Journal of Statistical Research *https://doi.org/10.3329/jsr.v59i1.83682*
2025, Vol. 59, No. 1, pp. 31-45 *ISSN 0256 - 422 X*

ON SOME GENERAL CLASS OF MODIFIED SKEW NORMAL DISTRIBUTION AND ITS PROPERTIES

C. SATHEESH KUMAR*
Department of Statistics, University of Kerala, Trivandrum 695 581, India
Email: drcsatheeshkumar@gmail.com

M. R. ANUSREE
Department of Statistics, St. Thomas College (Autonomous), Thrissur- 680 001, India
Email: anun74@gmail.com

SUMMARY

In this paper, we propose a new class of modified skew-normal distribution as a special case of the modified skew-normal distribution proposed by Kumar and Anusree (2014) is considered. We explore some of its key properties by deriving explicit expressions for its distribution function, characteristic function, and other relevant aspects using special functions and simplified moment expressions. Also, its distributional and structural properties are investigated. Additionally, we address the estimation of parameters for this general class of distributions.

Keywords and phrases: Entropy, Skew normal distribution, Moments, Special functions

1 Introduction

The normal or Gaussian distribution is symmetric and is widely accepted as the basis of many statistical work. But normal distribution is not suitable for the situation when the data is not symmetric. There has been a renewed interest in the development of asymmetric versions of normal distribution during the last three decades. For details, see Azzalini (1985), Genton (2004). Azzalini (1985) introduced the skew normal distribution and has received much attention due to its flexibility and mathematical tractability. The probability distribution of the Azzalini's skew normal distribution has the following form. For x $\in$ R, λ $\in$R,

$$f(x, \lambda) = 2\phi(x)\Phi(\lambda x), \tag{1.1}$$

where $\phi(\cdot)$and $\Phi(\cdot)$are respectively the probability density function (p.d.f.) and cumulative distribution function (c.d.f.) of a standard normal variate. The density is unimodal and is also having both positive and negative skewness. Following the work of Azzallini (1985) several generalizations came forward see, Azzalini (1986), Azzalini and Dalla-Valle (1996), Azzalini and Capitanio

* Corresponding author

(1999), Kumar and Anusree (2011, 2014), Shakil et. al (2014), Ahsanullah et. al (2015), Mondal et.al (2024) and several others. In this article, we study the key properties of a special class of distributions, which is a subclass of the skew normal distribution proposed by Kumar and Anusree (2014). Our attention relies on special functions and moment expressions, which play a crucial role in adjusting tail behavior and enhancing flexibility in modeling asymmetric data. These distributions have significant applications in risk modeling, actuarial science, and survival analysis, where skewness is an essential factor. Moreover, special functions enable controlled skewness while preserving the fundamental structure of the normal distribution.The paper is organized as follows: Section 2 introduces the modified skew normal distribution and discuss its fundamental properties, including the distribution function, characteristic function, moment-generating function, raw moments, skewness, and kurtosis. Section 3 explores information measures, such as Shannon entropy and Mille's ratio, for this class of distributions. Section 4 presents a location-scale extension of the distribution and examines the maximum likelihood estimation (MLE) of its parameters. Section 5, provides a numerical example to illustrate the practical usefulness of this new class of distributions, also summary and conclusion are included in section 6. Throughout our study, we utilize special functions and lemmas established by Gupta et al. (2013), which are essential for deriving key results.

(i) For the standard normal density function, the characteristic function is given by

$$\Psi(t) = e^{-\frac{t^2}{2}}$$

(ii) The special function $\mathrm{erf}(\cdot)$, which finds some relationship with ...

1. the normal distribution is $erf(z) = \frac{2}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}} dt$

2. the hypergeometric function is $erf(z) = \frac{2ze^{-\frac{z^2}{2}}}{\sqrt{\pi}} {}_1F_1(1, 3/2, z^2)$ or $erf(z) = \frac{2z}{\sqrt{\pi}} {}_1F_1(1/2, 3/2, -z^2)$, where ${}_1F_1(.)$is the hypergeometric function and is defined by

$$ {}_1F_1(a_1, a_2, \ldots, a_p; b_1, b_2, \ldots b_q; z) = \sum_{k=0}^{\infty} \frac{(a_1)_{\{k\}}(a_2)_{\{k\}} \ldots (a_p)_{\{k\}} z^k}{(b_1)_{\{k\}}(b_1)_{\{k\}} \ldots (b_q)_{\{k\}} k!},$$

   where $a_{\{k\}} = a(a+1)\ldots(a+k-1)$, $a^{\{k\}} = a(a-1)\ldots(a-k+1)$ with $a_{\{0\}} = a^{\{0\}} = 1$

3. the infinite power series is $erf(z) = \frac{2}{\sqrt{\pi}} \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k+1}}{k!(2k+1)}$ and $erf(z) = \frac{2}{\sqrt{\pi}} e^{-z^2} \sum_{k=0}^{\infty} \frac{2^k z^{2k+1}}{\prod_{j=0}^{k} k!(2j+1)}$

Lemma 1.1. *The c.d.f. of the normal distribution can be expressed as*

$$\Phi(x) = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sigma/\sqrt{2}}\right)\right],$$

where $\mu \in R,\ \sigma > 0.$

Lemma 1.2. *Let* $u = \exp(t^2/2)$*, which is the moment generating function of a standard normal variable, then*

$$E[Z^k] = \begin{cases} \frac{k!}{2^{k/2}(k/2)!}, & \text{if } k \text{ is even} \\ 0, & \text{if } k \text{ is odd} \end{cases}$$

Lemma 1.3. *Let* $u = exp(t^2/2)$*, the moment generating function of a standard normal variable, then*

$$E[Z^k] = \frac{1 + (-1)^k}{2} \frac{\Gamma(k+1)}{2^{k/2}\Gamma\left(k/2+1\right)}.$$

Lemma 1.4. *If* $u = \exp(t^2/(2\beta))$ *and* $\nu = exp(\ -\beta(z - t/\beta)^2/2),\ then\ \frac{d^k}{dt^k}(uv) = z^k uv$

2 Modified Skew Normal Distribution

In this section, we consider the definition of the modified generalized skew normal distribution and discuss the properties of its special case, known as the modified skew normal distribution. First we define the modified generalized skew normal distribution as in the following. A random variable Z is said to have a modified generalized skew normal distribution, (MGSND) if its p.d.f. takes the following form, in which $z \in R,\ \lambda_1 \in R$,$\lambda_2 \geq 0$ and , $\alpha \in [0, 1]$

Definition 2.1.

$$h(z; \lambda_1, \lambda_2, \alpha) = \phi(z) \left[\alpha + 2(1-\alpha)\Phi\left(\frac{\lambda_1 z}{\sqrt{1+\lambda_2^2 z^2}}\right)\right]. \tag{2.1}$$

Definition 2.2. A random variable Z is said to have a modified skew normal distribution, (MSND) if its p.d.f. takes the following form, in which $z \in R,\ \lambda \in R$and $\alpha \in [0, 1]$

$$h(z; \lambda, \alpha) = \phi(z) \left[\alpha + 2(1-\alpha)\Phi(\lambda z)\right]. \tag{2.2}$$

A distribution with p.d.f.(2.2) hereafter we written as MSND (λ, α). For particular values of λ and α, the MSND (λ, α) reduces to the following special cases.

Properties

1. $h(z; \lambda, \alpha)$ modifies the base pdf $\phi(z)$ by incorporating its own cdf $\Phi(\lambda z)$.
2. the term $\alpha + 2\left(1-\alpha\right)\Phi\left(\lambda z\right)$acts as a weight adjustment that depends on $\Phi(\lambda z)$
3. when α =1, $h(z; \lambda, \alpha) = \phi(z)$, meaning no modification and
4. when α =0, the weight depends purely on $\Phi(\lambda z)$, amplifying or suppressing parts of $\phi(z)$.

The p.d.f. of the MSND (λ, α) given in (2.2) is plotted for particular choice of λ and α and compared with the p.d.f. of normal distribution are presented in Figure 1.

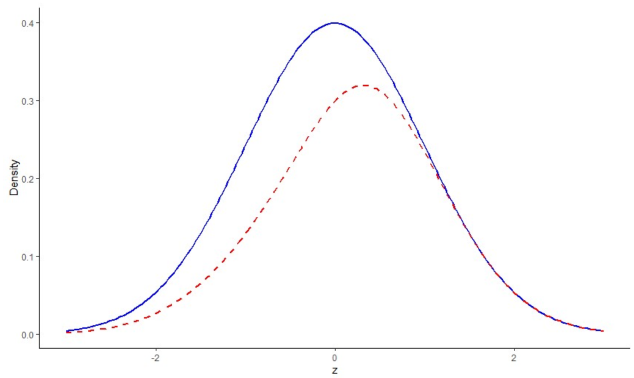

Figure 1: Probability plots of normal distribution and modified skew normal distribution

Result 2.1. If Z follows MSND (λ, α), then its c.d.f., $H(z;\lambda,\alpha)$ is the following

$$H(z;\lambda,\alpha) = \varPhi(z) + \frac{\lambda}{\pi}(1-\alpha)\,\varPsi\left(\lambda z/\sqrt{2}\right)$$

Where $\varPsi\left(\lambda z/\sqrt{2}\right) = \int_{-\infty}^{z} te^{-(1+\lambda^2)t^2/2}{}_1F_1(1,3/2,(\lambda t/\sqrt{2}))dt.$

Proof. The c.d.f. of MSND(λ,α) can be obtained in the following way

$$\begin{aligned}
H(z;\lambda,\alpha) &= \int_{-\infty}^{z} h(t;\lambda,\alpha)dt \\
&= \int_{-\infty}^{z} \phi(t)\left[\alpha + 2(1-\alpha)\varPhi(\lambda t)\right]dt \\
&= \alpha\int_{-\infty}^{z} \phi(t)dt + 2(1-\alpha)\int_{-\infty}^{z} \phi(t)\varPhi(\lambda t)dt \\
&= \alpha\varPhi(z) + 2(1-\alpha)\int_{-\infty}^{z} \frac{e^{-\frac{t^2}{2}}}{2\sqrt{2\pi}}\left[1 + \text{erf}\left(\frac{\lambda t}{\sqrt{2}}\right)\right]dt \\
&= \alpha\varPhi(z) + (1-\alpha)\varPhi(z) + (1-\alpha)\int_{-\infty}^{z} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\text{erf}\left(\frac{\lambda t}{\sqrt{2}}\right)dt \\
&= \varPhi(z) + (1-\alpha)\int_{-\infty}^{z} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\text{erf}\left(\frac{\lambda t}{\sqrt{2}}\right)dt \\
&= \varPhi(z) + (1-\alpha)\int_{-\infty}^{z} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}2\left(\frac{\lambda t}{\sqrt{2}}\right)e^{-\left(\frac{\lambda t}{\sqrt{2}}\right)^2}{}_1F_1\left(1,\frac{3}{2},\left(\frac{\lambda t}{\sqrt{2}}\right)^2\right)\text{erf}\left(\frac{\lambda t}{\sqrt{2}}\right)dt \\
&= \varPhi(z) + (1-\alpha)\frac{\lambda}{\pi}\varPsi\left(\frac{\lambda z}{\sqrt{2}}\right),
\end{aligned}$$

where $\varPsi\left(\lambda z/\sqrt{2}\right) = \int_{-\infty}^{z} te^{-\frac{t^2}{2}(1+\lambda^2)}{}_1F_1\left(1,3/2,\left(\frac{\lambda t}{\sqrt{2}}\right)^2\right)dt.$

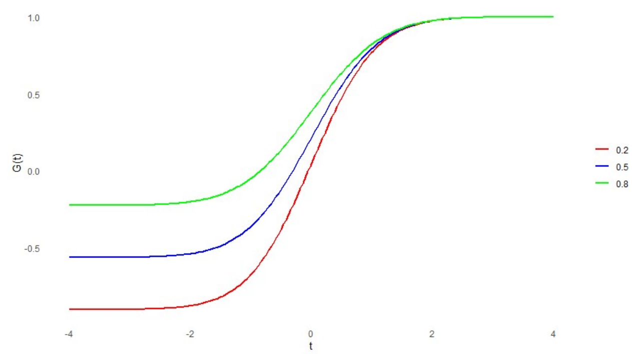

Figure 2: CDF plots of $g(t) = h(z; \lambda, \alpha)$ for $\lambda = 1$ and $\alpha = 0.2, 0.5,$ and 0.8

Result 2.2. If Z follows MSND (λ, α), then its hazard function $HD(z; \lambda, \alpha)$ is the following

$$HD(z; \lambda, \alpha) = \frac{\phi(z)\,[\alpha + 2(1-\alpha)\,\Phi(\lambda z)]}{1 - \Phi(z) + \frac{\lambda}{\pi}(1-\alpha)\,\Psi\left(\frac{\lambda z}{\sqrt{2}}\right)}$$

where $\Psi\left(\lambda z/\sqrt{2}\right) = \int_{-\infty}^{z} t e^{-(1+\lambda^2)t^2/2} {}_1F_1(1, 3/2, (\lambda t/\sqrt{2}))dt.$

Proof. Proof follows from the definition of hazard function and failure rate.

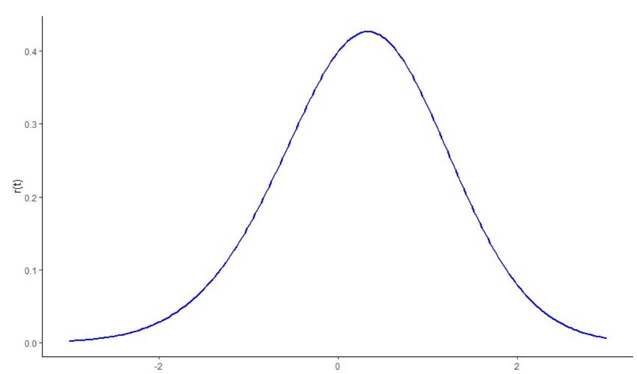

Figure 3: Plots of hazard function, r(t)= $HD(t; \lambda, \alpha)$

It exhibits a bell-shaped curve, indicating that the failure rate increases to a peak and then declines.

Result 2.3. If Z follows MSND (λ, α), then its characteristic function can take any of the following two forms,

$$(i).\ \Psi_1(t) = e^{-\frac{t^2}{2}} + i\tau(\delta t)e^{-\frac{t^2}{2}}\,(ii).\ \Psi_2(t) = e^{-\frac{t^2}{2}} + \frac{(1-\alpha)\lambda}{\pi}\Omega_1(t),$$

Where $\Omega_1(t) = \int_{-\infty}^{\infty} z e^{-\frac{(1+\lambda^2)}{2}\left(z-\frac{it}{1+\lambda^2}\right)^2} {}_1F_1\left(1, 3/2, {}^{\lambda^2 z^2}\!/_2\right) dz$

Proof. By the definition of characteristic function,

$$\begin{aligned}\Psi_1(t) &= \int_{-\infty}^{\infty} e^{itz}\phi(z)\left[\alpha + 2(1-\alpha)\Phi(\lambda z)\right] dz \\ &= \alpha e^{-\frac{t^2}{2}} + (1-\alpha)e^{-\frac{t^2}{2}}\left[1 + i\tau(\delta t)\right] \\ &= e^{-\frac{t^2}{2}} + i\tau(\delta t)e^{-\frac{t^2}{2}}.\end{aligned}$$

In the light of the expressions given in Pewsey (2000) and $\tau(x) = \int_0^x \sqrt{\frac{2}{\pi}} e^{-\frac{u^2}{2}} du,\ x > 0,\ \delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$, $\tau(-x) = \tau(x)$ which simplifies (i). Also,

$$\begin{aligned}\Psi_2(t) &= \int_{-\infty}^{\infty} e^{itz}\phi(z)\left[\alpha + 2(1-\alpha)\Phi(\lambda z)\right] dz \\ &= \alpha e^{-\frac{t^2}{2}} + (1-\alpha)\int_{-\infty}^{\infty} 2e^{itz}\phi(z)\Phi(\lambda z) dz \\ &= \alpha e^{-\frac{t^2}{2}} + (1-\alpha)\int_{-\infty}^{\infty} 2e^{itz}\phi(z)\left[1 + \text{erf}\left(\frac{\lambda z}{\sqrt{2}}\right)\right] dz \\ &= e^{-\frac{t^2}{2}} + (1-\alpha)\int_{-\infty}^{\infty} e^{itz}\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\frac{2}{\sqrt{\pi}}\left(\frac{\lambda z}{\sqrt{2}}\right) e^{-\frac{\lambda^2 z^2}{2}} {}_1F_1\left(1, \frac{3}{2}, \frac{\lambda^2 z^2}{2}\right) dz, \quad \text{by Lemma 1.4} \\ &= e^{-\frac{t^2}{2}} + (1-\alpha)\frac{\lambda}{\pi}\int_{-\infty}^{\infty} z e^{-\frac{1+\lambda^2}{2}\left(z-\frac{it}{1+\lambda^2}\right)^2} {}_1F_1\left(1, \frac{3}{2}, \frac{\lambda^2 z^2}{2}\right) dz \\ &= e^{-\frac{t^2}{2}} + (1-\alpha)\frac{\lambda}{\pi}\Omega_1(t).\end{aligned}$$

Simplifies (ii) in which $\Omega_1(t) = \int_{-\infty}^{\infty} z e^{-\frac{(1+\lambda^2)}{2}\left(z-\frac{it}{1+\lambda^2}\right)^2} {}_1F_1\left(1, 3/2, {}^{\lambda^2 z^2}\!/_2\right) dz$.

Result 2.4. If Z follows MSND(λ, α), then the rth moment of any integer order k>-2 is given by

$$\mu_r' = \begin{cases} \mu_1(k),\ k\ is\ even \\ \mu_2(k),\ k\ is\ odd \end{cases} \tag{2.3}$$

Proof. By differentiating the moment generating function 'r' times and then putting t=0 gives the rth raw moment.

$$\begin{aligned} E(Z^k) &= \frac{d^k}{dt^k}[M_Z(t)]_{t=0} \\ &= \frac{d^k}{dt^k}(e^{t^2/2}) + \frac{(1-\alpha)\lambda}{\pi}\frac{d^k}{dt^k}(e^{t^2/2})\int_R ze^{-\frac{(z-t)^2}{2}}{}_1F_1\left(1/2, 3/2, -\frac{\lambda^2 z^2}{2}\right)dz \\ &= I_1(t) + I_2(z, t; \lambda, \alpha) \end{aligned}$$

Now,

$$I_1(t) = \frac{d^k}{dt^k}(e^{\frac{t^2}{2}})_{t=0} = \mu_1(k) \ = \begin{cases} \frac{k!}{2^{\frac{k}{2}}(k/2)!}, & k \ is \ even \\ 0, & k \ is \ odd \end{cases} \tag{2.4}$$

$$\begin{aligned} I_2(z, t; \lambda, \alpha) &= (1-\alpha)\frac{\lambda}{\pi}\frac{d^k}{dt^k}\int_{\mathbb{R}} ze^{-\frac{(z-t)^2}{2}}e^{\frac{t^2}{2}}{}_1F_1\left(\frac{1}{2}, \frac{3}{2}, -\frac{\lambda^2 z^2}{2}\right)dz \\ &= (1-\alpha)\frac{\lambda}{\pi}\frac{d^k}{dt^k}\int_{\mathbb{R}} z{}_1F_1\left(\frac{1}{2}, \frac{3}{2}, -\frac{\lambda^2 z^2}{2}\right)e^{-\frac{(z-t)^2}{2}}e^{\frac{t^2}{2}}dz \\ &= (1-\alpha)\frac{\lambda}{\pi}\int_{\mathbb{R}} z{}_1F_1\left(\frac{1}{2}, \frac{3}{2}, -\frac{\lambda^2 z^2}{2}\right)\frac{d^k}{dt^k}e^{-\frac{(z-t)^2}{2}}e^{\frac{t^2}{2}}dz \\ &= (1-\alpha)\frac{\lambda}{\pi}\frac{d^k}{dt^k}\int_{\mathbb{R}} z{}_1F_1\left(\frac{1}{2}, \frac{3}{2}, -\frac{\lambda^2 z^2}{2}\right)z^k e^{-\frac{(z-t)^2}{2}}e^{\frac{t^2}{2}}dz, \quad \text{by Lemma 1.4} \end{aligned}$$

By Lemma 1.3, the expected value can be written in terms of gamma functions, therefore,

$$I_2(z, t; \lambda, \alpha)_{t=0} = \frac{(1-\alpha)\lambda\sqrt{2\pi}}{\pi}\sum_{j=0}^{\infty}\Gamma\,(j+1/2)\Gamma(3/2)\frac{(-\lambda^2)^j}{j!2^j}\frac{1+(-1)^{2j+k+1}}{2}\frac{\Gamma(2j+k+2)}{2^{(2j+k+1)/2}\Gamma[(2j+k+1)/2+1]}$$

using the gamma duplication formula

$$\Gamma(2y)\sqrt{\pi} = 2^{2y-1}\Gamma(y)\Gamma(y+1/2)$$

with $y = j + k/2 + 1$ to get,

$$\begin{aligned} I_2(z, t; \lambda, \alpha)_{t=0} &= \frac{(1-\alpha)\lambda\sqrt{2\pi}}{\pi}\sum_{j=0}^{\infty}\frac{\Gamma\,(j+1/2)\,\Gamma\,(3/2)}{\Gamma\,(1/2)\,\Gamma\,(j+3/2)}\frac{(-\lambda^2)^j}{j!}\frac{1+(-1)^{2j+k+1}}{2}\frac{2^{2(j+k/2+1)^{-1}}\Gamma(j+k/2+1)}{2^j 2^{j+k/2+1/2}\sqrt{\pi}} \\ &= \frac{(1-\alpha)\lambda}{\pi}2^{\frac{k}{2}+1}\Gamma\left(\frac{k}{2}+1\right)\frac{1+(-1)^{k+1}}{2}{}_2F_1(\{1/2, k/2+1\}, 3/2, -\lambda^2) = \mu_2(k) \end{aligned}$$

Substituting (2.5) and (2.6) in (2.4) gives (2.3).

Result 2.5. If Z follows MSND(λ, α), then the rth moment of any integer order k is the following, provided $\lambda > 0$.

$$E(Z^k) = E(Z_1) + (1-\alpha)\lambda\sqrt{\frac{2}{\pi}}\sum_{j=0}^{\infty}\frac{\mathit{\Gamma}(j+1/2)\mathit{\Gamma}(3/2)(-\lambda^2)^j E({Z_1}^{2j+k+1})}{\mathit{\Gamma}(1/2)\mathit{\Gamma}(j+3/2)j!2^j}, \tag{2.5}$$

where Z_1 follows the standard normal p.d.f.

Proof. For any real order k,

$$\begin{aligned} E(Z^k) &= \int_{-\infty}^{\infty} z^k\phi(z)\left[\alpha + 2(1-\alpha)\Phi(\lambda z)\right]dz \\ &= \alpha\int_{-\infty}^{\infty} z^k\phi(z)dz + (1-\alpha)\int_{-\infty}^{\infty} 2z^k\phi(z)\Phi(\lambda z)dz \\ &= \alpha E(Z_1^k) + (1-\alpha)\int_{-\infty}^{\infty} 2z^k\phi(z)\frac{1}{2}\left[1 + \text{erf}\left(\frac{\lambda z}{\sqrt{2}}\right)\right]dz \\ &= \alpha E(Z_1^k) + (1-\alpha)\int_{-\infty}^{\infty} z^k\phi(z)dz + (1-\alpha)\int_{-\infty}^{\infty} z^k\phi(z)\text{erf}\left(\frac{\lambda z}{\sqrt{2}}\right)dz \\ &= E(Z_1^k) + (1-\alpha)\int_{-\infty}^{\infty} z^k\phi(z)\frac{2\lambda z}{\sqrt{\pi}\sqrt{2}}\,{}_1F_1\left(\frac{1}{2},\frac{3}{2},-\frac{\lambda^2z^2}{2}\right)dz \\ &= E(Z_1^k) + (1-\alpha)\frac{\sqrt{2}}{\sqrt{\pi}}\sum_{j=0}^{\infty}\frac{\Gamma(j+1/2)\Gamma(3/2)(-\lambda^2)^j}{\Gamma(1/2)\Gamma(j+3/2)j!2^j}\int_{-\infty}^{\infty} z^{2j+k+1}\phi(z)dz \\ &= E(Z_1^k) + (1-\alpha)\frac{\sqrt{2}}{\sqrt{\pi}}\sum_{j=0}^{\infty}\frac{\Gamma(j+1/2)\Gamma(3/2)(-\lambda^2)^j}{\Gamma(1/2)\Gamma(j+3/2)j!2^j}E(Z_1^{2j+k+1}), \end{aligned} \tag{2.6}$$

which implies (2.7).

Result 2.6. The mean and variance of the random variable following MSND(λ, α) is given by $\mu = \sqrt{\frac{2}{\pi}}(1-\alpha)\delta$ and $\sigma^2 = 1 - \frac{2(1-\alpha)^2\delta^2}{\pi}$, where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

Result 2.7. The skewness and kurtosis of the random variable following MSND (λ, α) is given by

$$\text{skewness} = \frac{a(2a^2-\delta^2)}{(1-a^2)^{3/2}} \text{ and } \text{kurtosis} = \frac{3 - 3a^4 - a^2\left(2\delta^2 + \frac{6}{1+\lambda^2}\right)}{(1-a^2)^2},$$

where $a = \sqrt{\frac{2}{\pi}}(1-\alpha)\delta$.

Remark 2.1. When $\alpha = 1/\lambda = 0$, skewness is 0 and kurtosis is 3 which corresponds to the standard normal case.

Remark 2.2. When $\alpha = 0$, we get the skewness and kurtosis of skew normal distribution of Azzalini's (1985).

3 Shannon Entropy and Mills Ratio

Here in this section two tools associated with the MSND random variable are discussed.

3.1 Shanon Entropy

The entropy is the concept developed in the context of information theory. They are widely used in the case of normal, skew normal distributions. Here the Shannon entropy is extended for modified skew normal distribution to quantify how the skewness affects the uncertainty of the distribution compared to a standard normal distribution.

Result 3.1. The Shannon entropy of a random variable following MSND (λ, α) is given by,

$$SE_{MSND}(z) = -\alpha \int_{-\infty}^{\infty} ln(\ \phi(z))\phi(z)dz + (1-\alpha)SE_{SND}(z) - I(z, \lambda, \alpha) \tag{3.1}$$

where SE$_{SND}$ (z) corresponds to the Shannon entropy of the skew normal distribution of Azzalini's (1985) and $I(z, \lambda, \alpha) = \int_{-\infty}^{\infty} \phi(z)[\alpha + 2(1-\alpha)\Phi(\lambda z)]ln[\ \alpha + 2(1-\alpha)\Phi(\lambda z)]dz$.

Proof. Let Z follows MSND (λ, α). Then by the definition of Shannon entropy,

$$\begin{aligned}
SE(z) &= -E[\ln h(z)] \\
&= -\int_{-\infty}^{\infty} \ln\left\{\phi(z)[\alpha + 2(1-\alpha)\Phi(\lambda z)]\right\} \phi(z)[\alpha + 2(1-\alpha)\Phi(\lambda z)]dz \\
&= -\int_{-\infty}^{\infty} \left(\ln \phi(z) + \ln[\alpha + 2(1-\alpha)\Phi(\lambda z)]\right) \phi(z)[\alpha + 2(1-\alpha)\Phi(\lambda z)]dz \\
&= -\int_{-\infty}^{\infty} \ln(\phi(z))\phi(z)[\alpha + 2(1-\alpha)\Phi(\lambda z)]dz \\
&\quad - \int_{-\infty}^{\infty} \ln[\alpha + 2(1-\alpha)\Phi(\lambda z)]\phi(z)[\alpha + 2(1-\alpha)\Phi(\lambda z)]dz \\
&= -\alpha \int_{-\infty}^{\infty} \ln(\phi(z))\phi(z)dz - \int_{-\infty}^{\infty} \ln(\phi(z))2(1-\alpha)\phi(z)\Phi(\lambda z)dz - I(z, \lambda, \alpha) \\
&= -\alpha \int_{-\infty}^{\infty} \ln(\phi(z))\phi(z)dz + (1-\alpha)SE_{SND}(z) - I(z, \lambda, \alpha).
\end{aligned}$$

which implies (3.1).

3.2 Mills Ratio

Mills Ratio is defined as the ratio of the probability density function to the cumulative distribution function of a distribution. An important application of Mills Ratio are with regression analysis to account for selection bias, moreover it describes the tail behavior of a distribution. Here Mills Ratio of the MSND (λ, α) is discussed.

Result 3.2. The Mills Ratio of a random variable following MSND (λ, α) is given by,

$$m(z) = \frac{1 - erf\left(\frac{\lambda t}{\sqrt{2}}\right) - 2(1-\alpha)\int_{-\infty}^{z} e^{-\frac{t^2}{2}} erf\left(\frac{\lambda t}{\sqrt{2}}\right) dt}{\phi(z)\left[\alpha + (1-\alpha)\left(1 + + erf\left(\frac{\lambda t}{\sqrt{2}}\right)\right)\right]}$$

Proof. By the definition of the Mills Ratio, if h(z) is the p.d.f. and H(z) is the c.d.f. of a random variable Z, then the Mills Ratio denoted by m(z) is given by,

$$m(z) = \frac{1 - H(z)}{h(z)}$$

$$= \frac{1 - \Phi(z) + (1-\alpha)\int_{-\infty}^{z} e^{-\frac{t^2}{2}} \operatorname{erf}\left(\frac{\lambda t}{\sqrt{2}}\right) dt}{\phi(z)[\alpha + 2(1-\alpha)\Phi(\lambda z)]}$$

$$= \frac{1 - \Phi(z) - (1-\alpha)\int_{-\infty}^{z} e^{-\frac{t^2}{2}} \operatorname{erf}\left(\frac{\lambda t}{\sqrt{2}}\right) dt}{\phi(z)\left[\alpha + 2(1-\alpha)\frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{\lambda z}{\sqrt{2}}\right)\right)\right]}$$

$$= \frac{1 - \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{\lambda z}{\sqrt{2}}\right)\right) - (1-\alpha)\int_{-\infty}^{z} e^{-\frac{t^2}{2}} \operatorname{erf}\left(\frac{\lambda t}{\sqrt{2}}\right) dt}{\phi(z)\left[\alpha + (1-\alpha)\left(1 + \operatorname{erf}\left(\frac{\lambda z}{\sqrt{2}}\right)\right)\right]}$$

$$= \frac{1 - \operatorname{erf}\left(\frac{\lambda z}{\sqrt{2}}\right) - 2(1-\alpha)\int_{-\infty}^{z} e^{-\frac{t^2}{2}} \operatorname{erf}\left(\frac{\lambda t}{\sqrt{2}}\right) dt}{\phi(z)\left[\alpha + (1-\alpha)\left(1 + \operatorname{erf}\left(\frac{\lambda z}{\sqrt{2}}\right)\right)\right]}.$$

4 Maximum Likelihood Estimation

In this section, the maximum likelihood estimation of the parameters of location scale extension of MSND (λ, α), denoted as LSMSND $(\mu, \sigma; \lambda, \alpha)$ has been discussed and further a numerical illustration has been carried out. The likelihood of the sample X_1, X_2,…,X_n, size n from a population LSMSND $(\mu, \sigma; \lambda, \alpha)$ is

$$L = \frac{e^{-\sum_{i=1}^{n} \frac{(y_i-\mu)^2}{2\sigma^2}}}{\left(\sigma\sqrt{2\pi}\right)^n} \prod_{i=1}^{n}\left(\alpha + 2(1-\alpha)\Phi\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\right]\right).$$

On taking logarithm on both sides we get,

$$\log L = -\frac{n}{2}(\log\sigma^2 + \log 2\,\pi) - \sum_{i=1}^{n}\left[\frac{(y_i-\mu)^2}{(2\sigma^2)} - \log\left(\alpha + 2(1-\alpha)\Phi\big((y_i-\mu)\lambda/\sigma\big)\right)\right] \quad (4.1)$$

Differentiating (4.1) with respect to the parameters μ, σ^2, λ, α and equating to zero the following normal equations are obtained.

$$\frac{\partial lnL}{\partial\mu} = \sum_{i=1}^{n}\frac{(y_i-\mu)}{\sigma^2} - \frac{2(1-\alpha)\lambda}{\sigma}\sum_{i=1}^{n}\frac{\varphi\left[\frac{\lambda(y_i-\mu)}{\sigma}\right]}{\left[\alpha+2(1-\alpha)\Phi\left\{\lambda\left(\frac{y_i-\mu}{\sigma}\right)\right\}\right]} = 0$$

$$\frac{\partial \ln L}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\sigma^4} - \frac{\lambda(1-\alpha)}{\sigma^3}\sum_{i=1}^{n}\frac{(y_i-\mu)\varphi\left[\frac{\lambda(y_i-\mu)}{\sigma}\right]}{\left[\alpha+2(1-\alpha)\Phi\left\{\frac{\lambda(y_i-\mu)}{\sigma}\right\}\right]} = 0$$

$$\frac{\partial lnL}{\partial\lambda} = \sum_{i=1}^{n}\frac{2(1-\alpha)\varphi\left\{\lambda\left(\frac{y_i-\mu}{\sigma}\right)\right\}\left(\frac{y_i-\mu}{\sigma}\right)}{\left[\alpha+2(1-\alpha)\Phi\left\{\lambda\left(\frac{y_i-\mu}{\sigma}\right)\right\}\right]} = 0$$

$$\frac{\partial lnL}{\partial\alpha} = \sum_{i=1}^{n}\frac{\left[1-2\Phi\left(\frac{\lambda(y_i-\mu)}{\sigma}\right)\right]}{\left[\alpha+2(1-\alpha)\Phi\left\{\frac{\lambda(y_i-\mu)}{\sigma}\right\}\right]} = 0$$

which in turn reduces to

$$\sum_{i=1}^{n}\frac{(y_i-\mu)}{\sigma^2} = \frac{2(1-\alpha)\lambda}{\sigma}\sum_{i=1}^{n}w(y_i),$$

$$\sum_{i=1}^{n}(y_i-\mu)^2 = n\sigma^2 + 2(1-\alpha)\lambda\sigma\sum_{i=1}^{n}w(y_i)(y_i-\mu),$$

$$\frac{2(1-\alpha)}{\sigma}\sum_{i=1}^{n}W(y_i)(y_i-\mu) = 0$$

$$\sum_{i=1}^{n}\left[\alpha+2(1-\alpha)\Phi\left\{\frac{\lambda(y_i-\mu)}{\sigma}\right\}\right]^{-1} - 2\sum_{i=1}^{n}W(y_i) = 0.$$

for

$$w(y_i) = \frac{\varphi\left\{\frac{\lambda(y_i-\mu)}{\sigma}\right\}}{\left[\alpha+2(1-\alpha)\Phi\left\{\frac{\lambda(y_i-\mu)}{\sigma}\right\}\right]}, \;\; and \;\; W(y_i) = \frac{\Phi\left\{\frac{\lambda(y_i-\mu)}{\sigma}\right\}}{\left[\alpha+2(1-\alpha)\Phi\left\{\frac{\lambda(y_i-\mu)}{\sigma}\right\}\right]}.$$

On solving these non-linear equations (4.2) to (4.4) using some mathematical software's like MATHCAD, MATHLAB, MATHEMATICA etc, one can obtain the maximum likelihood estimators of the parameters of LSMSND $(\mu, \sigma; \lambda, \alpha)$.

5 Numerical Illustration

For illustrating the usefulness of the model LSMSND $(\mu, \sigma; \lambda, \alpha)$, we consider the following two real life data sets, among these the first data is on the heights (in centimeters) of 100 Australian athletes, given in Cook and Weisberg (1994) and second is on the IQ data set for 87 white males hired by a large insurance company in 1971 given in Roberts (1988).

Data Set 1:148.9 149 156 156.9 157.9 158.9 162 162 162.5 163 163.9 165 166.1 166.7 167.3 167.9 168 168.6 169.1 169.8 169.9 170 170 170.3 170.8 171.1 171.4 171.4 171.6 171.7 172 172.2 172.3 172.5 172.6 172.7 173 173.3 173.3 173.5 173.6 173.7 173.8 174 174 174 174.1 174.1 174.4 175 175 175 175.3 175.6 176 176 176 176 176.8 177 177.3 177.3 177.5 177.5 177.8 177.9 178 178.2 178.7 178.9 179.3 179.5 179.6 179.6 179.7 179.7 179.8 179.9 180.2 180.2 180.5 180.5 180.9 181 181.3 182.1 182.7 183 183.3 183.3 184.6 184.7 185 185.2186.2 186.3 188.7 189.7 193.4 195.9.

Data Set 2: 85 94 94 97 98 100 100 101 102 102 103 103 103 103 104 104 106 106 106 106 106 107 107 108 108 108 108 108 108 108 109 109 111 111 112 112 112 112 112 112 112 112 112 112 113 113 113 113 113 113 113 113 114 114 115 116 116 116 116 117 117 117 118 118 118 119 120 120 120 121 121 121 122 122 122 122 122 122 124 124 125 129 131 132 135 136 140

We obtained the MLEs of the parameters of the models N(μ, σ), ESN(μ, σ ; λ), LSMSND ($\mu, \sigma; \lambda, \alpha$) and obtained the best fit among them, based on certain information criterion such as the Akaike ' s Information Criterion (AIC), the Bayesian Information Criterion (BIC) and the corrected Akaike's Information Criterion (AICc). We have computed in all cases the measures-the loglikelihood(l), the AIC, the BIC and the AICc and included them in respective Tables 1 and 2. Also, it can be seen that LSMSND ($\mu, \sigma; \lambda, \alpha$) gives a better fit to the data sets compared to the existing models.

Table 1: Estimated Values of the Parameters l, AIC, BIC and AICc Values for the Models N(μ, σ), ESN(μ, σ ; λ), LSMSND ($\mu, \sigma; \lambda, \alpha$) for Data Set1.

Distribution:	**Normal**	**Skew Normal**	**LSMND**
	(μ, σ)	$(\mu, \sigma; \lambda)$	$(\mu, \sigma; \lambda, \alpha)$
μ	174.594	174.58	173
σ	8.24	8.20	8.8
λ	-	0.0016	0.7
α	-	-	0.05
l	-352.318	-352.318	-315.6867
AIC	708.64	710.64	639.37
BIC	713.85	718.45	649.19
AICc	708.76	710.89	639.867

It is seen that MSND (λ, α) provides a better fit for skewness but requires adjustments for tail behavior. Azzalini's model is more skewed, making it less suitable for datasets with mild asymmetry.

5.1 Key Comparison

The key comparisons between the MSND (λ, α) and Azzalini's Skew Normal Distribution is based on Graphical Comparisons (PDF, CDF, Q-Q plots). It is observed that MSND (λ, α) shows a

Table 2: Estimated Values of the Parameters l, AIC, BIC and AICc Values for the Models N(μ, σ), ESN(μ, σ ; λ), LSMSND ($\mu, \sigma; \lambda, \alpha$) for Data Set 2.

Distribution:	**Normal**	**Skew Normal**	**LSMND**
	(μ, σ)	$(\mu, \sigma; \lambda)$	$(\mu, \sigma; \lambda, \alpha)$
μ	112.86	105.78	111
σ	9.58	11.94	9.8
λ	-	1.14	0.9
α	-	-	0.05
l	-319.6	-319.29	-316
AIC	643.2	644.57	640.72
BIC	648.14	651.97	650.53
AICc	643.35	644.86	641.21

Table 3: Empirical Analysis of Skewness and Kurtosis for Data Sets 1 & 2

Model	**Skewness (Data Set 1)**	**Kurtosis (Data Set 1)**	**Skewness (Data Set 2)**	**Kurtosis (Data Set 2)**
Empirical Data	-0.560	1.197	0.172	0.684
Modified Skew Normal (MSND)	0.078	-0.512	0.043	-0.001
Azzalini's Skew Normal	0.157	0.089	0.344	-0.497

stronger skewness and heavier tail than Azzalini's model and MSND may be preferred when heavier tails are necessary. The empirical CDFs of both distributions show differences in tail behavior, Some deviations indicate that Azzalini's model does not perfectly capture the MSND (λ, α)'s tail behavior.

6 Summary and Conclusion

In this paper, we introduced the modified skew normal distribution and studied its fundamental properties such as distribution function, characteristic function, moment-generating function, raw moments, skewness, and kurtosis. Also, examined the information measures, such as Shannon entropy and Mille's ratio, which provides deeper insights into the distribution's behavior. We further discussed the maximum likelihood estimation of its parameters and its applicability in statistical modeling. Numerical examples were considered to identify the utility of the proposed distribution. Moreover , the modified skew normal distribution offers a flexible and robust framework for modeling asymmetric data, while preserving the structure of the normal distribution. Future research can explore its charecterisations, inferential aspects and its importance in applications in in finance, risk

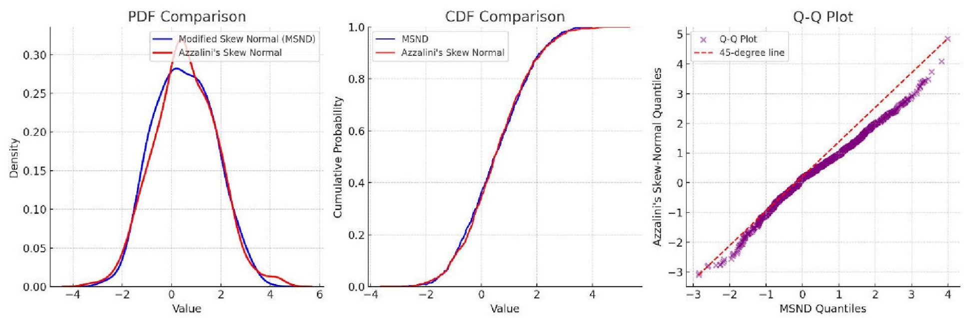

Figure 4: Graphical comparison of MSND (λ, α) and Azzalini's Skew Normal Distribution

modeling, and survival analysis where asymmetric data distribution is common.

Acknowledgement

The authors are grateful to the Editor in chief, Associate Editor and the anonymous Referees for carefully reading the paper and for the valuable comments and suggestions on an earlier version of the manuscript that greatly improved the quality and presentation of the paper.

References

Ahsanullah, M., Shakil, M., and Kibria, B. G. (2015). Characterizations of folded Student's t distribution. *Journal of Statistical Distributions and Applications*, *2*(1), 15.

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

Azzalini, A. (1986). Further results on the class of distributions which includes the normal ones. *Statistica*, *46*, 199–208.

Azzalini, A., and Capitanio, A. (1999). Statistical applications of the multivariate skew-normal distribution. *Journal of the Royal Statistical Society: Series B*, *61*, 579–602.

Azzalini, A., and Dalla-Valle, A. (1996). The multivariate skew normal distribution. *Biometrika*, *83*, 175–726.

Cook, R. D., and Weisberg, S. (1994). *An Introduction to Regression Analysis*. Wiley, New York.

Gupta, A. K., Omar, M. H., and Joarder, A. H. (2013). On a generalized mixture of standard normal and skew normal probability distributions. *Pan American Mathematical Journal*, *23*(1), 1–23.

Kumar, C. S., and Anusree, M. R. (2011). On a generalized mixture of standard normal and skew normal distributions. *Statistics and Probability Letters*, *81*, 1813–1821.

Kumar, C. S., and Anusree, M. R. (2014). On a modified class of generalized skew normal distribution. *South African Statistical Journal*, *31*, 111–124.

Ma, Y., and Genton, M. G. (2004). A flexible class of skew-symmetric distributions. *Scandinavian Journal of Statistics*, *31*, 459–468.

Mondal, S., Arellano-Valle, R. B., and Genton, M. G. (2024). A multivariate modified skew normal distribution. *Statistical Papers*, *65*, 511–555. https://doi.org/10.1007/s00362-023-01397-1

Pewsey, A. (2000). The characteristic functions of the skew normal and wrapped skew normal distributions. In *27th Congreso Nacional de Estadística e Investigación Operativa*.

Roberts, H. V. (1988). *Data Analysis for Managers with Minitab*. Scientific Press, Redwood City, CA.

Shakil, M., Ahsanullah, M., and Kibria, B. G. (2014). A characterization of skew normal distribution by truncated moment. *Applications and Applied Mathematics: An International Journal (AAM)*, *9*(1), Article 3. https://digitalcommons.pvamu.edu/aam/vol9/iss1/3

Received: November 22, 2024

Accepted: March 14, 2025