**RESEARCH**                                                                 **Open Access**

# *T*-normal family of distributions: a new approach to generalize the normal distribution

Ayman Alzaatreh[1*†], Carl Lee[2†] and Felix Famoye[2†]

*Correspondence:
alzaatreha@apsu.edu
†Equal contributors
[1] Department of Mathematics and Statistics, Austin Peay State University, Clarksville, TN 37044, USA
Full list of author information is available at the end of the article

## Abstract

The idea of generating skewed distributions from normal has been of great interest among researchers for decades. This paper proposes four families of generalized normal distributions using the *T-X* framework. These four families of distributions are named as *T*-normal families arising from the quantile functions of (i) standard exponential, (ii) standard log-logistic, (iii) standard logistic and (iv) standard extreme value distributions. Some general properties including moments, mean deviations and Shannon entropy of the *T*-normal family are studied. Four new generalized normal distributions are developed using the *T*-normal method. Some properties of these four generalized normal distributions are studied in detail. The shapes of the proposed *T*-normal distributions can be symmetric, skewed to the right, skewed to the left, or bimodal. Two data sets, one skewed unimodal and the other bimodal, are fitted by using the generalized *T*-normal distributions.

**AMS 2010 Subject Classification:** 60E05; 62E15; 62P10

**Keywords:** *T-X* distributions; Hazard function; Quantile function; Shannon entropy

## 1 Introduction

The normal distribution is perhaps the most commonly used probability distribution in both statistical theory and applications. The normal distribution was first used by de Moivre (1733) in the literature as an approximation to the binomial distribution. However, the development of the normal distribution by Gauss (1809, 1816) became the standard used in the modern statistics. Hence, the normal distribution is also commonly known as the Gaussian distribution. Properties of the normal distribution have been well developed (e.g., see Johnson et al. 1994; Patel and Read 1996). The distribution also plays an important role in generating new distributions.

Methods for developing generalized normal distributions seemed very limited until Azzalini (1985). A random variable $X_\lambda$ is said to follow the skewed normal distribution $SN(\lambda)$ if the probability density function (PDF) of $X_\lambda$ is $g(x|\lambda) = 2\phi(x)\Phi(\lambda x)$, where $\phi(x)$ and $\Phi(x)$ are $N(0, 1)$ PDF and cumulative distribution function (CDF) respectively. Various extensions of $SN(\lambda)$ have been proposed and studied (e.g., Arellano-Valle et al. 2004; Arnold and Beaver 2002; Arnold et al. 2007; Choudhury and Abdul 2011; Balakrishnan 2002; Gupta and Gupta 2004; Sharafi and Behboodian 2008; Yadegari et al. 2008). For reviews on skewed normal and its generalization, one may refer to Kotz and Vicari (2005) and Lee et al. (2013). Pourahmadi (2007) showed that the skewed normal distribution

$SN(\lambda)$ approaches half-normal as $\lambda \to \infty$. This explains why skewed normal distribution is limited in fitting real data. In order to allow for fitting diverse magnitudes of skewness, various works have been done by introducing different methods to capture the magnitude of the skewness.

Fernández and Steel (1998) introduced a two-piece PDF as $g(x) = \begin{cases} cf(\alpha x), x \geq 0, \\ cf(x/\alpha), x < 0. \end{cases}$, $c > 0$ and $\alpha > 0$, where $f$ is a symmetric PDF defined on $\Re$, which is unimodal and symmetric around 0. When $f$ is normal, it is a generalized skewed normal. Kotz and Vicari (2005) suggested that $\alpha$ and $1/\alpha$ be replaced by $\alpha_1$ and $\alpha_2$ respectively, in order to have more flexibility of controlling skewness. Another general framework that introduces a skew mechanism to symmetric distributions was given by Ferreira and Steel (2006). The corresponding skew family is $g(x|f, q) = f(x)q[F(x)]$, $x \in \Re$. The PDF $g(x|f, q)$ is a weighted version of $f(x)$, with the weight function given by $q[F(x)]$. If $q$ follows the uniform distribution, then, $g = f$. When $f$ is normal, this is a general framework for developing skewed normal distributions.

Eugene et al. (2002) introduced the beta-generated family of distributions with CDF

$$G(x) = \int_0^{F(x)} b(t)dt, \tag{1.1}$$

where $b(t)$ is the PDF of the beta random variable and $F(x)$ is the CDF of any random variable. The corresponding PDF to (1.1) is given by

$$g(x) = \frac{1}{B(\alpha, \beta)} f(x)F^{\alpha-1}(x)(1 - F(x))^{\beta-1}, \quad \alpha, \beta > 0. \tag{1.2}$$

If $F$ is $\Phi$, the CDF of the normal distribution, equation (1.2) defines the beta-normal distribution. If $\alpha$ and $\beta$ are integers, (1.2) is the $\alpha^{\text{th}}$ order statistic of the random sample of size $(\alpha + \beta - 1)$.

The beta-normal distribution can be unimodal or bimodal and it has been applied to fit a variety of real data including bimodal cases (Famoye et al. 2004). The main distinction between the method of skewed normal and the beta-generated normal is that the skewed normal method introduces a skewing mechanism into the normal distribution to generate skewed normal family. The skewness of the distribution is estimated by the skewing parameter. On the other hand, the beta-normal distribution is generated by adding more parameters using beta distribution as the generator. Thus, the skewness is not directly defined by a specific parameter; rather it is the combination of all shape parameters that play the role of measuring skewness. For detailed review about the methods for generating continuous distributions, including the normal distribution, one may refer to Lee et al. (2013).

Alzaatreh et al. (2013) extended the beta generated family and defined the $T\text{-}X(W)$ family. The CDF of the $T\text{-}X(W)$ distribution is $G(x) = \int_a^{W(F(x))} r(t)dt$, where $r(t)$ is the PDF of the random variable $T$ with support $(a, b)$ for $-\infty \leq a < b \leq \infty$. The function $W(F(x))$ is monotonic and absolutely continuous. Aljarrah et al. (2014) took $W(F(x))$ to be the quantile function of a random variable $Y$ and defined the $T\text{-}X\{Y\}$ family as

$$G(x) = \int_a^{Q_Y(F(x))} r(t)dt = R(Q_Y(F(x))), \tag{1.3}$$

where $Q_Y(p)$ is the quantile function of the random variable $Y$. In (1.1), $X$ is used as a random variable having CDF $F(x)$ and then as a random variable having CDF $G(x)$ which may be confusing. This article first gives a unified notation to re-define the $T$-$X\{Y\}$ as $T$-$R\{Y\}$ and proposes several different generalizations of the normal distribution using the $T$-$R\{Y\}$ framework.

Section 2 gives the unified definition of $T$-$R\{Y\}$ family and defines several new generalized normal families. Section 3 gives some general properties of the proposed generalized normal families. Section 4 defines some new generalized normal distributions and studies some of their properties. Section 5 provides some applications to numerical data sets and the paper ends with a short summary and conclusions.

## 2 *T*-normal families of distributions

Let $T$, $R$ and $Y$ be random variables with CDF $F_T(x) = P(T \leq x)$, $F_R(x) = P(R \leq x)$ and $F_Y(x) = P(Y \leq x)$. The corresponding quantile functions are $Q_T(p)$, $Q_R(p)$ and $Q_Y(p)$, where the quantile function is defined as $Q_Z(p) = \inf\{z : F_Z(z) \geq p\}$, $0 < p < 1$. If densities exist, we denote them by $f_T(x)$, $f_R(x)$ and $f_Y(x)$. Now assume the random variable $T \in (a, b)$ and $Y \in (c, d)$, for $-\infty \leq a < b \leq \infty$ and $-\infty \leq c < d \leq \infty$. Following the technique proposed by Aljarrah et al. (2014), the CDF of the random variable $X$ is defined as

$$F_X(x) = \int_a^{Q_Y(F_R(x))} f_T(t)dt = P\left[T \leq Q_Y(F_R(x))\right] = F_T(Q_Y(F_R(x))). \tag{2.1}$$

Note that (2.1) is an alternative expression to (1.3) without using $X$ in two different situations. Hereafter, the family of distributions in (2.1) will be called the $T$-$R\{Y\}$ family of distributions.

**Remark 1.** If $X$ follows the distribution in (2.1), it is easy to see that

(i)  $X \overset{d}{=} Q_R(F_Y(T))$,
(ii)  $Q_X(p) = Q_R(F_Y(Q_T(p)))$,
(iii)  If $T \overset{d}{=} Y$ then $X \overset{d}{=} R$ and
(iv)  If $Y \overset{d}{=} R$ then $X \overset{d}{=} T$.

The corresponding PDF associated with (2.1) is

$$f_X(x) = f_T(Q_Y(F_R(x))) \times Q'_Y(F_R(x)) \times f_R(x), \tag{2.2}$$

where $Q'_Y(F_R) = \frac{d}{dF_R}Q_Y(F_R)$. Using the fact that $Q_Y(F_Y(x)) = x$, it follows that $Q'_Y(F_Y(x)) \times f_Y(x) = 1$ so that $Q'_Y(p) = 1/f_Y(Q_Y(p))$. By taking $p = F_R(x)$, (2.2) reduces to

$$f_X(x) = f_R(x) \times \frac{f_T(Q_Y(F_R(x)))}{f_Y(Q_Y(F_R(x)))}. \tag{2.3}$$

From (2.1) and (2.3), the hazard function of the random variable $X$ can be written as

$$h_X(x) = \frac{f_X(x)}{1 - F_X(x)}$$

$$= \frac{f_R(x)}{1 - F_R(x)} \times \frac{1 - F_Y(Q_Y(F_R(x)))}{f_Y(Q_Y(F_R(x)))} \times \frac{f_T(Q_Y(F_R(x)))}{1 - F_T(Q_Y(F_R(x)))}$$

$$= h_R(x) \times \frac{h_T(Q_Y(F_R(x)))}{h_Y(Q_Y(F_R(x)))}. \tag{2.4}$$

One can see from (2.3) and (2.4) that

$$\frac{f_X(x)}{f_R(x)} = \frac{f_T(Q_Y(F_R(x)))}{f_Y(Q_Y(F_R(x)))} \quad \text{and} \quad \frac{h_X(x)}{h_R(x)} = \frac{h_T(Q_Y(F_R(x)))}{h_Y(Q_Y(F_R(x)))}.$$

Some general properties of the $T$-$R\{Y\}$ family were recently studied in the literature, for more details see Aljarrah et al. (2014). Equivalent expressions to (2.2) - (2.4) are given in Aljarrah et al. (2014) by using the $T$-$X\{Y\}$ notation. Table 1 gives some distributions of the $T$-$R\{Y\}$ families based on quantile functions of some standard forms of distribution and some commonly used random variables $T$. The explicit expression of a $T$-$R\{Y\}$ family can be obtained using (2.3) for different combinations of random variables $T$, $R$, and $Y$.

Several extensions from Table 1 can be made. First, one can use the quantile function of non-standard distributions, such as non-standard exponential, log-logistic, logistic, extreme value, and Weibull. For example, the quantile function of log-logistic is $Q_Y(p) = \alpha(p/(1-p))^{1/\beta}$, $\alpha, \beta > 0$. By using this $Q_Y$ function, two additional parameters corresponding to the log-logistic distribution may be added to the $T$-$R\{$log-logistic$\}$ family. Aljarrah et al. (2014) gave a more detailed list of $T$-$R\{Y\}$ distributions based on quantile functions of non-standard distributions. Secondly, one can introduce exponentiated and scale parameters by replacing $F_T(x)$ by $F_T^\delta(\alpha x)$, $\alpha, \delta > 0$ as well as for $F_R(x)$.

If $R$ is a normal random variable with PDF $f_R(x) = \phi(x)$ and CDF $F_R(x) = \Phi(x)$, then (2.1) gives the $T$-normal$\{Y\}$ family of distributions as

$$F_X(x) = \int_a^{Q_Y(\Phi(x))} f_T(t)dt = F_T(Q_Y(\Phi(x))). \tag{2.5}$$

The corresponding PDF associated with (2.5) is

$$f_X(x) = f_T(Q_Y(\Phi(x))) \times Q'_Y(\Phi(x)) \times \phi(x) = \phi(x) \times \frac{f_T(Q_Y(\Phi(x)))}{f_Y(Q_Y(\Phi(x)))}. \tag{2.6}$$

**Table 1 Families of *T-R{Y}* distributions based on different choices for the random variables *Y\** and *T***

| $Y$ | $Q_Y(p)$ | $T$ | $F_T(x)$ or $f_T(x)$ |
|---|---|---|---|
| (a) Uniform | $p$ | (1) Exponentiated-Exponential | $F_T(x) = (1 - e^{-\lambda x})^\alpha$ |
| (b) Exponential | $-\log(1-p)$ | (2) Weibull | $F_T(x) = 1 - e^{-(x/\gamma)^c}$ |
| (c) Log-logistic | $p/(1-p)$ | (3) Logistic | $F_T(x) = 1 - (1 + e^{-\lambda x})^{-1}$ |
| (d) Logistic | $\log(p/(1-p))$ | (4) Log-logistic | $F_T(x) = 1 - [1 + (x/\alpha)^\beta]^{-1}$ |
| (e) Extreme value | $\log(-\log(1-p))$ | (5) Pareto | $F_T(x) = 1 - (\alpha/x)^\beta$ |
| | | (6) Cauchy | $f_T(x) = \{\pi\beta(1 + [(x-\alpha)/\beta]^2)\}^{-1}$ |
| | | (7) Pascal | $f_T(x) = 0.5\lambda e^{-\lambda|x|}$ |
| | | (8) Gamma | $f_T(x) = x^{\alpha-1}e^{-x/\beta}/[\beta^\alpha\Gamma(\alpha)]$ |

*standard random variable $Y$.

The hazard function of the $T$-normal$\{Y\}$ family is given by $h_X(x) = h_\phi(x) \times \frac{h_T(Q_Y(\Phi(x)))}{h_Y(Q_Y(\Phi(x)))}$, where $h_\phi(x) = \phi(x)/(1 - \Phi(x))$.

The $T$-normal$\{Y\}$ family is a general framework for generating many different generalizations of the normal distribution. Various existing generalizations of normal distributions can be obtained based on this framework. The beta normal (Eugene et al. 2002), Kumaraswamy normal (Cordeiro and de Castro 2011), and generalized beta-generated normal (Alexander et al. 2012) belong to the $T$-normal$\{$standard uniform$\}$ families. The gamma-normal distribution studied by Alzaatreh, et al. (2014) is a member of $T$-normal$\{$standard exponential$\}$ family. For distribution "parsimony", we will focus on the quantile functions of standard distributions in order to limit the number of parameters. Generalizations from using non-standard quantile functions or adding exponentiated and/or scale parameters can be derived in a straightforward manner. In the following, we define the families of generalized normal (GN) distributions, $T$-normal$\{Y\}$, using the standard quantile functions (b)-(e) defined in Table 1.

### 2.1 Family of GN distributions from the quantile function of exponential distribution ($T$-$N\{$exponential$\}$)

By using the quantile function (b) in Table 1: $Q_Y(\Phi(x)) = -\log(1 - \Phi(x))$, the corresponding CDF to (2.5) is

$$F_X(x) = F_T\left\{-\log(1 - \Phi(x))\right\} = F_T\left(H_\phi(x)\right), \tag{2.7}$$

and the corresponding PDF is

$$f_X(x) = \frac{\phi(x)}{1 - \Phi(x)} f_T\left(-\log(1 - \Phi(x))\right) = h_\phi(x) f_T\left(H_\phi(x)\right), \tag{2.8}$$

where $h_\phi(x)$ and $H_\phi(x) = -\log[1 - \Phi(x)]$ are the hazard and cumulative hazard functions for the normal distribution, respectively. Thus, this family of GN distributions is denoted as $T$-$N\{$exponential$\}$, which arises from the "hazard function of the normal distribution".

### 2.2 Family of GN distributions from the quantile function of log-logistic distribution ($T$-$N\{$log-logistic$\}$)

By using the quantile function (c) in Table 1: $Q_Y(\Phi(x)) = \Phi(x)/(1 - \Phi(x))$, the corresponding CDF to (2.5) is

$$F_X(x) = F_T\left\{\Phi(x)/(1 - \Phi(x))\right\}, \tag{2.9}$$

and the corresponding PDF is

$$f_X(x) = \frac{\phi(x)}{(1 - \Phi(x))^2} f_T\left(\frac{\Phi(x)}{1 - \Phi(x)}\right). \tag{2.10}$$

The family of GN distributions in (2.9) is denoted as $T$-$N\{$log-logistic$\}$, which arises from the "odds of the normal distribution".

### 2.3 Family of GN distributions from the quantile function of logistic distribution ($T$-$N\{$logistic$\}$)

By using the quantile function (d) in Table 1: $Q_Y(\Phi(x)) = \log(\Phi(x)/(1 - \Phi(x)))$, the corresponding CDF to (2.5) is

$$F_X(x) = F_T\left\{\log[\Phi(x)/(1 - \Phi(x))]\right\}, \tag{2.11}$$

and the corresponding PDF is

$$f_X(x) = \frac{\phi(x)}{\Phi(x)(1 - \Phi(x))} f_T \left( \log \left\{ \Phi(x)/(1 - \Phi(x)) \right\} \right) = \frac{h_\phi(x)}{\Phi(x)} f_T \left( \log \left\{ \frac{\Phi(x)}{1 - \Phi(x)} \right\} \right).$$

(2.12)

The family of GN distributions in (2.11) is denoted as *T-N*{logistic}, which arises from the "logit function of the normal distribution".

### 2.4 Family of GN distributions from the quantile function of extreme value distribution (*T-N*{extreme value})

By using the quantile function (e) in Table 1: $Q_Y(\Phi(x)) = \log(-\log(1 - \Phi(x)))$, the corresponding CDF to (2.5) is

$$F_X(x) = F_T \left\{ \log[-\log(1 - \Phi(x))] \right\},$$

(2.13)

and the corresponding PDF is

$$f_X(x) = \frac{\phi(x)}{-(1 - \Phi(x)) \log(1 - \Phi(x))} f_T \left\{ \log \left( -\log(1 - \Phi(x)) \right) \right\} = \frac{h_\phi(x)}{H_\phi(x)} f_T (\log(H_\phi(x)).$$

(2.14)

The family of GN distributions in (2.13) is denoted as *T-N*{extreme value}, which arises from the "extreme value function of the normal distribution".

## 3 Some properties of the *T*-normal family of distributions

In this section, some of the general properties of the *T*-normal family will be discussed.

**Lemma 1** (**Transformation**). For any random variable *T* with PDF $f_T(x)$, then the random variable

(i)   $X = \Phi^{-1}(1 - e^{-T})$ follows the distribution of *T-N*{exponential} family in (2.7).
(ii)  $X = \Phi^{-1}(T/(1 + T))$ follows the distribution of *T-N*{log-logistic} family in (2.9).
(iii) $X = \Phi^{-1}\left(e^T/(1 + e^T)\right)$ follows the distribution of *T-N*{logistic} family in (2.11).
(iv)  $X = \Phi^{-1}\left(1 - e^{-e^T}\right)$ follows the distribution of *T-N*{extreme value} family in (2.13).

*Proof.* The result follows immediately from Remark 1(i). □

Lemma 1 gives the relationships between the random variable *X* and the random variable *T*. These relationships can be used to generate random samples from *X* by using *T*. For example, one can simulate the random variable *X* which follows the distribution of *T-N*{exponential} family in (2.7) by first simulating random variable *T* from the PDF $f_T(x)$ and then computing $X = \Phi^{-1}\left(1 - e^{-T}\right)$, which has the CDF $F_X(x)$.

**Lemma 2.** The quantile functions for the (i) *T-N*{exponential}, (ii) *T-N*{log-logistic}, (iii) *T-N*{logistic}, and (iv) *T-N*{extreme value} distributions, are respectively,

(i)  $Q_X(p) = \Phi^{-1}\left\{1 - e^{-Q_T(p)}\right\}$,
(ii) $Q_X(p) = \Phi^{-1}\left\{Q_T(p)/(1 + Q_T(p))\right\}$,

(iii)     $Q_X(p) = \Phi^{-1}\left\{e^{Q_T(p)}/\left(1 + e^{Q_T(p)}\right)\right\}$,

(iv)     $Q_X(p) = \Phi^{-1}\left\{1 - e^{-e^{Q_T(p)}}\right\}$.

*Proof.* The result follows directly from Remark 1(ii).                    □

**Theorem 1.** The mode(s) of the $T$-$N\{Y\}$ family are the solutions of the equation

$$x = \mu + \sigma^2 \phi(x) \left\{ \frac{Q''_Y(\Phi(x))}{Q'_Y(\Phi(x))} + \frac{f'_T(Q_Y(\Phi(x)))}{f_T(Q_Y(\Phi(x)))} Q'_Y(\Phi(x)) \right\}. \tag{3.1}$$

*Proof.* One can show the result in (3.1) by setting the derivative of the equation (2.6) to zero and then using the fact that $\phi'(x) = -\sigma^{-2}(x - \mu)\phi(x)$.                    □

**Corollary 1.** The mode(s) of the (i) $T$-$N\{$exponential$\}$, (ii) $T$-$N\{$log-logistic$\}$, (iii) $T$-$N\{$logistic$\}$, and (iv) $T$-$N\{$extreme value$\}$ distributions, respectively, are the solutions of the equations

$$\text{(i)} \ x = \mu + \sigma^2 h_\phi(x) \left\{ 1 + \frac{f'_T\left(H_\phi(x)\right)}{f_T\left(H_\phi(x)\right)} \right\}, \tag{3.2}$$

$$\text{(ii)} \ x = \mu + \sigma^2 h_\phi(x) \left\{ 2 + \frac{f'_T\left(\Phi(x)/(1 - \Phi(x))\right)}{(1 - \Phi(x))f_T\left(\Phi(x)/(1 - \Phi(x))\right)} \right\},$$

$$\text{(iii)} \ x = \mu + \sigma^2 \frac{h_\phi(x)}{\Phi(x)} \left\{ \frac{f'_T\left(\log\{\Phi(x)/(1 - \Phi(x))\}\right)}{f_T\left(\log\{\Phi(x)/(1 - \Phi(x))\}\right)} + 2\Phi(x) - 1 \right\},$$

$$\text{(iv)} \ x = \mu + \sigma^2 \frac{h_\phi(x)}{H_\phi(x)} \left\{ \frac{f'_T\{\log(H_\phi(x))\}}{f_T\{\log(H_\phi(x))\}} + H_\phi(x) - 1 \right\}.$$

Note that the results in Theorem 1 do not imply that the mode is unique. It is possible that there is more than one mode for some of these GN distributions. For example, the logistic-$N\{$logistic$\}$ distribution given in section 4 is a bimodal distribution. If $T$ follows the gamma distribution with parameters $\alpha$ and $\beta$, equation (3.2) can be simplified to

$$x = \mu + \sigma^2 h_\phi(x) \left\{ 1 + \frac{f'_T\left(H_\phi(x)\right)}{f_T\left(H_\phi(x)\right)} \right\} = \mu + \sigma^2 h_\phi(x) \left[ (\alpha - 1)/H_\phi(x) - \beta^{-1} + 1 \right].$$

This agrees with the result obtained by Alzaatreh et al. (2014) for the gamma-normal distribution.

The entropy of a random variable $X$ is a measure of variation of uncertainty (Rényi 1961). Shannon's entropy for a random variable $X$ with PDF $g(x)$ is defined as $E\left\{-\log\left(g(X)\right)\right\}$.

**Theorem 2.** The Shannon's entropies for the $T$-$N\{Y\}$ family is given by

$$\eta_X = \eta_T + E\left(\log f_Y(T)\right) + \log(\sigma\sqrt{2\pi}) + \frac{1}{2\sigma^2}E(X - \mu)^2. \tag{3.3}$$

*Proof.* Since $X \stackrel{d}{=} Q_R(F_Y(T))$, this implies that $T \stackrel{d}{=} Q_Y(F_R(X))$. Hence, from (2.3) we have $f_X(x) = \frac{f_T(t)}{f_Y(t)} \times f_R(x)$. This implies

$$\eta_X = \eta_T + E\left(\log f_Y(T)\right) - E\left(\log f_R(X)\right). \tag{3.4}$$

For the *T-N{Y}* family we have $f_R(x) = \phi(x)$, so

$$\log(\phi(x)) = -\log(\sigma\sqrt{2\pi}) - [(x - \mu)/\sigma]^2/2. \tag{3.5}$$

The result in (3.3) follows from (3.4) and (3.5). □

**Corollary 2.** The Shannon's entropies for the (i) *T-N{exponential}*, (ii) *T-N{log-logistic}*, (iii) *T-N{logistic}*, and (iv) *T-N{extreme value}* distributions, respectively, are given by

$$\text{(i)} \quad \eta_X = \log(\sigma\sqrt{2\pi}) - \mu_T + \eta_T + E(X - \mu)^2/(2\sigma^2), \tag{3.6}$$

$$\text{(ii)} \quad \eta_X = \log(\sigma\sqrt{2\pi}) - 2E(\log(1 + T)) + \eta_T + E(X - \mu)^2/(2\sigma^2),$$

$$\text{(iii)} \quad \eta_X = \log(\sigma\sqrt{2\pi}) - 2E(\log(1 + e^T)) + \mu_T + \eta_T + E(X - \mu)^2/(2\sigma^2),$$

$$\text{(iv)} \quad \eta_X = \log(\sigma\sqrt{2\pi}) - E(e^T) + \mu_T + \eta_T + E(X - \mu)^2/(2\sigma^2).$$

*Proof.* The results in (i)-(iv) can be easily shown using (3.3) and the fact that $f_Y(T) = e^{-T}$, $(1 + T)^{-2}$, $e^T(1 + e^T)^{-2}$ and $e^T e^{-e^T}$ for exponential, log-logistic, logistic and extreme value, respectively. □

**Theorem 3.** The $r^{\text{th}}$ non-central moments of the (i) *T-N{exponential}*, (ii) *T-N{log-logistic}*, (iii) *T-N{logistic}*, and (iv) *T-N{extreme value}* distributions, respectively, can be expressed as

$$\text{(i)} \quad E(X^r) = \sum_{j=0}^{r} \sum_{k_1,\cdots,k_j=0}^{\infty} \sum_{i=0}^{2s_j+j} 2^{j/2} \sigma^j \mu^{r-j} A(\underline{k}) \binom{r}{j} \binom{2s_j + j}{i} (-2)^i M_T(-i), \tag{3.7}$$

$$\text{(ii)} \quad E(X^r) = \sum_{j=0}^{r} \sum_{k_1,k_2,\cdots,k_j=0}^{\infty} 2^{j/2} \sigma^j \mu^{r-j} A(\underline{k}) \binom{r}{j} E\left\{\left(\frac{T-1}{T+1}\right)^{2s_j+j}\right\}, \tag{3.8}$$

$$\text{(iii)} \quad E(X^r) = \sum_{j=0}^{r} \sum_{k_1,k_2,\cdots,k_j=0}^{\infty} 2^{j/2} \sigma^j \mu^{r-j} A(\underline{k}) \binom{r}{j} E\left\{\left(\frac{e^T}{1+e^T}\right)^{2s_j+j}\right\}, \tag{3.9}$$

$$\text{(iv)} \quad E(X^r) = \sum_{j=0}^{r} \sum_{k_1,\cdots,k_j=0}^{\infty} \sum_{i=0}^{2s_j+j} 2^{j/2} \sigma^j \mu^{r-j} A(\underline{k}) \binom{r}{j} \binom{2s_j + j}{i} (-2)^i M_{e^T}(-i), \tag{3.10}$$

where $A(\underline{k}) = A(k_1, k_2, \cdots, k_j) = (\sqrt{\pi}/2)^{2s_j+j} a_{k_1} a_{k_2} \cdots a_{k_j}$, $s_j = k_1 + k_2 + \cdots + k_j$, $M_T(-i) = E(e^{-iT})$, $a_k = \frac{c_k}{2k+1}$, $c_k = \sum_{j=0}^{k-1} \frac{c_j c_{k-1-j}}{(j+1)(2j+1)}$, and $c_0 = 1$.

*Proof.* We first show (3.7). By using Lemma 1, the $r^{\text{th}}$ moments for the $T$-$N$\{exponential\} distribution can be written as $E(X^r) = E\big(\Phi^{-1}(1 - e^{-T})\big)^r$. Since

$$\Phi^{-1}(1 - e^{-T}) = \sqrt{2}\,\sigma\,\text{erf}^{-1}(1 - 2e^{-T}) + \mu,$$

the $r^{\text{th}}$ moments can be written as

$$E(X^r) = E\Big(\sqrt{2}\,\sigma\,\text{erf}^{-1}(1 - 2e^{-T}) + \mu\Big)^r = \sum_{j=0}^{r}\binom{r}{j}2^{j/2}\sigma^j\,E\left\{\big(\text{erf}^{-1}(1 - e^{-T})\big)^j\right\}\mu^{r-j}. \tag{3.11}$$

On using the series representation for $\text{erf}^{-1}(1 - 2e^{-T})$ (Wolfram.com, 2014), we get $\text{erf}^{-1}(1 - 2e^{-T}) = \sum_{k=0}^{\infty} a_k(\sqrt{\pi}/2)^{2k+1}(1 - 2e^{-T})^{2k+1}$, where $a_k = \frac{c_k}{2k+1}$, $c_k = \sum_{j=0}^{k-1}\frac{c_j c_{k-1-j}}{(j+1)(2j+1)}$, and $c_0 = 1$. This implies

$$\big(\text{erf}^{-1}(1 - 2e^{-T})\big)^j = \sum_{k_1, k_2, \cdots, k_j = 0}^{\infty} A(k_1, k_2, \cdots, k_j)(1 - 2e^{-T})^{2s_j + j}, \tag{3.12}$$

where $A(k_1, k_2, \cdots, k_j) = (\sqrt{\pi}/2)^{2s_j + j}a_{k_1}a_{k_2}\cdots a_{k_j}$ and $s_j = k_1 + k_2 + \cdots + k_j$. By using the binomial expansion on $(1 - 2e^{-T})^{2s_j + j}$, (3.12) can be written as

$$\big(\text{erf}^{-1}(1 - 2e^{-T})\big)^j = \sum_{k_1, k_2, \cdots, k_j = 0}^{\infty}\sum_{i=0}^{2s_j + j} A(k_1, k_2, \cdots, k_j)\binom{2s_j + j}{i}(-2)^i e^{-iT}. \tag{3.13}$$

The result of (3.7) follows by using equation (3.13) in equation (3.11). The results of (3.8)-(3.10) can be obtained by applying the same technique for (3.7). □

If $T$ follows the gamma distribution with parameters $\alpha$ and $\beta$ for the $T$-$N$\{exponential\}, we obtain the term $M_T(-i) = (1 + \beta i)^{-\alpha}$ in (3.7). Thus, (3.7) gives the $r^{\text{th}}$ non-central moment of gamma-$N$\{exponential\} distribution as shown in Alzaatreh et al. (2014).

The deviation from the mean and the deviation from the median are used to measure the dispersion and the spread in a population from the center. The mean deviation from the mean is denoted by $D(\mu)$, and the mean deviation from the median $M$ is denoted by $D(M)$.

**Theorem 4.** $D(\mu)$ and $D(M)$ for each of (i) $T$-$N$\{exponential\}, (ii) $T$-$N$\{log-logistic\}, (iii) $T$-$N$\{logistic\}, and (iv) $T$-$N$\{extreme value\} distributions, respectively, are

$$\text{(i)}\quad D(\mu) = \sqrt{2}\sigma\sum_{k=0}^{\infty}\sum_{i=0}^{2k+1} A(k)\binom{2k+1}{i}(-2)^{i+1}S_{e^{-u}}(\mu, 0, i), \tag{3.14}$$

$$D(M) = \sqrt{2}\sigma\sum_{k=0}^{\infty}\sum_{i=0}^{2k+1} A(k)\binom{2k+1}{i}(-2)^{i+1}S_{e^{-u}}(M, 0, i), \tag{3.15}$$

where $S_\xi(c, a, \alpha) = \int_a^{Q_Y(\Phi(c))}\xi^\alpha f_T(u)du$ and $Q_Y(\Phi(c)) = -\log(1 - \Phi(c))$.

$$\text{(ii)}\quad D(\mu) = -\sqrt{8}\sigma\sum_{k=0}^{\infty} A(k)S_{\frac{u-1}{u+1}}(\mu, 0, 2k+1), \tag{3.16}$$

$$D(M) = -\sqrt{8}\sigma\sum_{k=0}^{\infty} A(k)S_{\frac{u-1}{u+1}}(M, 0, 2k+1), \tag{3.17}$$

where $Q_Y(\Phi(c)) = \Phi(c)/(1 - \Phi(c))$.

$$\text{(iii)} \quad D(\mu) = -\sqrt{8}\sigma \sum_{k=0}^{\infty} A(k) S_{\frac{e^u}{1+e^u}} (\mu, -\infty, 2k+1), \qquad (3.18)$$

$$D(M) = -\sqrt{8}\sigma \sum_{k=0}^{\infty} A(k) S_{\frac{e^u}{1+e^u}} (M, -\infty, 2k+1), \qquad (3.19)$$

where $Q_Y(\Phi(c)) = \log\{\Phi(c)/(1 - \Phi(c))\}$.

$$\text{(iv)} \quad D(\mu) = \sqrt{2}\sigma \sum_{k=0}^{\infty} \sum_{i=0}^{2k+1} A(k) \binom{2k+1}{i} (-2)^i S_{e^{-e^u}} (\mu, -\infty, i), \qquad (3.20)$$

$$D(M) = \sqrt{2}\sigma \sum_{k=0}^{\infty} \sum_{i=0}^{2k+1} A(k) \binom{2k+1}{i} (-2)^i S_{e^{-e^u}} (M, -\infty, i), \qquad (3.21)$$

where $Q_Y(\Phi(c)) = \log\{-\log(1 - \Phi(c))\}$.

*Proof.* By definitions of $D(\mu)$ and $D(M)$,

$$D(\mu) = \int_{-\infty}^{\mu} (\mu - x) f_X(x) dx + \int_{\mu}^{\infty} (x - \mu) f_X(x) dx = 2 \int_{-\infty}^{\mu} (\mu - x) f_X(x) dx$$

$$= 2\mu F_X(\mu) - 2 \int_{-\infty}^{\mu} x f_X(x) dx. \qquad (3.22)$$

$$D(M) = \int_{-\infty}^{M} (M - x) f_X(x) dx + \int_{M}^{\infty} (x - M) f_X(x) dx$$

$$= 2 \int_{-\infty}^{M} (M - x) f_X(x) dx + E(X) - M$$

$$= \mu - 2 \int_{-\infty}^{M} x f_X(x) dx. \qquad (3.23)$$

We first proof the results (3.14) and (3.15) for the $T$-$N$\{exponential\} family. Defining the integral

$$I_c = \int_{-\infty}^{c} x f_X(x) dx = \int_{-\infty}^{c} \frac{x\phi(x)}{1 - \Phi(x)} f_T \{-\log(1 - \Phi(x)\} dx, \qquad (3.24)$$

and using the substitution $u = -\log(1 - \Phi(x))$, (3.24) can be written as

$$I_c = \int_{0}^{-\log(1-\Phi(c))} \Phi^{-1}(1 - e^{-u}) f_T(u) du. \qquad (3.25)$$

By using similar approach as in Theorem 3, the equation (3.25) can be written as

$$I_c = \mu F_X(c) + \sqrt{2}\sigma \sum_{k=0}^{\infty} \sum_{i=0}^{2k+1} A(k) \binom{2k+1}{i} (-2)^i S_{e^{-u}} (c, 0, i), \qquad (3.26)$$

where $A(k)$ is defined in the proof of Theorem 3, $S_{\xi}(c, a, \alpha) = \int_a^{Q_Y(\Phi(c))} \xi^{\alpha} f_T(u) du$ and $Q_Y(\Phi(c)) = -\log(1 - \Phi(c))$. The results in (3.14) and (3.15) follow by using (3.26) in (3.22) and (3.23). Applying the same techniques of showing (3.14) and (3.15), one can show the results of (3.16) and (3.17) for (ii), (3.18) and (3.19) for (iii), and (3.20) and (3.21) for (iv). □

## 4 Some examples of GN families of distributions with different *T* distributions

In this section different $T$ distributions are used to generate different GN distributions. In the following subsections, we present four new GN distributions namely, Weibull-$N$\{exponential\}, exponential-$N$\{log-logistic\}, logistic-$N$\{logistic\} and logistic-$N$\{extreme value\}. For illustrative purposes, we study some properties of the Weibull-$N$\{logistic\} distribution. To conserve space, properties of other GN distributions are not given. One can follow the same method to study the properties of other GN distributions.

### 4.1 The Weibull-*N*{exponential} distribution

If a random variable $T$ follows the Weibull distribution with parameters $c$ and $\gamma$, then $f_T(x) = c\gamma^{-1}(\frac{x}{\gamma})^{c-1}e^{-(\frac{x}{\gamma})^c}$, $c, \gamma > 0$. From (2.8), the PDF of the Weibull-$N$\{exponential\} is defined as

$$f_X(x) = \frac{c}{\gamma}\frac{\phi(x)}{1-\Phi(x)}\left\{\frac{-\log(1-\Phi(x))}{\gamma}\right\}^{c-1}\exp\left(-\left\{\frac{-\log(1-\Phi(x))}{\gamma}\right\}^c\right). \quad (4.1)$$

**Remark 2.**

(i) When $c = 1$, the Weibull-$N$\{exponential\} reduces to the exponential-normal distribution with $\theta = 1/\gamma$.

(ii) When $c = \gamma = 1$, the Weibull-$N$\{exponential\} reduces to the normal distribution.

(iii) When $c = 1$ and $\gamma^{-1} = n \in N$, the PDF in (4.1) reduces to the distribution of the minimum order statistics, $x_{(1)}$, from a normal random sample of size $n$.

By using (2.7), the CDF of the Weibull-$N$\{exponential\} is given by

$$F_X(x) = 1 - \exp\left(-\left\{-(1/\gamma)\log(1-\Phi(x))\right\}^c\right).$$

In Figures 1 and 2, various graphs of $f_X(x)$ when $\mu = 0$, $\sigma = 1$ and for various values of $c$ and $\gamma$ are provided. These Figures indicate that the Weibull-$N$\{exponential\} PDF can be left skewed, right skewed, or symmetric. Also, the Weibull-$N$\{exponential\} is left skewed whenever $\gamma > 1$ and right skewed whenever $\gamma < 1$. For fixed $\gamma$, the peak increases as $c$ increases.



**Figure 1 The PDF of Weibull-*N*{exponential} for various values of *c* and *γ*.**
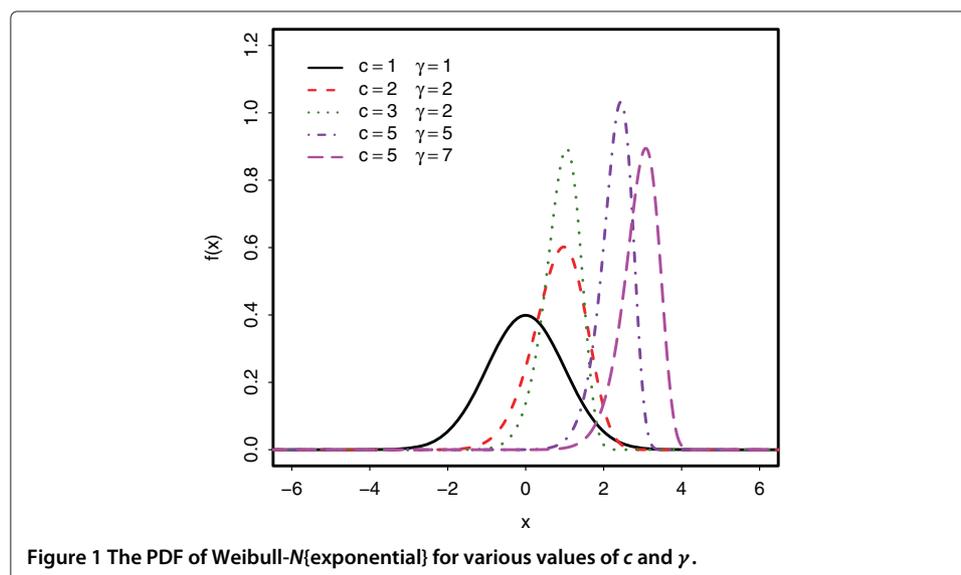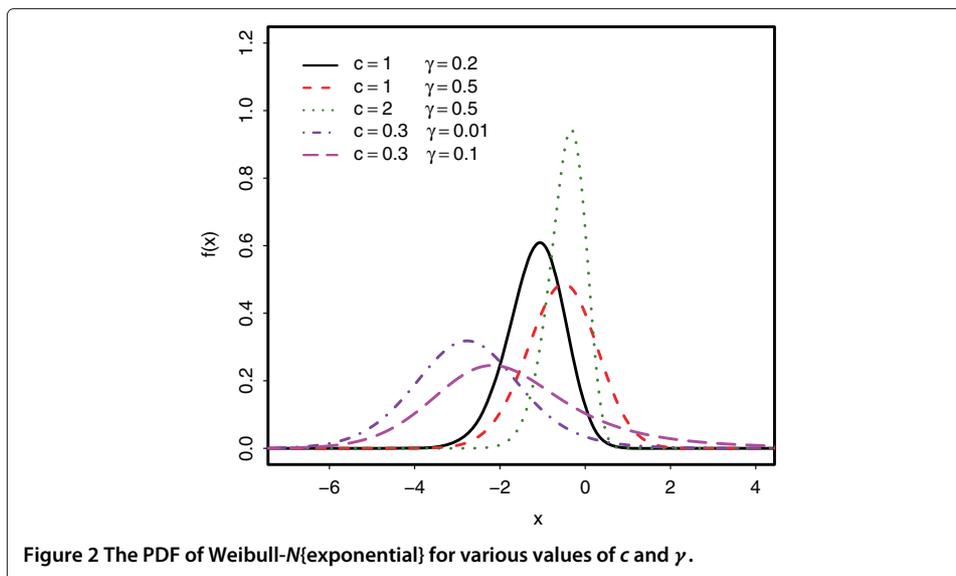
**Figure 2 The PDF of Weibull-*N*{exponential} for various values of *c* and *γ*.**

Some properties of the Weibull-*N*{exponential} are obtained in the following by using the general properties discussed in section 3.

(1) Quantile function: By using Lemma 2, the quantile function of the Weibull-*N*{exponential} distribution is given by
$Q_X(p) = \Phi^{-1}\left\{1 - \exp(-\gamma(-\log(1-p))^{1/c})\right\}$.

(2) Mode: By using Corollary 1, the mode of Weibull-*N*{exponential} distribution is the solution of the following equation

$$x = \mu + \sigma^2 h_\phi(x)\left\{\frac{c-1}{H_\phi(x)} - c\gamma^{-c}(H_\phi(x))^{c-1} + 1\right\}.$$

(3) Shannon entropy: By using Corollary 2 and the fact that $\mu_T = \gamma\Gamma(1+1/c)$ and $\eta_T = 1 + \xi(1-1/c) + \log(\gamma/c)$ (see Song 2001), one can easily obtain the Shannon entropy of Weibull-*N*{exponential} distribution as

$$\eta_X = \log(\sigma\sqrt{2\pi}) - \gamma\Gamma(1+1/c) + \xi(1-1/c) + \log(\gamma/c) + E(X-\mu)^2/(2\sigma^2) + 1.$$

(4) Moments: By using Theorem 3, a series representation of the $r^{\text{th}}$ moments of the Weibull-*N*{exponential} distribution can be obtained by replacing $M_T(-i)$ with

$$\sum_{k=0}^{\infty} \frac{(-1)^k \gamma^k}{k!}\Gamma\left(1 + \frac{k}{c}\right)$$

in equation (3.7).

(5) Mean deviations: By using Theorem 4, the mean deviation from the mean and the mean deviation from the median of Weibull-*N*{exponential} distribution can be obtained by replacing $S(\mu, 0, i)$ and $S(M, 0, i)$ with $\frac{c}{i^c\gamma^c}\sum_{k=0}^{\infty}\frac{(-1)^k}{\gamma^{ck}i^{ck}k!}\Gamma[c(k+1), -i$

$\log(1 - \Phi(\mu))]$ and $\frac{c}{i^c\gamma^c}\sum_{k=0}^{\infty}\frac{(-1)^k}{\gamma^{ck}i^{ck}k!}\Gamma[c(k+1), -i\log(1 - \Phi(M))]$ in

equations (3.14) and (3.15) respectively, where $\Gamma(\alpha, x) = \int_0^x u^{\alpha-1}e^{-u}du$ is the incomplete gamma function.

### 4.2 The exponential-*N* {log-logistic} distribution

If a random variable $T$ follows the exponential distribution with parameter $\lambda$, then $f_T(x) = \lambda e^{-\lambda x}$, $\lambda > 0$. From (2.10), the PDF of the exponential-*N*{log-logistic} is defined as

$$f_X(x) = \frac{\lambda \phi(x)}{(1 - \Phi(x))^2} \exp\left[\frac{-\lambda \Phi(x)}{1 - \Phi(x)}\right]. \tag{4.2}$$

From (2.9), the CDF of (4.2) is given by $F_X(x) = 1 - \exp\left[\frac{\lambda \Phi(x)}{1 - \Phi(x)}\right]$.

In Figure 3, various graphs of $f_X(x)$ when $\mu = 0$, $\sigma = 1$ and for various values of $\lambda$ are provided. These graphs indicate that the exponential-*N*{log-logistic} distribution is always left skewed. Also, the skewness increases as $\lambda$ decreases.

### 4.3 The logistic-*N*{logistic} distribution

If a random variable $T$ follows the logistic distribution with parameter $\lambda$, then $f_T(x) = \lambda e^{-\lambda x}(1 + e^{-\lambda x})^{-2}$, $\lambda > 0$. From (2.12), the PDF of logistic-*N*{logistic} distribution is defined as

$$f_X(x) = \frac{\lambda \phi(x) \Phi^{\lambda-1}(x)(1 - \Phi(x))^{\lambda-1}}{\left[\Phi^\lambda(x) + (1 - \Phi(x))^\lambda\right]^2}. \tag{4.3}$$

From (2.11), the CDF of (4.3) is given by $F_X(x) = \frac{\Phi^\lambda(x)}{\Phi^\lambda(x) + (1 - \Phi(x))^\lambda}$.

When $\lambda = 1$, (4.3) reduces to the normal distribution. In Figure 4, various graphs of $f_X(x)$ when $\mu = 0$, $\sigma = 1$ and for various values of $\lambda$ are provided. These graphs indicate that the PDF of logistic-*N*{logistic} can be bimodal and the bimodality occurs for small values of $\lambda$. Also, it is easy to see from the PDFs in (4.3) that the distribution is symmetric for all values of $\lambda$.

### 4.4 The logistic-*N* {extreme value} distribution

If a random variable $T$ follows the logistic distribution with parameter $\lambda$, then $f_T(x) = \lambda e^{-\lambda x}(1 + e^{-\lambda x})^{-2}$, $\lambda > 0$. From (2.14), the PDF of the logistic-*N*{extreme value} distribution is defined as
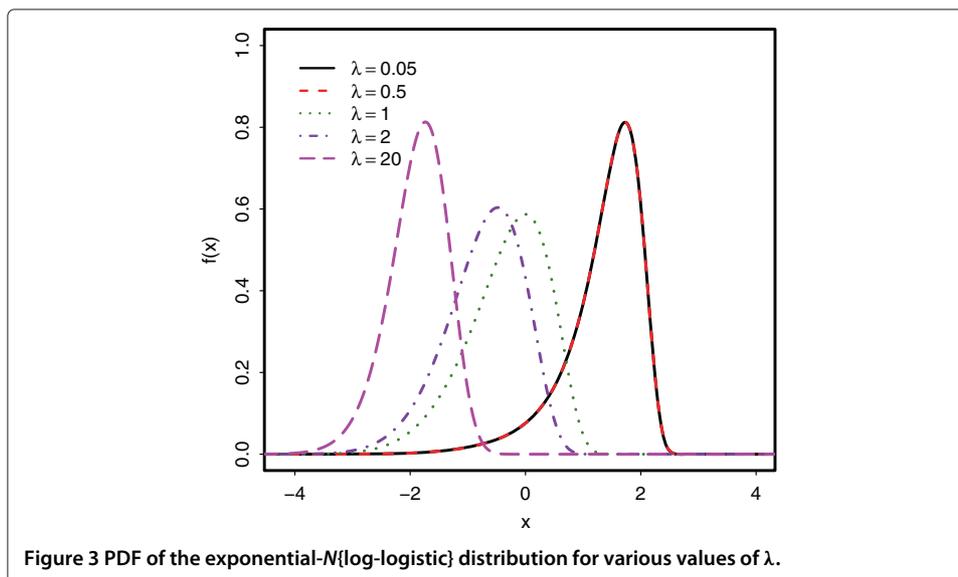


**Figure 3 PDF of the exponential-*N*{log-logistic} distribution for various values of $\lambda$.**

**Figure 4 PDF of logistic-*N*{logistic} distribution for various values of λ.**
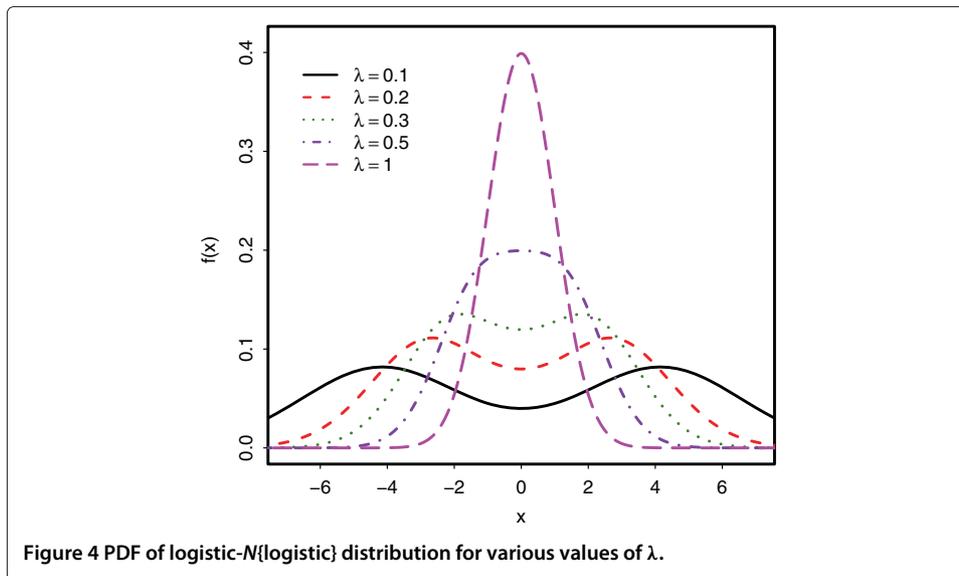
$$f_X(x) = \frac{\lambda h_\phi(x) H_\phi{}^{\lambda-1}(x)}{\left(1 + H_\phi{}^\lambda(x)\right)^2}. \tag{4.4}$$

From (2.13), the CDF of (4.4) is given by $F_X(x) = \frac{H_\phi{}^\lambda(x)}{1 + H_\phi{}^\lambda(x)}$.

In Figure 5, various graphs of $f_X(x)$ when $\mu = 0$, $\sigma = 1$ and for various values of $\lambda$ are provided. These graphs indicate that the distribution is always right skewed. Also, the skewness increases as $\lambda$ decreases.

## 5 Applications

To illustrate the flexibility of the GN distributions, we fit some GN distributions to a unimodal data set and a bimodal data set. The unimodal data with $n = 66$ in Table 2



**Figure 5 PDF of the logistic-*N*{extreme value} distribution for various values of λ.**
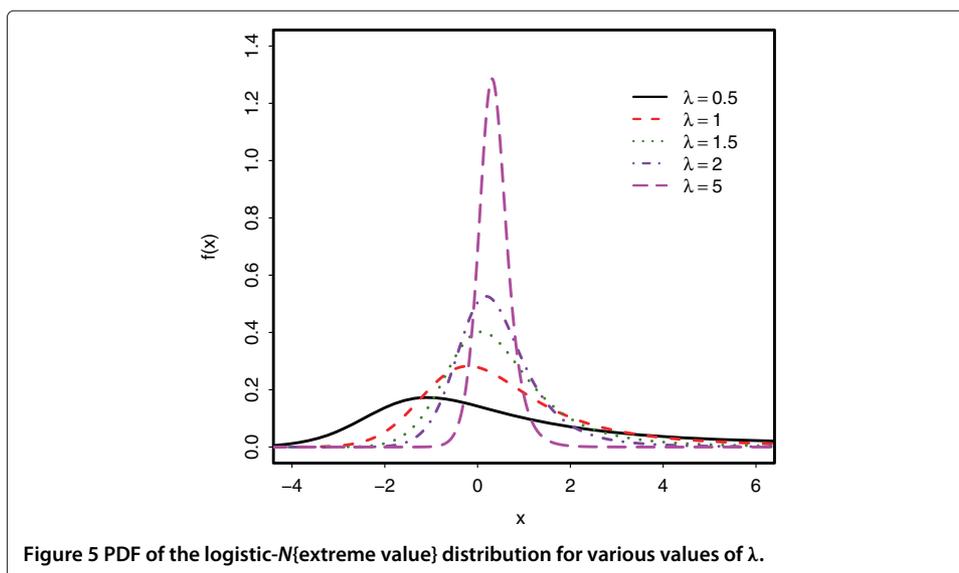
**Table 2 Breaking stress of carbon fibers data**

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 3.70 | 2.74 | 2.73 | 2.50 | 3.60 | 3.11 | 3.27 | 2.87 | 1.47 | 3.11 | 3.56 |
| 4.42 | 2.41 | 3.19 | 3.22 | 1.69 | 3.28 | 3.09 | 1.87 | 3.15 | 4.90 | 1.57 |
| 2.67 | 2.93 | 3.22 | 3.39 | 2.81 | 4.20 | 3.33 | 2.55 | 3.31 | 3.31 | 2.85 |
| 1.25 | 4.38 | 1.84 | 0.39 | 3.68 | 2.48 | 0.85 | 1.61 | 2.79 | 4.70 | 2.03 |
| 1.89 | 2.88 | 2.82 | 2.05 | 3.65 | 3.75 | 2.43 | 2.95 | 2.97 | 3.39 | 2.96 |
| 2.35 | 2.55 | 2.59 | 2.03 | 1.61 | 2.12 | 3.15 | 1.08 | 2.56 | 1.80 | 2.53 |

is obtained from Nichols and Padgett (2006) on the breaking stress of carbon fibers of 50 mm in length. Alzaatreh at el. (2013) fitted the data set to the gamma-normal distribution. They showed that the standard gamma-normal distribution with $\mu = 0$ and $\sigma = 1$ provides a good fit to the data set. The standard form of exponential-$N$\{exponential\}, exponentiated exponential-$N$\{exponential\} and Weibull-$N$\{exponential\} distributions with $\mu = 0$ and $\sigma = 1$ are applied to fit the data set and the results compared with the results from standard gamma-normal distribution. The maximum likelihood estimates, the log-likelihood value, the AIC (Akaike Information Criterion), the Kolmogorov-Smirnov (K-S) test statistic, and the p-value for the K-S statistic for the fitted distributions are reported in Table 3. The results in Table 3 show that all the generalized normal distributions give an adequate fit to the data. However, the K-S values indicate that the gamma-$N$\{exponential\} distribution provides the best fit among the distributions. Figure 6 displays the histogram and the fitted density functions for the data.

The second application is on a bimodal data set obtained from Emlet et al. (1987) on the asteroid and echinoid egg size. The data is available from the first author. The data consists of 88 asteroid species divided into three types; 35 planktotrophic larvae, 36 lecithotrophic larvae, and 17 brooding larvae. Since the logarithm of the egg diameters of the asteroids data has a bimodal shape, Famoye et al. (2004) applied the beta-normal distribution to the logarithm of the data set. We apply the logistic-$N$\{logistic\} distribution, which can be bimodal, to fit the same data. The results of the maximum likelihood estimates, the log-likelihood value, the AIC (Akaike Information Criterion), the Kolmogorov-Smirnov (K-S) test statistic, and the p-value for the K-S statistic for the fitted distributions are reported in Table 4. The results in Table 4 show

**Table 3 Parameter estimates for the carbon fibers data**

| Distribution | Exponential-$N$\{exponential\} | Exponentiated exponential-$N$\{exponential\} | Weibull-$N$\{exponential\} | gamma-$N$\{exponential\} |
|---|---|---|---|---|
| Parameter | $\hat{\theta} = 0.1612$ | $\hat{\alpha} = 6.0389$ | $\hat{c} = 2.4062$ | $\hat{\alpha} = 4.7966$ |
| Estimates | (0.0198)* | (1.3675) | (0.2226) | (0.8076) |
| | | $\hat{\lambda} = 0.3919$ | $\hat{\gamma} = 6.9991$ | $\hat{\beta} = 1.2932$ |
| | | (0.0434) | (0.3780) | (0.2296) |
| Log-likelihood | -114.8292 | -87.0385 | -86.0629 | -85.9070 |
| AIC | 231.6584 | 178.0770 | 176.1258 | 175.8140 |
| K-S | 0.2768 | 0.0854 | 0.0894 | 0.0693 |
| K-S p-value | 0.0001 | 0.7215 | 0.6676 | 0.9090 |

*standard error.

**Figure 6 PDF for fitted distributions of the breaking stress of carbon fibers data.**
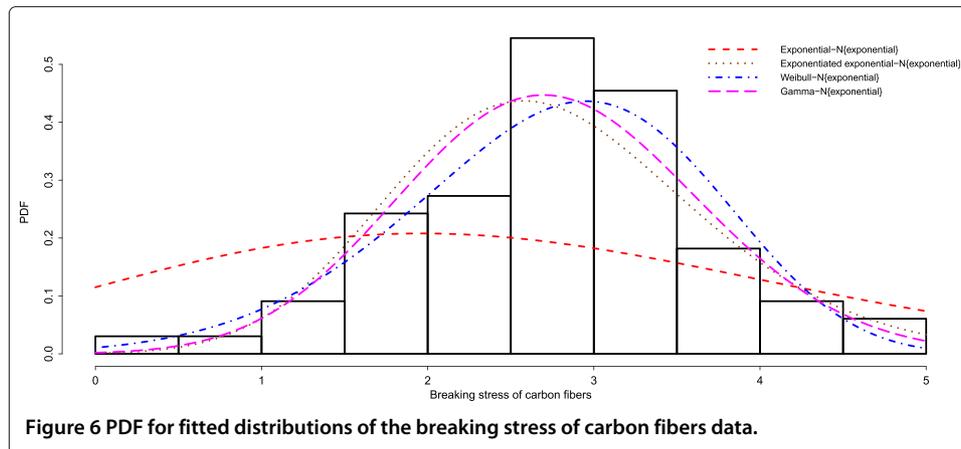
that both the beta-normal and logistic-$N$\{logistic\} distributions give an adequate fit to the data. However, the K-S values indicate that the logistic-$N$\{logistic\} distribution provides a better fit. Figure 7 displays the histogram and the fitted density functions for the data.

## 6   Summary and conclusions

The normal distribution is the most commonly used distribution in both statistical theory and applications. The generalization of the normal distribution is studied using the $T$-$X$ framework proposed by Alzaatreh et al. (2013). Four types of generalized normal families from the quantile functions of the (i) exponential, (ii) log-logistic, (iii) logistic, and (iv) extreme value distributions are proposed. Some general properties are studied. Four generalized normal distributions are described and some of their properties investigated. It is noticed that the shapes of GN distributions can be symmetric, skewed to the right, skewed to the left or bimodal. This gives the families some flexibility in fitting real world data. Because the GN distributions include the normal distribution as a special case, using the GN distributions to fit data enables one to check if the additional parameters characterize the deviation from the normal distribution. Many types of generalizations of the normal distribution can be derived using the methodology described in this paper. Due to the fact that GN distributions are

**Table 4 Parameter estimates for the asteroids data**

| Distribution | Beta-normal* | Logistic-$N$\{logistic\} |
|---|---|---|
| Parameter | $\hat{\alpha} = 0.0129$ | $\hat{\lambda} = 0.1498(0.0185)$ |
| Estimates | $\hat{\beta} = 0.0070$ | $\hat{\mu} = 6.0348(0.0685)$ |
| | $\hat{\mu} = 5.7466$ | $\hat{\sigma} = 0.2604(0.010)$ |
| | $\hat{\sigma} = 0.0675$ | |
| Log-likelihood | -109.4800 | -111.4287 |
| AIC | 226.9600 | 228.4974 |
| K-S | 0.1233 | 0.0988 |
| K-S p-value | 0.1377 | 0.3572 |

*From Famoye et al. (2004) and the MLE standard errors were not provided.

**Figure 7 PDF for the fitted distributions for the asteroids data.**
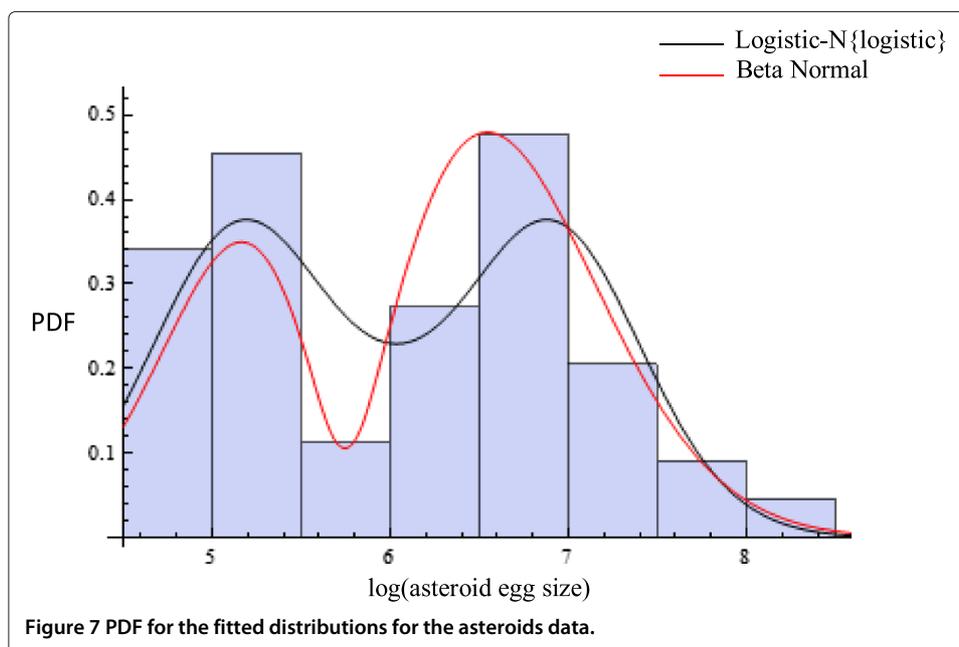
natural extensions from the normal distribution, statistical modeling by assuming the error term follows some form of GN distribution will be an interesting topic for future research.

**Authors' contributions**
The authors, viz AA, CL and FF with the consultation of each other carried out this work and drafted the manuscript together. All authors read and approved the final manuscript.

**Author details**
[1]Department of Mathematics and Statistics, Austin Peay State University, Clarksville, TN 37044, USA. [2]Department of Mathematics, Central Michigan University, Mount Pleasant, MI 48859, USA.

**References**
Alexander, C, Cordeiro, GM, Ortega, EMM, Sarabia, JM: Generalized beta-generated distributions. Comput. Stat. Data Anal. **56**(6), 1880–1897 (2012)
Aljarrah, MA, Lee, C, Famoye, F: A method of generating *T-X* family of distributions using quantile functions. J. Stat. Distributions Appl. **1**(2), 17 (2014)
Alzaatreh, A, Famoye, F, Lee, C: The gamma-normal distribution: Properties and applications. Comput. Stat. Data Anal. **69**(1), 67–80 (2014)
Alzaatreh, A, Lee, C, Famoye, F: A new method for generating families of continuous distributions. Metron. **71**(1), 63–79 (2013)
Arellano-Valle, RB, Gómez, HW, Quintanan, FA: A new class of skew-normal distributions. Commun. Stat.-Theory Methods. **33**, 1465–1480 (2004)
Arnold, BC, Beaver, RJ: Skewed multivariate models related to hidden truncation and/or selective reporting (with discussion). Test. **11**, 7–35 (2002)
Arnold, BC, Castillo, E, Jose, JM: Distributions with generalized skewed conditionals and mixtures of such distributions. Commun. Stat.-Theory Methods. **36**, 1493–1503 (2007)
Azzalini, A: A class of distributions which includes the normal ones. Scand. J. Stat. **12**, 171–178 (1985)

Balakrishnan, N: Discussion of skewed multivariate models related to hidden truncation and/or selective reporting by Arnold & Beaver. Test. **11**, 37–39 (2002)

Choudhury, K, Abdul, MM: Extended skew generalized normal distribution. Metron. **69**, 265–278 (2011)

Cordeiro, GM, de Castro, M: A new family of generalized distributions. J. Stat. Comput. Simulat. **81**(7), 883–898 (2011)

de Moivre, A: Approximatio ad summam ferminorum binomii $(a + b)^n$ in seriem expansi. Self-published pamphlet (1733)

Emlet, RB, McEdward, LR, Strathmann, RR: Echinoderm larval ecology viewed from the egg. In: Jangoux, M, Lawrence, JM (eds.) *Echinoderm Studies, volume 2*, pp. 55–136. AA Balkema, Rotterdam (1987)

Eugene, N, Lee, C, Famoye, F: The beta-normal distribution and its applications. Commun. Stat.-Theory Methods. **31**(4), 497–512 (2002)

Famoye, F, Lee, C, Eugene, N: Beta-normal distribution: bimodality properties and applications. J. Mod. Appl. Stat. Meth. **3**(1), 85–103 (2004)

Fernández, C, Steel, MFJ: On Bayesian modeling of fat tails and skewness. J. Am. Stat. Assoc. **93**, 359–371 (1998)

Ferreira, JTAS, Steel, MFJ: A constructive representation of univariate skewed distributions. J. Am. Stat. Assoc. **101**, 823–829 (2006)

Gauss, CF: Theoria Motus Corporum Coelestium, pp. 205–224. Perthes u. Besser, Hamburg (1809). Lib. 2, Sec. III

Gauss, CF: Bestimmung der genauigkeit der beobachtungen. Zeitschrift Astronomi. **1**, 185–197 (1816)

Gupta, RC, Gupta, RD: Generalized skew normal model. Test. **13**, 501–520 (2004)

Johnson, NL, Kotz, S, Balakrishnan, N: Continuous Univariate Distributions, volume 1. second edition. John Wiley & Sons, New York (1994)

Kotz, S, Vicari, D: Survey of developments in the theory of continuous skewed distributions. Metron. **63**, 225–261 (2005)

Lee, C, Famoye, F, Alzaatreh, A: Methods for generating families of univariate continuous distributions in the recent decades. WIREs Comput. Stat. **5**, 219–238 (2013)

Nichols, MD, Padgett, WJ: A bootstrap control for Weibull percentiles. Qual. Reliab. Eng. Int. **22**, 141–151 (2006)

Patel, JK, Read, CB, Handbook of the Normal Distribution. second edition. Marcel Dekker, New York (1996)

Pourahmadi, M: Construction of skew-normal random variables: are they linear combination of normals and half-normals? J. Stat. Theory Appl. **3**, 314–328 (2007)

Rényi, A: On measures of entropy and information. In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 547–561. University of California Press, Berkeley, CA (1961)

Sharafi, M, Behboodian, J: The Balakrishnan skew-normal density. Stat. Pap. **49**, 769–778 (2008)

Song, KS: Rényi information, loglikelihood and an intrinsic distribution measure. J. Stat. Plann. Infer. **93**, 51–69 (2001)

Wolfram. http://functions.wolfram.com/GammaBetaErf/InverseErf/06/01/02/0004/. Retrieved on May 29, 2014

Yadegari, I, Gerami, A, Khaledi, MJ: A generalization of the Balakrishnan skew-normal distribution. Stat. Probability Lett. **78**, 1165–1167 (2008)