Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/080301

An Extension of the Skew-Normal Distribution

Haroon M. Barakat[1,∗], *Abdallah W. Aboutahoun*[2] *and Naeema N. EL-kadar*[2,3]

[1] Department of Mathematics, Faculty of Science, Zagazig University, Zagazig, Egypt
[2] Faculty of Science, Alexandria University, Alexandria, Egypt
[3] Department of Statistics Faculty of Science, Benghazi University, Benghazi-Libya

Received: 26 Sep. 2018, Revised: 4 Jul. 2019, Accepted: 19 Sep. 2019
Published online: 1 Nov. 2019

Abstract: The stable symmetric family of distribution functions (DF's) suggested by [1] is a family that contains the reverse of every DF belonging to it. It is revealed that the stable families are capable of describing many types of statistical data. We introduce a new stable family via a mixture of the skew-normal distribution and its reverse, after inserting a scale parameter and its reciprocal to the skew-normal distribution and its reverse, respectively. We show that this family contains all the possible types of DFs. Besides, it has a very remarkable wide range of the indices of skewness and kurtosis. Computational technique using EM algorithm is implemented for estimating the model parameters. Moreover, an application with a real data set is presented.

Keywords: Parametric family, mixture of distributions, skew-normal distribution, skewness, kurtosis

1 Introduction

In many applied areas like lifetime analysis, finance, insurance and biology, there is a clear need to find an appropriate distribution that represents the data in the best way. Data modeling is a great challenge. Therefore, the distribution theory was widely studied. The generalized families of distributions are appeared to provide great flexibility to model various types of data. The generalized distributions are also useful in survival analysis, where the focus in this case is on the survival and hazard rate functions; while in the data modeling, the focus is on the indices of skewness and kurtosis.

It is known that the kurtosis and skewness uniquely determine the type of the vast majority of the DFs. Therefore, practically, any DF should belong (up to affine transformation) to one and only one of the following nine types:

(1) symmetric and mesokurtic, denoted by "00";
(2) symmetric and leptokurtic (positive excess kurtosis or the DF has a more acute peak around the mean and fatter tails than normal DF), denoted by "0+";
(3) symmetric and platykurtic (negative excess kurtosis or the DF has a lower, wider peak around the mean and thinner tails), denoted by "0−",
(4) positive symmetric and mesokurtic, denoted by "+0";
(5) positive symmetric and leptokurtic, denoted by "++";
(6) positive symmetric and platykurtic, denoted by "+−";
(7) negative symmetric and mesokurtic, denoted by "−0";
(8) negative symmetric and leptokurtic, denoted by "−+";
(9) negative symmetric and platykurtic, denoted by "−−".

Therefore, the most essential factor of the quality of any generalized family of DFs is how much number of aforesaid types of DFs are included in it. Usually, the generalized families of distributions are constructed by inserting one or more additional shape parameter to a baseline distribution. Moreover, the generalized family is more flexible to analyze data than its baseline distribution provided that the following conditions are fulfilled:

1. Inclusion a large number of the possible nine types $00, 0+, 0-, +0, ++, +-, -0, -+, --$ of DFs.
2. Wide range of the indices of skewness and kurtosis.

∗ Corresponding author e-mail: hbarakat2@hotmail.com

On the other hand, since the normal distribution is the only known distribution, which is of the type 00, then it is naturally to take it as the base, or at least the family should include it as a DF.

For the time being, there are many generalized families of distributions that can be found in the literatures. Among those famous generalized families is the Marshall and Olkin family, suggested by [2], which has the property that the minimum of a geometric number of independent random variables (RVs) with a common distribution in the family has a distribution again in the family. For more details about this family, the reader is referred to [3,4,5,6]. Moreover, one of the foremost of the generalized families is the exponentiated family, which is obtained by raising the baseline distribution to a positive power. The statistical properties of this family were studied in detail by [7], see also [8].

Another important generalized family is the beta-distribution family. This family was extensively studied by [9], it is considered as a generalization of the exponentiated family. This family is obtained by substituting the independent variable in the incomplete beta function by the baseline distribution. For more general results on this family, see [10,11,12,13].

Cordeiro and Castro [14] suggested the Kumaraswamy (abbreviated by Kum) family as a more flexible generalized family than beta-distribution family. The Kum-family is built by superseding the independent variable in the Kum-distribution by any baseline distribution. The Kum-distribution has a simple form, which makes it much better suited than the Beta distribution for computation-intensive activities like simulation modeling and the estimation of models by simulation-based methods. For more general results on this family, the reader is referred to [15,16,17].

Azzalini [18] proposed an appropriate family, known as the skew-normal distribution, for the analysis of data which is unimodal but exhibits some skewness. For more details about the Azzalini's family, see [15,19,20].

It is known that in the mixture model, which is a convex combination of two, or more, probability density functions (PDFs) is a powerful and flexible tool for modeling complex data for combining the properties of the individual PDFs. The word mixture is used because the PDF of a random observation is a mixture of several (distinct) component density functions of the form $f(x) = \sum_{i=1}^{m} a_i f_i(x)$, with $\sum_{i=1}^{m} a_i = 1$, $a_i > 0$, for all $i = 1,2,...,m$, and $f_i(x)$, $i = 1,2,...,m$, are distinct densities with known form. Moreover, each one of these densities is specified by one or more unknown parameters. Finally, each observation comes from one of the m (distinct) component distributions with unknown membership status.

Barakat [1] suggested a new generalized family, denoted by ASSN (additive stable-symmetric normal), by inserting a location parameter to the standard normal distribution, as a baseline DF, and inserting the same location parameter with different sign to the reverse of the baseline distribution. By this way the used location parameter turns to a shape parameter to maneuver the skewness and kurtosis of the resulted family. He also showed that the ASSN family contains all the possible types of DFs, except the type $0+$, besides it has very remarkable wide range of the indices of skewness and kurtosis. Therefore, it is capable of describing many types of statistical data than many other known families.

Barakat and Khaled [21] called any generalized family that contains all the possible types of DFs (nine types), a full family. Moreover, they suggested two full families, via a mixture of the ASSN family and the standard logistic or Laplace distributions (note that both of these DFs are symmetric leptokurtic, i.e., have the type 0+). More recently Barakat et al. [22] carried out a comparison of most capable families of distributions for modelling asymmetry. Kum-normal family, ASSN and the two full families suggested by [21] were chosen, where the quality of the fit, the flexibility and the amount of asymmetry parameters were factors used for comparison.

Barakat et al. [23] suggested a more tractable full family with three parameters, denoted by MSSN (multiplicative stable-symmetric normal) family via mixture of normal distribution $(\nu,1), \nu \neq 0$ and its reverse, after multiplying and dividing, respectively, them to the same scale parameter. They also showed that via the theoretical and practical studies the MSSN family is more capable of fitting different types of data than the ASSN family. On the other hand, Barakat et al. [24] introduced another full family with three parameters, denoted by ASSA (additive stable-symmetric Azzalini), by using the standard skew-normal distribution instead of the standard normal distribution, in the definition of ASSN family.

In this paper, we suggest new full family with four parameters, denoted by MSSA (multiplicative stable-symmetric Azzalini), via a mixture of the skew-normal distribution and its reverse, after inserting a scale parameter and its reciprocal to the skew-normal distribution and its reverse, respectively. We show that the MSSA family is full and it has a very remarkable wide range of the indices of skewness and kurtosis. Moreover, we show that it outperforms the ASSN, MSSN and ASSA families via a successful application of a real data set.

2 An extension of the Azzalini distribution-MSSA family

Azzalini [18] introduced the skew-normal distribution by adding a shape parameter to the normal DF. Namely, the PDF of the skew-normal distribution is given by

$$\mathscr{A}(x;\lambda) = 2\phi(x)\Phi(\lambda x), \ -\infty < \lambda < \infty, -\infty < x < \infty, \tag{1}$$

where ϕ and Φ are the standard normal PDF and DF, respectively. The DF of the skew-normal distribution is given by

$$A(x;\lambda) = 2\int_{-\infty}^{x}\int_{-\infty}^{\lambda t}\phi(t)\phi(u)dudt. \tag{2}$$

Clearly $\mathscr{A}(x;0) = \phi(x)$. Moreover, $\o vA(-x;\lambda) = 1 - A(-x;\lambda) = A(x;-\lambda)$, $A(x;1) = \Phi^2(x)$ (which is the DF of the maximum of i.i.d two RVs from Φ) and if $W \sim A(x;\lambda)$, then $W^2 = Z^2$, i.e., $|W| \sim |Z|$, where $Z \sim \Phi(x)$. This implies that the even moments of W and Z are identical. Also, this implies the existence of the odd moments of W. Moreover, Azzalini [18] derived the moment-generating function of W, which is given by

$$M_W(t) = 2\exp\left(\frac{t^2}{2}\right)\Phi(\delta t),\ \delta = \frac{\lambda}{\sqrt{1+\lambda^2}}. \tag{3}$$

A RV Y is said to have a location-skew-normal distribution, with location at μ, denoted by $Y \sim SN(\mu,1,\lambda)$ if its PDF is given by

$$f(y;\mu,\lambda) = 2\phi(y-\mu)\Phi(\lambda(y-\mu)),\ y \in \Re\ (\lambda,\mu \in \Re). \tag{4}$$

If $\mu = 0$, we obtain the standard skew-normal PDF (1), denoted by $SN(\lambda)$. Clearly, (4) yields the moment-generating function of the RV Y as

$$M_Y(t) = 2\exp(\mu t + \frac{t^2}{2})\Phi(\delta t),\ \forall\, t \in \Re. \tag{5}$$

Moreover, by using (5), we can easily derive the mean, variance, the third and the fourth central moments, as

$$\left.\begin{array}{l}
\mathrm{E}(Y) = \mu_Y = b\delta + \mu, b = \sqrt{\frac{2}{\pi}},\\
\mathrm{E}(Y-\mu_Y)^2 = \sigma_Y^2 = C_Y^{[2]} = 1-(b\delta)^2,\\
\mathrm{E}(Y-\mu_Y)^3 = C_Y^{[3]} = (2b^2-1)b\delta^3,\\
\mathrm{E}(Y-\mu_Y)^4 = C_Y^{[4]} = 3+(4\delta^2-3b^2\delta^2-6)b^2\delta^2.
\end{array}\right\} \tag{6}$$

For any $0 \le \bar{\alpha} = 1-\alpha \le 1$ and $c > 0$, the suggested new family, denoted by MSSA family, is defined via a mixture of the DF $A(x;-\lambda)$ and its reverse $\o vA(-x;\lambda) = A(x;-\lambda)$, by

$$\begin{aligned}
G(z;\alpha,c,\mu,\lambda) &= \alpha A(\frac{z}{c}-\mu,\lambda) + \bar{\alpha}\bar{A}(-cz-\mu,\lambda)\\
&= \alpha A(\frac{z}{c}-\mu,\lambda) + \bar{\alpha}A(cz+\mu,-\lambda).
\end{aligned} \tag{7}$$

In the sequel, we attach the DF G with its RV (say Z) and write $G_Z(z;\alpha,c,\lambda)$ instead of $G(z;\alpha,c,\lambda)$ (especially, when we talk about the moments of $G(z;\alpha,c,\lambda)$). In this case, the stochastic representation for the MSSA family is simply given by $Z = IY_1 + (1-I)Y_2$, where I is a Bernoulli distributed RV with parameter α and the components Y_1 (distributed as $A(\frac{z}{c}-\mu,\lambda)$) and Y_2 (distributed as $A(cz+\mu,-\lambda)$) are independent of I. The PDF and the survival function of $G_Z(z;\alpha,c,\lambda)$ are given by

$$\begin{aligned}
g_Z(z;\alpha,c,\mu,\lambda) &= \frac{\alpha}{c}\mathscr{A}(\frac{z}{c}-\mu,\lambda) + \bar{\alpha}c\mathscr{A}(cz+\mu,-\lambda)\\
&= \frac{2\alpha}{c}\phi(\frac{z}{c}-\mu)\Phi(\lambda(\frac{z}{c}-\mu)) + 2\bar{\alpha}c\phi(cz+\mu)\Phi(-\lambda(cz+\mu))
\end{aligned} \tag{8}$$

and

$$\begin{aligned}
\bar{G}_Z(z;\alpha,c,\mu,\lambda) &= 1 - G_Z(z;\alpha,c,\mu,\lambda) = \alpha + \bar{\alpha} - \alpha A(\frac{z}{c}-\mu,\lambda) - \bar{\alpha}\bar{A}(-cz-\mu,\lambda)\\
&= \alpha\bar{A}(\frac{z}{c}-\mu,\lambda) + \bar{\alpha}A(-cz-\mu,\lambda)\\
&= \alpha\bar{A}(\frac{z}{c}-\mu,\lambda) + \bar{\alpha}\bar{A}(cz+\mu,-\lambda),
\end{aligned}$$

respectively. The following lemma gives the mean, variance, coefficient of skewness and coefficient of kurtosis of the MSSA family.

Lemma 1. *For the MSSA family, we have*

1. The mean is $\mu_Z = (\alpha c - \frac{\bar{\alpha}}{c})(b\delta + \mu)$.

2.The variance is $\sigma_Z^2 = \frac{1}{c^2}[(1+c^2)^2(b\delta+\mu)^2\alpha\bar{\alpha} + (1-(b\delta)^2)(\alpha c^4+\bar{\alpha})]$.
3.The coefficient of skewness is

$$\gamma_Z^{[1]} = \frac{C_Z^{[3]}}{\sigma_Z^3} = \frac{(1+c^2)^3(b\delta+\mu)^3\alpha\bar{\alpha}(1-2\alpha)}{((1+c^2)^2(b\delta+\mu)^2\alpha\bar{\alpha}+(1-(b\delta)^2)(\alpha c^4+\bar{\alpha}))^{3/2}}$$
$$+\frac{3(1-(b\delta)^2)(1+c^2)(b\delta+\mu)\alpha\bar{\alpha}(c^4-1)+(2b^2-1)b\delta^3(\alpha c^6-\bar{\alpha})}{((1+c^2)^2(b\delta+\mu)^2\alpha\bar{\alpha}+(1-(b\delta)^2)(\alpha c^4+\bar{\alpha}))^{3/2}}.$$

4.The coefficient of kurtosis is

$$\gamma_Z^{[2]} = \frac{C_Z^{[4]}}{\sigma_Z^4}$$
$$= \frac{(3+(4\delta^2-3(b\delta)^2-6)(b\delta)^2)(\alpha c^8+\bar{\alpha})+(1+c^2)^4(b\delta+\mu)^4\alpha\bar{\alpha}(\bar{\alpha}^3+\alpha^3)}{((1+c^2)^2(b\delta+\mu)^2\alpha\bar{\alpha}+(1-(b\delta)^2)(\alpha c^4+\bar{\alpha}))^2}$$
$$+\frac{6(1-(b\delta)^2)(1+c^2)^2\alpha\bar{\alpha}(b\delta+\mu)^2(\bar{\alpha}c^4+\alpha)}{((1+c^2)^2(b\delta+\mu)^2\alpha\bar{\alpha}+(1-(b\delta)^2)(\alpha c^4+\bar{\alpha}))^2}.$$
$$+\frac{4(2b^2-1)b\delta^3(1+c^2)(b\delta+\mu)\alpha\bar{\alpha}(c^6+1)}{((1+c^2)^2(b\delta+\mu)^2\alpha\bar{\alpha}+(1-(b\delta)^2)(\alpha c^4+\bar{\alpha}))^2}.$$

*Proof.*The proof follows after some routine algebra by using the relation (6) and applying Theorem 2.1 in [23].

Theorem 1.*The family MSSA is full. Moreover, it has a very remarkable wide range of the indices of skewness and kurtosis, comparing with the skew-normal distribution.*

*Proof.*By putting $\lambda = 0$ in (7), we can easily see that the MSSA family is converted to MSSN family. On the other hand, since the family MSSN is full then the family MSSA is also full with wider range of the indices of skewness and kurtosis.

Table 1 presents some values of $\gamma_Z^{[1]}$ and $\gamma_Z^{[2]}$ that covered all possible types of DFs. For computing these moments, we have constructed a simple code by Matlab 8.2 on laptop Intel 1.8GHZ processor. Table 1 does not only confirm that the MSSA family contains all the possible types of DFs, i.e., it is a full family, but it possesses very remarkable wide range of the indices of skewness and kurtosis. For example, $\gamma_Z^{[2]}$ reaches to 23.6844, while the maximum value of the coefficient of kurtosis of Azzalini's family is 3.869. On the other hand, $-2.61 \leq \gamma_Z^{[1]} \leq 3.1007$, while for the skew-normal distribution, we have $-0.995 \leq$ the coefficient of skewness ≤ 0.995. Finally, since for the skew-normal distribution, we have $3 \leq$ the coefficient of kurtosis ≤ 3.869 (cf., [18]), then the skew-normal distribution contains at most $\{00, 0+, +0, ++, -0, -+\}$ types of DFs. We include the graph of the PDF, defined by (8), of the MSSA family for some selected values of $\gamma_Z^{[1]}$ and $\gamma_Z^{[2]}$ that covered all possible types of DFs. These graphs (Figure 1) show that this family is extremely fertile and rich.

3 Statistical inference for the parameters of the MSSA family

In this section, we carry out a simulation study with different sample sizes and different parameter values of MSSA family defined by (7). First we can write the family (7) in a more convenient form for the inferential purposes

$$G(z;\alpha,c,\mu,\lambda) = \alpha A(\frac{z}{c}-\mu,\lambda)+\bar{\alpha}\bar{A}(-cz-\mu,\lambda)$$
$$= \alpha A(\frac{z-c_1}{\sigma_1},\lambda)+\bar{\alpha}A(\frac{z-c_2}{\sigma_2},-\lambda), \tag{9}$$

where

$$c_1 = c\mu,\ c_2 = \frac{-\mu}{c},\ \sigma_1 = c \text{ and } \sigma_2 = \frac{1}{c} = \frac{1}{\sigma_1}, 0 \leqslant \alpha \leqslant 1. \tag{10}$$

In this simulation study we focus on the shape parameters c and λ, besides the location parameter μ. We compute the maximum likelihood estimates (MLE) of the parameters by using the simulated samples via the EM algorithm. The EM algorithm is an iterative method for approximating the maximum of a likelihood function. It is worth mentioning that the EM algorithm is guaranteed to monotonically converge to local optima under mild continuity conditions (cf. [25]).

Table 1: Some selected values of $\gamma_Z^{[1]}$ and $\gamma_Z^{[2]}$ that covered all possible types of DFs

Types	$-\infty < \lambda < \infty$	$-\infty < \mu < \infty$	$0 \le \alpha \le 1$	$0 < c < \infty$	μ_Z	σ_Z^2	$\gamma_Z^{[1]}$	$\gamma_Z^{[2]}$
00	0	1	0	0.5	-2	4	0	3
00	1	0	0.5	1	0	1	0	3
00	0	0.5	1	3	1.5	9	0	3
$0+$	0	0	0.2	0.5	0	3.25	0	3.6391
$0+$	0	0	0.1	2	0	0.625	0	12.72
$0+$	0	0	0.9	2	0	3.625	0	3.2889
$0-$	1	0.5	0.5	1	0	1.8142	0	2.329
$0-$	1	1	0.5	1	0	3.1284	0	1.8289
$0-$	0	1.5	0.1	0.5	-2.625	4.8906	0	2.5486
$+0$	0	-1	0.413160144	0.5	0.9671	3.966	0.7801	3
$+0$	0	-1	0.788675135	1	-0.5773	1.6667	0.3578	3
$+0$	0	1	0.211324865	1	-0.5774	1.6667	0.3578	3
$++$	0	1	0.1	3	0	1.25	3.1007	23.6844
$++$	0.5	0.5	0.1	1.5	-0.3856	0.8556	1.5818	7.8617
$++$	1	1	0.2	1	-0.9385	2.2476	0.8586	3.1881
$+-$	0.5	0.5	0.3	1	-0.3427	1.4894	0.2283	2.8151
$+-$	1	1	0.1	0.5	-2.7373	3.8474	0.1395	2.7198
$+-$	0	1	0.4	1	-0.1	1.24	0.0348	2.9313
-0	0	1	0.413160144	0.5	-0.9671	3.966	-0.7801	3
-0	0	1	0.788675135	1	0.5773	1.6667	-0.3578	3
-0	0	-1	0.211324865	1	0.5774	1.6667	-0.3578	3
$-+$	1	0	0.9	0.5	0.141	0.6051	-2.61	15.2648
$-+$	1	0.5	0.7	0.5	-0.266	2.4237	-1.602	5.0495
$-+$	1	1	0.9	1	1.2514	1.5625	-1.0971	4.7071
$--$	1	1	0.9	2	2.7373	3.8478	-0.1395	2.7198
$--$	0.5	1	0.1	0.5	-2.3744	4.199	-0.038	2.567
$--$	1	1	0.7	1	0.6257	2.7369	-0.5612	2.3598

It is worth mentioning that, the present study covers all the possible types of DFs. Actually, we select the second case of each type in Table 1 (with given values of μ, c, λ). Moreover, for each case, we choose appropriate values of the two locations parameters c_1 and c_2, in the model (9) (e.g, in the second case of the first type we should have $-\infty < c_1 = c_2 < \infty$). Finally, by using the relations defined by (10), we compute the values of the remaining parameters μ, σ_1. On the other hand, we could now generate three groups of random samples each of size 100, 1000 and 5000 from the family (9). Furthermore, we repeat this simulation 10 times for each group. For every replication in each group, we compute the MLE of the parameters $c_1, c_2, \sigma_1, \sigma_2, \lambda$ and again by using the relation defined by (10), we compute the estimate of μ. Moreover, we compute the average estimate values for each parameter and group (each average estimate is computed for the 10 replications of each group). Finally, we attach these averages (as the estimates of the parameters) with their mean square errors (MSE). All the computation are carried out by using the package mixsmsn in the R package. The summary of this study is shown in Table 2, where the MLEs average estimate of the parameters μ, c and λ are denoted by $\hat{\mu}$, $\hat{c}$ and $\hat{\lambda}$.

The results in Table 2 show that with increasing the sample size the estimated values become more close to the true values. Moreover, based on the MSE, the three best estimates (in order) are obtained for the types $0-, 00$ and $-+$, while the worst estimates are appeared in the types $0+$, $+0$ and $--$.

Remark 3.1 One major concern in estimating the parameters of the mixture DFs is the identifiability problem. In our case, the family (8) is not identifiable (for the definition, see [26]) if we have, for at least two values $\bar{\upsilon}'$ and $\bar{\upsilon}''$ of the vector $\bar{\upsilon} = (\alpha, c, \mu, \lambda)$, $g_Z(z; \bar{\upsilon}') = g_Z(z; \bar{\upsilon}'')$, while $\bar{\upsilon}' \neq \bar{\upsilon}''$, where $g_Z(z; \bar{\upsilon}) = g_Z(z; \alpha, c, \mu\lambda)$. It is worth mentioning that (cf. [26]) in the class of mixture of normal distributions is generally not identifiable (when we put $\lambda = 0$ in the family (8)). This directly implies that the family (8) is, in general, not identifiable. However, in our study, the absence of identifiability is not "bad" because our main focus is not the estimation problem itself, but to pick up a suitable family (with estimated parameters) to describe the given data (even if this family is not unique).

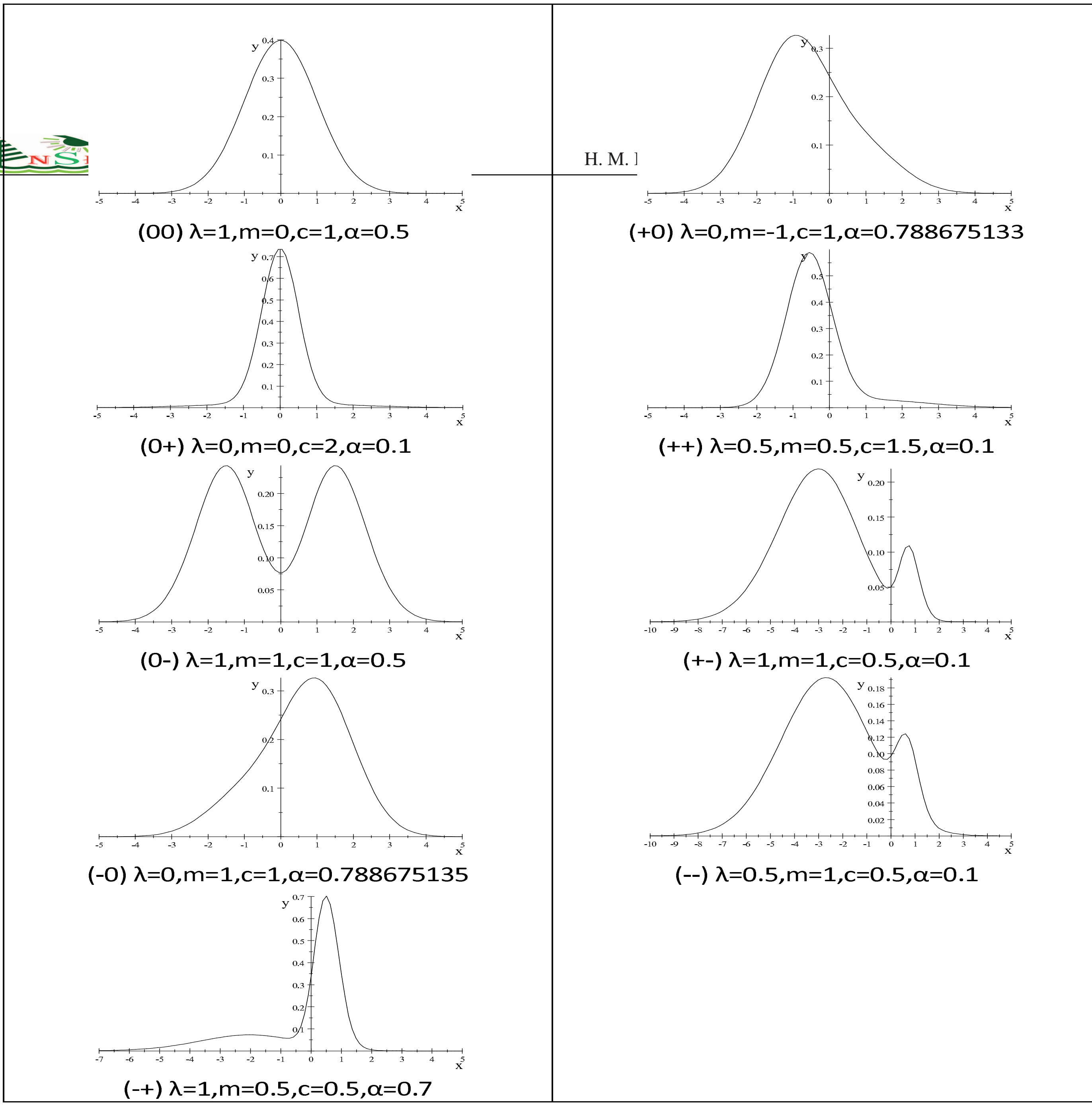

Fig. 1: MSSA family

4 Application

We have seen in Section 3, that at least theoretically the proposed MSSA family is capable of fitting a wide spectrum of real-world data set as it contains all the possible types of data (i.e., it is a full family). Moreover, it possesses a very wide range of the indices of skewness and kurtosis. In this section we can confirm its outperforming than the three competitors ASSN, MSSN and ASSA families via an example of a real data concerning the air pollution. The location-scale ASSA, ASSN and ASSM families are respectively defined by

$$G^{\langle ASSN\rangle}(x;\alpha,c,\mu,\sigma)=\alpha\Phi(\frac{x-c_1}{\sigma_1})+\bar{\alpha}\Phi(\frac{x-c_2}{\sigma_2}), \tag{11}$$

where $\sigma_1=\sigma_2=\sigma$, $c_1=\mu-c$, $c_2=\mu+c$ $0\leqslant\alpha\leqslant 1$, and

$$G^{\langle MSSN\rangle}(x;\alpha,c,\nu,\mu)=\alpha\Phi(\frac{x-c_1}{\sigma_1})+\bar{\alpha}\Phi(\frac{x-c_2}{\sigma_2}), \tag{12}$$

Table 2: Simulation study

Types	n	μ	$\hat{\mu}$	$mse(\hat{\mu})$	c	$\hat{c}$	$mse(\hat{c})$	λ	$\hat{\lambda}$	$mse(\hat{\lambda})$
00	100	0	0.1687	0.2159771	1	0.8362	0.3055952	1	1.3039	0.6483723
00	1000	0	0.0459	0.0550754	1	0.9601	0.06505306	1	1.3261	0.3612989
00	5000	0	0.0379	0.04448033	1	0.976	0.04210463	1	1.3116	0.317714
0+	100	0	-0.1226364	0.2661032	2	0.8576	1.25745	0	2.5368	2.870444
0+	1000	0	0.0055	0.4526762	2	1.1	0.9197888	0	1.9411	1.973493
0+	5000	0	-0.0648	0.3967876	2	1.0394	0.9623308	0	1.855	1.861411
0−	100	1	0.9527	0.1928422	1	1.146	0.463589	1	1.1935	0.6505823
0−	1000	1	1.01	0.03455431	1	1.0013	0.07467998	1	1.0081	0.2264215
0−	5000	1	1.0027	0.01583351	1	1.0077	0.03153569	1	0.9828	0.04612158
+0	100	-1	0.1604	1.209067	1	1.3427	0.9927521	0	1.0453	1.421823
+0	1000	-1	-0.0267	0.9868979	1	1.547	0.6998201	0	1.225	1.250351
+0	5000	-1	-0.0042	1.010152	1	2.853	0.6394886	0	1.1797	1.184315
++	100	0.5	0.3196	0.3352471	1.5	1.0431	0.7066134	0.5	1.7261	1.5392
++	1000	0.5	0.1389	0.3673964	1.5	1.4944	0.9336483	0.5	1.5601	1.244518
++	5000	0.5	0.1656	0.0336376	1.5	1.277	0.7470195	0.5	1.4738	1.121107
+−	100	1	0.9234	0.4502755	0.5	1.7864	1.507245	1	1.7325	1.992738
+−	1000	1	0.94	0.1204085	0.5	1.0439	0.9725704	1	0.6645	0.5526503
+−	5000	1	0.9233	0.1066897	0.5	0.9452	0.9013832	1	0.6261	0.5544521
−0	100	1	0.2689	0.7635499	1	1.5023	1.72884	0	1.5537	1.817578
−0	1000	1	0.1605	0.8442079	1	1.6078	0.7009128	0	1.3271	1.345603
−0	5000	1	0.1946	0.806707	1	1.3533	0.4359857	0	1.1661	1.170687
−+	100	0.5	0.7572	0.4314789	0.5	1.4727	1.342692	1	2.2115	1.668127
−+	1000	0.5	0.5995	0.1701285	0.5	1.3584	1.107632	1	1.0691	0.1647626
−+	5000	0.5	0.5106	0.0369107	0.5	1.4004	1.174064	1	1.0572	0.1558544
−−	100	1	0.4478	0.6891238	0.5	2.0671	1.890185	0.5	1.4973	1.593874
−−	1000	1	0.3976	0.631724	0.5	1.9716	1.581154	0.5	1.1119	0.896357
−−	5000	1	0.4859	0.5221564	0.5	2.0458	1.597412	0.5	1.0156	0.7762126

where $c = \sigma_1 = \frac{1}{\sigma_2}$, $c_1 = \mu + c\nu$, $c_2 = \mu - \frac{\nu}{c}$, $0 < \alpha < 1$.

$$G^{\langle ASSA\rangle}(x;\alpha,c,\lambda,\mu,\sigma) = \alpha A(\frac{x-c_1}{\sigma_1},\lambda) + \bar{\alpha}A(\frac{x-c_2}{\sigma_2},-\lambda), \tag{13}$$

where $\sigma_1 = \sigma_2 = \sigma$, $c_1 = \mu - \sigma c$, $c_2 = \mu + \sigma c$ and $\sigma > 0, 0 \leqslant \alpha, \ \leqslant 1$.

Example 4.1. We have compared the performances of the four families defined in (9), (11), (12) and (13) in fitting a data set for air pollution from the London Air Quality Network (LAQN). This network is a unified resource for air pollution measurements which are fundamental to support air quality administration. The greater part of London's 33 districts supply estimations to the system and also this information is progressively supplemented by estimations from neighborhood specialists encompassing London, in this manner giving a general point of view of air pollution in London and the home countries .

The data are taken from the site Greenwich Plumstead high street, that monitors Nitrogen oxides, sulphur dioxide, PM2.5 and climatology data. The hourly measurement of PM2.5 is recorded every hour. So a total of 53000 readings have been obtained in the period from 1-1-2015 to 31-12-2015. These data can be downloaded by any researcher in the form of a report every half hour, every hour or every day according to the type of study from the following site: "www.londonair.org.uk/london/asp/datadownload.asp ". The hourly of these data was used for application of these models. The summary statistics for data set is given in Table 3. We have compared the performances of the four families defined in (9), (11), (12) and (13) by using Akaike information criterion (AIC) (cf. [27]). This criterion is based on the likelihood value of the model, the number of observations and the number of parameters thereof. The computations of the estimates of the parameters and the values of AIC criterion are carried out by using the mix.print() function in the mixsmsn Package in the R-Package (see Table 4). Table 4 shows that the MSSA family has the best performance.

Table 3: summary statistics for data set

minimum	maximum	mean	variance	skewness	kurtosis
1.3	95.1	14.74992	100.1687	2.758921	10.8712

Table 4: Comparison between the ASSN, MSSN, ASSA and MSSA families

Family	Parameter estimations	log-likelihood estimate	AIC
ASSN	$\hat{c}_1 = 27.538$ $\hat{c}_2 = 11.078$ $\hat{\sigma} = 13.349$	-30217.37	60442.74
MSSN	$\hat{c}_1 = 11.083$ $\hat{c}_2 = 27.513$ $\hat{\sigma}_1 = 13.386 = \frac{1}{\hat{\sigma}_2}$	-28804.38	57618.76
ASSA	$\hat{c}_1 = 6.815$ $\hat{c}_2 = 14.799$ $\hat{\sigma} = 36.951$ $\hat{\lambda} = 2.744$	-34642.11	69294.22
MSSA	$\hat{c}_1 = 6.815$ $\hat{c}_2 = 14.799$ $\hat{\sigma} = 36.951 = \frac{1}{\hat{\sigma}_2}$ $\hat{\lambda} = 2.744$	-25984.84	51981.68

Acknowledgement

The authors are grateful to the editorial board and the referees for suggestions that improved the presentation substantially.

References

[1] H. M. Barakat, A new method for adding two parameters to a family of distributions with application to the normal and exponential families. Statist. Meth. & Appl. (SMA), 24(3), 359-372 (2015).
[2] A. W. Marshall and I. Olkin, A new method for adding a prameter to a family of distributions with application to the exponential and Weibull families. Biometrika, 84, 641-652 (1997).
[3] H. M. Barakat, M. E. Ghitany and E. K. AL-Hussaini, Asymptotic distributions of order statistics and record values under the Marshall-Olkin parameterization operation. Comm. Statist. Theory Methods, 38, 2267-2273 (2009).
[4] K. K. Jose, Marshall-Olkin family of distributions and their applications in reliability theory, time series modeling and stress strength analysis. Proc. ISI 58th World Statist. Congr. Int. Stat. Inst., 21st-26th August 2011, Dublin, 3918-3923 (2011).
[5] W. Barreto-Souza, A. J. Lemonte and G. M. Cordeiro, General results for Marshall and Olkin's family of distributions. An Acad. Bras. Cienc. 85, 3-21 (2013).
[6] G. M. Cordeiro, A. J. Lemonte and E. E. M. Ortiga, The Marshall-Olkin family of distributions: Mathematical properties and new models. J. Statat, Theor. Pract. 8, 343-366 (2014a).
[7] G. E. P. Box and G. C. Tiao, Bayesian inference in statistical analysis. Addision-Wesley, Reading, MA, 1973.
[8] E. K. AL-Hussaini and M. Ahsanullah, Exponentiated distributions. Atlantis Press, Paris-France, 2015.
[9] M. C. Jones, Families of distributions arising from distributions of order statistics. Test, 13(1), 1-43 (2004).
[10] N. Eugene, C. Lee and F. Famoye, Beta-normal distribution and its applications. Commun. Comm. Statist. Theory Methods, 31: 497-512 (2002).
[11] V. Mameli and M. Musio, A Generalization of the skew-normal distribution: The Beta skew-normal. Comm. Statist. Theory Methods, 42(12), 2229-2244 (2013).
[12] A. J. Lemonte, The beta log-logistic distribution. Braz. J. Probab. Stat. 28, 313-332 (2014).
[13] H. M. Barakat and E. M. Nigm, Asymptotic distributions of order statistics and record values arising from the class of beta-generated distributions. Statistics, 48(5), 1005-1012 (2014).
[14] G. M. Cordeiro and M. Castro, A new family of generalized distributions. J. Statist. Comp. & Simul. 81(7), 883-898 (2011).
[15] M. A. Correa, D. A. Nogueira and E. B. Ferreira, Kumaraswamy normal and Azzalini's skew normal modeling asymmetry. Sigmae, 1, 65-83 (2012).

[16] G. M. Cordeiro, E. E. M. Ortiga and G. O. Silva, The Kumaraswamy modified Weibull distribution: Theory and applications. J. Stat. Comput. Simul. 84, 1387-1411 (2014b).
[17] H. M. Barakat, E. M. Nigm and M. H. Harpy, Limit theorems of order statistics and record values from the gamma and Kumaraswamy-generated-distributions. Bull. Malays. Math. Sci. Soc., 40, 1055-1069 (2017).
[18] A. Azzalini, A class of distributions which includes the normal ones. Scand. J. Statist. 12, 171-178 (1985).
[19] A. Azzalini and A. Capitanio, The skew-normal and related families. New York: Cambridge University Press, 2014.
[20] B. V. Popović, G. M. Cordeiro and M. A. R. Pascoa, A new extended mixture normal distribution. Math. Comm. 22, 53-73 (2017).
[21] H. M. Barakat and O. M. Khaled, Towards the establishment of a family of distributions that best fits any data set. Comm. Statist.-Sim. Comp. 46(8), 6129-6143 (2017).
[22] H. M. Barakat, O. M. Khaled and Rakha, N. Khalil, Modern techniques in data analysis, with application to the water pollution. Proceedings of the Latvian Academy of science, Section B, 72(3)(714), 184-192(2018).
[23] H. M. Barakat, A. W. Aboutahounand and N. N. EL-kadar, On some generalized families arising from mixture normal distribution with applications. Comm. Statist.-Sim. Comp. "https://doi.org/10.1080/03610918.2018.1554110"(2019a).
[24] H. M. Barakat, A. W. Aboutahounand and N. N. EL-kadar, A new extended mixture skew-normal distribution, with applications. Revista Colombiana de Estadística, to appear (2019b).
[25] C. F. Wu, On the convergence properties of the EM algorithm. Ann. Statist. 11(1), 95-103 (1983).
[26] H. Teicher, Identifiability of finite mixtures. Ann. Math. Statist. 34, 1265-1269 (1963).
[27] H. Akaike, Information theory and the maximum likelihood principle. In: B. N. Petrov and F. Csaki (eds.), Second International Symposium on Information Theory, Akademiai Kiado, Budapest: 267-281, 1973.

Haroon M. Barakat is Professor of Mathematical Statistics at Zagazig University. He received the Ph.D. degree in 1986-Mathematical Statistics- Moscow State University. His research interests are in the areas of statistical ordered data and its applications. He has published research articles in reputed international journals of mathematical statistics. He is referee and editor of mathematical journals. In 1989 he was awarded the Egyptian State Prize in mathematics. He was awarded in 2000, the prize of National Egyptian Committee of Mathematics in pure mathematics. In 2007, he was awarded State Award for Excellence in basic science. In 2010 he was awarded Discretionary Zagazig University Prize and in 2012, he was awarded the International Fiza Al-Kharafi Prize in the field of mathematics.

Abdallah W. Aboutahoun received his PhD degree in Network Optimization from Department of Mathematics, Northeastern University, Boston, Massachusetts, USA. Currently, he is an Associate Professor at the Department of Applied Mathematics and Information Science, Zewail City of Science and Technology and Department of Mathematics, Faculty of Science, Alexandria University, Egypt. His research interests are in the areas of combinatorial optimization, location analysis, network reliability and multiple criteria decision making. He has published in journals such as Applied Mathematical Modeling, Int. J. Operational Research, Computational and Applied Mathematics, Journal of Computational Mathematics, Bulletin of the Polish Academy of Sciences, Applied Mathematics and Computation.

Naeema N. EL-kadar was a Ph.D. student at the Faculty of Science- Alexandria University. She received the Ph.D. degree in 2019-Mathematical Statistics. She got the bachelor of science (in Statistics) in 2000, University of Benghazi, Libya. She earned the M.Sc. degree in Statistics in 2008 from Benghazi University-Libya. Her research interests are in the areas of the distributions theory.