REVSTAT – Statistical Journal
Volume 15, Number 3, July 2017, 455–471

OPTIMAL B-ROBUST ESTIMATORS FOR THE PARAMETERS OF THE GENERALIZED HALF-NORMAL DISTRIBUTION

Authors: Fatma Zehra Doğru
– Department of Econometrics, Giresun University,
28100 Giresun, Turkey
fatma.dogru@giresun.edu.tr

Y. Murat Bulut
– Department of Statistics, Eskişehir Osmangazi University,
26480 Eskişehir, Turkey
ymbulut@ogu.edu.tr

Olcay Arslan
– Department of Statistics, Ankara University,
06100 Ankara, Turkey
oarslan@ankara.edu.tr

Received: December 2014 Revised: November 2015 Accepted: November 2015

Abstract:

- The purpose of this study is to propose robust estimators by using optimal B-robust (OBR) estimation method (Hampel *et al.* [5]) for the parameters of the generalized half-normal (GHN) distribution. After given the robust estimators, we provide a small simulation study to compare its performance with the estimators obtained from maximum likelihood (ML) estimation method. We also give a real data example to illustrate the performance of the proposed estimators.

Key-Words:

- *generalized half-normal; optimal B-robust; maximum likelihood.*

AMS Subject Classification:

- 62F35, 65C60.

1. INTRODUCTION

The GHN distribution was introduced by Cooray and Ananda [3] as an alternative lifetime distribution. It is observed by Cooray and Ananda [3] that the cumulative distribution function (cdf) of the new family is very similar to the cdf of the half-normal distribution. Thus, they called the new family as the "generalized half-normal (GHN) distribution". It can be seen that the GHN distribution is a special case of the three-parameter generalized gamma distribution given by Stacy [7] (Cooray and Ananda [3]).

Some distributional properties of the GHN distribution are given by Cooray and Ananda [3]. In their study, the parameters of the GHN distribution are estimated using the ML estimation method, and using real data sets the performance of the ML estimator is compared with the other commonly used failure time distributions such as Weibull, gamma, lognormal and Birnbaum–Saunders.

One way of estimating the parameters of a given distribution is to use the ML estimation method. However, this estimator can be very sensitive to the outliers. Thus, the robust estimators may be needed as an alternative to the ML estimators in the presence of outliers. In this paper, we will use the OBR estimation method to obtain robust estimators for the parameters of the GHN distribution. The OBR estimation method was introduced by Hampel *et al.* [5] and used by Victoria-Feser [8] and Victoria-Feser and Ronchetti [9] to estimate the shape parameters of the Pareto and the gamma distributions. Also, Doğru and Arslan [4] used the OBR estimation method to estimate the shape parameters of the Burr XII distribution. Our goal is to show that the OBR estimation method can be used as an alternative to the ML estimation method to obtain robust estimators for the parameters of the GHN distribution when the data includes outliers.

The paper is organized as follows. In Section 2, we briefly summarize the properties of the GHN distribution. In Section 3, we explore the estimation of the GHN distribution. We first give the ML estimation method and then we give the OBR estimation method. We also give the algorithm to obtain the OBR estimates. In Sections 4 and 5, we give a simulation study and a real data example to demonstrate the performance of the proposed estimators over the ML estimators. Some conclusions are given in Section 6.

2. GENERALIZED HALF-NORMAL DISTRIBUTION (GHN)

The probability density function (pdf) and the cdf of the GHN distribution are given by

$$(2.1)\qquad f(x;\alpha,\theta)=\begin{cases}\sqrt{\frac{2}{\pi}}\left(\frac{\alpha}{x}\right)\left(\frac{x}{\theta}\right)^{\alpha}e^{-\frac{1}{2}\left(\frac{x}{\theta}\right)^{2\alpha}}, & x>0,\ \alpha>0,\ \theta>0\\ 0 & ,\ x\le 0\end{cases}$$

and

$$(2.2)\qquad F(x;\alpha,\theta)=2\Phi\left(\left(\frac{x}{\theta}\right)^{\alpha}\right)-1\,,\qquad x\ge 0,\quad \alpha>0,\quad \theta>0$$

respectively, where $\Phi(\cdot)$ is the cdf of the standard normal distribution and α and θ are the shape and scale parameters of the GHN distribution.

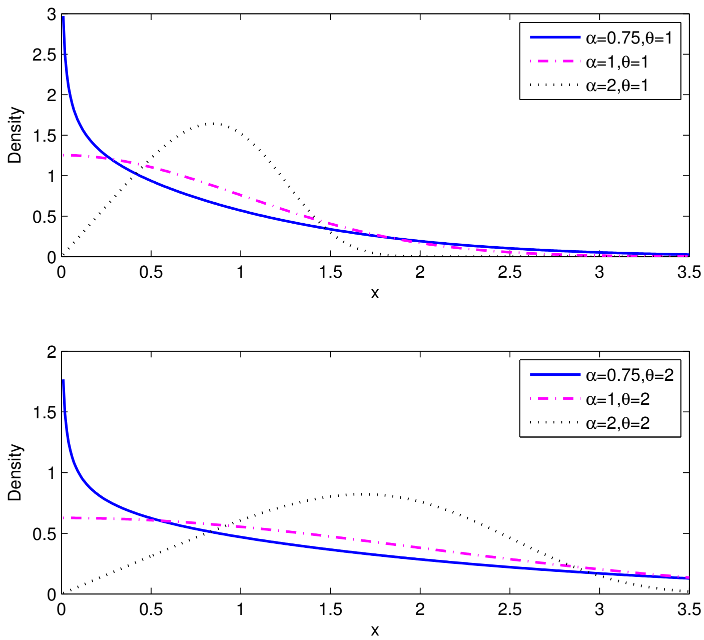

Figure 1: Examples of the GHN pdf for different values of α and θ.

The k-th moment, expected value and the variance are given by Cooray and Ananda [3] as follows:

$$E(X^k)=\sqrt{\frac{2^{\frac{k}{\alpha}}}{\pi}}\,\Gamma\left(\frac{k+\alpha}{2\alpha}\right)\theta^k\,,$$

$$E(X)=\sqrt{\frac{2^{\frac{1}{\alpha}}}{\pi}}\,\Gamma\left(\frac{1+\alpha}{2\alpha}\right)\theta$$

and

$$Var(X) = \frac{2^{\frac{1}{\alpha}}}{\pi}\left(\sqrt{\pi}\,\Gamma\!\left(\frac{2+\alpha}{2\alpha}\right) - \Gamma^2\!\left(\frac{1+\alpha}{2\alpha}\right)\right)\theta^2 ,$$

where $\Gamma(\cdot)$ is the gamma function. Figure 1 shows the plots of the pdf of the GHN distribution for some values of α and θ.

3. PARAMETER ESTIMATION

In this section, the parameters of the GHN distribution will be estimated using the ML and the OBR estimation methods.

3.1. ML estimation method

Let $X = (x_1, x_2, ..., x_n)$ be a random sample from GHN distribution. The log-likelihood function is

$$\log L(\alpha,\theta) = \frac{n}{2}\log\!\left(\frac{2}{\pi}\right) + n\log\alpha - n\alpha\log\theta + (\alpha-1)\sum_{i=1}^{n}\log(x_i) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i}{\theta}\right)^{2\alpha} . \tag{3.1}$$

Taking the derivatives of the log-likelihood function with respect to α and θ and setting to zero give the following equations

$$\frac{n}{\widehat{\alpha}} + \sum_{i=1}^{n}\log x_i - n\left(\sum_{i=1}^{n}x_i^{2\widehat{\alpha}}\log x_i\right)\left(\sum_{i=1}^{n}x_i^{2\widehat{\alpha}}\right)^{-1} = 0 \tag{3.2}$$

and

$$\widehat{\theta} = \left(\frac{1}{n}\sum_{i=1}^{n}x_i^{2\widehat{\alpha}}\right)^{\frac{1}{2\widehat{\alpha}}} . \tag{3.3}$$

Note that the same equations are also given by Cooray and Ananda [3]. There is not an analytical solution to the system formed by equations (3.2) and (3.3). This system can be only solved using numerical methods.

3.2. OBR estimation method

The OBR estimator introduced by Hampel *et al.* [5] belongs to the class of M-estimators (Huber [6]). Let $\boldsymbol{\eta} = (\alpha,\theta)$. The class of M-estimator for the parameter $\boldsymbol{\eta}$ is defined as the minimum of the following objective function

$$\sum_{i=1}^{n} \rho(x_i, \boldsymbol{\eta}) \;.$$

If the ρ function is differentiable the M-estimator will be the solution of the following equation

$$\sum_{i=1}^{n} \psi(x_i, \boldsymbol{\eta}) \;=\; 0 \;,$$

where $\psi = \rho'$ with $\psi\colon X \times \mathbb{R}^p \to \mathbb{R}^p$. There are many ρ functions in literature. In this paper, we will use the Huber's ρ function defined as

$$\rho_b(x) \;=\; \begin{cases} \frac{x^2}{2} \;, & |x| \le b \\ b\,|x| - \frac{1}{2}\,b^2 \,, & |x| > b \;. \end{cases}$$

Here, b is the robustness tuning constant and the derivative of ρ is $\psi_b(x)$ with

$$\psi_b(x) = \begin{cases} x \;, & |x| \le b \\ sgn(x)\,b \;, & |x| > b \;. \end{cases}$$

In general, the influence function (IF) for an M-estimator is defined as

$$IF \;=\; \frac{\psi(x, \boldsymbol{\eta})}{-\int \frac{\partial}{\partial \boldsymbol{\eta}} \psi(x, \boldsymbol{\eta})\, dF_{\boldsymbol{\eta}}(x)} \tag{3.4}$$

and it is used to measure the local robustness of an estimator. It is desired that IF is bounded. The estimators with bounded IF are called the OBR estimators. The IF of an ML estimator is

$$IF \;=\; J(\boldsymbol{\eta})^{-1}\, \mathbf{s}(x, \boldsymbol{\eta}) \;,$$

where $J(\boldsymbol{\eta})$ is the Fisher information matrix and $\mathbf{s}(x, \boldsymbol{\eta}) = \big(\frac{\partial}{\partial \boldsymbol{\eta}}\big) \log f_{\boldsymbol{\eta}}(x)$ is the score function. It can be seen that the IF of an ML estimator is proportional to the score functions, so the score function should be bounded for a bounded IF for the ML estimator.

Concerning the GHN distribution, we take logarithm of $f(x; \alpha, \theta)$ given in (2.1) to obtain the score functions

$$\log\big(f(x; \alpha, \theta)\big) \;=\; \log\left(\sqrt{\frac{2}{\pi}}\right) + \log(\alpha) - \log(x) - \frac{1}{2}\left(\frac{x}{\theta}\right)^{2\alpha} + \alpha \log\left(\frac{x}{\theta}\right) .$$

Then, taking the derivatives of the $\log\big(f(x; \alpha, \theta)\big)$ with respect to α and θ we obtain the following equations

$$\frac{\partial\Big(\log\big(f(x; \alpha, \theta)\big)\Big)}{\partial \alpha} \;=\; \frac{1}{\alpha} + \log\left(\frac{x}{\theta}\right) - \left(\frac{x}{\theta}\right)^{2\alpha} \log\left(\frac{x}{\theta}\right) , \tag{3.5}$$

(3.6)
$$\frac{\partial\Big(\log\big(f(x;\alpha,\theta)\big)\Big)}{\partial\theta} \;=\; -\frac{\alpha}{\theta} + \frac{\alpha}{\theta^{2\alpha+1}}\; x^{2\alpha}\;.$$

After some straightforward simplifications, the score functions for the parameters α and θ are given

$$\mathbf{s}(x;\alpha,\theta) \;=\; \begin{bmatrix} \frac{1}{\alpha} + \log\big(\frac{x}{\theta}\big)\Big(1-\big(\frac{x}{\theta}\big)^{2\alpha}\Big) \\ -\frac{\alpha}{\theta} + \frac{\alpha}{\theta^{2\alpha+1}}\, x^{2\alpha} \end{bmatrix}.$$

It is clear that the score functions are not bounded functions of x (see Figures (2) and (3)). Thus, the IF of the ML estimator for the GHN distribution will be unbounded. This implies that the ML estimators will be very sensitive to the outliers. Therefore, robust estimation methods will be needed to estimate the parameters of the GHN distribution.

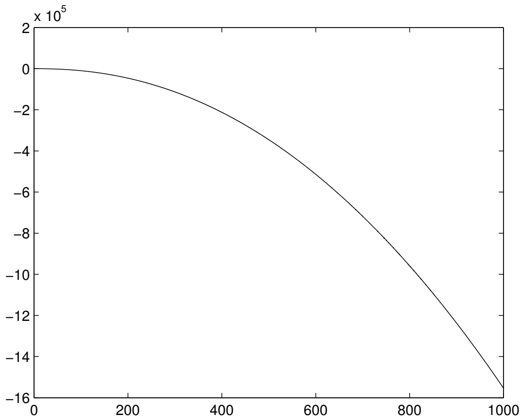

Figure 2: Score function for α parameter with $\alpha = 1$ and $\theta = 2$.

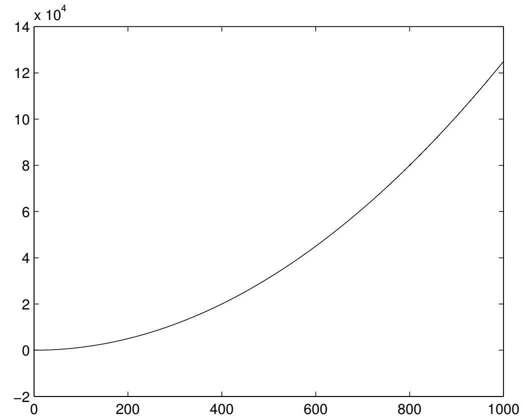

Figure 3: Score function for θ parameter with $\alpha = 1$ and $\theta = 2$.

There are several versions of the OBR estimators defined in Hampel *et al.* [5, p. 243], depending on the method of choosing to bound IF. In this study, we used the standardized OBR estimator which is defined as follows

$$\sum_{i=1}^{n}\psi\Big(\mathbf{A}(\boldsymbol{\eta})\left(\mathbf{s}(x_i,\boldsymbol{\eta})-\mathbf{a}(\boldsymbol{\eta})\right)\Big) \;=\; \sum_{i=1}^{n}W_b(x_i,\boldsymbol{\eta})\big\{\mathbf{s}(x_i,\boldsymbol{\eta})-\mathbf{a}(\boldsymbol{\eta})\big\} \;=\; \mathbf{0}\,, \tag{3.7}$$

where

$$W_b(x,\boldsymbol{\eta}) \;=\; \min\left\{1;\,\frac{b}{\big\|\mathbf{A}(\boldsymbol{\eta})\left\{\mathbf{s}(x_i,\boldsymbol{\eta})-\mathbf{a}(\boldsymbol{\eta})\right\}\big\|}\right\}, \tag{3.8}$$

is the weight function and $\|\cdot\|$ shows the Euclidian norm. Also the nonsingular $p\times p$ matrix $\mathbf{A}(\boldsymbol{\eta})$ and the $p\times 1$ vector $\mathbf{a}(\boldsymbol{\eta})$ are defined implicitly by

$$E\Big\{\psi(x,\boldsymbol{\eta})\,\psi(x,\boldsymbol{\eta})^T\Big\} = \Big\{\mathbf{A}(\boldsymbol{\eta})^T\mathbf{A}(\boldsymbol{\eta})\Big\}^{-1}, \tag{3.9}$$

$$E\Big\{\psi(x,\boldsymbol{\eta})\Big\} = \mathbf{0}\,. \tag{3.10}$$

The weight will be 1, if $\|\mathbf{A}(\boldsymbol{\eta})\left\{\mathbf{s}(x_i,\boldsymbol{\eta})-\mathbf{a}(\boldsymbol{\eta})\right\}\|\le b$, otherwise it will be $\frac{b}{\|\mathbf{A}(\boldsymbol{\eta})\left\{\mathbf{s}(x_i,\boldsymbol{\eta})-\mathbf{a}(\boldsymbol{\eta})\right\}\|}$, which bounds the score function for the outlying observations. Thus, the corresponding OBR estimator will be less sensitive to the outliers in the data.

To obtain the OBR estimates the following algorithm can be applied. Note that this algorithm was proposed by Victoria-Feser and Ronchetti [9].

Algorithm:

Step 1. Let ϵ be a stopping rule. Take initial values for the parameter $\boldsymbol{\eta}$. Set $\mathbf{a}=\mathbf{0}$ and $\mathbf{A}=\mathbf{J}^{\frac{1}{2}}(\boldsymbol{\eta})^{-T}$, where

$$\mathbf{J}(\boldsymbol{\eta}) \;=\; \int\mathbf{s}(x,\boldsymbol{\eta})\,\mathbf{s}(x,\boldsymbol{\eta})^T\,dF_{\boldsymbol{\eta}}(x)\,.$$

Step 2. Solve the following equations for $\mathbf{a}$ and $\mathbf{A}$

$$\mathbf{A}\,\mathbf{A}^T = \mathbf{M}_2^{-1}$$

and

$$\mathbf{a} \;=\; \frac{\int W_b(x,\boldsymbol{\eta})\,\mathbf{s}(x,\boldsymbol{\eta})\,dF_{\boldsymbol{\eta}}(x)}{\int W_b(x,\boldsymbol{\eta})\,dF_{\boldsymbol{\eta}}(x)}\,,$$

where

$$\mathbf{M}_k = \int W_b(x,\boldsymbol{\eta})^k\left\{\mathbf{s}(x,\boldsymbol{\eta})-\mathbf{a}\right\}\left\{\mathbf{s}(x,\boldsymbol{\eta})-\mathbf{a}\right\}^T dF_{\boldsymbol{\eta}}(x)\,, \quad k=1,2\,.$$

The current values of $\boldsymbol{\eta}$, $\mathbf{a}$ and $\mathbf{A}$ are used as starting values to solve the given equations.

Step 3. Compute $\mathbf{M_1}$ and $\Delta\boldsymbol{\eta} = \boldsymbol{M_1}^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} W_b(x,\boldsymbol{\eta})\left\{\mathbf{s}(x,\boldsymbol{\eta}) - \mathbf{a}\right\}\right)$.

Step 4. If $|\Delta\boldsymbol{\eta}| > \epsilon$ then $\boldsymbol{\eta} \longrightarrow \boldsymbol{\eta} + \Delta\boldsymbol{\eta}$ and return to step 2, else stop.

Note that for the finite sample case the integrals in the equations will be replaced by the summations.

The ML estimator can be taken as initial value for the parameter $\boldsymbol{\eta}$. In our simulation study, we have used several different initial points including the ML and true parameter values. We have also used robust starting values suggested by Victoria-Feser and Ronchetti [9]. As pointed out by Victoria-Feser and Ronchetti [9] the algorithm is convergent depending on the initial values. Other estimators such as moment estimators can also be used as starting values. However, for this distribution it is not possible to obtain explicit form of the moment estimators. Therefore, it is not tractable to use them as initial values for the algorithm.

4. SIMULATION STUDY

In this section, we will give a simulation study to compare the performance of the OBR estimators with the ML estimators with and without outliers in the data. The data are randomly generated from GHN distribution for different values of α and θ parameters. The data generation is conducted as follows. Let $U \sim \mathit{Uniform}(0,1)$. Then, $X \sim \theta\big(\Phi^{-1}\big(\frac{U+1}{2}\big)\big)^{\frac{1}{\alpha}}$ will have GHN distribution with the parameters α and θ. To evaluate the performance of the estimators bias and root mean square error (RMSE) are computed over 1000 replications for the sample sizes $n = 25, 50, 100$ and the parameter values $(\alpha,\theta) = (0.75,1), (1,1), (2,1), (0.75,2), (1,2), (2,2)$. Here, the bias and RMSE are defined as

$$\mathit{bias}(\widehat{\alpha}) = \overline{\alpha} - \alpha\ , \qquad \mathit{bias}(\widehat{\theta}) = \overline{\theta} - \theta\ ,$$

$$\mathit{RMSE}(\widehat{\alpha}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\widehat{\alpha}_i - \alpha)^2}\,, \qquad \mathit{RMSE}(\widehat{\theta}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\widehat{\theta}_i - \theta)^2}\,,$$

where $\overline{\alpha} = \frac{1}{N}\sum_{i=1}^{N}\widehat{\alpha}_i$, $\overline{\theta} = \frac{1}{N}\sum_{i=1}^{N}\widehat{\theta}_i$ and $N = 1000$. For all simulation cases, the stopping rule ϵ is taken as 10^{-6}. The simulation study and real data example are conducted using MATLAB R2013a.

In the OBR estimation method, the robustness tuning constant should be chosen in order to gain the desire efficiency. The most robust estimator can be obtained by choosing b as the squared of the number of parameters. In our case, we can take $b = \sqrt{2}$. For this value of b one can approximately gain 60% efficiency for the resulting estimators. When we increase the value of b, the efficiency will also increase. Therefore, we have taken $b = 2$ to have efficiency more than 60%. For more details about the selection of the robustness tuning constant, see Hampel *et al.* [5, p.247] and Victoria-Feser and Ronchetti [9].

Concerning the starting value for the algorithm given in Section 3.2 we use the robust starting values suggested by Victoria-Feser and Ronchetti [9] which is described as follows.

- **i)** Find the ML estimates.
- **ii)** Take $b = 3.5$ to get OBR estimates.
- **iii)** Use the OBR estimates obtained at step ii) as new initial values and set $b = 2$ to obtain OBR estimates again.

In this simulation study, we consider two types of outlier model to the right and the left in the X direction. These models are

$$\textit{Case I.}\quad (n-r)\ GHN(x;\alpha,\theta)\ +\ r\ Uniform\big(max(x)+5\sigma,\ max(x)+10\sigma\big)\ ,$$

$$\textit{Case II.}\quad (n-r)\ GHN(x;\alpha,\theta)\ +\ r\ Uniform(0, 0.0001)\ ,$$

where $max\,(x)$ is the largest observations in the sample, σ is standard deviation of a randomly generated sample from GHN distribution and r is chosen by multiplying the sample sizes by 0.1. That is, we add two outliers for $n = 25$, five outliers for $n = 50$ and ten outliers for $n = 100$. Further, in Case I, we add outliers in the upper tail of the distribution. In the second case, the outliers are added in the lower tail of the distribution to see the performance of the estimators for this type of outliers. As suggested by a referee, this type of outliers may represent severely shorted life-lengths.

Simulation results are given in Tables 1–3. In the tables, the estimates of the parameters, bias and RMSE are presented. Table 1 shows the results for the case without outliers in the data. From this table, we can observe that the performance of the ML seems slightly better than the performance of the OBR estimators. In Table 2, we give the simulation results for the outlier Case I. The results of this table show that the OBR estimators have smaller bias and RMSE values than the corresponding values for ML estimators in all the simulation configurations. Finally, in Table 3 the simulation results for the outlier Case II are displayed. Similar to the outlier Case I the OBR estimators outperform the ML estimators in terms of bias and RMSE in all the simulation scenarios. Overall when the data has outliers the OBR estimators should be used instead of ML estimators to obtain robust estimators for the parameters of interest.

Table 1: Estimates of parameters, bias and RMSE for different sample sizes without outlier.

n	θ	α	Parameter(α)			Parameter(θ)		
				ML	OBR		ML	OBR
25	1	0.75	$\overline{\alpha}$	0.7995	0.8808	$\overline{\theta}$	0.9915	0.9109
			Bias($\widehat{\alpha}$)	0.0495	0.1308	Bias($\widehat{\theta}$)	−0.0085	−0.0891
			RMSE($\widehat{\alpha}$)	0.1503	0.2061	RMSE($\widehat{\theta}$)	0.1978	0.2129
		1	$\overline{\alpha}$	1.0589	1.1677	$\overline{\theta}$	0.9931	0.9310
			Bias($\widehat{\alpha}$)	0.0589	0.1677	Bias($\widehat{\theta}$)	−0.0069	−0.0690
			RMSE($\widehat{\alpha}$)	0.2000	0.2741	RMSE($\widehat{\theta}$)	0.1548	0.1688
		2	$\overline{\alpha}$	2.1412	2.3654	$\overline{\theta}$	1.0005	0.9668
			Bias($\widehat{\alpha}$)	0.1412	0.3654	Bias($\widehat{\theta}$)	0.0005	−0.0332
			RMSE($\widehat{\alpha}$)	0.4112	0.5660	RMSE($\widehat{\theta}$)	0.0767	0.0854
	2	0.75	$\overline{\alpha}$	0.7942	0.8752	$\overline{\theta}$	2.0106	1.8402
			Bias($\widehat{\alpha}$)	0.0442	0.1252	Bias($\widehat{\theta}$)	0.0106	−0.1598
			RMSE($\widehat{\alpha}$)	0.1503	0.2031	RMSE($\widehat{\theta}$)	0.4057	0.4254
		1	$\overline{\alpha}$	1.0671	1.1782	$\overline{\theta}$	2.0029	1.8738
			Bias($\widehat{\alpha}$)	0.0671	0.1782	Bias($\widehat{\theta}$)	0.0029	−0.1262
			RMSE($\widehat{\alpha}$)	0.2035	0.2792	RMSE($\widehat{\theta}$)	0.3198	0.3416
		2	$\overline{\alpha}$	2.1282	2.3447	$\overline{\theta}$	1.9869	1.9207
			Bias($\widehat{\alpha}$)	0.1282	0.3447	Bias($\widehat{\theta}$)	−0.0131	−0.0793
			RMSE($\widehat{\alpha}$)	0.4032	0.5486	RMSE($\widehat{\theta}$)	0.1509	0.1738
50	1	0.75	$\overline{\alpha}$	0.7753	0.8578	$\overline{\theta}$	1.0051	0.9142
			Bias($\widehat{\alpha}$)	0.0253	0.1078	Bias($\widehat{\theta}$)	0.0051	−0.0858
			RMSE($\widehat{\alpha}$)	0.0990	0.1523	RMSE($\widehat{\theta}$)	0.1471	0.1653
		1	$\overline{\alpha}$	1.0245	1.1348	$\overline{\theta}$	0.9952	0.9265
			Bias($\widehat{\alpha}$)	0.0245	0.1348	Bias($\widehat{\theta}$)	−0.0048	−0.0735
			RMSE($\widehat{\alpha}$)	0.1267	0.1959	RMSE($\widehat{\theta}$)	0.1100	0.1316
		2	$\overline{\alpha}$	2.0495	2.2682	$\overline{\theta}$	0.9952	0.9598
			Bias($\widehat{\alpha}$)	0.0495	0.2682	Bias($\widehat{\theta}$)	−0.0048	−0.0402
			RMSE($\widehat{\alpha}$)	0.2551	0.3908	RMSE($\widehat{\theta}$)	0.0553	0.0696
	2	0.75	$\overline{\alpha}$	0.7722	0.8540	$\overline{\theta}$	2.0199	1.8468
			Bias($\widehat{\alpha}$)	0.0222	0.1040	Bias($\widehat{\theta}$)	0.0199	−0.1532
			RMSE($\widehat{\alpha}$)	0.0912	0.1442	RMSE($\widehat{\theta}$)	0.2936	0.3245
		1	$\overline{\alpha}$	1.0281	1.1389	$\overline{\theta}$	1.9890	1.8519
			Bias($\widehat{\alpha}$)	0.0281	0.1389	Bias($\widehat{\theta}$)	−0.0110	−0.1481
			RMSE($\widehat{\alpha}$)	0.1250	0.1966	RMSE($\widehat{\theta}$)	0.2157	0.2613
		2	$\overline{\alpha}$	2.0556	2.2803	$\overline{\theta}$	1.9988	1.9276
			Bias($\widehat{\alpha}$)	0.0556	0.2803	Bias($\widehat{\theta}$)	−0.0012	−0.0724
			RMSE($\widehat{\alpha}$)	0.2587	0.4011	RMSE($\widehat{\theta}$)	0.1100	0.1350
100	1	0.75	$\overline{\alpha}$	0.7631	0.8464	$\overline{\theta}$	1.0023	0.9102
			Bias($\widehat{\alpha}$)	0.0131	0.0964	Bias($\widehat{\theta}$)	0.0023	−0.0898
			RMSE($\widehat{\alpha}$)	0.0647	0.1209	RMSE($\widehat{\theta}$)	0.0993	0.1328
		1	$\overline{\alpha}$	1.0147	1.1271	$\overline{\theta}$	1.0004	0.9284
			Bias($\widehat{\alpha}$)	0.0147	0.1271	Bias($\widehat{\theta}$)	0.0004	−0.0716
			RMSE($\widehat{\alpha}$)	0.0859	0.1593	RMSE($\widehat{\theta}$)	0.0755	0.1038
		2	$\overline{\alpha}$	2.0420	2.2686	$\overline{\theta}$	0.9998	0.9627
			Bias($\widehat{\alpha}$)	0.0420	0.2686	Bias($\widehat{\theta}$)	−0.0002	−0.0373
			RMSE($\widehat{\alpha}$)	0.1827	0.3371	RMSE($\widehat{\theta}$)	0.0379	0.0545
	2	0.75	$\overline{\alpha}$	0.7617	0.8459	$\overline{\theta}$	1.9975	1.8107
			Bias($\widehat{\alpha}$)	0.0117	0.0959	Bias($\widehat{\theta}$)	−0.0025	−0.1893
			RMSE($\widehat{\alpha}$)	0.0658	0.1212	RMSE($\widehat{\theta}$)	0.2069	0.2772
		1	$\overline{\alpha}$	1.0179	1.1312	$\overline{\theta}$	1.9992	1.8563
			Bias($\widehat{\alpha}$)	0.0179	0.1312	Bias($\widehat{\theta}$)	−0.0008	−0.1437
			RMSE($\widehat{\alpha}$)	0.0837	0.1614	RMSE($\widehat{\theta}$)	0.1591	0.2141
		2	$\overline{\alpha}$	2.0287	2.2560	$\overline{\theta}$	1.9944	1.9192
			Bias($\widehat{\alpha}$)	0.0287	0.2560	Bias($\widehat{\theta}$)	−0.0056	−0.0808
			RMSE($\widehat{\alpha}$)	0.1720	0.3221	RMSE($\widehat{\theta}$)	0.0730	0.1107

Table 2: Estimates of parameters, bias and RMSE for different sample sizes for Case I.

n	θ	α	Parameter(α)			Parameter(θ)		
				ML	OBR		ML	OBR
25	1	0.75	$\overline{\alpha}$	0.5243	0.6119	$\overline{\theta}$	1.3713	1.0104
			Bias($\widehat{\alpha}$)	−0.2257	−0.1381	Bias($\widehat{\theta}$)	0.3713	0.0104
			RMSE($\widehat{\alpha}$)	0.2342	0.1588	RMSE($\widehat{\theta}$)	0.4695	0.2186
		1	$\overline{\alpha}$	0.6296	0.7430	$\overline{\theta}$	1.3592	1.0278
			Bias($\widehat{\alpha}$)	−0.3704	−0.2570	Bias($\widehat{\theta}$)	0.3592	0.0278
			RMSE($\widehat{\alpha}$)	0.3771	0.2723	RMSE($\widehat{\theta}$)	0.4174	0.1751
		2	$\overline{\alpha}$	0.9279	1.1280	$\overline{\theta}$	1.3321	1.0777
			Bias($\widehat{\alpha}$)	−1.0721	−0.8720	Bias($\widehat{\theta}$)	0.3321	0.0777
			RMSE($\widehat{\alpha}$)	1.0766	0.8813	RMSE($\widehat{\theta}$)	0.3486	0.1264
	2	0.75	$\overline{\alpha}$	0.5273	0.6148	$\overline{\theta}$	2.7406	2.0185
			Bias($\widehat{\alpha}$)	−0.2227	−0.1352	Bias($\widehat{\theta}$)	0.7406	0.0185
			RMSE($\widehat{\alpha}$)	0.2307	0.1550	RMSE($\widehat{\theta}$)	0.9204	0.4217
		1	$\overline{\alpha}$	0.6331	0.7475	$\overline{\theta}$	2.7256	2.0636
			Bias($\widehat{\alpha}$)	−0.3669	−0.2525	Bias($\widehat{\theta}$)	0.7256	0.0636
			RMSE($\widehat{\alpha}$)	0.3736	0.2676	RMSE($\widehat{\theta}$)	0.8336	0.3359
		2	$\overline{\alpha}$	0.9323	1.1327	$\overline{\theta}$	2.6779	2.1652
			Bias($\widehat{\alpha}$)	−1.0677	−0.8673	Bias($\widehat{\theta}$)	0.6779	0.1652
			RMSE($\widehat{\alpha}$)	1.0726	0.8773	RMSE($\widehat{\theta}$)	0.7115	0.2578
50	1	0.75	$\overline{\alpha}$	0.4948	0.5619	$\overline{\theta}$	1.4738	1.0876
			Bias($\widehat{\alpha}$)	−0.2552	−0.1881	Bias($\widehat{\theta}$)	0.4738	0.0876
			RMSE($\widehat{\alpha}$)	0.2584	0.1947	RMSE($\widehat{\theta}$)	0.5188	0.1880
		1	$\overline{\alpha}$	0.5922	0.6792	$\overline{\theta}$	1.4617	1.1146
			Bias($\widehat{\alpha}$)	−0.4078	−0.3208	Bias($\widehat{\theta}$)	0.4617	0.1146
			RMSE($\widehat{\alpha}$)	0.4103	0.3258	RMSE($\widehat{\theta}$)	0.4886	0.1719
		2	$\overline{\alpha}$	0.8816	1.0447	$\overline{\theta}$	1.4194	1.1690
			Bias($\widehat{\alpha}$)	−1.1184	−0.9553	Bias($\widehat{\theta}$)	0.4194	0.1690
			RMSE($\widehat{\alpha}$)	1.1201	0.9590	RMSE($\widehat{\theta}$)	0.4269	0.1844
	2	0.75	$\overline{\alpha}$	0.4952	0.5628	$\overline{\theta}$	2.9648	2.1899
			Bias($\widehat{\alpha}$)	−0.2548	−0.1872	Bias($\widehat{\theta}$)	0.9648	0.1899
			RMSE($\widehat{\alpha}$)	0.2582	0.1941	RMSE($\widehat{\theta}$)	1.0550	0.3858
		1	$\overline{\alpha}$	0.5941	0.6820	$\overline{\theta}$	2.9364	2.2459
			Bias($\widehat{\alpha}$)	−0.4059	−0.3180	Bias($\widehat{\theta}$)	0.9364	0.2459
			RMSE($\widehat{\alpha}$)	0.4086	0.3231	RMSE($\widehat{\theta}$)	0.9911	0.3639
		2	$\overline{\alpha}$	0.8813	1.0444	$\overline{\theta}$	2.8429	2.3387
			Bias($\widehat{\alpha}$)	−1.1187	−0.9556	Bias($\widehat{\theta}$)	0.8429	0.3387
			RMSE($\widehat{\alpha}$)	1.1206	0.9596	RMSE($\widehat{\theta}$)	0.8586	0.3711
100	1	0.75	$\overline{\alpha}$	0.4859	0.5522	$\overline{\theta}$	1.4969	1.0990
			Bias($\widehat{\alpha}$)	−0.2641	−0.1978	Bias($\widehat{\theta}$)	0.4969	0.0990
			RMSE($\widehat{\alpha}$)	0.2656	0.2009	RMSE($\widehat{\theta}$)	0.5205	0.1547
		1	$\overline{\alpha}$	0.5842	0.6704	$\overline{\theta}$	1.4783	1.1255
			Bias($\widehat{\alpha}$)	−0.4158	−0.3296	Bias($\widehat{\theta}$)	0.4783	0.1255
			RMSE($\widehat{\alpha}$)	0.4170	0.3321	RMSE($\widehat{\theta}$)	0.4925	0.1573
		2	$\overline{\alpha}$	0.8676	1.0268	$\overline{\theta}$	1.4262	1.1696
			Bias($\widehat{\alpha}$)	−1.1324	−0.9732	Bias($\widehat{\theta}$)	0.4262	0.1696
			RMSE($\widehat{\alpha}$)	1.1333	0.9752	RMSE($\widehat{\theta}$)	0.4301	0.1783
	2	0.75	$\overline{\alpha}$	0.4863	0.5525	$\overline{\theta}$	3.0158	2.2161
			Bias($\widehat{\alpha}$)	−0.2637	−0.1975	Bias($\widehat{\theta}$)	1.0158	0.2161
			RMSE($\widehat{\alpha}$)	0.2653	0.2006	RMSE($\widehat{\theta}$)	1.0648	0.3280
		1	$\overline{\alpha}$	0.5816	0.6675	$\overline{\theta}$	2.9503	2.2433
			Bias($\widehat{\alpha}$)	−0.4184	−0.3325	Bias($\widehat{\theta}$)	0.9503	0.2433
			RMSE($\widehat{\alpha}$)	0.4196	0.3349	RMSE($\widehat{\theta}$)	0.9773	0.3060
		2	$\overline{\alpha}$	0.8687	1.0303	$\overline{\theta}$	2.8523	2.3440
			Bias($\widehat{\alpha}$)	−1.1313	−0.9697	Bias($\widehat{\theta}$)	0.8523	0.3440
			RMSE($\widehat{\alpha}$)	1.1321	0.9715	RMSE($\widehat{\theta}$)	0.8595	0.3591

Table 3: Estimates of parameters, bias and RMSE for different sample sizes for Case II.

n	θ	α	Parameter(α)			Parameter(θ)		
				ML	OBR		ML	OBR
25	1	0.75	$\overline{\alpha}$	0.5872	0.6849	$\overline{\theta}$	0.6495	0.6564
			Bias($\widehat{\alpha}$)	−0.1628	−0.0651	Bias($\widehat{\theta}$)	−0.3505	−0.3436
			RMSE($\widehat{\alpha}$)	0.1780	0.1194	RMSE($\widehat{\theta}$)	0.3800	0.3772
		1	$\overline{\alpha}$	0.6906	0.8318	$\overline{\theta}$	0.7023	0.7167
			Bias($\widehat{\alpha}$)	−0.3094	−0.1682	Bias($\widehat{\theta}$)	−0.2977	−0.2833
			RMSE($\widehat{\alpha}$)	0.3191	0.2047	RMSE($\widehat{\theta}$)	0.3205	0.3104
		2	$\overline{\alpha}$	0.9338	1.2509	$\overline{\theta}$	0.7824	0.8082
			Bias($\widehat{\alpha}$)	−1.0662	−0.7491	Bias($\widehat{\theta}$)	−0.2176	−0.1918
			RMSE($\widehat{\alpha}$)	1.0690	0.7619	RMSE($\widehat{\theta}$)	0.2279	0.2050
	2	0.75	$\overline{\alpha}$	0.5722	0.6708	$\overline{\theta}$	1.3063	1.3240
			Bias($\widehat{\alpha}$)	−0.1778	−0.0792	Bias($\widehat{\theta}$)	−0.6937	−0.6760
			RMSE($\widehat{\alpha}$)	0.1890	0.1196	RMSE($\widehat{\theta}$)	0.7534	0.7460
		1	$\overline{\alpha}$	0.6686	0.8115	$\overline{\theta}$	1.3783	1.4127
			Bias($\widehat{\alpha}$)	−0.3314	−0.1885	Bias($\widehat{\theta}$)	−0.6217	−0.5873
			RMSE($\widehat{\alpha}$)	0.3390	0.2182	RMSE($\widehat{\theta}$)	0.6625	0.6358
		2	$\overline{\alpha}$	0.8923	1.2055	$\overline{\theta}$	1.5442	1.5966
			Bias($\widehat{\alpha}$)	−1.1077	−0.7945	Bias($\widehat{\theta}$)	−0.4558	−0.4034
			RMSE($\widehat{\alpha}$)	1.1098	0.8049	RMSE($\widehat{\theta}$)	0.4779	0.4320
50	1	0.75	$\overline{\alpha}$	0.5378	0.6285	$\overline{\theta}$	0.5867	0.6050
			Bias($\widehat{\alpha}$)	−0.2122	−0.1215	Bias($\widehat{\theta}$)	−0.4133	−0.3950
			RMSE($\widehat{\alpha}$)	0.2165	0.1364	RMSE($\widehat{\theta}$)	0.4245	0.4088
		1	$\overline{\alpha}$	0.6226	0.7527	$\overline{\theta}$	0.6420	0.6695
			Bias($\widehat{\alpha}$)	−0.3774	−0.2473	Bias($\widehat{\theta}$)	−0.3580	−0.3305
			RMSE($\widehat{\alpha}$)	0.3800	0.2562	RMSE($\widehat{\theta}$)	0.3669	0.3419
		2	$\overline{\alpha}$	0.8051	1.0570	$\overline{\theta}$	0.7423	0.7843
			Bias($\widehat{\alpha}$)	−1.1949	−0.9430	Bias($\widehat{\theta}$)	−0.2577	−0.2157
			RMSE($\widehat{\alpha}$)	1.1956	0.9453	RMSE($\widehat{\theta}$)	0.2621	0.2221
	2	0.75	$\overline{\alpha}$	0.5235	0.6171	$\overline{\theta}$	1.1691	1.2111
			Bias($\widehat{\alpha}$)	−0.2265	−0.1329	Bias($\widehat{\theta}$)	−0.8309	−0.7889
			RMSE($\widehat{\alpha}$)	0.2301	0.1457	RMSE($\widehat{\theta}$)	0.8524	0.8152
		1	$\overline{\alpha}$	0.6013	0.7334	$\overline{\theta}$	1.2692	1.3312
			Bias($\widehat{\alpha}$)	−0.3987	−0.2666	Bias($\widehat{\theta}$)	−0.7308	−0.6688
			RMSE($\widehat{\alpha}$)	0.4008	0.2741	RMSE($\widehat{\theta}$)	0.7462	0.6889
		2	$\overline{\alpha}$	0.7680	1.0162	$\overline{\theta}$	1.4733	1.5639
			Bias($\widehat{\alpha}$)	−1.2320	−0.9838	Bias($\widehat{\theta}$)	−0.5267	−0.4361
			RMSE($\widehat{\alpha}$)	1.2326	0.9856	RMSE($\widehat{\theta}$)	0.5351	0.4489
100	1	0.75	$\overline{\alpha}$	0.5347	0.6231	$\overline{\theta}$	0.5881	0.6071
			Bias($\widehat{\alpha}$)	−0.2153	−0.1269	Bias($\widehat{\theta}$)	−0.4119	−0.3929
			RMSE($\widehat{\alpha}$)	0.2176	0.1346	RMSE($\widehat{\theta}$)	0.4173	0.3995
		1	$\overline{\alpha}$	0.6210	0.7494	$\overline{\theta}$	0.6386	0.6686
			Bias($\widehat{\alpha}$)	−0.3790	−0.2506	Bias($\widehat{\theta}$)	−0.3614	−0.3314
			RMSE($\widehat{\alpha}$)	0.3802	0.2547	RMSE($\widehat{\theta}$)	0.3659	0.3371
		2	$\overline{\alpha}$	0.8044	1.0536	$\overline{\theta}$	0.7414	0.7869
			Bias($\widehat{\alpha}$)	−1.1956	−0.9464	Bias($\widehat{\theta}$)	−0.2586	−0.2131
			RMSE($\widehat{\alpha}$)	1.1959	0.9475	RMSE($\widehat{\theta}$)	0.2608	0.2164
	2	0.75	$\overline{\alpha}$	0.5217	0.6134	$\overline{\theta}$	1.1642	1.2106
			Bias($\widehat{\alpha}$)	−0.2283	−0.1366	Bias($\widehat{\theta}$)	−0.8358	−0.7894
			RMSE($\widehat{\alpha}$)	0.2301	0.1428	RMSE($\widehat{\theta}$)	0.8465	0.8029
		1	$\overline{\alpha}$	0.5996	0.7300	$\overline{\theta}$	1.2627	1.3292
			Bias($\widehat{\alpha}$)	−0.4004	−0.2700	Bias($\widehat{\theta}$)	−0.7373	−0.6708
			RMSE($\widehat{\alpha}$)	0.4014	0.2736	RMSE($\widehat{\theta}$)	0.7451	0.6810
		2	$\overline{\alpha}$	0.7640	1.0079	$\overline{\theta}$	1.4676	1.5634
			Bias($\widehat{\alpha}$)	−1.2360	−0.9921	Bias($\widehat{\theta}$)	−0.5324	−0.4366
			RMSE($\widehat{\alpha}$)	1.2363	0.9931	RMSE($\widehat{\theta}$)	0.5365	0.4426

5. REAL DATA EXAMPLE

In this section, we will analyze the data set used by Cooray and Ananda [3]. This data set contains the stress-rupture life of kevlar 49/ epoxy strands failure at 90% stress levels. The data set is given below (Andrews and Herzberg [1], Barlow *et al.* [2]).

Table 4: The failure times in hours.

0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.04	0.05	0.06	0.07	0.07	0.08
0.09	0.09	0.10	0.10	0.11	0.11	0.12	0.13	0.18	0.19	0.20	0.23	0.24
0.24	0.29	0.34	0.35	0.36	0.38	0.40	0.42	0.43	0.52	0.54	0.56	0.60
0.60	0.63	0.65	0.67	0.68	0.72	0.72	0.72	0.73	0.79	0.79	0.80	0.80
0.83	0.85	0.90	0.92	0.95	0.99	1.00	1.01	1.02	1.03	1.05	1.10	1.10
1.11	1.15	1.18	1.20	1.29	1.31	1.33	1.34	1.40	1.43	1.45	1.50	1.51
1.52	1.53	1.54	1.54	1.55	1.58	1.60	1.63	1.64	1.80	1.80	1.81	2.02
2.05	2.14	2.17	2.33	3.03	3.03	3.34	4.20	4.69	7.89			

Assume that this data set has a GHN distribution with the unknown parameters α and θ. We use the OBR estimation method to obtain the estimates for α and θ for the failure time data set. We also find the ML estimates for these parameters. Table 5 gives the summary of the estimates, standard error (SE) and the 95% confidence interval for the parameters of GHN distribution. The confidence intervals of the estimates are computed using the intervals given in Cooray and Ananda [3]. In their paper, they use the expected Fisher information matrix. For the ML estimators, we also use the expected Fisher information matrix to compute the standard errors and the confidence intervals. For the OBR estimators, we use the asymptotic covariance matrix given in Victoria-Feser and Ronchetti [9] to compute the standard errors and the confidence intervals.

Table 5: ML and OBR ($b = 2$) parameter estimates for the failure time data set.

Method	$\widehat{\alpha}$	SE	95% confidence interval of α	$\widehat{\theta}$	SE	95% confidence interval of θ
ML	0.7108	0.0584	(0.5964, 0.8252)	1.2238	0.1317	(0.9657, 1.4819)
OBR	0.7811	0.0574	(0.6685, 0.8937)	1.0540	0.0794	(0.8983, 1.2097)

Figure 4 (a) shows the boxplot of the failure time data set. After some preliminary examination of the data set, we can see from the boxplot that there may be four potential outliers in the data set. We give the histogram of the data set with the fitted densities obtained from ML and OBR estimates in Figure 4 (b).

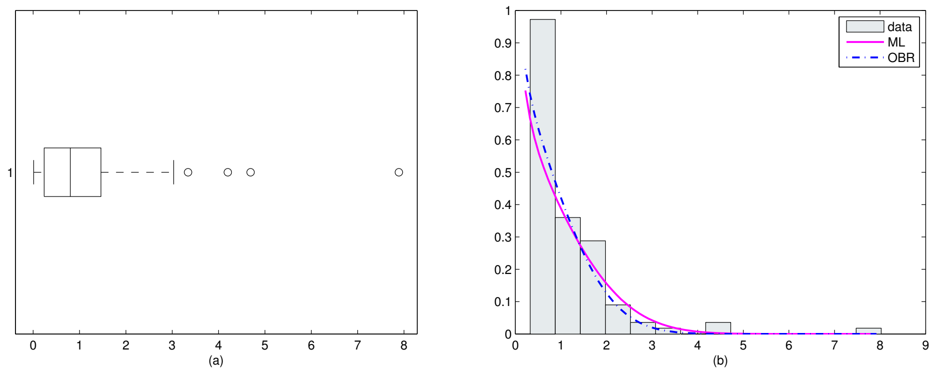

Figure 4: **(a)** Boxplot of the failure time data set; **(b)** Histogram with the fitted densities obtained from ML and OBR estimation methods.

We also give the Q-Q plots of the fitted distribution obtained from ML and OBR estimation methods in Figure 5. From this figure, we can see that the OBR estimates are not badly affected by the outliers. But we can clearly see that the ML estimators are influenced by the outliers. Furthermore, the Q-Q plot of the fitted distribution obtained from OBR estimators is well fitted contrary to the ML estimators.

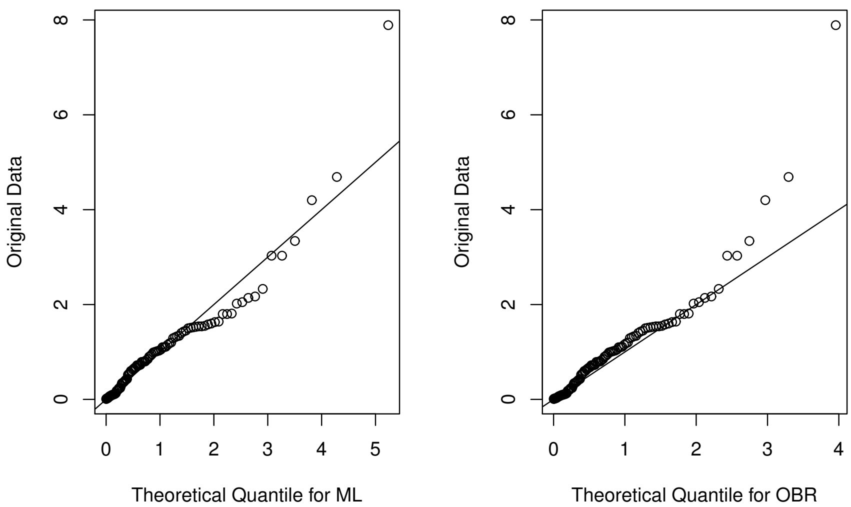

Figure 5: Q-Q plots for the failure time data set estimated by the ML and OBR estimation methods.

6. CONCLUSIONS

In this paper, we have proposed robust estimators for the parameters of the GHN distribution, which is proposed by Cooray and Ananda [3] as a flexible alternative lifetime distribution, using the OBR estimation method. Our limited simulation study has shown that the ML estimators are influenced by the outliers, but on the other hand, the OBR estimators are resistant to the outliers. The same results have been recorded from the real data example. Therefore, we can conclude that for this distribution the OBR estimators can be used as alternative estimators to the ML estimators.

ACKNOWLEDGMENTS

The authors would like to thank the referees and the associate editor for their constructive comments that led to substantial improvement of the paper.

REFERENCES

[1] Andrews, D.F. and Herzberg, A.M. (1985). *Data: A Collection of Problems from Many Fields for the Student and Research Worker*, New York: Springer Series in Statistics.

[2] Barlow, R.E.; Toland, R.H. and Freeman, T. (1984). *A Bayesian Analysis of Stress-rupture Life of Kevlar 49/epoxy Spherical Pressure Vessels*. In "Proceedings of the Canadian Conference in Applied Statistics, 1981" (T.D. Dwivedi, Ed.), Marcel Dekker, New York.

[3] Cooray, K. and Ananda, M.M.A. (2008). A generalization of the half-normal distribution with applications to lifetime data, *Communications in Statistics – Theory and Methods*, **37**, 1323–1337.

[4] Doğru, F.Z. and Arslan, O. (2016). Optimal B-robust estimators for the parameters of the Burr XII distribution, *Journal of Statistical Computation and Simulation*, **86**(6), 1133–1149.

[5] Hampel, F.R.; Ronchetti, E.M.; Rousseeuw, P.J. and Stahel, W.A. (1986). *Robust Statistics: The Approach Based on Influence Functions*, Wiley, New York.

[6] Huber, P.J. (1964). Robust estimation of a location parameter, *The Annals of Mathematical Statistics*, **35**, 73–101.

[7] Stacy, E.W. (1962). A generalization of the Gamma distribution, *The Annals of Mathematical Statistics*, **33**, 1187–1192.

[8] Victoria-Feser, M.P. (1993). *Robust methods for personal income distribution models*, Ph.D thesis in Econometrics and Statistics, Faculty of Economic and Social Sciences, University of Geneva, Switzerland. (GE-archives).

[9] Victoria-Feser, M.P. and Ronchetti, E. (1994). Robust methods for personal-income distribution models, *The Canadian Journal of Statistics*, **22**(3), 247–258.