

# ON THE COMPARISON OF FISHER INFORMATION OF THE WEIBULL AND GE DISTRIBUTIONS

Rameshwar D. Gupta <sup>†</sup>  
Debasis Kundu<sup>\*</sup>

## Abstract

In this paper we consider the Fisher information matrices of the generalized exponential (GE) and Weibull distributions for complete and Type-I censored observations. Fisher information matrix can be used to compute asymptotic variances of the different estimators. Although both distributions may provide similar data fit but the corresponding Fisher information matrices can be quite different. Moreover, the percentage loss of information due to truncation of the Weibull distribution is much more than the GE distribution. We compute the total information of the Weibull and GE distributions for different parameter ranges. We compare the asymptotic variances of the median estimators and the average asymptotic variances of all the percentile estimators for complete and Type-I censored observations. One data analysis has been preformed for illustrative purposes. When two fitted distributions are very close to each other and very difficult to discriminate otherwise, the Fisher information or the above mentioned asymptotic variances may be used for discrimination purposes.

KEYWORDS AND PHRASES: Fisher information matrix; Generalized exponential distribution; Hazard function; Median estimators; Model discrimination; Type-I censoring; Weibull distribution.

CORRESPONDING AUTHOR: <sup>\*</sup>Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, INDIA; Phone: 91-512-597141, Fax: 91-512-597500; e-mail: kundu@iitk.ac.in.

Postal address: <sup>†</sup> Department of Computer Science and Statistics, The University of New Brunswick at Saint John, New Brunswick, Canada E2L 4L5. Part of the work has been supported by a discovery grant from NSERC, CANADA.

# 1 INTRODUCTION

Recently the two-parameter generalized exponential (GE) distribution has been proposed by the authors. It has been studied extensively by Gupta and Kundu [6], [7], [8], [9], [10], Raqab [17] [18], Raqab and Ahsanullah [19] and Zheng [21]. For  $x > 0$ , the two-parameter generalized exponential distribution has the density function

$$f_{GE}(x; \alpha, \lambda) = \alpha \lambda e^{-\lambda x} (1 - e^{-\lambda x})^{\alpha-1}, \quad \alpha, \lambda > 0.$$

Here  $\alpha$  and  $\lambda$  are the shape and scale parameters respectively. For different values of the shape parameters, the density function can take different shapes. For  $0 < \alpha \leq 1$ , the density function is a decreasing function and for  $\alpha > 1$ , it becomes an uni-modal function. The different shapes of the density function can be seen in Gupta and Kundu [7]. The hazard function of the GE distribution can be increasing or decreasing depending on the shape parameter  $\alpha$ . It has been observed that the GE distribution can be used quite effectively to analyze skewed data set and it is a good alternative to the well known Weibull distribution. In many situations the GE distribution might provide a better data *fit* than the Weibull distribution. Since the distribution function of the GE distribution is in a closed form, therefore GE distribution also can be used very conveniently if the data are censored like Weibull distribution. Interestingly, even though the distribution function and the hazard function have convenient expressions for both Weibull and GE distributions, but the Fisher information matrices are not very convenient in both cases.

Computation of the Fisher information for any particular distribution is quite important and recently it has received some special attention, see for example Efron and Johnstone [3], Gertsbakh [4], Zheng and Gastwirth [20], Zheng [21] and the references therein. Fisher information can be used to compute the asymptotic variances of the different functions of the estimators. The problem is quite important when the data are censored. Sometimes

it is observed that two particular distributions may provide very similar data fit to a given data set. The distance between the two fitted distributions can be very small, and it may be very difficult to discriminate between them. The asymptotic variances of some function of the estimators which represent the common features of both distributions may be used to discriminate between them.

We compute the Fisher information matrices of both distributions under different conditions, namely for (i) complete observation and for (ii) Type-I censored observation. It is well known (Lehmann [14]) that the Fisher information matrix can be computed using the first and second derivative of the log-likelihood function. Efron and Johnstone [3] showed that the Fisher information matrix can be computed using hazard function also.

Recently it is observed by the authors (Gupta and Kundu [10]) that the GE and Weibull distributions provide a very similar data fit and one needs a large sample size to discriminate between the two distribution functions at least for certain ranges of the shape parameters. Interestingly, it is observed that although both GE and Weibull distributions can provide similar data fit but their Fisher information matrices can be quite different. We compute the Fisher information matrices of both distributions and study some interesting properties for complete and Type-I censored observations. Using the Fisher information, we compute the asymptotic variances of the median estimators for both distributions and they are used for model discrimination.

We denote the Weibull density function (for  $x > 0$ ) as

$$f_{WE}(x; \beta, \theta) = \beta\theta(\theta x)^{\beta-1}e^{-(\theta x)^\beta}; \quad \beta > 0, \theta > 0.$$

The generalized exponential distribution with the shape and scale parameters  $\alpha$  and  $\lambda$  respectively will be denoted by GE( $\alpha, \lambda$ ). Similarly the Weibull distribution with the shape and scale parameters  $\beta$  and  $\theta$  respectively will be denoted by WE( $\beta, \theta$ ). The rest of the

paper is organized as follows. In section 2, we give the preliminaries. In sections 3 and 4 we provide results for complete sample and Type-I censored sample respectively. We analyze one data set in section 5 and finally conclude the paper in section 6.

## 2 PRELIMINARIES

### 2.1 FISHER INFORMATION MATRIX FOR COMPLETE SAMPLE

Let  $X$  be a continuous random variable with the cumulative distribution function (CDF)  $F(x; \theta_1, \theta_2)$  and the probability density function (PDF)  $f(x; \theta_1, \theta_2)$ . For brevity, we consider only two parameters  $\theta_1$  and  $\theta_2$ , although the results are true for any finite dimensional vector. Under the standard regularity conditions (Lehmann [14]), the Fisher information matrix for the parameter vector  $\boldsymbol{\theta} = (\theta_1, \theta_2)$  based on an observation in terms of the expected values of the first and second derivatives of the log-likelihood function is provided in Lehmann [14].

Efron and Johnstone [3] observed that the Fisher information matrix can be obtained using the hazard function  $h(x, \boldsymbol{\theta}) = \frac{f(x, \boldsymbol{\theta})}{\bar{F}(x, \boldsymbol{\theta})} = -\frac{d}{dx} \ln \bar{F}(x; \boldsymbol{\theta})$  as follows

$$\mathbf{I}(\boldsymbol{\theta}) = E\left( \begin{bmatrix} \frac{\partial}{\partial \theta_1} \ln h(X, \boldsymbol{\theta}) \\ \frac{\partial}{\partial \theta_2} \ln h(X, \boldsymbol{\theta}) \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial \theta_1} \ln h(X, \boldsymbol{\theta}) & \frac{\partial}{\partial \theta_2} \ln h(X, \boldsymbol{\theta}) \end{bmatrix} \right). \quad (1)$$

The derivation of (1) is simple, and it mainly follows using the definition of the hazard function in terms of the density function (see Efron and Johnstone [3]). It has some interesting statistical and probabilistic implications also, noted by Efron and Johnstone [3].

### 2.2 FISHER INFORMATION MATRIX UNDER FIXED TIME CENSORING

If the observation of  $X$  is right censored at a fixed time point  $T$ , *i.e.* one observes  $\min(X, T)$ , the Fisher information for the parameter vector  $\boldsymbol{\theta}$  based on a censored observation (see

Gertsbakh [4] or Lawless [13]) is

$$\mathbf{I}_R^{(c)}(T, \boldsymbol{\theta}) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \frac{1}{\bar{F}(T, \boldsymbol{\theta})} \begin{bmatrix} \frac{\partial}{\partial \theta_1} \bar{F}(T, \boldsymbol{\theta}) \\ \frac{\partial}{\partial \theta_2} \bar{F}(T, \boldsymbol{\theta}) \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial \theta_1} \bar{F}(T, \boldsymbol{\theta}) & \frac{\partial}{\partial \theta_2} \bar{F}(T, \boldsymbol{\theta}) \end{bmatrix},$$

where for  $i, j = 1, 2$ ,

$$a_{ij} = \int_0^T \left( \frac{\partial}{\partial \theta_i} \ln f(x, \boldsymbol{\theta}) \right) \left( \frac{\partial}{\partial \theta_j} \ln f(x, \boldsymbol{\theta}) \right) f(x, \boldsymbol{\theta}) dx,$$

and  $\bar{F}(T, \boldsymbol{\theta}) = 1 - F(T, \boldsymbol{\theta})$ . It can be easily seen using the similar arguments of Gupta, Gupta and Sankaran [5] that

$$\mathbf{I}_R^{(c)}(T, \boldsymbol{\theta}) = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, \quad (2)$$

where for  $i, j = 1, 2$ ,

$$b_{ij} = \int_0^T \left( \frac{\partial}{\partial \theta_i} \ln h(x, \boldsymbol{\theta}) \right) \left( \frac{\partial}{\partial \theta_j} \ln h(x, \boldsymbol{\theta}) \right) f(x, \boldsymbol{\theta}) dx.$$

### 3 FISHER INFORMATION MATRICES : COMPLETE SAMPLE

Let us denote the Fisher information matrix of the GE parameters  $\alpha$  and  $\lambda$  for complete sample by

$$\mathbf{I}_G(\lambda, \alpha) = \begin{bmatrix} a_{11G} & a_{12G} \\ a_{21G} & a_{22G} \end{bmatrix}.$$

The elements are

$$\begin{aligned} a_{11G} &= \frac{1}{\alpha^2}, \\ a_{12G} &= a_{21G} = -\frac{\alpha}{\lambda} \int_0^\infty x e^{-2x} (1 - e^{-x})^{\alpha-2} dx, \\ a_{22G} &= \frac{1}{\lambda^2} + \frac{\alpha(\alpha-1)}{\lambda^2} \int_0^\infty x^2 e^{-2x} (1 - e^{-x})^{\alpha-3} dx. \end{aligned}$$

For  $\alpha \neq 1$ ,

$$a_{12G} = a_{21G} = \frac{1}{\lambda} \left[ \frac{\alpha}{\alpha-1} (\psi(\alpha) - \psi(1)) - (\psi(\alpha+1) - \psi(1)) \right]$$

and for  $\alpha = 1$ ,

$$a_{12G} = a_{21G} = -\frac{1}{\lambda} \sum_{j=0}^{\infty} \frac{1}{(2+j)^2}.$$

Similarly for  $\alpha \neq 2$ ,

$$a_{22G} = \frac{1}{\lambda^2} \left[ 1 + \frac{\alpha(\alpha-1)}{\alpha-2} \left( \psi'(1) - \psi'(\alpha-1) + (\psi(\alpha-1) - \psi(1))^2 \right) \right] - \frac{\alpha}{\lambda^2} \left[ \psi'(1) - \psi(\alpha) + (\psi(\alpha) - \psi(1))^2 \right]$$

and for  $\alpha = 2$

$$a_{22G} = \frac{1}{\lambda^2} + \frac{4}{\lambda^2} \sum_{j=0}^{\infty} \frac{1}{(2+j)^3}.$$

Now we compute the Fisher information matrix of the Weibull distribution. Let us denote

$$\mathbf{I}_W(\beta, \theta) = \begin{bmatrix} a_{11W} & a_{12W} \\ a_{21W} & a_{22W} \end{bmatrix},$$

as the Fisher information matrix of the Weibull distribution where the elements are as follows

$$a_{11W} = \frac{1}{\beta^2} \left( \psi'(1) + \psi^2(2) \right); \quad a_{12W} = a_{21W} = \frac{1}{\theta} (1 + \psi(1)); \quad a_{22W} = \frac{\beta^2}{\theta^2}.$$

Some of the interesting points are observed comparing the Fisher information matrices for the two distributions. First of all if we assume that the scale parameters are known in both cases then the Fisher information of the shape parameters are inversely proportional to the corresponding shape parameters. Similarly, if the shape parameters are assumed to be known then the Fisher information of the scale parameters are inversely proportional to the corresponding scale parameters. In this respect they are quite similar in nature.

Now let us consider the total information measure of a given Fisher information matrix. We consider the trace of the Fisher information matrix as its measure of the total information. It is similar to the E-optimality of the design of experiment problems. Note that the trace of the Fisher information matrix of the GE (Weibull) distribution is the sum of the information measure of  $\alpha$  ( $\beta$ ), when  $\lambda$  ( $\theta$ ) is known, and  $\lambda$  ( $\theta$ ), when  $\alpha$  ( $\beta$ ) is known.

It should be mentioned that although  $\alpha$ ,  $\beta$  and  $\lambda$ ,  $\theta$  are the shapes and scales parameters of the corresponding GE and Weibull distributions respectively but the parameters do not characterize the same features of the corresponding lifetime distributions. To compare the same features of both distributions, we compute the asymptotic variance (inversely proportional to the information measure) of the median estimators for both distributions. Note that the  $p$ -th ( $0 < p < 1$ ) percentile points of the GE and Weibull distributions are

$$P_{GE}(\alpha, \lambda) = -\frac{1}{\lambda} \ln(1 - p^{\frac{1}{\alpha}}) \quad \text{and} \quad P_{WE}(\beta, \theta) = \frac{1}{\theta} [-\ln(1 - p)]^{\frac{1}{\beta}}, \quad (3)$$

respectively. Therefore, the asymptotic variance of the  $p$ -th percentile estimator of the GE distribution, see Lewless [13], is

$$V_{GE}(p) = \begin{bmatrix} \frac{\partial P_{GE}}{\partial \alpha} & \frac{\partial P_{GE}}{\partial \lambda} \end{bmatrix} \begin{bmatrix} a_{11G} & a_{12G} \\ a_{21G} & a_{22G} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial P_{GE}}{\partial \alpha} \\ \frac{\partial P_{GE}}{\partial \lambda} \end{bmatrix},$$

and the asymptotic variance of the  $p$ -th percentile estimator of the Weibull distribution is

$$V_{WE}(p) = \begin{bmatrix} \frac{\partial P_{WE}}{\partial \beta} & \frac{\partial P_{WE}}{\partial \theta} \end{bmatrix} \begin{bmatrix} a_{11W} & a_{12W} \\ a_{21W} & a_{22W} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial P_{WE}}{\partial \beta} \\ \frac{\partial P_{WE}}{\partial \theta} \end{bmatrix}.$$

Now to compare the information measures of the two distributions, we feel it is reasonable to compare  $V_{GE}(0.5)$  with  $V_{WE}(0.5)$  or the comparison between  $\int_0^1 V_{GE}(p) dp$  and  $\int_0^1 V_{WE}(p) dp$  is meaningful. Note that  $\int_0^1 V_{GE}(p) dp$  and  $\int_0^1 V_{WE}(p) dp$  represent the average asymptotic variances of the percentile estimators over all percentile points.

We compute the different information measures of the two distribution functions at their *closest* values. The distance between two distribution functions can be defined in several ways. We consider the Kullback-Leibler distance, see Gupta and Kundu [10] for details, as the distance between two distribution functions. In the same paper, [10], we reported  $\tilde{\alpha}$ ,  $\tilde{\lambda}$  for a given  $\beta$  such that  $GE(\tilde{\alpha}, \tilde{\lambda})$  is *closest* to  $WE(\beta, 1)$  with respect to Kullback-Leibler distance. Similarly, we also reported  $\tilde{\beta}$ ,  $\tilde{\theta}$  for a given  $\alpha$  such that  $WE(\tilde{\beta}, \tilde{\theta})$  is *closest* to  $GE(\alpha, 1)$  with respect to the same distance function. To compute the total information of the

| $\beta$ | $\tilde{\alpha}$ | $\lambda$ | Trace<br>Weibull | Trace<br>GE | Total-Var<br>Weibull | Total-Var<br>GE |
|---------|------------------|-----------|------------------|-------------|----------------------|-----------------|
| 0.6     | 0.474            | 0.410     | 5.42             | 7.46        | 3.28                 | 0.77            |
| 0.8     | 0.722            | 0.721     | 3.49             | 3.35        | 2.11                 | 1.86            |
| 1.2     | 1.390            | 1.307     | 2.71             | 1.30        | 1.64                 | 6.19            |
| 1.4     | 1.823            | 1.565     | 2.89             | 0.98        | 1.75                 | 9.80            |
| 1.6     | 2.334            | 1.802     | 3.27             | 0.81        | 1.98                 | 16.2            |
| 1.8     | 2.885            | 2.006     | 3.80             | 0.72        | 2.30                 | 24.0            |
| 2.0     | 3.639            | 2.239     | 4.46             | 0.66        | 2.70                 | 33.0            |

Table 1: The trace of the Fisher information matrices of  $WE(\beta, 1)$  and  $GE(\tilde{\alpha}, \tilde{\lambda})$  are reported in columns 4 and 5. The total asymptotic variances of the shape and scale parameters for both distributions are also presented in columns 6 and 7. The values of  $\tilde{\alpha}$  and  $\tilde{\lambda}$  are obtained from Gupta and Kundu [10].

| $\alpha$ | $\tilde{\beta}$ | $\tilde{\theta}$ | Trace<br>Weibull | Trace<br>GE | Total-Var<br>Weibull | Total-Var<br>GE |
|----------|-----------------|------------------|------------------|-------------|----------------------|-----------------|
| 0.5      | 0.649           | 2.243            | 4.41             | 4.53        | 14.2                 | 2.96            |
| 1.5      | 1.257           | 0.735            | 4.07             | 1.86        | 1.35                 | 6.00            |
| 2.0      | 1.440           | 0.609            | 6.47             | 2.06        | 1.45                 | 10.3            |
| 2.5      | 1.585           | 0.537            | 9.44             | 2.33        | 1.65                 | 16.6            |
| 3.0      | 1.706           | 0.488            | 12.85            | 2.61        | 1.85                 | 26.1            |

Table 2: The trace of the Fisher information matrices of  $WE(\tilde{\beta}, \tilde{\theta})$  and  $GE(\alpha, 1)$  are reported in columns 4 and 5. The total asymptotic variances of the shape and scale parameters for both distributions are also presented in columns 6 and 7. The values of  $\tilde{\beta}$  and  $\tilde{\theta}$  are obtained from Gupta and Kundu [10].

Fisher information matrices we compute the traces of the corresponding Fisher information matrices for both distributions. The results are reported in Tables 1 and 2. We also compute  $V_{GE}(0.5)$ ,  $V_{WE}(0.5)$ ,  $\int_0^1 V_{GE}(p)dp$  and  $\int_0^1 V_{WE}(p)dp$ . The results are reported in Tables 3 and 4.

From Tables 1 and 2, it is observed that for both distributions when the scale parameter is fixed, the total information increases as the shape parameter moves away from one. The total asymptotic variances also show the expected trend. Comparing the Tables 3 and 4 it is observed that for the shape parameter less than one, the average asymptotic variance of



| $\beta \rightarrow$    | 0.6     | 0.8    | 1.2    | 1.4    | 1.6    | 1.8    | 2.0    |
|------------------------|---------|--------|--------|--------|--------|--------|--------|
| $V_{WE}(0.5)$          | 1.1287  | 0.8617 | 0.5197 | 0.4167 | 0.3406 | 0.2832 | 0.2389 |
| $V_{GE}(0.5)$          | 1.3190  | 0.8082 | 0.4332 | 0.3487 | 0.2908 | 0.2514 | 0.2145 |
| $\int_0^1 V_{WE}(p)dp$ | 45.8192 | 7.8307 | 1.4199 | 0.8705 | 0.5993 | 0.4451 | 0.3482 |
| $\int_0^1 V_{GE}(p)dp$ | 13.5675 | 4.4938 | 1.4248 | 1.0101 | 0.7721 | 0.6294 | 0.5104 |

Table 3: The asymptotic variances of the median estimators of  $WE(\beta,1)$ ,  $GE(\tilde{\alpha}, \tilde{\lambda})$  are reported and the average asymptotic variances over all the percentile points are also reported.

| $\alpha \rightarrow$   | 0.5    | 1.5    | 2.0    | 2.5    | 3.0    |
|------------------------|--------|--------|--------|--------|--------|
| $V_{WE}(0.5)$          | 0.2103 | 0.9015 | 1.0776 | 1.1985 | 1.2944 |
| $V_{GE}(0.5)$          | 0.2447 | 0.7736 | 0.8895 | 0.9673 | 1.0230 |
| $\int_0^1 V_{WE}(p)dp$ | 5.2905 | 2.2516 | 2.2134 | 2.1309 | 2.1331 |
| $\int_0^1 V_{GE}(p)dp$ | 2.2871 | 2.4457 | 2.4829 | 2.5172 | 2.5379 |

Table 4: The asymptotic variances of the median estimators of  $WE(\tilde{\beta}, \tilde{\theta})$ ,  $GE(\alpha, 1)$  and the average asymptotic variances over all the percentile points are also reported.

the percentile estimators is smaller for the GE model than the Weibull model, and for the shape parameter greater than one it is the other way. Asymptotic variances of the median estimators for the GE model are usually smaller than the Weibull model unless the shape parameter is small.

## 4 FISHER INFORMATION MATRICES: RIGHT CENSORING

In this case let us denote the Fisher information matrix of the GE parameters  $\alpha$  and  $\lambda$  by

$$\mathbf{I}_{G,R}^{(c)}(T, \alpha, \lambda) = \begin{bmatrix} b_{11G} & b_{12G} \\ b_{21G} & b_{22G} \end{bmatrix}.$$

Using (2), it can be easily seen that for  $p = (1 - e^{-\lambda T})^\alpha$

$$b_{11G} = \frac{1}{\alpha^2} \int_0^p \left[ 1 + \frac{\ln y}{1-y} \right]^2 dy = \frac{1}{\alpha^2} \left[ p + \frac{p}{1-p} (\ln p)^2 \right],$$

$$b_{12G} = b_{21G} = \frac{\alpha}{\lambda} \int_0^{p^{\frac{1}{\alpha}}} \left[ \frac{1}{\alpha} + \frac{\ln y}{1-y^\alpha} \right] \left[ 1 + \frac{\ln(1-y)}{y} - \frac{\alpha(1-y) \ln(1-y)}{y(1-y^\alpha)} \right] y^{\alpha-1} dy,$$

$$b_{22G} = \frac{\alpha}{\lambda^2} \int_0^{p^{\frac{1}{\alpha}}} \left[ 1 + \frac{\ln(1-y)}{y} - \frac{\alpha(1-y)\ln(1-y)}{y(1-y^\alpha)} \right]^2 y^{\alpha-1} dy.$$

Now let us define the Fisher information matrix for the Weibull parameters  $\beta$  and  $\theta$  by

$$\mathbf{I}_{W,R}^{(c)}(T, \beta, \theta) = \begin{bmatrix} b_{11W} & b_{12W} \\ b_{21W} & b_{22W} \end{bmatrix}.$$

In this case also using (2) and for  $\tilde{p} = 1 - e^{-(\theta T)^\beta}$ , we obtain

$$\begin{aligned} b_{11W} &= \frac{1}{\beta^2} \int_0^{-\ln(1-\tilde{p})} [1 + \ln y]^2 e^{-y} dy, \\ b_{12W} = b_{21W} &= \frac{1}{\theta} \int_0^{-\ln(1-\tilde{p})} [1 + \ln y] e^{-y} dy, \\ b_{22W} &= \frac{\beta^2}{\theta^2} \tilde{p}. \end{aligned}$$

Now we compute the total information measures of the two information matrices for different parameter values and for different  $T$ . Similarly as before we compute the different information measures at their closest values. We consider different  $T$ 's namely  $T \approx \text{mean}$  and  $T \approx \text{mean} + \text{standard deviation}$ .

There are mainly two reasons to choose these two  $T$  values. First of all once we consider these two  $T$  values then the corresponding  $p$  or  $\tilde{p}$  values become independent of the scale parameters. Secondly, once we truncate at the mean (mean + sd) then we observe at least 57%(86%) and 43%(87%) observations for GE and Weibull cases respectively. If we choose  $T$  lower than mean then the censoring becomes quite heavy.

We compute the loss of information (in percentage) in each case and the results are reported in Tables 5 and 6. Interestingly it is observed at  $T \approx \text{mean}$  the loss of information for Weibull distribution is approximately between 44% to 49% and for the GE distribution it is approximately 6% to 25%. When  $T \approx \text{mean} + \text{standard deviation}$  then the loss of information for Weibull is approximately 22% to 25% and for GE distribution it is approximately 2% to 8%. Therefore, for a given truncated data set the percentage loss of information for the Weibull parameters will be much more compared to the GE parameters.

| $\beta$ | $\tilde{\alpha}$ | $\lambda$ | Trace Weibull            | Trace GE                | Total-Var Weibull | Total-Var GE     |
|---------|------------------|-----------|--------------------------|-------------------------|-------------------|------------------|
| 0.6     | 0.474            | 0.410     | 2.85, 47%<br>(4.12, 24%) | 6.08, 19%<br>(6.98, 6%) | 4.26<br>(3.46)    | 1.48<br>(0.91)   |
| 0.8     | 0.722            | 0.721     | 1.77, 49%<br>(2.58, 26%) | 2.76, 18%<br>(3.14, 6%) | 3.16<br>(2.33)    | 3.53<br>(2.55)   |
| 1.2     | 1.390            | 1.307     | 1.42, 48%<br>(2.07, 24%) | 1.04, 20%<br>(1.21, 7%) | 3.20<br>(2.03)    | 10.55<br>(7.29)  |
| 1.4     | 1.823            | 1.565     | 1.54, 47%<br>(2.27, 22%) | 0.78, 20%<br>(0.92, 6%) | 3.75<br>(2.27)    | 16.87<br>(11.99) |
| 1.6     | 2.334            | 1.802     | 1.75, 47%<br>(2.62, 20%) | 0.64, 21%<br>(0.76, 6%) | 4.54<br>(2.66)    | 26.88<br>(19.43) |
| 1.8     | 2.885            | 2.006     | 2.03, 47%<br>(3.08, 19%) | 0.57, 21%<br>(0.67, 7%) | 5.54<br>(3.16)    | 41.24<br>(30.24) |
| 2.0     | 3.639            | 2.239     | 2.36, 47%<br>(3.65, 18%) | 0.52, 21%<br>(0.61, 7%) | 6.74<br>(3.76)    | 67.65<br>(50.88) |

Table 5: The trace of the Fisher information matrices of WE( $\beta,1$ ) and GE( $\tilde{\alpha}, \tilde{\lambda}$ ) are reported in columns 4 and 5, when the data are right truncated at  $T \approx \text{mean}$  and the loss of information from the complete sample are also presented in the same column. The corresponding total asymptotic variances of the shape and scale parameters for both distributions are presented in columns 6 and 7. The results corresponding to  $T \approx \text{mean} + \text{standard deviation}$  are reported within bracket in each case.

Comparing Tables 7 and 8 it is observed that because of truncation the asymptotic variances of the median estimators may not change much but the average asymptotic variances are changing quite significantly. It is observed that the average asymptotic variances for GE distribution are smaller than the Weibull distribution for all ranges of the shape parameters, although that is not true for the complete data. It shows that due to truncation the loss of information of the Weibull model is much more than the GE model.

Now we would like to compute the loss of information due to right truncation, in one parameter when the other parameter is known. First we consider the GE distribution. If the scale parameter is known then the loss of information of the shape parameter is

$$1 - \frac{b_{11G}}{a_{11G}} = 1 - \left[ p + \frac{p}{1-p} (\ln p)^2 \right], \quad (4)$$

| $\alpha$ | $\beta$ | $\theta$ | Trace Weibull             | Trace GE                | Total-Var Weibull | Total-Var GE     |
|----------|---------|----------|---------------------------|-------------------------|-------------------|------------------|
| 0.5      | 0.649   | 2.243    | 2.23, 49%<br>(3.16, 28%)  | 4.27, 6%<br>(4.42, 2%)  | 17.51<br>(14.23)  | 6.84<br>(3.98)   |
| 1.5      | 1.257   | 0.735    | 2.27, 44%<br>(1.70, 17%)  | 1.41, 24%<br>(1.72, 8%) | 2.76<br>(1.70)    | 9.38<br>(6.73)   |
| 2.0      | 1.440   | 0.609    | 3.64, 44%<br>(5.39, 17%)  | 1.54, 25%<br>(1.90, 8%) | 3.28<br>(1.93)    | 15.61<br>(11.50) |
| 2.5      | 1.585   | 0.537    | 5.29, 44%<br>(7.94, 16%)  | 1.76, 24%<br>(2.16, 7%) | 3.88<br>(2.23)    | 25.00<br>(18.54) |
| 3.0      | 1.706   | 0.488    | 7.14, 44%<br>(10.83, 16%) | 2.01, 23%<br>(2.44, 7%) | 4.49<br>(2.54)    | 37.83<br>(28.08) |

Table 6: The trace of the Fisher information matrices of  $WE(\tilde{\beta}, \tilde{\theta})$  and  $GE(\alpha, 1)$  are reported in columns 4 and 5 when the data are right truncated at  $T \approx \text{mean}$ . The corresponding total asymptotic variances of the shape and scale parameters for both distributions are presented in columns 6 and 7. The results corresponding to  $T \approx \text{mean} + \text{standard deviation}$  are reported within bracket in each case.

here  $p$  is same as before, *i.e.*  $p = (1 - e^{-\lambda T})^\alpha$ . Therefore, the ratio (4) depends on  $\lambda$  and  $\alpha$  through  $p$  only. We plot the ratio (4) in Figure 1, as a function of  $\alpha$  when  $T = \text{the mean of GE}$ . From Figure 1, it is clear that the loss of information is an increasing function of  $\alpha$ . However, more than 99% of the information on shape parameter is retained by considering right censoring at the mean, when the  $\alpha \leq 5$ .

Similarly, we look at the loss of information on the scale parameter, when the shape parameter is known. Clearly, the ratio  $1 - \frac{b_{22G}}{a_{22G}}$  is a function of  $p$  and  $\alpha$ . However, when  $T = \text{mean or mean} + \text{sd}$ , it is a function of  $\alpha$  only. We plot this ratio in Figure 2, as a function of  $\alpha$ , at  $T = \text{mean of GE}$ . It is clear that there is a big impact of  $\alpha$  on this loss of information of  $\lambda$ . Contrary to the shape parameter case, the loss of information is a decreasing function of the shape parameter.

Now, we consider the Weibull distribution. The loss of information of the shape parameter

| $\beta \rightarrow$    | 0.6                   | 0.8                  | 1.2                | 1.4                | 1.6                | 1.8                | 2.0                |
|------------------------|-----------------------|----------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| $V_{WE}(0.5)$          | 1.2027<br>(1.1324)    | 0.9582<br>(0.8670)   | 0.6244<br>(0.5252) | 0.5172<br>(0.4216) | 0.4350<br>(0.3450) | 0.3710<br>(0.2870) | 0.3202<br>(0.2423) |
| $V_{GE}(0.5)$          | 1.6343<br>(1.3696)    | 1.0395<br>(0.8533)   | 0.5700<br>(0.4574) | 0.4602<br>(0.3676) | 0.3840<br>(0.3061) | 0.3317<br>(0.2643) | 0.2825<br>(0.2251) |
| $\int_0^1 V_{WE}(p)dp$ | 106.2347<br>(61.2300) | 20.5022<br>(10.8776) | 4.0914<br>(1.9771) | 2.5386<br>(1.1965) | 1.7492<br>(0.8113) | 1.2919<br>(0.5934) | 1.0019<br>(0.4576) |
| $\int_0^1 V_{GE}(p)dp$ | 31.9521<br>(18.4462)  | 9.6217<br>(5.7080)   | 2.7345<br>(1.7031) | 1.8656<br>(1.1848) | 1.3809<br>(0.8927) | 1.0984<br>(0.7199) | 0.8695<br>(0.5785) |

Table 7: The variances of the median estimators of  $WE(\beta,1)$  and  $GE(\tilde{\alpha}, \tilde{\lambda})$  when the data are right truncated at  $T \approx \text{mean}$  are reported. The results corresponding to  $T \approx \text{mean} + \text{standard deviation}$  are reported within bracket below in each case.

if the scale parameter is known, is

$$1 - \frac{b_{11W}}{a_{11W}} = 1 - \frac{1}{\psi^2(2) + \psi'(1)} \int_0^{(\theta T)^\beta} (1 + \ln y)^2 e^{-y} dy. \quad (5)$$

When  $T = \text{mean}$  ( $\text{mean} + \text{sd}$ ), the ratio (5) is a function of  $\theta$ ,  $\beta$  and  $T$ , moreover it is an increasing function of  $\beta$ . We plot this ratio at  $T = \text{mean}$  of Weibull for different values of  $\beta$  in Figure 3.

When the shape parameter is known, the loss of information on the scale parameter due to truncation is

$$1 - \frac{b_{22W}}{a_{22W}} = 1 - \tilde{p} = e^{-(\theta T)^\beta}. \quad (6)$$

The ratio (6) is plotted in Figure 4. The loss of information is an increasing function of the  $\beta$ . However, the impact is relatively less than that for the shape parameter case.

## 5 DATA ANALYSIS AND DISCUSSIONS

For illustration purposes, we analyze one data set (Linhart and Zucchini [15], pp: 69), using both GE and Weibull models. The data set represents the failure times of the air conditioning

| $\alpha \rightarrow$   | 0.5                 | 1.5                | 2.0                | 2.5                | 3.0                |
|------------------------|---------------------|--------------------|--------------------|--------------------|--------------------|
| $V_{WE}(0.5)$          | 0.2264<br>(0.2111)  | 1.0934<br>(0.9113) | 1.3454<br>(1.0905) | 1.5278<br>(1.2137) | 1.6765<br>(1.3115) |
| $V_{GE}(0.5)$          | 0.3051<br>(0.2551)  | 1.0189<br>(0.8163) | 1.1740<br>(0.9372) | 1.2770<br>(1.0179) | 1.3498<br>(1.0753) |
| $\int_0^1 V_{WE}(p)dp$ | 12.7261<br>(7.1655) | 6.5220<br>(3.1252) | 6.3097<br>(2.9625) | 6.2225<br>(2.8880) | 6.2120<br>(2.8641) |
| $\int_0^1 V_{GE}(p)dp$ | 5.3113<br>(3.0746)  | 4.6409<br>(2.9068) | 4.5320<br>(2.9011) | 4.4640<br>(2.8986) | 4.4123<br>(2.8984) |

Table 8: The variances of the median estimators of WE( $\tilde{\beta}, \tilde{\theta}$ ) and GE( $\alpha, 1$ ) when the data are right truncated at  $T \approx$  mean are reported. The results corresponding to  $T \approx$  mean + standard deviation are reported within bracket below in each case.

system of an airplane.

When we use the Weibull model the MLEs are,  $\hat{\beta} = 0.8554$  and  $\hat{\theta} = 0.0183$ . Similarly, when we use the GE model, the MLEs are  $\hat{\alpha} = 0.8130$  and  $\hat{\lambda} = 0.0145$ . The estimated Fisher information matrices for the Weibull and GE models are

$$\mathbf{I}_W(\hat{\beta}, \hat{\theta}) = \begin{bmatrix} 2.6125 & 23.1033 \\ 23.1033 & 2184.9238 \end{bmatrix} \quad \mathbf{I}_G(\hat{\alpha}, \hat{\lambda}) = \begin{bmatrix} 1.5129 & -47.3124 \\ -47.3124 & 3952.2546 \end{bmatrix}$$

respectively. It may be mentioned that, here the estimated and the observed information matrices are very close to each other. Therefore, the estimated information matrix can be replaced by the observed information matrix, which is easier to calculate. The traces are 2187.5363 and 3953.7675. Moreover,  $V_{WE}(0.05) = 2394.3945$ ,  $\int_0^1 V_{WE}(p)dp = 16736.4023$ ,  $V_{GE}(0.50) = 2283.6055$  and  $\int_0^1 V_{GE}(p)dp = 11197.1875$ . Therefore, asymptotic variance of the median estimator and also the average asymptotic variance of the percentile estimators for the GE model are smaller than the Weibull model.

One point should be mentioned here that if we multiply the data by a constant  $c$ , then the information matrix changes for both distributions and therefore the trace also changes. Therefore, it is possible to choose  $c$  in such a manner that the trace of the new Weibull infor-

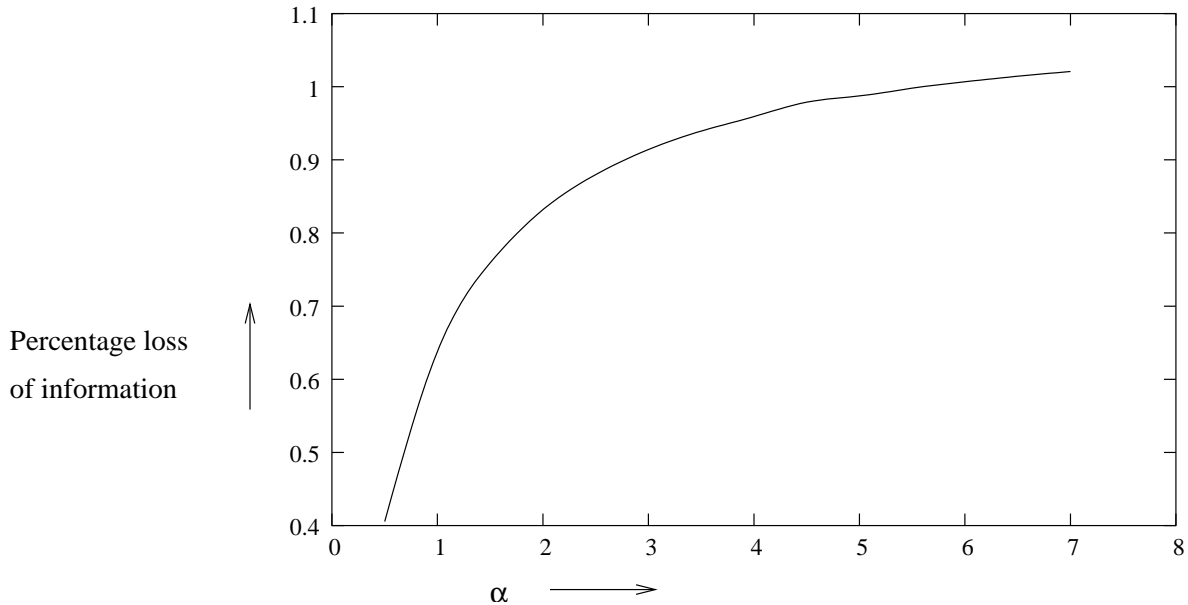


Figure 1: Percentage loss of information of the shape parameter of the GE distribution when the scale parameter is fixed and when the right truncation takes place at the mean value.

mation matrix becomes larger than the trace of the new GE information matrix. Although, the asymptotic variance of the percentile estimator also changes for the transformed data but the trend remains the same.

Note that the asymptotic variance of  $\hat{\alpha}$  is smaller than the asymptotic variance of  $\hat{\beta}$  and the asymptotic variance of  $\hat{\lambda}$  is larger than the asymptotic variance of  $\hat{\theta}$ . It is reflected in their profile likelihood surfaces also. We plot the profile likelihood surface of  $\beta(\alpha)$  for  $\theta = \hat{\theta}$  ( $\lambda = \hat{\lambda}$ ) in Figure 5. It clearly shows that the profile likelihood surface of  $\alpha$  is flatter than  $\beta$ . Similarly, we plot the profile likelihood surface of  $\theta(\lambda)$  for  $\beta = \hat{\beta}$  ( $\alpha = \hat{\alpha}$ ) in Figure 6. In this case the profile likelihood of  $\theta$  is flatter than  $\lambda$ .

Now let us look at the two fitted distributions. The maximum log-likelihood values are -152.007 and -152.264 for Weibull and GE models respectively. Similarly the Kolmogorv-Smirnov (K-S) distances between the empirical distribution function and the fitted distribution functions are as follows. For the Weibull case the K-S distance is 0.1539 and the

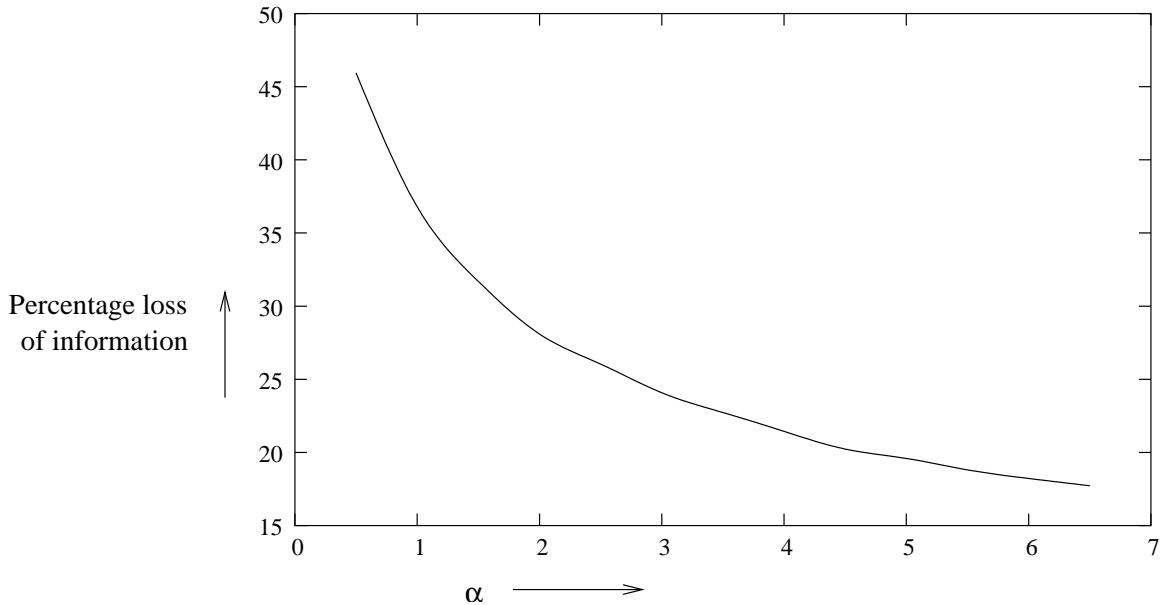


Figure 2: Percentage loss of information of the scale parameter of the GE distribution when the shape parameter is fixed and when the right truncation takes place at the mean value.

corresponding  $p$  value is 0.4755 and for the GE case the K-S distance is 0.1900 and the corresponding  $p$  values is 0.2286. We also plot the empirical survival function and the fitted survival functions in Figure 7. Both provide good fit. It is clear from the likelihood values, from the K-S distance and from Figure 7 that Weibull provides a marginally better fit than GE to the given data set. But since the two fitted distribution functions are very close to each other it is observed that the probability of correct selection (see Gupta and Kundu [10]) is also close to 0.5. In fact in this particular case, assuming that the data are coming from Weibull distribution the probability of correct selection is only 0.5224. Similarly, if we frame this model discrimination problem as a testing of hypothesis problem (see Kundu and Manglick [12]), then also we can not reject any of the null hypotheses. But comparing the asymptotic variances of the median estimators and also the average asymptotic variances, we prefer to use GE distribution rather than Weibull distribution for this particular data set.



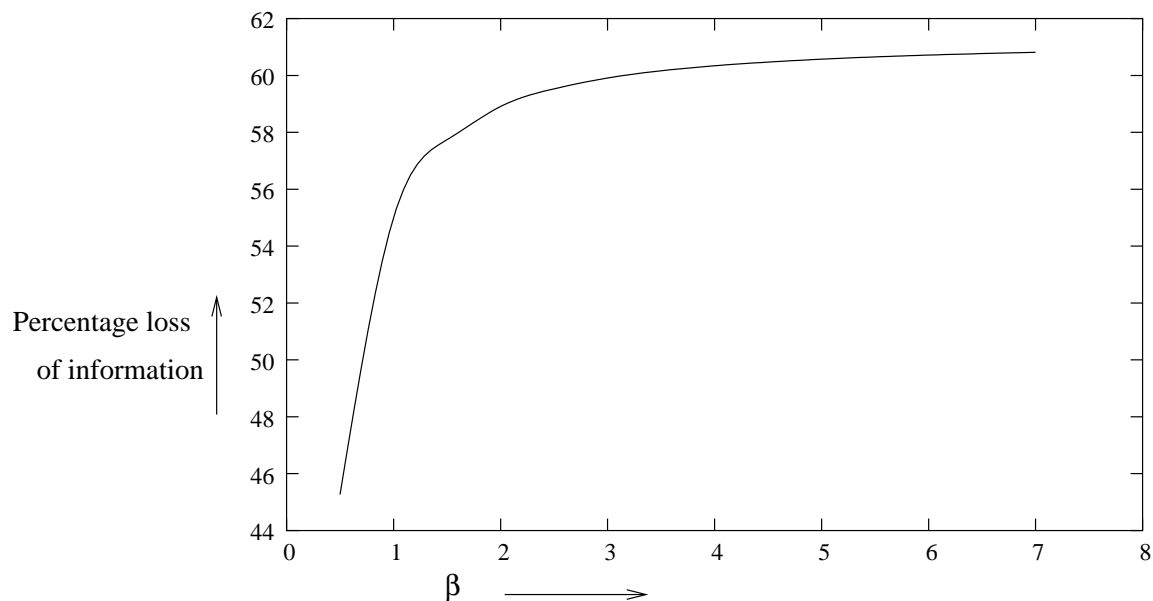


Figure 3: Percentage loss of information of the shape parameter of the Weibull distribution when the scale parameter is fixed and when the right truncation takes place at the mean value.

## 6 CONCLUSIONS

In this paper we compute the Fisher information matrices of the Weibull and GE distributions for both complete and Type-I censored data. We computed the Fisher information matrices using the likelihood function and hazard function approach. Recently, reversed hazard rate has been introduced and studied quite extensively by several authors, see for example Block, Savits and Singh [1], Chandra and Roy [2] and Gupta, Gupta and Sankaran [5]. It is observed by Gupta, Gupta and Sankaran [5] that the reversed hazard function also can be used to compute the Fisher information matrix. It is not pursued here, because it is observed that for complete or Type-I right censored data it may not be of much use. We have observed that the reversed hazard function can be used quite effectively if the data are left censored. Although, we have discussed only the complete and Type-I censored data, but similar results can be obtained even for Type-II data also. The work is in progress and it will be reported later.

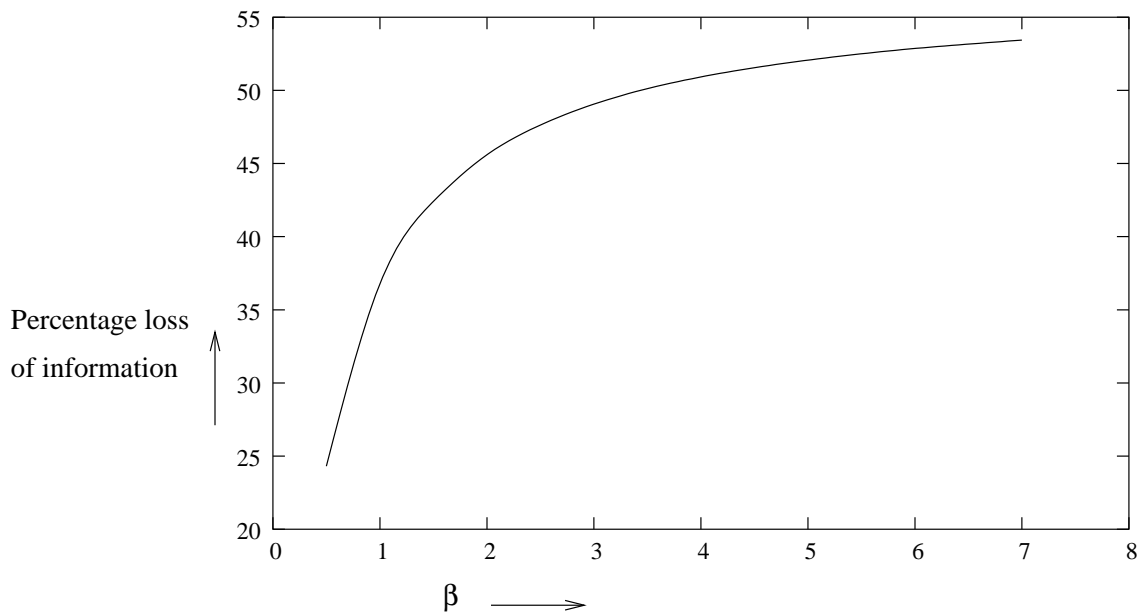


Figure 4: Percentage loss of information of the scale parameter of the Weibull distribution when the shape parameter is fixed and when the right truncation takes place at the mean value.

## ACKNOWLEDGMENTS

The authors would like to thank two referees and one of the associate editors for some very constructive comments.

## References

- [1] Block, H. W., Savits, T. H. and Singh, H. (1998), "On the reversed hazard rate function", *Probabilities in the Engineering and Information Sciences*, vol. 12, 69-90.
- [2] Chandra, N. N. and Roy, D. (2001), "Some results on reversed hazard rate", *Probabilities in the Engineering and Information Sciences*, vol. 15, 95-102.
- [3] Efron, B. and Johnstone, I. (1990), "Fisher information in terms of the hazard rate", *Annals of Statistics*, vol. 18, 38-62.

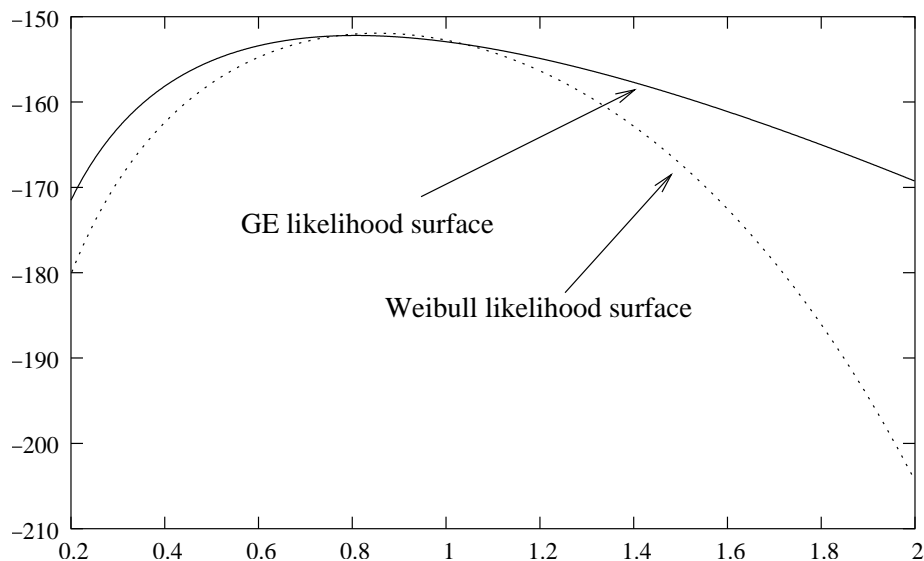


Figure 5: Profile likelihood of  $\beta$  and  $\alpha$  when  $\theta$  and  $\lambda$  are constant.

- [4] Gertsbakh, I. (1995), “On the Fisher Information in Type-I censored and quantal response data”, *Statistics and Probability Letters*, vol. 23, 297-306.
- [5] Gupta, R. D., Gupta, R.C. and Sankaran, P.G. (2004), “ Fisher Information in terms of the (reversed) hazard rate function”, to appear in the *Communications in Statistics - Theory and Methods*.
- [6] Gupta, R. D. and Kundu, D. (1999), “Generalized exponential distribution”, *Aust. NZ. Jour. Stat.*, vol. 41, 173-188.
- [7] Gupta, R. D. and Kundu, D. (2001a), “Exponentiated exponential family; an alternative to gamma and Weibull distributions”, *Biometrical Journal*, vol. 43, 117-130.
- [8] Gupta, R. D. and Kundu, D. (2001b), “Generalized exponential distribution: different methods of estimation”, *Journal of Statistical Computation and Simulation*, vol. 69, 315-337.
- [9] Gupta, R. D. and Kundu, D. (2002a), “Generalized exponential distribution: statistical inferences”, *Journal of Statistical Theory and Applications*, vol. 1, 101-118.

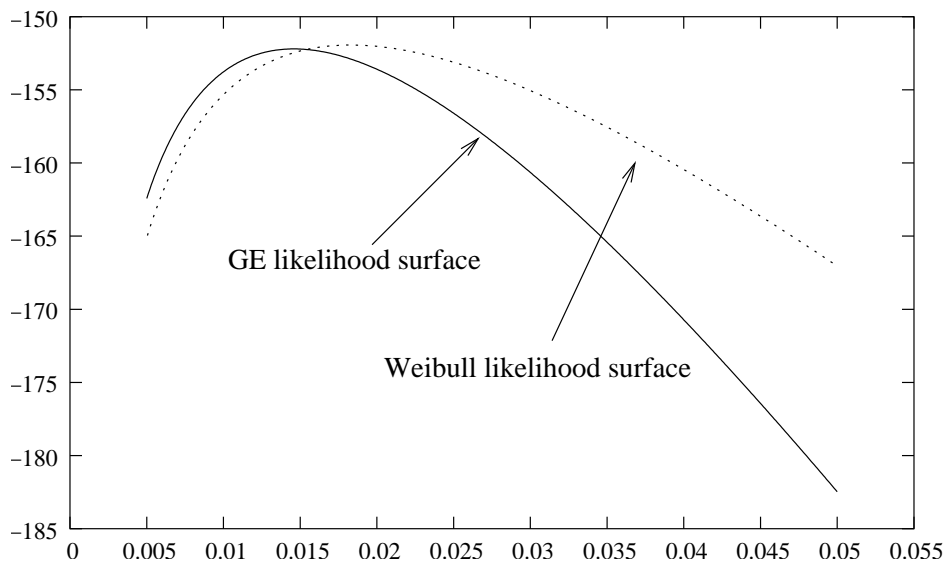


Figure 6: Profile likelihood of  $\theta$  and  $\lambda$  when  $\beta$  and  $\alpha$  are constant.

- [10] Gupta, R. D. and Kundu, D. (2003), “Discriminating between the Weibull and Generalized exponential distribution”, *Computational Statistics and Data Analysis*, vol. 43, 179-196.
- [11] Gupta, R. D. and Nanda, A. (2001), “Some results on reversed hazard rate ordering”, *Communications in Statistics - Theory and Methods*, vol. 30, 2447-2457.
- [12] Kundu, D. and Manglick, A. (2004), “Discriminating between the Weibull and log-normal distributions”, *Naval Research Logistics*, vol. 51, 893 - 905.
- [13] Lawless, J. F. (1982), *Statistical Models and Methods for Lifetime Data*, Wiley, New York.
- [14] Lehmann, E. L. (1991), *Theory of Point Estimation*, Wiley, New York.
- [15] Linhart, H. and Zucchini, W. (1986), *Model Selection*, Wiley, New York.
- [16] Park, S. (1996), “Fisher Information in order statistics”, *Journal of the American Statistical Association*, vol. 91, 385-390.

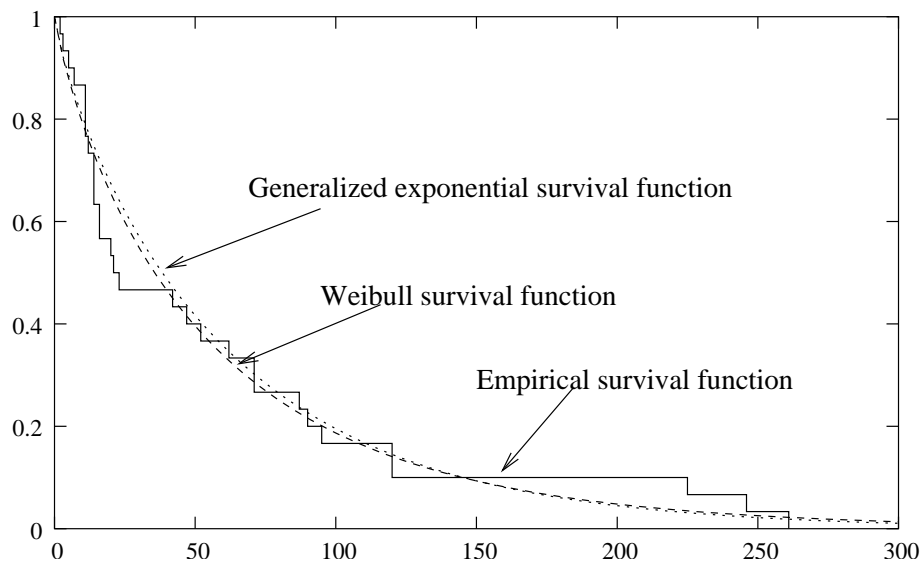


Figure 7: Empirical survival function and the fitted survival functions

- [17] Raqab, M. Z. (2002), “Inference for generalized exponential distribution based on record statistics”, *Journal of Statistical Planning and Inference*, vol. 104, 339-350.
- [18] Raqab, M. Z. (2004), “Generalized exponential distribution: moments of order statistics”, *Statistics*, vol. 38, 29 - 41.
- [19] Raqab, M.Z. and Ahsanullah, M. (2001), “Estimation of the location and scale parameters of the generalized exponential distribution based on order statistics”, *Journal of Statistical Computation and Simulation*, vol. 69, 109-124.
- [20] Zheng, G. and Gastwirth, J.L. (2000), “Where is the Fisher information in an ordered sample?”, *Statistica Sinica*, vol. 10, 1267 - 1280.
- [21] Zheng, G. (2002), “Fisher information matrix in Type-II censored data from exponentiated exponential family”, *Biometrical Journal*, vol. 44, 353-357.