Journal of Mahani Mathematical Research
Shahid Bahonar University of Kerman
Print ISSN: 2251-7952 Online ISSN: 2645-4505

TWO WEIGHTED MODELS FOR SKEWED DATA

P. Hasanalipour, M. Razmkhah ✉, and G.R. Mohtashami Borzadaran

Article type: Research Article
(Received: 25 October 2024, Received in revised form 10 April 2025)
(Accepted: 20 May 2025, Published Online: 21 May 2025)

Abstract. The skewed and weighted distributions may be considered as nice alternatives to fit a real data set in practical situations in which the well-known distributions are not suitable. The weighted distributions are first discussed and then two extensions of weighted models are proposed in this paper to analyze the skew data. The flexibility of the models is studied in view of the moment skewness coefficient for some cases. Finally, two real data sets are used to illustrate the results of the paper.

Keywords: Proportional hazard rate model, Skew symmetric distribution, Skewness coefficient, Weight function.
2020 MSC: 62E15, 62G30.

1. Introduction

The weighted distributions are generally used in the statistical modelling of data that some parts of them are missed or an unequal-chance sampling process is performed. The main idea of using weighted models was first expressed by Fisher [11] and then formally defined by Rao [28]. Weighted distributions were proposed to model the sampling methods in which the sampling probabilities are proportional to a weighted function. Using the theory of weighted distributions, flexibility of a distribution can be increased by considering an appropriate weight. The random variable X with probability density function (pdf) $f(\cdot)$ is said to have the structure of a weighted model if it has the following pdf

$$g(x) = \frac{w(x)f(x)}{E[w(X)]}, \; x \in R, \tag{1}$$

where $w(x)$ is a weight function with finite expectation, $E[w(X)] < \infty$. The weighted models with $w(x) = \varphi(F(x))$ are of great importance, where $F(\cdot)$ is the cumulative distribution function (cdf) of the underlying population.

The proportional hazard rate is a well-known special case of a weighted model with $w(x) = [\bar{F}(x)]^{\alpha}$. This model is also known as an α model in the literature. Similarly, the proportional reversed hazard rate models is derived by taking $w(x) = [F(x)]^{\alpha}$. As a more general weighted model Jones [21] considered a model with weight $w(x) = (F(x))^{a-1}(1 - F(x))^{b-1}$. In Jones model, if $a = i$ and $b = (n-i+1)$, the distribution of the ith order statistic in the sample of size

✉ razmkhah_m@um.ac.ir, ORCID: 0000-0002-6661-2822
https://doi.org/10.22103/jmmr.2025.24265.1714

Publisher: Shahid Bahonar University of Kerman
How to cite: P. Hasanalipour, M. Razmkhah, G.R. Mohtashami Borzadaran, *Two weighted models for skewed data*, J. Mahani Math. Res. 2025; 14(2): 353-372.

n is obtained. Many authors have studied special cases of the Jones model by using specific distributions instead of $F(x)$. Nadaraja and Kotz [26] introduced the exponential beta model by considering the exponential distribution. Also, the beta-normal model is obtained by considering $F(x)$ as the cdf of standard normal distribution; see Gupta and Nadaraja [13]. On the other hand, in the situations in which the observations are skewed, a more flexible distribution may be created by adding the appropriate weight to a symmetric distribution. Azzalini [5] introduced the skew-normal (SN) distribution with pdf

$$f(x;\lambda) = 2\phi(x)\Phi(\lambda x), \; x \in (-\infty, \infty), \tag{2}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ stand for the pdf and cdf of standard normal distribution, respectively, and λ is a tuning parameter of skewness. Indeed, the distribution (2) follows a weighted model (1) with $w(x) = 2\Phi(\lambda x)$. Many authors have introduced new models of skew distributions by considering different weight functions. For example, Arellano-Valle [2] proposed the skew-generalized normal distribution by taking $w(x) = \Phi(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}})$. The Kumaraswamy skew-normal distribution was introduced by Mameli [23]. Beranger et al. [8] studied some properties of the univariate extended SN distribution. Recently, some inferential analyses have been derived by some authors. Gui and Guo [14] considered the problem of estimating the location and scale parameters of the SN distribution. The Fisher information in order statistics of SN distribution was investigated by Hasanalipour and Razmkhah [15]. Testing skew-symmetry based on extreme ranked set sampling was also discussed by Hasanalipour and Razmkhah [16]. A more detailed literature review is presented in the next section.

Based on the idea of weighted distributions, two new models are proposed in this paper to deal with the skew data. These model put a specific pdf together with a different cdf in two various approaches. The first one looks like model (2), though, the second investigates that if a proportional hazard rate model might be used as an alternative of a skew distribution. This motivates us to study these two special weighted model as candidates for skew data. Their flexibility in skewness is investigated in detail when the underlying distributions are normal, Laplace or logistic.

The rest of the paper is organized as follows. A literature review of the relationship between weighted models and skew distributions is proposed in Section 2. Two extensions of weighted models are discussed in Section 3. To illustrate the results of the paper, a real data set is analyzed in Section 4. Eventually, some conclusions are stated in Section 5.

2. A literature review of weighted models and skew distributions

Most of proposed skew-symmetric distributions belong to class of weighted distributions. The weighted density functions of a number of skew models with different weights are given in Table 1. In this table, the $\phi(\cdot)$ and $\Phi(\cdot)$ stand for the pdf and cdf of standard normal distribution, respectively.

It is of great importance to note that there are different approaches to define a suitable weighted distribution for skew data. For example, Beta-skew distributions are derived by substituting skew distributions as $F(x)$ in Jones model. Mameli and Musio [24] studied the beta skew-normal distribution with parameter λ which has the pdf below

$$g^B_{\Phi(x;\lambda)}(x;\lambda,a,b) = \frac{2}{B(a,b)}\big(\Phi(x;\lambda)\big)^{a-1}\big(1-\Phi(x;\lambda)\big)^{b-1}\phi(x)\Phi(\lambda x),\ \ a,b>0.$$

The beta skew-generalized normal distribution was introduced by Hasanalipour et al. [17] with the pdf

$$\begin{aligned} g^B_{\Phi(x;\lambda_1,\lambda_2)}(x;\lambda_1,\lambda_2,a,b) &= \frac{2}{B(a,b)}\big(\Phi(x;\lambda_1,\lambda_2)\big)^{a-1}\big(1-\Phi(x;\lambda_1,\lambda_2)\big)^{b-1} \\ &\quad \times\phi(x)\Phi(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}), \end{aligned}$$

where λ_1 and λ_2 are the skewness parameters. Also, Basalamah et al. [7] introduced the beta skew-t model with pdf

$$f(x;\lambda,r,a,b) = \frac{1}{B(a,b)} f(x;\lambda,r)F^{a-1}(x;\lambda,r)(1-F(x;\lambda,r))^{b-1},$$

where $f(x;\lambda,r)$ and $F(x;\lambda,r)$ are the pdf and cdf of the t-student distribution with r degrees of freedom, respectively. Aleem et. al [1] introduced a class of Modified weighted Weibull distribution and its properties. A Brief review, perspective and characterizations of weighted distributions was introduced by Saghir et al. [30]. Moreover, the alpha-skew and the alpha-beta-skew distributions are obtained by taking $w(x) = (1-\alpha x)^2+1$ and $w(x) = (1-\alpha x-\beta x^3)^2+1$ in model (1), where α and β are real-valued parameters. Recently, Das et al. [9] introduced a new flexible alpha-skew normal distribution. Also, Hazarika et al. [19] proposed a multimodal alpha-skew normal distribution.

In the next section, we first propose two new extensions of weighted models which are useful for the skew data, then, their flexibilities are compared in view of skewness criterion.

Table 1. Some skew models with different weights.

Reference	model	Weight density function
Azzalini [5]	SN	$f(x;\lambda)=2\phi(x)\Phi(\lambda x)$
Arellano-Valleh [2]	Generalized SN	$f(x;\lambda_1,\lambda_2)=2\phi(x)\Phi(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}),\ \lambda_2\geq 0$
Gomez [12]	Skew-Laplace	$f(x;\lambda)=\frac{1}{2}e^{-\frac{\lvert x\rvert}{2}}\Phi(\lambda x)$
Gomez [12]	Skew-Logistic	$f(x;\lambda)=\frac{2e^{-x}}{(1+e^{-x})^2}\Phi(\lambda x)$
Gomez [12]	Skew-t	$f(x;\lambda)=\frac{2\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}}(1+\frac{x^2}{\nu})^{-\frac{(1+\nu)}{2}}\Phi(\lambda x)$
Sharafi and Behboodian [31]	The Balakrishnan SN	$f_n(x;\lambda)=c_n(\lambda)\phi(x)\Phi^n(\lambda x)$
Jamalizade et al. [20]	A two-parameter generalized SN	$f(x;\lambda_1,\lambda_2)=c(\lambda_1,\lambda_2)\phi(x)\Phi(\lambda_1 x)\Phi(\lambda_2 x)$
Bahrami et al. [6]	A two-parameter Balakrishnan SN	$f_{n,m}(x;\lambda_1,\lambda_2)=\frac{1}{c_{n,m}(\lambda_1,\lambda_2)}\phi(x)\Phi^n(\lambda_1 x)\Phi^m(\lambda_2 x)$
Kazemi et al. [22]	A generalization of the SN	$f(x;n,m)=mn\phi(x)\Phi^{m-1}(x)[1-\Phi^m(x)]^{n-1},\ x\in R,\ n,m>0$
Hasanalipour and Sharafi [18]	A new generalized balakrishnan SN	$f_n(x;\lambda_1,\lambda_2)=c_n(\lambda_1,\lambda_2)\phi(x)\Phi^n(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}})$
Fathi-Vajargah and Hasanalipour [10]	Bivariate generalized SN	$f_n(x,y;\lambda_1,\lambda_2)=c_n(\lambda_1,\lambda_2)\phi(x)\phi(y)\Phi^n(\frac{\lambda_1 xy}{\sqrt{1+\lambda_2(xy)^2}})$
Asgharzadeh et al. [3]	Generalized skew–Logistic	$f(x;\alpha,\lambda)=\frac{2}{B^2(\alpha,\alpha)}\frac{e^{-\alpha x}}{(1+e^{-x})^{2\alpha}}\int_0^{\frac{1}{1+e^{-\lambda x}}}t^{\alpha-1}(1-t)^{\alpha-1}dt,\ \alpha>0$
Rasekhi et al. [29]	A flexible generalization of SN	$f(x;\alpha,\lambda)=2\frac{1+\alpha x^2}{1+\alpha}\phi(x)\Phi(\lambda x),\ x,\lambda\in R,\ \alpha\geq 0$
Asgharzadeh et al. [4]	Balakrishnan skew–Logistic	$f_n(x;\lambda)=c_n(\lambda)\frac{e^{-x}}{(1+e^{-x})^2}(\frac{1}{1+e^{-\lambda x}})^n,\ n\geq 1$
Mirzadeh and Iranmanesh [25]	A new class of skew–Logistic	$f(x;\lambda)=\frac{\lambda e^{-x}}{[1-e^{-\lambda}][1+e^{-x}]^2}e^{-\frac{\lambda}{1+e^{-x}}}$
Naghibi et al. [27]	A New Weighted SN Model	$f(x;\lambda,\beta)=2\frac{\lvert x\rvert^\beta}{\Gamma(\frac{\beta+1}{2})2^{\frac{\beta+1}{2}}}e^{-\frac{x^2}{2}}\Phi(\lambda x)$

3. Models' description and their properties

Here, two weighted models are proposed which are useful in dealing with the skew data. Assume X has the pdf $f(\cdot)$, and $G(\cdot)$ is any continuous cdf. Let us define two extended weighted models with pdfs

$$g_1(x,\lambda) = \frac{f(x)G(\lambda x)}{E\big(G(\lambda X)\big)} \tag{3}$$

and

$$g_2(x,\lambda) = \frac{f(x)\big(G(x)\big)^{\lambda}}{E\Big(\big(G(X)\big)^{\lambda}\Big)} \tag{4}$$

which are correspond to the weight functions $w_1(x) = G(\lambda x)$ and $w_2(x) = \big(\big(G(x)\big)^{\lambda}$, respectively. By substituting the cdf and pdf of different distributions instead of $G(\cdot)$ and $f(\cdot)$, a class of weighted distributions is created. The SN distribution, introduced by Azzalini [5] is obtained by taking $f \equiv \phi$ and $G \equiv \Phi$ in the model (3), where $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of standard normal distribution, respectively. Also, in a special case where the pdf $f(\cdot)$ is the derivative of the cdf $G(\cdot)$, the model (4) for $\lambda > 0$ turns into a class of proportional reversed hazard rate distributions. Therefore, the proposed models are in fact some generalizations of proportional hazard rate model and skew-symmetric distributions. To compare the flexibility of the proposed models in view of skewness, the moment coefficient of skewness is used. For the weighted random variable X_w by weight $w(\cdot)$, this criterion is defined as

$$S_w(\lambda) = E\bigg(\frac{X_w - \mu_w}{\sigma_w}\bigg)^3, \tag{5}$$

where μ_w and σ_w are mean and standard deviation of X_w, respectively. So, the skewness of the models (3) and (4) are denoted by $S_{w_1}(\lambda)$ and $S_{w_2}(\lambda)$, respectively.

Note that the pdf $f(\cdot)$ and cdf $G(\cdot)$ of either symmetric or asymmetric distributions may be used in the models (3) and (4). So, different cases are investigated in the sequel.

3.1. Symmetric underlying distributions.

Here, some weighted distributions based on symmetric underlying distributions are studied. Consider the normal, Laplace and logistic distributions with the pdfs

$$\begin{aligned} f_1(x) &= \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},\ x \in R \\ f_2(x) &= \frac{1}{2} e^{-|x|},\ x \in R \\ f_3(x) &= \frac{e^{-x}}{(1+e^{-x})^2},\ x \in R \end{aligned}$$

respectively. Note that these distributions are all symmetric with different heavy tails. At first, we assume that both of $f(\cdot)$ and $G(\cdot)$ in the models (3) and (4) are the pdf and cdf of the same distribution. It is of interest to find how the parameter λ affect on the model skewness. Figure 1 shows the behavior of $S_{w_1}(\lambda)$ and $S_{w_2}(\lambda)$ with respect to the variation of the parameter λ when $f(\cdot)$ and $G(\cdot)$ are both follow normal, logistic or Laplace distributions.

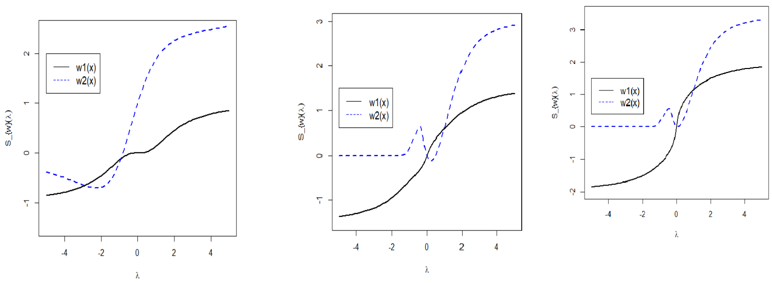

FIGURE 1. Skewness of the weighted models (3) and (4) when $f(\cdot)$ and $G(\cdot)$ follow normal (left), logistic (middle) or Laplace (right) distributions.

From Figure 1, the following results are deduced for the case that both $f(\cdot)$ and $G(\cdot)$ follow the same distribution:

- For all cases, the $S_{w_1}(\lambda)$ is an odd function. More precisely, for $\lambda > 0$ ($\lambda < 0$) the model (3) is skew to the right (left), and the absolute skewness is the same when λ takes away from zero to either left or right.
- In the model (4), the sign of the parameter λ is not directly influenced on the positive or negative values of skewness. That is, sometimes model skewness is negative (positive) for $\lambda > 0$ ($\lambda < 0$).
- For negative values of λ, the model (4) is not flexible such that $S_{w_2}(\lambda) \to 0$ when $\lambda \to -\infty$.
- In models (3) and (4), when the pdf and cdf of the normal distribution are used as $f(\cdot)$ and $G(\cdot)$, respectively, we have $S_{w_2}(\lambda) > S_{w_1}(\lambda)$, at least for positive values of λ. This means that the skewness of model (4) is greater than model (3) for $\lambda > 0$. But for Laplace and logistic distributions, $S_{w_2}(\lambda) > S_{w_1}(\lambda)$ for $\lambda > 1$.
- When $\mid \lambda \mid \to \infty$, the amount of skewness for both models tends to a constant limit.

To more investigations about the flexibility of the proposed models, let us determine the range of skewness of each model by the following interval

$$(S_w^l, S_w^u) = \left(\inf_{\lambda} S_w(\lambda), \sup_{\lambda} S_w(\lambda) \right),$$

where $S_w(\lambda)$ is computed using (5). The range of skewness are reported in Table 2, for some choices of $f(\cdot)$ and $G(\cdot)$.

Table 2. The range of skewness of the models (3) and (4) for some choices of $f(\cdot)$ and $G(\cdot)$.

$G(\cdot)$	$f(\cdot)$	$(S^l_{w_1}, S^u_{w_1})$	$(S^l_{w_2}, S^u_{w_2})$
Normal	Normal	(-0.995, 0.995)	(-0.931, 2.530)
Normal	Logistic	(-1.540, 1.540)	(-0.062, 2.735)
Normal	Laplace	(-2, 2)	(0, 3.110)
Logistic	Normal	(-0.995, 0.995)	(-0.223, 3.212)
Logistic	Logistic	(-1.540, 1.540)	(-0.125, 2.924)
Logistic	Laplace	(-2, 2)	(0, 3.523)
Laplace	Normal	(-0.995, 0.995)	(-0.314, 2.747)
Laplace	Logistic	(-1.540, 1.540)	(-0.165, 2.825)
Laplace	Laplace	(-2, 2)	(0, 3.264)

Since the models may be constructed based on different choices of pdfs and cdfs, to simply interpret the results the length of skewness interval is denoted by $LS(dist_1, dist_2)$ when the pdf of distribution $dist_1$ and the cdf of distribution $dist_2$ are used in a given model. Also, Nor, Lap and Log stand for normal, Laplace and logistic distributions, respectively. From Table 2, the following results are obtained:

- In model (3), when $w_1(x) = G(\lambda x)$, for fixed $G(\cdot)$, the range of skewness changes by varying $f(\cdot)$, but, this interval remains the same by varying $G(\cdot)$ when $f(\cdot)$ is fixed. Moreover, the flexibility of model (3) by using the pdf of Laplace distribution is more than logistic, and it is more than normal distribution. Precisely,

$$\begin{aligned}
LS\big(Nor, Nor\big) &= LS\big(Nor, Log\big) = LS\big(Nor, Lap\big),\\
LS\big(Log, Nor\big) &= LS\big(Log, Log\big) = LS\big(Log, Lap\big),\\
LS\big(Lap, Nor\big) &= LS\big(Lap, Log\big) = LS\big(Lap, Lap\big),\\
LS\big(Lap, Nor\big) &> LS\big(Log, Nor\big) > LS\big(Nor, Nor\big).
\end{aligned}$$

- In model (4), when $w_2(x) = G^{\lambda}(x)$, for fixed $G(\cdot)$, using the pdf of normal distribution provides more flexibility in skewness compared to pdfs of logistic and Laplace distributions.
- For the skew-to-right data, using the pdf of Laplace distribution along the cdf of either normal, logistic or Laplace distribution in model (4) may lead to a suitable skew model.
- For the skew-to-right data, model (4) is more flexible than model (3).
- The range of skewness in model (3) is symmetric around zero, but in model (4) it is asymmetric and tends to positive values.

Figures 2-4 show the plots of pdfs of skew distributions introduced in Table 2 for some choices of skewness parameter λ based on models (3) (left plots) and (4) (right plots) when $G(\cdot)$ is the cdf of normal, logistic or Laplace distributions, respectively. The figures show that by increasing the skewness parameter, model (4) is more flexible than model (3) for the skew-to right data.

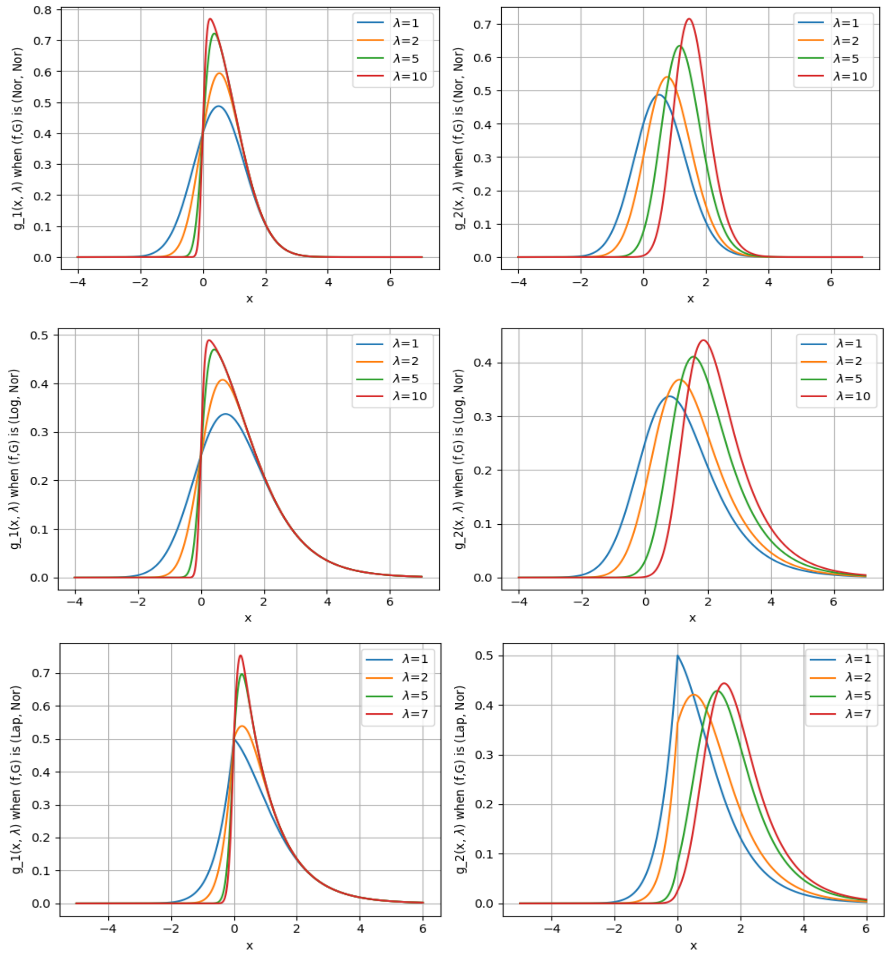

FIGURE 2. Plots of pdf of skew distributions introduced in Table 2 based on modes (3) (left) and (4) (right) when $G(\cdot)$ is the cdf of normal distribution.

3.2. **Asymmetric underlying distributions.** Here, we investigate about how the amount of skewness will change in models (3) and (4) when either $f(\cdot)$ or $G(\cdot)$ are the pdf or cdf of asymmetric distributions, respectively. Toward this end, we consider the standard exponential and Weibull distributions with the pdfs

$$f_4(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & x \le 0 \end{cases} \quad \text{and } f_5(x) = \begin{cases} 2xe^{-x^2}, & x > 0, \\ 0, & x \le 0, \end{cases}$$

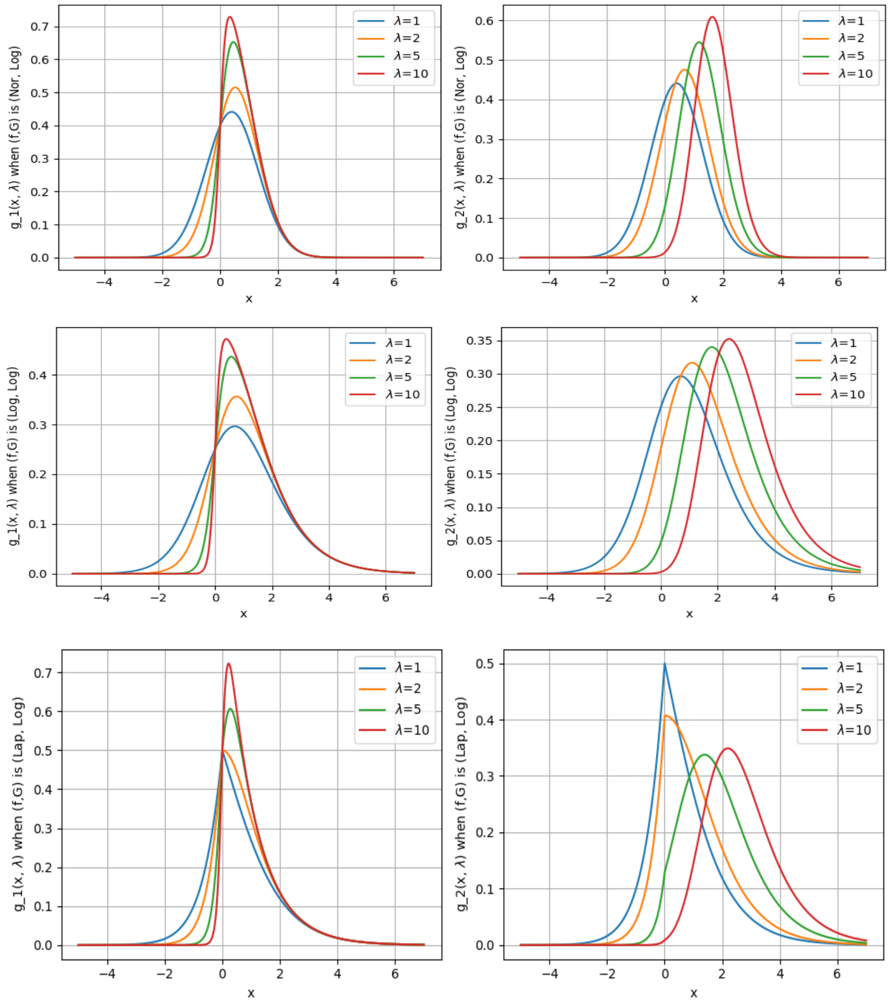

Figure 3. Plots of pdf of skew distributions introduced in Table 2 based on modes (3) (left) and (4) (right) when $G(\cdot)$ is the cdf of logistic distribution.

respectively. Then, the following three cases are studied:
Case I $f(\cdot)$ is a symmetric pdf and $G(\cdot)$ is the cdf of an asymmetric distribution.
Case II $f(\cdot)$ is an asymmetric pdf and $G(\cdot)$ is the cdf of a symmetric distribution.
Case III $f(\cdot)$ and $G(\cdot)$ are both the pdf and cdf of asymmetric distributions, respectively.

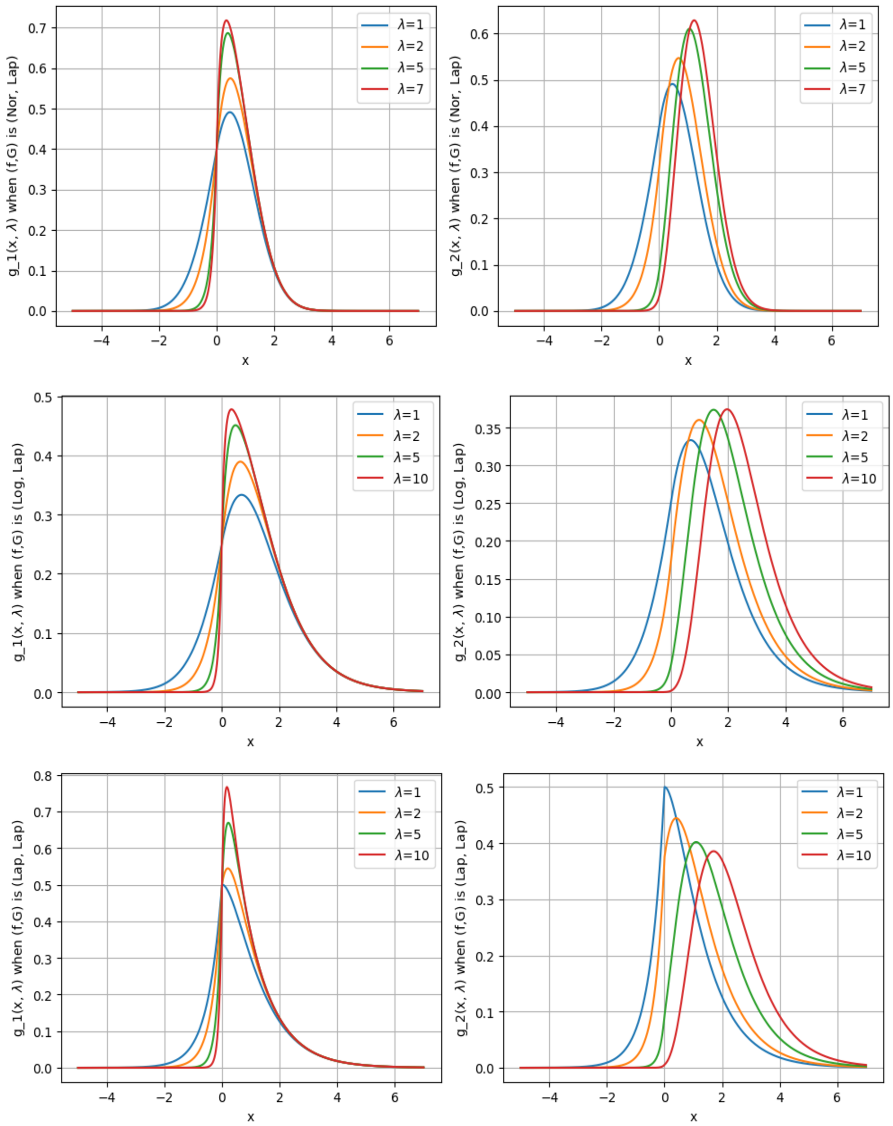

FIGURE 4. Plots of pdf of Skew distributions introduced in Table 2 based on modes (3) (left) and (4) (right) when $G(\cdot)$ is the cdf of Laplace distribution.

Table 3 shows ranges of skewness of models (3) and (4) for the Case I for different choices of $f(\cdot)$ and $G(\cdot)$. From this table, it is observed that:

- For fixed $G(\cdot)$ (either exponential or Weibull), the length of skewness range for model (3) by using the pdf of Laplace distribution as $f(\cdot)$ is larger than logistic and it is larger than normal distribution. That is,

$$\begin{aligned} LS\big(Lap, Exp\big) &> LS\big(Log, Exp\big) > LS\big(Nor, Exp\big), \\ LS\big(Lap, Wei\big) &> LS\big(Log, Wei\big) > LS\big(Nor, Wei\big), \end{aligned}$$

where Exp and Wei stand for exponential and Weibull distributions, respectively.
- The range of skewness of model (3) is symmetric around zero, though, it is totally positive for model (4). So, it is suggested to use model (4) if the data are skew-to-right.
- For fixed $f(\cdot)$, there is not a significant difference between using either exponential or Weibull distribution as $G(\cdot)$ in model (3), though, using exponential distribution as $G(\cdot)$ in model (4) is more flexible than using Weibull distribution.

Table 3. The range of skewness of models (3) and (4) for Case I.

$G(\cdot)$	$f(\cdot)$	$(S^l_{w_1}, S^u_{w_1})$	$(S^l_{w_2}, S^u_{w_2})$
Exponential	Normal	(-1.004, 1.004)	(0, 3.713)
Weibull	Normal	(-1.022, 1.022)	(0, 2.834)
Exponential	Laplace	(-2, 2)	(0, 3.640)
Weibull	Laplace	(-2, 2)	(0, 3.110)
Exponential	Logistic	(-1.555, 1.555)	(0, 2.972)
Weibull	Logistic	(-1.586, 1.586)	(0, 2.798)

The ranges of skewness values of models (3) and (4) for the Case II are shown in Table 4. From this table, the following results are observed

- The length of skewness interval for model (3) is more than model (4). In other words, the model (3) is more flexible than model (4).
- For model (3), we get

$$\begin{aligned} LS\big(Exp, Lap\big) &> LS\big(Exp, Nor\big) > LS\big(Exp, Log\big), \\ LS\big(Wei, Nor\big) &> LS\big(Wei, Lap\big) > LS\big(Wei, Log\big). \end{aligned}$$

- For model (4), we have

$$\begin{aligned} LS\big(Exp, Log\big) &> LS\big(Exp, Lap\big) > LS\big(Exp, Nor\big), \\ LS\big(Wei, Log\big) &> LS\big(Wei, Nor\big) > LS\big(Wei, Lap\big). \end{aligned}$$

- Using the pdf of weibull distribution and the cdf of normal distribution in model (3), creates the most flexibility in the model skewness. Also, in model (4), the most flexibility is related to the use of the pdf of exponential distribution and the cdf of logistic distribution.
- The ranges of skewness in both of models (3) and (4) are asymmetric around zero and tend to positive values. This is because an asymmetric pdf is used as the pdf $f(\cdot)$ in these models.

Table 4. The range of skewness of models (3) and (4) for Case II.

$f(\cdot)$	$G(\cdot)$	$(S^l_{w_1}, S^u_{w_1})$	$(S^l_{w_2}, S^u_{w_2})$
Exponential	Normal	(-0.450, 4.981)	(-0.573,2.524)
Exponential	Laplace	(-0.436,5.640)	(-0.595,2.638)
Exponential	Logistic	(-0.402,4.637)	(-0.606,2.826)
Weibull	Normal	(-0.995,5.933)	(-0.989,1.797)
Weibull	Laplace	(-0.992,5.132)	(-0.833,1.997)
Weibull	Logistic	(-0.997,4.771)	(-0.986,2.269)

From Table 5, which shows skewness values for the Case III, it is seen that:

- For fixed $f(\cdot)$ and $G(\cdot)$, the length of skewness interval in model (3) is larger than model (4). That is, the model (3) is more flexible than model (4).
- For all pairs of $\big(f(\cdot), G(\cdot)\big)$, models (3) and (4) both may be used for either skew-to-right or skew-to-left data, however, they show more flexibility for skew-to-right data.

The behavior of the skewness ranges in Tables 3 to 6 show that if a symmetric pdf $f(\cdot)$ is used in model (3), the skewness range of the model is also symmetric when $G(\cdot)$ is either symmetric or asymmetric. We prove this fact in the following theorem. Further, we show that $S_{w_1}(\lambda)$ is an even function.

Table 5. The range of skewness of models (3) and (4) for Case III.

$f(\cdot)$	$G(\cdot)$	$(S^l_{w_1}, S^u_{w_1})$	$(S^l_{w_2}, S^u_{w_2})$
Exponential	Exponential	(-0.319,4.224)	(-0.500,2.868)
Exponential	Weibull	(-0.335,4.137)	(-0.565,2.823)
Weibull	Exponential	(-0.974,3.184)	(-0.993,2.602)
Weibull	Weibull	(-0.995,5.666)	(-1.010,2.049)

The plots of pdfs of skew distributions in models (3) and (4) with the pdfs and cdfs as introduced in Tables 3-5 may be plotted similar to Figures 2-4, which are omitted from here to avoid similarity.

Theorem 3.1. *Assume that in model* (3), $f(\cdot)$ *is a symmetric pdf around zero. For any continuous cdf* $G(\cdot)$, *we have*

$$S_{w_1}(\lambda) = S_{w_1}(-\lambda).$$

where λ *is the model parameter, such that* $w_1(x) = G(\lambda x)$.

Proof. Using (3) and (5), we get

$$S_{w_1}(\lambda) = \int_{-\infty}^{\infty} \left(\frac{x - \mu_{w_1}(\lambda)}{\sigma_{w_1}(\lambda)} \right)^3 \frac{f(x)G(\lambda x)}{\psi(\lambda)} dx, \tag{6}$$

where

$$\psi(\lambda) = \int_{-\infty}^{\infty} G(\lambda x) f(x) dx.$$

Note that $\psi(\lambda)$ is an even function when $f(\cdot)$ is symmetric, because

$$\begin{aligned}
\psi(-\lambda) &= \int_{-\infty}^{\infty} G(-\lambda x)f(x)dx \\
&= \int_{-\infty}^{\infty} G(\lambda y)f(-y)dy \\
&= \int_{-\infty}^{\infty} G(\lambda y)f(y)dy \\
&= \psi(\lambda).
\end{aligned}$$

Similarly, it can be shown that

$$\begin{aligned}
\mu_{w_1}(-\lambda) &= \int_{-\infty}^{\infty} \frac{xf(x)G(-\lambda x)}{\psi(-\lambda)}dx \\
&= -\int_{-\infty}^{\infty} \frac{yf(-y)G(\lambda y)}{\psi(-\lambda)}dy \\
&= -\int_{-\infty}^{\infty} \frac{yf(y)G(\lambda y)}{\psi(\lambda)}dy \\
&= -\mu_{w_1}(\lambda),
\end{aligned}$$

where the third equality is obtained from the symmetry of $f(\cdot)$ and evenness of $\psi(\cdot)$. So, $\mu_{w_1}(\lambda)$ is an odd function. Therefore,

$$\begin{aligned}
\sigma^2_{w_1}(-\lambda) &= \int_{-\infty}^{\infty} \Big(x - \mu_{w_1}(-\lambda)\Big)^2 \frac{f(x)G(-\lambda x)}{\psi(-\lambda)}dx \\
&= \int_{-\infty}^{\infty} \Big(-y - \mu_{w_1}(-\lambda)\Big)^2 \frac{f(-y)G(\lambda y)}{\psi(-\lambda)}dy \\
&= -\int_{-\infty}^{\infty} \Big(y - \mu_{w_1}(\lambda)\Big)^2 \frac{f(y)G(\lambda y)}{\psi(\lambda)}dy \\
&= -\sigma^2_{w_1}(\lambda)
\end{aligned}$$

That is, $\sigma^2_{w_1}(\lambda)$ is also an odd function. Using (6) and transforming $y = -x$, it can be shown that

$$\begin{aligned}
S_{w_1}(\lambda) &= \int_{-\infty}^{\infty} \left(\frac{-y - \mu_{w_1}(\lambda)}{\sigma_{w_1}(\lambda)}\right)^3 \frac{f(-y)G(-\lambda y)}{\psi(\lambda)}dy \\
&= \int_{-\infty}^{\infty} -\left(\frac{y - \mu_{w_1}(-\lambda)}{-\sigma_{w_1}(-\lambda)}\right)^3 \frac{f(y)G(-\lambda y)}{\psi(-\lambda)}dy \\
&= S_{w_1}(-\lambda).
\end{aligned}$$

Hence the proof is complete. □

4. **Application to real data sets**

In this section, two real data sets are used to illustrate the proposed models. The strategy is to use the proposed weighted models for skew-symmetric models.

4.1. **Strength of carbon fibers data.** The first data set is originally considered by Badar and Priest (1982). These data contain the strength of single carbon fibers. Single fibers were tested under tension at gauge lengths of 1, 10, 20 and 50 mm. Here, we consider the single fibers data set of 10 mm in gauge lengths with sample size 63. The data are as follows:
1.901, 2.132, 2.203, 2.228, 2.257, 2.350, 2.361, 2.396, 2.397, 2.445, 2.454, 2.474, 2.518, 2.522, 2.525, 2.532, 2.575, 2.614, 2.616, 2.618, 2.624, 2.659, 2.675, 2.738, 2.740, 2.856, 2.917, 2.928, 2.937, 2.937, 2.977, 2.996, 3.030, 3.125, 3.139, 3.145, 3.220, 3.223, 3.235, 3.243, 3.264, 3.272, 3.294, 3.332, 3.346, 3.377, 3.408, 3.435, 3.493, 3.501, 3.537, 3.554, 3.562, 3.628, 3.852, 3.871, 3.886, 3.971, 4.024, 4.027, 4.225, 4.395, 5.020.
The sample mean, variance and skewness are 3.0593, 0.3855 and 0.6278, respectively. So, the data are skew to the right. Gupta and Kunda (2010) showed that skew-logistic distribution is appropriate for these data. From Table 2, it was deduced that there is no difference in flexibility of models (3) and (4) if the cdf of logistic, Laplace or normal distributions are used as $G(\cdot)$ in these models. So, we fix to use the cdf of logistic distribution and investigate the pdf of logistic, Laplace and normal distributions as $f(\cdot)$ in these models. Since the support of all symmetric distributions in Table 2 are real line, the studentized data are used to get better fitted distribution. To compare the performance of the proposed weighted models, values of parameter estimate $\hat{\lambda}$, and some goodness of fit criteria such as AIC (Akaike Information Criterion), KSS (Kolmogorov-Smirnov statistic), the p-value are presented in Table 6. Moreover, the inverse of observed Fisher information measure $I^{-1}(\hat{\lambda})$ has been reported in Table 6 as an approximation to the variance of MLE. Note that under regularity conditions, the MLE is asymptotically normal with mean equal to the true parameter and variance given by the inverse of the Fisher Information. Further, the Q-Q plots for all different cases discussed in Table 6 have been drawn in Figure 5.

Table 6. Some goodness of fit criteria for strength of carbon fibers data and some choices of $f(\cdot)$ in Models (3) and (4).

Model	$f(\cdot)$	$\hat{\lambda}$	$I^{-1}(\hat{\lambda})$	AIC	KSS	p-value
Model (3)	*Log*	0.00008	0.1149	204.6037	0.1694	0.0475
	Lap	2.28417	0.2808	189.6569	0.1448	0.1291
	Nor	0.00006	0.6423	179.7863	0.0971	0.5592
Model (4)	*Log*	0.24382	0.1200	201.8116	0.2087	0.0069
	Lap	2.24381	0.0773	201.8116	0.2087	0.0069
	Nor	0.00772	0.5136	179.7853	0.0985	0.5407

From Table 6 and Figure 5, it is observed that using the pdf of normal distribution along with the cdf of logistic distribution in model (3) provides the best fit among all other cases. However, using the same pdf and cdf in model (4) does not lead to a significant difference.

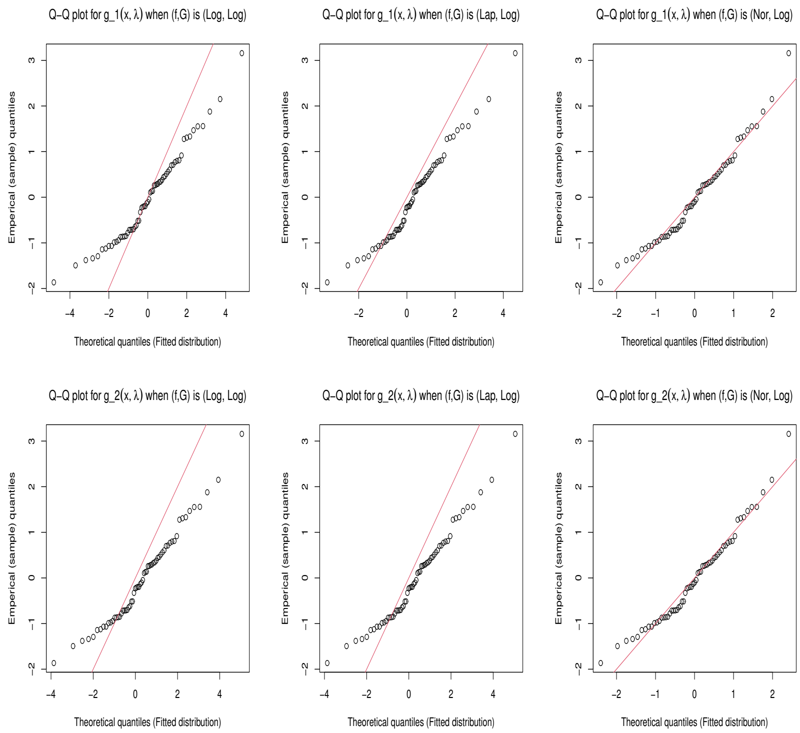

FIGURE 5. Q-Q plots of skew distributions based on models (3) and (4) for the strength of single carbon fibers data.

4.2. **Breaking stress of carbon fibers data.** The second data set is obtained from Nichols and Padget (2006). These data, as presented below, include 100 observations on breaking stress of carbon fibers.
3.70, 2.74, 2.73, 2.50, 3.60, 3.11, 3.27, 2.87, 1.47, 3.11, 4.42, 2.41, 3.19,3.22, 1.69, 3.28, 3.09, 1.87, 3.15, 4.90, 3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.53, 2.67, 2.93, 3.22, 3.39, 2.81, 4.20, 3.33, 2.55, 3.31, 3.31, 2.85, 2.56, 3.56, 3.15, 2.35, 2.55, 2.59, 2.38, 2.81, 2.77, 2.17, 2.83,1.92, 1.41, 3.68, 2.97, 1.36, 0.98, 2.76, 4.91, 3.68, 1.84, 1.59, 3.19, 1.57, 0.81, 5.56, 1.73, 1.59, 2.00,1.22, 1.12, 1.71, 2.17, 1.17, 5.08, 2.48, 1.18, 3.51, 2.17, 1.69, 1.25, 4.38, 1.84, 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.70, 2.03, 1.80, 1.57, 1.08, 2.03, 1.61, 2.12, 1.89, 2.88, 2.82, 2.05, 3.65.
The sample mean, variance and skewness are 2.6214, 1.0279 and 1.0138, respectively. So, these data are skew to the right. Hassan and Abd-Allah (2018)

fitted the exponentiated-Lomax distribution to them. Now, we compare the performance of some weighted models for these data when $f(\cdot)$ or $G(\cdot)$ are the pdf and cdf of either a symmetric or an asymmetric distribution. The logistic, Laplace and normal distributions are used as some candidates for symmetric distributions and exponential distribution is considered as an asymmetric distribution. The results for the Weibull distribution are similar and so they are omitted. Values of parameter estimate $\hat{\lambda}$, the inverse of observed Fisher information and some goodness of fit criteria are presented in Table 7. Moreover, the Q-Q plots for all different cases have been drawn in Figure 6.

Table 7. Some goodness of fit criteria for breaking stress of carbon fibers data for some choices of $f(\cdot)$ and $G(\cdot)$ in Models (3) and (4).

Model	$f(\cdot)$	$G(\cdot)$	$\hat{\lambda}$	$I^{-1}(\hat{\lambda})$	AIC	KSS	p-value
Model (3)	Exp	Log	0.94967	0.2559	473.6023	0.59741	0.00000
	Exp	Lap	0.57997	0.1489	478.3610	0.60438	0.00000
	Exp	Nor	0.58243	0.2170	471.8882	0.59502	0.00000
	Log	Exp	0.10005	0.0288	321.8044	0.22991	0.00005
	Lap	Exp	1.00644	0.0190	381.0012	0.33162	0.00000
	Nor	Exp	0.00006	0.0571	615.4765	0.57843	0.00000
Model (4)	Exp	Log	10.58986	0.5970	292.8269	0.12008	0.11180
	Exp	Lap	15.49155	0.7451	292.2572	0.09723	0.30123
	Exp	Nor	16.35712	1.0513	299.042	0.16911	0.00656
	Log	Exp	5.29202	2.0802	293.4271	0.09632	0.31160
	Lap	Exp	1.82942	2.4125	375.7508	0.38978	0.00000
	Nor	Exp	15.64054	4.9164	414.1316	0.31276	0.00000

From Table 7 and Figure 6, we can see that the model (4) with $(f, G) = (Exp, Lap)$ or (Log, Exp) may be suggested as an almost suitable distribution for the data.

5. Conclusions

In this paper, the use of weighted distributions for modelling skew data was investigated. First, a literature review regarding relationship between weighted models and skew distributions was presented. Then, two weighted models were proposed and their flexibility were discussed for some special cases. The results of the paper are summarized as follows:

- In model (3), for fixed $G(\cdot)$, the range of skewness changes by varying $f(\cdot)$, but, it remains the same by varying $G(\cdot)$ when $f(\cdot)$ is fixed.
- For the skew-to-right data, model (4) is more flexible than model (3). So, it is suggested to use model (4) if the data are skew-to-right.
- For a symmetric pdf $f(\cdot)$ and any continuous cdf $G(\cdot)$ (symmetric or asymmetric), the range of skewness in model (3) is symmetric around zero, but in model (4) it is asymmetric and tends to positive values.
- The skewness range of model (3) is symmetric if a symmetric pdf $f(\cdot)$ is used in the model regardless that $G(\cdot)$ is either symmetric or asymmetric.
- Using the pdf or cdf of either symmetric or asymmetric distributions in models (3) and (4) may be used for either skew-to-right or skew-to-left

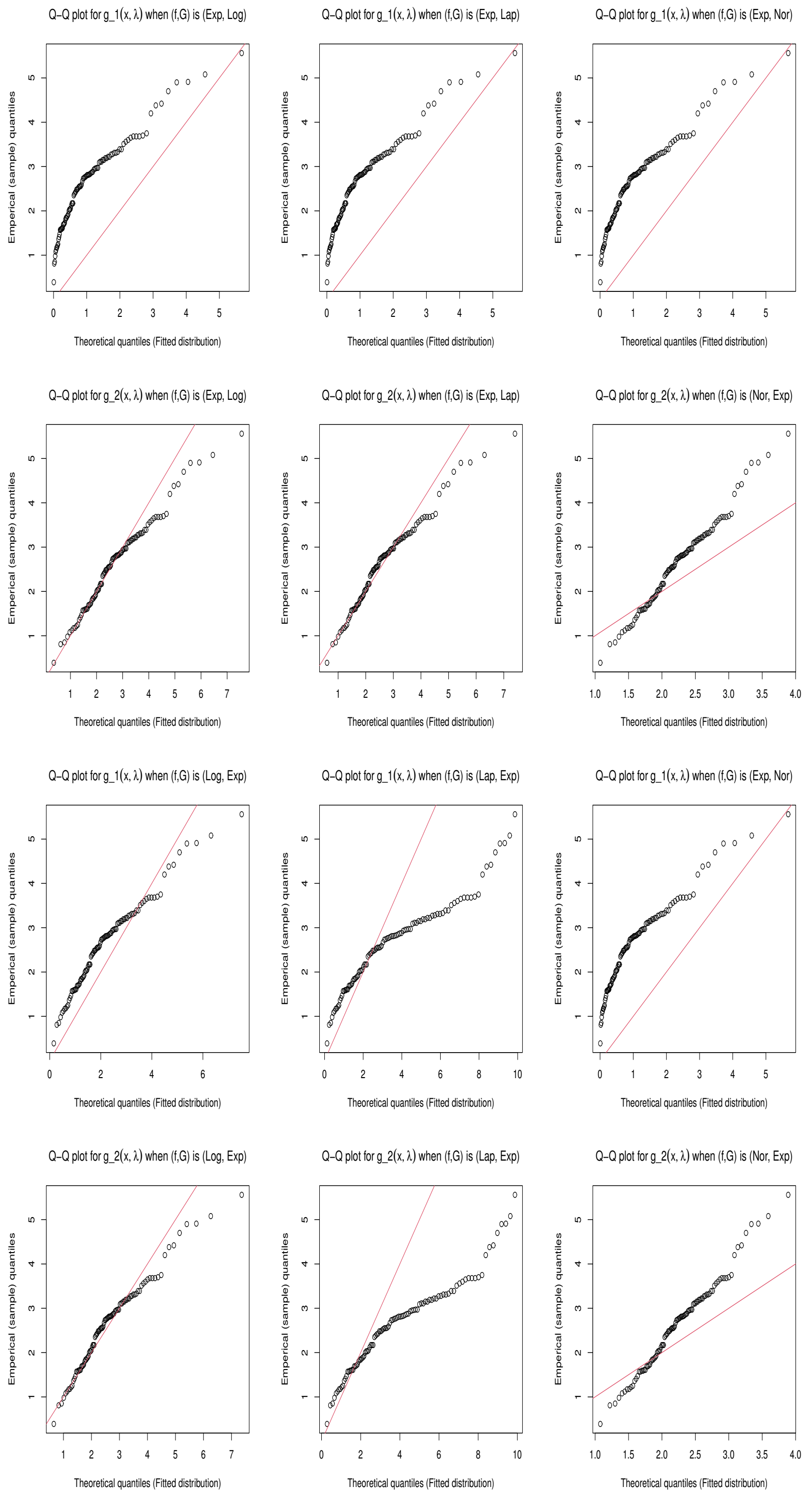

FIGURE 6. Q-Q plots of skew distributions based on models (3) and (4) for breaking stress of carbon fibers data.

data, however, using asymmetric distributions shows more flexibility for skew-to-right data.

- Two real data sets were studied to show the performance of the proposed models. In the first data set, both models could be appropriate. However, only model (4) provided a reasonable fit for the second data set.

Some properties of the proposed weighted model may be considered as future research works. For example, the Fisher information or entropy of these models as well as their order statistics may be investigated. Also, estimating the unknown parameters and testing hypotheses may be of great importance.

Acknowledgement

The authors would like to thank anonymous referees and the associate editor for their useful comments and constructive criticisms on the original version of this manuscript which led to this considerably improved version.

References

[1] Aleem, M., Sufyan, M., & Khan, NS. (2013). A Class of Modified Weighted Weibull Distribution and Its Properties. American Review of Mathematics and Statistics, 1(1), 29-37. https://api.semanticscholar.org/CorpusID:125623946

[2] Arellano-Valle, RB., Gomez, HW., & Quintana, FA. (2004). A new class of skew-normal distribution. Communication in Statistics–Theory and Methods, 33, 1465-1480. https://doi.org/10.1081/STA-120037254

[3] Asgharzadeh, A., Esmaeili, L., & Shih, SH. (2013). A generalized skew logistic distribution. REVSTAT-Statistical Journal, 11(3), 317-338. https://doi.org/10.57805/revstat.v11i3.140

[4] Asgharzadeh, A., Esmaeili, L., & Nadarajah, S. (2016). Balakrishnan skew logistic distribution. Communications in Statistics-Theory and Methods, 42(2), 444-464. https://doi.org/10.1080/03610926.2013.823205

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

[6] Bahrami, W., Agahi, H., & Rangin, H. (2009). A two-parameter Balakrishnan skew-normal distribution. Journal of Statistical Research of Iran, 6(2), 231-242. https://doi.org/10.18869/acadpub.jsri.6.2.231

[7] Basalamah, D., Ning, W., & Gupta, A. (2018). The beta skew t distribution and its properties. Journal of Statistical Theory and Practice, 12(4), 837-860. https://doi.org/10.1080/15598608.2018.1481468

[8] Beranger, B., Padoan, SA., & Sisson, SA. (2019). Extremal properties of the univariate extended skew-normal distribution. Statistics & Probability Letters, 147, 73-82. https://doi.org/10.1016/j.spl.2018.09.018

[9] Das, J., Pathak, D., Hazarika, PJ., Chakraborty, S., & Hamedani, GG. (2023). A new flexible alpha skew normal distribution. Journal of the Indian Society for Probability and Statistics, 24, 485-507. https://doi.org/10.1007/s41096-023-00163-8

[10] Fathi-Vajargah, B., & Hasanalipour, P. (2013). Introducing a novel bivariate generalized skew-symmetric normal distribution. Journal of Mathematics and Computer Science, 61, 266-271. https://doi.org/10.22436/jmcs.07.04.05

[11] Fisher, AA. (1934). The effect of methods of ascertainment upon the estimation of frequencies. Annals of eugenics, 6(1), 13-25. https://doi.org/10.1111/j.1469-1809.1934.tb02105.x

[12] Gomez, HW., Venegas, O., & Bolfarin, H. (2007). Skew-symmetric distributions generated by the distribution function of the normal distribution. Environmetrics: The official journal of the International Environmetrics Society, 18(4), 395-407. https://doi.org/10.1002/env.817

[13] Gupta, A., & Nadarajah, S. (2005). On the moments of the beta normal distribution. Communications in Statistics-Theory and Methods, 31(1), 1-13. https://doi.org/10.1081/STA-120026573

[14] Gui, W., & Guo, L. (2018). Statistical Inference for the Location and Scale Parameters of the Skew Normal Distribution. Indian Journal of Pure and Applied Mathematics, 49, 633-650. https://doi.org/10.1007/s13226-018-0291-6

[15] Hasanalipour, P., & Razmkhah, M. (2021). Inference on skew-normal distribution based on Fisher information in order statistics. Communications in Statistics-Simulation and Computation, 51(4), 1525-1541. https://doi.org/10.1080/03610918.2019.1674869

[16] Hasanalipour, P., & Razmkhah, M. (2022). Testing skew-symmetry based on extreme ranked set sampling. Statistical Papers, 62, 2311-2332. https://doi.org/10.1007/s00362-020-01183-3

[17] Hasanalipour, P., Hassani Oskouei, L., Jabbari Khamnei, H., & Khorshidian, K. (2017). The beta skew-generalized normal distribution. Communications in Statistics-Theory and Methods, 46, 2270-2276. https://doi.org/10.1080/03610926.2015.1041980

[18] Hasanalipour, P., & Sharafi, M. (2012). A new generalized Balakrishnan skew-normal distribution. Statistical Papers, 53, 219-228. https://doi.org/10.1007/s00362-010-0330-9

[19] Hazarika, PJ.,Shah, S., Chakraborty, S., Alizadeh, M., & Hamedani, G. (2023). Multimodal alpha skew normal distribution: A new distribution to model skewed multimodal observations. Pakistan Journal of Statistics and Operation Research, 19, 747-763. https://doi.org/10.18187/pjsor.v19i4.4232

[20] Jamalizadeh, A., Behboodian, J., & Balakrishnan, N. (2008). A two-parameter generalized skew-normal distribution. Statistics & Probability Letters, 78(13), 1722-1726. https://doi.org/10.1016/j.spl.2008.01.006

[21] Jones, M. (2004). Families of distributions arising from distributions of order statistics. Test, 13(1), 1-43. https://doi.org/10.1007/BF02602999

[22] Kazemi, MR., Haghbin, H., & Behboodian, J. (2011). Another generalization of the skew normal distribution. World Applied Sciences Journal, 12(7), 1034-1039.

[23] Mameli, V. (2015). The Kumaraswamy skew-normal distribution. Statistics & Probability Letters, 104, 75-81. https://doi.org/10.1016/j.spl.2015.04.031

[24] Mameli, V., & Musio, M. (2013). A generalization of the skew-normal distribution: the beta skew-normal. Communications in Statistics-Theory and Methods, 42(12), 2229-2244. https://doi.org/10.1007/s10440-006-9017-6

[25] Mirzadeh, S., & Iranmanesh, A. (2019). A new class of skew-logistic distribution. Mathematical Sciences, 13, 375-385. https://doi.org/10.1007/s40096-019-00306-8

[26] Nadarajah, S., & Kotz, S. (2006). Skew distributions generated from different families. Acta Applicanda Mathematica, 91(1), 1-37. https://doi.org/10.1007/s10440-006-9017-6

[27] Naghibi, F., Alavi, S.M.R ., & Chinipardaz, R. (2022). A New Weighted Skew Normal Model. STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING, 10, 1126-1142. https://doi.org/10.19139/soic-2310-5070-1108

[28] Rao, CR. (1965). On discrete distributions arising out of methods of ascertainment. Sankhya: The Indian Journal of Statistics, Series A, 27(2), 311-324.

[29] Rasekhi, M., Chinipardaz, R., & Alavi, SMR. (2016). A flexible generalization of the skew normal distribution based on a weighted normal distribution. Statistical Methods & Applications, 25, 375-394. https://doi.org/10.1007/s10260-015-0337-4

[30] Saghir, A., Hamedani, G., Tazeem, S., & Khadim, A. (2018). Weighted Distributions: A Brief Review, Perspective and Characterizations. International Journal of Statistics and Probability, 7(1), 39-71. https://doi.org/10.5539/ijsp.v6n3p109

[31] Sharafi, M., & Behboodian, J. (2008). The Balakrishnan skew-normal density. Statistical Papers, 49, 769-778. https://doi.org/10.1007/s00362-006-0038-z

P. Hasanalipour
Orcid number: 0000-0002-6989-305x
Department of Statistics
Faculty of Mathematical Sciences
Ferdowsi University of Mashhad
Mashhad, Iran
Email address: P.hasanalipour116919@gmail.com

M. Razmkhah
Orcid number: 0000-0002-6661-2822
Department of Statistics
Faculty of Mathematical Sciences
Ferdowsi University of Mashhad
Mashhad, Iran
Email address: razmkhah_m@um.ac.ir

G.R. Mohtashami Borzadaran
Orcid number: 0000-0002-8841-1386
Department of Statistics
Faculty of Mathematical Sciences
Ferdowsi University of Mashhad
Mashhad, Iran
Email address: grmohtashami@um.ac.ir