A Consistent Method of Estimation for Three-parameter Generalized Exponential Distribution

Kiran Prajapat*[1], Sharmishtha Mitra[1], and Debasis Kundu[1]

[1]Department of Mathematics, Indian Institute of Technology Kanpur, Kanpur 208016, India

January 6, 2021

Abstract

In this article, we provide a consistent method of estimation for the parameters of a three-parameter generalized exponential distribution which avoids the problem of unbounded likelihood function. The method is based on a maximum likelihood estimation of the shape parameter, which uses location and scale invariant statistic, originally proposed by Nagatsuka et al. (2013). It has been shown that the estimators are unique and consistent for the entire range of the parameter space. We also present a Monte-Carlo simulation study along with the comparisons with some prominent estimation methods in terms of the bias and root mean square error. For the illustration purpose, the data analysis of a real lifetime data set has been reported.

Key words: Invariant statistic; generalized exponential distribution; likelihood; estimation; maximum likelihood estimation; consistency.

*Corresponding author: Kiran Prajapat, kiranp@iitk.ac.in

1 Introduction

The three-parameter gamma and three-parameter Weibull are the most popular distributions for analyzing lifetime data or skewed data. In both the distributions, the three parameters represent location, scale and shape, and because of these parameters both the distributions have quite a bit of flexibility for analyzing skewed data. Unfortunately both distributions also have certain drawbacks. Few of the these drawbacks, we will discuss later in this section. A three-parameter generalized exponential (GE) distribution is a particular case of the exponentiated Weibull distribution when the location parameter is not present. The exponentiated Weibull distribution is originally proposed by Mudholkar et al. (1995). It has been shown that the GE model can be used as an alternative to the gamma model or the Weibull model in many situations, see Gupta and Kundu (1999) , Gupta and Kundu (2001) and Gupta and Kundu (2007) for more details.

Suppose a random variable from three-parameter GE distribution has the following cumulative distribution function (CDF)

$$F(x;\alpha,\beta,\gamma) = \begin{cases} 0, & if x < \gamma \\ \left(1 - exp(-\frac{x-\gamma}{\alpha})\right)^{\beta}, & if\ x > \gamma, \end{cases} \quad \alpha > 0, \beta > 0, \gamma \in \mathbb{R}. \tag{1}$$

Let the three-parameter GE distribution be denoted by $GE(\alpha,\beta,\gamma)$, where α, β and γ denote the scale, shape and location parameters, respectively. Here all the three parameters are unknown. Let us denote the corresponding probability density function (PDF) by $f(.;\alpha,\beta,\gamma)$, then

$$f(x;\alpha,\beta,\gamma) = \begin{cases} \frac{\beta}{\alpha} exp(-\frac{x-\gamma}{\alpha})\left(1 - exp(-\frac{x-\gamma}{\alpha})\right)^{\beta-1}, & if\ x > \gamma \\ 0, & \text{otherwise}, \end{cases} \quad \alpha > 0, \beta > 0, \gamma \in \mathbb{R}. \tag{2}$$

It is well known that for the similar three-parameter distributions, e.g., lognormal, gamma, Weibull and inverse Gaussian, etc., in which the location parameter is unknown, the regularity conditions are not satisfied for the estimation method of the well-known maximum likelihood (ML) since the support of the PDF depends on the unknown location parameter and therefore the ML estimation may face problems. Most of the time, the maximum likelihood estimator (MLE) do not exist for the entire range of the parameter space because it does not exist for some range of the parameter space. In such cases, the likelihood becomes unbounded. This is one of the very well known drawbacks mentioned before. Many authors like Cohen and Whitten (1982), Cohen and Whitten (1980), Cohen et al. (1984), Hall and Wang (2005), Nagatsuka et al. (2013), Nagatsuka and Balakrishnan (2012), and Nagatsuka and Balakrishnan (2013) and several others addressed this issue.

The same problem arises in the case of three-parameter GE distribution, including location parameter, as well. Gupta and Kundu (1999) discussed the ML estimation for three-parameter GE distribution in detail. When the shape parameter $\beta < 1$, they have shown that the MLE does not exist since the likelihood function becomes unbounded as the location parameter γ tends to the smallest observation in observed sample, elsewhere for $\beta > 2$, all the asymptotic results have been presented.

Among all the existing estimation methods, two estimation methods proposed by Nagatsuka et al. (2013) and Nagatsuka and Balakrishnan (2012) which are abbreviated as the location scale parameter free (LSPF) method and location parameter free (LPF) method, respectively, are the most recently proposed methods. These methods provide the estimators that exist uniquely and that are consistent for the complete range of the parameter space.

In this study, we discuss the LSPF method of estimation in detail for the three-parameter GE distribution and present the bias and root mean square error (RMSE) of the estimator through a Monte-Carlo simulation. We also do some comparisons with some prominent estimation methods in terms of the biases and RMSEs of the estimators. We also present a real lifetime data set to illustrate the LSPF method.

Rest of the article is organized as follows. Section 2 includes the estimation procedure of the proposed LSPF method for the three-parameter GE distribution. In this Section, we also present some properties of the estimators, such as the existence, uniqueness and consistency. In Section 3, we perform a Monte-Carlo simulation study for evaluation of the LSPF method and for comparisons with some prominent methods. The LSPF method has been illustrated through a real lifetime data set in Section 4, then concluding remarks in Section 5.

2 Estimation Procedure

Suppose that $X_1, X_2, \ldots, X_n$ are n independent and identically distributed random variables from three-parameter GE distribution with the following common CDF given in the equation (1) and $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ are the order statistics of $X_1, X_2, \ldots, X_n$.

Throughout the paper, we have the following two assumptions to hold:

1. $n \geq 3$,
2. $X_{(i)} \neq X_{(j)}, \forall\ i \neq j, i = 1, 2, \ldots, n, j = 1, 2, \ldots, n$.

Let us consider the following statistic in order to start the estimation procedure:

$$W_{(i)} = \frac{X_{(i)} - X_{(1)}}{X_{(n)} - X_{(1)}},\ \ i = 1, 2, \ldots, n. \tag{3}$$

The probability distribution of $W_{(i)}$'s are free from the location and scale parameter. It may be noted that $W_{(1)} = 0$ and $W_{(n)} = 1$. Based on the transformed data $W_{(1)}, W_{(2)}, \ldots, W_{(n)}$, whose joint probability distribution depends only on shape parameter, the shape parameter is estimated. It is further used in the estimation of other parameters. Subsection 2.1 and Subsection 2.2 include a detailed discussion on the estimation of the shape parameter and estimation of the location and scale parameters, respectively.

2.1 Estimation of Shape Parameter

We estimate the shape parameter β based on the transformed random variables $W_{(1)}, W_{(2)}, \ldots, W_{(n)}$. The likelihood function of β based on the transformed data might be bounded as it does not depend on the location parameter. We obtain the expression for the likelihood function of β before we check its boundedness. Let $w_1, w_2, \ldots, w_n$ be the realizations of $W_{(1)}, W_{(2)}, \ldots, W_{(n)}$, respectively. Note that we must have $w_1 = 0$ and $w_n = 1$.

Theorem 2.1. *Likelihood function of β given $w_2, w_3, \ldots, w_{n-1}$ is given by*

$$\ell_w(\beta|w_2, \ldots, w_{n-1}) = n!\beta^n \int_0^\infty \int_0^\infty v^{n-2} e^{-\sum_{i=1}^{n}(u+vw_i)} \prod_{i=1}^{n} \left(1 - e^{-(u+vw_i)}\right)^{\beta-1} \, dvdu, \;\; \beta > 0, \tag{4}$$

with $0 \leq w_2 \leq \cdots \leq w_{n-1} \leq 1, w_1 = 0, w_n = 1.$

Proof. See Appendix A.1. □

Now, we show that the likelihood function in Theorem 2.1 is a bounded function of β. For the sake of simplicity, we will use sometime the notation $\ell_w(\beta)$ for $\ell_w(\beta|w_2, \ldots, w_{n-1})$.

Proposition 2.1. *Likelihood function $\ell_w(\beta)$ is a bounded function of parameter β with* $0 \leq w_2 \leq \cdots \leq w_{n-1} \leq 1, w_1 = 0, w_n = 1.$

Proof. See Appendix A.2. □

We rewrite the likelihood function $\ell_w(\beta)$ as following

$$\ell_w(\beta) = n! \int_0^\infty \int_0^\infty e^{h_w(\beta; u,v)} \, dvdu, \tag{5}$$

where

$$h_w(\beta; u, v) = n \log \beta + (n-2) \log v - \sum_{i=1}^{n} (u + vw_i) + (\beta - 1) \sum_{i=1}^{n} \log(1 - e^{-(u+vw_i)}). \quad (6)$$

In proposition 2.2, it has been shown that the likelihood function $\ell_w(\beta)$ is defferentiable with respect to β.

Proposition 2.2. *For $\beta > 0$, the likelihood function $\ell_w(\beta)$, with $0 \leq w_2 \leq \cdots \leq w_{n-1} \leq 1, w_1 = 0, w_n = 1$, can be differentiated with respect to β and its derivative is given by*

$$\begin{aligned} \ell'_w(\beta) =& n!\beta^n \int_0^\infty \int_0^\infty v^{n-2} e^{-\sum_{i=1}^n (u+vw_i)} \left\{ \prod_{i=1}^{n} \left(1 - e^{-(u+vw_i)}\right)^{\beta-1} \right\} \\ &\times \left\{ \frac{n}{\beta} + \sum_{i=1}^{n} \log(1 - e^{-(u+vw_i)}) \right\} dvdu, \ \beta > 0, \end{aligned}$$

Proof. See Appendix A.3. □

We present two main results in the upcoming theorems where first result is regarding the unique maximum of the likelihood function $\ell_w(\beta)$ and another is about the consistency of the unique maximum. First we focus on the uni-modality for the likelihood function of the shape parameter given the transformed data $0 \leq w_2 \leq \cdots \leq w_{n-1} \leq 1$.

Theorem 2.2. *For $\beta > 0$ and any given $0 \leq w_2 \leq \cdots \leq w_{n-1} \leq 1$, the likelihood equation $\ell'_w(\beta) = 0$ always has an unique solution and it maximizes the likelihood function $\ell_w(\beta)$.*

Proof. See Appendix A.4. □

Now, we have an unique estimator of β, denoted by $\widehat{\beta}_w$, from the theorem 2.2. To show the consistency of the estimator $\widehat{\beta}_w$, we need a lemma which is presented in Appendix A.5. Its proof is quite similar to lemma 1 of both Nagatsuka et al. (2013) and Nagatsuka et al. (2014), hence it is omitted.

Theorem 2.3. *Estimator $\widehat{\beta}_w$ is a consistent estimator for $\beta > 0$.*

Proof. Theorem 3.7 of Chapter 6 in Lehmann and Casella (2006) along with the lemma reported in Appendix A.5 prove the result. □

Once we have $\widehat{\beta}_w$ based on the location and scale invariant data $w_2, w_3, \ldots, w_{n-1}$, we use the estimator $\widehat{\beta}_w$ for estimating the remaining location and scale parameters elaborated below in Section 2.2.

2.2 Estimation of Location and Scale Parameters

2.2.1 Estimation Procedure

Estimation of location and scale parameters is a sequential procedure in which we estimate the location parameter γ first and then we use it in the estimation of the scale parameter α. We reuse these estimates of γ and α in order to get better estimates further. We discuss the complete procedure below in detail. The likelihood function of (α, β, γ) given $(x_1, x_2, \ldots, x_n)$ is given by

$$\ell(\alpha, \beta, \gamma) = \Big(\frac{\beta}{\alpha}\Big)^n e^{-\sum_{i=1}^n \frac{x_i-\gamma}{\alpha}} \prod_{i=1}^n \big(1 - e^{-\frac{x_i-\gamma}{\alpha}}\big)^{\beta-1} \qquad \alpha > 0, \beta > 0, \gamma < x_{(1)}. \tag{7}$$

Suppose that we already have $\widehat{\beta}_w$ using Section 2.1. Now, assuming that the shape parameter is known, proceed further to estimate γ and α using $\widehat{\beta}_w$ in place of β in equation (7). In this case, it is reasonable to take $X_{(1)}$ as an estimator of the location parameter γ. Let us denote it by $\widehat{\gamma}_{init}$ *i.e.*, $\widehat{\gamma}_{init} = X_{(1)}$. For a known $0 < \beta < 1$, it is not feasible to maximize the likelihood function in order to use the ML estimation because the likelihood is unbounded when $\gamma = \widehat{\gamma}_{init}$. But, once $\widehat{\gamma}_{init}$ is obtained, by shifting each observation (to left or right according to the sign of γ) with length of $\widehat{\gamma}_{init}$ as suggested by Pasari and Dikshit (2014), we can avoid the problem of unbounded likelihood. Let the shifted data points be $t_2, t_3, \ldots, t_n$, then $t_i = x_i - \widehat{\gamma}_{init} = x_i - x_{(1)},\ i = 2, 3, \ldots, n$. For known $\widehat{\beta}_w$ and $\widehat{\gamma}_{init}$, the likelihood function of α based on the $(n-1)$ shifted data points $(t_2, t_3, \ldots, t_n)$ is given by

$$\ell_t(\alpha) = \big(\frac{\widehat{\beta}_w}{\alpha}\big)^{n-1} e^{-\sum_{i=2}^n \frac{t_i}{\alpha}} \prod_{i=2}^{n} \big(1 - e^{-\frac{t_i}{\alpha}}\big)^{\widehat{\beta}_w - 1}$$
$$\implies \log \ell_t(\alpha) \propto -(n-1)\log\alpha - \sum_{i=2}^{n} \frac{x_i - \widehat{\gamma}_{init}}{\alpha} + (\widehat{\beta}_w - 1)\sum_{i=2}^{n} \log\big(1 - e^{-\frac{x_i - \widehat{\gamma}_{init}}{\alpha}}\big). \tag{8}$$

Now, we estimate α by maximizing the log-likelihood function $\log \ell_t(\alpha)$ with respect to α. Solving the equation $\frac{\partial}{\partial\alpha}\log(\ell_t(\alpha)) = 0$ is equivalent to solve the following equation

$$h_1(\alpha, \widehat{\beta}_w, \widehat{\gamma}_{init}) = \alpha, \tag{9}$$

where $h_1(\alpha, \widehat{\beta}_w, \widehat{\gamma}_{init}) = \frac{1}{n-1}\sum_{i=2}^{n}(x_i - \widehat{\gamma}_{init})\Big(\frac{1-\widehat{\beta}_w e^{-\frac{x_i - \widehat{\gamma}_{init}}{\alpha}}}{1-e^{-\frac{x_i - \widehat{\gamma}_{init}}{\alpha}}}\Big)$. Here $\log(\ell_t(\alpha))$ is an uni-modal function which we prove later in Appendix A.6. Hence, the equation (9) has unique solution which maximizes the log-likelihood and equation (9) is easily solvable by any of the numerical iterative methods. Let us denote the unique root of the equation (9) by $\widehat{\alpha}_{init}$.

Here, it may be noted that $\widehat{\gamma}_{init}$ is a biased estimator of the parameter γ because $E(X_{(1)}) = \gamma + \alpha\int_0^\infty (1 - F_Z(y;\beta))^n \, dy$ with $Z \sim GE(1, \beta, 0)$. Therefore, it is feasible to consider a bias-corrected estimator of γ, denoted by $\widehat{\gamma}_w$, as $X_{(1)} - \widehat{\alpha}_{init}\int_0^\infty (1 - F_Z(y;\widehat{\beta}_w))^n \, dy$. It is also feasible to update the estimator of α using $\widehat{\gamma}_w$. Note that, for known $\beta = \widehat{\beta}_w$ and $\gamma = \widehat{\gamma}_w$, α can be directly estimated from sample $(x_1, x_2, \ldots, x_n)$ by maximizing the likelihood function of α given $x_1, x_2, \ldots, x_n$ as now there is no problem of unbounded likelihood function. Hence, the newly updated estimator of α, denoted by $\widehat{\alpha}_w$, can be obtained along the same line as obtained earlier by solving the $h_2(\alpha, \widehat{\beta}_w, \widehat{\gamma}_w) = \alpha$ with $h_2(\alpha, \widehat{\beta}_w, \widehat{\gamma}_w) = \frac{1}{n}\sum_{i=1}^n (x_i - \widehat{\gamma}_w)\big(\frac{1-\widehat{\beta}_w e^{-\frac{x_i - \widehat{\gamma}_w}{\alpha}}}{1-e^{-\frac{x_i - \widehat{\gamma}_w}{\alpha}}}\big)$. It can be proved that $h_2(\alpha, \widehat{\beta}_w, \widehat{\gamma}_w) = \alpha$ has unique solution which maximizes the likelihood function of α.

For the brief overview of this section the complete estimation procedure discussed above can be summarized in the following steps:

Step 1: Obtain $\widehat{\beta}_w$.

Step 2: Take $\widehat{\gamma}_{init} = X_{(1)}$ and obtain $\widehat{\alpha}_{init}$ by solving $h_1(\alpha, \widehat{\beta}_w, \widehat{\gamma}_{init}) = \alpha$.

Step 3: Use $\widehat{\gamma}_{init}$ and $\widehat{\alpha}_{init}$ of Step 2, to find bias-corrected $\widehat{\gamma}_w$ as following:

$$\widehat{\gamma}_w = X_{(1)} - \widehat{\alpha}_{init} \int_0^\infty (1 - (1 - e^{-y})^{\widehat{\beta}_w})^n \, dy,$$

and obtain $\widehat{\alpha}_w$ by solving $h_2(\alpha, \widehat{\beta}_w, \widehat{\gamma}_w) = \alpha$.

2.2.2 Uniqueness and Consistency

Here, we try to show the following two properties of provided estimators of location and scale parameters:

1. $\widehat{\gamma}_w$ and $\widehat{\alpha}_w$ exist uniquely.
2. $\widehat{\gamma}_w$ and $\widehat{\alpha}_w$ are consistent estimators of γ and α, respectively.

It can be seen that $\widehat{\gamma}_{init}$ is unique for a given sample $(x_1, x_2, \dots, x_n)$, whereas the uniqueness of estimator $\widehat{\alpha}_{init}$ follows from the uni-modality of $\ell_t(\alpha)$. By Theorem 2.2, $\widehat{\beta}_w$ is also unique and so bias-corrected estimators $\widehat{\gamma}_w$ and $\widehat{\alpha}_w$.

In order to show the consistency of $\widehat{\gamma}_{init}$, let us consider the following probability for an arbitrary $\epsilon > 0$:

$$Pr(|\widehat{\gamma}_{init} - \gamma| > \epsilon) = Pr(X_{(1)} - \gamma > \epsilon) = Pr(n^{1/\beta} \frac{X_{(1)} - \gamma}{\alpha} > n^{1/\beta} \frac{\epsilon}{\alpha})$$

Now the using Theorem 8 of Gupta and Kundu (1999), it can be seen that the above probability converges to 0 as the sample size n tends to ∞, which proves that $\widehat{\gamma}_{init}$ is consistent for γ. Proof for the consistency of $\widehat{\alpha}_{init}$ is provided in Appendix A.7.

In order to show the consistency of the bias-corrected estimators, recall that $\widehat{\gamma}_w = X_{(1)} - \widehat{\alpha}_{init} \int_0^\infty (1 - (1 - e^{-y})^{\widehat{\beta}_w})^n \, dy$. Here, using the Slutskey's theorem and the facts that $\widehat{\alpha}_{init}$ and $\widehat{\beta}_w$ are consistent for α and β, respectively, $\widehat{\alpha}_{init} \int_0^\infty (1 - (1 - e^{-y})^{\widehat{\beta}_w})^n \, dy$ converges to 0 in probability. Hence, again, by using the Slutskey's theorem, $\widehat{\gamma}_w$ is consistent for γ.

Furthermore, one can show that $\widehat{\alpha}_w$ is also consistent for α along the same line as we used for the consistency of $\widehat{\alpha}_{init}$. It is obvious that the estimators will preserve the uniqueness property after bias-correction as well.

3 Evaluations of the Proposed Method

We perform a Monte-Carlo simulation study which evaluates the performance of the proposed estimator. The proposed estimation method is LSPF method. We compare the performance of LSPF method with three other modified maximum likelihood estimation (MMLE) methods as these methods provide estimators which exist for entire parameter space. We also compare LSPF method with MLE method when shape parameter is greater than 1. The comparisons are based on the biases and RMSE of the estimators. These three modified maximum likelihood estimators are termed as MMLE I, MMLE II and MMLE III, respectively. The estimation procedures for these methods are as follows:

MMLE I: This is the most convenient method of estimation. Recently, Pasari and Dikshit (2014) and Raqab et al. (2008) have also used the similar method in their work. This method is simply performed by estimating the location parameter γ by $\widehat{\gamma} = X_{(1)}$ and then, with a known $\widehat{\gamma}$, the other parameters α and β are estimated by maximizing the likelihood function of (α, β) based on the $(n-1)$ data points $(x_2 - \widehat{\gamma}, x_3 - \widehat{\gamma}, \ldots, x_n - \widehat{\gamma})$. Here, the estimation approach gets rid of the problem of unbounded likelihood function because of the shift of the data points by $\widehat{\gamma}$. Therefore, we maximize the log-likelihood function of (α, β), which is given by

$$\begin{aligned}\log(\ell_1(\alpha, \beta)) =& (n-1)\big(\log(\beta) - \log(\alpha)\big) - \sum_{i=2}^{n} \frac{x_i - x_{(1)}}{\alpha} \\ &+ (\beta - 1)\sum_{i=2}^{n} \log\left(1 - e^{-\frac{x_i - x_{(1)}}{\alpha}}\right), \alpha > 0, \beta > 0. \qquad (10)\end{aligned}$$

By using the notation from Subsection 2.2, $t_i = x_i - \widehat{\gamma}$, for computing the estimators of

β and α, we solve the following two equations simultaneously:

$$\frac{\partial}{\partial\beta}\log(\ell_1(\alpha,\beta)) = \frac{(n-1)}{\beta} + \sum_{i=2}^{n}\log\left(1-e^{-\frac{t_i}{\alpha}}\right) = 0,$$
$$\frac{\partial}{\partial\alpha}\log(\ell_1(\alpha,\beta)) = -\frac{n-1}{\alpha} + \frac{\sum_{i=2}^{n} t_i}{\alpha^2} - \frac{\beta-1}{\alpha^2}\sum_{i=2}^{n}\frac{t_i e^{-\frac{t_i}{\alpha}}}{\left(1-e^{-\frac{t_i}{\alpha}}\right)} = 0.$$

After simplification, the first equation gives $\beta = -\frac{n-1}{\sum_{i=2}^{n}\log\left(1-e^{-\frac{t_i}{\alpha}}\right)}$, say, $\widehat{\beta}(\alpha)$. Replacing this expression in place of β in equation (10), log-likelihood becomes function of parameter α, say $\log(\ell_1^*(\alpha))$, and hence can be maximized with respect to α in order to get the estimator of α, say, $\widehat{\alpha}$. Basically, this method has been reduced to one dimensional optimization problem. Hence, the estimators of (α,β,γ) using MMLE I are $(\widehat{\alpha},\widehat{\beta}(\widehat{\alpha}),\widehat{\gamma})$.

MMLE II: This method is proposed by Clifford Cohen and Jones Whitten (1982), Cohen and Whitten (1980) and Cohen and Whitten (1982). It is performed by simply replacing log-likelihood equation with respect to the location parameter, *i.e.*, $\frac{\partial}{\partial\gamma}\log(\ell) = 0$ with an alternative functional relationship $E(F(X_{(r)})) = F(x_{(r)})$, where r can take values $1, 2, \ldots, n$ and $E(F(X_{(r)})) = \frac{r}{n+1}$. Here we take $r = 1$ throughout the paper. For computing the estimators using MMLE II, we solve the following three equations simultaneously:

$$\frac{r}{n+1} = \left(1-e^{-\frac{x_{(r)}-\gamma}{\alpha}}\right)^{\beta}, \tag{11}$$
$$\frac{\partial}{\partial\beta}\log(\ell(\alpha,\beta,\gamma)) = \frac{n}{\beta} + \sum_{i=1}^{n}\log\left(1-e^{-\frac{x_i-\gamma}{\alpha}}\right) = 0, \tag{12}$$
$$\frac{\partial}{\partial\alpha}\log(\ell(\alpha,\beta,\gamma)) = -\frac{n}{\alpha} + \frac{\sum_{i=1}^{n}(x_i-\gamma)}{\alpha^2} - \frac{\beta-1}{\alpha^2}\sum_{i=1}^{n}\frac{(x_i-\gamma)e^{-\frac{x_i-\gamma}{\alpha}}}{\left(1-e^{-\frac{x_i-\gamma}{\alpha}}\right)} = 0. \tag{13}$$

The equation (11) gives $\gamma = x_{(r)} + \alpha\log\left(1-\left(\frac{r}{n+1}\right)^{1/\beta}\right)$, say, $\widehat{\gamma}(\alpha,\beta)$. Now, replacing γ with $\widehat{\gamma}(\alpha,\beta)$ in the remaining two equations and solving them simultaneously for α and β, gives the estimators $\widehat{\alpha}$ and $\widehat{\beta}$ and hence $\widehat{\gamma}(\widehat{\alpha},\widehat{\beta})$ can be obtained. Alternatively, an iterative procedure may be followed where we begin with a moment estimate $\widehat{\gamma}(\alpha,\beta)$ of γ for some (α,β). Then, $\widehat{\alpha}$ and $\widehat{\beta}$ can be obtained by maximizing $\log(\ell(\alpha,\beta,\widehat{\gamma}(\alpha,\beta))$,

Table 1: Biases and RMSEs of the estimators while varying sample size n based on 10000 simulations.

β	β_U	n	Method	Shape		Scale		Location		p
				Bias	RMSE	Bias	RMSE	Bias	RMSE	
0.50	2.00	20	LSPF	0.1468	0.2450	-0.1220	0.3310	-0.0078	0.0154	0.0036
			MMLE I	0.1526	0.2572	-0.0882	0.3494	0.0044	0.0103	0.0008
			MMLE II	-0.2941	0.3755	1.2851	1.6972	0.0028	0.0109	0.0018
			MMLE III	0.0709	0.2128	-0.0368	0.3601	0.0005	0.0103	0.0039
		50	LSPF	0.0885	0.1265	-0.0968	0.2204	-0.0016	0.0025	0.0000
			MMLE I	0.0641	0.1181	-0.0535	0.2249	0.0007	0.0018	0.0000
			MMLE II	-0.0794	0.1262	0.1353	0.3299	0.0005	0.0016	0.0000
			MMLE III	0.0217	0.0977	-0.0134	0.2294	0.0003	0.0017	0.0000
		100	LSPF	0.0701	0.0904	-0.0843	0.1631	-0.0009	0.001	0.0000
			MMLE I	0.0354	0.0740	-0.0306	0.1605	0.0002	0.0005	0.0000
			MMLE II	-0.0456	0.0776	0.0673	0.1933	0.0001	0.0004	0.0000
			MMLE III	0.0086	0.0631	-0.0054	0.1623	0.0001	0.0004	0.0000
0.75	5	20	LSPF	0.1916	0.4600	-0.0576	0.3076	-0.0174	0.0445	0.0094
			MMLE I	0.1850	0.3754	-0.0633	0.3099	0.0211	0.0343	0.0001
			MMLE II	-0.2135	0.5073	0.8193	1.5493	0.0063	0.0438	0.0099
			MMLE III	0.1202	0.4467	-0.0113	0.3354	0.0010	0.0413	0.0099
		50	LSPF	0.0857	0.1898	-0.0446	0.1958	-0.0042	0.0113	0.0000
			MMLE I	0.0796	0.1735	-0.0344	0.1995	0.0063	0.0106	0.0000
			MMLE II	-0.0700	0.1879	0.0760	0.2508	0.0018	0.0106	0.0001
			MMLE III	0.0253	0.1703	-0.0021	0.2080	0.0022	0.0096	0.0000
		100	LSPF	0.0586	0.1169	-0.0384	0.1392	-0.0015	0.0041	0.0000
			MMLE I	0.0395	0.1092	-0.0169	0.1406	0.0025	0.0043	0.0000
			MMLE II	-0.0435	0.1130	0.0429	0.1599	0.0007	0.0037	0.0000
			MMLE III	0.0076	0.1033	0.0032	0.1453	0.0011	0.0036	0.0000
1.00	8	20	LSPF	0.2975	0.8620	-0.0264	0.2980	-0.0246	0.0853	0.0252
			MMLE I	0.1753	0.4672	-0.0387	0.2965	0.0501	0.0703	0.0001
			MMLE II	-0.1207	0.8193	0.4744	1.1903	0.0050	0.1018	0.0276
			MMLE III	0.2259	0.8532	-0.0030	0.3225	-0.0031	0.0911	0.0280
		50	LSPF	0.1001	0.3174	-0.0241	0.1849	-0.0072	0.0288	0.0002
			MMLE I	0.0561	0.2196	-0.0117	0.1852	0.0202	0.0283	0.0000
			MMLE II	-0.0491	0.3215	0.0541	0.2260	0.0005	0.0310	0.0000
			MMLE III	0.0398	0.3037	0.0059	0.2020	0.0047	0.0281	0.0001
		100	LSPF	0.0534	0.1649	-0.0192	0.1276	-0.0023	0.013	0.0000
			MMLE I	0.0285	0.1441	-0.0079	0.1323	0.0100	0.0141	0.0000
			MMLE II	-0.0382	0.1671	0.0295	0.1472	0.0008	0.0124	0.0000
			MMLE III	0.0104	0.1616	0.0060	0.1381	0.0033	0.0118	0.0000

say $\log(\ell^*(\alpha,\beta))$, with respect to (α,β) which a two-dimensional optimization. Then a moment estimate $\widehat{\gamma}(\widehat{\alpha},\widehat{\beta})$ is obtained with $(\widehat{\alpha},\widehat{\beta})$. Therefore, taking some initial values of $(\alpha,\ \beta)$, this iteration $\widehat{\gamma}(\alpha,\beta) \to (\widehat{\alpha},\widehat{\beta}) \to \widehat{\gamma}(\widehat{\alpha},\widehat{\beta}) \to \ldots$ is continued until it converges. Such iteration was demonstrated by Wang (2005)

MMLE III: This method is proposed by Hall and Wang (2005). In order to under-

Table 2: Biases and RMSEs of the estimators while varying sample size n based on 10000 simulations.

β	β_U	n	Method	Shape		Scale		Location		p
				Bias	RMSE	Bias	RMSE	Bias	RMSE	
1.5	12	20	LSPF	0.4277	1.5597	0.0111	0.2937	-0.014	0.1566	0.0756
			MLE	-0.4644	0.7732	0.2503	0.4257	0.1141	0.1571	0.0700
			MMLE I	0.0245	0.6465	0.0068	0.2868	0.1321	0.1627	0.0001
			MMLE II	-0.0311	1.4164	0.2667	0.7935	0.0077	0.2087	0.0806
			MMLE III	0.3272	1.4925	0.0396	0.3148	0.0060	0.1893	0.1010
		50	LSPF	0.1952	0.8489	-0.0014	0.1851	-0.0116	0.0816	0.0036
			MLE	-0.1575	0.5464	0.0836	0.2255	0.046	0.082	0.0066
			MMLE I	-0.0698	0.3390	0.0203	0.1818	0.0692	0.0844	0.0000
			MMLE II	0.0796	0.8968	0.0459	0.2149	-0.0158	0.1103	0.0061
			MMLE III	0.1359	0.8503	0.0107	0.1940	0.0045	0.0949	0.0047
		100	LSPF	0.0630	0.3743	0.0004	0.1274	-0.0041	0.0434	0.0000
			MLE	-0.1108	0.3225	0.0464	0.1486	0.0274	0.0469	0.0010
			MMLE I	-0.0826	0.2367	0.0226	0.1264	0.0429	0.0523	0.0000
			MMLE II	0.0129	0.4032	0.0189	0.1381	-0.0070	0.0512	0.0001
			MMLE III	0.0196	0.3634	0.0086	0.1343	0.0081	0.0445	0.0000
2.0	20	20	LSPF	0.5276	2.4641	0.0361	0.3009	0.0266	0.2175	0.1310
			MLE	-0.4570	1.8233	0.2683	0.4606	0.1577	0.2692	0.0748
			MMLE I	-0.2141	0.8696	0.0286	0.2840	0.2222	0.2564	0.0000
			MMLE II	0.0862	2.4753	0.2432	0.6724	0.0186	0.3310	0.1477
			MMLE III	0.2357	1.8799	0.0769	0.3164	0.0462	0.2685	0.1851
		50	LSPF	0.3895	1.7103	0.0035	0.1884	-0.0086	0.1380	0.0198
			MLE	-0.1315	1.1409	0.0810	0.2287	0.0648	0.1552	0.0168
			MMLE I	-0.3036	0.5344	0.0445	0.1814	0.1353	0.1551	0.0000
			MMLE II	0.3645	1.9027	0.0426	0.2031	-0.0456	0.2174	0.0268
			MMLE III	0.3126	1.6384	0.0172	0.1930	0.0018	0.1778	0.0238
		100	LSPF	0.1358	0.8422	0.0040	0.1302	-0.0029	0.0867	0.0010
			MLE	-0.1168	0.5871	0.0386	0.1445	0.0405	0.0917	0.0050
			MMLE I	-0.2785	0.4159	0.0424	0.1299	0.0934	0.1065	0.0000
			MMLE II	0.1922	1.0170	0.0126	0.1365	-0.0317	0.1252	0.0007
			MMLE III	0.1020	0.8177	0.0078	0.1319	0.0066	0.1012	0.0009
3.0	30	20	LSPF	0.4588	3.8851	0.0617	0.3138	0.1546	0.3366	0.2088
			MLE	-0.7827	3.2839	0.3096	0.5051	0.2866	0.4529	0.1408
			MMLE I	-0.8651	1.4790	0.0592	0.2888	0.4077	0.4457	0.0000
			MMLE II	-0.2881	3.5134	0.2608	0.6219	0.1226	0.4705	0.2537
			MMLE III	-0.3581	2.3205	0.1287	0.3328	0.1860	0.4048	0.3161
		50	LSPF	0.8014	3.6369	0.0064	0.2014	0.0416	0.2287	0.0744
			MLE	-0.0205	2.7964	0.0873	0.2295	0.1066	0.3061	0.0554
			MMLE I	-0.9325	1.1292	0.0772	0.1925	0.2830	0.3062	0.0000
			MMLE II	0.9456	4.0872	0.0580	0.1946	-0.0732	0.3983	0.1163
			MMLE III	0.5161	3.0517	0.0382	0.1866	0.0225	0.3155	0.1033
		100	LSPF	0.5027	2.4756	-0.0048	0.1428	0.0150	0.1708	0.0174
			MLE	0.0504	1.8775	0.0350	0.1393	0.0475	0.2143	0.0156
			MMLE I	-0.8631	0.9838	0.0734	0.1438	0.2171	0.2340	0.0000
			MMLE II	0.9534	3.2747	0.0131	0.1290	-0.0960	0.3112	0.0314
			MMLE III	0.5147	2.4750	0.0082	0.1266	-0.0131	0.2404	0.0191

stand this method, suppose the PDF has the following form

$$h_1(x|\mu,\eta) = (x-\mu)^{\delta-1} h_2(x|\mu,\eta),\ \ x>\mu, \delta>0, \tag{14}$$

where η is vector of parameters other than μ and the corresponding likelihood function is given by

$$l_2(\mu,\eta) = \prod_{i=1}^{n}(x_i-\mu)^{\delta-1}\prod_{i=1}^{n} h_2(x_i|\mu,\eta),\ \ x_{(1)}>\mu, \delta>0. \tag{15}$$

According to Hall and Wang (2005), define a new function denoted by l_3 as follows

$$\begin{aligned} l_3(\mu,\eta) =& \frac{x_{(1)}-\mu}{x_{(2)}-\mu} l_2(\mu,\eta) \\ =& \frac{x_{(1)}-\mu}{x_{(2)}-\mu}\prod_{i=1}^{n}(x_i-\mu)^{\delta-1}\prod_{i=1}^{n} h_2(x_i|\mu,\eta), \end{aligned} \tag{16}$$

then, maximize $l_3(\mu,\eta)$ with respect to the parameters instead of maximizing $l_2(\mu,\eta)$ in order to get the maximum likelihood estimators. For a fixed μ, maximize $l_3(\mu,\eta)$ with respect to η and get $\widehat{\eta}(\mu)$, Now maximize $l_3(\mu,\widehat{\eta}(\mu))$ with respect to μ to get $\widehat{\mu}$, then obtain $\widehat{\eta}(\widehat{\mu})$. In our case, $\mu=\gamma$, $\delta=\beta$ and $\eta=(\alpha,\beta)$. The PDF of GE distribution in (2) can be written in the form of (14) by expanding $\big(1-e^{-\frac{x_i-\gamma}{\alpha}}\big)$ in terms of a polynomial. Hence, the above methodology is applicable for the estimation of the parameters. Therefore, $\log(\ell_3)$ after some adjustment for the convenient purpose, in our case, is as follows:

$$\begin{aligned} \log(\ell_3(\alpha,\beta,\gamma)) =& n\big(\log(\beta)-\log(\alpha)\big) - \sum_{i=1}^{n}\frac{x_{(i)}-\gamma}{\alpha} + (\beta-1)\sum_{i=2}^{n}\log\left(1-e^{-\frac{x_{(i)}-\gamma}{\alpha}}\right) \\ & -\log(x_{(2)}-\gamma) + \log\left((x_{(1)}-\gamma)\Big(1-e^{-\frac{x_{(1)}-\gamma}{\alpha}}\Big)^{\beta-1}\right) \end{aligned} \tag{17}$$

Now, we maximize the $\log(\ell_3(\alpha,\beta,\gamma))$ in (17) with respect to the parameters α,β and γ. Since the equation $\frac{\partial \log(\ell_3(\alpha,\beta,\gamma))}{\partial\beta}=0$, after simplification, gives $\beta=-\frac{n}{\sum_{i=1}^{n}\log\left(1-e^{-\frac{x_i-\gamma}{\alpha}}\right)}$, say, $\widehat{\beta}(\alpha,\gamma)$. Now we find $(\widehat{\alpha},\widehat{\gamma})$ which maximizes the $\log(\ell_3(\alpha,\widehat{\beta}(\alpha,\gamma),\gamma))$, say $\log(\ell_3^*(\alpha,\gamma))$, with respect to (α,γ) and hence, $\widehat{\beta}(\widehat{\alpha},\widehat{\gamma})$ can be obtained. Further, a similar kind of iterative procedure earlier suggested for MMLE II can be followed, where we begin with an estimate of β using some (α, γ). Therefore, the iteration $\widehat{\beta}(\alpha, \gamma) \to (\widehat{\alpha}, \widehat{\gamma}) \to \widehat{\beta}(\widehat{\alpha}, \widehat{\gamma}) \to \ldots$ with some (α, γ) is continued until it converges.

We obtain all the results by performing Monte-Carlo simulation for the values of shape parameter β as $0.50, 0.75, 1.00, 1.5, 2, 3$ and the sample size n are taken to be $20, 50$ and 100. We consider scale parameter α equal to 1 and location parameter γ to be 0. All the simulation results are presented in tables 1-2. These results are based on the 10000 simulations. We provide bias and RMSE for the estimators of location, scale and shape parameters.

While performing the simulation some proportion of generated data sets has been rejected which is reported in the last column of the tables. These data sets are rejected because they provide very bad estimate values of the parameters. For example, 10^7 as an estimate of the shape parameter, hence, it is needed to be rejected while performing the simulation study. In order to do so, we choose an upper bound for the estimates of shape parameter, say β_U, which can be determined looking into the histogram of the estimates. If we take the proportion of rejected samples into consideration, we observe that MMLE I and LSPF method perform better than any other method as the proportion of rejected samples is small under the application of these methods.

Now, we try to summarize the results of the performed simulation study. These are the following observations based on the reported simulation results:

- As the shape parameter β takes small values or the sample size increases, the performance of every method for estimating all the parameters improve.
- MMLE I underestimates the scale parameter when $\beta \leq 1$ and overestimates the scale parameter when $\beta > 1$.
- For $\beta \approx 0.5$:
  - MMLE III performs better than any other method in terms of RMSE and bias.

- When $n \approx 20$, LSPF gives second best performance after MMLE III for estimating shape and scale parameter, while estimating location, it does not perform that good as compared to other methods for any sample size.

- For $0.5 < \beta \leq 1$:
  - When $n \approx 20$, MMLE I performs very good in terms of RMSE while MMLE III and LSPF are the next choices after MMLE I.
  - When $n \geq 50$, MMLE III is advisable to use, but other methods also perform equally good.
- For $1 < \beta \leq 2$:
  - For the estimation of shape parameter, it is advisable to use MMLE I (first preference) and MLE (second preference).
  - For estimation of location parameter, LSPF performs better than any other method both in terms of RMSE and bias when $\beta \approx 2$.
  - For $\beta \approx 1.5$, if the small RMSE is the main concern to keep in mind, MMLE I is the only one which estimates all the three-parameters with minimum RMSE when $n \leq 50$.
- For $\beta \approx 3$:
  - For the estimation of shape parameter, it is strictly advisable to use MMLE I as a first preference and then either of MLE or MMLE III as a second preference if RMSE is main concern. MMLE I has very large bias, therefore MMLE III is preferred over MMLE I when $n \approx 20$ and it is MLE which is preferred over MMLE I when $n \geq 50$.
  - For estimation of location parameter, LSPF performs very good irrespective of any sample size in terms of both RMSE and bias and when $\beta \approx 3$, MLE can be used as a second choice. Also, as β increases, MMLE III starts performing better than LSPF with smaller bias values of the estimates of the location parameter.

- MMLE I is the only method which estimates the parameters with approximately zero proportion of rejection irrespective of any value of β and any sample size. Therefore, we advise to use this method as a first preference based on the reported proportion of rejection. The second preference is LSPF for $0.5 \leq \beta < 3$ and MLE for comparatively large value of β ($\beta \geq 3$). When β is very small ($\beta \approx 0.5$), it has been observed that all the methods have very small (≈ 0) proportion of rejection for sample size $n \geq 50$ and therefore any method can be recommended but when $n \leq 20$, we recommend to use LSPF.

4 Data Analysis

We consider a data set for the data analysis which represent the lifetimes in months of electrical parts, see Bain and Engelhardt (1987) (page 162). A simple data analysis by plotting the histogram of the data set indicates that the shape of the density is reverse 'J' shaped.

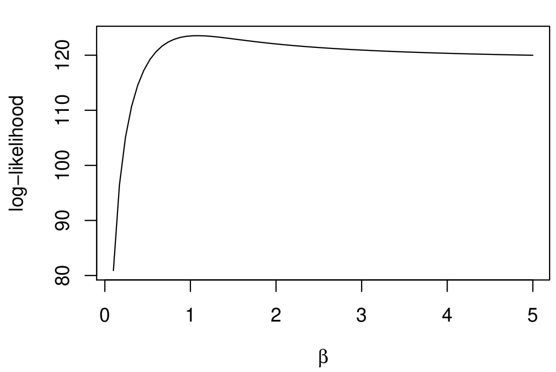
4.1 Data Set

Table 3: Data analysis for data set.

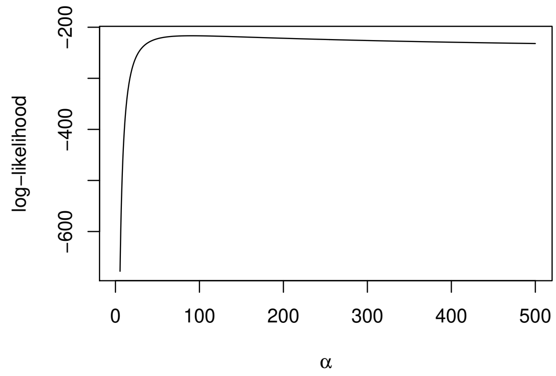
Data Set:
0.15, 2.37, 2.90, 7.39, 7.99, 12.05, 15.17, 17.56,
22.40, 34.84, 35.39, 36.38, 39.52, 41.07, 46.50, 50.52,
52.54, 58.91, 58.93, 66.71, 71.48, 71.84, 77.66, 79.31,
80.90, 90.87, 91.22, 96.35, 108.92, 112.26, 122.71, 126.87,
127.05, 137.96, 167.59, 183.53, 282.49, 335.33, 341.19, 409.97

Method	Shape	Scale	Location	CvM-Statistics	P-value
LSPF	1.0799	91.2659	-2.7673	0.0415	0.9252
MMLE I	1.0786	90.8277	0.1500	0.0489	0.8826
MMLE II	0.8919	99.2703	-1.4062	0.0739	0.7277
MMLE III	1.0408	92.1752	-1.4577	0.0419	0.9229

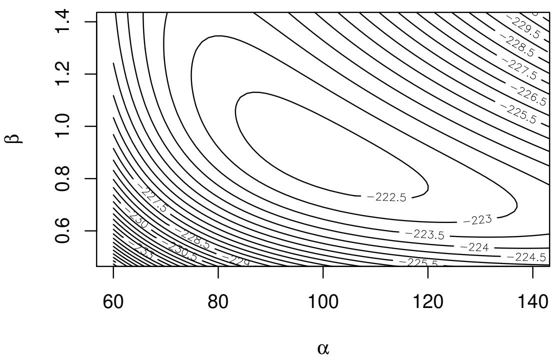
We consider the data originally reported by Bain and Engelhardt (1987) (see page 162) and analyzed by Harter (1973). The data represents the observed lifetimes in months



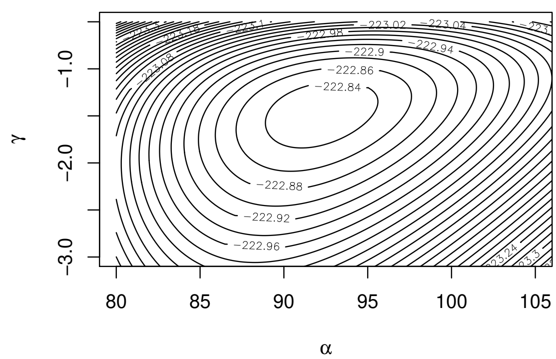
(a) Plot of the log-likelihood function $\log(\ell_w(\beta))$ under LSPF method



(b) Plot of the log-likelihood function $\log(\ell_1^*(\alpha))$ under MMLE I method



(c) Contour plot of the log-likelihood function $\log(\ell^*(\alpha, \beta))$ under MMLE II method



(d) Contour plot of the log-likelihood function $\log(\ell_3^*(\alpha, \gamma))$ under MMLE III method

Figure 1: Plots for various estimation methods for data analysis of data set.

Table 4: Bootstrap confidence intervals for data set based on 10000 simulations.

Method	95% CI			99% CI			p
	Shape	Scale	Location	Shape	Scale	Location	
LSPF	(0.7462, 2.3137)	(56.7866, 130.8569)	(-12.8434, 4.5948)	(0.6765, 3.7231)	(48.3591, 147.1999)	(-19.4118, 8.3977)	0.0020
MMLE I	(0.7415, 1.7730)	(57.3981, 131.9950)	(0.2530, 10.3114)	(0.6630, 2.0775)	(50.3146, 147.2998)	(0.1764, 14.9580)	0.0000
MMLE II	(0.4676, 1.5387)	(63.1960, 166.0368)	(-7.1885, 4.3296)	(0.4075, 2.3868)	(53.7078, 193.6294)	(-15.4015, 7.6560)	0.0004
MMLE III	(0.6568, 2.1126)	(57.0883, 137.9524)	(-8.8587, 6.7738)	(0.5831, 3.1162)	(48.3295, 157.8311)	(-17.9056, 10.3762)	0.0014

of a random sample of electrical parts. The data set is presented in Table 3 along with the estimates of the parameters based on LSPF method and other methods. Cramér-von Mises (CvM) test is implemented for every method. We report the CvM statistic along with the P-value corresponding to each method. In order to demonstrate the

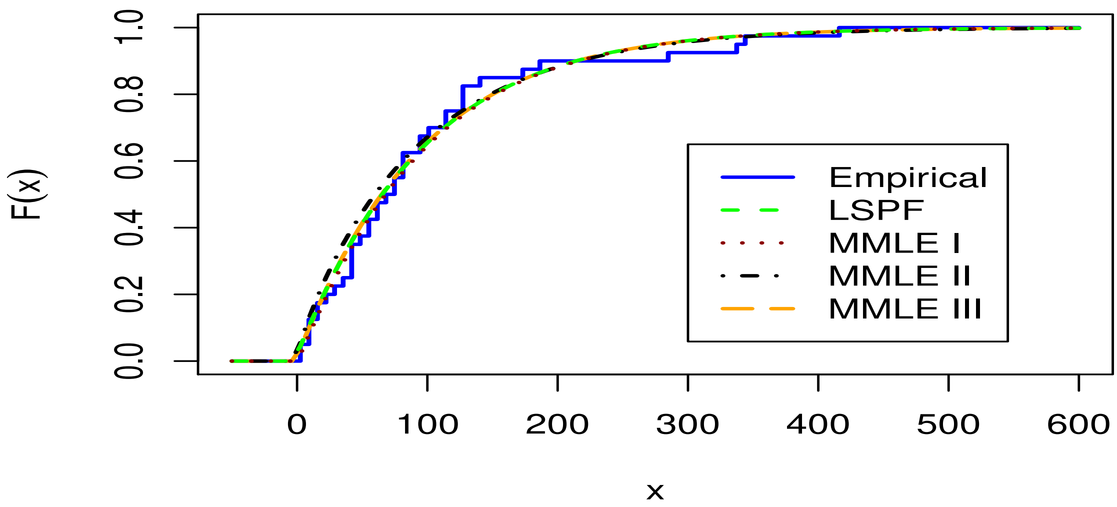


Figure 2: Fitted CDF plots along with the empirical CDF of data set

LSPF method, we maximize the likelihood function $\ell_w(\beta)$ in equation (4) with respect to the β and get the estimate of β as 1.0799. Using the estimate of β, we obtain estimates of α and γ as 91.2659 and -2.7673, respectively, using the steps mentioned in Section 2. Based on the reported CvM statistics and P-values in **Table 3**, we observe that LSPF and MMLE III perform very good for this particular data set. Plots of $\mathbf{\log(\ell_w(\beta)),\ \log(\ell_1^*(\alpha)),\ \log(\ell^*(\alpha,\beta))}$ and $\mathbf{\log(\ell_3^*(\alpha,\gamma))}$ in **Figure 1** show that the likelihood functions are uni-modal, *i.e.*, the obtained estimates maximize these likelihoods globally. The plots of fitted CDF and the empirical CDF have been presented in Figure 2. In order to find the bootstrap confidence intervals, we choose $\beta_U = 12$. Table 4 presents 95% and 99% bootstrap confidence intervals for each method based on this data set.

5 Conclusion

In this article, we have proposed a method of estimation for a three-parameter GE distribution proposed earlier and known as LSPF method. We presented some properties of the proposed estimators, like, existence; uniqueness and consistency for the entire range of the parameter space. A Monte-Carlo simulation study has been carried out for evaluation of the performance of the proposed LSPF method in comparisons with other existing prominent methods. In simulation study, we reported the bias and the RMSE for all the three estimators of GE distribution. We also have reported the required proportion of rejected samples during the simulation study in order to estimate the parameters. For $\beta > 1$, LSPF method performs better than any other method, in terms of both bias and RMSE, for the estimation of the location and scale parameters but there are other methods, such as MMLE I, MLE and sometimes MMLE III as well, which have very good and consistent performance for the estimation of shape parameter. For the other case when $\beta < 1$, LSPF method is advisable to use for the estimation of the shape parameter when sample size is small. When $n \geq 50$, MMLE III is advisable if β is small enough ($\beta \approx 0.5$) and any method can be good if $0.5 < \beta \leq 1$ as they all have equal performances for all the parameters. For the data analysis, LSPF and MMLE I are advisable to use before the use of any method which is quite clear from the reported results on the proportion of rejected samples during the simulation study. The method has been illustrated based on a real lifetime data set.

References

Bain, L. J. and Engelhardt, M. (1987). *Introduction to probability and mathematical statistics*. Brooks/Cole.

Billingsley, P. (2008). *Probability and measure*. John Wiley & Sons.

Clifford Cohen, A. and Jones Whitten, B. (1982). Modified moment and maximum likelihood estimators for parameters of the three-parameter gamma distribution. *Communications in Statistics-Simulation and Computation*, 11(2):197–216.

Cohen, A. C. and Whitten, B. J. (1980). Estimation in the three-parameter lognormal distribution. *Journal of the American Statistical Association*, 75(370):399–404.

Cohen, A. C., Whitten, B. J., and Ding, Y. (1984). Modified moment estimation for the three-parameter weibull distribution. *Journal of Quality Technology*, 16(3):159–167.

Cohen, C. A. and Whitten, B. (1982). Modified maximum likelihood and modified moment estimators for the three-parameter weibull distribution. *Communications in Statistics-Theory and Methods*, 11(23):2631–2656.

Ghitany, M., Al-Jarallah, R., and Balakrishnan, N. (2013). On the existence and uniqueness of the mles of the parameters of a general class of exponentiated distributions. *Statistics*, 47(3):605–612.

Gupta, R. D. and Kundu, D. (1999). Theory & methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, 41(2):173–188.

Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family: an alternative to gamma and weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 43(1):117–130.

Gupta, R. D. and Kundu, D. (2007). Generalized exponential distribution: Existing results and some recent developments. *Journal of Statistical Planning and Inference*, 137(11):3537–3547.

Hall, P. and Wang, J. Z. (2005). Bayesian likelihood methods for estimating the end point of a distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(5):717–729.

Harter, H. L. (1973). Likelihood ratio test for discrimination between two models with unknown location and scale parameters: Discussion no. 2. *Technometrics*, 15(1):29–31.

Lehmann, E. L. and Casella, G. (2006). *Theory of point estimation*. Springer Science & Business Media.

McCool, J. (1974). Inferential techniques for weibull populations. aerospace research laboratories report. Technical report, ARL TR 74-0180, Wright-Patterson AFB, Ohio.

Mudholkar, G. S., Srivastava, D. K., and Freimer, M. (1995). The exponentiated weibull family: A reanalysis of the bus-motor-failure data. *Technometrics*, 37(4):436–445.

Nagatsuka, H. and Balakrishnan, N. (2012). Parameter and quantile estimation for the three-parameter gamma distribution based on statistics invariant to unknown location. *Journal of Statistical Planning and Inference*, 142(7):2087–2102.

Nagatsuka, H. and Balakrishnan, N. (2013). A consistent method of estimation for the parameters of the three-parameter inverse gaussian distribution. *Journal of Statistical Computation and Simulation*, 83(10):1915–1931.

Nagatsuka, H., Balakrishnan, N., and Kamakura, T. (2014). A consistent method of estimation for the three-parameter gamma distribution. *Communications in Statistics-Theory and Methods*, 43(18):3905–3926.

Nagatsuka, H., Kamakura, T., and Balakrishnan, N. (2013). A consistent method of estimation for the three-parameter weibull distribution. *Computational Statistics & Data Analysis*, 58:210–226.

Pasari, S. and Dikshit, O. (2014). Three-parameter generalized exponential distribution in earthquake recurrence interval estimation. *Natural hazards*, 73(2):639–656.

Raqab, M. Z., Madi, M. T., and Kundu, D. (2008). Estimation of p (y< x) for the three-parameter generalized exponential distribution. *Communications in Statistics—Theory and Methods*, 37(18):2854–2864.

Wang, J. Z. (2005). A note on estimation in the four-parameter beta distribution. *Communications in Statistics-Simulation and Computation*, 34(3):495–501.

A Proofs

A.1 Proof of Theorem 2.1

First of all, we find joint PDF of the random variables $W_{(2)}, W_{(3)}, \ldots, W_{(n)}$ for $\beta > 0$ using the transformation of random variables . Define $Z_{(i)} = \frac{X_{(i)}-\gamma}{\alpha}, i = 1, 2, \ldots, n$. Here, $Z_1, Z_2, \ldots, Z_n$ are independent and identical random variables from standard GE distribution $GE(1, \beta, 0)$ with shape parameter β. For the sake of convenience, denote the PDF and CDF of $GE(1, \beta, 0)$ by $g(.;\beta)$ and $G(.;\beta)$, respectively. Now, $W_{(i)}$ in the equation (3) can be re-expressed in terms of $Z_{(i)}$'s as follows:

$$W_{(i)} = \frac{Z_{(i)} - Z_{(1)}}{Z_{(n)} - Z_{(1)}},\ i = 2, 3, \ldots, n$$
$$\implies Z_{(i)} = Z_{(1)} + W_{(i)}(Z_{(n)} - Z_{(1)}),\ i = 2, 3, \ldots, n.$$

Let $U = Z_{(1)}$ and $V = Z_{(n)}$. Therefore $Z_{(i)} = U + W_{(i)}(V - U),\ i = 2, 3, \ldots, n$. It can be shown that the Jocobian of the transformation $J = \frac{\partial(Z_{(1)}, Z_{(2)}, \ldots, Z_{(n-1)}, Z_{(n)})}{\partial(U, W_{(2)}, \ldots, W_{(n-1)}, V)} = (V - U)^{(n-2)}$. If we use a notation $f_{Y_1, Y_2, \ldots, Y_p}(.)$ to denote the joint PDF of $Y_1, Y_2, \ldots, Y_p$, we have

$$\begin{aligned} &f_{U, W_{(2)}, \ldots, W_{(n-1)}, V}(u, w_2, \ldots, w_{n-1}, v) \\ &\quad = |J| f_{Z_{(1)}, Z_{(2)}, \ldots, Z_{(n-1)}, Z_{(n)}}(u, u + w_2(v-u), \ldots, u + w_{n-1}(v-u), v) \\ &\quad = n!(v-u)^{n-2} g(u;\beta) g(v;\beta) \Big\{ \prod_{i=2}^{n-1} g(u + (v-u)w_i;\beta) \Big\}, \end{aligned}$$

$0 \le w_2 \le \cdots \le w_{n-1} \le 1$ and $0 < u < v < \infty$. Therefore, the joint PDF

$$\begin{aligned} &f_{W_{(2)}, \ldots, W_{(n-1)}}(w_2, \ldots, w_{n-1}) \\ &\quad = \int_0^\infty \int_u^\infty n!(v-u)^{n-2} g(u;\beta) g(v;\beta) \Big\{ \prod_{i=2}^{n-1} g(u + (v-u)w_i;\beta) \Big\}\, dv du \qquad (18) \\ &\quad = n!\beta^n \int_0^\infty \int_0^\infty v^{n-2} e^{-\sum_{i=1}^n (u+vw_i)} \prod_{i=1}^n \left(1 - e^{-(u+vw_i)}\right)^{\beta-1} dv du, \end{aligned}$$

with $w_1 = 0$ and $w_n = 1$. Hence, the likelihood function of β given $w_2, w_3, \ldots, w_{n-1}$ is given by

$$\ell_w(\beta) = f_{W_{(2)},\ldots,W_{(n-1)}}(w_2, \ldots, w_{n-1}), \beta > 0.$$

which proves the proposition.

A.2 Proof of Proposition 2.1

By equation (18), we have

$$\ell_w(\beta) = n! \int_0^\infty \int_u^\infty (v-u)^{n-2} g(u;\beta) g(v;\beta) \Big\{ \prod_{i=2}^{n-1} g(u + (v-u)w_i;\beta) \Big\}, \beta > 0. \tag{19}$$

It can be shown for $\beta > 0,\ 0 \leq w_2 \leq \cdots \leq w_{n-1} \leq 1$ and $0 < u < v < \infty$ that $\frac{(n-2)! \prod_{i=2}^{n-1} (v-u) g(u+(v-u)w_i;\beta)}{(G(v;\beta)-G(u;\beta))^{n-2}}$ is bounded , *i.e.*, $\exists$ an $M \in \mathbb{R}$ such that $(n-2)! \prod_{i=2}^{n-1} (v-u)$ $\times g(u + (v-u)w_i;\beta) < M(G(v;\beta) - G(u;\beta))^{n-2}$ $\forall$ $\beta > 0,\ 0 \leq w_2 \leq \cdots \leq w_{n-1} \leq 1$ and $0 < u < v < \infty$.

$$\begin{aligned}
\implies \int_0^\infty \int_u^\infty n!(v-u)^{n-2} g(u;\beta) g(v;\beta) \prod_{i=2}^{n-1} g(u + (v-u)w_i;\beta)\ dvdu \\
< M \int_0^\infty \int_u^\infty n(n-1) g(u;\beta) g(v;\beta) (G(v;\beta) - G(u;\beta))^{n-2}\ dvdu = M,
\end{aligned}$$

since $\int_0^\infty \int_u^\infty n(n-1) g(u;\beta) g(v;\beta)(G(v;\beta) - G(u;\beta))^{n-2}\ dvdu = 1$. Hence, using the above inequality and from equation (19), it is clear that likelihood function $\ell_w(\beta)$ is a bounded function of β.

A.3 Proof of Proposition 2.2

To show that the likelihood function is differentiable, we need to show that the partial derivative, with respect to β, can be taken inside the double integrals in the equation (5).

Given $0 \leq w_2 \leq \cdots \leq w_{n-1} \leq 1, w_1 = 0, w_n = 1$, we show:

1. $\frac{\partial}{\partial \beta} e^{h_w(\beta; u,v)}$ exists,

2. $\left| \frac{\partial}{\partial \beta} e^{h_w(\beta; u,v)} \right| < h_1(u,v)$ for some positive function h_1 and $\forall (u,v) \in (0,\infty) \times (0,\infty)$ such that $\int_0^\infty \int_0^\infty h_1(u,v)\, dvdu < \infty$.

Since exponential, logarithmic and polynomials are well-known smooth functions, therefore $e^{h_w(\beta;u,v)}$ is differentiable with respect to β and it is given by

$$\frac{\partial}{\partial \beta} e^{h_w(\beta;u,v)} = h'_w(\beta; u, v) e^{h_w(\beta;u,v)}, \tag{20}$$

where $h'_w(\beta; u, v) = \frac{\partial}{\partial \beta} h_w(\beta; u, v) = \frac{n}{\beta} + \sum_{i=1}^n \log(1 - e^{-(u+vw_i)})$. For every $\beta > 0, 0 \leq w_2 \leq \cdots \leq w_{n-1} \leq 1$ and $0 < u < v < \infty$, $\left| h'_w(\beta; u, v) e^{h_w(\beta;u,v)} \right|$ is bounded above. That is, $\exists\ M_2 > 0$ such that $\left| h'_w(\beta; u, v) e^{h_w(\beta;u,v)} \right| < M_2$. Now consider

$$\left| h'_w(\beta; u, v) e^{h_w(\beta;u,v)} \right| = \left| h'_w(\beta; u, v) e^{h_w(\beta;u,v)/2} e^{h_w(\beta;u,v)/2} \right| \leq M_2 e^{h_w(\beta;u,v)/2}$$

Say, $h_1(u,v) = M_2 e^{h_w(\beta;u,v)/2}$. Then

$$\begin{aligned} \int_0^\infty \int_0^\infty h_1(u,v)\, dvdu &= M_2 \int_0^\infty \int_0^\infty e^{h_w(\beta;u,v)/2}\, dvdu \\ &\leq M_2 \Big(\int_0^\infty \int_0^\infty e^{h_w(\beta;u,v)}\, dvdu \Big)^{1/2} < \infty \end{aligned},$$

since the likelihood function $\ell_w(\beta)$ is bounded. Now, part (ii) of theorem 16.8 of Billingsley (2008) implies that the derivative of $\ell_w(\beta)$ is given by

$$\begin{aligned} \ell'_w(\beta) &= n! \int_0^\infty \int_0^\infty \frac{\partial}{\partial \beta} e^{h_w(\beta;u,v)}\, dvdu \\ &= n! \int_0^\infty \int_0^\infty h'_w(\beta; u, v) e^{h_w(\beta;u,v)}\, dvdu \end{aligned} \tag{21}$$

Hence, $\ell'_w(\beta)$ can be obtained by replacing $h'_w(\beta; u, v)$ and $e^{h_w(\beta;u,v)}$ by their respective expressions in equation (21).

A.4 Proof of Theorem 2.2

Recall from (21) that

$$\ell'_w(\beta) = n! \int_0^\infty \int_0^\infty h'_w(\beta; u, v) e^{h_w(\beta;u,v)}\ dvdu, \tag{22}$$

where $h'_w(\beta; u, v) = \frac{n}{\beta} + \sum_{i=1}^n \log(1 - e^{-(u+vw_i)})$. First we show that likelihood equation $\ell'_w(\beta) = 0$ has at least one solution. Note that, for any choice of u, v and w, $e^{h_w(\beta;u,v)} > 0$ and also the integrals are on positive support, hence the change in the sign of $\ell'_w(\beta)$ directly depends on the change in the sign of $h'_w(\beta; u, v)$. $h'_w(\beta; u, v) \to \infty$ as $\beta \downarrow 0$ and $h'_w(\beta; u, v) < 0$ as $\beta \to \infty$. Moreover, $h''_w(\beta; u, v) = -\frac{n}{\beta^2} < 0\ \forall\ \beta$, *i.e.*, $h'_w(\beta; u, v)$ changes sign from positive to negative and the change in sign is only once. Hence $\ell'_w(\beta)$ changes sign in the similar way and $\ell'_w(\beta) = 0$ has an unique solution which maximizes the likelihood function $\ell_w(\beta)$. Suppose the change in sign occurs at $\widehat{\beta}_w$, then $\widehat{\beta}_w$ is unique maxima of likelihood function $\ell_w(\beta)$.

A.5

Lemma A.1. *For any fixed $\beta \neq \beta_0$, where β_0 is true value of the parameter β,*

$$\lim_{n\to\infty} Pr\Big(\frac{\ell_w(\beta; W_{(2)}, \ldots, W_{(n-1)})}{\ell_w(\beta_0; W_{(2)}, \ldots, W_{(n-1)})} < 1\Big) = 1.$$

A.6 Uni-modality of $\log(\ell_t(\alpha))$

Showing the uni-modality of likelihood function in equation (8) is equivalent to show that the equation (9) has an unique solution. From (9), we have

$$\alpha = \frac{1}{n-1} \sum_{i=2}^n t_i \Big(\frac{1 - \widehat{\beta}_w e^{-\frac{t_i}{\alpha}}}{1 - e^{-\frac{t_i}{\alpha}}}\Big). \tag{23}$$

Let us denote the right hand side of above equation by $H_1(\alpha)$. Now, we will show that equation $\alpha = H_1(\alpha)$ has exactly one solution. Which means it is enough to show that the functions $H_1(\alpha)$ and α meet each other exactly once. $H_1(\alpha)$ further can be simplified as following

$$H_1(\alpha) = \frac{1}{n-1}\sum_{i=2}^{n} t_i\Big(1 + \frac{1-\widehat{\beta}_w}{e^{\frac{t_i}{\alpha}}-1}\Big). \tag{24}$$

When $\widehat{\beta}_w > 1$, $H_1(\alpha)$ is a strictly decreasing in α and it decreases from $\frac{1}{n-1}\sum\limits_{i=2}^{n} t_i$ to $-\infty$, whereas α is strictly increasing in α. $H_1(\alpha)$ is constant function of α whenever $\widehat{\beta}_w = 1$. It is easy to see that $H_1(\alpha)$ and α meet exactly once whenever $\widehat{\beta}_w \geq 1$. Again, when $\widehat{\beta}_w < 1$, $H_1(\alpha)$ is a strictly increasing in α and it increases from $\frac{1}{n-1}\sum\limits_{i=2}^{n} t_i$ to ∞, whereas α is also strictly increasing in α. Therefore, there is possibility that $H_1(\alpha)$ and α meet at most once, but if we show that $\frac{\partial H_1(\alpha)}{\partial \alpha} < 1$, they have to meet exactly once. Let us consider

$$\frac{\partial H_1(\alpha)}{\partial \alpha} = \frac{1-\widehat{\beta}_w}{(n-1)\alpha^2}\sum_{i=2}^{n}\left(\frac{t_i^2 e^{-\frac{t_i}{\alpha}}}{(1-e^{-\frac{t_i}{\alpha}})^2}\right).$$

Since $\frac{s^2 e^{-s}}{(1-e^{-s})^2} < 1 \;\forall\; s > 0$ (see part (i) of Lemma 2 in Ghitany et al. (2013)), we have $\frac{\partial H_1(\alpha)}{\partial \alpha} < 1$. Hence, equation $\alpha = H_1(\alpha)$ has unique solution. Now, if it can be shown that $\lim\limits_{\alpha\to 0}\ell_t(\alpha) = 0$ and $\lim\limits_{\alpha\to\infty}\ell_t(\alpha) = 0$, it will imply that $\ell_t(\alpha)$ is uni-modal because $\ell_t(\alpha) > 0 \;\forall\; \alpha > 0$. Now, we show that $\lim\limits_{\alpha\to 0}\ell_t(\alpha) = 0$ and $\lim\limits_{\alpha\to\infty}\ell_t(\alpha) = 0$ which will complete the proof. Since

$$\ell_t(\alpha) \propto \frac{1}{\alpha^{n-1}} e^{-\sum_{i=2}^{n}\frac{t_i}{\alpha}} \prod_{i=2}^{n}\big(1-e^{-\frac{t_i}{\alpha}}\big)^{\widehat{\beta}_w - 1}. \tag{25}$$

It is easy to see from (25) that $\lim\limits_{\alpha\to 0}\ell_t(\alpha) = 0$ because the convergence rate of exponential functions is faster than the polynomial function. Rewrite equation (25) as following

$$\ell_t(\alpha) \propto \prod_{i=2}^{n}\frac{1}{t_i}\frac{\frac{t_i}{\alpha}e^{-\frac{t_i}{\alpha}}}{(1-e^{-\frac{t_i}{\alpha}})}\big(1-e^{-\frac{t_i}{\alpha}}\big)^{\widehat{\beta}_w}. \tag{26}$$

Since $\lim_{s\to 0} \frac{se^{-s}}{(1-e^{-s})} = 1$ and $\lim_{s\to 0}(1-e^{-s})^{\widehat{\beta}_w} = 0$ for any $\widehat{\beta}_w$. Hence, $\lim_{\alpha\to\infty} \ell_t(\alpha) = 0$.

A.7 Consistency of $\widehat{\alpha}_{init}$

Recall that $\widehat{\alpha}_{init}$ is obtained by maximizing the log-likelihood function of α based on $(t_2, t_3, \ldots, t_n)$ assuming that γ and β are known and replaced by $\widehat{\gamma}_{init}$ and $\widehat{\beta}_w$, respectively. Assume that α_0, β_0 are γ_0 are true values of the unknown parameters. In order to the consistency, we first show that

$$\lim_{n\to\infty} Pr\Big(\frac{\log \ell_t(\alpha_0) - \log \ell_t(\alpha)}{n-1} > 0\Big) = 0 \;\forall\; \alpha > 0. \tag{27}$$

Consider the quantity

$$\begin{aligned}
\frac{1}{n-1}\log \ell_t(\alpha_0) =& \log \widehat{\beta}_w - \log \alpha_0 - \frac{1}{n-1}\sum_{i=2}^{n} \frac{x_i - \widehat{\gamma}_{init}}{\alpha_0} + \frac{\widehat{\beta}_w - 1}{n-1}\sum_{i=2}^{n} \log\left(1 - e^{-\frac{x_i - \widehat{\gamma}_{init}}{\alpha_0}}\right) \\
=& \log \widehat{\beta}_w - \log \alpha_0 - \frac{1}{n-1}\sum_{i=2}^{n}\Big\{\frac{x_i-\gamma_0}{\alpha_0} + \frac{\gamma_0 - \widehat{\gamma}_{init}}{\alpha_0}\Big\} \\
&+ \frac{\widehat{\beta}_w - 1}{n-1}\sum_{i=2}^{n} \log\Big\{1 - A_n e^{-\frac{x_i-\gamma_0}{\alpha_0}}\Big\},
\end{aligned}$$

where $A_n = e^{-\frac{\gamma_0 - \widehat{\gamma}_{init}}{\alpha_0}}$. Suppose that the term "convergence in probability" is denoted by $\xrightarrow{P}$. Here, $\widehat{\beta}_w \xrightarrow{P} \beta$ and $\widehat{\gamma}_{init} \xrightarrow{P} \gamma$. Also, by using the Weak Law of Large Numbers,

$$\frac{1}{n-1}\sum_{i=2}^{n} \log\Big\{1 - A_n e^{-\frac{x_i-\gamma_0}{\alpha_0}}\Big\} \xrightarrow{P} E\Big(\log\Big\{1 - e^{-Z_1}\Big\}\Big) \tag{28}$$

$$\frac{1}{n-1}\sum_{i=2}^{n} \frac{x_i - \gamma_0}{\alpha_0} \xrightarrow{P} E(Z_1), \tag{29}$$

Where $Z_1 \sim GE(1, \beta_0, 0)$ as defined earlier in A.1 and the result stated in equation (28) is possible because $A_n \xrightarrow{P} 1$. Also, it can be verified from Gupta and Kundu (1999) that $E(Z_1) = \psi(\beta_0 + 1) - \psi(1)$ and $e^{-Z_1} \sim Beta(1, \beta_0)$, where $\psi(.)$ denote the digamma function. It can be easily obtained that $E\Big(\log\Big\{1 - e^{-Z_1}\Big\}\Big) = -1/\beta_0$ Therefore, by using the Slutsky's Theorem and above facts, we have

$$\frac{1}{n-1}\log \ell_t(\alpha_0) \xrightarrow{P} \log\beta_0 - \log\alpha_0 - \big(\psi(\beta_0+1)-\psi(1)\big) - \frac{\beta_0-1}{\beta_0}. \tag{30}$$

Now, consider the quantity

$$\begin{aligned}\frac{1}{n-1}\log \ell_t(\alpha) =& \log\widehat{\beta}_w - \log\alpha - \frac{1}{n-1}\sum_{i=2}^{n}\frac{x_i-\widehat{\gamma}_{init}}{\alpha} + \frac{\widehat{\beta}_w-1}{n-1}\sum_{i=2}^{n}\log\left(1-e^{-\frac{x_i-\widehat{\gamma}_{init}}{\alpha}}\right)\\ =& \log\widehat{\beta}_w - \log\alpha - \frac{1}{n-1}\Big(\frac{\alpha_0}{\alpha}\Big)\sum_{i=2}^{n}\Big\{\frac{x_i-\gamma_0}{\alpha_0}+\frac{\gamma_0-\widehat{\gamma}_{init}}{\alpha_0}\Big\}\\ &+\frac{\widehat{\beta}_w-1}{n-1}\sum_{i=2}^{n}\log\Big\{1-B_n e^{-\left(\frac{\alpha_0}{\alpha}\right)\left(\frac{x_i-\gamma_0}{\alpha_0}\right)}\Big\},\end{aligned}$$

where $B_n = e^{-\frac{\gamma_0-\widehat{\gamma}_{init}}{\alpha}}$. From the Weak Law of Large Numbers and by using the fact that $B_n \xrightarrow{P} 1$, we have

$$\frac{1}{n-1}\sum_{i=2}^{n}\log\Big\{1-B_n e^{-\left(\frac{\alpha_0}{\alpha}\right)\left(\frac{x_i-\gamma_0}{\alpha_0}\right)}\Big\} \xrightarrow{P} E\Big(\log\Big\{1-e^{-\frac{\alpha_0}{\alpha}Z_1}\Big\}\Big). \tag{31}$$

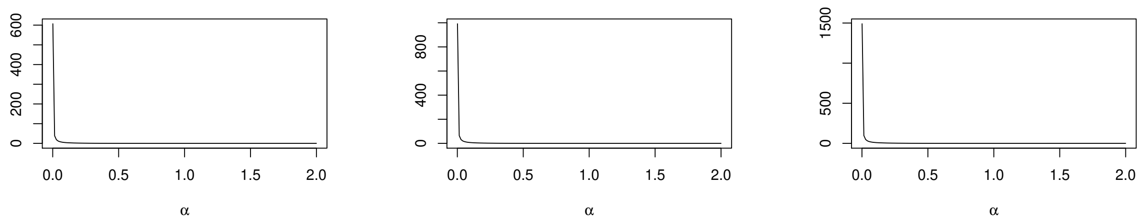
. It can be obtained that $E\Big(\log\Big\{1-e^{-\frac{\alpha_0}{\alpha}Z_1}\Big\}\Big) = -\beta_0\sum_{i=1}^{\infty}\frac{B(\beta_0, i\alpha_0/\alpha+1)}{i}$. Therefore, by using the Slutsky's Theorem, we have

$$\frac{1}{n-1}\log \ell_t(\alpha) \xrightarrow{P} \log\beta_0 - \log\alpha - \frac{\alpha_0}{\alpha}\big(\psi(\beta_0+1)-\psi(1)\big) - \beta_0(\beta_0-1)\sum_{i=1}^{\infty}\frac{B(\beta_0, i\alpha_0/\alpha+1)}{i}. \tag{32}$$

Hence,

$$\begin{aligned}\frac{1}{n-1}\Big(\log \ell_t(\alpha_0) - \log \ell_t(\alpha)\Big) \xrightarrow{P} -\Big[&\log\Big(\frac{\alpha_0}{\alpha}\Big) - \Big(1-\frac{\alpha_0}{\alpha}\Big)\big(\psi(\beta_0+1)-\psi(1)\big)\\ &+(\beta_0-1)\Big(\frac{1}{\beta_0}-\beta_0\sum_{i=1}^{\infty}\frac{B(\beta_0, i\alpha_0/\alpha+1)}{i}\Big)\Big].\end{aligned} \tag{33}$$

It is not easy to show theoretically that the right hand side of (33) is positive for all values of α. But numerically, but it has been verified using the well-known mathematical software mathematica. Without lose of generality, assuming that $\alpha_0 = 1$, some plots of the expression in right hand side of (33) for various values of β_0 are presented in Figure 3. Therefore, it confirms the results in (27). Theorem 3.7 of Chapter 6 in Lehmann and Casella (2006) along with this result prove the consistency of $\widehat{\alpha}_{init}$.



$\beta_0 = 0.5$ $\beta_0 = 1$ $\beta_0 = 2$

Figure 3: Plots of the expression in right hand side of (33) for various values of β_0.