alpha - \frac{2}{\beta}\right)$$

$$= -\frac{4k(k-1)}{\beta^2}\left[\frac{1}{(n+\alpha)(n+\alpha-2k/\beta)(n+\alpha-2/\beta)}\right] < 0.$$

Taking partial derivative with respect to parameter β, we obtain

$$\frac{\partial \ln I_k(\beta)}{\partial \alpha} == 2k[\varphi(\alpha - 2/\beta) - \varphi(\alpha - 2k/\beta)] < 0.$$

From

$$\Gamma(nt) = \frac{n^{nt}}{(2\pi)^{(n-1)/2}\sqrt{n}}\Gamma(t)\Gamma(t+1/n)\cdots\Gamma[t+(n-1)/n],$$

we have

$$
\lim_{\beta \longrightarrow 0^+} h_k(\beta) = \lim_{t \longrightarrow +\infty} h_k(t) \qquad \text{let } t = \frac{1}{\beta}
$$

$$
= \lim_{t \longrightarrow +\infty} \frac{[\Gamma(t)]^{k-1} \Gamma(t) \Gamma[t + 1/(2k+1)] \cdots \Gamma[t + 2k/(2k+1)]}{[\Gamma(t) \Gamma(t + 1/3) \Gamma(t + 2/3)]^k}
$$

$$
\times \frac{(2k+1)^{(2k+1)t} / ((2\pi)^k \sqrt{2k+1})}{3^{3t} / (2\pi \sqrt{3})}
$$

since $\displaystyle \lim_{t \longrightarrow +\infty} \frac{\Gamma(t+a)}{\Gamma(t) t^a} = 1$, the above equation can be rewritten as

$$
= \lim_{t \longrightarrow +\infty} \frac{(2k+1)^{(2k+1)t} / ((2\pi)^k \sqrt{2k+1})}{3^{3t} / (2\pi \sqrt{3})} \times \frac{[\Gamma(t)]^{k-1} [\Gamma(t)]^{2k+1} t^{\sum_{i=1}^{2k} i/(2k+1)}}{[\Gamma(t)]^k [\Gamma(t) t^{1/3}]^k [\Gamma(t) t^{2/3}]^k}
$$

$$
= \lim_{t \longrightarrow +\infty} \frac{(2k+1)^{(2k+1)t}}{\sqrt{2k+1} \times (3^{3t})^k} = +\infty.
$$

On the same way,

$$
\lim_{\beta \longrightarrow +\infty} h_k(\beta) = \lim_{t \longrightarrow 0^+} h_k(t)
$$

$$
= \frac{3^{k/2}}{\sqrt{2k+1}} \times \frac{\prod_{i=1}^{2k} \Gamma[t + i/(2i+1)]}{[\Gamma(1/3) \Gamma(2/3)]^k}.
$$

Besides that, it is easy to prove that

$$
\lim_{\beta \longrightarrow +\infty} I_k(\alpha, \beta) = 1, \qquad \lim_{\alpha \longrightarrow +\infty} I_k(\alpha, \beta) = 1, \qquad \lim_{\alpha \longrightarrow (2k/\beta)^+} I_k(\alpha, \beta) = +\infty.
$$

With the above conclusions, the result of the theorem is obvious. $\square$

Cauculation for the asymptotic variances and covariances of the estimates under EM algorithm

The second partial derivatives of the complete likelihood function is given by

$$
\frac{\partial^2 \ln L(\theta|Z)}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{\beta^2(\beta+1)}{4\sigma^2} \sum_{i=1}^{n} \left(\frac{|y_i|}{\sigma} \right) U_i, \qquad \frac{\partial^2 \ln L(\theta|Z)}{\partial \sigma \partial \alpha} = 0,
$$

$$
\frac{\partial^2 \ln L(\theta|Z)}{\partial \beta \partial \sigma} = \frac{\beta}{4\sigma} \sum_{i=1}^{n} \left[\frac{|y_i|}{\sigma} \right]^{\beta} U_i - \frac{\beta^2}{4\sigma} \sum_{i=1}^{n} \left[\frac{|y_i|}{\sigma} \right]^{\beta} \left(\ln \frac{|y_i|}{\sigma} \right) U_i - \frac{\beta}{4} \sum_{i=1}^{n} \left[\frac{|y_i|}{\sigma} \right]^{\beta-1} \left(\ln \frac{|y_i|}{\sigma} \right) U_i,
$$

$$
\frac{\partial^2 \ln L(\theta|Z)}{\partial \alpha^2} = -n \varphi'(\alpha), \qquad \frac{\partial^2 \ln L(\theta|Z)}{\partial \alpha \partial \beta} = 0,
$$

$$
\frac{\partial^2 \ln L(\theta|Z)}{\partial \beta^2} = - \frac{2n}{\beta^3} \left[-\ln \beta + (1+\beta) + 2\ln 2 + \varphi\left(\frac{1}{\beta}\right) \right] - \frac{n}{\beta^2} \left[\frac{1}{\beta} - 1 + \frac{\varphi\left(\frac{1}{\beta}\right)}{\beta^2} \right] + \frac{2}{\beta^3} \sum_{i=1}^{n} \ln U_i
$$

$$
- \frac{1}{2} \sum_{i=1}^{n} \left[\frac{|y_i|}{\sigma} \right]^{\beta} \left(\ln \frac{|y_i|}{\sigma} \right) U_i - \frac{\beta}{4} \sum_{i=1}^{n} \left[\frac{|y_i|}{\sigma} \right]^{\beta} \left(\ln \frac{|y_i|}{\sigma} \right)^2 U_i.
$$

From Proposition 2, let $V = \frac{Y}{\sigma}$, the joint density function of (U, V) is given by

$$
h_{U,V}(u, v) = \frac{\beta^{1+\frac{1}{\beta}}}{2^{\alpha+\frac{2}{\beta}} \Gamma(\frac{1}{\beta}) \Gamma(\alpha)} u^{\alpha + \frac{1}{\beta} - 1} \exp \left(-\frac{\beta u v^{\beta}}{4} - \frac{u}{2} \right), \quad u, v > 0.
$$

then through the above density function, the expectation can be calculated as

$$E(V^\beta U) = \frac{\beta^{1+\frac{1}{\beta}}}{2^{\alpha+\frac{2}{\beta}}\Gamma(\frac{1}{\beta})\Gamma(\alpha)} \int_0^{+\infty} u^{\alpha+\frac{1}{\beta}} \exp\left(-\frac{u}{2}\right) \int_0^{+\infty} v^\beta \exp\left(-\frac{\beta u v^\beta}{4}\right) dvdu \text{ let } t = \frac{\beta u v^\beta}{4}$$

$$= \frac{4}{\beta^2} ,$$

$$E(V^{\beta-1} U) = \frac{\beta^{1+\frac{1}{\beta}}}{2^{\alpha+\frac{2}{\beta}}\Gamma(\frac{1}{\beta})\Gamma(\alpha)} \int_0^{+\infty} u^{\alpha+\frac{1}{\beta}} \exp\left(-\frac{u}{2}\right) \int_0^{+\infty} v^{\beta-1} \exp\left(-\frac{\beta u v^\beta}{4}\right) dvdu \text{ let } t = \frac{\beta u v^\beta}{4}$$

$$= \frac{n\beta^{\frac{1}{\beta}-1}}{2^{\frac{1}{\beta}-2} B(\alpha, 1/\beta)} ,$$

$$E[(\ln V) V^\beta U] = \frac{\beta^{1+\frac{1}{\beta}}}{2^{\alpha+\frac{2}{\beta}}\Gamma(\frac{1}{\beta})\Gamma(\alpha)} \int_0^{+\infty} u^{\alpha+\frac{1}{\beta}} \exp\left(-\frac{u}{2}\right) \int_0^{+\infty} (\ln v) v^\beta \exp\left(-\frac{\beta u v^\beta}{4}\right) dvdu .$$

The inner integral can be calculated as

$$\int_0^{+\infty} (\ln v) v^\beta \exp\left(-\frac{\beta u v^\beta}{4}\right) dv = \frac{1}{\beta^2}\left(\frac{4}{\beta u}\right)^{1+1/\beta} \int_0^{+\infty} \left(\ln\frac{4t}{\beta u}\right) t^{1/\beta} e^{-t} dt \text{ let } t = \frac{\beta u v^\beta}{4}$$

$$= \frac{1}{\beta^2}\left(\frac{4}{\beta u}\right)^{1+1/\beta} [(\ln 4 - \ln \beta - \ln u)\Gamma(1+1/\beta) - \beta^2 \Gamma'(1+1/\beta)] .$$

By substituting the above expression into the double integral, we have

$$E[(\ln V) V^\beta U] = \frac{4}{\beta^3}\left[\ln\frac{2}{\beta} - \varphi(\alpha) - \beta^2 \varphi\left(1+\frac{1}{\beta}\right)\right]$$

On the same way, we obtain

$$E[(\ln V)^2 V^\beta U] = \frac{\beta^{1+\frac{1}{\beta}}}{2^{\alpha+\frac{2}{\beta}}\Gamma(\frac{1}{\beta})\Gamma(\alpha)} \int_0^{+\infty} u^{\alpha+\frac{1}{\beta}} \exp\left(-\frac{u}{2}\right) \int_0^{+\infty} (\ln v)^2 v^\beta \exp\left(-\frac{\beta u v^\beta}{4}\right) dvdu$$

$$= \frac{4}{\beta^4}\left\{(\ln 4)^2 + (\ln 2\beta)^2 + 2\varphi(\alpha)\ln 2\beta + \varphi'(\alpha) + [\varphi(\alpha)]^2\right\} + 4\left\{\varphi'\left(1+\frac{1}{\beta}\right) + \left[\varphi\left(1+\frac{1}{\beta}\right)\right]^2\right\}$$

$$- \frac{8ln4}{\beta^2}\left[\varphi\left(1+\frac{1}{\beta}\right) - \frac{\ln 2\beta}{\beta^2} - \frac{\varphi(\alpha)}{\beta^2}\right] + \frac{8}{\beta^2}\varphi\left(1+\frac{1}{\beta}\right)[\varphi(\alpha) + \ln 2\beta]$$

Proof. Proof of Lemma 2

First of all, we decide to prove that the first and the second equation always has a unique solution for fixed (β, α) and (σ, β), respectively.

Let $L_1(\sigma) = n - (\alpha\beta + 1) \sum_{i=1}^{n} \frac{\beta|x_i|^\beta}{2\sigma^\beta + \beta|x_i|^\beta}$.

Obviously, this function is monotonically increasing with respect to σ and $\lim_{\sigma \to 0} L_1(\sigma) = -n\alpha\beta < 0$, $\lim_{\sigma \to +\infty} L_1(\sigma) = n > 0$. Therefore, there must exist a root for $L_1(\sigma) = 0$ in $(0, +\infty)$.

On the same way, let $L_2(\alpha) = n\left[\varphi(\alpha + 1/\beta) - \varphi(\alpha)\right] - \sum_{i=1}^{n} \ln\left[1 + \frac{\beta|x_i|^\beta}{2\sigma^\beta}\right]$ and it is easy to prove that this function is monotonically decreasing with respect to α.

From Taylor's theorem we obtain

$$L_2(\alpha) \approx \frac{n}{\beta}\varphi'(\alpha) - \sum_{i=1}^{n} \ln\left[1 + \frac{\beta|x_i|^\beta}{2\sigma^\beta}\right],$$

obviously, for fixed β and σ, $\lim\limits_{\alpha \to 0} L_2(\alpha) > 0$ and $\lim\limits_{\alpha \to +\infty} L_2(\alpha) < 0$. Therefore, there must exist a root for $L_2(\alpha) = 0$ in $(0, +\infty)$.

After that, we decide to prove the convergence of the iterative algorithm. Assuming that (σ^*, α^*) denote the solution of the equations and $(\sigma_{(p)}, \alpha_{(p)})$ represent the values of p-th iteration, we have

Case I. $\alpha_{(0)} < \alpha_{(1)}$

The proof is conducted by mathematical induction, for $p = 1$, the conclusion clearly holds.

Assume that for $p = i$, we have: $\alpha_{(i)} < \alpha_{(i+1)}$, $\sigma_{(i)} < \sigma_{(i+1)}$, then $\sigma_{(i+2)}$ and $\sigma_{(i+1)}$ are obtained from the following two equations, respectively

$$n - (\alpha_{(i+1)}\beta + 1)\sum_{i=1}^{n}\frac{\beta|x_i|^\beta}{2\sigma_{(i+2)}^\beta + \beta|x_i|^\beta} = 0,$$

$$n - (\alpha_{(i)}\beta + 1)\sum_{i=1}^{n}\frac{\beta|x_i|^\beta}{2\sigma_{(i+1)}^\beta + \beta|x_i|^\beta} = 0.$$

Obviously, we can conclude that

$$(\alpha_{(i)}\beta + 1)\sum_{i=1}^{n}\frac{\beta|x_i|^\beta}{2\sigma_{(i+1)}^\beta + \beta|x_i|^\beta} - (\alpha_{(i)}\beta + 1)\sum_{i=1}^{n}\frac{\beta|x_i|^\beta}{2\sigma_{(i+2)}^\beta + \beta|x_i|^\beta} > 0,$$

and this is equivalent to $\sigma_{(i+1)} < \sigma_{(i+2)}$.

$\alpha_{(i+1)}$ and $\alpha_{(i+2)}$ are obtained from the following two equations, respectively

$$n\left[\varphi(\alpha_{(i+1)} + 1/\beta) - \varphi(\alpha_{(i+1)})\right] - \sum_{i=1}^{n}\ln\left[1 + \frac{\beta|x_i|^\beta}{2\sigma_{(i+1)}^\beta}\right] \approx \frac{n}{\beta}\varphi'(\alpha_{(i+1)}) - \sum_{i=1}^{n}\ln\left[1 + \frac{\beta|x_i|^\beta}{2\sigma_{(i+1)}^\beta}\right] = 0,$$

$$n\left[\varphi(\alpha_{(i+2)} + 1/\beta) - \varphi(\alpha_{(i+1)})\right] - \sum_{i=1}^{n}\ln\left[1 + \frac{\beta|x_i|^\beta}{2\sigma_{(i+2)}^\beta}\right] \approx \frac{n}{\beta}\varphi'(\alpha_{(i+2)}) - \sum_{i=1}^{n}\ln\left[1 + \frac{\beta|x_i|^\beta}{2\sigma_{(i+2)}^\beta}\right] = 0,$$

and this is equivalent to $\alpha_{(i+1)} < \alpha_{(i+2)}$.

Since the two squences $\sigma_{(i)}$ $\alpha_{(i)}$ statisfy: $0 < \sigma_{(j)} < +\infty$, $0 < \alpha_{(j)} < +\infty$, the conclusion holds.

Case II. $\alpha_{(0)} < \alpha_{(1)}$

The proof in this case is similar to the above case. $\square$

In a word, this conclusion can be summarized as follows:

1. if $\alpha_0 < \alpha^*$, the sequence $\{\alpha_{(i)}\}$, $\{\sigma_{(i)}\}$ are both monotonically increasing and converge from the left side to the solution,
2. if $\alpha_0 > \alpha^*$, the sequence $\{\alpha_{(i)}\}$, $\{\sigma_{(i)}\}$ are both monotonically decreasing and converge from the right side to the solution.

Appendix A.4. Part IV

Table A1. Estimation value under $\sigma = 1$, $\alpha = 3$, $\beta = 2$.

Method	Sample Size	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
MMOM	$n = 25$	1.584	7.83	2.314
	$n = 50$	1.528	6.615	2.147
	$n = 100$	1.463	5.331	2.105
	$n = 200$	1.265	4.584	2.071
MLE via EM	$n = 25$	1.18	5.174	3.228
	$n = 50$	1.132	4.421	2.482
	$n = 100$	1.104	3.814	2.364
	$n = 200$	1.036	3.341	2.201
IPWM	$n = 25$	1.109	3.659	3.766
	$n = 50$	1.062	3.49	2.882
	$n = 100$	1.033	3.3885	2.528
	$n = 200$	1.022	3.308	2.35

Table A2. Estimation value under $\sigma = 1$, $\alpha = 3$, $\beta = 3$.

Method	Sample Size	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
MMOM	$n = 25$	1.373	7.830	5.339
	$n = 50$	1.314	6.615	3.659
	$n = 100$	1.263	5.331	3.176
	$n = 200$	1.217	4.584	2.98
MLE via EM	$n = 25$	1.046	5.073	5.207
	$n = 50$	1.032	4.119	4.524
	$n = 100$	1.021	3.531	3.884
	$n = 200$	0.997	3.181	3.481
IPWM	$n = 25$	1.059	4.087	2.569
	$n = 50$	1.044	3.408	2.691
	$n = 100$	1.037	3.182	2.737
	$n = 200$	1.025	3.05	2.825

Table A3. Estimation value under $\sigma = 1$, $\alpha = 3$, $\beta = 4$.

Method	Sample Size	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
MMOM	$n = 25$	1.298	5.403	6.281
	$n = 50$	1.267	4.933	5.124
	$n = 100$	1.231	4.508	4.619
	$n = 200$	1.197	3.718	4.358
MLE via EM	$n = 25$	1.067	5.214	7.179
	$n = 50$	1.022	4.155	5.794
	$n = 100$	1.002	3.513	5.039
	$n = 200$	0.999	3.115	4.402
IPWM	$n = 25$	0.887	4.056	3.672
	$n = 50$	0.919	3.224	3.496
	$n = 100$	0.932	3.075	3.318
	$n = 200$	0.954	2.976	3.132

Table A4. Estimation value for PLA using the EM algorithm under different α (fixed $\mu = 1$, $\sigma = 1$, $\beta = 2$).

Value of α	Sample Size	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
	$n = 25$	1.0082	1.4411	4.3878	4.8281
	$n = 50$	0.9986	1.3755	3.7841	3.3574
$\alpha = 2$	$n = 100$	0.9997	1.2325	3.0712	2.6304
	$n = 200$	0.9999	1.1233	2.5676	2.2731
	$n = 25$	1.0036	1.3598	4.9423	4.6686
	$n = 50$	1.0029	1.2715	4.0851	3.4920
$\alpha = 2.5$	$n = 100$	0.9994	1.1692	3.4725	2.6501
	$n = 200$	1.0006	1.0876	2.962	2.3084
	$n = 25$	0.9741	1.2494	5.0373	4.6552
	$n = 50$	0.9896	1.1827	4.3119	3.4574
$\alpha = 3$	$n = 100$	1.0005	1.1058	3.7419	2.6380
	$n = 200$	0.954	1.0006	3.3182	2.2532

References

1. McDonald, J.B.; Newey, W.K. Partially Adaptive Estimation of Regression Models via the Generalized t Distribution. *Econom. Theory* **1988**, *4*, 428–457.
2. Nadarajah, H.N. On the Generalized *t* (Gt) Distribution. *Statistics* **2008**, *42*, 467–473.
3. Galbraith, J.W.; Zhu, D. A Generalized Asymmetric Student-t Distribution with Application to Financial Econometrics. *J. Econom.* **2010**, *157*, 297–305.
4. Harvey, A.; Lange, R.J. Volatility Modeling with a Generalized t Distribution. *J. Time Ser. Anal.* **2017**, *38*, 175–190.
5. Peel, D.; McLachlan, G.J. Robust Mixture Modelling Using the t Distribution. *Stat. Comput.* **2000**, *10*, 339–348.
6. Arslan, O.; Genç, A.İ. Robust Location and Scale Estimation Based on the Univariate Generalized t (GT) Distribution. *Commun.-Stat.-Theory Methods* **2003**, *32*, 1505–1525.
7. Paula, G.A.; Leiva, V.; Barros, M.; Liu, S. Robust Statistical Modeling Using the Birnbaum Saunders t Distribution Applied to Insurance. *Appl. Stoch. Model. Bus. Ind.* **2012**, *150*, 169–185.
8. Çankaya, M.N.; Arslan, O. On the Robustness Properties for Maximum Likelihood Estimators of Parameters in Exponential Power and Generalized *t* Distributions. *Commun.-Stat.-Theory Methods* **2018**, *49*, 607–630.
9. Arslan, O.; Genç, A.İ. The Skew Generalized *t* Distribution as the Scale Mixture of a Skew Exponential Power Distribution and Its Applications in Robust Estimation. *Statistics* **2009**, *43*, 481–498.
10. Venegas, O.; Rodríguez, F.; Gómez, H.W.; Olivares-Pacheco, J.F.; Bolfarine, H. Robust Modeling Using the Generalized Epsilon Skew t Distribution. *J. Appl. Stat.* **2012**, *39*, 2685–2698.
11. Papastathopoulos, I.; Tawn, J.A. A Generalised Student's *t*-Distribution *Stat. Probab. Lett.* **2013**, *83*, 70–77.
12. Acitas, S.; Senoglu, B.; Arslan, O. Alpha-Skew Generalized *t* Distribution. *Rev. Colomb. Estad.* **2015**, *38*, 371–384.
13. Lak, F.; Alizadeh, M.; Monfared, M.E.D.; Esmaeili, H. The Alpha-Beta Skew Generalized *t* Distribution: Properties and Applications. *Pak. J. Stat. Oper. Res.* **2019**, *46*, 605–616.
14. Genç, A.İ. The Generalized T Birnbaum–Saunders Family. *Statistics* **2013**, *47*, 613–625.
15. Balakrishnan, N.; Kundu, D. Birnbaum-Saunders Distribution: A Review of Models, Analysis, and Applications. *Appl. Stoch. Model. Bus. Ind.* **2019**, *35*, 4–49.
16. Govindarajulu, Z. Relationships Among Moments of Order Statistics in Samples from Two Related Populations. *Technometrics* **1963**, *5*, 514–518.
17. David, H.A.; Nagaraja, H.N. *Order Statistics*, 3rd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2003.
18. Exton, H. *Handbook of Hypergeometric Integrals: Theory, Applications, Tables, Computer Programs*; Ellis Horwood: Chichester, UK, 1978.
19. Arnold, B.C.; Balakrishnan, N.; Nagaraja, H.N. *A First Course in Order Statistics*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2008.
20. Balakrishnan, N.; Alam, F.M.A. Maximum Likelihood Estimation of the Parameters of Student's *t* Birnbaum-Saunders Distribution: A Comparative Study. *Commun.-Stat.-Simul. Comput.* **2019**, *2*, 1–30.
21. Poursadeghfard, T.; Jamalizadeh, A.; Nematollahi, A. On the Extended Birnbaum—Saunders Distribution Based on the Skew-*t*-Normal Distribution. *Iran. J. Sci. Technol. Trans. Sci.* **2019**, *43*, 1689–1703.
22. Dempster, A.P. Maximum Likelihood from Incomplete Data via the EM Algorithm. *J. R. Stat. Soc.* **1977**, *39*, 1–38.
23. Louis, T.A. Finding the Observed Information Matrix When Using the EM Algorithm. *J. R. Stat. Soc. Ser. Stat. Methodol.* **1982**, *44*, 226–233.
24. Greenwood, J.A.; Landwehr, J.M.; Matalas, N.C.; Wallis, J.R. Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form. *Water Resour. Res.* **1979**, *15*, 1049–1054.

25. Wang, Q.J. Estimation of the GEV Distribution from Censored Samples by Method of Partial Probability Weighted Moments. *J. Hydrol.* **1990**, *120*, 103–114.
26. Shen, Z.; Chen, Y.; Shi, R. Modeling Tail Index with Autoregressive Conditional Pareto Model. *J. Bus. Econ. Stat.* **2020**, 1–9, doi:10.1080/07350015.2020.1832504.
27. Barakat, H.M.; Abdelkader, Y.H. Computing the Moments of Order Statistics from Nonidentical Random Variables. *Stat. Methods Appl.* **2004**, *13*, 15–26.