I164T001070 . 164
T001070d.164

J. Statist. Comput. Simul., 2000, Vol. 00, pp. 1–22
Reprints available directly from the publisher
Photocopying permitted by license only



# GENERALIZED EXPONENTIAL DISTRIBUTION: DIFFERENT METHOD OF ESTIMATIONS

RAMESHWAR D. GUPTA[a,*] and DEBASIS KUNDU[b,†]

[a]*Department of Applied Statistics and Computer Science, The University of New Brunswick, Saint John, Canada, E2L 4L5;*
[b]*Department of Mathematics, Indian Institute of Technology, Kanpur, India*

*(Received 1 October 1999; In final form 14 November 2000)*

Recently a new distribution, named as generalized exponential distribution has been introduced and studied quite extensively by the authors. Generalized exponential distribution can be used as an alternative to gamma or Weibull distribution in many situations. In a companion paper, the authors considered the maximum likelihood estimation of the different parameters of a generalized exponential distribution and discussed some of the testing of hypothesis problems. In this paper we mainly consider five other estimation procedures and compare their performances through numerical simulations.

*Keywords and Phrases*: Bias; Mean squared errors; Unbiased estimators; Method of moment estimators; Least squares estimators; Weighted least squares estimators; Percentiles estimators; L-estimators; Simulations

## 1. INTRODUCTION

Recently a new two-parameter distribution, named as Generalized Exponential (*GE*) Distribution has been introduced by the authors (Gupta and Kundu, 1999a). The *GE* distribution has the distribution function;

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

*Part of the work was supported by a grant from the Natural Sciences and Engineering Research Council, Canada.
†Corresponding author. e-mail: kundu@iitk.ac.in

I164T001070 . 164
T001070d.164



Therefore, *GE* distribution 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 $\alpha$ is the shape parameter and $\lambda$ is the scale parameter. *GE* distribution with the shape parameter $\alpha$ and the scale parameter $\lambda$ will be denoted by $GE(\alpha,\lambda)$. $GE(1,\lambda)$ represents the exponential distribution with the scale parameter $\lambda$.

It is observed in Gupta and Kundu (1999a) that the two-parameter $GE(\alpha,\lambda)$ can be used quite effectively in analyzing many lifetime data, particularly in place of two-parameter gamma and two-parameter Weibull distributions. The two-parameter $GE(\alpha,\lambda)$ can have increasing and decreasing failure rates depending on the shape parameter.

The main aim of this paper is to study how the different estimators of the unknown parameter/parameters of a *GE* distribution behave for different sample sizes and for different parameter values. Recently in Gupta and Kundu (1999b), we studied the properties of the maximum likelihood estimators (MLE's) in great details. In this paper, we mainly compare the MLE's with the other estimators like method of moment estimators (MME's), estimators based on percentiles (PCE's), least squares estimators (LSE's), weighted least squares estimators (WLSE's) and the estimators based on the linear combinations of order statistics (LME's), mainly with respect to their biases and mean squared errors (MSE's) using extensive simulation techniques.

The rest of the paper is organized as follows. In Section 2, we briefly discuss the MLE's and their implementations. In Sections 3 to 6 we discuss other methods. Simulation results and discussions are provided in Section 7.

## 2. MAXIMUM LIKELIHOOD ESTIMATORS

In this section the maximum likelihood estimators of $GE(\alpha, \lambda)$ are considered. We consider two different cases. First consider estimation of $\alpha$ and $\lambda$ when both are unknown. If $x_1, \ldots, x_n$ is a random sample from $GE(\alpha, \lambda)$, then the log-likelihood function, $L(\alpha, \lambda)$, is

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

The normal equations become:

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

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

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

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

Putting $\hat{\alpha}(\lambda)$ in (2.1), we obtain

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

Therefore, MLE of $\lambda$, say $\hat{\lambda}_{MLE}$, can be obtained by maximizing (2.5) with respect to $\lambda$. It is observed in Gupta and Kundu (1999b) that $g(\lambda)$ is a unimodal function and the $\hat{\lambda}_{MLE}$ which maximizes (2.5) can be obtained from the fixed point solution of

$$h(\lambda) = \lambda, \quad (2.6)$$

where

$$h(\lambda) = \left[\frac{\sum_{i=1}^{n}((x_i e^{-\lambda x_i})/(1 - e^{-\lambda x_i}))}{\sum_{i=1}^{n} ln(1 - e^{-\lambda x_i})} + \frac{1}{n}\sum_{i=1}^{n} \frac{x_i}{(1 - e^{-\lambda x_i})}\right]^{-1}. \quad (2.7)$$

Very simple iterative procedure can be used to find a solution of (2.6) and it works very well. Once we obtain $\hat{\lambda}_{MLE}$, the MLE of $\alpha$ say $\hat{\alpha}_{MLE}$ can be obtained from (2.4) as $\hat{\alpha}_{MLE} = \hat{\alpha}(\hat{\lambda}_{MLE})$.

Now we state the asymptotic normality results to obtain the asymptotic variances of the different parameters. It can be stated as follows:

$$\left[\sqrt{n}(\hat{\alpha}_{MLE} - \alpha), \sqrt{n}(\hat{\lambda}_{MLE} - \lambda)\right] \rightarrow N_2(\mathbf{0}, \mathbf{I}^{-1}(\alpha, \lambda)) \tag{2.8}$$

where $\mathbf{I}(\alpha, \lambda)$ is the Fisher Information matrix, *i.e.*,

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

The elements of the Fisher Information matrix are as follows, for $\alpha > 2$;

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

and for $0 < \alpha \leq 2$,

$$\begin{aligned} E\left(\frac{\partial^2 L}{\partial \alpha^2}\right) &= -\frac{n}{\alpha^2}, \quad E\left(\frac{\partial^2 L}{\partial \alpha \partial \lambda}\right) \\ &= \frac{n\alpha}{\lambda}\int_0^\infty x e^{-2x}(1 - e^{-x})^{\alpha_2} dx < \infty \\ E\left(\frac{\partial^2 L}{\partial \lambda^2}\right) &= -\frac{n}{\lambda^2} - \frac{n\alpha(\alpha - 1)}{\lambda^2}\int_0^\infty x^2 e^{-2x}(1 - e^{-x})^{\alpha - 3} dx < \infty. \end{aligned}$$

Now consider the MLE of $\alpha$, when the scale parameter $\lambda$ is known.

Without loss of generality we can take $\lambda = 1$. If $\lambda$ is known the MLE of $\alpha$, say $\hat{\alpha}_{MLESCK}$, is

$$\hat{\alpha}_{MLESCK} = -\frac{n}{\sum_{i=1}^{n} ln(1 - e^{-\lambda x_i})}. \tag{2.9}$$

The distribution of $\hat{\alpha}_{MLESCK}$ is same as the distribution of $(n\alpha/Y)$, where $Y$ follows $Gamma(n, 1)$. Therefore, for $n > 2$

$$E(\hat{\alpha}_{MLESCK}) = \frac{n}{n-1}\alpha,$$
$$Var(\hat{\alpha}_{MLESCK}) = \frac{n^2}{(n-1)^2(n-2)}\alpha^2,$$
$$MSE(\hat{\alpha}_{MLESCK}) = \frac{n+2}{(n-1)(n-2)}\alpha^2.$$

Clearly $\hat{\alpha}_{MLESCK}$ is not an unbiased estimator of $\alpha$, although asymptotically it is unbiased.

From the expression of the expected value, we consider the following unbiased estimator of $\alpha$, say $\hat{\alpha}_{USCK}$,

$$\hat{\alpha}_{USCK} = \frac{n-1}{n}\hat{\alpha}_{MLESCK} = -\frac{n-1}{\sum_{i=1}^{n} ln(1 - e^{-x_i})} \tag{2.10}$$

where

$$Var(\hat{\alpha}_{USCK}) = MSE(\hat{\alpha}_{USCK}) = \frac{\alpha^2}{n-2} \tag{2.11}$$

Therefore, $V(\hat{\alpha}_{USCK})$ is closer to the Cramer-Rao lower bound $(= (\alpha^2/n))$ compared to the MLE.

Now consider the MLE of $\lambda$ when the shape parameter $\alpha$ is known. For known $\alpha$ the MLE of $\lambda$ say $\hat{\lambda}_{MLESHK}$ can be obtained by maximizing

$$u(\lambda) = nln(\lambda) + (\alpha - 1)\sum_{i=1}^{n} ln(1 - e^{-\lambda x_i}) - \lambda\sum_{i=1}^{n} x_i \tag{2.12}$$

with respect to $\lambda$. It can be easily shown that $u(\lambda)$ is a unimodal function of $\lambda$ and $\hat{\lambda}_{MLESHK}$ which maximizes $u(\lambda)$ can be obtained as the fixed point solution of

$$v(\lambda) = \lambda \tag{2.13}$$

where

$$v(\lambda) = \left[\frac{1}{n}\sum_{i=1}^{n}\frac{x_i(1-\alpha e^{-\lambda x_i})}{1-e^{-\lambda x_i}}\right]^{-1}.$$

## 3. METHOD OF MOMENT ESTIMATORS

In this section we provide the method of moment estimators of the parameters of a *GE* distribution. First we consider the case when both the parameters are unknown. If $X$ follows $GE(\alpha, \lambda)$, then

$$\mu = E(X) = \frac{1}{\lambda}(\psi(\alpha+1) - \psi(1)), \tag{3.1}$$

$$\sigma^2 = V(X) = -\frac{1}{\lambda^2}(\psi'(\alpha+1) - \psi'(1)) \tag{3.2}$$

see Gupta and Kundu (1999a). Here $\psi(\cdot)$ denotes the digamma function and $\psi'(\cdot)$ denotes the derivative of $\psi(\cdot)$. From (3.1) and (3.2), we obtain the coefficient of variation $(C.V.)$ as

$$\frac{\sigma}{\mu} = C.V. = \frac{\sqrt{V(X)}}{E(X)} = \frac{\sqrt{\psi'(1) - \psi'(\alpha+1)}}{\psi(\alpha+1) - \psi(1)}. \tag{3.3}$$

The $C.V.$ is independent of the scale parameter $\lambda$. Therefore, equating the sample $C.V.$ with the population $C.V.$, we obtain

$$\frac{S}{\bar{X}} = \frac{\sqrt{\psi'(1) - \psi'(\alpha+1)}}{\psi(\alpha+1) - \psi(1)}, \tag{3.4}$$

where $S^2 = ((\sum_{i=1}^{n}(X_i - \bar{X})^2)/(n-1))$ and $\bar{X} = (1/n)\sum_{i=1} X_i$. We need to solve (3.4) to obtain the MME of $\alpha$, say $\hat{\alpha}_{MME}$. Once we estimate $\alpha$, we can use (3.1) to obtain the MME of $\lambda$. We need to use some iterative procedure to solve (3.4). The extensive table of the population $C.V.$ for different values $\alpha$ can be obtained from the authors on request. The table can be used to obtain an initial estimate of $\alpha$. Note that the MME's of $\alpha$ and $\lambda$ say $\hat{\alpha}_{MME}$ and $\hat{\lambda}_{MME}$ have the following asymptotic property.

$$\left[\sqrt{n}(\hat{\alpha}_{MME} - \alpha), \sqrt{n}(\hat{\lambda}_{MME} - \lambda)\right] \rightarrow N_2[\mathbf{0}, \mathbf{D}\mathbf{A}^{-1}\mathbf{C}\mathbf{A}^{-1}\mathbf{D}], \tag{3.5}$$

where

$$\mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} -\psi'(\alpha+1) & \psi''(\alpha+1) \\ \psi(\alpha+1) - \psi(1) & 2(-\psi'(\alpha+1) + \psi'(1)) \end{bmatrix}$$

and

$$\mathbf{C} = \begin{bmatrix} \psi'(1) - \psi'(\alpha+1) & \psi''(\alpha+1) - \psi'(1) \\ \psi''(\alpha+1) - \psi''(1) & (\psi^{(3)} - \psi^{(3)}(\alpha+1)) \\ & -(\psi'(1) - \psi'(\alpha+1))^2 \end{bmatrix}.$$

Here $\psi''(\cdot)$ and $\psi^{(3)}$ are the second and the third derivative of the $\psi(\cdot)$. The proof of (3.5) is provided in the appendix.

If the scale parameter is known, then the MME of $\alpha$ can be obtained by solving the non-linear equation

$$\lambda\bar{X} = \psi(\alpha+1) - \psi(1). \tag{3.6}$$

We need to solve the non-linear Eq. (3.6) by some iterative technique. Since the population $C.V.$ is independent of $\lambda$, therefore the $C.V.$ table can be used to obtain an initial MME of $\alpha$ even if $\lambda$ is known.

Now consider the case when the shape parameter $\alpha$ is known. If the shape parameter is known, then the MME of $\lambda$ is

$$\hat{\lambda}_{MMESHK} = \frac{\psi(\alpha+1) - \psi(1)}{\bar{X}}. \tag{3.7}$$

Note that (3.7) follows easily from (3.1). Although $\hat{\lambda}_{MMESHK}$ is not an unbiased estimator of $\lambda$ but $(1/\hat{\lambda}_{MMESHK})$ is an unbiased estimator of $(1/\lambda)$ and also

$$V\left(\frac{\lambda}{\hat{\lambda}_{MMESHK}}\right) = \frac{\psi'(1) - \psi'(\alpha+1)}{n(\psi(\alpha+1) - \psi(1))^2}.$$

## 4. ESTIMATORS BASED ON PERCENTILES

Among the most easily obtained estimators of the parameters of the Weibull distribution are the graphical approximation to the best linear unbiased estimators. It can be obtained by fitting a straight line to the theoretical points obtained from the distribution function and the sample percentile points. This method was originally explored by Kao

(1958, 1959), see also Mann, Schafer and Singpurwalla (1974) and Johnson, Kotz and Balakrishnan (1994). It is possible for the Weibull case because of the nature of its distribution function.

In case of a *GE* distribution also it is possible to use the same concept to obtain the estimators of $\alpha$ and $\lambda$ based on the percentiles, because of the structure of its distribution function. First let's consider the case, when both the parameters are unknown. Since

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

therefore

$$-\frac{1}{\lambda} ln(1 - [F(x;\alpha,\lambda)]^{(1/\alpha)}) = x. \tag{4.1}$$

If $p_i$ denotes some estimate of $F(x_{(i)};\alpha,\lambda)$ then the estimate of $\alpha$ and $\lambda$ can be obtained by minimizing

$$\sum_{i=1}^{n} \left[x_{(i)} + \lambda^{-1} ln(1 - p_i^{(1/\alpha)})\right]^2 \tag{4.2}$$

with respect to $\alpha$ and $\lambda$. (4.2) is a non-linear function of $\alpha$ and $\lambda$. It is possible to use some non-linear regression techniques to estimate $\alpha$ and $\lambda$ simultaneously. These estimators we call as percentile estimators (PCE's). It may be mentioned that approximating the true least squares estimators one tacitly and incorrectly assumes that the covariance matrix of the vector of order statistics is some constant times the identity matrix, which is not a correct assumption. It is possible to use several $p_i$'s as estimators of $F(x_{(i)})$. For example $p_i = (i/(n+1))$ is the most used estimator of $F(x_{(i)})$, as $(i/(n+1))$ is the expected value of $F(x_{(i)})$. We have also used this $p_i$ here. Some of the other choices of $p_i$'s are $p_i = ((i-(3/8))/(n+(1/4)))$ or $p_i = ((i-(1/2))/n)$, (see Mann, Schafer and Singpurwalla (1974)) although they have not pursued here.

Now let's consider the case when one parameter is known. If the shape parameter $\alpha$ is known, then the estimator of $\lambda$ can be obtained by minimizing (4.2) with respect to $\lambda$ only. The percentile estimator of $\lambda$ for known $\alpha$, say $\hat{\lambda}_{PCESHK}$, becomes

$$\hat{\lambda}_{PCESHK} = -\frac{\sum_{i=1}^{n} [ln(1 - p_i^{(1/\alpha)})]^2}{\sum_{i=1}^{n} x_{(i)} ln(1 - p_i^{(1/\alpha)})}. \tag{4.3}$$

It is an explicit form unlike MLE.

If the scale parameter $\lambda$ is known, without loss of generality we can assume $\lambda = 1$. With $\lambda = 1$, the distribution function of $GE(\alpha, 1)$ becomes

$$F(x;\alpha) = (1 - e^{-x})^{\alpha}$$

or

$$ln(F(x;\alpha)) = \alpha ln(1 - e^{-x}). \tag{4.4}$$

Therefore, similarly as before the percentile estimator of $\alpha$ for known $\lambda$, say $\hat{\alpha}_{PCESCK}$, can be obtained by minimizing

$$\sum_{i=1}^{n} [ln(p_i) - \alpha ln(1 - e^{-x_{(i)}})]^2 \tag{4.5}$$

with respect $\alpha$ and it becomes

$$\hat{\alpha}_{PCESCK} = \frac{\sum_{i=1}^{n} ln(p_i) ln(1 - e^{-x_{(i)}})}{\sum_{i=1}^{n} [ln(1 - e^{-x_{(i)}})]^2}. \tag{4.6}$$

Interestingly $\hat{\alpha}_{PCESCK}$ is also in a closed form like $\hat{\alpha}_{MLESCK}$. Note that, as one of the referees had properly mentioned, the percentile estimator of $\alpha$ for known $\lambda$ is very much like a least sqaures estimator.

## 5. LEAST SQUARES AND WEIGHTED LEAST SQUARES ESTIMATORS

In this section we provide the regression based method estimators of the unknown parameters, which was originally suggested by Swain, Venkatraman and Wilson (1988) to estimate the parameters of beta distributions. It can be used some other cases also. Suppose $Y_1, \ldots, Y_n$ is a random sample of size $n$ from a distribution function $G(\cdot)$ and suppose $Y_{(i)}$; $i = 1, \ldots, n$ denotes the ordered sample. The proposed method uses the distribution of $G(Y_{(i)})$. For a sample of size $n$, we have

$$E(G(Y_{(j)})) = \frac{j}{n+1}, \qquad V(G(Y_{(j)})) = \frac{j(n-j+1)}{(n+1)^2(n+2)}$$

and

$$Cov(G(Y_{(j)}), G(Y_{(k)})) = \frac{j(n-k+1)}{(n+1)^2(n+2)}; \quad \text{for } j < k,$$

see Johnson, Kotz and Balakrishnan (1995). Using the expectations and the variances, two variants of the least squares methods can be used.

**Method 1 (Least Squares Estimators)** Obtain the estimators by minimizing

$$\sum_{j=1}^{n}\left(G(Y_{(j)}) - \frac{j}{n+1}\right)^2. \tag{5.1}$$

with respect to the unknown parameters. Therefore in case of $GE$ distribution the least squares estimators of $\alpha$ and $\lambda$, say $\hat{\alpha}_{LSE}$, $\hat{\lambda}_{LSE}$ respectively, can be obtained by minimizing

$$\sum_{j=1}^{n}\left((1 - e^{-\lambda x_{(j)}})^{\alpha} - \frac{j}{n+1}\right)^2. \tag{5.2}$$

with respect to $\alpha$ and $\lambda$.

**Method 2 (Weighted Least Squares Estimators)** The weighted least squares estimators can be obtained by minimizing

$$\sum_{j=1}^{n} w_j\left(G(Y_{(j)}) - \frac{j}{n+1}\right)^2, \tag{5.3}$$

with respect to the unknown parameters, where $w_j = (1/(V(G(Y_{(j)})))) = ((n+1)^2(n+2))/(j(n-j+1)))$. Therefore, in case of $GE$ distribution the weighted least squares of $\alpha$ and $\lambda$, say $\hat{\alpha}_{WLSE}$ and $\hat{\lambda}_{WLSE}$ respectively, can be obtained by minimizing

$$\sum_{j=1}^{n} w_j\left((1 - e^{-\lambda x_{(j)}})^{\alpha} - \frac{j}{n+1}\right)^2. \tag{5.4}$$

with respect to $\alpha$ and $\lambda$ only.

## 6. L-MOMENT ESTIMATORS

In this section we propose a method of estimating the unknown parameters of a $GE$ distribution based on the linear combination of order statistics, see for example David (1981) or Hosking (1990). The estimators obtained by this method are popularly known as L-moment estimators (LME's). The LME's are analogous to the conventional moment estimators but can be estimated by linear combinations of order statistics, *i.e.*, by L-statistics. The LME's have theoretical advantages over conventional moments of being more robust to the presence of outliers in the data. It is observed that LME's are less subject to bias in estimation and sometimes more accurate in small samples than even the MLE's.

First we discuss the case how to obtain the LME's when both the parameters of a *GE* distribution are unknown. If $x_{(1)} < \cdots < x_{(n)}$ denote the ordered sample then using the same notation as Hosking (1990), we obtain first and second sample L-moments as

$$l_1 = \frac{1}{n}\sum_{i=1}^{n} x_{(i)}, \qquad l_2 = \frac{2}{n(n-1)}\sum_{i=1}^{n}(i-1)x_{(i)} - l_1. \tag{6.1}$$

and first two population L-moments are

$$\lambda_1 = \frac{1}{\lambda}[\psi(\alpha+1) - \psi(1)], \qquad \lambda_2 = \frac{1}{\lambda}[\psi(2\alpha+1) - \psi(\alpha+1)], \tag{6.2}$$

respectively. Note that (6.2) follows from the distribution function of the $i$th order statistic of a *GE* random variable (see Gupta and Kundu, 1999a). Now to obtain the LME's of the unknown parameters $\alpha$ and $\lambda$, we need to equate the sample L-moments with the population L-moments. Therefore, the LME's can be obtained from

$$l_1 = \frac{1}{\lambda}[\psi(\alpha+1) - \psi(1)], \tag{6.3}$$

$$l_2 = \frac{1}{\lambda}[\psi(2\alpha+1) - \psi(\alpha+1)]. \tag{6.4}$$

First we obtain LME of $\alpha$, say $\hat{\alpha}_{LME}$ as solution of the following non-linear equation

$$\frac{\psi(2\alpha+1) - \psi(\alpha+1)}{\psi(\alpha+1) - \psi(1)} = \frac{l_2}{l_1}. \tag{6.5}$$

Once $\hat{\alpha}_{LME}$ is obtained, the LME of $\lambda$, say, $\hat{\lambda}_{LME}$, can be obtained from (6.3) as

$$\hat{\lambda}_{LME} = \frac{\psi(\hat{\alpha}_{LME} + 1) - \psi(1)}{l_1}. \tag{6.6}$$

It is interesting to note that if $\alpha$ or $\lambda$ is known, then the LME of $\lambda$ or $\alpha$ is same as the corresponding moment estimator.

## 7. NUMERICAL EXPERIMENTS AND DISCUSSIONS

It is very difficult to compare the theoretical performances of the different estimators proposed in the previous sections. Therefore, we perform extensive simulations to compare the performances of the different methods mainly with respect to their biases and mean squared errors (MSE's), for different sample sizes and for different parametric values. All the computations are performed at the University of New Brunswick, Saint John, using the Sun Work-stations. Note that the generation of the $GE(\alpha, \lambda)$ is very simple. If $U$ follows uniform distribution in $[0, 1]$, then $X = ((-ln(1 - U^{(1/\alpha)})/\lambda)$ follows $GE(\alpha, \lambda)$. Therefore, if one has a good uniform random number generator, then the generation of $GE$ random deviate is immediate. We use the random deviate generator of Press *et al.* (1993) for uniform generator. We also use the subroutines of Press *et al.* (1993) for computing the minimization or maximization of a function and Psi function computations.

We consider different sample sizes ranging from very small to large. Since $\lambda$ is the scale parameter and all the estimators are scale invariant, we take $\lambda = 1$ in all our computations and we consider different values of $\alpha$. We report the average relative estimates and average relative MSE's over 10,000 replications for different cases. Note that it will give the accuracy in the order $\pm\,(10{,}000)^{-.5} = \pm\,.01$ (Karian and Dudewicz, 1999). Therefore, we report all the results up to three decimal places. First we observe how the different methods perform in estimating $\alpha$ if $\lambda$ is known.

### 7.1. Estimation of $\alpha$ when $\lambda$ is Known

If the scale parameter $\lambda$ is known, the MLE's, PCE's and the unbiased estimators (UBE's) of $\alpha$ can be obtained directly from (2.9), (4.6) and (2.10) respectively. The MME's of $\alpha$ can be obtained by solving the non-linear Eq. (3.6). The LSE's and WLSE's can be obtained by minimizing (5.2) and (5.4) respectively with respect to $\alpha$ only. If $\hat{\alpha}$ is an estimate (any one of those) then we report the average values of $(\hat{\alpha}/\alpha)$ and also the average MSE's of $(\hat{\alpha}/\alpha)$. We report the results for $\alpha = .2, .6, 1.0, 2.0, 2.5$ and for $n = 10$, (small sample) 20 (moderate sample), 30, 50 and 100 (very large sample). The results are reported in Table I. For each method the average value of $(\hat{\alpha}/\alpha)$ is reported in each box and the corresponding MSE is reported within parenthesis.

Some of the points are very clear from Table I. For each method the average relative biases and the average relative MSE's decrease as sample size increases. It indicates that all the methods provide asymptotically unbiased and consistent estimators of the shape parameter $\alpha$ for known $\lambda$. For almost all the methods (except for

TABLE I Average relative estimates and average relative mean squared errors of $\alpha$ when $\lambda$ is known

| $n$ | *Method* | $\alpha = .2$ | $\alpha = .6$ | $\alpha = 1.0$ | $\alpha = 2.0$ | $\alpha = 2.5$ |
|---|---|---|---|---|---|---|
| | MLE | 1.109(0.168) | 1.115(0.175) | 1.114(0.171) | 1.114(0.171) | 1.115(0.174) |
| | MME | 1.096(0.829) | 1.079(0.425) | 1.091(0.371) | 1.093(0.318) | 1.086(0.296) |
| $n = 10$ | PCE | 0.922(0.127) | 0.928(0.129) | 0.921(0.127) | 0.928(0.129) | 0.928(0.129) |
| | LSE | 1.088(0.230) | 1.091(0.244) | 1.089(0.216) | 1.084(0.230) | 1.081(0.237) |
| | WLSE | 1.078(0.211) | 1.074(0.224) | 1.073(0.203) | 1.082(0.228) | 1.074(0.224) |
| | UBE | 0.998(0.137) | 1.004(0.142) | 1.002(0.138) | 1.002(0.138) | 1.004(0.142) |
| | MLE | 1.056(0.066) | 1.049(0.062) | 1.052(0.063) | 1.052(0.064) | 1.049(0.062) |
| | MME | 1.027(0.329) | 1.037(0.179) | 1.036(0.141) | 1.037(0.120) | 1.040(0.120) |
| $n = 20$ | PCE | 0.929(0.063) | 0.923(0.061) | 0.929(0.062) | 0.923(0.061) | 0.923(0.061) |
| | LSE | 1.033(0.080) | 1.036(0.082) | 1.038(0.083) | 1.035(0.086) | 1.040(0.087) |
| | WLSE | 1.029(0.073) | 1.032(0.076) | 1.035(0.076) | 1.032(0.076) | 1.032(0.076) |
| | UBE | 1.003(0.060) | 0.997(0.056) | 1.000(0.057) | 1.000(0.057) | 0.997(0.056) |
| | MLE | 1.030(0.038) | 1.035(0.040) | 1.037(0.040) | 1.037(0.040) | 1.035(0.040) |
| | MME | 1.028(0.220) | 1.024(0.109) | 1.021(0.093) | 1.022(0.078) | 1.026(0.073) |
| $n = 30$ | PCE | 0.930(0.042) | 0.935(0.042) | 0.930(0.041) | 0.935(0.042) | 0.935(0.042) |
| | LSE | 1.028(0.052) | 1.024(0.051) | 1.026(0.051) | 1.026(0.052) | 1.021(0.051) |
| | WLSE | 1.025(0.047) | 1.022(0.046) | 1.025(0.046) | 1.018(0.048) | 1.022(0.046) |
| | UBE | 0.996(0.036) | 1.000(0.037) | 1.002(0.038) | 1.002(0.038) | 1.000(0.037) |
| | MLE | 1.020(0.022) | 1.020(0.022) | 1.020(0.023) | 1.020(0.023) | 1.020(0.022) |
| | MME | 1.012(0.125) | 1.016(0.063) | 1.014(0.051) | 1.014(0.043) | 1.016(0.041) |
| $n = 50$ | PCE | 0.944(0.025) | 0.946(0.026) | 0.944(0.026) | 0.946(0.026) | 0.946(0.026) |
| | LSE | 1.014(0.029) | 1.012(0.028) | 1.013(0.029) | 1.016(0.029) | 1.018(0.029) |
| | WLSE | 1.013(0.025) | 1.011(0.025) | 1.013(0.026) | 1.014(0.026) | 1.012(0.025) |
| | UBE | 1.000(0.021) | 0.999(0.021) | 1.000(0.022) | 1.000(0.022) | 0.999(0.021) |
| | MLE | 1.008(0.010) | 1.009(0.011) | 1.009(0.010) | 1.009(0.010) | 1.010(0.010) |
| | MME | 1.008(0.060) | 1.009(0.031) | 1.007(0.025) | 1.009(0.021) | 1.007(0.021) |
| $n = 100$ | PCE | 0.960(0.013) | 0.961(0.013) | 0.962(0.013) | 0.961(0.013) | 0.961(0.013) |
| | LSE | 1.008(0.014) | 1.007(0.014) | 1.005(0.014) | 1.008(0.014) | 1.006(0.014) |
| | WLSE | 1.007(0.013) | 1.006(0.013) | 1.004(0.012) | 1.007(0.013) | 1.006(0.012) |
| | UBE | 0.998(0.010) | 0.999(0.010) | 0.999(0.010) | 0.999(0.010) | 1.000(0.010) |

MME) the average values of $(\hat{\alpha}/\alpha)$ and the MSE's of $(\hat{\alpha}/\alpha)$ remain constant for different choices of $\alpha$ at different sample sizes. The same phenomena is observed even for gamma and Weibull also. Most of the methods (except PCE) usually overestimate $\alpha$, whereas PCE underestimates $\alpha$ in all the cases considered. Although, we would say that the biases are not that severe in all the methods considered. For example, when the sample size is small ($n=10$), the relative bias varies between .2% to 11.5%, and when the sample size is very large ($n=100$), the relative bias varies between 0% to 4%.

Comparing all the six methods, it is clear that as far as bias is concerned, UBE works the best for all choices of $\alpha$ and for all sample sizes. With respect to the minimum relative mean squared errors, PCE works the best for small ($n=10$) sample size and otherwise the UBE outperforms the rest. The performance of MME is quite bad particularly in terms of the MSE's. Between LSE and WLSE, WLSE works better than LSE in all cases, as expected. Now as far as computations are concerned, UBE's, MLE's and PCE's are easiest to implement (do not involve any non-linear equation solving), whereas, MME's, LSE's and WLSE's involve either non-linear equation solving or minimizing certain function. Performances of the UBE's are quite good in general and they are quite close to the best performed cases in most of the times. Considering all the points, we recommend to use UBE's for estimating $\alpha$, when $\lambda$ is known.

### 7.2. Estimation of $\lambda$ when $\alpha$ is Known

In this subsection we present the results of the different methods of estimating $\lambda$, for known $\alpha$. In this situation, the MLE of $\lambda$ can be obtained by maximizing (2.12) or equivalently solving for the fixed point solution of (2.13). The MME and PCE can be obtained directly from (3.7) and (4.3) respectively. The LSE and the WLSE can be obtained by minimizing (5.2) and (5.4) respectively with respect to $\lambda$ only. Since $\lambda$ is the scale parameter, we take $\lambda=1$ throughout without loss of generality. We consider different sample sizes namely, $n=10, 20, 30, 50, 100$ and different $\alpha=.5$, 1.5, 2.5. For a given sample we estimate $\lambda$ by different methods and report the average values of $\hat{\lambda}$ and the average MSE's of $\hat{\lambda}$ over 10,000 replications. The results are reported in Table II. Similarly as in Table I, in each box corresponds

TABLE II Average relative estimates and average relative mean squared errors of $\lambda$ when $\alpha$ is known

| $n$ | *Method* | $\alpha = .5$ | $\alpha = 1.5$ | $\alpha = 2.5$ |
|---|---|---|---|---|
| | MLE | 1.244(0.524) | 1.073(0.100) | 1.043(0.057) |
| | MME | 1.234(0.510) | 1.076(0.101) | 1.049(0.058) |
| $n = 10$ | PCE | 0.996(0.321) | 0.967(0.084) | 0.969(0.054) |
| | LSE | 1.234(1.174) | 1.053(0.122) | 1.029(0.065) |
| | WLSE | 1.215(1.125) | 1.049(0.115) | 1.028(0.062) |
| | MLE | 1.111(0.162) | 1.032(0.040) | 1.023(0.026) |
| | MME | 1.107(0.161) | 1.034(0.044) | 1.027(0.026) |
| $n = 20$ | PCE | 0.959(0.122) | 0.960(0.040) | 0.972(0.027) |
| | LSE | 1.091(0.251) | 1.023(0.049) | 1.016(0.029) |
| | WLSE | 1.081(0.219) | 1.021(0.046) | 1.016(0.028) |
| | MLE | 1.072(0.089) | 1.025(0.026) | 1.015(0.017) |
| | MME | 1.069(0.089) | 1.026(0.026) | 1.016(0.017) |
| $n = 30$ | PCE | 0.955(0.076) | 0.968(0.028) | 0.973(0.018) |
| | LSE | 1.062(0.145) | 1.017(0.031) | 1.011(0.019) |
| | WLSE | 1.056(0.124) | 1.016(0.029) | 1.010(0.018) |
| | MLE | 1.043(0.047) | 1.013(0.015) | 1.008(0.009) |
| | MME | 1.041(0.047) | 1.013(0.015) | 1.009(0.009) |
| $n = 50$ | PCE | 0.958(0.044) | 0.972(0.017) | 0.978(0.011) |
| | LSE | 1.037(0.073) | 1.007(0.018) | 1.005(0.010) |
| | WLSE | 1.034(0.062) | 1.007(0.016) | 1.005(0.010) |
| | MLE | 1.022(0.021) | 1.006(0.007) | 1.004(0.005) |
| | MME | 1.022(0.021) | 1.007(0.007) | 1.005(0.004) |
| $n = 100$ | PCE | 0.968(0.022) | 0.980(0.008) | 0.985(0.006) |
| | LSE | 1.020(0.033) | 1.004(0.009) | 1.003(0.005) |
| | WLSE | 1.020(0.029) | 1.004(0.008) | 1.003(0.005) |

to different methods the average values of $\hat{\lambda}$ are reported and their MSE's are reported within brackets.

Some of the points are clear from the experiments. As sample size increases for all the methods the biases and the MSE's decrease. It indicates that all the methods provide asymptotically unbiased and consistent estimators of $\lambda$ when $\alpha$ is known. If $\hat{\lambda}$ denotes any one of those estimators, then it is clear that $(\hat{\lambda}/\lambda)$ is not independent of $\alpha$. The bias and the MSE decrease as $\alpha$ increases. For small $\alpha$, the biases are quite severe for all the methods. For example, when the sample size is 10, the bias of MLE is around 24%. Some bias correction techniques, like Jackknifing, may be used to reduce the bias, although it is not pursued here. When $\alpha$ is large or when the sample size is large the bias is not that severe. From Table II, it is clear that for known $\alpha$, the estimation of $\lambda$ is less accurate for small values of $\alpha$, particularly if $\alpha < 1$. It is observed that all the methods (except PCE) over estimate $\lambda$, where as PCE under estimates $\lambda$.

Comparing the biases, it is observed that PCE works the best for small ($n = 10$) and moderate ($n = 20$) sample sizes and for all values of $\alpha$. For large or very large sample sizes ($n \geq 30$) and if $\alpha > 1$, WLSE works very well. But for $\alpha < 1$, PCE performs better than the other proposed methods. Similarly comparing the MSE's, it is observed that PCE works the best for small and moderate sample sizes and also for all choices of $\alpha$. PCE works very well even for large sample sizes if $\alpha < 1$. If $\alpha > 1$, for large sample sizes MME works better than the other methods. Since PCE and MME are very simple to implement and there is not much difference in the performances between PCE and WLSE for large sample and when $\alpha > 1$, we recommend as follows. Use PCE for small and moderate sample sizes at all the times. If sample size is large and $\alpha < 1$, use PCE and if $\alpha > 1$, use MME.

### 7.3. Estimation of $\alpha$ and $\lambda$ when Both are Unknown

In this subsection, we present the results of the different methods when both the parameters are unknown. The $\hat{\lambda}_{MLE}$ can be obtained from the fixed point solution of (2.6) and $\hat{\alpha}_{MLE}$ can be obtained from (2.4). The $\hat{\alpha}_{MME}$ or $\hat{\alpha}_{LME}$ can be obtained by solving the non-linear Eq. (3.4) or (6.5), and then $\hat{\lambda}_{MME}$ or $\hat{\lambda}_{LME}$ can be obtained from (3.7) or (6.6). The PCE's, LSE's and WLSE's can be obtained by minimizing (4.2), (5.2) and (5.4) respectively with respect to $\alpha$ and $\lambda$. We consider different sample sizes and different values of $\alpha$. We take $n = 15$, 20, 30, 50, 100 and $\alpha = .2$, .5, 1.0, 2.0, 2.5. Throughout, we consider $\lambda = 1$. For each combination of $n$ and $\alpha$ we generate a sample of size $n$ from $GE(\alpha, 1)$ and estimate $\alpha$ and $\lambda$ by different methods. We report the average values of $(\hat{\alpha}/\alpha)$ and $\hat{\lambda}$ and also the corresponding average MSE's. All the reported results are based on 10,000 replications. The results are presented in Tables III and IV. In Table III, we report the average values of $(\hat{\alpha}/\alpha)$ for each method and the corresponding MSE's are reported within brackets. Similar results for $\hat{\lambda}$ are reported in Table IV. For a quick understanding, the relative biases and the relative MSE's of the different estimators of the scale parameter when the shape parameter is also unknown is presented in Figures 1–4 for sample sizes 20 and 100. The other cases are similar in nature and not provided for space restrictions.

From the Tables III and IV it is immediate that even when both the parameters are unknown the average biases and the average MSE's decrease as sample size increases. It verifies the asymptotic unbiasedness and the consistency of all the estimators. It is observed that the average biases and the average MSE's of $(\hat{\alpha}/\alpha)$ and $\hat{\lambda}$ depend on $\alpha$. For all the methods as $\alpha$ increases the average relative MSE's of $\hat{\lambda}$ decrease and the same thing is true for the average biases also for most of the methods. On the other hand there is no pattern observed for the average biases of $(\hat{\alpha}/\alpha)$ and the corresponding average MSE's. It may be mentioned that for most of the methods the biases are quite severe for small sample sizes ($n = 15$) for both $(\hat{\alpha}/\alpha)$ and $\hat{\lambda}$. Considering only MSE's it can be said that the estimation of $\alpha$'s are more accurate for

TABLE III Average relative estimates and average relative mean squared errors of $\alpha$ when $\lambda$ is unknown

| $n$ | *Method* | $\alpha = .2$ | $\alpha = .5$ | $\alpha = 1.0$ | $\alpha = 2.0$ | $\alpha = 2.5$ |
|---|---|---|---|---|---|---|
| | MLE | 1.145(0.153) | 1.180(0.230) | 1.237(0.436) | 1.338(0.857) | 1.356(0.907) |
| | MME | 1.602(0.990) | 1.396(0.679) | 1.383(0.871) | 1.455(1.624) | 1.465(1.466) |
| $n = 15$ | PCE | 1.192(0.670) | 1.043(0.416) | 1.019(0.420) | 1.015(0.508) | 1.016(0.637) |
| | LSE | 1.059(0.217) | 1.065(0.362) | 1.118(1.952) | 1.161(1.880) | 1.214(9.838) |
| | WLSE | 1.068(0.178) | 1.069(0.316) | 1.123(2.433) | 1.168(1.575) | 1.205(2.099) |
| | LME | 1.064(0.334) | 1.089(0.306) | 1.145(0.446) | 1.226(0.842) | 1.235(0.800) |
| | MLE | 1.108(0.102) | 1.132(0.141) | 1.160(0.199) | 1.223(0.343) | 1.244(0.424) |
| | MME | 1.458(0.644) | 1.314(0.448) | 1.285(0.479) | 1.318(1.675) | 1.338(0.791) |
| $n = 20$ | PCE | 1.113(0.464) | 1.009(0.309) | 0.988(0.277) | 0.980(0.306) | 0.972(0.328) |
| | LSE | 1.044(0.110) | 1.033(0.154) | 1.064(0.316) | 1.103(0.565) | 1.113(0.622) |
| | WLSE | 1.058(0.111) | 1.041(0.132) | 1.069(0.248) | 1.102(0.597) | 1.120(0.907) |
| | LME | 1.041(0.230) | 1.071(0.203) | 1.101(0.245) | 1.149(0.376) | 1.167(0.454) |
| | MLE | 1.063(0.053) | 1.084(0.075) | 1.100(0.100) | 1.129(0.146) | 1.145(0.178) |
| | MME | 1.305(0.354) | 1.218(0.248) | 1.191(0.242) | 1.197(0.293) | 1.213(0.357) |
| $n = 30$ | PCE | 1.033(0.315) | 0.969(0.206) | 0.957(0.181) | 0.955(0.190) | 0.952(0.190) |
| | LSE | 1.029(0.068) | 1.021(0.089) | 1.034(0.124) | 1.050(0.209) | 1.054(0.229) |
| | WLSE | 1.045(0.062) | 1.030(0.078) | 1.048(0.184) | 1.063(0.166) | 1.071(0.251) |
| | LME | 1.014(0.136) | 1.050(0.113) | 1.062(0.126) | 1.082(0.164) | 1.099(0.204) |
| | MLE | 1.039(0.028) | 1.048(0.038) | 1.054(0.045) | 1.077(0.068) | 1.082(0.077) |
| | MME | 1.203(0.199) | 1.134(0.131) | 1.112(0.117) | 1.123(0.145) | 1.125(0.157) |
| $n = 50$ | PCE | 0.983(0.207) | 0.949(0.136) | 0.955(0.115) | 0.946(0.111) | 0.944(0.113) |
| | LSE | 1.017(0.036) | 1.010(0.043) | 1.016(0.060) | 1.023(0.073) | 1.027(0.101) |
| | WLSE | 1.030(0.032) | 1.018(0.037) | 1.025(0.053) | 1.032(0.074) | 1.044(0.086) |
| | LME | 1.012(0.079) | 1.028(0.063) | 1.031(0.062) | 1.052(0.082) | 1.056(0.090) |
| | MLE | 1.019(0.012) | 1.022(0.016) | 1.027(0.020) | 1.035(0.027) | 1.038(0.030) |
| | MME | 1.107(0.090) | 1.065(0.060) | 1.056(0.053) | 1.058(0.061) | 1.061(0.066) |
| $n = 100$ | PCE | 0.949(0.120) | 0.948(0.078) | 0.952(0.063) | 0.954(0.060) | 0.954(0.061) |
| | LSE | 1.011(0.017) | 1.006(0.021) | 1.009(0.027) | 1.011(0.033) | 1.016(0.043) |
| | WLSE | 1.023(0.015) | 1.013(0.018) | 1.017(0.023) | 1.019(0.028) | 1.024(0.036) |
| | LME | 1.005(0.039) | 1.011(0.029) | 1.016(0.029) | 1.023(0.036) | 1.026(0.039) |

TABLE IV Average relative estimates and average relative mean squared errors of $\lambda$ when $\alpha$ is unknown

| $n$ | *Method* | $\alpha=.2$ | $\alpha=.5$ | $\alpha=1.0$ | $\alpha=2.0$ | $\alpha=2.5$ |
|---|---|---|---|---|---|---|
| | MLE | 1.644(3.034) | 1.264(0.455) | 1.187(0.222) | 1.148(0.148) | 1.134(0.135) |
| | MME | 2.108(6.064) | 1.390(0.716) | 1.239(0.317) | 1.170(0.194) | 1.153(0.173) |
| $n=15$ | PCE | 1.219(4.577) | 0.946(0.392) | 0.925(0.215) | 0.928(0.144) | 0.925(0.125) |
| | LSE | 1.488(9.842) | 1.078(0.556) | 1.021(0.256) | 1.005(0.156) | 1.006(0.140) |
| | WLSE | 1.465(7.920) | 1.088(0.501) | 1.034(0.235) | 1.019(0.148) | 1.019(0.131) |
| | LME | 1.493(2.925) | 1.167(0.431) | 1.112(0.216) | 1.088(0.142) | 1.077(0.129) |
| | MLE | 1.421(1.184) | 1.132(0.287) | 1.160(0.143) | 1.222(0.097) | 1.244(0.086) |
| | MME | 1.730(2.334) | 1.306(0.461) | 1.179(0.221) | 1.129(0.134) | 1.114(0.118) |
| $n=20$ | PCE | 1.062(1.107) | 0.924(0.281) | 0.920(0.162) | 0.925(0.104) | 0.917(0.096) |
| | LSE | 1.306(2.956) | 1.044(0.324) | 1.064(0.171) | 1.103(0.113) | 1.113(0.100) |
| | WLSE | 1.320(2.585) | 1.058(0.284) | 1.029(0.153) | 1.019(0.097) | 1.014(0.086) |
| | LME | 1.300(1.153) | 1.134(0.284) | 1.081(0.152) | 1.065(0.098) | 1.057(0.088) |
| | MLE | 1.256(0.477) | 1.121(0.141) | 1.088(0.082) | 1.068(0.056) | 1.063(0.049) |
| | MME | 1.459(0.936) | 1.201(0.239) | 1.127(0.129) | 1.086(0.081) | 1.077(0.073) |
| $n=30$ | PCE | 0.961(0.486) | 0.912(0.177) | 0.919(0.108) | 0.927(0.073) | 0.931(0.063) |
| | LSE | 1.173(1.051) | 1.027(0.173) | 1.013(0.099) | 1.001(0.065) | 0.998(0.061) |
| | WLSE | 1.114(0.965) | 1.041(0.151) | 1.025(0.088) | 1.016(0.059) | 1.011(0.054) |
| | LME | 1.180(0.501) | 1.083(0.151) | 1.057(0.089) | 1.040(0.059) | 1.037(0.054) |
| | MLE | 1.138(0.189) | 1.074(0.073) | 1.046(0.041) | 1.043(0.030) | 1.037(0.026) |
| | MME | 1.268(0.381) | 1.126(0.127) | 1.072(0.067) | 1.057(0.047) | 1.047(0.041) |
| $n=50$ | PCE | 0.925(0.258) | 0.917(0.110) | 0.931(0.068) | 0.940(0.045) | 0.942(0.040) |
| | LSE | 1.092(0.339) | 1.017(0.094) | 1.001(0.054) | 0.997(0.038) | 0.997(0.034) |
| | WLSE | 1.111(0.278) | 1.031(0.080) | 1.016(0.046) | 1.008(0.032) | 1.011(0.029) |
| | LME | 1.098(0.214) | 1.052(0.083) | 1.028(0.047) | 1.028(0.034) | 1.022(0.029) |
| | MLE | 1.068(0.072) | 1.035(0.030) | 1.025(0.019) | 1.020(0.014) | 1.019(0.012) |
| | MME | 1.136(0.143) | 1.061(0.054) | 1.038(0.033) | 1.027(0.022) | 1.024(0.020) |
| $n=100$ | PCE | 0.916(0.121) | 0.935(0.058) | 0.948(0.036) | 0.958(0.024) | 0.960(0.021) |
| | LSE | 1.045(0.128) | 1.010(0.044) | 1.004(0.026) | 1.003(0.021) | 1.001(0.017) |
| | WLSE | 1.062(0.103) | 1.020(0.036) | 1.013(0.022) | 1.010(0.018) | 1.009(0.014) |
| | LME | 1.049(0.091) | 1.024(0.038) | 1.016(0.023) | 1.012(0.016) | 1.012(0.015) |

smaller $\alpha$, whereas the estimation of $\lambda$'s are more accurate for larger $\alpha$. Most of the estimators overestimate both $\alpha$ and $\lambda$ at all the times considered, except PCE, which usually underestimates the corresponding parameters, particularly for moderate and large sample sizes.

Now comparing the performances of all the methods it is clear that as far as the minimum bias is concerned, PCE works the best in almost all the cases considered for estimating both $\alpha$ and $\lambda$, followed by the LSE's and WLSE's. The performances of the MME's are the worst as far as the bias is concerned. Now with respect to the MSE's it is clear that MLE's work the best in almost all the cases considered for estimating both the parameters. The performances of the LME's are also quite close to that of the MLE's. In fact for small sample sizes,

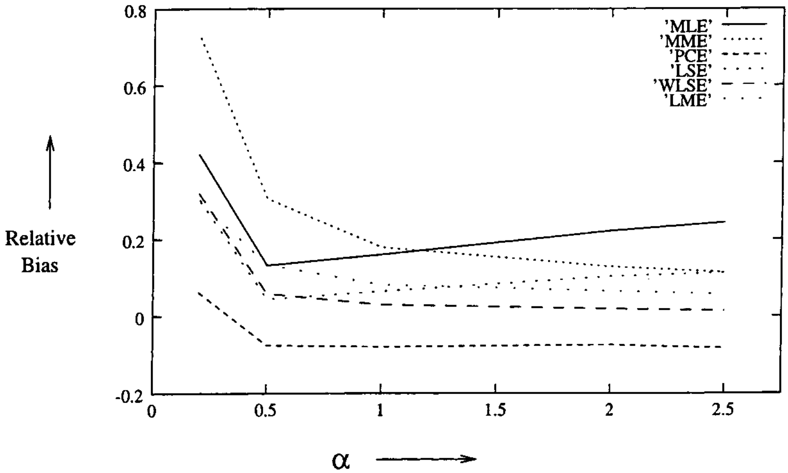


FIGURE 1 Average relative biases of the different estimators of $\lambda$ for sample size 20.

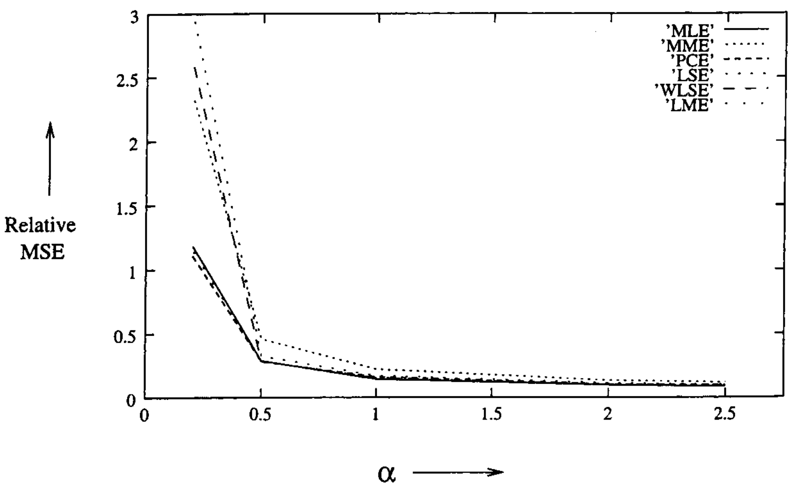


FIGURE 2 Average relative MSE's of the different estimators of $\lambda$ for sample size 20.

sometimes the performances of the LME's are better than MLE's. The MSE's of the PCE's are also quite close to that of the MLE's and LME's and they are sometimes smaller than both of them at least for the small sample sizes. The MSE's of the WLSE's are usually smaller than that of the LSE's, where as the MSE's of the MME's are usually

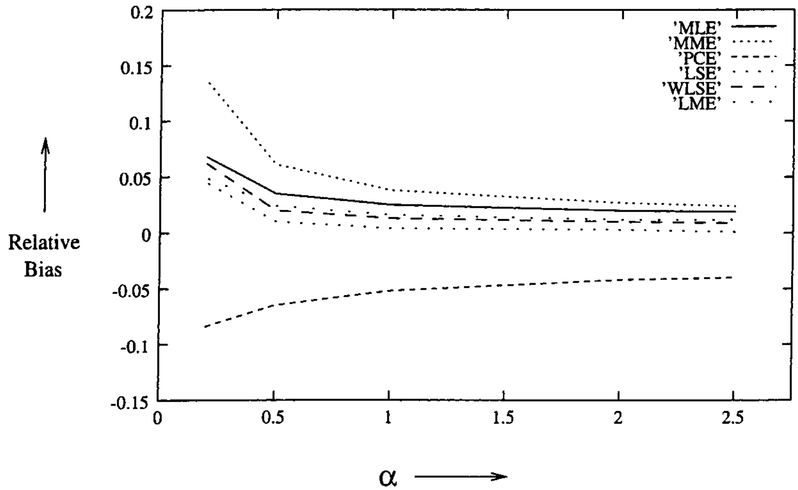


FIGURE 3 Average relative biases of the different estimators of $\lambda$ for sample size 100.

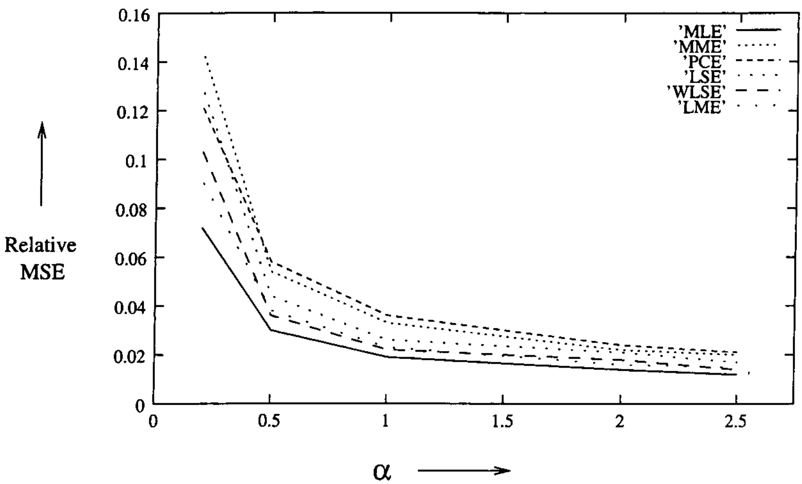


FIGURE 4 Average relative MSE's of the different estimators of $\lambda$ for sample size 100.

the maximum in most of the cases. Now if we consider the computational complexities, it is observed that MLE's, MME's and LME's involve one dimensional minimization, where as PCE's, LSE's and WLSE's involve two dimensional minimization. Another point it is worth mentioning that to compute MME's or LME's we need to evaluate $\psi(\cdot)$ function at different points. Therefore, we need to use some series expansion for this purpose. Considering all the points, we recommend to use PCE's for small sample sizes, where as MLE's for moderate or large sample sizes.

### *Acknowledgement*

The authors would like to thank one referee for some constructive suggestions.

### *References*

David, H. A. (1981) *Order Statistics*, 2nd edition, New York, Wiley.

Gupta, R. D. and Kundu, D. (1999a) "Generalized Exponential Distributions", *Australian and New Zealand Journal of Statistics*, **41**(2), 173–188.

Gupta, R. D. and Kundu, D. (1999b) "Generalized Exponential Distributions: Statistical Inferences", *Technical Report*, The University of New Brunswick, Saint John.

Hosking, J. R. M. (1990) "L-Moment: Analysis and estimation of distributions using linear combinations of order statistics", *Journal of Royal Statistical Society*, *Ser. B*, **52**(1), 105–124.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994) *Continuous Univariate Distribution*, Vol. 1, 2nd edition, New York, Wiley.

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

Kao, J. H. K. (1958) "Computer methods for estimating Weibull parameters in reliability studies", *Transaction of IRE-Reliability and Quality Control*, **13**, 15–22.

Kao, J. H. K. (1959) "A graphical estimation of mixed Weibull parameters in life testing electron tubes", *Technometrics*, **1**, 389–407.

Karian, Z. A. and Dudewicz, E. J. (1999) *Modern Statistical Systems and GPSS Simulations*, 2nd edition, CRC Press, Florida.

Mann, N. R., Schafer, R. E. and Singpurwalla, N. D. (1974) *Methods for Statistical Analysis of Reliability and Life Data*, New York, Wiley.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1993) *Numerical Recipes in FORTRAN*, Cambridge University Press, Cambridge.

Swain, J., Venkatraman, S. and Wilson, J. (1988) "Least squares estimation of distribution function in Johnson's translation system", *Journal of Statistical Computation and Simulation*, **29**, 271–297.

## APPENDIX

The asymptotic distribution of the MME's can be obtained as follows. In the appendix we use $\hat{\alpha} = \hat{\alpha}_{MME}$ and $\hat{\lambda} = \hat{\lambda}_{MME}$ for brevity. Let's define

$$f_1(\alpha, \lambda) = \bar{X} - \frac{\psi(\alpha+1) - \psi(1)}{\lambda},$$
$$f_2(\alpha, \lambda) = S^2 - \frac{-\psi'(\alpha+1) + \psi'(1)}{\lambda^2}.$$

Consider $f(\alpha, \lambda) = (f_1(\alpha, \lambda), f_2(\alpha, \lambda))$ and expand $f(\hat{\alpha}, \hat{\lambda})$ by Taylor series around the true value of $(\alpha, \lambda)$, we obtain

$$f(\hat{\alpha}, \hat{\lambda}) - f(\alpha, \lambda) = [\hat{\alpha} - \alpha, \hat{\lambda} - \lambda] \begin{bmatrix} (\partial f_1/\partial\alpha) & (\partial f_2/\partial\alpha) \\ (\partial f_1/\partial\lambda) & (\partial f_2/\partial\lambda) \end{bmatrix}_{(\alpha,\lambda)=(\bar{\alpha},\bar{\lambda})}.$$

Here $(\bar{\alpha}, \bar{\lambda})$ is a point between $(\hat{\alpha}, \hat{\lambda})$ and $(\alpha, \lambda)$. Note that $f(\hat{\alpha}, \hat{\lambda}) = 0$ and as $n \to \infty$, $(\hat{\alpha}, \hat{\lambda}) \to (\alpha, \lambda)$, so $(\bar{\alpha}, \bar{\lambda}) \to (\alpha, \lambda)$. Therefore as $n \to \infty$, the distribution of $[\sqrt{n}(\hat{\alpha} - \alpha), \sqrt{n}(\hat{\lambda} - \lambda)]$ is same as

$$-[\sqrt{n}(\bar{X} - E(X)), \sqrt{n}(S^2 - E(S^2))] \begin{bmatrix} (\partial f_1/\partial\alpha) & (\partial f_2/\partial\alpha) \\ (\partial f_1/\partial\lambda) & (\partial f_2/\partial\lambda) \end{bmatrix}^{-1}.$$

Now using the central limit theorem we obtain

$$\sqrt{n}(\bar{X} - E(\bar{X})) \to N\left(0, \frac{\psi'(1) - \psi'(\alpha+1)}{\lambda^2}\right)$$
$$\sqrt{n}(S^2 - E(S^2)) \to N\left(0, \frac{(\psi^{(3)}(1) - \psi^{(3)}(\alpha+1)) - (\psi'(1) - \psi'(\alpha+1))^2}{\lambda^4}\right)$$
$$cov(\sqrt{n}(\bar{X} - E(\bar{X})), \sqrt{n}(S^2 - E(S^2))) \to N\left(0, \frac{\psi''(\alpha+1) - \psi''(1)}{\lambda^3}\right),$$

therefore, (3.5) follows immediately.