A Test for skewness within the Univariate and Multivariate Epsilon Skew Laplace Distributions

Howraa Al-Mousawi *   Hassan Elsalloukh †   Jose H. Guardiola‡

Abstract
In the univariate case, the popular measures of skewness, β_1, and kurtosis, β_2, have been proved to be useful measures in developing a test for normality and investigating the robustness of the standard normal theory procedures. In the multivariate case, we have a p-dimensional skewness vector, β_{12}, introduced by Mardia in 1970, as multivariate skewness measure. In this work, the skewness measure has been derived for the Multivariate Epsilon Skew Laplace Distribution (MESL), the MESL is the multivariate version of the Epsilon Skew Laplace distribution (ESL) that have been introduce recently by Elalloukh (2008) and revised by Al-Mousawi et al (2012A, 2012B) . The MESL is an asymmetric distribution that can handle both symmetric, asymmetric, and heavy tail data. The p-dimensional skewness vector is introduced by using the Mardia's measures of skewness. Moreover, we provide a test for goodness of fit test to pick distributions that can fit the data correctly. We provide theoretical proofs and a Monte carlo simulation study to compare the ESL distributions to normal and Epsilon Skew Normal distributions (ESN), in the univariate cases when modeling data

Key Words: Epsilon skew Laplace distribution, Measure of skewness, Multivariate distributions, heavy tail distributions

1. Introduction

It is prudent before doing any statistical modeling to verify that the data satisfy the underlying distribution assumption. The multivariate normal distribution plays an important role in the analysis of multivariate data and multivariate regression models, especially when the multivariate analysis of variance (MANOVA) is used. Most of the techniques of multivariate statistical analysis are based on the assumption that the data are generated from multivariate normal distributions because the normal distribution is a useful approximation to the true population distribution. Hence, it is important to determine whether the data that are being used for statistical inference are from multivariate normal or other distributions such as skewed or asymmetric distributions. Based on published work, testing multivariate normality or symmetry can be done through using the well known Mardia's measures (1970) of skewness and kurtosis or other tests for normality such as Henze-Zirkler et al (1990).

Mardia (1970), introduced the multivariate skewness and kurtosis and are denoted by $\beta_{1,p}$ and $\beta_{2,p}$, respectively. They became the standard characteristics of a multivariate distribution. These measures are defined as follows in Mardia (1970, 1974), let X be a random p-vector with $E(X) = \mu$ and variance-covariance matrix Σ, then the p-variate skewness and kurtosis are, respectively

$$\beta_{1,p}(X) = \sum_{i,j,k} \{E(X_i X_j X_k)\}^2 \quad \text{and} \quad \beta_{2,p}(X) = \{E(X_i X_j X_k X_\ell)\} - p(p+2),$$

*Arkansas Department of Health, Vital Statistics Section, 4815 West Markham Street, Little Rock, AR 72205

†University of Arkansas at Little Rock, Department of Mathematics and Statistics, 2801 South University Avenue, Little Rock, AR 72204-1099

‡Department of Computing and Mathematical Sciences, Texas AM University at Corpus Christi, TX

or, equivalently

$$\beta_{1,p}(X) = E\left((X-\mu)'\Sigma^{-1}(X-\mu)\right)^3 \tag{1}$$

and

$$\beta_{2,p}(X) = E\left((X-\mu)'\Sigma^{-1}(X-\mu)\right)^2. \tag{2}$$

Gutjahr (1999) studied Mardia's formula (1) and pointed out that the skewness measure $\beta_{1,p}$ equals zero in case of multivariate normality and other symmetric distributions. Recently, several new skew distribution families have been introduced for modeling skew data, such as the multivariate skew normal distribution introduced by Azzalini (1996), the multivariate skew t-distribution (2003), and the asymmetric multivariate Laplace distribution introduced by Kotz (2001). These families are presented by three parameters, in addition to the location parameter vector and the variance-covariance scale parameter matrix, there is also a skewness vector. In recent years, several suggestions and generalizations for Mardia's measure have been made to modify multivariate skewness and kurtosis measures. For example, Móri, Rohatgi and Śzekely (1993) defined the multivariate skewness as a p-vector and multivariate kurtosis as $p \times p$-matrix, respectively, as

$$\beta_1(X) = E\left(\|Y\|^2 Y\right) \quad \text{and} \quad \beta_2(X) = E\left(YY'YY'\right) - (p+2)I_p,$$

where $E(X) = \mu$ and Σ are the mean and the variance-covariance matrix, respectively. Moreover, $Y = \Sigma^{-1/2}(X-\mu)$ and $\Sigma^{-1/2}$ is a square root of Σ. Kollo and Srivastave (2004) represented $\beta_{1,p}$ via the third order of multivariate moments as

$$\beta_{1,p}(X) = tr\left(m_3'(Y)m_3(Y)\right), \tag{3}$$

where m_3 represents the third moments of Y. Kollo (2008) gave $b(X) = 1_{p\times p} \star m_3(Y)$, where $(\star)$ represents the star product between two matrices, as another measure of the skewness of X. Another way to find the skewness or kurtosis is to use the cumulant function. The cumulant function of a random variable (or random p-vector) is defined as

$$\psi_X(t) = \ln \phi_X(t), \tag{4}$$

where $\phi_X(t)$ is the characteristic function of X. We can find the skewness or kurtosis of a random p-vector X by using the third and fourth cumulants, where the kth cumulant is obtained from the partial differentiation of the cumulant function matrix (4), that is

$$C_k = \frac{1}{i^k} \frac{\partial^k \psi_X(t)}{\partial t \partial t' \ldots \partial t}\bigg|_{t=0}. \tag{5}$$

Note that, the relationship between the cumulants and the moment is one-to-one and is given by the next proposition.

Proposition 1. *Kollo (1991), Let X be a random p-vector with $m_4(X) < \infty$, then*
1- $C_1(X) = E(X)$.
2- $C_2(X) = \bar{m}_2(X)$.
3- $C_3(X) = \bar{m}_3(X)$.
where $\bar{m}_k(X)$ presents the centered moment.

In this paper, the skewness measure of a multivariate p-dimensional MESL is introduced by using the Mardia's measures of skewness. Moreover, we provide a test for goodness of fit to pick distributions that can fit the data correctly. In particular, we generalizes Kundu's (2004) procedure to skewed data. We provide theoretical proofs and a Monte carlo simulation study to compare normal to ESL distributions and ESN to ESL distributions, in the univariate cases when modeling data.

The rest of the paper is organized as follows. In Section 2, a brief introduction of the MESL distribution is provided along with some of its properties. In Section 3, we derive the skewness measure for the MESL distribution. In Section 4, we derive a goodness of fit test for univariate cases, in particular, fitting the ESL distribution to normal and ESN distributions. Finally, in Section 5, simulation study demonstrates the test for fitness is provided.

2. Introduction and Basic Information of The MESL Distribution

Al-Mousawi et al (2011, 2012B) provided a new asymmetric multivariate distribution called the Multivariate Epsilon-Skew Laplace distribution (MESL), with density function defined by:

Definition 2.1. *: Al-Mousawi et al (2012B), A random vector X in R^p, for $p \geq 1$, is said to have p-dimensional Multivariate Epsilon-Skew Laplace distribution denoted by $X \sim$ MESL$(\Theta, \Sigma, \Upsilon)$, if its density function is given by*

$$f(x, \Theta, \Sigma, \Upsilon) = \frac{|\Sigma|^{-1/2}}{2^{3P/2}} \exp\left\{ -\frac{1}{\sqrt{2}} \left\{ \left[(x-\Theta)' (I-\Upsilon)^{-2} W \Sigma^{-1} (x-\Theta) \right]^{1/2} + \left[(x-\Theta)' (I+\Upsilon)^{-2} W_1 \Sigma^{-1} (x-\Theta) \right]^{1/2} \right\} \right\}, \tag{6}$$

where $x \in R^p$, $\Theta \in R^p$ is the location vector, $\Upsilon \in R^p$ is the diagonal skewness parameters matrix, Σ is the positive definite scatter matrix parameter, I is a $p \times p$ identity matrix, and W and W_1 are the indicator diagonal matrices defined as

$$W = \begin{bmatrix} I_{(x_1 \geq \theta_1)} & 0 & 0 & 0 & 0 \\ 0 & I_{(x_2 \geq \theta_2)} & 0 & 0 & 0 \\ 0 & 0 & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & I_{(x_p \geq \theta_p)} \end{bmatrix} \tag{7}$$

and

$$W_1 = \begin{bmatrix} I_{(x_1 < \theta_1)} & 0 & 0 & 0 & 0 \\ 0 & I_{(x_2 < \theta_2)} & 0 & 0 & 0 \\ 0 & 0 & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & I_{(x_p < \theta_p)} \end{bmatrix}, \tag{8}$$

where $I_{(z)}$ is the indicator function defined as

$$I_{(z)} = \begin{cases} 1 & \text{for } z \text{ is true} \\ 0 & \text{for } z \text{ is false.} \end{cases} \tag{9}$$

Note that the pdf (6) has fatter tail regions than that of the multivariate normal distribution, and hence, it can be used as an alternative when the data are highly skewed. When $\Upsilon = 0$, (6) reduces to the multivariate Laplace distribution defined by Kotz (2001). When $p = 1$, (6) reduces to the Epsilon Skew Laplace (ESL) distribution defined by Elsalloukh (2005, 2008), with density function given by

$$f_{ESL}(x) = \frac{1}{2\sqrt{2}\sigma} e^{\left\{ \left(\frac{-|x-\theta|}{\sqrt{2}(1-\varepsilon)\sigma} \right) I_{(x \geq \theta)} + \left(\frac{-|x-\theta|}{\sqrt{2}(1+\varepsilon)\sigma} \right) I_{(x < \theta)} \right\}}, \tag{10}$$

where θ, σ, and ε, are the location, scale, and skewness parameters, respectively, and $I_{(z)}$ is the indicator function as defined in (9). The MESL distribution is the multivariate version of the ESL distribution that can handle both symmetric, asymmetric, and heavy tail data. The multivariate moment generating function (MMGF) for the MESL, is given by AL-Mousawi at el (2012B) by

$$M_X(t) = \frac{e^{t'\Theta}}{1 - 2t'(I - \Upsilon^2)\Sigma t + 2\sqrt{2}t'\Sigma^{1/2}\Upsilon \cdot 1}, \tag{11}$$

where $t' = (t_1, t_2, ..., t_p)$ and $1' = (1, 1, ..., 1)$ is a vector of one's. While the mean and variance are respectively given by

$$E(X) = \Theta - 2\sqrt{2}\Sigma^{1/2}\Upsilon \cdot 1 \tag{12}$$

and

$$\text{Var}(X) = 4\,(I + (2\,p - 1)\Upsilon^2)\Sigma. \tag{13}$$

Proposition 2. *Al-Mousawi et al (2012B), If $X \sim MESL_p(\Theta, \Sigma, \Upsilon)$ and $Y = AX + b$, where $A \in R^{q\times p}$ and $b \in R^q$, then $Y \sim MESL_q(A\Theta + b, A\Sigma A', \Upsilon)$.*

3. Skewness Measurement of the MESL Distribution

In this section, we derive the multivariate skewness parameters for the MESL using Mardia measures of skewness (1970). First, we transform the random variable X to standard random variable Z, we then use the random variable $Z_1, Z_2, ..., Z_p$ to measure the skewness, where $Z = \text{var}(X)^{-1/2}(X - E(X))$. The transformation leads to $E(Z) = 0$ and $E(Z'Z) = \text{var}(Z) = I_p$.

Proposition 3. *Let X be a random p-vector from MESL$(0, \Sigma, \Upsilon)$, with $E(X)$ and variance-covariance matrix var(X) as defined in (12) and (13) respectively. Let $Z = (var(X))^{-1/2}$ $(X - E(X))$, then the random p-vector Z has MESL distribution given by: $Z \sim MESL(\sqrt{2}(I + (2p-1)\Upsilon^2)^{-1/2}\Upsilon\, 1, \left(4(I + (2p-1)\Upsilon^2)\right)^{-1}, \Upsilon)$.*

Proof. Let $Z = (4(I+(2p-1)\Upsilon^2)\Sigma)^{-1/2}(X+2\sqrt{2}\Sigma^{1/2}\Upsilon\, 1)$, equivalently, we can write Z as $Z = AX + B$, where

$$A = (4(I + (2p-1)\Upsilon^2)\Sigma)^{-1/2} = 2^{-1}(I + (2p-1)\Upsilon^2)^{-1/2}\Sigma^{-1/2}$$

and

$$B = (4(I + (2p-1)\Upsilon^2)\Sigma)^{-1/2}2\sqrt{2}\Sigma^{1/2}\Upsilon\, 1 = \sqrt{2}(I + (2p-1)\Upsilon^2)^{-1/2}\Upsilon\, 1\,,$$

where $A \in R^{q\times p}$ and $B \in R^q$. Note that A and B satisfy proposition (2), thus $Z \sim$ MESL$_q(A\Theta + B, A\Sigma A', \Upsilon)$. Since $\Theta = 0$, and

$$\begin{aligned} A\Sigma A' &= 2^{-2}(I + (2p-1)\Upsilon^2)^{-1/2}\Sigma^{-1/2}\Sigma\Sigma^{-1/2}(I + (2p-1)\Upsilon^2)^{-1/2} \\ &= \left(4(I + (2p-1)\Upsilon^2)\right)^{-1}, \end{aligned}$$

it follows that $Z \sim \text{MESL}(\sqrt{2}(I+(2p-1)\Upsilon^2)^{-1/2}\Upsilon\, 1, \left(4(I + (2p-1)\Upsilon^2)\right)^{-1}, \Upsilon)$. □

Note that, the random vector Z has $E(Z) = 0$, $Var(Z) = I$.

Proposition 4. *Let $Z \sim$ MESL$(\sqrt{2}AB, 4CAA', \Upsilon)$, then the third cumulants of Z is*

$$C_3(z) = -\sqrt{2}\left((AB \otimes CA^2) + 4(AB \otimes AB(AB)') + vecA^2CAB \right.$$
$$\left. -(CA^2 \otimes AB)\right) \,, \tag{14}$$

where

$$\begin{aligned} A_{p\times p} &= (I + (2p-1)\Upsilon^2)^{-1/2} \\ B_{p\times 1} &= \Upsilon\, 1 \\ C_{p\times p} &= (I - \Upsilon^2) \end{aligned} \tag{15}$$

Proof. Using the MMGF (11), the MMGF of Z is

$$M_z(t) = \frac{\exp(\sqrt{2}\, t'\, AB)}{1 - \frac{1}{2}\, t' CAA't + \sqrt{2}\, t'AB} \;=\; \exp(\sqrt{2}\, t'\, AB) w(t)\,, \tag{16}$$

where A, B, and C are as defined in (15), and $w(t)$ is

$$w(t) = \frac{1}{1 - \frac{1}{2}\, t'\, CAA'\, t + \sqrt{2}\, t'AB}.$$

Thus, the characteristic function of Z is

$$\Psi_z(t) = \exp(i\sqrt{2}\, t'\, AB) w(i\, t)$$

and the cumulant function (4), becomes

$$\psi_z(t) = i\sqrt{2}\, t'\, AB - \ln(1 + \frac{1}{2}\, t'CA^2\, t + \sqrt{2}\, i\, t'\, AB)\,. \tag{17}$$

To find the third cumulant, we need to find the third derivative of (17) with respect to t, so that

$$\frac{\partial \psi_z(t)}{\partial t} = i\,\sqrt{2}\, AB - \left(t'\, CA^2 + i\,\sqrt{2}\, AB\right) w(i\, t)\,,$$

$$\begin{aligned} \frac{\partial^2 \psi_z(t)}{\partial t \partial t'} &= -\, CA^2\, w(it) + (t'\, CA^2 + i\,\sqrt{2}\, AB)(A^2C't + i\,\sqrt{2}\, B'A')w^2(it) \\ &= i^2 \Big(CA^2 w(it) - (t'CA^2AC't - 2AB(AB)' \\ &\quad + i\,\sqrt{2}(t'\, CA^2(AB)' + ABA^2Ct))\, w^2(it) \Big)\,, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^3 \psi_z(t)}{\partial t \partial t' \partial t} &= i^2 \left\{ \frac{\partial (w(it) \otimes (CA^2 + 2ABw(it)(AB)'))}{\partial t} - \frac{\partial (tw^2(it))}{\partial t}\, A^2C^2A^2t' \right. \\ &\quad - (I_p \otimes tw^2(it))\, \frac{\partial A^2C^2A^2t'}{\partial t} - i\,\sqrt{2}\, \left[\frac{\partial A^2CABt}{\partial t}\, w^2(it) \right. \\ &\quad + (I_p \otimes ABA^2Ct) \frac{\partial w^2(it)}{\partial t} + \frac{\partial w^2(it)}{\partial t}\, t'CA^2(AB)' \\ &\quad \left.\left. + \; (I_p \otimes w^2(it)) \frac{\partial t'CA^2(AB)'}{\partial t} \right] \right\} \end{aligned}$$

$$
\begin{aligned}
= \;& i^2 \left\{ \frac{\partial w(it)}{\partial t} \otimes (CA^2 + 2AB2w(it)(AB)') - \frac{\partial t}{\partial t}\, w^2(it)\, A^2C^2A^2t' \right. \\
& - (I_p \otimes t)\, \frac{\partial w^2(it)}{\partial t}\, A^2C^2A^2t' - (I_p \otimes tw^2(it))\, A^2C^2A^2 \\
& - i\,\sqrt{2}\, \Big[vec\ A^2CABw^2(it) - (I_p \otimes A^2CABt)\, 2\, w^3(it)(t'CA^2 + i\,\sqrt{2}\, AB) \\
& - ((t'CA^2 + i\,\sqrt{2}\, AB) \otimes I_p)\, 2\, w^3(it)t'\, CA^2(AB)' \\
& \left. + (I_p \otimes I_p)\, w^2(it)\, CA^2(AB) \Big] \right\}
\end{aligned}
$$

$$
\begin{aligned}
= \;& i^2 \Big\{ - w^2(it)(t'CA^2 + i\,\sqrt{2}\, AB) \otimes (CA^2 + 4\, AB\, w(it)(AB)') \\
& - w^2(it)\, A^2C^2A^2t' + 2\, w^3(it)\, (I_p \otimes t)\, (t'CA^2 + i\,\sqrt{2}\, AB)\, A^2C^2A^2t' \\
& - w^2(it)\, (I_p \otimes t)\, A^2C^2A^2 - i\,\sqrt{2}\, \Big[w^2(it) vec A^2CAB \\
& - 2\, w^3(it)\, ((t\, CA^2 + i\,\sqrt{2}\, AB) \otimes AB\, t\, CA^2) - \; 2\, w^3(it) \\
& ((t'\, CA^2 + i\,\sqrt{2}\, AB) \otimes I_p) t' CA^2(AB)' + (CA^2(AB)' \otimes I_p)\, w^2(it) \Big] \Big\}.
\end{aligned}
$$

When $t = 0$, we have $w(it) = 1$, and

$$
\begin{aligned}
C_3(z) &= \frac{1}{i^3}\, \frac{\partial^3 \psi_z(t)}{\partial t \partial t' \partial t}\Big|_{t=0} \\
&= -\sqrt{2}\, (AB \otimes CA^2) - 4\,\sqrt{2}\, (AB \otimes AB(AB)') \\
&\quad -\sqrt{2}\ vec\ A^2CAB + \sqrt{2}\, (CA^2 \otimes AB) \\
&= -\sqrt{2}\, \Big[(AB \otimes CA^2) + 4\, (AB \otimes AB(AB)') + vec\, A^2CAB \\
&\quad - (CA^2 \otimes AB) \Big].
\end{aligned}
$$

□

Proposition 5. *Let* $X \sim$ *MESL*$(0, \Sigma, \Upsilon)$*, then the skewness parameter for the MESL is*

$$\beta_{1,p}(X) = 2\, a\, \Big[16\, a^2 + 8\, a + (3\, p - 2) \Big], \tag{18}$$

where $a = 1\, \Upsilon (I - \Upsilon^2)^{-1}\, \Upsilon\, 1$.

Proof. Using equation (3), we have

$$\beta_{1,p} = \mathrm{tr}\Big[C_3'(z)\, C_3(z) \Big]. \tag{19}$$

Let

$$h = (A^2C)^{-1/2}\, AB = C^{-1/2}\, B = (I - \Upsilon^2)^{-1/2}\, \Upsilon\, 1$$

and

$$a = h'\, h = B'\, C^{-1/2}\, C^{-1/2}\, B = 1\ \Upsilon\, (I - \Upsilon^2)^{-1}\, \Upsilon\ 1\,,$$

where A, B, C, and $C_3(z)$ are as defined in (15) and (14) respectively. Substituting the value of h and a in (14), we have

$$C_3(z) = -\sqrt{2} \Big[(h \otimes I_p) + 4\, (h\, h' \times h) - (I_p \otimes h) + vec\ I_p\, h' \Big]\,.$$

To find the skewness of the multivariate distribution, we first simplify (19)

$$
\begin{aligned}
C_3'(z)\, C_3(z) &= 2\Big[(h' \otimes I_p) + 4\,(h\,h' \otimes h') - (I_p \otimes h') + (h'\ vec\ I_p)\Big] \\
&\quad \times \Big[(h \otimes I_p) + 4\,(h\,h' \otimes h) - (I_p \otimes h) + (vec\ I_p\, h')\Big] \\
&= 2\Big[D' + 4\,E' - F' + G'\Big] \times \Big[D + 4\,E - F + G\Big] \\
&= 2\Big[G'G + 16\,E'E + D'D + F'F + 2\,G'D + 8\,G'E + 8\,E'D \\
&\quad - 2\,G'F - 8\,E'F - 2\,F'D\Big].
\end{aligned}
$$

Using $G = vec\ I_p\, h'$, $E = h\,h' \otimes h$, $F = I_p \otimes h$ and $D = h \otimes I_p$. Then

$$
\begin{aligned}
\beta_{1,p} &= 2\text{tr}\Big[G'G + 16\,E'E + D'D + F'F + 2\,G'D + 8\,G'E + 8\,E'D \\
&\quad - 2\,G'F - 8\,E'F - 2\,F'D\Big].
\end{aligned}
$$

Note that

$$F'F = (I \otimes h')\,(I \otimes h) = (I \otimes h'\,h) \implies \text{tr}\,(F'F) = \text{tr}\,(I \otimes h'\,h) = a\,p,$$

$$\text{tr}\,\big(G'G\big) = \text{tr}\,\big(F'F\big) = \text{tr}\,\big(D'D\big) = \text{tr}\,\big(I \otimes h'\,h\big) = a\,p\,,$$

$$\text{tr}\,(E'E) = \text{tr}\,(h\,h' \otimes h')\,(h\,h' \otimes h) = \text{tr}\,(h\,h'\,h\,h' \otimes h\,h') = \text{tr}\,(h\,h')^3 = a^3\,,$$

and

$$\text{tr}\,(G'E) = \text{tr}\,(E'F) = \text{tr}\,(E'D) = \text{tr}\,(h\,h' \otimes h\,h') = \big(\text{tr}\,(h\,h')\big)^2 = a^2$$

while

$$G'F = (vec\ I_p\, h')\,(I_p \otimes h) = h\,h' \qquad \text{and} \qquad F'D = (I_p \otimes h)\,(h \otimes I_p) = h\,h'\,,$$

with

$$\text{tr}\,(G'F) = \text{tr}\,(G'D) = \text{tr}\,(F'D) = \text{tr}\,(h\,h') = a\,.$$

Then (19) becomes

$$
\begin{aligned}
\beta_{1,p}(x) &= 2\Big[a\,p + 16\,a^3 + a\,p + a\,p + 2\,a + 8\,a^2 - 2\,a - 8\,a^2 + 8\,a^2 - 2\,a\Big] \\
&= 2\,a\Big[16\,a^2 + 8\,a + (3\,p - 2)\Big].
\end{aligned}
$$

□

Note that, when $\Upsilon = 0$, the skewness parameter in (18) will be zero, which indicates the multivariate distribution is symmetric.

4. Goodness of Fit Test for Univariate Distribution

One of the important assumptions when dealing with data is choosing the best distribution that fits the data. The normal and Laplace distributions are used to fit symmetric data. The laplace distribution is used to model long tails data, while the normal distribution is used to model data with short tails. Although these two distributions may provide similar data fit for moderate sample size, the problem is which distribution is best and more closely fits the data.

Cox (1961) studied the accuracy of goodness of fit among two probability distributions. Kundu (2004) considered a study between the normal and Laplace distributions, using the Ratio of the Maximized Likelihoods entitle "RML", the algorithm of RML depends on finding the value of G, where G is defined as

$$G = \ln \left\{ \frac{L_N(\hat{\mu}, \hat{\sigma})}{L_L(\hat{\theta}, \hat{\phi})} \right\},$$

where $L_N(\hat{\mu}, \hat{\sigma})$ and $L_L(\hat{\theta}, \hat{\phi})$ represent the likelihood functions for a random sample from a normal and Laplace distributions, respectively. The $(\hat{\mu}, \hat{\sigma})$ and $(\hat{\theta}, \hat{\phi})$ are the MLE of (μ, σ) and (θ, ϕ) respectively, which are the MLE of the normal and Laplace distributions parameters, respectively. The procedure that was used is to choose the normal distribution in favor of the Laplace distribution when the value of $G > 0$.

4.1 A Goodness of Fit Test for ESL VS Normal Distributions

Let $X_1, ..., X_n$ be a random sample from an ESL (10) or a normal distribution, then the likelihood function, respectively,

$$L_{ESL}(\theta, \varepsilon, \sigma) = \Pi_{i=1}^{n} f_{ESL}(x_i, \theta, \varepsilon, \sigma), \tag{20}$$

or

$$L_N(\mu, \phi) = \Pi_{i=1}^{n} f_N(x_i, \mu, \phi)\,.$$

Define the statistics

$$T = \ln \left\{ \frac{L_N(\hat{\mu}, \hat{\phi})}{L_{ESL}(\hat{\theta}, \hat{\varepsilon}, \hat{\sigma})} \right\}, \tag{21}$$

where, $(\hat{\mu}, \hat{\phi})$ and $(\hat{\theta}, \hat{\varepsilon}, \hat{\sigma})$ represent the MLEs of (μ, ϕ) and $(\theta, \varepsilon, \sigma)$ of the normal and ESL distribution, respectively. The MLEs of the normal distribution parameters are

$$\hat{\mu} = \bar{X} \quad \text{and} \qquad \hat{\phi}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \hat{\mu})^2\,. \tag{22}$$

While Al-Mousawi et al (2012A) showed that the maximum likelihood estimator (MLE) for the ESL parameters θ, ε and σ, are respectively given by

$$\hat{\theta}_k = X_{j(n)}, \tag{23}$$

$$\hat{\varepsilon}_k = \frac{\sqrt{\sum_{i=k+1}^{n} (x_{(i)} - \hat{\theta})^-} - \sqrt{\sum_{i=1}^{k} (x_{(i)} - \hat{\theta})^+}}{\sqrt{\sum_{i=1}^{k} (x_{(i)} - \hat{\theta})^+} + \sqrt{\sum_{i=k+1}^{n} (x_{(i)} - \hat{\theta})^-}}, \tag{24}$$

and

$$\hat{\sigma}_k^2 = \frac{\sum_{i=1}^{k} (x_{(i)} - \hat{\theta})^+}{\sqrt{2}n(1 - \hat{\varepsilon})} + \frac{\sum_{i=k+1}^{n} (x_{(i)} - \hat{\theta})^-}{\sqrt{2}n(1 + \hat{\varepsilon})}, \tag{25}$$

where where $x_{(i)}$ represents the order statistics, $j(n) = [n(1 + \varepsilon)/2] + 1$, $[x]$ denotes the integer part of x, k is a random integer such that $x_{(k)} < \theta < x_{(k+1)}$ and satisfies $[0 \le k \le n, k = k(x_{(1)}, x_{(2)}, ..., x_{(n)}, \theta)$, and ranges over $(0, 1, ..., n)]$.
Note that, when $k = 0$ or $k = n$, the MLE for θ, ε are given respectively by $x_{(1)}$ or $x_{(n)}$ and -1 or 1.

Using the MLE of ESL distribution (23), (24), and (25) with the MLE of the normal distribution (22) in equation (21) we get

$$T = \ln \left\{ \frac{(2\pi)^{-n/2}\hat{\phi}^{-n} e^{\frac{-1}{2\hat{\phi}^2}\sum_{i=1}^{n}(x_i-\hat{\mu})^2}}{(2\sqrt{2})^{-n}\hat{\sigma}^{-n} e^{\frac{-1}{\sqrt{2}\hat{\sigma}^2}\left[\frac{\sum_{i=1}^{k}(x_i-\hat{\theta})^+}{(1-\hat{\varepsilon})}+\frac{\sum_{i=k+1}^{n}(x_i-\hat{\theta})^-}{1+\hat{\varepsilon}}\right]}} \right\}. \tag{26}$$

Thus

$$T = \frac{1}{n} + n\,\ln(2) - \frac{n}{2}\,\ln\pi - n\,\ln\hat{\phi} + n\,\ln\hat{\sigma}. \tag{27}$$

If $T > 0$, we choose the normal distribution, otherwise, we choose the ESL distribution as the preferred model.

4.2 A Goodness of Fit Test for ESL VS ESN Distributions

The Epsilon Skew Normal (ESN) distribution introduced by Mudholkar and Hutson (1999), with a pdf given by

$$f_{ESN}(x) = \frac{1}{2\sqrt{2}\phi}\exp\left\{\left(\frac{-(x-\mu)^2}{2(1-\tau)^2\phi^2}\right)\mathrm{I}_{(x\geq\mu)} + \left(\frac{-(x-\mu)^2}{2(1+\tau)^2\phi^2}\right)\mathrm{I}_{(x<\mu)}\right\}, \tag{28}$$

where θ, σ and ε are location, scale, and skewness parameters, respectively.

In this section, we provide a test to be able to determine which of ESL (10) and ESN (28) best fits a data set. Let $X_1, ..., X_n$ be a random sample from an ESL or an ESN distribution. The likelihood function of the ESL is defined in (20), while the likelihood function of the ESN is defined by

$$L_{ESN}(\mu,\tau,\phi) = \Pi_{i=1}^{n} f_{ESN}(x_i,\mu,\tau,\phi),$$

where $f_{ESN}(x)$ are define in (28), define the statistic

$$T = \ln\left\{\frac{L_{ESN}(\hat{\mu},\hat{\tau},\hat{\phi})}{L_{ESL}(\hat{\theta},\hat{\varepsilon},\hat{\sigma})}\right\}, \tag{29}$$

where $(\hat{\mu},\hat{\tau},\hat{\phi})$ and $(\hat{\theta},\hat{\varepsilon},\hat{\sigma})$ represent the MLEs of the ESN and ESL distributions, respectively. Mudholkar et al (1999) show that, the MLEs of the ESN parameters are,

$$\hat{\tau}^2 = \frac{\left[\sum_{i=1}^{k}(x_{(i)}-\hat{\mu})^2\right]^{(1/3)} - \left[\sum_{i=1}^{k}(x_{(i)}-\hat{\mu})^2\right]^{(1/3)}}{\left[\sum_{i=1}^{k}(x_{(i)}-\hat{\mu})^2\right]^{(1/3)} + \left[\sum_{i=1}^{k}(x_{(i)}-\hat{\mu})^2\right]^{(1/3)}} \tag{30}$$

and

$$\hat{\phi}^2 = \frac{1}{n}\left[\frac{\sum_{i=1}^{k}(x_{(i)}-\hat{\mu})^2}{(1+\hat{\tau})^2} + \frac{\sum_{i=k+1}^{n}(x_{(i)}-\hat{\mu})^2}{(1-\hat{\tau})^2}\right], \tag{31}$$

where $X_{(i)}$ represents the order statistics. Using (30), (31), and (23), (24), and (25) in equation (29), we have

$$T = -\frac{n}{2}\,\ln 2 - \frac{n}{2}\ln\pi - n\,\ln\hat{\phi} - \frac{n\,\hat{\phi^2}}{2\,\hat{\phi^2}} + n\,\ln 2 - \frac{n}{2}\,\ln 2 + n\,\ln\hat{\sigma} + \frac{n\,\hat{\sigma^2}}{\hat{\sigma^2}},$$

$$= \frac{n}{2} - \frac{n}{2}\,\ln\pi - n\,\ln\hat{\phi} + n\,\ln\hat{\sigma}. \tag{32}$$

If $T > 0$, we choose the ESN distribution, otherwise we choose the ESL distribution as the preferred model.

5. Simulation Study

In this section, we provide a discriminant analysis simulation study between the ESL and the ESN or normal distributions. Using PROC IML and PROC NLP, in SAS, we generate data from normal, ESL, and ESN distributions. We use (27) to test for better model fitting between normal, ESL, and (32) between ESN and ESL. The simulation algorithm for the program is as follows

Step (1): Generate Data of 1000 observations from ESL, ESN and Normal distribution. Find the MLEs for the parameters using PROC NLP and PROC IML and equations (25) for ESL, (31) for ESN, and (22) for Normal distribution.

Step (2): calculate the test statistics T (27) and (32).

Step (3): Simulation Results: The simulation result show that,

- Using the data generated from the normal distribution, the test statistic T, (27), was greater than zero, which indicates that the normal distribution is a better fit for the data than the ESL distribution
- Using the data generated from ESL, the test statistic T, (27), was less than zero, which indicates that the ESL distribution is a better fit of the data than the normal distribution.
- Using the data generated from ESL, the test statistic T, (32), was less than zero, which indicates that ESL is a better fit of the data than the ESN
- Using the data generated from ESN, the test statistic T, (32), was grater than zero, indicating that the ESN is a better fit than the ESL distribution.

REFERENCES

Al-Mousawi, H., Elsalloukh, H., and Guardiola, H. (2012A), "Further Analysis of the Epsilon-Skew Laplace Distribution" *Universal Journal of Mathematics and Mathematical Sciences*, Vo. 2, No.2, 181-194.

Al-Mousawi, H., Elsalloukh, H., Guardiola, H., and McMillan, T. (2012B), "The Multivariate Epsilon Skew Laplace distribution", *Advances and Applications in Statistics Journal*, Vo. 27, No.1, 9-26.

Al-Mousawi H., Elsalloukh H., and McMillan T. (2011), "A New Class of Multivariate Power Distribution Family", *American Statistical Association*, JSM proceeding, Biometrics Section, Miami, FL, 1656-1662.

Azzalini A. and Capitanio A., (2003), "Distribution generated by perturbation of symmetry with emphasis on a multivariate skew t distribution", *Journal Roy. Stat. Soc.*, B 65:367389.

Azzalini A. and Valle D. D., (1996), "The multivariate skew-normal distribution", *Biometrika*, 83(4):715726.

Cox D. R., (1961), "Tests of separate families of hypotheses", *University of California Press*, Berkely, Proceedings of the Fourth Berkely Symposium in Mathematical Statistics and Probability, 105-123.

Elsalloukh H., (2008), "The Epsilon-Skew Laplace Distribution," In B. Section, editor, The 2008 Proceedings of the American Statistical Association.

Elsalloukh H., Guardiola H. J., and Young M. D., (2005), " The Epsilon-Skew Exponential Power Distribution Family." *Far East Journal of Theoretical Statistics,* 16(1), 97112.

Gutjahr S., Henze N. , and Folkers M., (1999) "Shortcomings of generalized affine invariant skewness measures", *Journal of Multivariate Analyses*, 71:123.

Henze N. and Zirkler B., (1990), "A class of invariance consistent test for multivariate normality", *Commun. Statistics - Theory Meth.*, 19:35953617.

Kollo T., (1991) "Matrix derivative for multivariate statistic", Masters thesis, Tratu University Press, Tartu, in Russian.

Kollo T., (2008), "Multivariate skewness and kurtosis measures with and application in ica", *Journal of Multivariate Analysis*, 99:23282338.

Kollo T. and Srivastava M., (2004), "Estimation and testing of parameters in multivariate laplace distribution", *Commun Stat Theory Methods*, 33:23632387.

Kotz S., Kozubowski T. J., and Podgorski K., (2001), *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*, Birkhauser, Boston .

Kundu D., (), "Discriminating between the normal and the laplace distributions", Report, Indian Institute of Technology, Department of Mathematics, Kanpur, India, 2004.

Mardia K. V., (1970), "Measures of multivariate skewness and kurtosis with applications", *Biometrika*, 57:512530.

Mardia K. V., (1974), "Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies", *Sankhya Ser.*, B36:115128.

Móri T. F., Rohatge V. K., and Székely G. J., (1993), "On multivariate skewness and kurtosis", *Theory Probab. Appl.*, 38:547551.

Mudholkar G. and Hutson A. D., (1999), "The epsilon-skew-normal distribution for analyzing near-normal data", *Journal of Statistical Planning and Inference*, 83:291309.