**Journal of Sciences, Islamic Republic of Iran 21(4): 343-351 (2010)** http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

# Generalizations of the Skew t-Normal Distribution and their Properties

A. Kheradmandi,[1] M. Mohammadzadeh,[1,*] and N. Sanjari Farsipour[2]

[1] *Department of Statistics, Faculty of Science, Tarbiat Modares University, Tehran, Islamic Republic of Iran*
[2] *Department of Mathematics, Faculty of Basic Science, Alzahra University, Tehran, Islamic Republic of Iran*

Received: 13 June 2010 / Revised: 25 August 2010 / Accepted: 31 August 2010

## Abstract

In this paper we consider several generalizations of the skew t-normal distribution, and some of their properties. Also, we represent several theorems for constructing each generalized skew t-normal distribution. Next, we illustrate the application of the proposed distribution studying the ratio of two heavy metals, Nickel and Vanadium, associated with crude oil in Shadgan wetland in the south-west of Iran at the head of the Persian Gulf.

**Keywords:** Skewness; Skew normal; Skew t-normal; Generalized skew t-normal



## Introduction

In practice, some data sets exhibit noticeable asymmetry such that the normal distribution is not appropriate for their modeling. However, one can use the skew-symmetric distributions generated by normal distribution to describe their characteristics. Azzalini [3] introduced the skew-normal (SN) distribution, denoted by $SN(\lambda)$, where the parameter $\lambda$, controls the skewness of the distribution. This distribution and its variations have been discussed by several authors including [2,4,7,8,10,14]. The SN distribution has a number of properties resembling those of the normal one, but it has constraints in asymmetry and kurtosis coefficients. It will be useful to consider an alternative to the SN distribution which is both heavy tailed and skew. Jones and Faddy [15], based on a beta density transformation, proposed a skew student's t-distribution (St) which includes extensions of the t-distribution with non-zero skewness. Azzalini [6] proposed an alternative form of St distribution that is more naturally linked to the SN distribution. Gomez *et al* [13] introduced the skew student-t-normal distribution (StN), showing that it is a good alternative to model heavy tailed data with strong asymmetrical nature, especially because it has a larger range of skewness than other traditional skewed versions of normal and student's distributions. The StN density function is defined by $f(x;\nu,\lambda) = 2\psi(x;\nu)\Phi(\lambda x)$, $x \in R$, $\lambda \in R$, where $\psi(x;\nu)$ is the density of $t(\nu)$, i.e. student's t-distribution function with $\nu > 0$ degree of freedom and $\Phi$ is the standard normal cumulative distribution function. When X has StN distribution with parameters $\nu$ and $\lambda$, we denote it by $X \sim StN(\nu,\lambda)$. Kheradmandi *et al* [16] considered some properties of the StN distribution. The skew-normal distribution has been generalized by some researchers, for example, [1,2,17,19]. In this paper we use their ideas to extend and present new classes of the generalized StN distribution. Next, the important

* Corresponding author, Tel.: +98(21)82883483, Fax: +98(21)82883483, E-mail: mohsen_m@modares.ac.ir

properties of all presented classes and their construction methods are described. Then, all skewed distributions considered in this paper are applied to fit a suitable probability distribution to the oil pollution data of Shadegan wetland in south-west of Iran.

## Generalized StN Distributions

### *Type One*

A limitation of the StN($\nu$, $\lambda$) model is that for moderate values of $\lambda$ nearly all the mass accumulates either on the positive or negative numbers, as determined by the sign of $\lambda$. Similar to generalization given by [1] for the SN distribution, we define a generalization for StN distribution and present some of its properties. This distribution exhibits a better behavior, especially on the side with smaller mass. Also we define a specific form of this family. In the sequel $X \overset{D}{=} Y$ denotes that $X$ and $Y$ are identically distributed, and $X \mid Y$ denotes $X$ given $Y$. It shoud be noted that the probability density function of all generalized StN distributions presented in the sequal are integrated to one by the following lemma.

***Lemma 1.*** (Azzalini [5]) If $f_0$ is a one-dimensional probability density function symmetric about 0, and G is a one dimensional distribution function such that $G_0$ exists and is a density symmetric about 0, then $f(z) = 2f_0(z)G(w(z))$, $z \in R$ is a density function for any odd function $w(\cdot)$.

***Definition 1.*** A random variable X with density function

$$f(x;\nu,\lambda_1,\lambda_2) = 2\psi(x;\nu)\Phi\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right),\ x \in R,$$

is said to have a generalized StN distribution of type one with parameters $\lambda_1 \in R$, $\lambda_2 \geq 0$, and it is denoted by X ~ StN($\nu$,$\lambda$).

Figure 1(a) illustrates several possible shapes obtained from $G_1StN(\nu,\lambda_1,\lambda_2)$ under various choices of $(\lambda_1,\lambda_2)$ and $\nu = 5$. It can be seen that when $\lambda_2$ increases, the resulting $G_1StN$ density gets heavier or thinner right tail depending on $\lambda_1 < 0$ or $\lambda_1 > 0$. This distribution has the following properties:

- If $\lambda_1 = 0$, then $X \sim t(\nu)$,
- If $\lambda_2 = 0$, then $X \sim StN(\nu,\lambda_1)$,
- $-X \sim G_1StN(\nu,-\lambda_1,\lambda_2)$,
- $f(x;\nu,-\lambda_1,\lambda_2)+f(x;\nu,\lambda_1,\lambda_2) = 2\psi(x;\nu),\ \forall x \in R$,
- $\lim_{\lambda_1 \to \infty} f(x;\nu,\lambda_1,\lambda_2) = 2\psi(x;\nu)I_{(x>0)}$,
- $\lim_{\lambda_1 \to -\infty} f(x;\nu,\lambda_1,\lambda_2) = 2\psi(x;\nu)I_{(x<0)}$,
- $\lim_{\lambda_2 \to \infty} f(x;\nu,\lambda_1,\lambda_2) = \psi(x;\nu),\ \forall x \in R$, ,
- $F(x;\nu,\lambda_1 = 0,\lambda_2) = \Psi(x;\nu),\ \forall x \in R,\ \lambda_2 \geq 0$,
- $F(-x;\nu,\lambda_1,\lambda_2) = 1 - F(x;\nu,-\lambda_1,\lambda_2)$,
- If $T \sim t(\nu)$, then $|X| \overset{D}{=} |Y|$,

where $F$ and $\Psi$ are the cumulative distribution functions of $G_1StN(\nu,\lambda_1,\lambda_2)$ and $t(\nu)$, respectively. These properties are straightforward to verify; for the item x note that the random variable $|T|$ has half-t distribution, and setting $U = |X|$ implies $f_U(u)$ $= 2\psi(u;\nu)[\Phi(\frac{\lambda_1 u}{\sqrt{1+\lambda_2 u^2}}) + \Phi(\frac{-\lambda_1 u}{\sqrt{1+\lambda_2 u^2}})] = 2\psi(u;\nu)I_{(U \geq 0)}$.

Now we present three theorems for simulating a random variable from $G_1StN(\nu,\lambda_1,\lambda_2)$.

***Theorem 1.*** Let $T \sim t(\nu)$ be independent of $Z \sim N(0,1)$ and $\lambda_1 \in R$, $\lambda_2 \geq 0$, then

$$T \mid \{Z \leq \frac{\lambda_1 T}{\sqrt{1+\lambda_2 T^2}}\} \sim G_1StN(\nu,\lambda_1,\lambda_2)$$

***Proof.***

$$f\left(t \mid Z \leq \frac{\lambda_1 T}{\sqrt{1+\lambda_2 T^2}}\right) = \frac{P\left(Z \leq \frac{\lambda_1 t}{\sqrt{1+\lambda_2 t^2}}\right)\psi(t;\nu)}{P\left(Z \leq \frac{\lambda_1 T}{\sqrt{1+\lambda_2 T^2}}\right)}$$

$$= \frac{\Phi\left(\frac{\lambda_1 t}{\sqrt{1+\lambda_2 t^2}}\right)\psi(t;\nu)}{\int_{-\infty}^{\infty}\Phi\left(\frac{\lambda_1 t}{\sqrt{1+\lambda_2 t^2}}\right)\psi(t;\nu)dt}$$

$$= 2\psi(t;\nu)\Phi\left(\frac{\lambda_1 t}{\sqrt{1+\lambda_2 t^2}}\right).$$

Now a random value x of $G_1StN(\nu,\lambda_1,\lambda_2)$ distribution can be generated by the following algorithm.

**(a)** Generated $t$ from $t(\nu)$,

**(b)** Generated $z$ from $N(0,1)$,

**(c)** If $z \le \frac{\lambda_1 t}{\sqrt{1+\lambda_2 t^2}}$, then $x = t$, otherwise go to (a).

***Theorem 2.*** Let $T \sim t(\nu)$ be independent of $Z \sim N(0,1)$ and $\lambda_1 \in R$, $\lambda_2 \ge 0$. If we define $X$ to be equal to $T$ when $Z \le \frac{\lambda_1 t}{\sqrt{1+\lambda_2 t^2}}$, otherwise equal to $-T$, then $X \sim G_1StN(\nu, \lambda_1, \lambda_2)$

***Proof.***

$$f_X(x) = P(Z \le \frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}})\psi(x;\nu)$$

$$+P(Z > \frac{-\lambda_1 x}{\sqrt{1+\lambda_2 x^2}})\psi(-x;\nu)$$

$$= \psi(x;\nu)[\Phi(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}) + 1 - \Phi(\frac{-\lambda_1 x}{\sqrt{1+\lambda_2 x^2}})]$$

$$= 2\psi(x;\nu)\Phi(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}).$$

The $G_1StN(\nu, \lambda_1, \lambda_2)$ distribution might be generated via mixtures on the asymmetry parameter of a StN($\nu, \lambda$) distribution, when $\lambda$ has normal distribution. This form provides a Bayesian interpretation of $G_1StN(\nu, \lambda_1, \lambda_2)$ distribution.

Now a random value x of $G_1StN(\nu, \lambda_1, \lambda_2)$ distribution can be generated by the following algorithm.

**(a)** Generated $t$ from $t(\nu)$,

**(b)** Generated $z$ from $N(0,1)$,

**(c)** If $z \le \frac{\lambda_1 t}{\sqrt{1+\lambda_2 t^2}}$, then $x = t$, otherwise $x = -t$.

***Theorem 3.*** If $X | \Lambda = \lambda \sim StN(\nu, \lambda)$ and $\Lambda \sim N(\lambda_1, \lambda_2)$, then $X \sim G_1StN(\nu, \lambda_1, \lambda_2)$.

***Proof.***

$$f_X(x;\nu,\lambda_1,\lambda_2) = \frac{2}{\sqrt{\lambda_2}}\int_{-\infty}^{+\infty}\psi(x;\nu)\Phi(x\lambda)\phi(\frac{\lambda-\lambda_1}{\sqrt{\lambda_2}})d\lambda$$

$$= \int_{-\infty}^{+\infty} 2\psi(x;\nu)\Phi(\sqrt{\lambda_2}xz + \lambda_1 x)\phi(z)dz$$

$$= 2\psi(x;\nu)E(\Phi(\sqrt{\lambda_2}xZ + \lambda_1 x)]$$

$$= 2\psi(x;\nu)\Phi(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}),$$

because $E[\Phi(Y)] = \Phi(\frac{\mu}{\sqrt{1+\sigma^2}})$, when $Y \sim N(\mu, \sigma^2)$ (Ellison, [11]).

Now a random value $x$ of $G_1StN(\nu, \lambda_1, \lambda_2)$ distribution can be generated by the following algorithm.

**(a)** Generated $\lambda \sim N(\lambda_1, \lambda_2)$,

**(b)** Generated $X \sim StN(\nu, \lambda)$.

If we choose $\lambda_1 = \sqrt{\lambda_2} = \lambda$ in $G_1StN$ distribution, then its density function is given by $f(x;\nu,\lambda) = 2\psi(x;\nu)\Phi(\frac{\lambda x}{\sqrt{1+\lambda^2 x^2}})$, that is called skew-curved t-normal distribution and is denoted by $SCtN(\nu, \lambda)$. The plots of $SCtN(\nu, \lambda)$ for various values of $\lambda$ are shown in Figure 1(b). Note that for $\lambda = 0$, it reduces to t-distribution. But it does not include the skew t-normal distribution, because we are not able to find a value for $\lambda$ that $\frac{\lambda x}{\sqrt{1+\lambda^2 x^2}}$ gets the necessary form of $\lambda x$. If $f_{SCtN}(\cdot;\nu,\lambda)$ and $F_{SCtN}(\cdot;\nu,\lambda)$ are the density and distribution functions of $SCtN(5, \lambda)$ then it is straightforward to show the following properties:

- $\lim_{\lambda\to\infty} f_{SCtN}(x;\nu,\lambda) = 2\psi(x;\nu)\Phi(\pm sign(x)),$
- $F_{SCtN}(-x;\nu,\lambda) = 1 - F_{SCtN}(x;\nu,-\lambda),$
- $\lim_{\lambda\to\infty} F_{SCtN}(x;\nu,\lambda) = \begin{cases} 2(1-\Phi(1))\Psi(x;\nu) & x < 0 \\ 1-2(1-\Psi(x;\nu))\Phi(1) & x \ge 0 \end{cases}$

  $\lim_{\lambda\to -\infty} F_{SCtN}(x;\nu,\lambda)$
- $= \begin{cases} 2\Phi(1)\Psi(x;\nu) & x < 0 \\ 2[\Phi(1) + (1-\Phi(1))\Psi(x;\nu) - \frac{1}{2}] & x \ge 0 \end{cases}$

The first item shows, when $\lambda$ grows, the resulting SCtN($\nu, \lambda$) density has heavier tail to the left than the StN distribution.

***Type Two***

Ma and Gentone [17] generalized the Azzalini SN distribution, which includes multimodal SN distributions. In this section we define similar generalization for StN distribution and present some of its properties.

***Definition 2.*** A random variable X with density function; $f(x;\nu,\lambda_1,\lambda_2) = 2\psi(x;\nu)\Phi(\lambda_1 x + \lambda_2 x^3)$;

$x \in R$ is said to have a generalized skew t-normal of type two with parameters $\lambda_1, \lambda_2 \in R$, and it is denoted by $X \sim G_2StN(\nu, \lambda_1, \lambda_2)$.

The plot of $G_2StN(\nu, \lambda_1, \lambda_2)$ is shown in Figure 1(c). This distribution has the following properties:

- If $\lambda_2 = 0$, then $X \sim StN(\nu, \lambda_1)$,
- If $\lambda_1 = \lambda_2 = 0$, then $X \sim t(\nu)$,
- $\lim_{\lambda_i \to \infty} f(x; \nu, \lambda_1, \lambda_2) = 2\psi(x;\nu) I_{(x>0)} \;\; i = 1, 2,$
- $\lim_{\lambda_i \to -\infty} f(x; \nu, \lambda_1, \lambda_2) = 2\psi(x;\nu) I_{(x<0)} \;\; i = 1, 2.$

We present two theorems for simulating a random variable from $G_2StN(\nu, \lambda_1, \lambda_2)$.

***Theorem 4.*** Let $T \sim t(\nu)$ be independent of $U \sim U(0,1)$ and $\lambda_1, \lambda_2 \in R$, then

$$T \mid \{U \le \Phi(\lambda_1 T + \lambda_2 T^3)\} \sim G_2StN(\nu, \lambda_1, \lambda_2).$$

***Proof.***

$$\begin{aligned} & f(t \mid U \le \Phi(\lambda_1 T + \lambda_2 T^3)) \\ &= \frac{P(U \le \Phi(\lambda_1 t + \lambda_2 t^3))\psi(t;\nu)}{P(U \le \Phi(\lambda_1 T + \lambda_2 T^3))} \\ &= \frac{\Phi(\lambda_1 t + \lambda_2 t^3)\psi(t;\nu)}{\int_{-\infty}^{\infty} \Phi(\lambda_1 t + \lambda_2 t^3)\psi(t;\nu)dt} \\ &= 2\Phi(\lambda_1 t + \lambda_2 t^3)\psi(t;\nu). \end{aligned}$$

***Theorem 5.*** Let $T \sim t(\nu)$ be independent of $U \sim U(0,1)$ and $\lambda_1 \in R$, $\lambda_2 \ge 0$. If we define $X$ to be equal to $T$ when $U \le \Phi(\lambda_1 T + \lambda_2 T^3)$, otherwise equal to $-T$, then $X \sim G_2StN(\nu, \lambda_1, \lambda_2)$.

***Proof.***

$$\begin{aligned} f_X(x) &= P(U \le \Phi(\lambda_1 x + \lambda_2 x^3))\psi(x;\nu) \\ &\quad + P(U > \Phi(-\lambda_1 x - \lambda_2 x^3))\psi(-x;\nu) \\ &= [\Phi(\lambda_1 x + \lambda_2 x^3) + 1 - \Phi(-\lambda_1 x - \lambda_2 x^3)]\psi(x;\nu) \\ &= 2\Phi(\lambda_1 T + \lambda_2 x^3)\psi(x;\nu). \end{aligned}$$

### *Type Three*

Balakrishnan [9], as a discussant of [2], generalized the Azzalini SN distribution and [12] and [18] consider some of its properties. In this section we give similar generalization for StN distribution and present some of its properties.

***Definition 3.*** A random variable X has generalized StN distribution of type three, if it has the density of the form $f_n(x;\nu,\lambda) = c_n(\nu,\lambda)\psi(x;\nu)\Phi^n(\lambda x)$ $x \in R, \lambda \in R$, where $n$ is a non-negative integer, $\psi(\cdot,\nu)$ is the student's t-density function with $\nu$ degrees of freedom and $\Phi(\cdot)$ is the standard normal cumulative distribution function. Here the coefficient $c_n(\nu,\lambda)$ is given by $[c_n(\nu,\lambda)]^{-1} = \int_{-\infty}^{+\infty} \psi(x;\nu)\Phi^n(\lambda x)dx = E(\Phi^n(\lambda T))$, where $T \sim t(\nu)$. When X has generalized StN distribution of

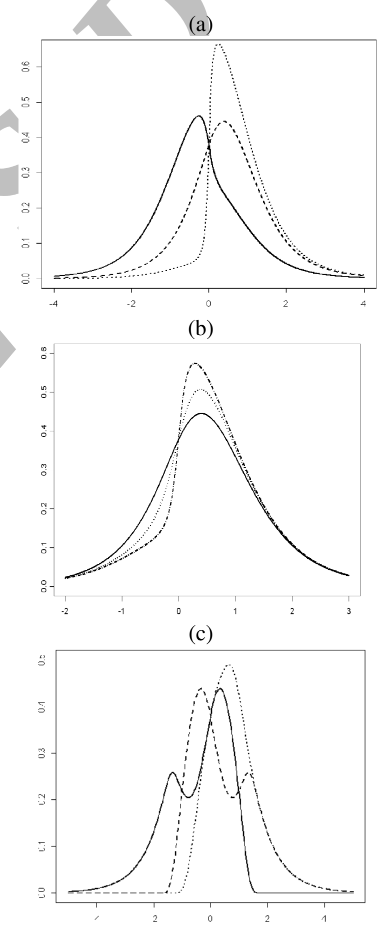



**Figure 1**. (a) Plots of $G_1StN(\nu, \lambda_1, \lambda_2)$ for $\nu = 5$ and (solid line) $(\lambda_1, \lambda_2) = (-2, 20)$, (dashed line) $(\lambda_1, \lambda_2) = (1,1)$, (dotted line) $(\lambda_1, \lambda_2) = (10, 40)$ (b) Plots of $SCtN(\nu, \lambda)$ for $\nu = 5$ and (solid line) $\lambda = 1$, (dashed line) $\lambda = 2$, (dotted line) $\lambda = 5$ and (c) Plots of $G_2StN(\nu, \lambda_1, \lambda_2)$ for $\nu = 5$ and (solid line) $(\lambda_1, \lambda_2) = (1, -1)$, (dashed line) $(\lambda_1, \lambda_2) = (-1, 1)$, (dotted line) $(\lambda_1, \lambda_2) = (1,1)$.

type three we denote it by $X \sim G_3StN(\nu,\lambda,n)$, then it is straightforward to verify the following properties:

- If $\lambda=0$ then $X \sim t(\nu)$,
- If $n=1$ then $X \sim StN(\nu,\lambda)$,
- $\lim_{\lambda\to\infty} f_n(x;\nu,\lambda) = 2\psi(x;\nu)I_{(x>0)}$,
- $\lim_{\lambda\to -\infty} f_n(x;\nu,\lambda) = 2\psi(x;\nu)I_{(x<0)}$,

The curves of $G_3StN(\nu;\lambda,n)$ for $\nu=5$ and various values of $n$ and $\lambda$ are presented in Figure 2. Figure 2(a) shows that, when $\lambda$ and $\nu$ are fixed, increasing $n$ causes heavier tailed distribution. Also Figure 2(b) shows the effect of $\lambda$ on the skewness of $G_3$ StN distribution. Now we present two theorems that can be used for simulating a random variable from $G_3StN(\nu,\lambda,n)$.

***Theorem 6.*** Let $T \sim t(\nu)$ be independent of the random sample $Z_1,\cdots,Z_n$ from N(0,1) and $\lambda\in R$. If $Z_{(n)} = max\{Z_1,\cdots,Z_n\}$, then $T \mid \{Z_{(n)} \le \lambda T\} \sim G_3StN(\nu,\lambda,n)$.

***Proof.*** Let $X \equiv T \mid \{Z_{(n)} \le \lambda T\}$, then $P(X \le x) = \dfrac{P(T \le x, Z_1 \le \lambda T,\cdots,Z_n \le \lambda T)}{P(Z_1 \le \lambda T,\cdots,Z_n \le \lambda T)}$. Since

$$P(T \le x, Z_1 \le \lambda T,\cdots,Z_n \le \lambda T)$$
$$= \int_{-\infty}^{\infty} P(T \le x \mid t) P(Z_1 \le \lambda t,\cdots,Z_n \le \lambda t)\psi(t;\nu)dt$$
$$= \int_{-\infty}^{x} \Phi^n(\lambda t)\psi(t;\nu)dt,$$

and

$$P(Z_1 \le \lambda T,\cdots,Z_n \le \lambda T)$$
$$= \int_{-\infty}^{+\infty} P(Z_1 \le \lambda t,\cdots,Z_n \le \lambda t)\psi(t;\nu)dt$$
$$= \int_{-\infty}^{+\infty} \Phi^n(\lambda t)\psi(t;\nu)dt = [c_n(\nu,\lambda)]^{-1},$$

we conclude that $P(X \le x) = \int_{-\infty}^{x} c_n(\nu;\lambda)\psi(t;\nu)\Phi^n(\lambda t)dt$.

***Theorem 7.*** If $Z \sim N(0,1)$ be independent of $X \sim G_3StN(\nu,\lambda,n-1)$, then $X \mid \{\lambda X > Z\} \sim G_3StN(\nu,\lambda,n)$.

***Proof.*** Let $Y \equiv X \mid \{\lambda X > Z\}$, then $P(Y \le y) = \dfrac{P(X \le y, \lambda X > Z)}{P(\lambda X > Z)}$. Since

$$P(X \le y, \lambda X > Z)$$
$$= \int_{-\infty}^{+\infty} P(X \le y, \lambda X > Z \mid x) f_{(n-1)}(x;\nu,\lambda)dx$$
$$= \int_{-\infty}^{\infty} I_{(-\infty,y]}(x) P(Z \le \lambda x) f_{(n-1)}(x;\nu,\lambda)dx$$
$$= \int_{-\infty}^{y} c_{n-1}(\nu,\lambda)\psi(x;\nu)\Phi^n(\lambda x)dx,$$

and

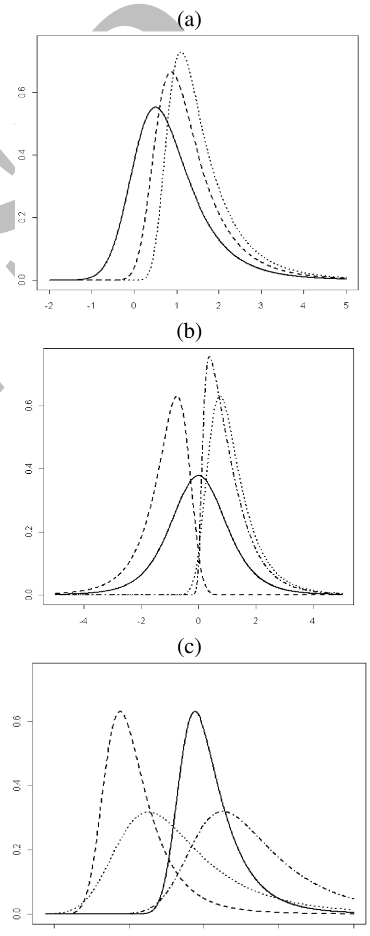



**Figure 2**. Plots of $G_3StN(\nu,\lambda,n)$ for (a) $\lambda=2$ and (solid line) n=1, (dashed line) n=2, (dotted line) n=5 (b) $n=3$ and (solid line) $\lambda=0$, (dashed line) $\lambda=-2$, (dotted line) $\lambda=2$, (dashed-dotted line) $\lambda=6$ and (c) Plots of $G_3StN(\xi,\omega,\nu,\lambda,n)$ for $\nu=5$, $\lambda=2$, $n=3$ and various values of location and scale parameters: (solid line) $(\xi,\omega)=(1,1)$, (dashed line) $(\xi,\omega)=(-1,1)$, (dotted line) $(\xi,\omega)=(-1,2)$, (dashed-dotted line) $(\xi,\omega)=(1,2)$.

$$P(\lambda X > Z)$$
$$= \int_{-\infty}^{+\infty} P(\lambda X > Z \mid x) f_{(n-1)}(x;\nu,\lambda)dx$$
$$= \int_{-\infty}^{\infty} P(Z \le \lambda x) c_{n-1}(\nu,\lambda)\psi(x;\nu)\Phi^{n-1}(\lambda x)dx$$
$$= \int_{-\infty}^{y} c_{n-1}(\nu,\lambda)\psi(x;\nu)\Phi^{n}(\lambda x)dx = \frac{c_{n-1}(\nu,\lambda)}{c_n(\nu,\lambda)},$$

we conclude that $P(Y \le y) = \int_{-\infty}^{y} c_n(\nu,\lambda)\psi(x;\nu)\Phi^n(\lambda x)dx$.

***Definition 4.*** The location-scale $G_3StN(\nu,\lambda)$ distribution is defined as the distribution of $Y = \omega X + \xi$, where $X \sim G_3StN(\nu,\lambda,n)$, $\omega > 0$, $\xi \in$ R. Its density is given by

$$f_Y(y;\xi,\omega,\nu,\lambda)$$
$$= \frac{1}{\omega} c_n(\nu,\lambda)\psi(\frac{y-\xi}{\omega};\nu)\Phi^n(\lambda(\frac{y-\xi}{\omega})),$$

and we denote it by $Y \sim G_3StN(\xi,\omega,\nu,\lambda)$.

The impact of location and scale parameters of the curves of location-scale $G_3$StN distribution, which determine the shift and dispersion of the distribution, respectively, is shown in Figure 2(c).

### *Type Four*

In this section we generalize StN to a new family of distributions, which are multimodal and their modes can be controlled by two parameters $\lambda_1$ and $\lambda_2$.

***Definition 5.*** A random variable X with density function

$$f(x;\nu,\lambda_1,\lambda_2) = 2\psi(x;\nu)\Phi(\frac{\lambda_1 x + \lambda_2 x^3}{\sqrt{1+\lambda_2 x^2}}),$$

is said to have a generalized StN of type four with parameters $\lambda_1 \in R$, $\lambda_2 \ge 0$, and it is denoted by $X \sim G_4StN(\nu,\lambda_1,\lambda_2)$.

The plots of this distribution for various values of $\lambda_1$ and $\lambda_2$ are shown in Figure 3. Let $X \sim G_4StN(\nu,\lambda,n)$, then it is straightforward to prove the following properties:

- If $\lambda_2 = 0$, then $X \sim StN(\nu,\lambda)$,
- If $\lambda_1 = \lambda_2 = 0$, then $X \sim t(\nu)$,
- $\lim_{\lambda_1 \to \infty} f(x;\nu,\lambda_1,\lambda_2) = 2\psi(x;\nu)I_{(x>0)}$,
- $\lim_{\lambda_1 \to -\infty} f(x;\nu,\lambda_1,\lambda_2) = 2\psi(x;\nu)I_{(x<0)}$,
- If $T \sim t(\nu)$ and $\tau(\cdot)$ is an even function, then $\tau(X) \overset{D}{=} \tau(T)$,
- If $T \sim t(\nu)$, then $|X| \overset{D}{=} |T|$.

Now we present four theorems for simulating a random variable from $G_4StN(\nu,\lambda_1,\lambda_2)$.

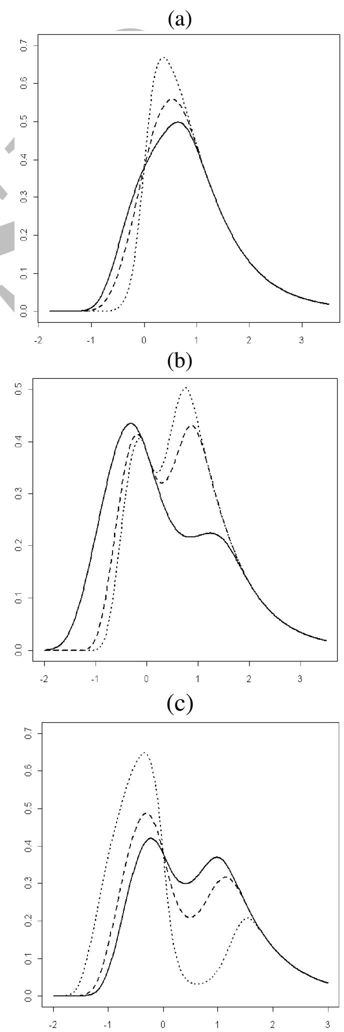


**Figure 3**. Plots of $G_3StN(\nu,\lambda,n)$ for $\nu = 5$ and (a) $(\lambda_1,\lambda_2) = (1,3),(2,3),(5,3)$, (b) $(\lambda_1,\lambda_2) = (-1,3),(-2,3),(-5,3)$, (c) $(\lambda_1,\lambda_2) = (-1,1),(-1,5),(-1,10)$ denoted by solid, dashed and dotted lines, respectively.

***Theorem 8.*** Let $T \sim t(\nu)$ be independent of $Z \sim N(0,1)$ and $\lambda_1 \in R$, $\lambda_2 \geq 0$, then

$$T \mid \{Z \leq \frac{\lambda_1 T + \lambda_2 T^3}{\sqrt{1+\lambda_2 T^2}}\} \sim G_4StN(\nu, \lambda_1, \lambda_2).$$

***Proof.***

$$f(t \mid Z \leq \frac{\lambda_1 T + \lambda_2 T^3}{\sqrt{1+\lambda_2 T^2}}) = \frac{P(Z \leq \frac{\lambda_1 T + \lambda_2 T^3}{\sqrt{1+\lambda_2 T^2}} \mid t)\psi(t;\nu)}{P(Z \leq \frac{\lambda_1 T + \lambda_2 T^3}{\sqrt{1+\lambda_2 T^2}})}$$

$$= \frac{\Phi(\frac{\lambda_1 t + \lambda_2 t^3}{\sqrt{1+\lambda_2 t^2}})\psi(t;\nu)}{\int_{-\infty}^{\infty}\Phi(\frac{\lambda_1 t + \lambda_2 t^3}{\sqrt{1+\lambda_2 t^2}})\psi(t;\nu)dt}$$

$$= 2\psi(t;\nu)\Phi(\frac{\lambda_1 t + \lambda_2 t^3}{\sqrt{1+\lambda_2 t^2}}).$$

***Theorem 9.*** Let $T \sim t(\nu)$ be independent of $Z \sim N(0,1)$ and $\lambda_1 \in R$, $\lambda_2 \geq 0$. If we define $X$ to be equal to $T$ when $Z \leq \frac{\lambda_1 T + \lambda_2 T^3}{\sqrt{1+\lambda_2 T^2}}$, otherwise it is equal to $-T$, then $X \sim G_4StN(\nu, \lambda_1, \lambda_2)$.

***Proof.***

$$f_X(x) = \psi(x;\nu)[\Phi(\frac{\lambda_1 x + \lambda_2 x^3}{\sqrt{1+\lambda_2 x^2}}) + 1 - \Phi(-\frac{\lambda_1 x + \lambda_2 x^3}{\sqrt{1+\lambda_2 x^2}})]$$

$$= 2\psi(x;\nu)\Phi(\frac{\lambda_1 x + \lambda_2 x^3}{\sqrt{1+\lambda_2 x^2}}).$$

***Theorem 10.*** Let $T \sim t(\nu)$ be independent of $U \sim U(0,1)$ and $\lambda_1 \in R$, $\lambda_2 \geq 0$, then

$$T \mid \{U \leq \Phi(\frac{\lambda_1 T + \lambda_2 T^3}{\sqrt{1+\lambda_2 T^2}})\} \sim G_4StN(\nu, \lambda_1, \lambda_2).$$

***Proof.***

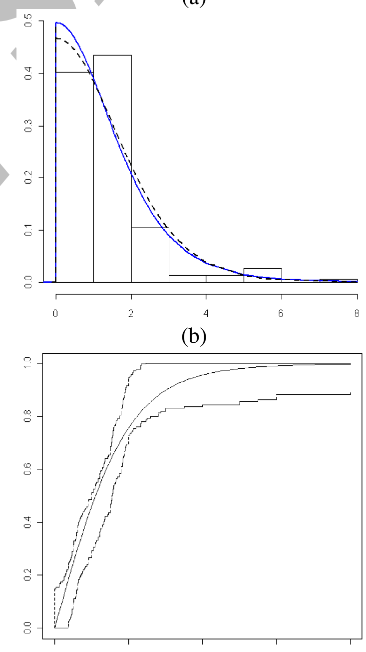
$$f(t \mid U \leq \Phi(\frac{\lambda_1 T + \lambda_2 T^3}{\sqrt{1+\lambda_2 T^2}}))$$

$$= \frac{P(U \leq \Phi(\frac{\lambda_1 T + \lambda_2 T^3}{\sqrt{1+\lambda_2 T^2}}) \mid t)\psi(t;\nu)}{P(U \leq \Phi(\frac{\lambda_1 T + \lambda_2 T^3}{\sqrt{1+\lambda_2 T^2}}))}$$

$$= \frac{\Phi(\frac{\lambda_1 t + \lambda_2 t^3}{\sqrt{1+\lambda_2 t^2}})\psi(t;\nu)}{\int_{-\infty}^{\infty}\Phi(\frac{\lambda_1 t + \lambda_2 t^3}{\sqrt{1+\lambda_2 t^2}})\psi(t;\nu)dt}$$

$$= 2\Phi(\frac{\lambda_1 t + \lambda_2 t^3}{\sqrt{1+\lambda_2 t^2}})\psi(t;\nu).$$

(a)



(b)

**Figure 4**. (a) Histogram of Shadgan wetland data, together with fitted distributions using maximum likelihood estimators: (solid line) $G_2SN(\hat{\xi},\hat{\omega},\hat{\nu},\hat{\lambda}_1,\hat{\lambda}_2)$ ; (dashed line) $G_4SN(\hat{\xi},\hat{\omega},\hat{\nu},\hat{\lambda}_1,\hat{\lambda}_2)$ and (b) Plot of 95% confidence bands for cumulative distribution of $G_2StN$.

**Table 1.** Descriptive statistics for the $Ni\ /V$

| n | Mean | SD | Asymmetry Coeff. | Kurtosis Coeff. |
|---|---|---|---|---|
| 154 | 1.4634 | 1.2587 | 2.2553 | 6.9446 |

**Table 2.** The MLE of the model parameters and their K-S statistic and Log-likelihood values

| | SN | St | StN | $G_1$StN | $G_2$StN | $G_3$StN | $G_4$StN |
|---|---|---|---|---|---|---|---|
| ν | - | 35.3233 | 14.0007 | 14.058 | 5.5793 | 200.6456 | 10.7331 |
| $\lambda_1$ | 3208343 | 2916707 | 3208320 | 3529135 | 3208346 | 4297559 | 3208314 |
| $\lambda_2$ | - | - | - | 4.17778 | -15.6099 | - | 40.8125 |
| n | - | - | - | - | - | 278993 | - |
| ξ | -3.19e-6 | -3.92e-6 | -2.93e-6 | -2.60e-6 | -2.29e-6 | -1.41e-3 | -7.48e-4 |
| ω | 2.2339 | 2.2332 | 1.9999 | 1.9999 | 1.5337 | 1.9175 | 1.6461 |
| K-S | 0.2083 | 0.2142 | 0.1719 | 0.1719 | 0.0818 | 0.1357 | 0.0883 |
| AIC | -433.76 | -435.18 | -426.10 | -428.10 | -418.66 | -430.74 | -420.49 |

***Theorem 11.*** Let $T \sim t(\nu)$ be independent of $U \sim U(0,1)$ and $\lambda_1 \in R$, $\lambda_2 \geq 0$. If we define $X$ to be equal to $T$ when $U \leq \Phi(\frac{\lambda_1 T + \lambda_2 T^3}{\sqrt{1+\lambda_2 T^2}})$, otherwise equal to $-T$, then $X \sim G_4StN(\nu, \lambda_1, \lambda_2)$.

***Proof.*** Proof is similar to Theorem 9.

***Model Fitting of Shadegan Data***

In order to present an application of GStN distributions, we carry out the estimation and model fitting to a real data set by maximizing the likelihood functions corresponding to the skewed distributions considered in this paper. Then we propose the best fitted distribution to the data by comparing the log-likelihood values of the models. Shadegan wetland extends over an area of 296,000 hectares in the south-west of Iran at the head of the Persian Gulf. A great variety of plants cover parts of it. This creates a suitable habitat for a diverse group of migrating waterfowls, which fly to this area from Northern Europe, Russia and Siberia in autumn. As such, Shadegan wetland plays a significant hydrological and ecological role in the natural functioning of the Northern Persian Gulf. But shipping oil and related activities of the oil industry has caused chronic oil pollution in the Persian Gulf region. The release of oil from destroyed oil terminals and damaged oil tankers following the occupation of Kuwait by Iraq in 1991 introduced catastrophic quantities of petroleum into the Persian Gulf. Precipitation of black rain and deposition of soot resulting from burning Kuwait oilwell caused widespread contamination in many parts of Iran including the Shadegan wetland. Nickel ($Ni$) and Vanadium ($V$) are two heavy metals associated with crude oil. The ratio $Ni/V$ is usually used for identification of oil source. Where $Ni/V < 3$, Iran's oil resources are recognized as the main cause of the pollution. Otherwise, the oil resources of the Persian Gulf countries would be counted so. In addition, the probability distribution of this ratio is applicable in estimation of the pattern of pollution load and planning the future sampling points in the field following best decision of environmental manager. So, we fit a suitable distribution to the data of ratio $Ni/V$ in Shadgan wetland.

Table 1 presents basic descriptive statistics for the Shadegan data set. The sample asymmetry and kurtosis coefficients show the skewness of distribution of the data. The maximum likelihood estimates (MLE) of the parameters, the values of the Kolmogorov-Smirnov (K-S) goodness-of-fit test statistics and the model selection criterion AIC(twice the log-likelihood minus twice the number of parameters) for Azzalani's SN and St distributions together with StN and four GStN distributions are presented in Table 2. Note that, the critical point of K-S goodness-of-fit test with confidence level $\alpha = 0.05$ is equal to $0.1096$. Therefore, goodness-of-fit test for all of these models except $G_2$StN and $G_4$StN distributions are significant. The AIC values in Table 2 also show that fitting a SN model to these data would be inadequate and this is reflected in that the fitted St model is almost indistinguishable from a SN model. The histogram of the data together with the fitted densities of $G_2$StN and $G_4$StN are illustrated in Figure 4(a). Based on AIC values, the $G_2$StN distribution fits much better than $G_4$StN. Therefore, we conclude that for Shadegan data, $G_2$StN distribution is the most appropriate fitted model among all other skewed distributions considered here. Figure 4(b) shows the 95% confidence bands, determined by the K-S goodness-of-fit test, for cumulative distribution of $G_2$StN. It illustrates that the fitted distribution is entirely in the confidence region, so this model suits the data well. Based on the $G_2StN$ and SN models the pollution probability obtained by other countries are 0.1 and 0.18 respectively, whereas the empirical pollution probability provided by the Shadgan data is 0.07. It is clear that the fitted $G_2StN$ model gives pollution probability more closely to the empirical value than the SN model.

## Results and Discussion

It is typical in some applications that data follow asymmetric distributions. In this paper four generalizations of StN distribution are introduced and their properties with methods for random number generation presented. The application section shows that the SN and St distributions are not suitable enough to model the skewness and kurtosis in the Shadegan data, as is the case of the GStN distributions proposed in this paper. therefore, this model can be more suitable than SN and St distributions in fitting data sets with strong degrees of asymmetry and kurtosis.

## Acknowledgement

The authors wish to thank the referees and the editor for their useful comments and suggestions. Partial support from Ordered and Spatial Data Center of Excellence of Ferdowsi University of Mashhad is acknowledged.

## References

1. Arellano-Valle, R.B., Gomez, H. W., and Quintana, F. A., A New Class of Skew-normal Distribution. Communications in Statistics: *Theory and Methods*., **33**: 1465-1480 (2004).
2. Arnold B.C., and Beaver, R.J., Skewed Multivariate Models Related to Hidden Truncation and/or Selective Reporting. *Test*, **11**: 7-54 (2002).
3. Azzalini, A., A class of distribution which includes the normal ones. *Scandinavian Journal of Statistics,* **12**: 171-178 (1985).
4. Azzalini, A., Further Results on a Class of Distributions which Includes the Normal Ones. *Statistica,* **46**: 199-208 (1986).
5. Azzalini, A., The Skew-normal Distribution and Related Multivariate Families. *Scandinavian Journal of Statistics,* **32**: 159-188 (2005).
6. Azzalini, A. and Capitanio, A., Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t Distribution. *Journal of the Royal Statistical Society: Series B.,* **65**: 367-389 (2003).
7. Azzalini, A. and Chiogna, M., Some Results on the Stress-strength Model for Skew Normal Variate. *Metron LXII.,* 315-326 (2004).
8. Azzalini, A., and Dalla Valla, A., The Multivariate Skew-normal Distribution. *Biometrika.* **83**: 715-726. (1996).
9. Balakrishnan, N., Discussion on "Skew Multivariate Models Related to Hidden Truncation and/or Selective Reporting" by B. C. Arnold and R. J. Beaver. *Test.,* **11**: 37-39 (2002).
10. Branco, M., and Dey, D.K., A General Class of Multivariate Elliptical Distribution. *Journal of Multivariate Analysis.,* **79**: 99-113 (2001).
11. Ellison, B., Two Theorems for Inferences about the Normal Distribution with Applications in Acceptance Sampling. *Journal of the American Statistical Association*., **59**: 89-95 (1964).
12. Gupta, R.C. and Gupta, R.D., Generalized Skew Normal Model*, Test,* **13**: 501-524 (2004).
13. Gomez, H.W., Venegas, O. and Bolfarine, H., Skew-symmetric Distributions Generated by the Distribution Function of the Normal Distribution. *Environmetrics.***18**: 395-407 (2007).
14. Henze, N.A., A Probabilistic Representation of the Skew-normal Distribution. *Scandinavian Journal of Statistics,* **13**: 271-275 (1986).
15. Jones, M.C., and Faddy, M. J., A Skew Extension of the t-distribution, with Applications. *Journal of the Royal Statistical Society*, Series B., **65**: 159-174 (2003).
16. Kheradmandi, A., Mohammadzadeh, M. and Sanjari Farsipour N., Properties of Skew t-Normal Distribution and Modeling of Shadegan Wetland Pollution Data, Proceedings of the 10th Iranian Statistical Conference, Tabriz, Iran, 164-174 (2010).
17. Ma, Y., and Genton, M.G., Flexible Class of Skew-symmetric Distribution. *Scandinavian Journal of Statistics,* **31**: 459-468 (2004).
18. Sharafi, M., and Behboodian, J., The Balakrishnan Skew-normal Density, *Statistical Papers,* **49**: 769-778 (2008).
19. Yadegari, I., Gerami, A. and Khaledi, M. J., A generalization of the Balakrishnan skew-normal distribution. *Statistics and Probability Letters,* **78**: 1165-1167 (2008).