Journal of Probability and Statistical Science
7(2), 153-171, Aug. 2009

Introduction of Shape/Skewness Parameter(s) in a Probability Distribution

Rameshwar D. Gupta
University of New Brunswick

Debasis Kundu
Indian Institute of Technology

ABSTRACT In this paper we discuss six different methods to introduce a shape/skewness parameter in a probability distribution. It should be noted that all these methods may not be new, but we provide new interpretations to them and that might help the partitioner to choose the correct model. It is observed that if we apply any one of these methods to any probability distribution, it may produce an extra shape/skewness parameter to that distribution. Structural properties of these skewed distributions are discussed. For illustrative purposes, we apply these methods when the base distribution is exponential, which resulted in five different generalizations of the exponential distribution. It is also observed that if we combine two or more than two methods successively, then it may produce more than one shape/skewness parameters. Several known distributions can be obtained by these methods and various new distributions with more than one shape parameters may be generated. Some of these new distributions have several interesting properties.

Keywords Hazard function; Reversed hazard function; Cumulative distribution function; Odds ratio; Cumulant generating function.

1. Introduction

Recently there has been quite a bit of interests to define skewed distributions in various manners. Azzalini [1, 2] first introduced the skew-normal distribution and then a series of papers appeared on skewed distribution both in the univariate and multivariate set up. Arnold

□Received September 2008, revised January 2009, in final form February 2009.
□Rameshwar D. Gupta is affiliated with the Department of Computer Science and Applied Statistics at The University of New Brunswick, Saint John, Canada E2L 4L5. Debasis Kundu (corresponding author) is affiliated with the Department of Mathematics and Statistics at Indian Institute of Technology, Kanpur Pin 208016, India; email: kundu@iitk.ac.in, web page: http://www.iitk.ac.in/math/faculty/kundu.
□Part of the first author's work was supported by a grant from The Natural Sciences and Engineering Research Coulcil.

 ISSN 1726-3328

and Bever [3], provided a nice interpretations of Azzalini's skew normal distribution as a hidden truncation model. For other univariate and multivariate skewed distributions which have been defined along the same line of Azzalini [1], the readers are referred to Arnold and Bever [4], Gupta and Gupta [9], the recent monograph by Genton [8] and the references cited there.

Another set of skewed distributions which has been developed using the probability transformation was studied by Ferreira and Steel [7]. Ferreira and Steel [7] generalized the idea of Jones [17], who had mainly used the beta probability integration transformation. Although both Azzalini's and Ferreira and Steel's families of skewed distributions can be viewed as weighted distributions, but the two classes do not overlap. Moreover, it may be mentioned that although Azzalini's family of distributions has some practical interpretations, but the skewed family developed through probability integration transformation do not have any proper practical interpretations till date.

The main aim of this paper is to discuss six different general methods which may be used to introduce an additional shape/ skewness parameter to a base probability model. The base probability model may or may not have any shape/ skewness parameter, but applying one of these methods an additional shape/ skewness parameter may be introduced. Moreover, if we combine two or more methods sequentially, more than one shape/ skewness parameter (s) may be introduced.

As we have already mentioned that, not all these methods are completely new. Some of these methods have already been applied to some specific distributions and some of them have been discussed in more general form in the literature. But in this paper we try to provide some interpretations from the reliability or survival theory point of view. Although, all the six methods can be applied to any base distribution with arbitrary support, but we mainly provide the interpretations and restrict our discussions for the distributions with non-negative support only. It is observed that many generalized distributions, which have been already discussed in the literature can be obtained by using one of these methods or combining more than one methods sequentially.

We discuss the structural properties of the different families in general. It is observed that from these general properties, several existing known properties can be obtained as particular cases. If the base distribution is completely known then the inference on the shape parameter can be obtained very easily in most of the cases. If the base distribution also has unknown parameter(s) then the estimation of the unknown parameters becomes more complicated mainly from the computational point of view.

For illustrative purposes, we have demonstrated our methods using exponential distribution as the base distribution. We discuss and compare their different properties. It is observed that out of six methods, five methods produce an extra shape parameter to the base exponential distribution, which resulted five different generalization of the exponential distribution. Gamma, Weibull, generalized exponential, geometric extreme distributions are obtained by using different methods. In all cases except one it is observed that the exponential distribution is a member of the generalized family, but in one case it is observed that the exponential distribution can be obtained as a limiting member of the generalized family. Finally, we have applied more than one methods sequentially when the base distribution is exponential and it is observed that many familiar distributions and several new distributions can be obtained by this procedure.

The rest of the paper is organized as follows. In section 2, we discuss all the six different methods and provide interpretations. Structural properties of the different families are discussed in section 3. In section 4, we discuss basic inferential procedures for the different families. Different examples with exponential distribution as the base distribution are provided in section 5. In section 6, we discuss the effect of applying more than one method sequentially to a particular probability model. Finally we conclude the paper in section 7.

2. Introduction of a Shape Parameter

Let X be a random variable with the probability density function (PDF) $f_X(\cdot)$ and the distribution function $F_X(\cdot)$. We will use the following notation for the rest of the paper. The survival function, the hazard function, the reversed hazard function, the characteristic function and the odds ratio will be defined by $\bar{F}_X(x) = 1 - F_X(x)$, $h_X(x) = f_X(x)/\bar{F}_X(x)$, $r_X(x) = f_X(x)/F_X(x)$, $\phi_X(t) = Ee^{itX}$, $\psi_X(x) = \bar{F}_X(x)/F_X(x) = r_X(x)/h_X(x)$ respectively. It is assumed unless otherwise mentioned that $F_X(\cdot)$ is the base distribution and $F_X(\cdot)$ does not have any other parameters. Note that in all our developments, it is always possible to introduce the location and scale parameters in $F_X(\cdot)$. It is further assumed that Y is the random variable obtained from the random variable X after using one of the following six methods. Now we provide the different methods which may introduce a shape/ skewness parameter in the distribution function of Y. The notations of the density function, distribution function, survival function, hazard function, reversed hazard function, characteristics function and odds ratio of the random variable Y are same as those of X, except replacing the subscript X by Y.

2.1 Method 1: Proportional Hazard Model

The random variables X and Y satisfy the proportional hazard model (PHM) with proportionality constant $\alpha > 0$, if

$$h_Y(x) = \alpha h_X(x), \quad \text{for all } x. \tag{1}$$

Note that if two random variables X and Y satisfy (1), then their survival and density functions satisfy the following relation

$$\bar{F}_Y(x) = \left(\bar{F}_X(x)\right)^{\alpha} \quad \text{and} \quad f_Y(x) = \alpha\left(\bar{F}(x)\right)^{\alpha-1} f_X(x) \tag{2}$$

respectively. Here α might be called the exponentiated parameter. By performing this operation for $\alpha > 0$, the transformed variable may have a shape parameter α.

Proportional hazard model has been in the statistical literature since the work of Cox [6]. The proportional hazard model is one of the most celebrated model in the survival analysis. Although Cox introduced the proportional hazard model to introduce the covariates, but the same concept can be used to introduce an additional shape/ skewness parameter to the base distribution. Among the different distributions, the well known Burr XII, see for example Johnson *et al.* [15], distribution can be obtained as a proportional hazard model from the base distribution function

$$F_X(x) = \frac{x^c}{1+x^c}, \quad \text{for} \quad x > 0, c > 0. \tag{3}$$

Some of the other proportional hazard models are Lomax or Pareto distributions. Note that in all cases, if α is an integer, say m, then the distribution function $F_Y(\cdot)$ represents the distribution function of the system with m independent identically distributed ($i.i.d.$) parallel components each having distribution function $F_X(\cdot)$.

2.2 Method 2: Proportional Reversed Hazard Model

Although, hazard function has been used in the reliability or survival analysis for quite some times, but the reversed hazard function has been introduced by Shaked and Shantikumar [22] not that long back. It is observed that this function can be successfully used in Forensic studies where time since failure plays an important role. The reversed hazard function of a random variable is defined as the ratio of the density function to its distribution function, as mentioned earlier. Now, we define the proportional reversed hazard model similarly as the PHM.

Suppose, X and Y are two random variables. The distribution function of Y is said to be proportional reversed hazard model (PRHM) of the distribution function of X with proportionality constant $\alpha > 0$, if

$$r_Y(x) = \alpha\, r_X(x), \quad \text{for all} \quad x > 0. \tag{4}$$

Note that if two random variables X and Y satisfy (4), then the corresponding distribution and density functions respectively become,

$$F_Y(x) = (F_X(x))^{\alpha} \quad \text{and} \quad \alpha f_X(x)\,(F_X(x))^{\alpha-1}. \tag{5}$$

Here also α might be called as the exponentiated parameter.

Recently, several exponentiated distributions have been studied quite extensively, since the work of Mudholkar and Srivastava [20] on exponentiated Weibull distribution. It is observed by Mudholkar and Srivastava [20] that the exponentiated Weibull distribution has two shape parameters and because of that the density function and the hazard function can take different shapes. Although, two-parameter Weibull distribution can have increasing or decreasing hazard function, depending on the shape parameter, but the exponentiated Weibull distribution can have increasing, decreasing, bath-tub or inverted bath-tub hazard functions.

Some of the other exponentiated distributions which have been found in the literature are exponentiated gamma, exponentiated exponential, exponentiated Rayleigh, exponentiated

Pareto, see for example the review articles by Gupta and Kundu [14] and Gupta and Gupta [10] and the references cited their in these respects. If Y is the PRHM of X and α is an integer say m, then the distribution function $F_Y(\cdot)$ represents the distribution function of a series system with m, $i.i.d.$ components each having the distribution function $F_X(\cdot)$.

2.3 Method 3: Proportional Cumulants Model

Similar to the PHM and PRHM model, we can also define the proportional cumulants model (PCM) as follows. Suppose X and Y are two random variables. The distribution function of Y is said to be PCM of the distribution function of X with proportionality constant α if

$$\eta_Y(x) = \alpha \eta_X(x) \quad \text{for all } \; x > 0, \tag{6}$$

where $\eta_Y(\cdot)$ and $\eta_X(\cdot)$ are the cumulants generating functions of Y and X respectively. Note that (6) is equivalent as

$$\phi_Y(x) = (\phi_X(x))^{\alpha}, \tag{7}$$

where $\phi_Y(\cdot)$ and $\phi_X(\cdot)$ are the characteristic functions of Y and X respectively. It may be mentioned that if $\phi_X(\cdot)$ is a characteristic function, then for any $\alpha > 0$, $(\phi_X(\cdot))^{\alpha}$ may not be always a characteristic function. But if $\phi_X(\cdot)$ is symmetric or α is an integer, then $(\phi_X(\cdot))^{\alpha}$ will be always a characteristic function. Moreover, if X belongs to a natural exponential family, then also $(\phi_X(\cdot))^{\alpha}$ will be a characteristic function for any $\alpha > 0$.

Some of the PCM are already well studied in the literature. For example, if X follows an exponential distribution, then Y follows gamma with the shape parameter α. Therefore, gamma is a PCM. Although the base distribution X does not have any shape parameter, but in this case Y has a shape parameter. Note that if the base distribution is symmetric, then the transformation (7) may not produce a shape parameter in Y. For example, if the base distribution is normal, student-t, Laplace, Cauchy or Linnik, then the transformation (7) does not produce any extra shape parameter in Y.

If α is an integer, say m, then Y can be observed as the convolution of m $i.i.d.$ components of X. Equivalently, it can be thought of as the total lifetime of a system with $m-1$ $i.i.d.$ spare components, where each spare components and the original component have the lifetime distribution function $F_X(\cdot)$.

2.4 Method 4: Proportional Odds Model

The odds ratio plays an important role in the survival and reliability analysis. It is mainly used to compare two groups. For a random variable X, the odds ratio on surviving beyond t are given by the "odds function" $\psi_X(\cdot)$. Along the same line as PHM, PRHM, PCM, it is possible to define the proportional odds model (POM) as follows. Suppose X and Y are two random variables. The distribution function of Y is said to be proportional odds model of the distribution function of X, with proportionality constant α, if

$$\psi_Y(t) = \alpha \; \psi_X(t), \quad \text{for all } \; t. \tag{8}$$

Note that (8) can be written equivalently as

$$\psi_Y(t) = \frac{\bar{F}_Y(t)}{F_Y(t)} = \alpha\,\frac{\bar{F}_X(t)}{F_X(t)} = \alpha\ \psi_X(t) \quad \text{or} \quad \frac{r_Y(t)}{h_Y(t)} = \alpha\,\frac{r_X(t)}{h_X(t)}. \tag{9}$$

Therefore, the PDF and CDF of Y become

$$f_Y(x) = \frac{\alpha f_X(x)}{\left(F_X(x) + \alpha \bar{F}_X(x)\right)^2} \quad \text{and} \quad F_Y(x) = \frac{F_X(x)}{F_X(x) + \alpha \bar{F}_X(x)}, \tag{10}$$

respectively. Marshal and Olkin [19] first proposed the general method (10) to incorporate a shape parameter in the base random variable X and studied specifically when the base distribution is either exponential or Weibull. They have shown the flexibility of the transformed model due to the introduction of the shape parameter α. It is observed that this family of models is geometric-extreme stable, *i.e.*, if $Y_1, Y_2, \ldots$ is a sequence of independent copies of Y and if N is a geometric distribution having support on $\{1, 2, \ldots\}$, then $\min\{X_1, \ldots, X_N\}$ and $\max\{X_1, \ldots, X_N\}$ have the distribution functions from that family.

2.5 Method 5: Power Transformed Model

The transformation of the observed data has been used quite regularly for statistical data analysis. Among the different transformations, the power transformation becomes the most popular one. The most celebrated power transformation method was introduced by Box and Cox [5] and it is well known as Box-Cox transformation. Although, Box and Cox [5] introduced the power transformation method in regression analysis when the variances are not homogeneous, but power transformation technique can be used to introduce a shape/ skewness parameter in a probability model also.

Suppose X is a non-negative random variable, then for $\alpha > 0$, consider a new random variable Y such that

$$Y = X^{\frac{1}{\alpha}}. \tag{11}$$

If two non-negative random variables X and Y satisfy (11) then the corresponding distribution functions and PDFs satisfy

$$F_Y(x) = F_X(x^\alpha) \quad \text{and} \quad f_Y(x) = \alpha x^{\alpha-1} f_X(x^\alpha), \tag{12}$$

respectively. Moreover, the hazard functions and reversed hazard functions satisfy respectively

$$h_Y(x) = \alpha x^{\alpha-1} h_X(x^\alpha), \quad \text{and} \quad r_Y(x) = \alpha x^{\alpha-1} r_X(x^\alpha). \tag{13}$$

If X is the base random variable, then the class of random variables Y, which are related through (11) is called power transformed models (PTMs).

One of the most important class of PTMs is the Weibull class. For example, if X follows an exponential distribution then Y follows Weibull distribution with the shape parameter α. Although, the exponential distribution does not have a shape parameter, then Y, the PTM of X, has a shape parameter. It may be mentioned that if the base distribution already has a shape parameter, then the PTMs may not have an additional shape parameter. For example, if X has Weibull or Log-normal distribution, then Y also has Weibull or Log-normal distribution respectively.

2.6 Method 6: Azzalini's Skewed Model

Azzalini [1] introduced skew-normal distribution as a natural extension of the classical normal density to accommodate asymmetry. He proposed the skew-normal distribution which has the following density function;

$$f_{SN}(x;\alpha) = 2\Phi(\alpha x)\phi(x); \quad -\infty < x < \infty. \tag{14}$$

Here $-\infty < \alpha < \infty$ is the shape/ skewness parameter, $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and distribution function of the standard normal random variable. A motivation of the above model has been elegantly exhibited by Arnold and Beaver [3]. Azzalini [1] discussed the following method by which skew-normal density function of the form (14) can be obtained.

Consider two independent and identically distributed standard normal random variables U and V. Now define Y to be equal to U conditionally on the event $\{\alpha U > V\}$. The resulting density function of Y will be (14). Now using the same idea of Azzalini's skew-normal distribution, it is possible to introduce a skewness/ shape parameter to any distribution, which may or may not have any shape parameter originally. Although, it is possible to extend the model for a random variable X with arbitrary support, but here we restrict ourselves only for $X \ge 0$.

Suppose, X is a random variable with PDF $f_X(\cdot)$ and CDF $F_X(\cdot)$, then consider a new random variable Y, for any $\alpha > 0$, with the following PDF;

$$f_Y(y;\alpha) = \frac{1}{P(\alpha X_1 > X_2)} f_X(y) F_X(\alpha y); \quad y > 0. \tag{15}$$

Here X_1 and X_2 are two independent and identically distributed random variables having the distribution function $F_X(\cdot)$. If X is the base random variable, the class of random variables Y, which are related through (15) for $\alpha > 0$, is called Azzalini's skewed models (ASMs).

It can be easily seen that if X is symmetric random variable on $(-\infty, \infty)$ then $P(\alpha X_1 > X_2) = 1/2$. Several authors discussed the properties of the random variables having density function of the form (15), when X is symmetric and not necessarily standard normal. For example, skew-Cauchy, skew-t and skew-Laplace have been discussed by several authors, see Genton [8] and the references cited there.

3. Structural Properties

3.1 Proportional Hazard Model

In case of PHMs, it is observed that if X has the PDF $f_X(\cdot)$, and if the random variable Y belongs to the class of PHM of X, then the PDF of Y is given by (2). Therefore, it is clear that the PDF of Y is a weighted version of X, where the weight function $w(x) = \left(\bar{F}_X(x)\right)^{\alpha-1}$. The weight function $w(\cdot)$ is a decreasing function if $\alpha > 1$ and it is an increasing function if $\alpha < 1$. It implies, if the base random variable X has a decreasing PDF, then for any $\alpha > 1$, Y also has a decreasing PDF. If X has an unimodal PDF, then the Mode(Y) $<$ ($>$) Mode (X), if $\alpha >$ ($<$)1. Comparing the two PDFs, it is easily observed that if X has a log-concave (convex) density function then the PDF of Y will be log-concave (convex) if $\alpha > 1$ and vice versa for $\alpha < 1$. Using the log-concavity (convexity) property of the PDF, the monotonicity property of the hazard function can be easily obtained.

If x_p and y_p are the $p-th$ percentile points of X and Y respectively, then $y_p < (>) x_p$, if $\alpha > (<)1$. It shows Y has heavier tail than X if $\alpha > 1$ and vice versa. The $r-th$ moment of the random variable Y can be written as

$$E(Y^r) = \alpha \int_0^\infty x^r \left(\bar{F}_X(x)\right)^{\alpha-1} f_X(x)dx = \int_0^1 \alpha \left[F_X^{-1}(u)\right]^r (1-u)^{\alpha-1} du. \tag{16}$$

The explicit expression of (16) may not be always possible. The corresponding L-moments similarly as PRHM (see next sub section) also can be obtained in this case. Moreover, from (1) it is clear that the shape of the hazard function of Y is same as the shape of the hazard function of X. Let us write the weight function $w(x) = \left(\bar{F}(x)\right)^{\alpha-1}$ as $w(x,\alpha)$. It can be easily verified that $w(x,\alpha)$ is a reverse rule of order two (RR_2) on $x > 0$ and $\alpha > 0$. Therefore, the PHM family is a RR_2 family. It is well known, see Shaked and Shantikumar [22], that if a family is total positivity of order two (TP_2) or RR_2, then it has all the ordering properties, namely likelihood ratio ordering, hazard rate ordering, stochastic ordering, etc. Therefore, PHM family members enjoy all these properties.

3.1 Proportional Reversed Hazard Model

It is defined in the previous section that if the random variable X has the PDF $f_X(\cdot)$, then the random variable Y belongs to the class of PRHM if Y has the PDF (5). The PDF of Y is a weighted version of the PDF of X and the weight function in this case is $w(x) = (F_X(x))^{\alpha-1}$. The weight function is an increasing or decreasing function if $\alpha > 1$ or $\alpha < 1$ respectively. Therefore, if X has a decreasing PDF then for $\alpha < 1$, Y also has a decreasing PDF. If $f_X(\cdot)$ is unimodal, then Mode(Y) $>$ ($<$) Mode(X), if $\alpha >$ ($<$)1. Observe that if the PDF of X is log-concave (convex), then the PDF of Y will be log-concave (convex) if $\alpha >$ ($<$) 1. Moreover, the family is a TP_2 family.

Also if x_p and y_p are the p-th percentile points of X and Y respectively, then $y_p < (>) x_p$ if $\alpha < (>)1$. It is immediate that Y has a heavier or lighter tail than X provided $\alpha > 1$ or $\alpha < 1$ respectively. The r-th moment of Y is

$$E(Y^r) = \alpha \int_0^\infty f_X(x) x^r (F_X(x))^{\alpha-1} dx = \alpha \int_0^1 \left(F^{-1}(u)\right)^r u^{\alpha-1} du = M(\alpha, r) \text{ (say).} \tag{17}$$

Explicit expression of $M(\alpha, r)$ is possible in some cases. The L-moment can be expressed in terms of $M(\alpha, 1)$. To observe this first let us define the r-th weighted moment as

$$\beta_r = E\left(Y(F_Y(Y))^r\right) = \frac{1}{r+1} M(\alpha(r+1), 1). \tag{18}$$

It is well known that the L-moments can be defined as the linear combinations of probability weighted moments as follows;

$$\lambda_{k+1} = \sum_{j=0}^{k} (-1)^{k-j} \frac{(k+j)!}{(k-j)!(j!)^2} \beta_j; \quad k = 1, 2, \ldots. \tag{19}$$

Therefore, the first few L-moments of Y are as follows;

$$\begin{aligned} \lambda_1 &= \beta_0 = M(\alpha, 1) \\ \lambda_2 &= 2\beta_2 - \beta_0 = M(2\alpha, 1) - M(\alpha, 1) \\ \lambda_3 &= 6\beta_2 - 6\beta_1 + \beta_0 = 2M(3\alpha, 1) - 3M(2\alpha, 1) + M(\alpha, 1). \end{aligned}$$

Again studying the weight function, it can be easily observed that the PRHM family is a TP_2, and therefore the members in this family satisfy all the ordering properties with respect to α. The different monotonicity properties of the hazard function have been discussed in details by Gupta *et al.* [11]. For more properties of the PRHM model, the readers are referred to the recent review article by Gupta and Gupta [10].

3.3 Proportional Cumulant Model

We have defined the class of PCMs through the relation (7). It is already observed that if X is the base random variable with PDF $f_X(\cdot)$, and the characteristic function of the random variable with PDF $f_X(\cdot)$ and the characteristic function of the random variable Y satisfies (7), it may not be possible always to obtain explicitly the PDF of Y. Although we may not obtain the PDF of Y, but the mean, $\mu(\cdot)$, variance, $\sigma^2(\cdot)$, coefficient of variation, $CV(\cdot)$, measure of skewness, $\gamma_1(\cdot)$, and kurtosis, $\gamma_2(\cdot)$ of Y can be obtained in terms of the corresponding quantities of X. For example, $\mu(Y) = \alpha\mu(X)$, $\sigma^2(Y) = \alpha\sigma^2(X)$, $CV(Y) = CV(X)/\sqrt{\alpha}$, $\gamma_1(Y) = \gamma_1(X)/\sqrt{\alpha}$ and $\gamma_2(Y) = \gamma_2(X)/\alpha$. Some of the points are very clear from these relations. If X is symmetric about zero, Y will also be symmetric about zero for all α. If X is skewed, Y is approximately symmetric for large α.

3.4 Proportional Odd Model

The random variable Y belongs to the POM for the base line random variable X, if their PDF and CDF satisfy (10). Therefore, the PDF of Y is the weighted version of the PDF of X with the weight function

$$w(x) = \left(\frac{1}{F_X(x) + \alpha \bar{F}_X(x)}\right)^2. \tag{20}$$

Since $w(x)$ is an increasing or decreasing function of x for $\alpha > (<)1$, therefore, the shape of PDFs of X and Y satisfy similar relations as the PRHM. From (10), it is observed that the hazard function of X and Y satisfy the following relation;

$$h_Y(x) = \sqrt{w(x)}h_X(x). \tag{21}$$

Therefore, if X has a constant or increasing hazard function, then for $\alpha > 1$, Y also has increasing hazard function. If X has a constant or decreasing hazard function, Y also has a decreasing hazard function if $\alpha < 1$. Studying the weight function (20), it can be observed by simple calculation that the POM family is a TP_2 family. Therefore, several ordering properties, like hazard rate ordering, likelihood ratio ordering or stochastic ordering easily follow from this result. Several ordering results, which were established by Marshal and Olkin [19] can be obtained as special cases of this general result. For further results on ordering for this model see Kirmani and Gupta [18]. Another interesting point can be observed by taking $t \to \infty$ that the tail behavior of $f_X(\cdot)$ and $f_Y(\cdot)$ or $h_X(\cdot)$ and $h_Y(\cdot)$ are very similar.

3.5 Power Transformation Model

If X is the base line distribution, then Y belongs to the PTM of X if their distribution functions or density functions satisfy relations (12). From (12), it is clear that if $f_X(\cdot)$ is a decreasing function, then $f_Y(\cdot)$ is also a decreasing function if $\alpha < 1$. From (13) it is clear that Y has an increasing (decreasing) hazard function if X has an increasing (decreasing) hazard function, provided $\alpha > (<)1$. Similarly, if X has a decreasing hazard function then Y also has a decreasing hazard function of $\alpha < 1$. It may have non-monotone hazard function also. Comparing the two distribution functions, as given in (12), we can conclude that Y has a thinner tail than X if $\alpha > 1$ and vice versa if $\alpha < 1$. The p-th percentile points of X and Y have the relation $x_p = y_p^{\alpha}$. Moreover, it can be easily seen that

$$E(Y^{\beta}) = E(X^{\frac{\beta}{\alpha}}); \quad \beta > 0. \tag{22}$$

Therefore, if the moments of X are known, the moments of Y can also easily obtained. The corresponding L-moments also can be easily obtained similarly as PRHM.

3.6 Azzalini's Skewed Model

For the base random variable X, ASM is defined through (15). If the random variable Y belongs to ASM, then the PDF of Y is a weighted version of the PDF of X and the weight function is

$$w(x) = F_X(\alpha x). \tag{23}$$

Therefore, the weight function is an increasing function for all α. It implies that if the PDF of X is unimodal then Mode$(Y) >$ Mode(X). In this case even if X is symmetric, Y will be always skewed. It is observed that the skewness increases as α increases to ∞. It easily follows that the family of Y is RR_2. Moreover, if the PDF of X is log-concave, then the PDF of Y also will be always log-concave. Therefore the hazard function of Y will be always an increasing function. Note that $\alpha = 1$ for ASM, matches with the corresponding PRHM when $\alpha = 2$. It is not very surprising, because both of them can be obtained as the maximum of the two $i.i.d.$ random variables.

4. Estimation of the Unknown Parameter(s)

In this section we briefly discuss about the estimation of the unknown parameters of the six models generated by different methods. We will consider two cases separately, namely when (i) base distribution is completely known, (ii) base distribution also has some unknown parameter (s). First let us consider the case when the base distribution is completely known.

When the base distribution is known, note that in case of PHM or PRHM we make the data transformation as $Y = \bar{F}(X)$ or $Y = F(X)$ respectively. In both the cases Y follows Beta distribution with one unknown parameter α. Extensive work has been done in the literature on the statistical inferences of the unknown parameter (s) of the Beta distribution, see for example Johnson *et al.* [16]. Note that in both cases, the MLE of α can be obtained in explicit form in terms of the observed sample $x_1, \ldots, x_n$ as

$$\widehat{\alpha} = -\frac{n}{\sum_{i=1}^{n} \ln \bar{F}(x_i)} \qquad \text{and} \qquad \widehat{\alpha} = -\frac{n}{\sum_{i=1}^{n} \ln F(x_i)}$$

respectively. The distribution of $\widehat{\alpha}$ and the corresponding exact confidence interval can be easily obtained using the properties of Beta distribution.

In case of PCM, the MLE may not be obtained, since the PDF of Y is not known explicitly. But if the first moment of the base distribution exists, the moment estimator of α can be obtained by solving non-linear equation. In case of POM or PTM, since the PDF of Y can be obtained in terms of the PDF of X, the MLE of α can be obtained. But in most of the cases the MLE has to be obtained by non-linear optimization method. Since in both the cases, the distribution function of Y also can be written in terms of the distribution function of X, therefore, percentile estimator (PE), least squares estimator (LSE) or the weighted least squares estimator (WLSE) also can be obtained by solving non-linear equation in both the cases, see for example Gupta and Kundu [13]. In case of ASM, it is well known that the estimation of the unknown parameter is a challenging problem. The MLE may not even exist with positive probability moreover it is not known how the other estimators behave in this case.

Now we discuss the case, when the base distribution may have unknown parameter (s). In case of PHM or PRHM, the MLEs can be obtained only by non-linear optimization method. In most of the cases, the explicit solutions may not be possible. In these cases it is possible to obtain PEs, LSEs or WLSEs again by using some non-linear optimization method. As before, in case PCM, the moment estimators can be obtained by solving non-linear equations if the corresponding moments of the base distribution exist. The corresponding L-moment estimators also can be obtained. In case of POM or PTM, it is possible to obtain the MLEs, PEs, LSE and WLSEs. Of course, in case of ASM, similar problem as mentioned before also exist in this case also.

5. Example

Just for illustrative purposes, in this section we consider an example taking the base distribution as one-parameter exponential distribution and apply all the six methods to generate different models. We can take any other base distribution also. It is assumed that X has a exponential distribution with mean $\frac{1}{\lambda}$, *i.e.*,

$$f_X(x) = \lambda e^{-\lambda x}; \quad x > 0. \tag{24}$$

Note that if the base line PDF is (24), then the PDF of Y for the PHM becomes

$$f_Y(y) = \alpha\lambda e^{-\alpha\lambda y}; \quad y > 0. \tag{25}$$

In this case Y also has exponential distribution and therefore, by this method no shape parameter can be introduced in this case. If we consider PRHM, then the PDF of Y becomes

$$f_Y(y) = \alpha\lambda \left(1 - e^{-\lambda y}\right)^{\alpha-1} e^{-\lambda y}; \quad y > 0. \tag{26}$$

The distribution of Y is known as the generalized exponential(GE) distribution and it was originally introduced by the authors in Gupta and Kundu [12]. It is well known that α plays the role of the shape parameter and λ is the scale parameter. It is observed that the GE distribution can be used quite effectively for analyzing any skewed data in place of gamma or Weibull distribution. It may have both increasing and decreasing hazard functions. It has many properties which are quite close to the gamma distribution. The GE family is a TP_2 family. Extensive work has been done on GE distribution, the readers are referred to the review article by Gupta and Kundu [14] in this respect.

In case of PCM, since $\phi_X(t) = (1 - it/\lambda)^{-1}$ for $-\infty < t < \infty$, therefore

$$\phi_Y(t) = \left(1 - \frac{it}{\lambda}\right)^{-\alpha}. \tag{27}$$

Clearly Y follows gamma distribution with shape and scale parameters as α and λ respectively. Therefore, if the base line distribution is exponential, then PCM introduces a new shape parameter. The properties of Y has been extensively studied in the literature, see for example Johnson *et al.* [15].

Now let us consider the POM. In this case the PDF of Y becomes

$$f_Y(y) = \frac{\alpha\lambda e^{-\lambda y}}{\left(1 - (1-\alpha)e^{-\lambda y}\right)^2}; \quad y > 0. \tag{28}$$

Recently, (28) has been introduced by Marshall and Olkin [19]. Shapes of $f_Y(\cdot)$ for different α are provided in Figure 1. The shapes of the PDFs can be either decreasing or unimodal. It is clear that like gamma, GE or Weibull PDFs, (28) can also take different shapes. Therefore, it also can be used for analyzing skewed data. Interestingly unlike GE, Weibull or gamma, when the PDF is unimodal, here the PDF does not start at 0. Therefore, if the data consists of high early observations, then this might be a very useful model. It may have also increasing and decreasing hazard functions. The shapes of the hazard function, different moments and different ordering properties have been discussed in details by Marshall and Olkin [19]. Examining the weight function, it can be easily verified that it is a TP_2 family. The MLEs of α and λ can be obtained by non-linear optimization only. Since the distribution function of Y in this case

$$F_Y(y) = \frac{1 - e^{-\lambda y}}{1 - e^{-\lambda y} + \alpha e^{-\lambda y}}; \quad y > 0 \tag{29}$$

is in explicit form, therefore PEs, LSEs and WLSEs of α and λ can be obtained similarly as Gupta and Kundu [13]. The properties of the different estimators are not known. Extensive simulations are needed similarly as Gupta and Kundu [13] to compare the performances of the different estimators. Work is in progress it will be reported later.

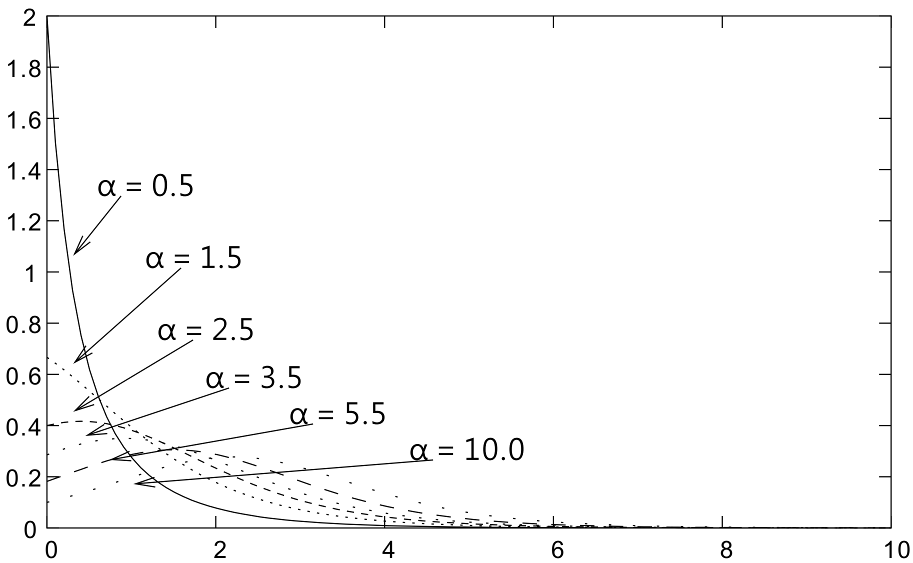

Figure 1 The plot of (28) for different values of α.

Now we consider the PTM. In this case the PDF of Y becomes;

$$f_Y(y) = \alpha\lambda y^{\alpha-1} e^{-\lambda y^{\alpha}}; \quad y > 0. \tag{30}$$

Therefore, Y follows Weibull distribution with the shape and scale parameters as α and λ respectively. Clearly, for the exponential base distribution PTM also introduces a shape parameter. Extensive work has been done on Weibull distribution, the details are available in Johnson *et al.* [15].

Finally we consider ASM, and when the base distribution is exponential, the PDF of Y becomes;

$$f_Y(y) = \frac{\alpha+1}{\alpha}\lambda e^{-\lambda y}\left(1 - e^{-\alpha\lambda y}\right); \quad y > 0. \tag{31}$$

The model (31) has not been studied in the literature. Note that the model (31) for $\alpha = 1$, matches with the GE distribution with $\alpha = 2$. The shapes of the density functions are provided in Figure 2.

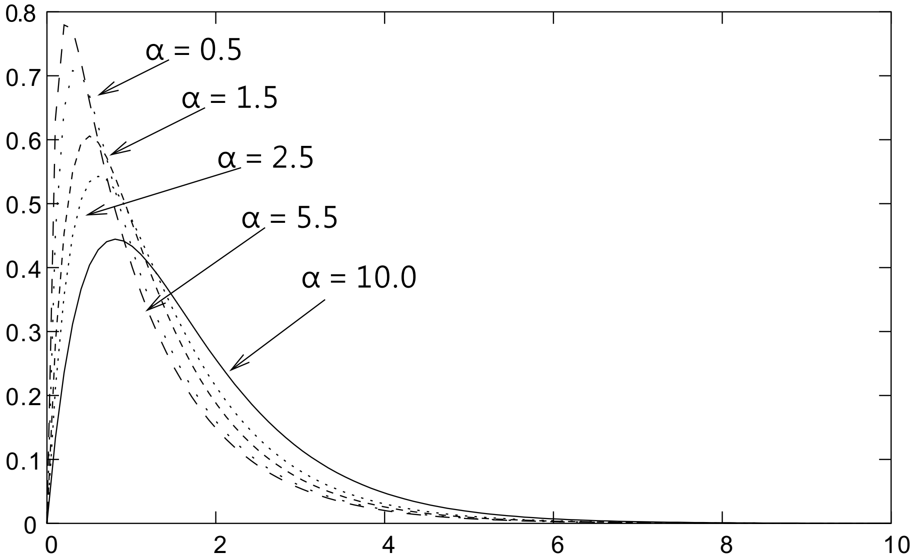

Figure 2 The plot of (31) for different values of α.

Clearly, ASM introduces a shape parameter in the model. From the shapes of the PDFs it is clear that this model also can be used to analyze skewed data. The PDF is always unimodal and the mode is at $\log(1+\alpha)/(\lambda\alpha)$. As $\alpha \to \infty$, the PDF of (31) converges to exponential (base) distribution. Unlike the other previous families, in this case the base distribution is not a member of the family. It can be obtained only as a limiting case.

The hazard function of Y in this case is

$$h(y) = \frac{(\alpha+1)\lambda\left(1 - e^{-\alpha\lambda y}\right)}{\left(\alpha + 1 - \lambda e^{-\alpha\lambda y}\right)}; \quad y > 0 \tag{32}$$

and as it has been observed before that the hazard function of Y is always an increasing function. The moment generating function of Y is

$$m_Y(t) = E(e^{tY}) = \frac{\alpha+1}{\alpha}\left[\left(1 - \frac{t}{\lambda}\right)^{-1} - \frac{1}{1+\alpha}\left(1 - \frac{t}{\lambda(1+\alpha)}\right)^{-1}\right]$$

for $-\lambda(1+\alpha) < t < \lambda(1+\alpha)$. Therefore, all the moments of Y exist. Moreover, the family is RR_2, therefore all the ordering properties hold in this case. The distribution function of Y becomes;

$$F_Y(y) = \frac{\alpha+1}{\alpha}\left[1 - e^{-\lambda y} - \frac{1}{\alpha}\left(1 - e^{-\lambda(1+\alpha)y}\right)\right],$$

which is in explicit form. The MLEs of α and λ can not be obtained in explicit form. It has to be obtained by solving two non-linear equations. The moment estimators, L-moment estimators, PEs, LSEs and WLSEs can be obtained. The detailed comparison of the different estimators are in progress and it will be reported elsewhere.

Now for comparison purposes, we have provided in a tabular form the different characteristics of the six families generated from the exponential base distribution in Table 1. We have assumed $\lambda = 1$, without loss of generality.

Table 1 The different characteristics of the six families of distributions. Here D = decreasing, I = increasing, U = unimodal, C = constant, N = neither

Model	Range of α	Shapes of the Density Function	Shapes of the Hazard Function	Mean	Mode	Ordering Relation
PHM	$\alpha > 0$	D (from α to 0)	C (at α)	$1/\alpha$	0	RR_2
PRHM	$0 < \alpha < 1$ $\alpha = 1$ $\alpha > 1$	D (from ∞ to 0) D (from α to 0) U (starts from 0)	D (from ∞ to 1) C (at α) I (from 0 to 1)	$\psi(\alpha+1) - \psi(1)$	0 0 $\log(\alpha)$	TP_2
PCM	$0 < \alpha < 1$ $\alpha = 1$ $\alpha > 1$	D (from ∞ to 0) D (from α to 0) U (starts from 0)	D (from ∞ to 1) C (at α) I (from 0 to 1)	α	0 0 $\alpha - 1$	TP_2
POM	$0 < \alpha < 1$ $1 \le \alpha < 2$ $\alpha \ge 2$	D (from $1/\alpha$ to 0) U (starts from $1/\alpha$) U (starts from $1/\alpha$)	D (from $1/\alpha$ to 0) I (from $1/\alpha$ to 1) I (from $1/\alpha$ to 1)	$-\dfrac{\alpha\log(\alpha)}{1-\alpha}$	0 0 $\log(\alpha - 1)$	TP_2
PTM	$0 < \alpha < 1$ $\alpha = 1$ $\alpha > 1$	D (from ∞ to 0) D (from α to 0) U (starts from 0)	D (from ∞ to 0) C (at α) I (from 0 to ∞)	$\Gamma(1 + 1/\alpha)$	0 0 $((\alpha-1)/\alpha)^{1/\alpha}$	N
ASM	$\alpha > 0$	U (starts from 0)	I (from 0 to 1)	$(\alpha+2)/(\alpha+1)$	$\log(1+\alpha)/\alpha$	RR_2

6. Applying the Methods More Than Once

So far we have discussed six different methods to introduce a shape parameter in a model. Now in this section we will discuss what will happen if we apply same methods twice or different methods sequentially.

It may be easily observed that Method 1 to Method 5, all have stability property, *i.e.*, if the methods are applied more than once then it will be within the same families. In case of Method 6, it is very difficult to apply the method more than once in general. Moreover, it may not be stable always. Note that for the exponential ($\lambda = 1$), base distribution if Method 6 is applied twice, with parameters α and β respectively, then the new model has the density function;

$$f_Z(z) = ce^{-y}\left(1 - e^{-\alpha y}\right)\left(1 - e^{-y\beta} - \frac{1}{\alpha}\left(1 - e^{-(1+\alpha)\beta y}\right)\right), \tag{33}$$

where c is the normalizing constant. Clearly, (31) and (33) are not the same families.

Now we apply different methods sequentially taking the base distribution as exponential with $\lambda = 1$. Note that if we apply Method 1 first then in this case it does not introduce any new shape parameter. Therefore, applying Method 1 first and then applying any other methods will not produce more than one shape parameter. Now let us discuss few examples which will produce more than one shape parameter, other examples may be similarly obtained. In all cases considered below, we apply the first method with the shape parameter $\alpha > 0$ and the second method with the shape parameter $\beta > 0$.

Method 2 & Method 5: If we apply Method 2 first and then Method 5, the resulting model has the PDF,

$$f_Z(z) = \alpha\beta \left(1 - e^{-z^{\beta}}\right)^{\alpha-1} e^{-z^{\beta}} z^{-\beta-1}, \qquad z > 0. \tag{34}$$

This is the exponentiated Weibull model introduced by Mudholkar and Srivastava [20]. It has two shape parameters, several properties of this distribution have been discussed by several authors, see for example Mudholkar and Srivastava [20], Mudholkar *et al.* [21]. Interestingly, in this case if we apply the Method 5 first and then Method 2, then also it will produce the same family. Therefore, Method 2 and Method 5 are commutative in this case, although in general any two methods may not be so.

Method 3 & Method 2: If the Methods 3 and 2 are applied sequentially, then the resulting model has the PDF

$$f_Z(z) = \frac{\beta}{(\Gamma(\alpha))^{\beta}} \left(\Gamma(\alpha, z)\right)^{\alpha\beta-1} z^{\alpha-1} e^{-z}; \quad z > 0. \tag{35}$$

Here

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \quad \text{and} \quad \Gamma(\alpha, z) = \int_0^{z} x^{\alpha-1} e^{-x} dx,$$

are the gamma function and the incomplete gamma function respectively. The PDF (35) in this case represents the exponentiated gamma distribution introduced by Gupta *et al.* (1998). Here also α and β are two shape parameters. It may be mentioned that here two methods are not commutative.

Method 3 & Method 5: Applying Method 3 first and then Method 5, the PDF of the resulting model becomes;

$$f_Z(z) = \frac{\beta}{\Gamma(\alpha)} z^{\alpha\beta-1} e^{-z^{\beta}}; \quad z > 0. \tag{36}$$

The distribution of Z is known as the generalized gamma distribution. It has also two shape parameters and this model has been well studied in the literature, see for example Johnson *et al.* [16]. In this case also the two methods are not commutative.

Method 5 & Method 4: If the Method 5 is applied first and then Method 4, the resulting model will have the following PDF;

$$f_Z(z) = \frac{\beta\alpha z^{\alpha-1} e^{-z^{\alpha}}}{\left(e^{-z^{\alpha}} + \beta(1 - e^{-z^{\alpha}})\right)^2}, \quad z > 0. \tag{37}$$

This model was introduced and several properties were discussed by Marshal and Olkin [19]. It also has two shape parameters and in this the two methods are commutative.

Method 2 & Method 4: First Method 2 and then Method 4 provides the following PDF;

$$f_Z(z) = \frac{\alpha\beta\,(1-e^{-z})^{\alpha-1}\,e^{-z}}{\left((1-e^{-z})^{\alpha} + \beta(1-(1-e^{-z})^{\alpha})\right)^2}, \quad z > 0. \tag{38}$$

The model (38) has two shape parameters. Now if we apply Method 4 first and then Method 2, the resulting PDF becomes;

$$f_Z(z) = \alpha\beta e^{-z}\frac{(1-e^{-z})^{\beta-1}}{(1-e^{-z}+\alpha e^{-z})^{\beta+1}}; \quad z > 0. \tag{39}$$

Both the models (38) and (39) are new two-parameter models. It is natural that both of them can be very good alternatives to exponentiated Weibull or exponentiated gamma distributions. Their properties need to be investigated.

From the above discussions it is clear that combining two different methods several other new models with two different shape parameters can be obtained. It not surprising that more than two methods also can be used sequentially to generate models with more than two shape parameters. For example applying Method 2, Method 4 and Method 5 sequentially or Method 4, Method 2 and Method 5 sequentially, for $\gamma > 0$, we obtain the following two models;

$$f_Z(z) = \frac{\alpha\beta\gamma z^{\gamma-1}\left(1-e^{-z^{\gamma}}\right)^{\alpha-1} e^{-z^{\gamma}}}{\left((1-e^{-z})^{\alpha} + \beta(1-(1-e^{-z^{\gamma}})^{\alpha})\right)^2}, \quad z > 0, \tag{40}$$

and

$$f_Z(z) = \alpha\beta\gamma z^{\gamma-1} e^{-z^{\gamma}}\frac{\left(1-e^{-z^{\gamma}}\right)^{\beta-1}}{(1-e^{-z^{\gamma}}+\alpha e^{-y^{\gamma}})^{\beta+1}}; \quad z > 0. \tag{41}$$

The families (40) and (41) have three shape parameters each. Similarly several other new distribution with different shape parameters can be generated.

7. Conclusions

In this paper we have considered six different methods of introducing a shape parameter in a model. It is observed that each of them has its own interpretations. Therefore, a practitioner may use his/ her prior knowledge to choose a particular family. We have also discussed structural properties of the different models. We have illustrated different methods by taking a specific example with the base distribution as exponential. Several interesting properties have been observed in these six families. It is observed that all the six families except the Weibull family are either TP_2 or RR_2. Their PDFs and hazard functions have several common features but they have their own characteristics also.

Moreover, it is observed that it is possible to apply our methods more than once and it may produce more than one shape parameters. Interestingly it is observed that most of the methods are not commutative. But in general it is observed that Method 5 is commutative with Method 1 and Method 2. We have seen several existing models can be obtained by combining different methods and various new models with more than one shape parameters can be generated. Preliminary investigation indicates that some of these models have quite interesting properties. More investigations are in progress on these new distributions.

References

[1] Azzalini, A. (1985). A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, **12**, 171-178.

[2] Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones, *Statistica*, **46**, 199 - 208.

[3] Arnold, B. C. and Beaver, R. J. (2000). Hidden truncation models, *Sankhya*, Ser A, **62**, 23-35.

[4] Arnold, B. C. and Bever, R. J. (2002). Skewed multivariate models related to hidden truncation and/or selected reporting, *Test*, **11**, 7-54.

[5] Box, G. and Cox, D. (1964). An analysis of transformations, *Journal of the Royal Statistical Society*, Ser B, **26**, 211-246.

[6] Cox, D. (1972). Regression models and life tables, *Journal of the Royal Statistical Society*, Ser B, **34**, 187-200.

[7] Ferrera, J. T. A. S and Steel, M. F. J. (2006). A constructive representation of univariate skewed distributions, *Journal of the American Statistical Association*, **101**, 823-829.

[8] Genton, M. G. (2004). Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality, Editor, Chapman & Hall/CRS, New York.

[9] Gupta, R. C. and Gupta, R. D. (2004). Generalized skew normal distribution, *Test*, **13**, 510-524.

[10] Gupta, R. C. and Gupta, R. D. (2007). Proportional reversed hazard model and its applications, *Journal of Statistical Planning and Inference*, **137**, 3525 - 3536.

[11] Gupta, R. C., Gupta, P. L., and Gupta, R. D. (1998). Modeling failure time data by Lehmann alternatives, *Communications in Statistics-Theory and Methods*, **33**, 3009-3031.

[12] Gupta, R. D. and Kundu, D. (1999). Generalized exponential distributions, *Australian and New Zealand Journal of Statistics*, **41**, 173-188.

[13] Gupta, R. D. and Kundu, D. (2003). Generalized exponential distributions: different methods of estimation, *Journal of Statistical computation and Simulation*, **69**, 315-338.

[14] Gupta, R. D. and Kundu, D. (2007). Generalized exponential distribution; existing methods and recent developments, *Journal of Statistical Planning and Inference*, **137**, 3537-3547.

[15] Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). Continuous Univariate Distribution, Vol. 1, second edition, John Wiley and Sons, New York.

[16] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994), Continuous Univariate Distribution, Vol. 2, second edition, John Wiley and Sons, New York.

[17] Jones, M. C. (2004). Families of distributions arising of order statistics, *Test*, **13**, 1-43.

[18] Kirmani, S. N. U. A. and Gupta, R. C. (2001). On the proportional odds model in survival analysis, *Annals of the Institute of Statistical Mathematics*, **53**, 203 -216.

[19] Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distribution with application to the exponential and Weibull families, *Biometrika*, **84**, 641-652.

[20] Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure data, *IEEE Transactions on Reliability*, **42**, 299-302.

[21] Mudholkar, G. S., Srivastava, D. K. and Freimer, M. (1995). The exponentiated Weibull family: a reanalysis of the bus motor failure data, *Technometrics*, **37**, 436-445.

[22] Shaked, M. and Shantikumar, J. G. (1994). Stochastic Ordering and Their Applications, Academic Press, San Diego.