ELSEVIER Computational Statistics & Data Analysis 43 (2003) 179–196

COMPUTATIONAL STATISTICS & DATA ANALYSIS

www.elsevier.com/locate/csda

Discriminating between Weibull and generalized exponential distributions

Rameshwar D. Gupta[a,1], 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, Kanpur 208016, India

Received 1 November 2001; received in revised form 1 April 2002

Abstract

Recently the two-parameter generalized exponential (GE) distribution was introduced by the authors. It is observed that a GE distribution can be considered for situations where a skewed distribution for a non-negative random variable is needed. The ratio of the maximized likelihoods (RML) is used in discriminating between Weibull and GE distributions. Asymptotic distributions of the logarithm of the RML under null hypotheses are obtained and they are used to determine the minimum sample size required in discriminating between two overlapping families of distributions for a user specified probability of correct selection and tolerance limit.

Keywords: Asymptotic distributions; Generalized exponential distribution; Kolmogorov–Smirnov distance; Likelihood ratio statistic; Weibull distribution

1. Introduction

Recently, the two-parameter generalized exponential (GE) distribution has been introduced and studied quite extensively by the authors (Gupta and Kundu, 1999, 2001a,b). The two-parameter GE distribution has the distribution function

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

* Corresponding author. Tel.: 91-512-597141; fax: 91-512-597500.
E-mail address: kundu@iitk.ac.in (D. Kundu).

[1] Part of the work was supported by a grant from the Natural Sciences and Engineering Research Council.

PII: S0167-9473(02)00206-2

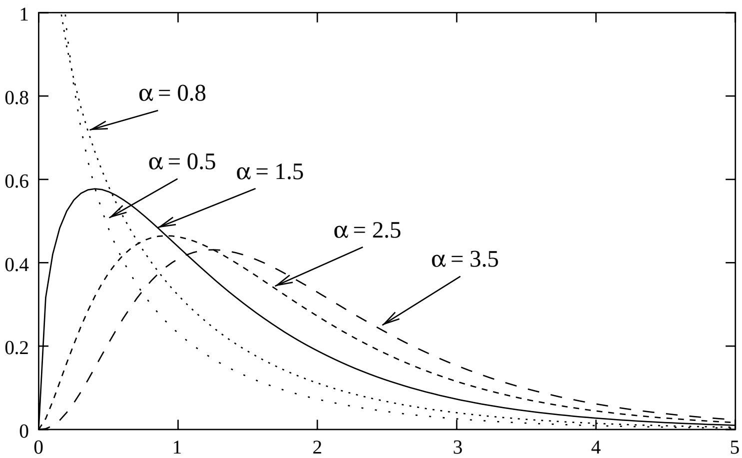

Fig. 1. Density functions of the GE distribution for different values of α when λ is constant.

density function

$$f_{GE}(x;\alpha,\lambda) = \alpha\lambda(1-e^{-\lambda x})^{\alpha-1}e^{-\lambda x}, \quad \alpha,\lambda > 0, \tag{1.2}$$

survival function

$$S_{GE}(x;\alpha,\lambda) = 1-(1-e^{-\lambda x})^{\alpha}, \quad \alpha,\lambda > 0 \tag{1.3}$$

and hazard function

$$h_{GE}(x;\alpha,\lambda) = \frac{\alpha\lambda(1-e^{-\lambda x})^{\alpha-1}e^{-\lambda x}}{1-(1-e^{-\lambda x})^{\alpha}}, \quad \alpha,\lambda > 0. \tag{1.4}$$

Here α and λ are shape and scale parameters respectively. Naturally the shape of the density function does not depend on λ. For different values of α, when $\lambda = 1$, we provide density functions of the GE distribution in Fig. 1.

It is clear that they can take different shapes and they are quite similar to Weibull density functions. When $\alpha = 1$, it coincides with the exponential distribution. The hazard function of a GE distribution can be increasing, decreasing or constant depending on the shape parameter similarly as a Weibull distribution. Therefore, GE and Weibull distributions are both generalization of an exponential distribution in different ways. If it is known or apparent from the histogram that data are coming from a right tailed distribution, then a GE distribution can be used quite effectively. It is observed that in many situations GE distribution provides *better fit* than a Weibull distribution. Therefore to analyze a skewed lifetime data an experimenter might wish to choose one of the two models. Although, these two models may provide similar data fit for moderate sample sizes but it is still desirable to select the correct or more nearly correct model, since the inferences based on the model will often involve tail probabilities where the affect of the model assumptions will be more crucial. Therefore, even if large sample sizes are not available it is still important to make the best possible decision based on whatever data are available. Discriminating between two distributions have been well studied in the statistical literature, see for example the work of Dumonceaux and Antle (1973), Dumonceaux et al. (1973), Quesenberry and Kent (1982), Balasooriya and Abeysinghe (1994) and the references cited there.

Recently the ratio of the maximized likelihoods (RML) was used by Gupta et al. (2001) in discriminating between two overlapping families of distributions. The idea was originally proposed by Cox (1961, 1962) in discriminating between two separate models and Bain and Engelhardt (1980) used it in discriminating between Weibull and gamma distributions. In this paper we obtain asymptotic distributions of the RML under null hypotheses. It is observed by a Monte Carlo simulation study that these asymptotic distributions work quite well even when sample size is not too large. Using these asymptotic distributions and the distance between two distribution functions, we determine the minimum sample size needed to discriminate between them at a user specified protection level. Two real life data are analyzed to see how the proposed method works in practice.

The rest of the paper is organized as follows. We briefly describe the likelihood ratio tests in Section 2. In Section 3, we obtain asymptotic distributions of the RML under null hypotheses. In Section 4, these asymptotic distributions are used to determine the minimum sample size needed to discriminate between two distributions at a user specified protection level and a tolerance level. Some numerical experiments are performed to observe how these asymptotic results behave for finite samples in Section 5. Two real life data sets are analyzed in Section 6 and finally the conclusion appears in Section 7.

2. Likelihood ratio test

Suppose $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d) random variables from any one of the two distribution functions. The density of a Weibull random variable with shape parameter β and scale parameter θ will be denoted by

$$f_{WE}(x;\beta,\theta) = \beta\theta^{\beta}x^{\beta-1}e^{-(x\theta)^{\beta}}. \tag{2.1}$$

A GE distribution with shape parameter α and scale parameter λ will be denoted by $GE(\alpha,\lambda)$, similarly a Weibull distribution with shape parameter β and scale parameter θ will be denoted by $WE(\beta,\theta)$. Let us define the likelihood functions assuming that the data are coming from a $GE(\alpha,\lambda)$ or from a $WE(\beta,\theta)$ as

$$L_{GE}(\alpha,\lambda) = \prod_{i=1}^{n} f_{GE}(x_i;\alpha,\lambda), \qquad L_{WE}(\beta,\theta) = \prod_{i=1}^{n} f_{WE}(x_i;\beta,\theta),$$

respectively. The RML is defined as

$$L = \frac{L_{GE}(\hat{\alpha},\hat{\lambda})}{L_{WE}(\hat{\beta},\hat{\theta})}. \tag{2.2}$$

Here $(\hat{\alpha}, \hat{\lambda})$ and $(\hat{\beta}, \hat{\theta})$ are the maximum likelihood estimators of (α, λ) and (β, θ), respectively. The logarithm of RML can be written as

$$T = n\left[\ln\left(\frac{\hat{\alpha}\hat{\lambda}\tilde{X}}{\hat{\beta}}\right) - \frac{\hat{\alpha}-1}{\hat{\alpha}} - \hat{\lambda}\bar{X} - \hat{\beta}\ln(\hat{\theta}\tilde{X}) + 1\right], \tag{2.3}$$

where $\bar{X} = (1/n)\sum_{i=1}^{n} X_i$, $\tilde{X} = (\prod_{i=1}^{n} X_i)^{1/n}$. Moreover, $\hat{\alpha}$ and $\hat{\lambda}$ (Gupta and Kundu, 2001a) have the following relation:

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

In case of a Weibull distribution, $\hat{\theta}$ and $\hat{\beta}$ satisfy the following relation:

$$\hat{\theta} = \left(\frac{n}{\sum_{i=1}^{n} X_i^{\hat{\beta}}}\right)^{1/\hat{\beta}}. \tag{2.5}$$

Gupta et al. (2001) proposed the following discrimination procedure. Choose the GE distribution if $T > 0$, otherwise choose the Weibull distribution as the preferred model. It can be easily seen that if the data come from a GE distribution then the distribution of T depends only on α and independent of λ and similarly if the data come from a Weibull distribution, then the distribution of T depends only of β.

3. Asymptotic properties of the RML under null hypotheses

In this section, we obtain asymptotic distributions of the RML statistics under null hypotheses in two different cases. From now on we denote the almost sure convergence by *a.s.*

Case 1: The data are coming from a Weibull distribution and the alternative is a GE distribution.

Let us assume that the n data points are from a Weibull distribution with shape parameter β and scale parameter θ as given in (2.1). $\hat{\beta}$, $\hat{\theta}$, $\hat{\alpha}$ and $\hat{\lambda}$ are same as defined earlier. We use the following notations. For any Borel Measurable function $h(.)$, $E_{WE}(h(U))$ and $V_{WE}(h(U))$ denote mean and variance of $h(U)$ under the assumption that U follows $WE(.,.)$. Similarly we can define $E_{GE}(h(U)$ and $V_{GE}(h(U))$ as mean and variance of $h(U)$ under the assumption that U follows $GE(.,.)$. If $g(.)$ and $h(.)$ are two Borel Measurable functions, we can define along the same line $cov_{WE}(g(U), h(U)) = E_{WE}(g(U)h(U)) - E_{WE}(g(U))E_{WE}(h(U))$ and $cov_{GE}(g(U), h(U)) = E_{GE}(g(U)h(U)) - E_{GE}(g(U))E_{GE}(h(U))$, where U follows $WE(.,.)$ and $GE(.,.)$ respectively. Now we have the following lemma.

Lemma 1. *Suppose data are from* $WE(\beta, \theta)$, *then as* $n \to \infty$, *we have*
(i) $\hat{\beta} \to \beta$ *a.s.*, $\hat{\theta} \to \theta$ *a.s.*, *where*

$$E_{WE}[\ln(f_{WE}(X;\beta,\theta))] = \max_{\bar{\beta},\bar{\theta}} E_{WE}[\ln(f_{WE}(X;\bar{\beta},\bar{\theta}))].$$

(ii) $\hat{\alpha} \to \tilde{\alpha}$ *a.s.*, $\hat{\lambda} \to \tilde{\lambda}$ *a.s.*, *where*

$$E_{WE}[\ln(f_{GE}(X;\tilde{\alpha},\tilde{\lambda}))] = \max_{\alpha,\lambda} E_{WE}[\ln(f_{GE}(X;\alpha,\lambda))].$$

Note that $\tilde{\alpha}$ and $\tilde{\lambda}$ may depend on β and θ but we do not make it explicit for brevity. Let us denote $T^* = \ln(L_{GE}(\tilde{\alpha},\tilde{\lambda})/L_{WE}(\beta,\theta))$.

(iii) $n^{-1/2}[T - E_{WE}(T)]$ is asymptotically equivalent to $n^{-1/2}[T^* - E_{WE}(T^*)]$.

Proof. The proof follows using the similar argument of White (1982, Theorem 1) and therefore it is omitted. □

Theorem 1. *Under the assumption that the data are from a Weibull distribution, the distribution of* T *is approximately normally distributed with mean* $E_{WE}(T)$ *and variance* $V_{WE}(T)$.

Proof. Using the Central Limit Theorem, it can be easily shown that $n^{-1/2}[T^* - E_{WE}(T^*)]$ is asymptotically normally distributed. Therefore the proof immediately follows from part (iii) of Lemma 1 and using the Central Limit Theorem. □

Now we discuss how to obtain $\tilde{\alpha}$, $\tilde{\lambda}$, $E_{WE}(T)$ and $V_{WE}(T)$. Let us define

$$\begin{aligned} g(\alpha,\lambda) &= E_{WE}[\ln(f_{GE}(X;\alpha,\lambda))] \\ &= E_{WE}[\ln(\alpha) + \ln(\lambda) - \lambda X + (\alpha-1)\ln(1-e^{-\lambda X})] \\ &= \ln(\alpha) + \ln(\lambda) - \frac{\lambda}{\theta}\Gamma\left(1+\frac{1}{\beta}\right) + (\alpha-1)v\left(\beta,\frac{\lambda}{\theta}\right), \end{aligned} \tag{3.1}$$

where

$$v(x,y) = \int_0^\infty \ln(1-e^{-yz^{1/x}})e^{-z}\,dz. \tag{3.2}$$

Therefore, $\tilde{\alpha}$ and $\tilde{\lambda}$ can be obtained as solutions of

$$\frac{1}{\tilde{\alpha}} + v\left(\beta,\frac{\tilde{\lambda}}{\theta}\right) = 0 \tag{3.3}$$

and

$$\frac{1}{\tilde{\lambda}} - \frac{1}{\theta}\Gamma\left(1+\frac{1}{\beta}\right) + \frac{\tilde{\alpha}-1}{\theta}v_2\left(\beta,\frac{\tilde{\lambda}}{\theta}\right) = 0. \tag{3.4}$$

Here $v_2(x,y)$ is the derivative of $v(x,y)$ with respect to the second argument y, i.e.

$$v_2(x,y) = \int_0^\infty e^{-z}\frac{z^{1/x}e^{-yz^{1/x}}}{(1-e^{-yz^{1/x}})}\,dz. \tag{3.5}$$

From (3.3), it is immediate that $(\tilde{\lambda}/\theta)$ is a function of $\tilde{\alpha}$ and β. Therefore, from (3.4) it follows that $\tilde{\alpha}$ is a function of β only and in turn $(\tilde{\lambda}/\theta)$ is also a function of β only. Now we provide the expressions for $E_{WE}(T)$ and $V_{WE}(T)$. Note that $\lim_{n\to\infty} E_{WE}(T)/n$

and $\lim_{n\to\infty} V_{WE}(T)/n$ exist. Suppose we denote $\lim_{n\to\infty} E_{WE}(T)/n = AM_{WE}(\beta)$ and $\lim_{n\to\infty} V_{WE}(T)/n = AV_{WE}(\beta)$, therefore for large n,

$$\frac{E_{WE}(T)}{n} \approx AM_{WE}(\beta) = E_{WE}[\ln(f_{GE}(X;\tilde{\alpha},\tilde{\lambda})) - \ln(f_{WE}(X;\beta,\theta))]$$

$$= \ln(\tilde{\alpha}) + \ln\left(\frac{\tilde{\lambda}}{\theta}\right) + (\tilde{\alpha}-1)E_{WE}[\ln(1-\mathrm{e}^{-(\tilde{\lambda}/\theta)Y})] - \ln(\beta)$$

$$-\left(\frac{\tilde{\lambda}}{\theta}\right)\Gamma\left(1+\frac{1}{\beta}\right) - (\beta-1)\frac{\psi(1)}{\beta} + 1. \tag{3.6}$$

Here Y is a random variable such that Y follows $WE(\beta,1)$ and $\psi(.)$ is the digamma function. We also have

$$\frac{V_{WE}(T)}{n} \approx AV_{WE}(\beta) = V_{WE}[\ln(f_{GE}(X;\tilde{\alpha},\tilde{\lambda})) - \ln(f_{WE}(X;\beta,\theta))]$$

$$= V_{WE}\left[(\tilde{\alpha}-1)\ln(1-\mathrm{e}^{-(\tilde{\lambda}/\theta)Y}) - \left(\frac{\tilde{\lambda}}{\theta}\right)Y - (\beta-1)\ln(Y) + Y^{\beta}\right]$$

$$= (\tilde{\alpha}-1)^2 V_{WE}[\ln(1-\mathrm{e}^{-(\tilde{\lambda}/\theta)Y})] + \left(\frac{\tilde{\lambda}}{\theta}\right)^2\left[\Gamma\left(\frac{2}{\beta}+1\right) - \Gamma^2\left(\frac{1}{\beta}+1\right)\right]$$

$$+1 + (\beta-1)^2\,\frac{\psi'(1)}{\beta^2} - 2(\tilde{\alpha}-1)\,\frac{\tilde{\lambda}}{\theta}\,cov_{WE}(\ln(1-\mathrm{e}^{-(\tilde{\lambda}/\theta)Y}), Y)$$

$$-2(\tilde{\alpha}-1)(\beta-1)cov_{WE}(\ln(1-\mathrm{e}^{-(\tilde{\lambda}/\theta)Y}), \ln(Y))$$

$$+2(\tilde{\alpha}-1)cov_{WE}(\ln(1-\mathrm{e}^{-(\tilde{\lambda}/\theta)Y}), Y^{\beta})$$

$$+2\left(\frac{\tilde{\lambda}}{\theta}\right)\frac{(\beta-1)}{\beta}\left[\psi\left(\frac{1}{\beta}+1\right)\Gamma\left(\frac{1}{\beta}+1\right) - \Gamma\left(\frac{1}{\beta}+1\right)\psi(1)\right]$$

$$-2\left(\frac{\tilde{\lambda}}{\theta}\right)\left[\Gamma\left(\frac{(\beta+1)}{\beta}+1\right) - \Gamma\left(\frac{1}{\beta}+1\right)\right]$$

$$+2\,\frac{(\beta-1)}{\beta}[\psi(2)-\psi(1)]. \tag{3.7}$$

Case 2: The data are coming from a GE distribution and the alternative is a Weibull distribution.

Let us assume that a sample $X_1,\ldots,X_n$ of size n is obtained from $GE(\alpha,\lambda)$ and the alternative is $WE(\beta,\theta)$. We denote $\hat{\alpha}$, $\hat{\lambda}$, $\hat{\beta}$ and $\hat{\theta}$ as the MLEs of α, λ, β and θ, respectively. In this case we have the following lemma.

Lemma 2. *Under the assumption that the data are from a GE distribution and as* $n \to \infty$, *we have*

(i) $\hat{\alpha} \to \alpha$ *a.s.*, $\hat{\lambda} \to \lambda$ *a.s.*, *where*

$$E_{GE}[\ln(f_{GE}(X;\alpha,\lambda))] = \max_{\bar{\alpha},\bar{\lambda}} E_{GE}[\ln(f_{GE}(X;\bar{\alpha},\bar{\lambda}))].$$

(ii) $\hat{\beta} \to \tilde{\beta}$ *a.s.*, $\hat{\theta} \to \tilde{\theta}$ *a.s.*, *where*

$$E_{GE}[\ln(f_{WE}(X;\tilde{\beta},\tilde{\theta}))] = \max_{\beta,\theta} E_{GE}[\ln(f_{WE}(X;\beta,\theta))].$$

Note that here also $\tilde{\beta}$ and $\tilde{\theta}$ may depend on α and λ but we do not make it explicit for brevity. Let us denote $T_* = \ln(L_{GE}(\alpha,\lambda)/L_{WE}(\tilde{\beta},\tilde{\theta}))$.

(iii) $n^{-1/2}[T - E_{GE}(T)]$ is asymptotically equivalent to $n^{-1/2}[T_* - E_{GE}(T_*)]$.

Theorem 2. *Under the assumption that the data are from a GE distribution*, T *is approximately normally distributed with mean* $E_{GE}(T)$ *and variance* $V_{GE}(T)$.

Now to obtain $\tilde{\beta}$ and $\tilde{\theta}$, let us define

$$\begin{aligned} h(\beta,\theta) &= E_{GE}[\ln(f_{WE}(X;\beta,\theta))] \\ &= E_{GE}[\ln(\beta) + \beta\ln(\theta) + (\beta-1)\ln(X) - (X\theta)^{\beta}] \\ &= \ln(\beta) + \beta\left[\ln\left(\frac{\theta}{\lambda}\right) + E_{GE}(\ln(Z))\right] - E_{GE}(\ln(X)) - \alpha w\left(\beta, \frac{\theta}{\lambda}\right), \end{aligned} \tag{3.8}$$

here

$$w(x,y) = y^x \int_0^\infty u^x(1-\mathrm{e}^{-u})^{\alpha-1}\mathrm{e}^{-u}\,\mathrm{d}u, \tag{3.9}$$

X follows $GE(\alpha,\lambda)$ and Z follows $GE(\alpha,1)$. Therefore, $\tilde{\beta}$ and $\tilde{\theta}$ can be obtained as solutions of

$$\frac{1}{\tilde{\beta}} + \ln\left(\frac{\tilde{\theta}}{\lambda}\right) + E_{GE}(\ln(Z)) - \alpha w_1\left(\tilde{\beta}, \frac{\tilde{\theta}}{\lambda}\right) = 0 \tag{3.10}$$

and

$$\frac{\tilde{\beta}}{\tilde{\theta}} - \frac{\alpha}{\lambda} w_2\left(\tilde{\beta}, \frac{\tilde{\theta}}{\lambda}\right) = 0. \tag{3.11}$$

Here $w_1(x,y)$ and $w_2(x,y)$ are the derivatives of $w(x,y)$ with respect to x and y, respectively, i.e.,

$$w_1(x,y) = \int_0^\infty (yu)^x(\ln(y)+\ln(u))(1-\mathrm{e}^{-u})^{\alpha-1}\mathrm{e}^{-u}\,\mathrm{d}u$$

and

$$w_2(x,y) = xy^{x-1}\int_0^\infty u^x(1-\mathrm{e}^{-u})^{\alpha-1}\mathrm{e}^{-u}\,\mathrm{d}u.$$

Note that as before $\tilde{\theta}/\lambda$ and $\tilde{\beta}$ both are functions of α only. Now we provide the expressions for $E_{GE}(T)$ and $V_{GE}(T)$. Similarly as before, we observe that $\lim_{n\to\infty} E_{GE}(T)/n$ and $\lim_{n\to\infty} V_{GE}(T)/n$ exist. Suppose we denote $\lim_{n\to\infty} E_{GE}(T)/n = AM_{GE}(\alpha)$ and $\lim_{n\to\infty} V_{GE}(T)/n = AV_{GE}(\alpha)$ then for large n,

$$\frac{E_{GE}(T)}{n} \approx AM_{GE}(\alpha) = E_{GE}[\ln(f_{GE}(X;\alpha,\lambda)) - \ln(f_{WE}(X;\tilde{\beta},\tilde{\theta}))]$$
$$= \ln(\alpha) - \frac{(\alpha-1)}{\alpha} - (\psi(\alpha+1) - \psi(1)) - \ln(\tilde{\beta}) - (\tilde{\beta}-1)E_{GE}(\ln(Z))$$
$$- \tilde{\beta}\ln\left(\frac{\tilde{\theta}}{\lambda}\right) + \left(\frac{\tilde{\theta}}{\lambda}\right)^{\tilde{\beta}} E_{GE}(Z^{\tilde{\beta}}), \tag{3.12}$$

here X and Z are same as defined before. Also

$$\frac{V_{GE}(T)}{n} \approx AV_{GE}(\alpha) = V_{GE}[\ln(f_{GE}(X;\alpha,\lambda)) - \ln(f_{WE}(X;\tilde{\beta},\tilde{\theta}))]$$
$$= V_{GE}\left[(\alpha-1)\ln(1-e^{-Z}) - Z - (\tilde{\beta}-1)\ln(Z) + \left(\frac{\tilde{\theta}}{\lambda}\right)^{\tilde{\beta}} Z^{\tilde{\beta}}\right]$$
$$= \frac{(\alpha-1)^2}{\alpha^2} + (\psi'(1) - \psi'(\alpha+1)) + (\tilde{\beta}-1)^2 V_{GE}(\ln(Z))$$
$$+ \left(\frac{\tilde{\theta}}{\lambda}\right)^{2\tilde{\beta}} V_{GE}(Z^{\tilde{\beta}}) - 2(\alpha-1)cov_{GE}(\ln(1-e^{-Z}), Z)$$
$$+ 2(\tilde{\beta}-1)(\alpha-1)cov_{GE}(\ln(1-e^{-Z}), \ln(Z))$$
$$+ 2(\alpha-1)\left(\frac{\tilde{\theta}}{\lambda}\right)^{\tilde{\beta}} cov_{GE}(\ln(1-e^{-Z}), Z^{\tilde{\beta}}) + 2(\tilde{\beta}-1)cov_{GE}(Z, \ln(Z))$$
$$- 2\left(\frac{\tilde{\theta}}{\lambda}\right)^{\tilde{\beta}} cov_{GE}(Z, Z^{\tilde{\beta}}) - 2(\tilde{\beta}-1)\left(\frac{\tilde{\theta}}{\lambda}\right)^{\tilde{\beta}} cov_{GE}(\ln(Z), Z^{\tilde{\beta}}). \tag{3.13}$$

Note that $\tilde{\alpha}$, $\tilde{\lambda}$, $AM_{WE}(\beta)$, $AV_{WE}(\beta)$, $\tilde{\beta}$, $\tilde{\theta}$, $AM_{GE}(\alpha)$ and $AV_{GE}(\alpha)$ are quite difficult to compute numerically. We present $\tilde{\alpha}$, $\tilde{\lambda}$ as obtained from (3.3), (3.4) and also $AM_{WE}(\beta)$ and $AV_{WE}(\beta)$ for different values of β (note that they are independent of θ) in Table 1. We also present $\tilde{\beta}$, $\tilde{\theta}$ as obtained from (3.10) and (3.11) and also $AM_{GE}(\alpha)$ and $AV_{GE}(\alpha)$ for different values of α in Table 2 for convenience.

Table 1
Different values of $AM_{WE}(\beta)$, $AV_{WE}(\beta)$, $\tilde{\alpha}$ and $\tilde{\lambda}$ for different β

β	$AM_{WE}(\beta)$	$AV_{WE}(\beta)$	$\tilde{\alpha}$	$\tilde{\lambda}$
0.6	−0.0192	0.0950	0.474	0.410
0.8	−0.0032	0.0090	0.722	0.721
1.2	−0.0021	0.0061	1.390	1.307
1.4	−0.0072	0.0165	1.823	1.565
1.6	−0.0137	0.0296	2.334	1.802
1.8	−0.0209	0.0439	2.885	2.006
2.0	−0.0284	0.0587	3.639	2.239

Table 2
Different values of $AM_{GE}(\alpha)$, $AV_{GE}(\alpha)$, $\tilde{\beta}$ and $\tilde{\theta}$ for different α

α	$AM_{GE}(\alpha)$	$AV_{GE}(\alpha)$	$\tilde{\beta}$	$\tilde{\theta}$
0.5	0.0117	0.0248	0.649	2.243
1.5	0.0034	0.0095	1.257	0.735
2.0	0.0093	0.0236	1.440	0.609
2.5	0.0153	0.0378	1.585	0.537
3.0	0.0209	0.0511	1.706	0.488

4. Determination of sample size

In this section, we propose a method to determine the minimum sample size needed to discriminate between Weibull and GE distributions, for a given user specified probability of correct selection (PCS). There are several ways to measure the closeness or the distance between two distribution functions, for example, the Kolmogorov–Smirnov (K–S) distance or Hellinger distance, etc. Intuitively, it is clear that if two distributions are very close, one needs a very large sample size to discriminate between them for a given probability of correct selection. On the other hand if two distribution functions are quite different, then one may not need very large sample size to discriminate between two distribution functions. It is also true that if two distribution functions are very close to each other, then one may not need to differentiate the two distributions from a practical point of view. Therefore, it is expected that the user will specify before hand the PCS and also the tolerance limit in terms of the distance between two distribution functions. The tolerance limit simply indicates that the user does not want to make the distinction between two distribution functions if their distance is less than the tolerance limit. Based on the PCS and the tolerance limit, the required minimum sample size can be determined. In this paper we use K–S distance to discriminate between two distribution functions but similar methodology can be developed using the Hellinger distance also, which is not pursued here.

We observed in Section 3 that the RML statistics follow normal distribution approximately for large n. Now it will be used with the help of K–S distance to determine

Table 3
The minimum sample size $n = z_{0.70}^2 AV_{WE}(\beta)/(AM_{WE}(\beta))^2$, using (4.5), for $p^* = 0.7$ and when the null distribution is Weibull is presented. The K–S distance between $WE\ (\beta, 1)$ and $GE(\tilde{\alpha},\ \tilde{\lambda})$ for different values of β is reported

$\beta \rightarrow$	0.6	0.8	1.2	1.4	1.6	1.8	2.0
$n \rightarrow$	71	242	381	88	44	28	20
K–S	0.036	0.016	0.013	0.022	0.029	0.036	0.039

Table 4
The minimum sample size $n = z_{0.70}^2 AV_{GE}(\alpha)/(AM_{GE}(\alpha))^2$, using (4.6), for $p^* = 0.7$ and when the null distribution is GE is presented. The K–S distance between $GE\ (\alpha, 1)$ and $WE\ (\tilde{\beta}, \tilde{\theta})$ for different values of α is reported

$\alpha \rightarrow$	0.5	1.5	2.0	2.5	3.0
$n \rightarrow$	50	431	75	45	32
K–S	0.032	0.010	0.019	0.025	0.030

the required sample size n such that the PCS achieves a certain protection level p^* for a given tolerance level D^*. We explain the procedure assuming Case 1, Case 2 follows exactly along the same line.

Since T is asymptotically normally distributed with mean $E_{WE}(T)$ and variance $V_{WE}(T)$, therefore the PCS is

$$PCS(\beta) = P[T < 0|\beta] \approx \Phi\left(\frac{-E_{WE}(T)}{\sqrt{V_{WE}(T)}}\right) = \Phi\left(\frac{-n \times AM_{WE}(\beta)}{\sqrt{n \times AV_{WE}(\beta)}}\right). \tag{4.3}$$

Here Φ is the distribution function of the standard normal random variable. $AM_{WE}(\beta)$ and $AV_{WE}(\beta)$ are same as defined in (3.6) and (3.7), respectively. Now to determine the sample size needed to achieve at least a p^* protection level, equate

$$\Phi\left(\frac{-n \times AM_{WE}(\beta)}{\sqrt{n \times AV_{WE}(\beta)}}\right) = p^* \tag{4.4}$$

and solve for n. It provides

$$n = \frac{z_{p^*}^2 AV_{WE}(\beta)}{(AM_{WE}(\beta))^2}. \tag{4.5}$$

Here z_{p^*} is the $100p^*$ percentile point of a standard normal distribution. For $p^* = 0.7$ and for different β, the values of n are reported in Table 3. Similarly for Case 2, we need

$$n = \frac{z_{p^*}^2 AV_{GE}(\alpha)}{(AM_{GE}(\alpha))^2}. \tag{4.6}$$

Here $AM_{GE}(\alpha)$ and $AV_{GE}(\alpha)$ are as defined in (3.12) and (3.13), respectively. We report n, with the help of Table 2 for different values of α when $p^* = 0.7$ in Table 4.

From Tables 3 and 4 it is immediate that as α and β move away from 1, for a given PCS, the required sample size decreases as expected. From (4.5) and (4.6) it is clear that if one knows the range of the shape parameter of the null distribution then the minimum sample size can be obtained using (4.5) or (4.6) and using the fact that n increases as the shape parameter moves away from 1. But unfortunately in practice it may be completely unknown. Therefore, to have some idea of the shape parameter of the null distribution we make the following assumptions. It is assumed that the experimenter would like to choose the minimum sample size needed for a given protection level when the distance between two distribution functions is greater than a pre-specified tolerance level. The distance between two distribution functions is defined by the K–S distance. The K–S distance between two distribution functions, say $F(x)$ and $G(x)$ is defined as

$$\sup_x |F(x) - G(x)|. \tag{4.7}$$

We report K–S distance between $WE(\beta, 1)$ and $GE(\tilde{\alpha}, \tilde{\lambda})$ for different values of β in Table 3. Here $\tilde{\alpha}$ and $\tilde{\lambda}$ are same as defined in (3.3) and (3.4) and they have been reported in Table 1. Similarly, K–S distance between $GE(\alpha, 1)$ and $WE(\tilde{\beta}, \tilde{\theta})$ for different values of α is reported in Table 4. Here $\tilde{\beta}$ and $\tilde{\theta}$ are same as defined in (3.10) and (3.11) and they have been reported in Table 2. From Tables 3 and 4 it is clear that in both cases K–S distance between two distribution functions increases as the shape parameter moves away from 1.

Now we explain how we can determine the minimum sample size required to discriminate between Weibull and GE distributions for a user specified protection level and for a given tolerance level between the two distribution functions. Suppose the protection level is $p^* = 0.7$ and the tolerance level is given in terms of K–S distance as $D^* = 0.036$. Here tolerance level $D^* = 0.036$ means that the practitioner wants to discriminate between a Weibull distribution function and a GE distribution function only when their K–S distance is more than 0.036. From Table 3, it is clear that for case 1, K–S distance will be more than 0.036 if $\beta \leqslant 0.6$ or $\beta \geqslant 1.8$. Similarly from Table 4, it is clear that for case 2, K–S distance will be more than 0.036 if $\alpha < 0.5$ or $\alpha > 3.0$. Therefore, if the null distribution is Weibull, then for the tolerance level $D^* = 0.036$, one needs $n = \max(71, 28) = 71$ to meet the PCS, $p^* = 0.7$. Similarly if the null distribution is GE then one needs at most $n = \max(32, 50) = 50$ to meet the above protection level $p^* = 0.7$ and when the tolerance level $D^* = 0.036$. Therefore, for the given tolerance level 0.036 one needs $\max(50, 71) = 71$ to meet the protection level $p^* = 0.7$ simultaneously for both the cases.

5. Numerical experiments

In this section we perform some numerical experiments to observe how these asymptotic results derived in Section 3 work for finite sample sizes. All computations are performed at the Indian Institute of Technology Kanpur, using Pentium-II processor. We use the random deviate generator of Press et al. (1993) and all the programs are written in FORTRAN. They can be obtained from the authors on request. We

Table 5
The PCS based on Monte Carlo Simulations and also based on asymptotic results when the null distribution is Weibull. The element in the first row in each box represents the results based on Monte Carlo Simulations (10,000 replications) and the number in bracket immediately below represents the result obtained by using asymptotic results

$\beta \downarrow n \rightarrow$	20	40	60	80	100
0.6	0.58 (0.61)	0.66 (0.65)	0.70 (0.68)	0.73 (0.71)	0.75 (0.75)
0.8	0.53 (0.56)	0.56 (0.58)	0.59 (0.60)	0.61 (0.62)	0.62 (0.63)
1.2	0.56 (0.55)	0.58 (0.57)	0.59 (0.58)	0.60 (0.59)	0.61 (0.60)
1.4	0.63 (0.60)	0.66 (0.64)	0.67 (0.67)	0.71 (0.69)	0.72 (0.71)
1.6	0.66 (0.64)	0.71 (0.69)	0.74 (0.73)	0.77 (0.76)	0.82 (0.82)
1.8	0.70 (0.67)	0.75 (0.74)	0.78 (0.78)	0.82 (0.81)	0.84 (0.84)
2.0	0.71 (0.70)	0.79 (0.77)	0.82 (0.82)	0.86 (0.85)	0.88 (0.88)

compute the PCS based on simulations and we also compute it based on asymptotic results derived in Section 3. We consider different sample sizes and also different shape parameters of the null distributions. The details are explained below.

First we consider the case when the null distribution is Weibull and the alternative is GE. In this case we consider $n = 20, 40, 60, 80, 100$ and $\beta = 0.6, 0.8, 1.2, 1.4, 1.6, 1.8$ and 2.0. For a fixed β and n we generate a random sample of size n from $WE(\beta, 1)$, we finally compute T as defined in (2.3) and check whether T is positive or negative. We replicate the process 10,000 times and obtain an estimate of the PCS. We also compute the PCSs by using these asymptotic results as given in (4.3). The results are reported in Table 5. Similarly, we obtain the results when the null distribution is GE and the alternative is Weibull. In this case we consider the same set of n and $\alpha = 0.5, 1.5, 2.0, 2.5, 3.0$. The results are reported in Table 6. In each box the first row represents the results obtained by using Monte Carlo simulations and the second row represents the results obtained by using the asymptotic theory.

It is quite clear from Tables 5 and 6 that as the sample size increases the PCS increases as expected. It is also clear that as the shape parameter moves away from 1, the PCS increases. Even when the sample size is 20, asymptotic results work quite well for both the cases for all possible parameter ranges. From the simulation study

Table 6
The PCS based on Monte Carlo Simulations and also based on asymptotic results when the null distribution is GE. The element in the first row in each box represents the results based on Monte Carlo Simulations (10,000 replications) and the number in bracket immediately below represents the result obtained by using asymptotic results

$\alpha \downarrow n \rightarrow$	20	40	60	80	100
0.5	0.66 (0.63)	0.70 (0.68)	0.71 (0.72)	0.74 (0.75)	0.76 (0.77)
1.5	0.53 (0.56)	0.57 (0.59)	0.60 (0.61)	0.62 (0.62)	0.64 (0.64)
2.0	0.57 (0.60)	0.63 (0.65)	0.68 (0.68)	0.70 (0.71)	0.73 (0.73)
2.5	0.62 (0.64)	0.68 (0.69)	0.72 (0.73)	0.77 (0.76)	0.79 (0.79)
3.0	0.63 (0.66)	0.72 (0.72)	0.75 (0.76)	0.81 (0.80)	0.82 (0.82)

it is recommended that asymptotic results can be used quite effectively even when the sample size is as small as 20 for all possible choices of the shape parameters.

6. Data analysis

In this section we analyze two data sets and use our method to discriminate between two populations.

Data Set 1: The first data set is as follows; (Lawless, 1982, p. 228). The data given 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, 68.44, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, 173.40.

When we use a GE model, the MLEs of the different parameters are $\hat{\alpha}=5.2589$, $\hat{\lambda}=0.0314$ and $\ln(L_{GE}(\hat{\alpha},\hat{\lambda}))=-112.9763$. Similarly, if we use a Weibull model, the MLEs of the different parameters are $\hat{\beta}=2.1050$, $\hat{\theta}=0.0122$ and $\ln(L_{WE}(\hat{\beta},\hat{\theta}))=-113.6887$. Therefore, $T=-112.9763+113.6887=0.7124>0$, which indicates to choose the GE model. In Fig. 2, we provide the histogram of the data and the two fitted densities. From the fitted density functions it appears that generalized exponential distribution provides a *better fit* than Weibull distribution in this case.

If the distribution were GE with $\alpha=5.2589=\hat{\alpha}$ and $\lambda=0.0314=\hat{\lambda}$, then we compute PCS by computer simulations (based on 10,000 replications) similarly as in Section 5 and we obtain $PCS=0.7059$. It implies that PCS will be more than 70%. On the other hand if the choice of GE were wrong and the original distribution was Weibull

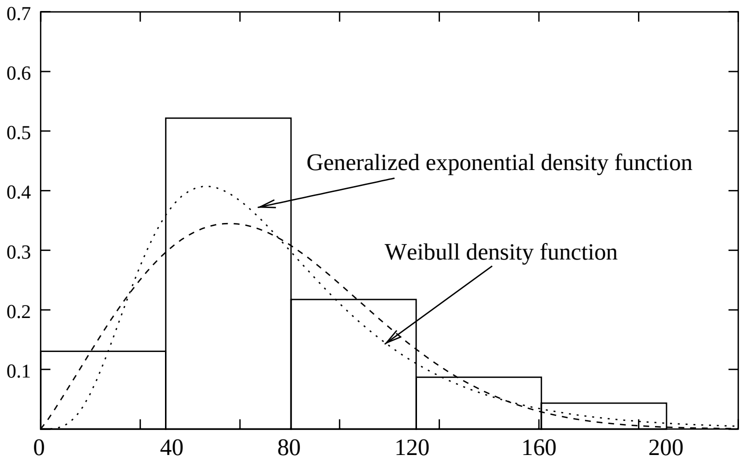

Fig. 2. The histogram of the data set 1 and the fitted density functions.

with shape parameter $\beta = 2.1050 = \hat{\beta}$ and scale parameter $\theta = 0.0122 = \hat{\theta}$, then similarly as before based on 10,000 replications we obtain $PCS = 0.7121$, yielding an estimated risk less than approximately 30% to choose the wrong model. Now we compute the PCSs based on large sample approximations. Assuming that the data are coming from GE, we obtain $AM_{GE}(5.2589) = 0.0300$ and $AV_{GE}(5.2589) = 0.0763$, it implies from (3.12) and (3.13) that $E_{GE}(T) \approx 0.6910$ and $V_{GE}(T) \approx =1.7554$. Therefore, assuming that the data are from GE, T is approximately normally distributed with mean $= 0.6910$, variance $= 1.7554$ and $PCS = 1 - \Phi(-0.5215) = \Phi(0.5215) \approx 0.70$, which is almost equal to the above simulation result. Moreover under the same assumption that the data are from a GE, we obtain the approximate p value of the observed $T = 0.7124$ is 0.49. Similarly, assuming that the data are coming from a Weibull, we compute $AM_{WE}(2.1050) = -0.0297$ and $AV_{WE}(2.1050) = 0.0646$. Using, (3.6) and (3.7) we have $E_{WE}(T) \approx -0.6794$ and $V_{WE}(T) \approx =1.4860$. Therefore, assuming that the data are from a Weibull distribution the probability of miss classification $(1 - PCS)$ is $1 - \Phi(0.5573) \approx 0.28$, which is also very close to the simulated results. In this case the approximate p value of the observed T is 0.13. Comparing the two p values also, we would like to say that the data are coming from a GE distribution and the probability correct selection is at least $\min(0.70, 0.72) = 0.70$ in this case.

Data Set 2: The second data set (Linhart and Zucchini, 1986, p. 69) represents the failure times of the air conditioning system of an airplane: 23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 95.

Under the assumption that the data are from a GE distribution, the MLEs of the different parameters are $\hat{\alpha} = 0.8130$ and $\hat{\lambda} = 0.0145$, also $\ln(L_{GE}(\hat{\alpha}, \hat{\lambda})) = -152.264$. Similarly under the assumption that the data are from a Weibull distribution, the MLEs of the different Weibull parameters are $\hat{\beta} = 0.8554$ and $\hat{\theta} = 0.0183$. We

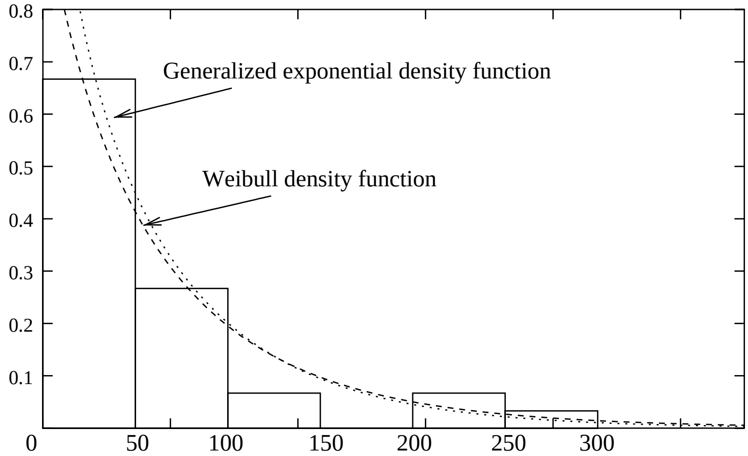

Fig. 3. The histogram of the data set 2 and the fitted density functions.

provide the histogram of the data set 2 and the two fitted densities in Fig. 3. From the figure it appears that both the fits are quite close to each other. In this case $\ln(L_{WE}(\hat{\beta},\hat{\theta})) = -152.007$ and that provides $T = -152.264 + 152.007 = -0.257 < 0$. Therefore, we choose the Weibull model in this case. Under the assumptions that the data are from $WE(0.8554, 0.0183)$, we obtain $PCS = 0.5224$ based on simulation results. Moreover, under the assumptions that the data are from $GE(0.8130, 0.0145)$, $PCS = 0.5380$. We obtain $AM_{WE}(0.8554) = -0.0055$, $AV_{WE}(0.8554) = 0.0184$, $AM_{GE}(0.8130) = 0.0004$, $AV_{GE}(0.8130) = 0.0061$. From (3.12), (3.13), (3.6) and (3.7) we have $E_{WE}(T) \approx -0.1658$, $V_{WE}(T) \approx =0.5515$, $E_{GE}(T) \approx =0.0013$ and $V_{GE} \approx =0.1823$. Therefore using large sample approximation, under the assumption that the data are coming from a Weibull distribution, $PCS = \Phi(0.2025) \approx 0.5871$ and using simulations we obtain $PCS = 0.5224$. The approximate p value of the observed T is 0.43. Similarly, under the assumption that the data are coming from a GE, using the large sample approximation we obtain $PCS = \Phi(0.0030) \approx 0.50$ and using simulations $PCS = 0.5345$. The corresponding approximate p value of the observed T is 0.15. Therefore, the p value also suggests to choose a Weibull model for data set 2, but interestingly PCS is only around 50% in this case.

From the two examples it is clear that not only the sample size but the model parameters also play a very important role in choosing between two overlapping distributions. For comparison purposes we compute K–S distances in both cases and plot the two fitted distribution functions for data set 1 and data set 2 in Figs. 4 and 5, respectively. It is observed that for data set 1, the K–S distance between the two fitted distributions is 0.039 and for data set 2, the corresponding K–S distance is 0.022. For data set 2, it is very clear that two fitted distribution functions are very close to each other and therefore discriminating between them is very difficult. It also shows that the distance between the two fitted distributions is very important in discriminating two overlapping families.

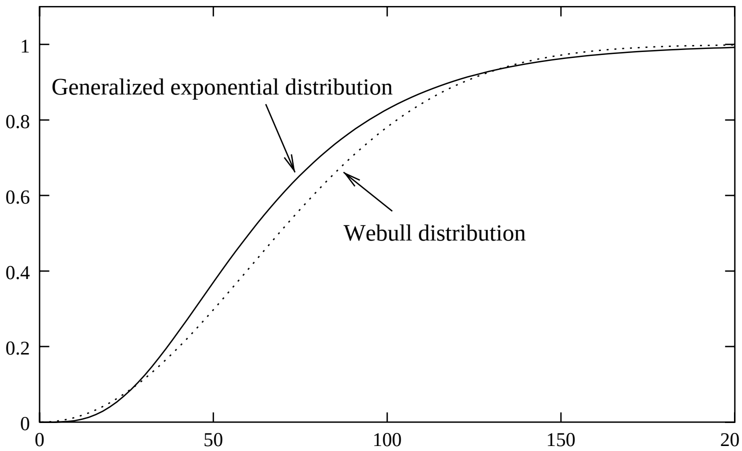

Fig. 4. The two fitted distribution functions for data set 1.

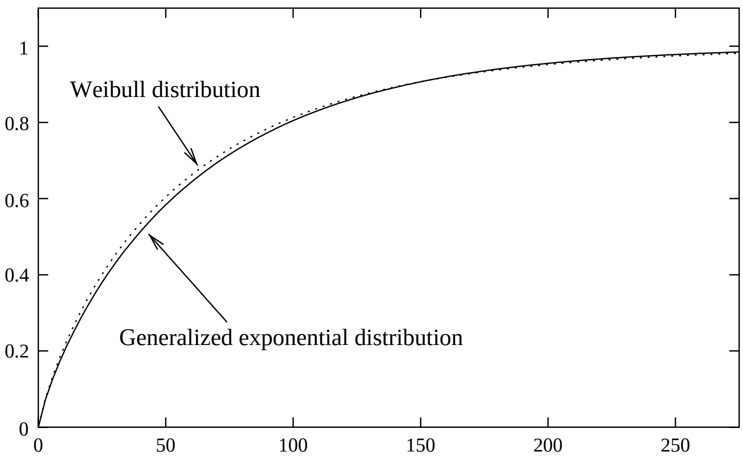

Fig. 5. The two fitted distribution functions for data set 2.

7. Conclusions

In this paper we consider the problem of discriminating between two overlapping families of distribution functions, namely Weibull and GE families. We consider the statistic based on the RML and obtain asymptotic distributions of the test statistics under null hypotheses. Using a Monte Carlo simulation we compare the probability of correct selection with these asymptotic results and it is observed that even when the sample size is very small these asymptotic results work quite well for a wide range of the parameter space. Therefore, these asymptotic results can be used to estimate the PCS. We use these asymptotic results to calculate the minimum sample size needed to discriminate between two distribution functions for a user specified probability of correct selection. We use the concept of tolerance level based on the distance between two distribution functions. For a particular D^* tolerance level the required minimum sample size is obtained for a given user specified protection level. Two small tables are provided for the protection level 0.70 but for the other protection level the tables can be easily used as follows. For example if we need the protection level $p^* = 0.8$, then all the entries corresponding to the row of n, will be multiplied by $z_{0.8}^2/z_{0.7}^2$, because of (4.6). Therefore, Tables 3 and 4 can be used for any given protection level. We have just presented two small tables for illustration purposes, extensive tables for different values of β and α can be obtained from the authors on request. It may be mentioned that similar methodologies can be developed in discriminating between GE and gamma distributions or between GE and log-normal distributions. More work is needed in that direction.

Acknowledgements

The authors would like to thank two referees and one associate editor for very constructive suggestions. The authors would also like to thank the co-editor Professor Dr. Erricos John Kontoghiorghes for encouragements.

References

Bain, L.J., Engelhardt, M., 1980. Probability of correct selection of Weibull versus gamma based on likelihood ratio. Comm. Statist. Ser. A. 9, 375–381.

Balasooriya, C.P., Abeysinghe, T., 1994. Selecting between gamma and Weibull distributions: approach based on prediction of order statistics. J. Appl. Statist. 21 (3), 17–27.

Cox, D.R., 1961. Tests of separate families of hypotheses. Proceedings of the Fourth Berkeley Symposium in Mathematical Statistics and Probability, Berkeley, University of California Press, pp. 105–123.

Cox, D.R., 1962. Further results on tests of separate families of hypotheses. J. Roy. Statist. Soc. Ser. B 24, 406–424.

Dumonceaux, R., Antle, C.E., 1973. Discrimination between the log-normal and the Weibull distributions. Technometrics 15 (4), 923–926.

Dumonceaux, R., Antle, C.E., Haas, G., 1973. Likelihood ratio test for discrimination between two models with unknown location and scale parameters. Technometrics 15 (1), 19–27.

Gupta, R.D., Kundu, D., 1999. Generalized exponential distributions. Austral. N. Z. J. Statist. 41 (2), 173–188.

Gupta, R.D., Kundu, D., 2001a. Exponentiated exponential distribution: an alternative to gamma and Weibull distributions. Biometrical J. 43 (1), 117–130.

Gupta, R.D., Kundu, D., 2001b. Generalized exponential distributions: different methods of estimations. J. Statist. Comput. Simulations 69 (4), 315–338.

Gupta, R.D., Kundu, D., Manglick, A., 2001. Probability of correct selection of Gamma versus GE or Weibull versus GE based on likelihood ratio test. Technical Report, The University of New Brunswick, Saint John.

Lawless, J.F., 1982. Statistical Models and Methods for Lifetime Data, Wiley, New York.
Linhart, H., Zucchini, W., 1986. Model Selection. Wiley, New York.
Press, et al., 1993. Numerical Recipes in FORTRAN, Cambridge University Press, Cambridge.
Quesenberry, C.P., Kent, J., 1982. Selecting among probability distributions used in reliability. Technometrics 24 (1), 59–65.
White, H., 1982. Regularity conditions for Cox's test of non-nested hypotheses. J. Econometrics 19, 301–318.