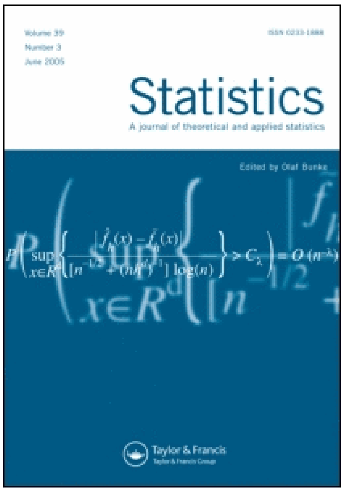

Statistics
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713682269

A new class of weighted exponential distributions

Rameshwar D. Gupta [a]; Debasis Kundu [b]
[a] Department of Computer Science and Applied Statistics, The University of New Brunswick, Saint John, Canada [b] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur, India

First Published:December2009

To cite this Article Gupta, Rameshwar D. and Kundu, Debasis(2009)'A new class of weighted exponential distributions',Statistics,43:6,621 — 634

To link to this Article: DOI: 10.1080/02331880802605346

URL: http://dx.doi.org/10.1080/02331880802605346

Statistics, Vol. 43, No. 6, December 2009, 621–634

A new class of weighted exponential distributions

Rameshwar D. Gupta[a] and Debasis Kundu[b]*

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

(*Received 7 September 2007; final version received 7 October 2008*)

Introducing a shape parameter to an exponential model is nothing new. There are many ways to introduce a shape parameter to an exponential distribution. The different methods may result in variety of weighted exponential (WE) distributions. In this article, we have introduced a shape parameter to an exponential model using the idea of Azzalini, which results in a new class of WE distributions. This new WE model has the probability density function (PDF) whose shape is very close to the shape of the PDFS of Weibull, gamma or generalized exponential distributions. Therefore, this model can be used as an alternative to any of these distributions. It is observed that this model can also be obtained as a hidden truncation model. Different properties of this new model have been discussed and compared with the corresponding properties of well-known distributions. Two data sets have been analysed for illustrative purposes and it is observed that in both the cases it fits better than Weibull, gamma or generalized exponential distributions.

Keywords: hazard function; log-concavity; RR_2 ordering; maximum likelihood estimators; moment estimators; Fisher information; moment generating function

1. Introduction

Different methods may be used to introduce a shape parameter to an exponential model and they may result in a variety of weighted exponential (WE) distributions. For example, the gamma distribution and the generalized exponential distribution are different weighted versions of the exponential distribution. In this article we have used the idea of Azzalini [1] to introduce a shape parameter to an exponential distribution which results in a new class of WE distributions.

Azzalini [1] first introduced the skew-normal distribution to incorporate a shape/skewness parameter to a normal distribution. Since then extensive work has been done to introduce a skewness parameter to a symmetric distribution. For example, skew-t, skew-Cauchy, skew-Laplace, skew-logistic and in general skew-symmetric distribution have been defined and several properties and their inference procedures have been discussed, see for example, Arnold and Beaver [2], Gupta and Kundu [3] and the recent monograph by Genton [4]. Arnold and Beaver [5] provided a nice interpretation of Azzalini's skew-normal model as a hidden truncation model, although the same interpretation may not be true for other skewed distributions.

*Corresponding author. Email: kundu@iitk.ac.in

ISSN 0233-1888 print/ISSN 1029-4910 online

DOI: 10.1080/02331880802605346
http://www.informaworld.com

Interestingly, although Azzalini's method has been used extensively for several symmetric distributions, an attempt has not been made to use the idea for non-symmetric distribution. In this article, it is observed that if we apply Azzalini's method to the exponential distribution, then it produces a new class of WE distributions that has a shape parameter. From now on we denote a member of this new class of weighted distributions as WE distribution. This shape parameter governs the shape of the probability density function (PDF) of the WE distribution. It is observed that the shape of the PDF of WE is very similar to the other generalization of the exponential distribution, for example gamma, Weibull or generalized exponential distribution.

The main aim of this article is to introduce this model and study its different properties. It is observed that although this model has been obtained as a WE distribution, it has several other interpretations. It can be observed as a hidden truncation model, as was observed in the case of the skew-normal distribution by Arnold and Beaver [5]. This model can also be obtained by beta transformation as it was proposed by Jones [6]. Moreover, this model can also be represented as a sum of two independent but non-identical exponential distributions. These different representations have been used to derive different properties of the WE distribution and also to generate WE random deviate.

It is observed that the WE distribution function has a compact form and all the moments can be computed explicitly. Therefore, mean, variance, skewness, kurtosis, coefficient of variation, hazard function (HF), and mean residual lifetime, all can be computed explicitly. Interestingly, the distribution of the sum of independent n WE random variables can also be obtained. The maximum likelihood estimators (MLEs) of the two unknown parameters of the WE distribution cannot be obtained in explicit form, but they can be obtained as a solution of a non-linear fixed-point type equation. The moment estimators (MEs) of the unknown parameters can be obtained explicitly. The asymptotic distributions of the MLEs are provided and they can be used for constructing the asymptotic confidence intervals or for testing purposes.

The rest of the article is organized as follows. In Section 2, we provide the definition and different interpretations. Different properties are discussed in Section 3. Statistical inferences of the unknown parameters are carried out in Section 4. Two data sets have been analysed in Section 5 and finally we conclude the article in Section 6.

2. Definition, interpretations and generation

In this section, first we provide the definition of the WE model and then provide four different interpretations. It may be observed that different interpretations can be used very easily to generate WE random deviates.

DEFINITION *The random variable X is said to have* WE *distribution, with the shape and scale parameters as $\alpha > 0$ and $\lambda > 0$, respectively, if the PDF of X is*

$$f_X(x;\alpha,\lambda) = \frac{\alpha+1}{\alpha}\lambda e^{-\lambda x}(1-e^{-\alpha\lambda x}); \quad x > 0 \tag{1}$$

and 0 otherwise. We will denote it as $\mathrm{WE}(\alpha,\lambda)$. *Note that in the model* (1) *the location parameter can be easily incorporated.*

INTERPRETATION 1 *This model can be obtained from two independent identically distributed* (i.i.d.) *random variables exactly the same way Azzalini* [1] *obtained the skew-normal distribution from two* i.i.d. *normal distributions. Suppose X_1 and X_2 are two* i.i.d. *random variables, with the PDF $f_Y(\cdot)$ and cumulative distribution function (CDF) $F_Y(\cdot)$, then for any $\alpha > 0$, consider a new* *random variable* $X = X_1$ *given that* $\alpha X_1 > X_2$. *Note that the PDF of the new random variable* X *is*

$$f_X(x) = \frac{1}{P(\alpha X_1 > X_2)} f_Y(x) F_Y(\alpha x); \quad x > 0. \tag{2}$$

Now Equation (1) *can be obtained from Equation* (2) *by replacing* $f_Y(x) = \mathrm{e}^{-\lambda x}$ *and* $F_Y(x) = 1 - \mathrm{e}^{-\lambda x}$.

INTERPRETATION 2 *This model can be obtained as a hidden truncation model as it was observed by Arnold and Beaver* [5] *in case of skew-normal distribution. Suppose Z and Y are two dependent random variables with the joint PDF as given below for* $\lambda > 0$;

$$f_{Z,Y}(z, y) = \lambda^2 z \mathrm{e}^{-\lambda z(1+y)}; \quad z > 0, y > 0. \tag{3}$$

The contour and the joint PDF of Z and Y are plotted in Figure 1. *Consider a new random variable* $X = Z$ *given that* $Y \le \alpha$. *It is easily observed that the PDF of X is Equation* (1). *Therefore*, *it can be interpreted as the hidden truncation model as follows. Suppose* Z *and* Y *are two correlated random variables with the joint PDF* (3) *and we can not observe* Y, *we only can observe* Z, *if* $Y \le \alpha$, *then the observed sample can be regarded as drawn from a distribution with the PDF* (1).

INTERPRETATION 3 *Recently Jones* [6], *see also Nadarajah and Kotz* [7], *obtained a class of distribution functions from a Beta*(*a*, *b*) *distribution as follows. Suppose* U *has a Beta*(*a*, *b*) *distribution, then for any* $c > 0$ *consider a new random variable* V, *such that* $U = \mathrm{e}^{-cV}$, *which has the PDF*

$$f_V(v; a, b, c) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} c \mathrm{e}^{-acv}(1 - \mathrm{e}^{-cv})^{b-1} \quad v > 0. \tag{4}$$

Therefore, the PDF of WE Equation (1) *can be obtained as a special case of Jones' model by taking* $a = 1/\alpha$, $b = 2$ *and* $c = \alpha\lambda$.

INTERPRETATION 4 *Suppose U and V are two independent* $\exp(\lambda)$ *and* $\exp(\lambda(1+\alpha))$, *i.e. exponential random variables with mean* $1/\lambda$ *and* $1/\lambda(1+\alpha)$, *respectively*, *then it can be easily observed that if*

$$X = U + V, \tag{5}$$

then X has the PDF (1). *In fact the stochastic representation of X can be used as a characterization of WE distribution. The following result can be stated without proof*, *as it can be easily obtained using the moment generating function* (*MGF*).

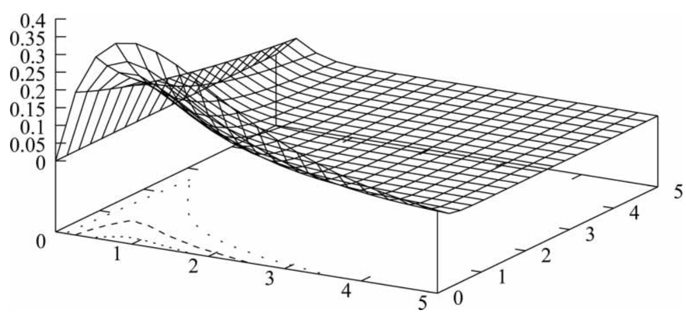

Figure 1. The contour plot and the joint PDF of Equation (3) are plotted for $\lambda = 1$.

Result The random variable X has a WE(α, λ) distribution if and only if X can be written as Equation (5), where U and V are same as defined earlier.

GENERATION *Note that all the above four interpretations can be used for* WE *random deviates generation purposes. The simplest way to generate* WE *random deviate is to use the stochastic representation* (5), *i.e. generate two independent exponential random variables with parameters 1 and* $1 + \alpha$, *respectively. The sum of the two has the* WE$(\alpha, 1)$ *distribution.*

3. Properties

In this section we study the different properties of the WE distribution. Without loss of generality we assume $\lambda = 1$ and for brevity we call it as WE(α). The PDF of the WE distribution is provided in Equation (1). The graph of PDFs of different WE(α) is provided in Figure 2. Since the PDF of X is always log-concave, and $f_X(x; \alpha)$ vanishes as $x \to 0$ or $x \to \infty$, therefore, the PDF is always unimodal with mode at $\ln(1 + \alpha)/\alpha$. In fact it will be seen later that the Pearsonian measure of skewness, i.e. (mean-mode)/σ, and the measure of skewness $\beta_1 = \mu_3/\sigma^3$, see Equation (12) for details, both are always positive.

As $\alpha \to \infty$, WE(α) converges to exp(1) and as $\alpha \to 0$, it converges to a gamma distribution with shape parameter 2. When $\alpha = 1$, WE(α) coincides with the generalized exponential distribution, see Gupta and Kundu [8], with shape parameter 2, which represents the lifetime distribution of a parallel system with two i.i.d. exp(1) lifetime distributions. Although, in general WE distribution represents the lifetime of a system with one spare component, the lifetime distributions of the original and standby components are independently distributed with distributions exp(1) and exp$(1 + \alpha)$, respectively.

The distribution function and the HF of X can be written as

$$F_x(x; \alpha) = \frac{\alpha + 1}{\alpha}\left[1 - \mathrm{e}^{-x} - \frac{1}{1+\alpha}(1 - \mathrm{e}^{-(1+\alpha)x})\right]; \quad x > 0, \tag{6}$$

and

$$h_X(x; \alpha) = \frac{(\alpha + 1)(1 - \mathrm{e}^{-\alpha x})}{(1 + \alpha - \mathrm{e}^{-\alpha x})}; \quad x > 0, \tag{7}$$

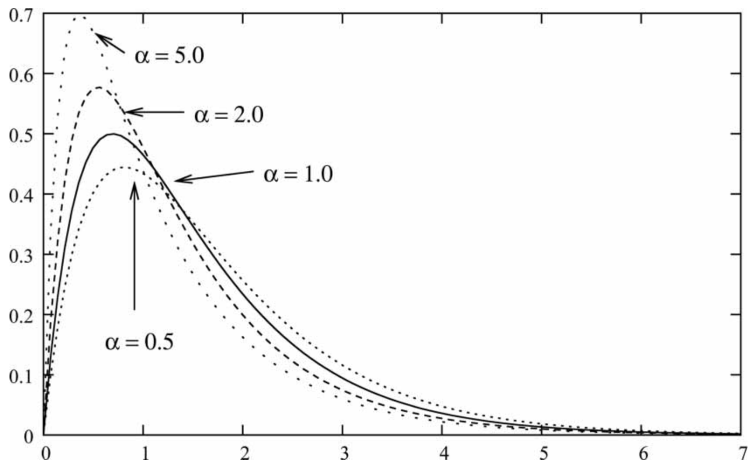

Figure 2. The plot of Equation (1) for different values of α, when $\lambda = 1$.

respectively. Since $f_X(x, \alpha)$ is always log-concave, therefore $h_X(x; \alpha)$ will be an increasing function x for all $\alpha > 0$, see Jones [6] also in this respect. For any α, the HF will start at 0 and will converge to 1 as $x \to \infty$. Since the HF is increasing, this is suitable for modelling lifetime data when wear-out or ageing is present.

Ordering of distributions, particularly among the lifetime distributions, plays an important role in the statistical literature. It can be easily seen that the WE family is a reverse rule of order two (RR_2) family, see Shaked and Shantikumar [9]. Since it has the RR_2 ordering, therefore it has the likelihood ratio ordering, which implies that has the hazard rate ordering and the stochastic ordering. As it has the likelihood ratio ordering, this implies there exists a uniformly power test (UMP) or uniformly most powerful unbiased test (UMPU) for one-sided or two-sided hypothesis on the shape parameter, when the scale parameter is known. Also from the stochastic ordering, it is clear that if $\alpha_1 < \alpha_2$, then WE(α_2) has thinner tail than WE(α_1).

Now let us consider different moments of WE(α) distribution. If X follows WE(α), then the MGF of X for $-1 < t < 1$ can be obtained as

$$M_X(t) = E\mathrm{e}^{tX} = \frac{\alpha+1}{\alpha}\left[(1-t)^{-1} - \frac{1}{1+\alpha}\left(1-\frac{t}{1+\alpha}\right)^{-1}\right] = \left(1-\frac{t}{1+\alpha}\right)^{-1}(1-t)^{-1}. \tag{8}$$

The MGF exists in a compact form, and moreover all the moments exist in this case. It might be easier to work with $\ln M_X(t)$ to compute the cumulants. Since

$$\ln M_X(t) = -\ln(1-t) - \ln\left(1-\frac{t}{1+\alpha}\right), \tag{9}$$

therefore differentiating $\ln M_X(t)$ and having $t = 0$, we obtain;

$$\mu = E(X) = 1 + \frac{1}{1+\alpha}, \quad \sigma^2 = V(X) = 1 + \frac{1}{(1+\alpha)^2}, \quad E(X-\mu)^3 = 2\left[1 + \frac{1}{(1+\alpha)^3}\right] \tag{10}$$

and the k-th cumulant τ_k can be obtained as

$$\tau_k = (k-1)!\left[1 + \frac{1}{(1+\alpha)^k}\right], \quad \text{for } k = 3, \ldots \tag{11}$$

Some other measures, like coefficient of variation (CV) and skewness (β_1) can also be easily obtained in explicit forms, as

$$\mathrm{CV} = \sqrt{\left(1 - \frac{2(1+\alpha)}{(2+\alpha)^2}\right)} \quad \text{and} \quad \beta_1 = \sqrt{\frac{4((1+\alpha)^3+1)^2}{((1+\alpha)^2+1)^3}}. \tag{12}$$

Note that the mean and variance are both decreasing functions of α and both of them decrease from 2 to 1. Both the CV and skewness are functions of α and the CV increases from $1/\sqrt{2}$ to 1, whereas skewness increases from $\sqrt{2}$ to 2. Although, it is not possible to compute the median explicitly, but from Figure 3 it is observed that

$$\text{mode} = \frac{\log(1+\alpha)}{\alpha} < \text{median} < \frac{\alpha+2}{\alpha+1} = \text{mean}.$$

The mean residual life at the time point t is

$$\mu_X(t) = E(X - t | X \geq t) = \frac{1}{1 - 1/(1+\alpha)\mathrm{e}^{-\alpha t}}\left[1 - \frac{\mathrm{e}^{-\alpha t}}{(\alpha+1)^2}\right] - t,$$

and it is a decreasing function of t.

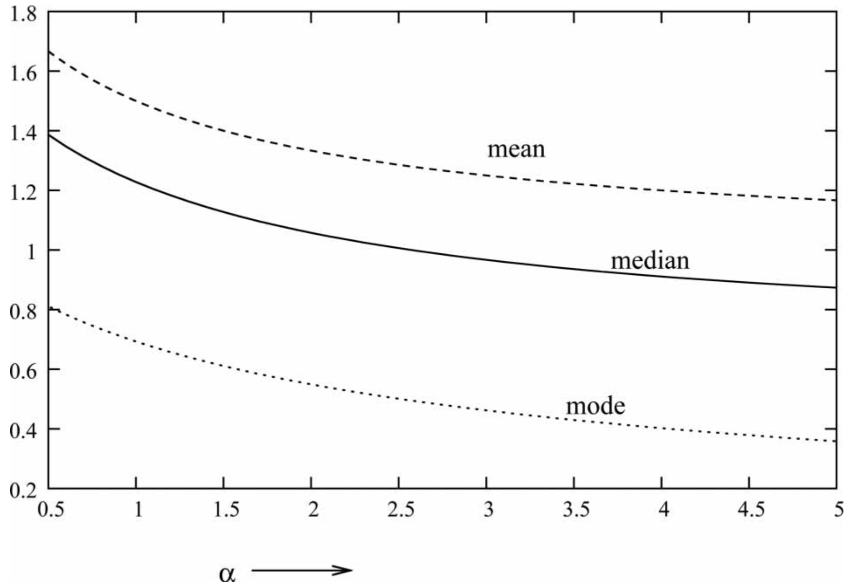

Figure 3. The mean, median and mode of WE for different values of α, when $\lambda = 1$.

Now we discuss the convolution of the WE distributions. We address the following questions. Suppose X_i follows $\mathrm{WE}(\alpha_i)$ for $i = 1, \ldots, n$ and is independent, then what will be the distribution of $\sum_{i=1}^{n} X_i$? Note that it is better to look at the problem from the stochastic representation point of view. Suppose we write $X_i = U_i + V_i$, for $i = 1, \ldots, n$, where U_i and V_i are independent exp(1) and $\exp(1 + \alpha_i)$, respectively. Then studying the distribution is equivalent to studying the distribution of $S + T$, where S and T are independent random variables, S follows gamma(n, 1) and T is the sum of n independent non-identical exponential random variables with parameters $1 + \alpha_1, \ldots, 1 + \alpha_n$, respectively. Several authors have discussed the distribution of T, see for example Johnson *et al.* [10]. Once we get the distribution of T, the distribution of $W = S + T$ can be easily obtained. For completeness purposes, we provide two representations of the distribution function of W, when all the α_is are equal to α. One representation is an infinite mixture of gamma distributions and the other as the convolution of two gamma distributions. The density function of W for $w > 0$ can be written as follows;

Representation 1

$$f_W(w) = \sum_{k=0}^{\infty} \frac{\delta_k}{(1+\alpha)^n} \mathrm{gamma}(2n + k - 1, 1), \tag{13}$$

where

$$\delta_0 = 1, \quad \delta_{k+1} = \frac{1}{k+1} \sum_{i=1}^{k+1} i \gamma_i \delta_{k+1-i}, \quad \gamma_k = \frac{n}{k}\left(1 - \frac{1}{1+\alpha}\right)^k; \quad k = 0, 1, 2, \ldots \tag{14}$$

and gamma(a, 1) represents the gamma density function with shape and scale parameters as a and 1, respectively.

REPRESENTATION 2

$$f_W(w) = \frac{1}{(\alpha+1)^n(\Gamma(n))^2}\int_0^w (w-x)^{n-1}e^{-(w-x)/(\alpha+1)}x^{n-1}e^{-x}\,dx$$
$$= \frac{e^{-w/(1+\alpha)}}{(\Gamma(n))^2(1+\alpha)^n}\sum_{j=0}^{n-1}\frac{\Gamma(n)}{\Gamma(j+1)\Gamma(n-j)}(-1)^j w^{n-j-1}\int_0^w x^{n+j-1}e^{-\alpha x/(1+\alpha)}\,dx$$
$$= \frac{e^{-w/(1+\alpha)}}{(\Gamma(n))(1+\alpha)^n}\sum_{j=0}^{n-1}(-1)^j\frac{\Gamma(n+j)}{\Gamma(j+1)\Gamma(n-j)}w^{n-1-j}\left(\frac{1+\alpha}{\alpha}\right)^{n+j}$$
$$\times\left(1-\sum_{i=0}^{j+n-1}\frac{e^{-(w\alpha/(\alpha+1))}(w\alpha/1+\alpha)^i}{\Gamma(i+1)}\right).$$

Note that in the first case it is an infinite mixture of gamma distributions. Since the mixture coefficients go to zero at the exponential rate therefore the convergence is quite fast. In the second case it is a finite representation, but the coefficients can take both positive and negative values.

Now we briefly discuss the similarity/dissimilarity between WE and truncated skew-Laplace distribution, as suggested by a reviewer. It is expected that WE and truncated skew-Laplace distributions should behave in a similar manner. The skew-Laplace distribution was originally proposed by Balakrishnan and Ambagaspitiya [11] and later it was studied by several authors, see for example Gupta *et al.* [12] and the references cited therein. The skew-Laplace distribution has the PDF, see Gupta *et al.* [12]

$$f_{\mathrm{SL}}(x;\alpha,\lambda) = \begin{cases}\lambda e^{(\alpha+\lambda)x} & \text{if } x<0\\ \lambda e^{-\lambda x}\left(1-\frac{1}{2}e^{-\alpha\lambda x}\right) & \text{if } x\ge 0.\end{cases} \tag{15}$$

Therefore, the truncated skew-Laplace distribution has the PDF

$$f_{\mathrm{TSL}}(x;\alpha,\lambda) = \frac{1+\alpha}{2\alpha+1}\lambda e^{-\lambda x}\left(1-\frac{1}{2}e^{-\alpha\lambda x}\right) \quad \text{for } x\ge 0. \tag{16}$$

It may appear that the PDFs (1) and (16) are of similar forms. We have provided the PDFs and HFs of $f_{\mathrm{TSL}}(x;\alpha,\lambda)$ of the truncated Laplace distribution for different values of α when $\lambda = 1$, in Figure 4. The PDF of truncated skew-Laplace distribution is either unimodal ($\alpha > 1$) or a decreasing ($\alpha \le 1$) function. Moreover at point 0, it will always be positive, unlike WE distribution. The HF of the truncated skew Laplace distribution is either inverted bathtub or a decreasing function.

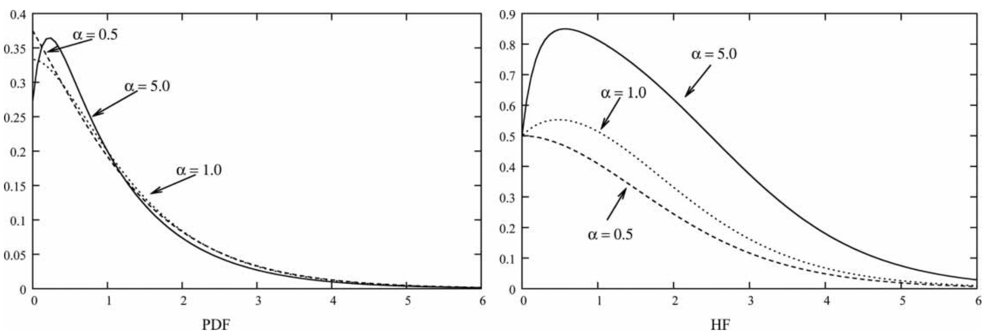

Figure 4. The PDFs and HFs of the truncated skew-Laplace distribution for different values of α, when $\lambda = 1$.

4. Statistical inferences

In this section, we consider the statistical inferences of the unknown parameters. First we calculate the MLEs of the unknown parameters and then we consider their asymptotic distribution.

4.1. *Maximum likelihood estimates*

In this case we re-parametrize α as $\beta = \alpha\lambda$ mainly for computational ease. By making this re-parametrization it is possible to reduce the two-dimensional optimization problem to a one-dimensional optimization problem. The log-likelihood function based on the observed sample $\{x_1, \ldots, x_n\}$ is

$$l(x_1, \ldots, x_n|\beta, \lambda) = n\ln(\beta + \lambda) - n\ln\beta + n\ln\lambda - \lambda\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\ln(1 - \mathrm{e}^{-\beta x_i}). \tag{17}$$

We need to maximize Equation (17) with respect to β and λ to obtain the MLEs. It can be easily observed that for fixed β, the function (17) is maximized for the following λ;

$$\widehat{\lambda}(\beta) = \frac{1}{2\bar{x}}\left(\sqrt{(\beta\bar{x} - 2)^2 + 4\beta\bar{x}} - (\beta\bar{x} - 2)\right). \tag{18}$$

Substituting $\widehat{\lambda}(\beta)$ in Equation (17) we obtain the profile log-likelihood of β and we can obtain the MLE of β by maximizing the profile log-likelihood of β with respect to β. The MLE of β can be obtained as a solution of the following fixed-point type equation;

$$h(\beta) = \beta, \tag{19}$$

where

$$h(\beta) = \left(\frac{1}{\beta + \widehat{\lambda}(\beta)} + \frac{1}{n}\sum_{i=1}^{n}\frac{x_i\mathrm{e}^{-\beta x_i}}{1 - \mathrm{e}^{-\beta x_i}}\right)^{-1}. \tag{20}$$

The solution of Equation (20) can be obtained by a very simple iterative procedure. Suppose we start with an initial guess $\beta_{(0)}$, then the next iterate $\beta_{(1)}$ can be obtained as $\beta_{(1)} = h(\beta_{(0)})$, similarly, $\beta_{(2)} = h(\beta_{(1)})$ and so on. Finally the iterative procedure should be stopped when $|\beta_{(i)} - \beta_{(i+1)}| < \epsilon$, where ϵ is a preassigned tolerance limit. Once we get the MLE of β, the MLE of λ can be obtained as $\widehat{\lambda}(\widehat{\beta})$ from Equation (18). From the MLE of β, MLE of α can also be easily obtained.

Now we discuss the MEs of α and λ. If $\bar{x}$ and s^2 denote the sample mean and sample variance of the observed sample, then solving and equating the two moments, we obtain the MEs of α and λ as

$$\widetilde{\alpha} = \frac{-(\bar{x}^2 - 2s^2) + \sqrt{(\bar{x}^2 - 2s^2)^2 - 2(\bar{x}^2 - s^2)(\bar{x}^2 - 2s^2)}}{\bar{x}^2 - s^2} \tag{21}$$

and

$$\widetilde{\lambda} = \frac{1}{\bar{x}}\left(1 + \frac{1}{1 + \widetilde{\alpha}}\right). \tag{22}$$

It should be mentioned that the ME of α is obtained by solving a quadratic equation. It may have two positive roots, but since the sample mean and sample variance are consistent estimators of the population mean and population variance, respectively, and since $1/\sqrt{2} < \mathrm{CV} < 1$, therefore, for large n with high probability the estimate of α given by Equation (21) will be the only positive

Table 1. Probability of obtaining feasible roots.

$n \to \alpha \downarrow$	50	100	250	500	1000
0.25	0.413	0.449	0.483	0.499	0.536
0.50	0.449	0.503	0.582	0.621	0.718
0.75	0.507	0.590	0.680	0.775	0.872
1.00	0.551	0.664	0.776	0.871	0.959
1.50	0.635	0.769	0.904	0.972	0.997
2.00	0.689	0.843	0.962	0.995	0.999
5.00	0.819	0.943	0.993	1.000	1.000

root. But for small n both of them might be positive or there may be infeasible roots also. It can be easily proved that the MEs exist and is feasible if and only if $s^2 < \bar{x}^2 < 2s^2$. We provide in Table 1 for different values of α and n indicating the probability of obtaining feasible MEs. It is clear from Table 1 that the MEs do not work well if α is small, even for very large sample sizes.

4.2. *Asymptotic distributions: both parameters unknown*

Now we provide the asymptotic distributions of the MLEs when both parameters are unknown. It immediately follows from the Fisher information matrix of λ and β that if $\widehat{\lambda}$ and $\widehat{\beta}$ are the MLEs of λ and β, then;

$$\sqrt{n}((\widehat{\lambda} - \lambda), (\widehat{\beta} - \beta)) \longrightarrow N_2(0, \Sigma_1), \tag{23}$$

where

$$\Sigma_1 = \begin{bmatrix} \dfrac{1}{(\beta+\lambda)^2} + \dfrac{1}{\lambda^2} & \dfrac{1}{(\beta+\lambda)^2} \\ \dfrac{1}{(\beta+\lambda)^2} & \dfrac{1}{(\beta+\lambda)^2} - \dfrac{1}{\beta^2} + \dfrac{\lambda(\beta+\lambda)}{\beta^4} u\left(\dfrac{\lambda}{\beta}\right) \end{bmatrix}^{-1}$$

and for $a > 0$,

$$u(a) = \int_0^1 (\ln(1-y))^2 (1-y)^a y^{-1}\, \mathrm{d}y.$$

Once we have the asymptotic distribution of $(\widehat{\lambda}, \widehat{\beta})$, we can easily obtain the asymptotic distribution of $(\widehat{\lambda}, \widehat{\alpha})$ by using the δ-method and the result can be stated as follows;

$$\sqrt{n}((\widehat{\lambda} - \lambda), (\widehat{\alpha} - \alpha)) \longrightarrow N_2(0, A\Sigma_1 A^T), \tag{24}$$

where

$$A = \begin{bmatrix} 1 & 0 \\ \alpha\lambda & 1/\lambda \end{bmatrix}.$$

Note that the asymptotic distributions of the MEs can also be obtained, but is not presented here. The detailed comparison of the MLEs, MEs and other estimators and their asymptotic properties are in progress and will be reported elsewhere.

Another interesting point (one referee has suggested) will be to compare the asymptotic correlation of $\widehat{\alpha}$ and $\widehat{\lambda}$, and asymptotic correlation of $\widehat{\beta}$ and $\widehat{\lambda}$, particularly when λ is close to one. It is difficult to compare them theoretically because of the complicated nature of the correlations. We have provided the two correlations and also the ratio of the correlations for different values of α in Figure 5. It is clear from Figure 5 that the correlation between $\widehat{\alpha}$ and $\widehat{\lambda}$ or between $\widehat{\beta}$ and $\widehat{\lambda}$ is always negative. They are almost perfectly correlated (very close to -1) when α is small, but the correlation decreases (absolute value) as α increases. Moreover, the ratio of the two correlations is very close to 1 for small α, but it gradually decreases as α increases.

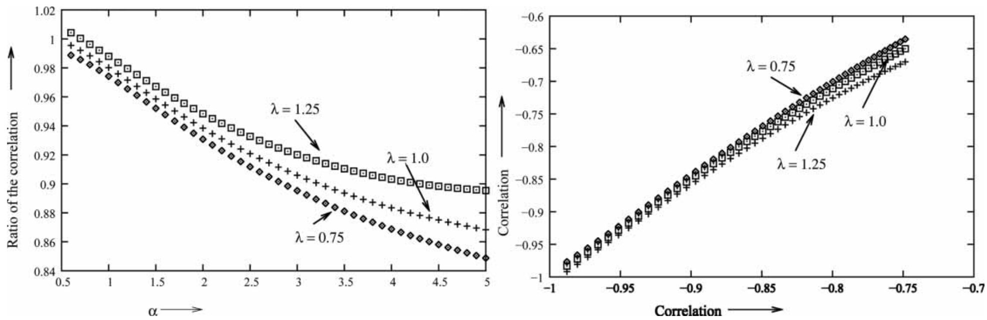

Figure 5. Ratio of Corr($\widehat{\beta}, \widehat{\lambda}$)/ Corr($\widehat{\alpha}, \widehat{\lambda}$) for different values of α and the two correlations are plotted separately.

5. Data analysis

Data Set 1: The data set consists of survival times of guinea pigs injected with different amount of tubercle bacilli and was studied by Bjerkedal [13]. Guinea pigs are known to have high susceptibility of human tuberculosis, which is one of the reasons for choosing this species. We consider only the study in which animals in a single cage are under the same regimen. The data represents the survival times of Guinea pigs in days. The data are given below:

12 15 22 24 24 32 32 33 34 38 38 43 44 48 52 53 54 54 55 56 57 58 58 59 60 60 60 60 61 62 63 65 65 67 68 70 70 72 73 75 76 76 81 83 84 85 87 91 95 96 98 99 109 110 121 127 129 131 143 146 146 175 175 211 233 258 258 263 297 341 341 376.

In this case $n = 72$, the mean $\bar{x} = 99.82$ and standard deviation $s = 80.55$. Since, $1 < \bar{x}^2/s^2 = 1.536 < 2$, therefore, the MEs of α and λ exist. They are $\widetilde{\alpha} = 2.6124$ and $\widetilde{\lambda} = 0.0128$. We plot the profile likelihood of β in Figure 6. Since it is an unimodal function, it has the unique maximum and we obtain $\widehat{\beta} = 0.0224$, $\widehat{\lambda} = 0.0138$ and $\widehat{\alpha} = \widehat{\beta}/\widehat{\lambda} = 1.6232$. Just to see how good fit it is we have plotted the empirical distribution function and the fitted WE distribution functions in Figure 7 and the fitted PDF and relative histogram in Figure 8. The Kolmogorov–Smirnov (K–S)

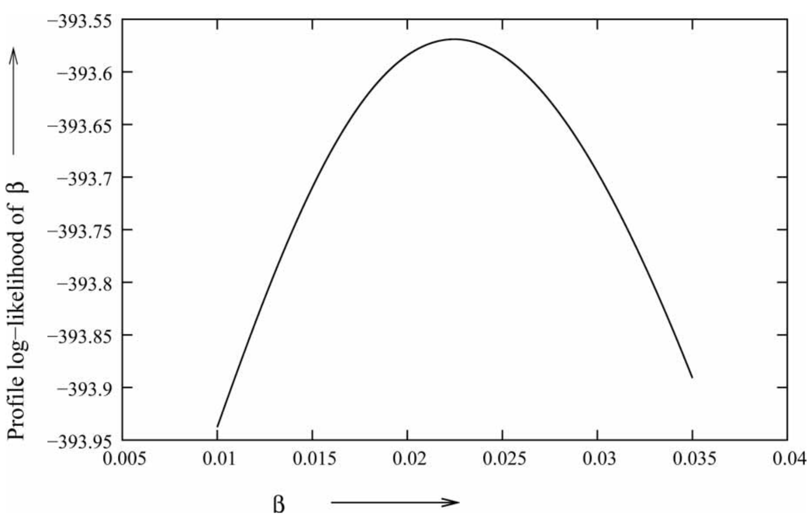

Figure 6. The profile log-likelihood function of β for the Guinea pig data set.

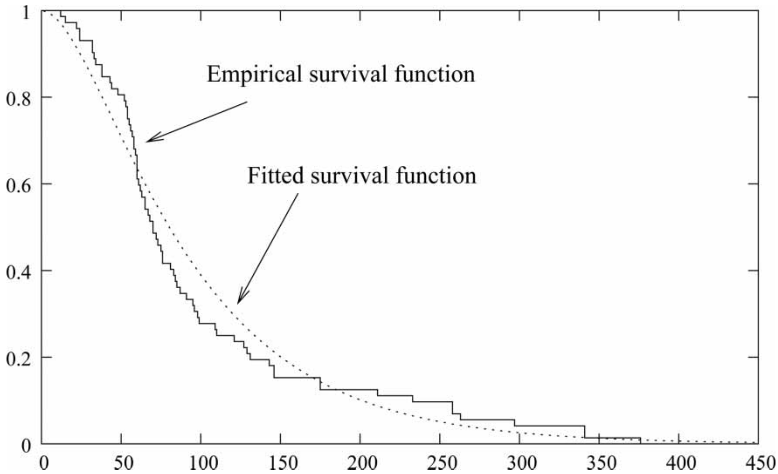

Figure 7. The empirical and the fitted survival functions for the Guinea pig data set.

distance between the empirical and fitted distribution functions is 0.1173 and the corresponding p-value is 0.2748. It clearly indicates that the WE distribution provides a good fit to the data. For comparison purposes, we have fitted Weibull, gamma and generalized exponential distributions also. It is observed that the K–S distances are 0.1490, 0.1392 and 0.1349, and the corresponding p-values are 0.0816, 0.1122 and 0.1349, respectively.

DATA SET 2: The Indian Institute of Technology Kanpur is one of the famous technical institutes of India. Here the admission to the different disciplines in the first year, is performed through the Joint Entrance Examination (JEE). Approximately 450 students join each year in the first year. Since the programme is very tough and the competition is also very high, the students can choose a slow pace programme that works as follows. In the first year all the courses are compulsory for everybody. In the first mid-semester examination if students get below than certain mark (not fixed in each year and vary from one subject to others) in a particular subject they can go for a slow pace programme which may extend their programme duration. In 2003, the marks of the slow pace students in Mathematics in the final examination were as follows;

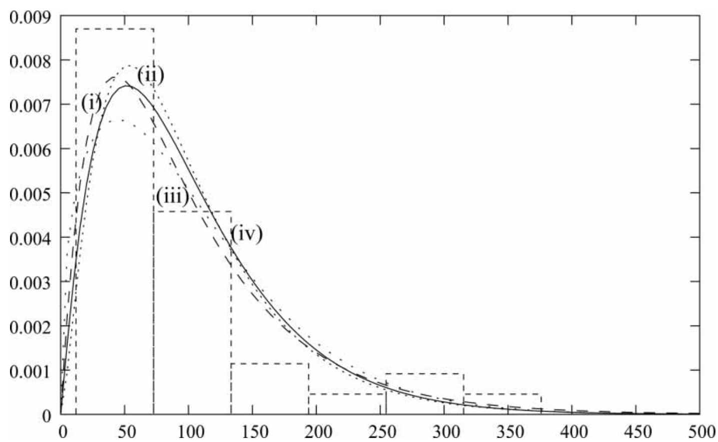

Figure 8. The fitted PDFs and the relative histogram for the Guinea pig data set. (i) WE, (ii) GE (iii) Weibull and (iv) gamma.

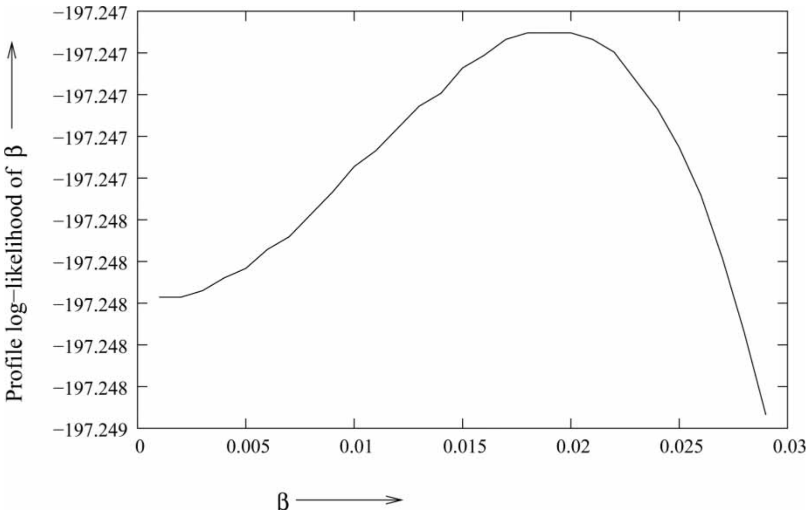

Figure 9. The profile log-likelihood function of β for the marks set.

29 25 50 15 13 27 15 18 7 7 8 19 12 18 5 21 15 86 21 15 14 39 15 14 70 44 6 23 58 19 50 23 11 6 34 18 28 34 12 37 4 60 20 23 40 65 19 31.

We do not have the access to the first semester marks, so we do not know the cut-off. In this case $n = 48$, the mean $\bar{x} = 25.89$ and standard deviation $s = 18.60$. Since, $1 < \bar{x}^2/s^2 = 1.937 < 2$, therefore, in this case also the MEs of α and λ exist. The MEs are $\tilde{\alpha} = 0.4384$ and $\tilde{\lambda} = 0.0655$. The profile log-likelihood of β can be seen as a unimodal function (Figure 9). The MLEs are obtained as follows: $\widehat{\beta} = 0.0200$, $\widehat{\lambda} = 0.0685$ and $\widehat{\alpha} = \widehat{\beta}/\widehat{\lambda} = 0.2919$. The empirical survival function and the fitted survival function are plotted in Figure 10. The fitted PDF and the relative histogram are plotted in Figure 11. The Kolmogorov–Smirnov distance between the empirical and fitted distribution functions is 0.0932 and the corresponding p-value is 0.7986. The K–S distances between the empirical distribution function and the fitted Weibull, gamma and

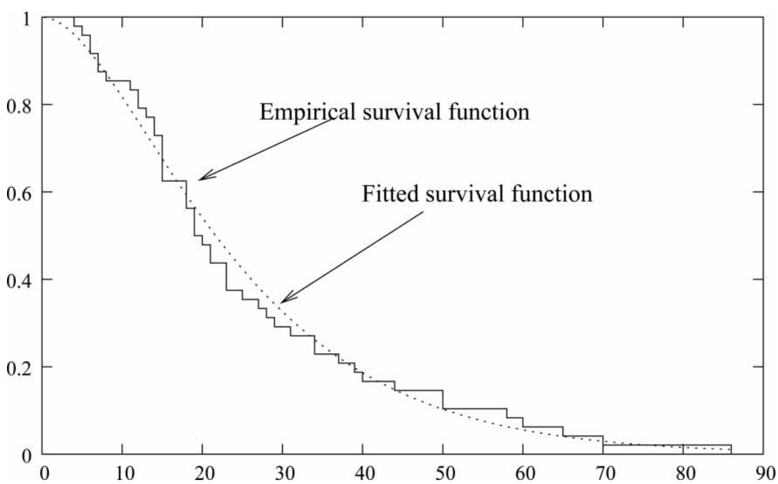

Figure 10. The empirical and fitted survival functions of the marks data.

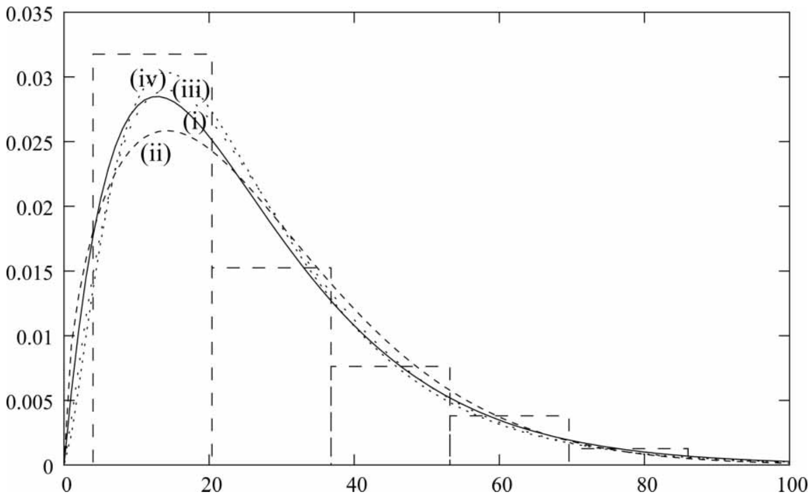

Figure 11. The fitted PDF and the relative histogram for the marks data. (i) WE, (ii) Weibull (iii) gamma and (iv) GE.

generalized exponential are 0.1177, 0.1031, 0.0937 and the corresponding p-values are 0.5186, 0.6638 and 0.7929, respectively. Interestingly in both the cases it is observed that WE fits better than the Weibull, gamma or generalized exponential distributions, although it may not always be true. Moreover, the tail behaviour in both the cases for all the four fitted distributions are very similar in nature.

It may be argued (as one referee pointed out) that how are we fitting the WE distribution to the marks data, when it is known that the maximum marks can be only 100. But it is not very uncommon to fit a normal distribution when it is known that the data are bell shaped even when it has a finite range. Moreover, in this case the MLE of $P(X > 100) \leq 10^{-41}$. This justifies that even if X is bounded above, it is possible to use WE distribution in this case, since the fit is also very good otherwise.

6. Conclusions

In this article we have considered a WE distribution obtained using the method of Azzalini. The proposed WE distribution is a generalization of the exponential distribution similar to the well-known Weibull, gamma or generalized exponential distributions. Although for all the above three cases, the exponential distribution can be obtained as a special case, in case of WE distribution, the exponential distribution can be obtained only as a limiting distribution. It is observed that the PDF of the WE distribution is very similar to the corresponding PDFs of the well-known Weibull, gamma or generalized exponential distributions. Therefore, the WE can be used quite effectively to analyse positively skewed data similar to those above skewed distributions. It can be observed as a hidden truncation model, and therefore if it is known that the data are coming from a hidden truncation model, then it can be used quite effectively instead of other skewed distributions.

Now for comparison purposes, we have provided the different characteristics of Weibull, gamma, generalized exponential and WE models in Table 2. We have assumed in all the cases that α represents the shape parameter and the scale parameter is 1, without loss of generality. One very important and natural question is how close or different these four distributions are? It is not very easy to answer, and more work is needed in this direction.

Table 2. The different characteristics of the four families of distributions.

Model	Range of α	Shapes of the Density function	Shapes of the Hazard function	Mean	Mode	Ordering relation
Generalized exponential	$0 < \alpha < 1$	D (from ∞ to 0)	D (from ∞ to 1)		0	
	$\alpha = 1$	D (from α to 0)	C (at α)	$\psi(\alpha+1) - \psi(1)$	0	TP_2
	$\alpha > 1$	U (starts from 0)	I (from 0 to 1)		$\log(\alpha)$	
Gamma	$0 < \alpha < 1$	D (from ∞ to 0)	D (from ∞ to 1)		0	
	$\alpha = 1$	D (from α to 0)	C (at α)	α	0	TP_2
	$\alpha > 1$	U (starts from 0)	I (from 0 to 1)		$\alpha - 1$	
Weibull	$0 < \alpha < 1$	D (from ∞ to 0)	D (from ∞ to 0)		0	
	$\alpha = 1$	D (from α to 0)	C (at α)	$\Gamma\left(1+\frac{1}{\alpha}\right)$	0	N
	$\alpha > 1$	U (starts from 0)	I (from 0 to ∞)		$\left(\frac{\alpha-1}{\alpha}\right)^{1/\alpha}$	
WE	$\alpha > 0$	U (starts from 0)	I (from 0 to 1)	$\frac{\alpha+2}{\alpha+1}$	$\frac{\log(1+\alpha)}{\alpha}$	RR_2

D = decreasing, I = increasing, U = unimodal, C = constant, N = neither.

Acknowledgements

The authors would like to thank two anonymous referees for their valuable comments. Part of the work of the first author was supported by a grant from the Natural Sciences and Engineering Research Council, CANADA and part of the work of the second author was supported by a grant from the Department of Science and Technology, Government of India.

References

[1] A. Azzalini, *A class of distributions which includes the normal ones*, Scand. J. Stat. 12 (1985), pp. 171–178.
[2] B.C. Arnold and R.J. Beaver, *The skew Cauchy distribution*, Stat. Probab. Lett. 49 (2000), pp. 285–290.
[3] R.D. Gupta and D. Kundu, *Skew-Logistic distribution*, Tech. Rep., Indian Institute of Technology Kanpur, Kanpur, India, 2007.
[4] M.G. Genton (ed.), *Skew-Elliptical Distributions and their Applications: A Journey beyond Normality*, Chapman & Hall/CRS, New York, 2004.
[5] B.C. Arnold and R.J. Beaver, *Hidden truncation models*, Sankhya, Ser. A 62 (2000), pp. 23–35.
[6] M.C. Jones, *Families of distributions arising from distributions of order statistics*, Test 13 (2004), pp. 1–43.
[7] S. Nadarajah and S. Kotz, *The beta exponential distribution*, Reliab. Eng. Syst. Saf. 91 (2006), pp. 689–697.
[8] R.D. Gupta and D. Kundu, *Generalized exponential distributions*, Aust. N. Z. J. Stat. 41 (1999), pp. 173–188.
[9] M. Shaked and J.G. Shantikumar, *Stochastic Ordering and Their Applications*, Academic Press, San Diego, 1994.
[10] N.L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distribution*, Vol. 1, 2nd ed., John Wiley and Sons, New York, 1994.
[11] N. Balakrishnan and R.S. Ambagaspitiya, *On skew-Laplace distributions*, Tech. Rep. Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada, 1994.
[12] A.K. Gupta, F.C. Chang, W.J. Huang, *Some skew-symmetric models*, Random Oper. Stoch. Equ. 10 (2002), pp. 133–140.
[13] T. Bjerkedal, *Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli*, Amer. J. Hyg. 72 (1960), pp. 130–148.