.

GENERALIZED EXPONENTIAL DISTRIBUTION: STATISTICAL INFERENCES

RAMESHWAR D. GUPTA
DEPARTMENT OF APPLIED STATISTICS AND COMPUTER SCIENCE
THE UNIVERSITY OF NEW BRUNSWICK
SAINT JOHN, PO BOX. 5050
CANADA E2L 4L5

AND

DEBASIS KUNDU
DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLGY KANPUR
PIN 208016, INDIA

SUMMARY

Recently a new distribution, named as generalized exponential distribution or Exponentiated Exponential Distribution has been introduced and studied quite extensively by the authors. It is observed that a generalized exponential distribution is quite similar to a gamma or a Weibull distribution and it can be used as an alternative to a gamma model or a Weibull model in many situations. In this paper we consider the maximum likelihood estimation of the unknown parameters of a generalized exponential distribution for complete sample as well as for censored sample. We also consider the testing of hypotheses and the confidence intervals construction of the unknown parameters. We analyze one data set and see how the different methods work in practice.

Key words and phrases: Confidence interval, Fisher Information matrix, generalized exponential distribution, maximum likelihood estimation, test of hypothesis, uniformly most powerful test.

1. INTRODUCTION

Two-parameter gamma and two-parameter Weibull distribution are the most popular distributions for analyzing any lifetime data. The two parameters of a gamma or a Weibull distribution represent the scale and the shape parameters. Because of the shape and the scale parameters both distributions have quite a bit of flexibility for analyzing any positive real data set. Both distributions have increasing and decreasing hazard rate depending on the shape parameter. It gives an extra edge over an exponential distribution, which has only constant failure rate. Unfortunately, both distributions have certain drawbacks also. One major disadvantage of a gamma distribution is that the distribution function or the survival function can't be expressed in a closed form if the shape parameter is not an integer. Since it is in terms of a incomplete gamma function, one needs to obtain distribution function, survival function or hazard function by numerical integration. This makes gamma distribution unpopular compare to a Weibull distribution, which has a nice distribution function, survival function or hazard function. However, the Weibull distribution also has its own disadvantages. For example, it has been pointed out by Bain and Engelhardt (1991) that the maximum likelihood estimators of a Weibull distribution may not behave properly for all parameter ranges.

Recently a new distribution, named as generalized exponential distribution or Exponentiated Exponential Distribution has been introduced, and several properties were investigated by the authors (Gupta and Kundu; 1999, 2001a, 2001b), see also Raqab and Ahsanullah (2001). The generalized exponential (GE) distribution is defined in the following way. The distribution function $F(x;\alpha,\lambda)$ of GE is

$$F(x;\alpha,\lambda) = (1-e^{-\lambda x})^{\alpha}; \qquad \alpha,\lambda,x > 0. \tag{1.1}$$

Therefore, GE has a density function

$$f(x,\alpha,\lambda) = \alpha\lambda(1-e^{-\lambda x})^{\alpha-1}e^{-\lambda x}, \tag{1.2}$$

survival function

$$S(x;\alpha,\lambda) = 1-(1-e^{-\lambda x})^{\alpha}, \tag{1.3}$$

and a hazard function

$$h(x;\alpha,\lambda) = \frac{\alpha\lambda(1-e^{-\lambda x})^{\alpha-1}e^{-\lambda x}}{1-(1-e^{-\lambda x})^{\alpha}}. \tag{1.4}$$

Here α is the shape parameter and λ is the scale parameter. GE distribution with the shape parameter α and the scale parameter λ will be denoted by GE(α,λ). The location parameter also can be easily introduced (see Gupta and Kundu; 1999). When $\alpha = 1$, it represents an exponential distribution. Several properties were investigated by the authors and it was observed that it has more similarity to a gamma family than to a Weibull family in terms of hazard function. It has an increasing hazard function if $\alpha > 1$, which increases from 0 to λ. It has an decreasing hazard function if $\alpha < 1$, which decreases from ∞ to λ, exactly like gamma distribution. Several ordering relations, different moments, the distribution of the sum of independent and identically distributed GE variables and the distribution of the extreme values were also obtained in Gupta and Kundu (1999). One major advantage of a GE distribution is its simple structure of the distribution function or the survival function. Due to the simple structure of the distribution function and the survival function, GE can be used quite effectively in analyzing any lifetime data particularly in presence of censoring or if the data are grouped.

The main aim of this paper is to present different inference procedures for the parameters of a GE family. The rest of the paper is organized as follows. We obtain the maximum likelihood estimators (MLE's) of the parameters for complete sample in Section 2 and for censored sample in Section 3. Consistency and asymptotic normality results of the MLE's under different conditions are obtained in Section 4. Different testing of hypotheses and confidence intervals are considered in Section 5. Finally we analyze one data set and see how the different methods work for that particular data set in Section 6.

2. MAXIMUM LIKELIHOOD ESTIMATION FOR COMPLETE SAMPLES

In this section we discuss the maximum likelihood estimators of the unknown parameters of the GE model when the complete sample is available. If $x_1, \ldots, x_n$ is a random sample from GE(α, λ), then the log-likelihood function, $L(\alpha, \lambda)$, is

$$L(\alpha,\lambda) = nln(\alpha) + nln(\lambda) + (\alpha - 1)\sum_{i=1}^{n} ln(1 - e^{-\lambda x_i}) - \lambda \sum_{i=1}^{n} x_i. \tag{2.1}$$

The likelihood equations become

$$\frac{\partial L}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} ln(1 - e^{-\lambda x_i}) = 0, \tag{2.2}$$

$$\frac{\partial L}{\partial \lambda} = \frac{n}{\lambda} + (\alpha - 1)\sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{(1 - e^{-\lambda x_i})} - \sum_{i=1}^{n} x_i = 0. \tag{2.3}$$

From (2.2), we obtain the MLE of α as a function of λ, say $\hat{\alpha}(\lambda)$,

$$\hat{\alpha}(\lambda) = -\frac{n}{\sum_{i=1}^{n} ln(1 - e^{-\lambda x_i})}. \tag{2.4}$$

Therefore, if the scale parameter is known, the MLE of the shape parameter, $\hat{\alpha}$, can be obtained directly from (2.4). If both the parameters are unknown, first the estimate of the scale parameter can be obtained by maximizing directly

$$g(\lambda) = L(\hat{\alpha}(\lambda), \lambda) = C - nln(-\sum_{i=1}^{n} ln(1 - e^{-\lambda x_i})) + nln(\lambda) - \sum_{i=1}^{n} ln(1 - e^{-\lambda x_i}) - \lambda \sum_{i=1}^{n} x_i \tag{2.5}$$

with respect to λ. Here C is a constant independent of λ. Once $\hat{\lambda}$ is obtained, $\hat{\alpha}$ can be obtained from (2.4) as $\hat{\alpha}(\hat{\lambda})$. It is observed empirically that for all $x_1, \ldots, x_n > 0$, $L(\hat{\alpha}(\lambda), \lambda)$ is an unimodal function. Although, we could not prove the general result, the following result can be obtained.

Theorem 1: If $n = 1$,

$$g(\lambda) = L(\hat{\alpha}(\lambda), \lambda) = C - ln(-ln(1 - e^{-\lambda x})) + ln(\lambda) - ln(1 - e^{-\lambda x}) - \lambda x$$

is an unimodal function of λ.

Proof: See Appendix.

Therefore, most of the maximization routine will work well from a reasonable guess value of λ. To maximize (2.5), we can differentiate $g(\lambda)$ with respect to λ and equate it zero,

$$g'(\lambda) = -n\frac{\sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{(1 - e^{-\lambda x_i})}}{\sum_{i=1}^{n} ln(1 - e^{-\lambda x_i})} + \frac{n}{\lambda} - \sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{1 - e^{-\lambda x_i}} - \sum_{i=1}^{n} x_i = 0. \tag{2.6}$$

Note that (2.6) can be written as

$$h(\lambda) = \lambda \tag{2.7}$$

where

$$h(\lambda) = \left[\frac{\sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{(1 - e^{-\lambda x_i})}}{\sum_{i=1}^{n} ln(1 - e^{-\lambda x_i})} + \frac{1}{n}\sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{1 - e^{-\lambda x_i}} + \frac{1}{n}\sum_{i=1}^{n} x_i \right]^{-1}. \tag{2.8}$$

Therefore, if $\hat{\lambda}$ is the MLE of λ, then $\hat{\lambda}$ is a fixed point solution of (2.7). Very simple iterative procedure can be used to find a solution of the equation (2.7). For example

$$\lambda_{j+1} = h(\lambda_j) \tag{2.9}$$

where λ_j is the estimate of $\hat{\lambda}$ at the j^{th} iterate. Continue the iterative process unless some stopping criterion is satisfied. It is also possible to use Newton-Raphson type procedure as follows

$$\lambda_{j+1} = \lambda_j - \frac{g'(\lambda_j)}{g''(\lambda_j)}, \tag{2.10}$$

where $g'(.)$ is same as defined in (2.6) and $g''(.)$ is the derivative of $g'(.)$. It is also possible to use any standard one dimensional minimization/ maximization routine to compute the MLE of λ.

Now we describe how to obtain initial guesses of the unknown parameters. Since the coefficient of variation (C.V.) is independent of the scale parameter, the sample C.V. gives some idea about the shape parameter. The exact expression of the mean and the C.V. can be obtained from the moment generating function of a $GE(\alpha, \lambda)$, derived in Gupta and Kundu (1999). If X is a $GE(\alpha, \lambda)$ random variable, then the exact expression of mean and C.V. of X are as follows;

$$\mu = E(X) = \frac{1}{\lambda}(\psi(\alpha+1) - \psi(1)), \qquad \text{C.V.} = \frac{\sqrt{\psi'(1) - \psi'(1+\alpha)}}{\psi(\alpha+1) - \psi(1)}. \tag{2.11}$$

Here $\psi(.)$ denotes the digamma function and $\psi'(.)$ denotes the first derivative of $\psi(.)$. Tabulated values are available in the literature for digamma and polygamma functions (see Johnson, Kotz and Balakrishnan; 1995). Many statistical packages, for example MAPLE has the inbuilt digamma and polygamma functions. We provide a small table (Table 1) for mean and C.V. for different sample sizes and for different values of the shape parameter for immediate use. Extensive tables are available on request from the authors. From the sample C.V. and from Table 1, we can have some idea about the shape parameter. From the shape parameter and from the sample mean we can get some idea about the scale parameter. The details will be explained in section 6. Another way, to have an idea about the initial guess of the scale parameter λ, is to plot the concentrated log-likelihood function $g(\lambda)$ as a function of λ and get an idea about the global maximum.

3. MAXIMUM LIKELIHOOD ESTIMATION FOR INCOMPLETE SAMPLES

In this section we discuss the MLE's of the unknown parameters of the GE distribution, given that the data are either Type I or Type II censored or in a grouped form. When the lifetime distribution is Weibull, there is a considerable amount of literature available for the estimation of the unknown parameters. In case of gamma distribution, if the data are censored or in a grouped form the estimation becomes quite complicated computationally.

First consider the usual situation leading to Type I censored sample, where T_i represents the lifetime and L_i represents the fixed censoring time of the i^{th} individual in a random sample of n individuals. One observes only $x_i = \min\{T_i, L_i\}$ and whether the observation is a lifetime or the censored time. In presence of Type I censoring, the log-likelihood function becomes;

$$L_1(\alpha, \lambda) = r ln(\alpha) + r ln(\lambda) + (\alpha - 1)\sum_{i \in D} ln(1 - e^{-\lambda x_i}) - \lambda \sum_{i \in D} x_i + \sum_{i \in D^*} ln(1 - (1 - e^{-\lambda x_i})^{\alpha}). \tag{3.1}$$

Here D and D^* denote the sets consisting of the individuals whose lifetimes are not censored and censored respectively, r denotes the number of elements in D. The likelihood equations take the forms

$$\frac{\partial L_1}{\partial \alpha} = \frac{r}{\alpha} + \sum_{i \in D} ln(1 - e^{-\lambda x_i}) - \sum_{i \in D^*} \frac{(1 - e^{-\lambda x_i})^{\alpha} ln(1 - e^{-\lambda x_i})}{1 - (1 - e^{-\lambda x_i})^{\alpha}} = 0, \tag{3.2}$$

$$\frac{\partial L_1}{\partial \lambda} = \frac{r}{\lambda} + (\alpha - 1)\sum_{i\in D} \frac{x_i e^{-\lambda x_i}}{(1-e^{-\lambda x_i})} - \sum_{i \in D} x_i - \sum_{i\in D^*} \frac{\alpha(1-e^{-\lambda x_i})^{\alpha-1} x_i e^{-\lambda x_i}}{1-(1-e^{-\lambda x_i})^{\alpha}} = 0. \tag{3.3}$$

We need to solve (3.2) and (3.3) to obtain the MLE's of α and λ. Note that (3.2) and (3.3) can be written as

$$h(\alpha, \lambda) = (h_1(\alpha,\lambda), h_2(\alpha,\lambda)) = (\alpha, \lambda) \tag{3.4}$$

where

$$h_1(\alpha,\lambda) = r\left[g_1(\alpha,\lambda) - \sum_{i\in D} ln(1-e^{-\lambda x_i})\right]^{-1},$$

$$h_2(\alpha,\lambda) = r\left[g_2(\alpha,\lambda) + \sum_{i\in D} x_i - (\alpha-1)\sum_{i\in D}\frac{x_i e^{-\lambda x_i}}{1-e^{-\lambda x_i}}\right]^{-1}$$

and

$$g_1(\alpha,\lambda) = \sum_{i\in D^*} \frac{(1-e^{-\lambda x_i})^{\alpha} ln(1-e^{-\lambda x_i})}{[1-(1-e^{-\lambda x_i})^{\alpha}]}, \tag{3.5}$$

$$g_2(\alpha,\lambda) = \sum_{i\in D^*} \frac{(1-e^{-\lambda x_i})^{\alpha-1} \alpha x_i e^{-\lambda x_i}}{[1-(1-e^{-\lambda x_i})^{\alpha}]}. \tag{3.6}$$

Very simple iterative procedure, like (2.9) can be used to obtain the MLE's of α and λ. It is also possible to use the two dimensional Newton-Raphson type procedure to obtain the solutions of (3.2) and (3.3).

Now consider the situation when the data are Type II censored. The first r ordered observations, $x_{(1)}, \ldots, x_{(r)}$, are only available out of the total n observations. Here r is some pre-chosen fixed number. The log-likelihood function becomes

$$L_2(\alpha,\lambda) = \sum_{i=n-r+1}^{n} ln(i) + rln(\alpha) + rln(\lambda) + (\alpha-1)\sum_{i=1}^{r} ln(1-e^{-\lambda x_{(i)}}) - \lambda \sum_{i=1}^{r} x_{(i)}$$
$$+ (n-r) ln(1-(1-e^{-\lambda x_{(r)}})^{\alpha}). \tag{3.7}$$

Therefore the likelihood equations become

$$\frac{\partial L_2}{\partial \alpha} = \frac{r}{\alpha} + \sum_{i=1}^{r} ln(1-e^{-\lambda x_{(i)}}) - (n-r)\frac{ln(1-e^{-\lambda x_{(r)}}))(1-e^{-\lambda x_{(r)}})^{\alpha}}{1-(1-e^{-\lambda x_{(r)}})^{\alpha}} = 0, \tag{3.8}$$

$$\frac{\partial L_2}{\partial \lambda} = \frac{r}{\lambda} + (\alpha-1)\sum_{i=1}^{r}\frac{x_{(i)}e^{-\lambda x_{(i)}}}{1-e^{-\lambda x_{(i)}}} - \sum_{i=1}^{r} x_{(i)} - (n-r)\frac{\alpha x_{(r)} e^{-\lambda x_{(r)}}(1-e^{-\lambda x_{(r)}})^{\alpha-1}}{1-(1-e^{-\lambda x_{(r)}})^{\alpha}} = 0. \tag{3.9}$$

In this case also (3.8) and (3.9) can be solved similarly as (3.4), *i.e.*,

$$h(\alpha, \lambda) = (h_1(\alpha,\lambda), h_2(\alpha,\lambda)) = (\alpha, \lambda) \tag{3.10}$$

where

$$h_1(\alpha,\lambda) = r\left[\frac{(n-r)ln(1-e^{-\lambda x_{(r)}}))(1-e^{-\lambda x_{(r)}})^{\alpha}}{1-(1-e^{-\lambda x_{(r)}})^{\alpha}} - \sum_{i=1}^{r} ln(1-e^{-\lambda x_{(i)}})\right]^{-1},$$

and

$$h_2(\alpha,\lambda) = r\left[\frac{\alpha(n-r)x_{(r)}e^{-\lambda x_{(r)}}(1-e^{-\lambda x_{(r)}})^{\alpha-1}}{1-(1-e^{-\lambda x_{(r)}})^{\alpha}} - (\alpha-1)\sum_{i=1}^{r}\frac{x_{(i)}e^{-\lambda x_{(i)}}}{1-e^{-\lambda x_{(i)}}} + \sum_{i=1}^{r} x_{(i)}\right]^{-1}.$$

Now we discuss the maximum likelihood estimation of the unknown parameters of the GE distribution when the data are available in a grouped form. We don't have the full information about the sample. We only have the information that how many observations are within a group or in a interval. The available data will be of the following form;

Table 2

Class	Observed Frequency
$a_1 - a_2$	n_1
$\vdots$	$\vdots$
$a_k - a_{k+1}$	n_k

where the end interval(s) may be of infinite range(s). The log-likelihood function of the observed sample is

$$L_3(\hat{\alpha}_1(\lambda), \lambda) = \sum_{i=1}^{k} n_i ln \left[(1 - e^{-\lambda a_{i+1}})^{\alpha} - (1 - e^{-\lambda a_i})^{\alpha} \right]. \tag{3.11}$$

The likelihood equations become

$$\frac{\partial L_3}{\partial \alpha} = \sum_{i=1}^{k} n_i \frac{(1 - e^{-\lambda a_{i+1}})^{\alpha} ln(1 - e^{-\lambda a_{i+1}}) - (1 - e^{-\lambda a_i})^{\alpha} ln(1 - e^{-\lambda a_i})}{(1 - e^{-\lambda a_{i+1}})^{\alpha} - (1 - e^{-\lambda a_i})^{\alpha}} = 0, \tag{3.12}$$

$$\frac{\partial L_3}{\partial \lambda} = \sum_{i=1}^{k} n_i \frac{(1 - e^{-\lambda a_{i+1}})^{\alpha-1} e^{-\lambda a_{i+1}} a_{i+1} - (1 - e^{-\lambda a_i})^{\alpha-1} e^{-\lambda a_i} a_i}{(1 - e^{-\lambda a_{i+1}})^{\alpha} - (1 - e^{-\lambda a_i})^{\alpha}} = 0. \tag{3.13}$$

We need to solve (3.12) and (3.13) simultaneously to obtain the MLE's of α and λ. Alternatively, one may use some two dimensional maximization algorithm to obtain the MLE's of α and λ directly, in this case.

4. FISHER INFORMATION MATRIX

In this section we provide the consistency and the asymptotic normality results of the MLE's of the GE parameters if the data are complete. We also provide the elements of the Fisher Information matrix for Type I censored data and if the data are in a grouped form. First let's consider the case when the data are not censored. Consider the second derivatives of $L(\alpha, \lambda)$ given in (2.1).

$$\frac{\partial^2 L}{\partial \alpha^2} = -\frac{n}{\alpha^2}, \qquad \frac{\partial^2 L}{\partial \alpha \partial \lambda} = \sum_{i=1}^{n} \frac{x_i e^{-x_i \lambda}}{(1 - e^{-\lambda x_i})},$$

$$\frac{\partial^2 L}{\partial \lambda^2} = -\frac{n}{\lambda^2} - (\alpha - 1) \sum_{i=1}^{n} \frac{x_i^2 e^{-\lambda x_i}}{(1 - e^{-\lambda x_i})^2}.$$

Therefore, for $\alpha > 2$, if X is a GE(α, λ) random variable, then

$$\begin{aligned} E\left(\frac{\partial^2 L}{\partial \alpha^2}\right) &= -\frac{n}{\alpha^2}, \\ E\left(\frac{\partial^2 L}{\partial \alpha \partial \lambda}\right) &= nE\left(\frac{X e^{-X\lambda}}{(1 - e^{-\lambda X})}\right) = nE\left(X\left[\frac{1}{1 - e^{-\lambda X}} - 1\right]\right) \end{aligned}$$

$$
\begin{aligned}
& = && n\left[\frac{\alpha}{\lambda(\alpha-1)}(\psi(\alpha)-\psi(1))-\frac{1}{\lambda}(\psi(\alpha+1)-\psi(1))\right], \\
E\left(\frac{\partial^2 L}{\partial \lambda^2}\right) & = && -n\left[\frac{1}{\lambda^2}+(\alpha-1)E\left(\frac{X^2 e^{-X\lambda}}{(1-e^{-\lambda X})^2}\right)\right] \\
& = && -n\left[\frac{1}{\lambda^2}+(\alpha-1)E\left(\frac{X^2}{(1-e^{-\lambda X})^2}-\frac{X^2}{(1-e^{-\lambda X})}\right)\right] \\
& = && -n[\frac{1}{\lambda^2}+\frac{\alpha(\alpha-1)}{(\alpha-2)\lambda^2}(\psi'(1)-\psi'(\alpha-1)+(\psi(\alpha-1)-\psi(1))^2) \\
& - && \frac{\alpha}{\lambda^2}(\psi'(1)-\psi'(\alpha)+(\psi(\alpha)-\psi(1))^2)].
\end{aligned}
$$

For $0 < \alpha \leq 2$,

$$
E\left(\frac{\partial^2 L}{\partial \alpha^2}\right) = -\frac{n}{\alpha^2}, \qquad E\left(\frac{\partial^2 L}{\partial \alpha \partial \lambda}\right) = n\frac{\alpha}{\lambda}\int_0^\infty x e^{-2x}(1-e^{-x})^{\alpha-2}dx < \infty,
$$

$$
E\left(\frac{\partial^2 L}{\partial \lambda^2}\right) = -\frac{n}{\lambda^2} - \frac{n\alpha(\alpha-1)}{\lambda^2}\int_0^\infty x^2 e^{-2x}(1-e^{-x})^{\alpha-3}dx < \infty.
$$

For $\alpha > 0$, the GE family satisfies all the regularity conditions (see Bain and Engelhardt; 1991) and therefore, we have the following result:

Theorem 2: For $\alpha > 0$, the maximum likelihood estimators, $(\hat{\alpha}, \hat{\lambda})$, of (α, λ) are consistent and $\sqrt{n}(\hat{\alpha}-\alpha, \hat{\lambda}-\lambda)$ is asymptotically normal with mean vector zero and dispersion matrix $\mathbf{I}^{-1}$, where

$$
\mathbf{I} = -\frac{1}{n}\begin{bmatrix} E\left(\frac{\partial^2 L}{\partial \alpha^2}\right) & E\left(\frac{\partial^2 L}{\partial \alpha \partial \lambda}\right) \\ E\left(\frac{\partial^2 L}{\partial \alpha \partial \lambda}\right) & E\left(\frac{\partial^2 L}{\partial \lambda^2}\right) \end{bmatrix}.
$$

We provide the elements of the Fisher Information matrix, when the data are type I censored. Note that it is possible to obtain the Fisher Information matrix in terms of

$$
E(X|X \leq L) = \psi(\alpha, L) \qquad \text{(say) and}
$$

$$
E(X^2|X \leq L) = \tilde{\psi}(\alpha, L) \qquad \text{(say)},
$$

where X is a $GE(\alpha, 1)$ random variable. The explicit expressions of $\psi(\alpha, L)$ and $\tilde{\psi}(\alpha, L)$ are as follows;

$$
\psi(\alpha, L) = \frac{\alpha}{(1-e^{-L})^\alpha}\sum_{j=0}^{\infty}(-1)^j c(\alpha-1, j)\left[\frac{1}{(j+1)^2}(1-e^{-(j+1)L}) - \frac{Le^{-(j+1)L}}{j+1}\right]
$$

and

$$
\tilde{\psi}(\alpha, L) = \frac{\alpha}{(1-e^{-L})^\alpha}\sum_{j=0}^{\infty}(-1)^j c(\alpha-1, j)\left[\frac{2}{(j+1)^2}(1-e^{-(j+1)L}) - \frac{2Le^{-(j+1)L}}{(j+1)^2} - \frac{L^2 e^{-(j+1)L}}{j+1}\right]
$$

where

$$
c(\alpha, i) = \frac{\alpha(\alpha-1)\ldots(\alpha-i+1)}{i!}.
$$

Then,

$$
E\left(\frac{\partial^2 L_1}{\partial \alpha^2}\right) \quad = \quad -\frac{r}{\alpha^2} + g_1'(\alpha, \lambda),
$$

$$
\begin{aligned}
E\left(\frac{\partial^2 L_1}{\partial\alpha\partial\lambda}\right) &= \frac{1}{\lambda}\sum_{i\in D}\left[\frac{\alpha}{(\alpha-1)(1-e^{-L_i\lambda})}\psi(\alpha-1,L_i\lambda)-\psi(\alpha,L_i\lambda)\right]+g_1'(\alpha,\lambda),\\
E\left(\frac{\partial^2 L}{\partial\lambda^2}\right) &= -\frac{r}{\lambda^2}+g_2'(\alpha,\lambda)-\frac{(\alpha-1)}{\lambda^2}\times\\
&\sum_{i\in D}\left[\frac{\alpha}{(\alpha-2)(1-e^{-L_i\lambda})^2}\tilde{\psi}(\alpha-2,L_i\lambda)-\frac{\alpha}{(\alpha-1)(1-e^{-L_i\lambda})}\tilde{\psi}(\alpha-1,L_i\lambda)\right].
\end{aligned}
$$

The elements of the Fisher Information matrix when the data are type II censored can be found in Zheng (2002).

Now we provide the elements of the Fisher Information matrix for the grouped data in terms of the cell probabilities. Note that (3.11) can be rewritten as

$$
L_3(\alpha,\lambda)=\sum_{i=1}^{k} n_i ln\left[(1-e^{-\lambda a_{i+1}})^{\alpha}-(1-e^{-\lambda a_i})^{\alpha}\right]=\sum_{i=1}^{k} n_i ln(P_i(\alpha,\lambda)) \quad \text{(say)}.
$$

Then

$$
\begin{aligned}
E\left(\frac{\partial^2 L_1}{\partial\alpha^2}\right) &= n\sum_{i=1}^{k}P_i(\alpha,\lambda)\frac{\partial^2}{\partial\alpha^2}ln(P_i(\alpha,\lambda)),\\
E\left(\frac{\partial^2 L_1}{\partial\alpha\partial\lambda}\right) &= n\sum_{i=1}^{k}P_i(\alpha,\lambda)\frac{\partial^2}{\partial\alpha\partial\lambda}ln(P_i(\alpha,\lambda)),\\
E\left(\frac{\partial^2 L}{\partial\lambda^2}\right) &= n\sum_{i=1}^{k}P_i(\alpha,\lambda)\frac{\partial^2}{\partial\lambda^2}ln(P_i(\alpha,\lambda)).
\end{aligned}
$$

5. CONFIDENCE INTERVALS AND TESTING OF HYPOTHESES

It is well known that (Rao; 1973) that the testing of hypotheses and confidence interval problems are equivalent to each other. In this section, we mainly address these two problems for both the scale and the shape parameters. Unfortunately, unlike gamma distribution, generalized exponential distribution doesn't belong to an exponential family. On the other hand it does not posses some of the nice extreme value distribution properties like Weibull distribution. For a gamma distribution, it is always convenient to work with a complete sufficient statistic and for a Weibull distribution, it can be transferred to a location-scale family (extreme value distribution) for which an ancillary statistic can be easily obtained. On the testing point of view, GE distribution is not similar to either gamma or Weibull distribution. We propose to use the likelihood ratio test (LRT) or use the asymptotic normality results as it was obtained in the previous section, when both the parameters are unknown. If the scale parameter is known, there exists a uniformly most powerful (UMP) test or uniformly most powerful unbiased (UMPU) test, depending on the alternatives. We consider the different cases separately. First let's assume that both the parameters are unknown, then we consider the case when one parameter is known.

5.1 Both the Parameters are Unknown

First let's consider the following testing problem on the shape parameter α, when the scale parameter is also unknown.

$$
H_0: \alpha=\alpha_0, \qquad H_1: \alpha\neq\alpha_0 \tag{5.1}
$$

Note that, when $\alpha_0 = 1$, it tests exponentiality. It is an important problem in practice. It is not possible to obtain UMP test or UMPU test in this case. We propose to use the LRT as follows:

$$\Lambda_1 = -2\left[L(\alpha_0, \tilde{\lambda}) - L(\hat{\alpha}, \hat{\lambda})\right] \tag{5.2}$$

where $\tilde{\lambda}$ is the restricted MLE of λ, when $\alpha = \alpha_0$. It can be obtained by maximizing $L(\alpha_0, \lambda)$ in (2.1) with respect to λ. Here $\hat{\alpha}$ and $\hat{\lambda}$ are the MLE's of α and λ respectively. For large sample, the distribution of Λ_1 under H_0 is approximately chi-square with one degree of freedom, χ^2_1. Since λ is a scale parameter, it is known (see Bain and Engelhardt; 1991) that the exact distribution of Λ_1 under H_0 does not depend on λ. An approximate $(1-\gamma)100\%$ confidence interval of α consists of the set of values α_0 such that $\Lambda_1(\alpha_0) \le \chi^2_{1,1-\gamma}$, where $\chi^2_{1,1-\gamma}$ is the $(1-\gamma)100^{th}$ percentile point of the χ^2_1. It can be obtained by trial and error of the different values of α_0, which satisfy $\Lambda_1(\alpha_0) \le \chi^2_{1,1-\gamma}$. It is possible to use the asymptotic normality results of the MLE's (Theorem 2) to construct approximate confidence interval or to test the hypothesis.

Similarly, we can consider the testing of hypothesis on the scale parameter λ. Consider the following testing problem on the scale parameter λ;

$$H_0: \lambda = \lambda_0, \qquad H_1: \lambda \ne \lambda_0. \tag{5.3}$$

The LR statistic is;

$$\Lambda_2 = -2\left[L(\tilde{\alpha}, \lambda_0) - L(\hat{\alpha}, \hat{\lambda})\right], \tag{5.4}$$

here $\hat{\alpha}$ and $\hat{\lambda}$ are the MLE's of α and λ respectively, $\tilde{\alpha}$ is the restricted MLE of α when $\lambda = \lambda_0$ and $\tilde{\alpha}$ can be obtained from (2.4) by putting $\lambda = \lambda_0$. We can approximate Λ_2 under H_0 as χ^2_1 for large n. Since λ is the scale parameter, the exact distribution of Λ_2 is independent of λ_0. It is possible to construct the confidence interval of λ similarly as α or we can use the asymptotic normality results to test the hypothesis (5.3).

We can consider the simultaneous testing of α and λ as follows. The hypothesis is,

$$H_0: \alpha = \alpha_0, \lambda = \lambda_0 \qquad H_1: \text{at least one is not true.} \tag{5.5}$$

The LR statistic is:

$$\Lambda_3 = -2\left[L(\alpha_0, \lambda_0) - L(\hat{\alpha}, \hat{\lambda})\right]. \tag{5.6}$$

In this case also $\hat{\alpha}$ and $\hat{\lambda}$ are the MLE's of α and λ respectively and the exact distribution of Λ_3 under H_0 is independent of λ_0. For large n, the distribution of Λ_3 under H_0 is approximately chi-square with two degrees of freedom, χ^2_2 and it is also possible to use the asymptotic normality results to test the hypothesis (5.5).

An alternative to the exact numerical calculation of tests or confidence intervals is the estimation of results by graphical means. We may not need any sophisticated programming tool in this case. The only requirement is to have a good graphical tool. Originally this procedure was suggested by Lawless (1982) for a gamma distribution. It can be used more conveniently even for GE parameters also. We explain the procedure below and it will be illustrated using data set in Section 6.

First obtain the MLE's of α and λ and then plot contours of

$$R(\alpha, \lambda) = -2\left[L(\alpha, \lambda) - L(\hat{\alpha}, \hat{\lambda})\right].$$

These contours plotting now a days can be done in most of the Statistical packages. These contours plots can be used to obtain the approximate testing or to form confidence intervals of the unknown parameters. Suppose, we want to obtain an approximate $(1-\gamma)100$ % confidence bounds for α or λ or to test a hypothesis at a significance level γ , the plot should include the contour

$$R(\alpha, \lambda) - \chi^2_{1,1-\gamma} = 0. \tag{5.7}$$

All points (α, λ) inside this contour will have $R(\alpha, \lambda) \le \chi^2_{1,1-\gamma}$. Suppose we want to test the following hypothesis on α, namely $H_0 : \alpha = \alpha_0$ vs. $H_1 : \alpha \neq \alpha_0$ at the 5 % significance level. The appropriate LR statistic is given in (5.2). We know Λ_1 is the maximum value of $R(\alpha, \lambda)$ on the line $\alpha = \alpha_0$. If one draws a line $\alpha = \alpha_0$ on the contour, this value can be easily obtained. Therefore, it can be easily checked whether $R(\alpha_0, \tilde{\lambda})$ is inside or outside a give contour and hence whether H_0 is accepted or rejected at a given significance level. Similarly an approximate $(1-\gamma)$ 100 % confidence interval of α can be obtained graphically by locating all values of α_0 such that $R(\alpha_0, \tilde{\lambda}) \le \chi^2_{1,1-\gamma}$.

It is possible to test the hypothesis (5.5) graphically also. In this case the plot should include the contour

$$R(\alpha, \lambda) - \chi^2_{2,1-\gamma} = 0.$$

Therefore, if we want 5 % significant level to test (5.5) then we should reject the hypothesis if (α_0, λ_0) falls outside the contour

$$R(\alpha, \lambda) - 5.991 = 0.$$

5.2 Scale Parameter is Known

In this subsection we consider the testing of hypothesis problem on the shape parameter, when the scale parameter is known. Suppose we want to test one of the following hypotheses:

$$H_0 : \alpha = \alpha_0, \qquad H_1 : \alpha \neq \alpha_0; \tag{5.8}$$

$$H_0 : \alpha = \alpha_0, \qquad H_1 : \alpha < \alpha_0; \tag{5.9}$$

$$H_0 : \alpha = \alpha_0, \qquad H_1 : \alpha > \alpha_0. \tag{5.10}$$

Consider the hypothesis (5.10), the others will follow similarly. Since λ is known, without loss of generality we can assume $\lambda = 1$. The likelihood ratio test for testing $H_0 : \alpha = \alpha_0, H_1 : \alpha > \alpha_0$, takes the following form: Reject H_0 if

$$\prod_{i=1}^{n}(1 - e^{-X_i}) > c, \tag{5.11}$$

accept otherwise. Now to get the constant c, we need to obtain the exact distribution of $\prod_{i=1}^{n}(1 - e^{-X_i})$ under the null hypothesis that $X_1, \ldots, X_n$ are *i.i.d.* from GE$(\alpha, 1)$. Note that $(1 - e^{-X_i}), i = 1, \ldots, n$ are *i.i.d.* beta random variables (see Gupta and Kundu; 1999). The distribution of the product of beta random variables are well studied in the literature (see Johnson, Kotz and Balakrishnan; 1995). Therefore the constant c and the power function can be obtained explicitly. Since the family had the monotone likelihood ratio property (see Gupta and Kundu; 1999), therefore it is the UMP test also. Similarly we can obtain the uniformly most powerful (UMP) test for (5.9) and uniformly most powerful unbiased (UMPU) test for (5.8). Note that the mean and the variance of $ln(1 - e^{-X_i})$ exist for any $\alpha > 0$ and they are as follows;

$$E(ln(1 - e^{-X_i})) = -\frac{1}{\alpha}, \qquad Var(ln(1 - e^{-X_i})) = \frac{1}{\alpha^2}.$$

Therefore, (5.11) is equivalent to

$$U = \left(\frac{\alpha_0 \sum_{i=1}^{n} ln(1 - e^{-X_i})}{\sqrt{n}}\right) + \sqrt{n} > c. \tag{5.12}$$

The distribution of U under H_0 is same as

$$U = -\frac{Y}{\sqrt{n}} + \sqrt{n}$$

where $2Y$ has a χ^2 distribution with $2n$ degrees of freedom. Therefore, if the scale parameter is known, $2Y = -2\alpha \sum_{i=1}^{n} ln(1-e^{-X_i})$ can be used as a pivotal quantity to obtain the confidence interval of α. When λ is known the exact 95 % confidence interval for α is

$$\left(\frac{\chi^2_{2n,.025}}{-2\sum_{i=1}^{n} ln(1-e^{-\lambda x_i})}, \frac{\chi^2_{2n,.975}}{-2\sum_{i=1}^{n} ln(1-e^{-\lambda x_i})} \right) \tag{5.13}$$

where $\chi^2_{2n,.975}$ and $\chi^2_{2n,.025}$ are the 97.5^{th} and 2.5^{th} percentiles of the χ^2 distribution with $2n$ degrees of freedom. For large n, the distribution of U can be approximated by the standard normal.

It may be mentioned that when α is known, then the UMP or UMPU test does not exist. We can use the LRT in this case to test the hypothesis on λ as mentioned before.

6. DATA ANALYSIS

In this section we analyze one data set originally discussed in Lawless (1982, page 228). The data given here arose in tests on endurance of deep groove ball bearings. The data are the number of million revolutions before failure for each of the 23 ball bearings in the life test and they are

17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.80, 51.84, 51.96, 54.12, 55.56, 67.80, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, 173.40.

The data has the sample mean = 72.2383, the sample variance = 1404.72 and the sample C.V. = .5188. Since the C.V. is independent of the scale parameter, from Table 1 and using sample C.V. it seems α should be between 5.00 to 5.50. We take the initial guess value of α to be 5.0. When $\alpha = 5.0$ and $\lambda = 1$, from Table 1, the mean of a $GE(5,1)$ is 2.2833. Since the mean of a $GE(\alpha,\lambda)$ is same as the mean of $GE(\alpha, 1)$ divided by λ, we obtain the guess value of $\lambda = \frac{2.2833}{72.2383} = .0316$. Figure 1 provides the plot of the function (2.5). It also indicates that the maximum should occur between .03 to .04. Using the iterative procedure (2.9) with the guess value of $\lambda = .0316$, we obtain the MLE, $\hat{\lambda}$, of $\lambda = .0322965$ in two iterations. With the same initial value of λ and using the Newton-Raphson type iterative procedure (2.10), we obtain $\hat{\lambda} = .0322960$ with the same number of iterations. Using the maximization routine of Press et al. (1993), and giving the information that the maximum should be between .03 and .04, we obtain the solution as .03230. Therefore, it seems it matches up to four decimal places in all three cases. We take $\hat{\lambda} = .0323$, and from (2.4), we obtain $\hat{\alpha} = 5.2775$. The corresponding log likelihood becomes -112.9778.

We obtain the approximate 95 % confidence bound on α using the trial and error method given in section 5.1 as follows. We start with $\alpha_0 = \hat{\alpha} = 5.2775$ and compute Λ_1 as given in (5.2) and then we take $\alpha_0 = \alpha_0 + h$, where $h = .0001$. We repeat the process unless $|\Lambda_1(\alpha_0) - \chi^2_{1,.95}| < h$ and that value of α_0 becomes the upper bound of the confidence interval of α. Similarly taking $\alpha_0 = \alpha_0 - h$, we can obtain the lower bound of the confidence interval of α. In this case we obtain the 95 % confidence interval of α as (2.4351,11.1550). Using the same procedure and the same step size we obtain the 95 % confidence interval of λ as (.0207, .0458). It seems both the confidence intervals are quite large.

Now we would like to obtain the confidence intervals using the asymptotic normality results. From theorem 2, we have;

$$\sqrt{n}\left(\hat{\alpha}-\alpha, \hat{\lambda}-\lambda\right) \approx N_2 \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 29.1594 & 4.867 \times 10^{-3} \\ 4.867 \times 10^{-3} & 1.812 \times 10^{-5} \end{bmatrix} \right\}, \tag{6.1}$$

where N_2 denotes the bivariate normal distribution and $\approx$ means approximate distribution. Here we obtain the asymptotic dispersion matrix of $(\hat{\alpha}-\alpha, \hat{\lambda}-\lambda)$ from theorem 2 and by replacing the Fisher Information matrix by its estimate. Therefore, an approximate 95 % confidence interval of α is $\hat{\alpha} \pm 1.96 \times \sqrt{29.1594/23} = (3.0706, 7.4844)$ and similarly an approximate 95 % confidence interval of λ is (.0306, .0340). The asymptotic confidence intervals are considerably smaller than the other ones. It seems more investigation is needed in terms of the coverage probabilities in both the cases.

It is possible to obtain the simultaneous confidence region of α and λ using (6.1). An approximate 95 % confidence ellipse of α and λ is

$$n(\hat{\alpha}-\alpha, \hat{\lambda}-\lambda) \begin{bmatrix} 29.1594 & 4.867 \times 10^{-3} \\ 4.867 \times 10^{-3} & 1.812 \times 10^{-5} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\alpha}-\alpha \\ \hat{\lambda}-\lambda \end{bmatrix} - \chi^2_{2,.95} = 0$$

or

$$n(\hat{\alpha}-\alpha, \hat{\lambda}-\lambda) \begin{bmatrix} .0359 & -9.6440 \\ -9.6440 & 57774.535 \end{bmatrix} \begin{bmatrix} \hat{\alpha}-\alpha \\ \hat{\lambda}-\lambda \end{bmatrix} - 5.991 = 0. \tag{6.2}$$

The ellipse (6.2) is provided in Figure 3. The ellipse (6.2) can be used to test the hypothesis (5.5).

Now we consider the testing of hypothesis only on α. We consider the following testing problem on α,

$$H_0 : \alpha = 1, \qquad H_1 : \alpha \neq 1. \tag{6.3}$$

We obtain $\tilde{\lambda} = .0138$, the corresponding log-likelihood becomes -118.567 and $\Lambda_1 = 11.1820$. Since the p value $= .0016 < .05$ we reject the null hypothesis. We can use the graphical technique mentioned in subsection 5.1 for testing (6.3). From the Figure 2, it is immediate that the point $\mathrm{R}(\alpha_0, \tilde{\lambda}) =$ R(1.0, .0138) is not inside the contour (5.7), therefore, we reject H_0 at 5 % significant level. Hence, it seems the exponential distribution may not be a good fit for this data.

We consider the same testing problem (6.3) on α assuming λ is known. Let $\lambda = \hat{\lambda} = .0323$. We obtain, the value of U defined in (5.12), by replacing $x_i = \lambda x_i$ and $\alpha_0 = 1$, as 3.8883. The corresponding p value ≈ 0 , and therefore we reject the null hypothesis H_0.

For illustration purpose we consider the following testing problem of α assuming λ is known;

$$H_0 : \alpha = 4, \qquad H_1 : \alpha \neq 4. \tag{6.4}$$

The idea to consider the testing problem (6.4) is as follows. If we know that the data are coming from a parallel system with m components, where each component has the exponential distribution with the same scale parameter, λ, then it is known (Gupta and Kundu; 1999) that the lifetime distribution of the parallel system is $GE(m, \lambda)$. Therefore, testing (6.4) means to test whether the data came from a parallel system with four *i.i.d.* exponential components with known scale parameter. For testing (6.4), we obtain the value of U as 1.1656 and the corresponding p value is 0.228. Therefore, based on the data we don't reject the null hypothesis H_0 with 5 % level of significance.

Finally, we obtain the exact confidence interval of α when $\lambda = .0323$. Using (5.13), the exact 95 % confidence interval of α is (3.3498,7.6526). Note that although the exact confidence interval of α is not symmetric around $\hat{\alpha}$, but it is quite close to the asymptotic confidence interval of α. Therefore, it may not be unreasonable to use theorem 2 for construction of the confidence intervals or for testing purposes even for moderate sample sizes.

ACKNOWLEDGMENT

The authors would like to thank one of the Associate editors, one anonymous referee and the Editor Professor M. Ahsanullah for some very constructive suggestions. Part of the work is supported by the Natural Sciences and Engineering Research Council, Canada.

REFERENCES

Bain, L.J. and Engelhardt, M. (1991). *Statistical Analysis of Reliability and Life-Testing Models.* 2nd. Edition. Marcel and Decker. New York.

Gupta, R.D. and Kundu, D. (1999). Generalized exponential distribution. Australian and New Zealand Journal of Statistics. 41, 2, 173-188.

Gupta, R.D. and Kundu, D. (2001a). Exponentiated exponential distribution; an alternative to gamma or Weibull distribution. Biometrical Journal. 43, 1, 117-130.

Gupta, R.D. and Kundu, D. (2001b). Generalized exponential distribution; Different method of estimations. Journal of Statistical Computations and Simulations. 69, 4, 315-337.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distribution.* Vol. 2. 2nd Edition. Wiley. New York.

Lawless, (1982). *Statistical Models and Methods for Lifetime Data.* New York. Wiley.

Press et al. (1993). *Numerical Recipes in FORTRAN.* Cambridge University Press. Cambridge.

Raqab, M.Z. and Ahsanullah, M. (2001). Estimation of the location and scale parameters of generalized exponential distribution based on order statistics. Journal of Statistical Computations and Simulations. 69, 2, 109-124.

Rao, C.R. (1973), *Linear Statistical Inference and its Applications*, 2nd Edition, Wiley, New York.

Zheng, G. (2002). On the Fisher Information matrix in type-II censored data from the exponentiated exponential family. Biometrical Journal. (To appear).

APPENDIX

Proof of Theorem 1:

We want to prove that

$$g(\lambda) = C - ln(-ln(1 - e^{-\lambda x})) + ln(\lambda) - ln(1 - e^{-\lambda x}) - \lambda x$$

is a unimodal function. Note that it is equivalent to prove that

$$g(\lambda) = -ln(-ln(1 - e^{-\lambda}) + ln(\lambda) - ln(1 - e^{-\lambda}) - \lambda$$

is a unimodal function of λ. Consider the second derivative of $g(\lambda)$,

$$g''(\lambda) = [\frac{e^{-2\lambda}}{(1 - e^{-\lambda})^2[ln(1 - e^{-\lambda})]^2} + \frac{e^{-\lambda}}{(1 - e^{-\lambda})^2[ln(1 - e^{-\lambda})]}]$$

$$+\left[\frac{e^{-\lambda}}{(1-e^{-\lambda})^2}-\frac{1}{\lambda^2}\right]=g_1(\lambda)+g_2(\lambda) \qquad \text{(say)}$$

We show that $g_1(\lambda)\leq 0$ and $g_2(\lambda)\leq 0$. It would imply that $g(\lambda)$ is a concave function. Now the result follows from the fact that $g(0)\to-\infty$ and $g(\infty)\to-\infty$, therefore $g(.)$ has to be unimodal. Therefore, the proof will be complete if we can show that $g_1(\lambda)\leq 0$ and $g_2(\lambda)\leq 0$.

Now to prove $g_1(\lambda)\leq 0$, it is enough to prove that for $\lambda\geq 0$

$$\frac{e^{-\lambda}}{ln(1-e^{-\lambda})}+1\ \geq 0$$

$$<=> u(x)=x+ln(1-x)\leq 0; \qquad 0\leq x\leq 1$$

Since $u(x)$ is a decreasing function and $u(0)=0$, implies $u(x)\leq 0$. Now to prove $g_2(\lambda)\leq 0$, it is enough to prove that for all $\lambda\geq 0$

$$\frac{e^{-\frac{\lambda}{2}}}{(1-e^{-\lambda})}\leq\frac{1}{\lambda}$$

$$<=> 1-e^{-\lambda}-\lambda e^{-\frac{\lambda}{2}}\geq 0$$

$$<=> u(\lambda)=e^{\frac{\lambda}{2}}-e^{-\frac{\lambda}{2}}-\lambda\geq 0$$

Since

$$u'(\lambda)=\frac{1}{2}e^{-\frac{\lambda}{2}}(1-e^{-\frac{\lambda}{2}})^2\geq 0$$

and $u(0)=0$, therefore $u(\lambda)\geq 0$. So $g_2(\lambda)\leq 0$.

Some Comments: Note that in the general case

$$g''(\lambda)=n\frac{\sum_{i=1}^n\frac{{x_i}^2e^{-\lambda x_i}}{(1-e^{-\lambda x_i})^2}}{\sum_{i=1}^n ln(1-e^{-\lambda x_i})}+n[\frac{\sum_{i=1}^n\frac{x_ie^{-\lambda x_i}}{(1-e^{-\lambda x_i})}}{\sum_{i=1}^n ln(1-e^{-\lambda x_i})}]^2-\frac{n}{\lambda}^2+\sum_{i=1}^n\frac{{x_i}^2e^{-\lambda x_i}}{(1-e^{-\lambda x_i})^2}$$

It is not very easy to prove that $g''(\lambda)<0$ for all $\lambda>0$ in the general case.

Table 1

Table for population mean and C.V. of a $GE(\alpha, 1)$ random variable for different values of α

α	mean	C.V.
0.5	0.6135	1.3731
1.0	1.0000	1.0000
1.5	1.2804	0.8392
2.0	1.5000	0.7454
2.5	1.6804	0.6823
3.0	1.8333	0.6364
3.5	1.9661	0.6010
4.0	2.0833	0.5727
4.5	2.1883	0.5494
5.0	2.2833	0.5298
5.5	2.3701	0.5130
6.0	2.4500	0.4985
6.5	2.5240	0.4856
7.0	2.5929	0.4742
7.5	2.6573	0.4640
8.0	2.7179	0.4547
8.5	2.7750	0.4463
9.0	2.8290	0.4386
9.5	2.8802	0.4316
10.0	2.9290	0.4250