Shahid Bahonar University of Kerman
Journal of Mahani Mathematical Research
Print ISSN: 2251-7952 Online ISSN: 2645-4505

A NEW WEIGHTED DISTRIBUTION BASED ON THE MIXTURE OF ASYMMETRIC LAPLACE FAMILY WITH APPLICATION IN SURVIVAL ANALYSIS

N. Gilani and R. Pourmousa

Article type: Research Article
(Received: 12 March 2023, Received in revised form 07 June 2023)
(Accepted: 24 June 2023, Published Online: 24 June 2023)

Abstract. The generalization of asymmetric Laplace (AL) distribution has recently received considerable attention in dealing with skewed and long-tailed data. In this article, we introduce a new family of distributions based on the location mixture of asymmetric Laplace (LM-AL) distribution. Some properties of this family, such as density function, moments, skewness and kurtosis coefficients are derived. We show that this family of distributions is quite flexible because it has wider ranges of skewness and kurtosis than the other skew distributions introduced in the literature.

We also generalize the weighted exponential distribution introduced by Gupta and Kundo (2009) and show that truncated LM-AL distribution in zero belongs to this family. This family of distributions represents a suitable alternative to existing models such as Weibull, log-normal, log-logistic, gamma, and Lindley distributions. The performance and applicability of the proposed model in survival analysis are illustrated by analyzing a simulation study and two real data sets. To compute the maximum likelihood (ML) estimation of the parameters in the LM-AL distribution, an EM-type algorithm is developed and the estimation of the parameters of the model in survival analysis is performed using a maximization algorithm, due to the problem complexity.

Keywords: Asymmetric Laplace distribution, Location mixture distribution, EM-algorithm, Survival analysis.
2020 MSC: 62Nxx.

1. Introduction

Few real-world phenomena that need to be studied statistically are symmetric. Thus, the symmetric models will not be useful for studying all phenomena. In the last few decades, data sets in financial studies, image processing, signal processing, reliability, and survival analysis have shown that following the rules of normal distribution is an exception rather than a rule. The existence of asymmetric observations in various scientific fields has led researchers to construct and generalize distributions that can provide a good fit for modeling this type of phenomenon.

During the last three decades, several important classes of distributions have been discussed extensively in this regard, such as the families of mean mixture, variance mixture, and mean-variance mixture of normal distributions (see Pourmousa et al.

✉ pourm@uk.ac.ir, ORCID: 0000-0001-6412-7260
DOI: 10.22103/jmmr.2023.21226.1418
Publisher: Shahid Bahonar University of Kerman
© the Author(s)
How to cite: N. Gilani, R. Pourmousa, *A new weighted distribution based on the mixture of asymmetric Laplace family with application in survival analysis*, J. Mahani Math. Res. 2024; 13(1): 229-249.

(2015), Naderi et al. (2018), and Naderi et al. (2023)). In variance mixtures of normal distributions, it is assumed that the variance is not constant for all members of the population. But in some cases, in addition to the inequality of variance, it may also have a non-constant mean. To model such data sets variance mixture of normal distributions were extended to variance–mean mixture of normal distributions. Recently Negarestani et al. (2018) introduced a general class of skewed distributions based on mean-mixtures of normal distribution.

This idea became a motivation for us to extend this family to long-tailed distributions such as Laplace. Among the family of important symmetric distributions such as normal, logistic, uniform, Laplace, t-student, etc., the Laplace distribution is very popular due to its towering peak, heavy tails and attractive possibilities. A generalized family of Laplace distribution is the asymmetric Laplace distribution. In this paper, we use a version of the asymmetric Laplace distribution that possesses many important statistical properties such as infinite divisibility, geometric infinite divisibility, and stability with respect to geometric summation and can accommodate asymmetry, peakedness, and tail heaviness. One of the aims of this paper is to increase the flexibility and introduce a specific class of skewed distributions based on the location mixtures of asymmetric Laplace (LM-AL) distribution, using a similar idea to Negarestani et al. (2018).

Another goal of this article is to generalize the family of weighted exponential distributions using the location mixture of exponential distribution and find its relationship with the LM-AL distribution, which is our main motivation for writing this paper. It should be noted that this distribution, which is a generalization of the distribution introduced by Gupta and Kundo (2009), has a much better fit than other common models in survival analysis.

The weighted exponential distribution, which is a competitor to the Weibull, gamma, and generalized exponential distributions, has a significant application in engineering and medical fields, which was proposed by Gupta and Kundo (2009) as a generalization of the exponential distribution. The importance of this distribution, in addition to its application, is in its introduction method. Different methods may be used to introduce a shape parameter in the generalization of the exponential model and may lead to a variety of weighted exponential distributions. Gupta and Kundo (2009) used the idea of Azzalini (1985) to introduce a shape parameter to an exponential distribution, which led to a new class of weighted exponential distributions that are more flexible than other models. We show that this distribution can also be produced based on the location mixture of the exponential distribution in addition to Azzalini's method. The weight function in this distribution is based on the cumulative distribution function. We also were able to generalize this function in terms of the survival function, so that the distribution introduced by Gupta and Kundo (2009) becomes a special case of it. This feature has made our model more flexible and adaptable than the model introduced by Gupta and Kundo (2009). We also showed that the zero-truncated distribution of the LM-AL distribution belongs to this family.

The rest of this paper is organized as follows: In Section 2, the asymmetric Laplace (AL) distribution and several of its related representations and properties were explored. In Section 3, we first discussed the location mixture of the exponential and the weighted exponential in terms of the survival function (WES) distributions, and then use it to introduce the location mixture of the asymmetric Laplace (LM-AL) family. An ECM algorithm is developed to get parameter estimates of the LM-AL model in Section 4. In Section 5, we have examined the application of the LM-AL family in survival analysis and introduced a new model for the analysis of time-to-event data, which belongs to the family of WES distribution. We also develop an MCECM algorithm for maximum likelihood estimation of the parameters of the WES distribution in Section 5. In Section 6, two simulation studies are conducted to examine the performance of the maximum likelihood estimators of the parameters of the models and finally, the applicability of the proposed model in survival analysis is demonstrated using two real data sets.

2. Asymmetric Laplace Distribution

The Laplace distribution and its various generalizations have been widely applied in many different aspects of life sciences, economics, finance, and risk analysis. When dealing with asymmetric data, skewed models are an essential tool for the ensuing analysis. Several asymmetric forms of Laplace distribution have appeared in the literature; see, for example, Kotz et al. (2001). Their method, based on the idea of Steel's, converts the Laplace distribution into a skew-Laplace by postulating a scale parameter. In this paper, we discuss a version of asymmetric Laplace distribution that has attracted the attention of many researchers due to its attractive probabilistic properties.

Definition 2.1. The random variable Y has an asymmetric Laplace distribution, denoted by AL(β), if its PDF is given by

$$f(y;\beta) = \frac{1}{2\alpha}\, e^{-\alpha|y|+\beta y}, \qquad y \in \mathbb{R}, \tag{1}$$

where $\beta \in \mathbb{R}$ is the skewness parameter and $\alpha = \sqrt{\beta^2+1}$.

The asymmetric Laplace distribution introduced in (1) possesses many important statistical properties such as infinite divisibility, geometric infinite divisibility, and stability with respect to geometric summation. Moreover, this model has considerable flexibility and can accommodate asymmetry, peakedness and tail heaviness, which are commonly encountered in many data sets. The skew-Laplace distributions can also be obtained as the difference between two exponential distributions which is introduced in the following theorem. It can be easily shown that many of the skew Laplace distributions introduced in the literature become special cases of it.

Proposition 2.2. *Let X_1 and X_2 be independent exponential random variables with mean λ. Then, for any positive-valued number τ, the random variable $Y = X_2 - \tau X_1$*

has an asymmetric distribution, denoted by $SL(\lambda,\tau)$, about the origin with PDF

$$f(y;\lambda,\tau)=\frac{1}{\lambda(1+\tau)}\begin{cases}e^{\frac{y}{\lambda\tau}}, & y<0,\\ e^{-\frac{y}{\lambda}}, & y\geq 0.\end{cases} \tag{2}$$

Remark 2.3. The probability density function (PDF) in (2) is a skewed Laplace distribution, denoted by $SL(\lambda,\tau)$, and the cases $0<\tau<1$ *and* $\tau>1$ correspond to positive and negative skewness, respectively. We also observe that when $\tau=1$, (2) reduces to the symmetric Laplace distribution.

There are several alternative parameterizations for the PDF in (2). For example, the skew Laplace distributions introduced by Fernandez and Steel (1998) are special cases of this family. Note that the skew Laplace distribution introduced in Proposition 2.2 is a special case of the asymmetric Laplace distribution if $\tau=\frac{\alpha-\beta}{\alpha+\beta}$ and $\lambda=\frac{1}{\alpha-\beta}$. Some applications of this AL distribution can be found in Fernandez and Steel (1998). Upon adding location and scale parameters, $\mu\in\mathbb{R}$ and $\sigma>0$, respectively, and denoting it by $Y\sim AL(\mu,\sigma,\beta)$, we have its PDF as

$$f(y;\mu,\sigma,\beta)=\frac{1}{2\alpha\sigma}e^{-\frac{\alpha}{\sigma}|y-\mu|+\frac{\beta}{\sigma}(y-\mu)},\qquad y\in\mathbb{R}.$$

Figure 1 presents the PDFs for different parameter values of the AL distribution, from which it can be seen that β is a parameter that impacts skewness and kurtosis of the AL distribution in the same and opposite directions, respectively.

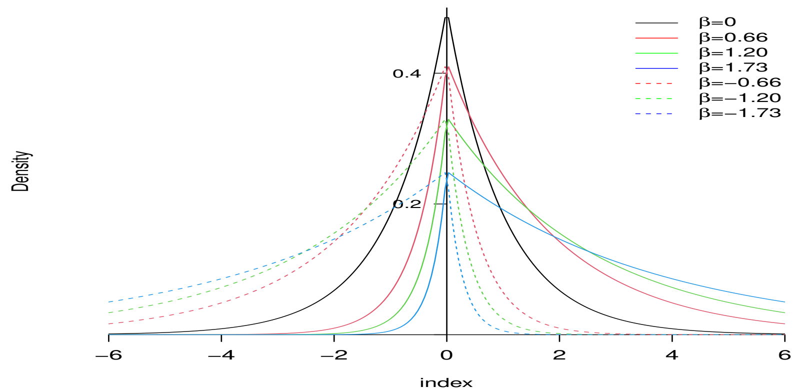

FIGURE 1. The asymmetric Laplace PDF for different values of β (with $\mu=0,\ \sigma=1$).

2.1. **Representations and Properties of AL distribution.** In this subsection, we present some important representations of the asymmetric Laplace distribution that can be used to simulate this distribution.

(I) **Mean-mixture representation**: If $W_1 \overset{d}{=} W_2 \sim Exp(\alpha+\beta)$ and $W_1 \perp W_2$, then

$$Y \overset{d}{=} W_2 - (\alpha-\beta)^2 W_1 \sim AL(\beta).$$

(II) **Variance mixture representation**: Let U be a discrete random variable with pdf

$$g(u;\beta) = \frac{1}{2\alpha}\begin{cases} \alpha+\beta & \text{if} \quad u=\beta+\alpha; \\ \alpha-\beta & \text{if} \quad u=\beta-\alpha. \end{cases}$$

If $W \sim Exp(1)$ and $W \perp U$, then, Y admits the stochastic representation $Y = UW$.

(III) **Mean-variance mixture representation**: Let $Z \sim N(0,1)$ and $W \sim \chi^2_{(2)}$, where $Z \perp W$. Then, the random variable Y can be generated by the representation

$$Y = \beta W + \sqrt{W} Z.$$

3. Location Mixture Distribution

Suppose X has an arbitrary distribution with PDF $g(x;\tau)$, where $\tau \in U \subset \mathbb{R}^k$ is a scalar or vector parameter indexing the distribution, and W, independently of X, has some distribution on $(0,\infty)$ with CDF $H(w;\nu)$, $\nu \in V \subset \mathbb{R}^k$. Then, we say that $Y = \delta W + \lambda X$ is a location mixture of X with a location mixing (CDF) $H(w;\nu)$, for $\delta \in \mathbb{R}$, if its PDF is given by

$$f(y;\delta,\lambda,\tau,\nu) = \frac{1}{\lambda}\int_0^\infty g\left(\frac{y-\delta w}{\lambda};\tau\right)\, dH(w;\nu).$$

It is important to mention that a stochastic representation for the skew-normal random variable Y, with parameter $\lambda \in \mathbb{R}$, has been presented by Azzalini (1985) and Henze (1986) as follows:

$$Y \overset{d}{=} \delta W + \sqrt{1-\delta^2}\, Z,$$

where W and Z are independent half-normal and standard normal random variables, respectively. This representation means that the skew-normal distribution is a mean-mixture of normal distribution in which the mixing distribution is the half-normal distribution. Convolution of normal and a positive random variable has been considered by many authors including Silver et al. (2009), and Krupskii et al. (2018). Recently, Negarestani et al. (2018) introduced a general class of skewed distributions based on the mean mixture of normal distribution, which motivates the present work. We introduce a location mixture of Laplace and asymmetric Laplace distributions and discuss their densities and properties. For this purpose, we first need to review the location mixture of the exponential distribution, which subsequently enables the definition of a new family of weighted distributions.

3.1. Location Mixture of Exponential Distribution.

Definition 3.1. Let X and W be independent exponential random variables with mean 1. Then, for $\delta > 0$, the random variable $Y = \delta W + \lambda X$ has a location mixture of exponential distribution with PDF

$$f(y;\delta,\lambda) = \lambda^* f(y;\lambda)\, F\left(\frac{y}{\delta};\lambda^*\right), \qquad y > 0, \tag{3}$$

where $\lambda^* = \frac{\lambda}{\lambda-\delta}$, and $f(.)$ and $F(.)$ are the PDF and CDF of a standard exponential random variable, respectively.

Remark 3.2. The weighted exponential distribution introduced by Gupta and Kundo (2009) is a location mixture of the exponential distribution upon taking $\delta = \frac{\lambda}{\theta+1}$ in (3), where denoted by $WE(\theta,\lambda)$.

The weight function in (3) is defined in terms of the distribution function. Instead, we now change this structure and express it in terms of some other functions. Specifically, in the following theorem we propose a new family of distributions, denoted as WES distribution. Then, in the next sections, we will show that truncated LM-AL belongs to this family of distributions.

Theorem 3.3. *Let $X_1 \sim SL(\lambda_1,\tau)$ be as defined in (2) and $X_2 \sim Exp(\lambda_2)$ be independent of X_1. Then, for $\theta > 0$, the PDF of $Y \overset{d}{=} X_2|(0 < X_1 \leq \theta X_2)$ is given by*

$$f(y;\lambda_1,\lambda_2,\theta,\tau) = c\, f(y;\lambda_2)\, \left(1 - \frac{1}{1+\tau} S(\theta y;\lambda_1)\right), \tag{4}$$

where $c = \left(1 + \frac{\lambda_1}{\tau\lambda_1+\theta\lambda_2(1+\tau)}\right)$, $f(.)$ and $S(.)$ are the PDF and survival function of the exponential random variable, respectively.

Proposition 3.4. *The random variable Y, with PDF in (4), is said to have the weighted exponential distribution in terms of the survival function, with parameters $\lambda_1,\lambda_2,\theta,\tau \in \mathbb{R}^+$. We will denote it by $Y \sim WES(\lambda_1,\lambda_2,\theta,\tau)$.*

Remark 3.5. If $Y \sim WES(\lambda_1,\lambda_2,\theta,\tau)$, then

- The $Exp(\lambda_2)$ distribution is a limiting case of the $WES(\lambda_1,\lambda_2,\theta,\tau)$, when $\theta \to 0$ or $\theta \to \infty$.
- The weighted exponential distribution introduced by Gupta and Kundo (2009) is belong to the family of WES$(\lambda_1,\lambda_2,\theta,\tau)$ distributions, when $\lambda_1 = \lambda_2 = \lambda$ and $\tau = 0$.
- The location mixture of exponential distribution in (3) is a special case of the $WES(\lambda_1,\lambda_2,\theta,\tau)$, when $\tau = 0$, $\lambda = \lambda_2$, and $\delta = \frac{\lambda_1\lambda_2}{\theta\lambda_2+\lambda_1}$.

So far, we have defined a location mixture of exponential distributions and used it to introduce a new family of weighted distributions in terms of the survival function. Now, we will use it to get a location mixture of asymmetric Laplace distribution.

3.2. **Location Mixture of Asymmetric Laplace Distribution.** Let $X \sim AL(\beta)$, defined in (1), and W, independently of X, has some distribution on $(0,\infty)$ with CDF $H(w;\nu)$. Then, we say that $Y = \delta W + \lambda X$ has a location mixture of asymmetric Laplace distribution with a location mixing (CDF) H(w;ν) for $\delta \in \mathbb{R}$ and $\lambda > 0$, denoted by $Y \sim LM - AL(\beta,\delta,\lambda)$, if its PDF is given by

$$f(y;\beta,\lambda,\delta,\nu) = \frac{1}{2\alpha\lambda}\int_0^\infty e^{-\frac{\alpha}{\lambda}|y-\delta w|+\frac{\beta}{\lambda}(y-\delta w)}\, dH(w;\nu). \tag{5}$$

The above integral has a closed form for some expression mixing distributions, such as exponential and Lindley distributions. We consider here the distributions of W to be exponential with mean 1. After some algebra, the PDF of Y is obtained as

$$f(y;\beta,\lambda,\delta) = \frac{1}{2\alpha(\lambda+\beta\delta+\alpha|\delta|)}\begin{cases} e^{(\frac{sign(\delta)\alpha+\beta}{\lambda})y}, & \delta y < 0 \\ \frac{(\lambda+\beta\delta+\alpha|\delta|)e^{-(\frac{sign(\delta)\alpha-\beta}{\lambda})y}-2|\delta|\alpha e^{-\frac{y}{\delta}}}{(\lambda-\alpha|\delta|+\beta\delta)}, & \delta y \geq 0, \end{cases} \tag{6}$$

where $sign(.)$ denotes the sign function. To eliminate the dependence between parameters, we set $\delta = \frac{\lambda}{sign(\delta)\alpha+\beta}$. Subsequently, the PDF of location mixture of asymmetric Laplace distribution, denoted by $LM - AL(\beta,\lambda)$, can be expressed as

$$f(y;\beta,\lambda) = \frac{1}{4\alpha\lambda}\begin{cases} e^{(\frac{sign(\delta)\alpha+\beta}{\lambda})y}, & \delta y < 0 \\ \frac{(\alpha+sign(\delta)\beta)e^{-(\frac{sign(\delta)\alpha-\beta}{\lambda})y}-\alpha e^{-(\frac{sign(\delta)\alpha+\beta}{\lambda})y}}{sign(\delta)\beta}, & \delta y \geq 0. \end{cases} \tag{7}$$

Proposition 3.6. *If $Y \sim LM - AL(\beta,\lambda)$ with PDF as in (7), the following stochastic representations hold:*

(I) *If W_1, W_2, and W_3 are three independent random variables such that $W_1 \sim Exp(\alpha - sign(\delta)\beta)$ and $W_2, W_3 \sim Exp(\alpha+\beta)$, then a representation of Y is*

$$Y \stackrel{d}{=} \lambda(W_1 + sign(\delta)W_2 - (\alpha-\beta)^2 W_3) \sim LM - AL(\beta,\lambda);$$

(II) *If $W_1 \sim Exp(\alpha - sign(\delta)\beta)$ and $W_2 \sim \chi^2_{(2)}$ is independent of W_1, then Y has the stochastic representation*

$$Y \stackrel{d}{=} \lambda(sign(\delta)W_1 + \beta W_2 + \sqrt{W_2}Z) \sim LM - AL(\beta,\lambda);$$

(III) *If N is a discrete random variable taking values $\beta - \alpha$ and $\beta + \alpha$ with probability $\frac{\alpha-\beta}{2\alpha}$ and $\frac{\alpha+\beta}{2\alpha}$, respectively. Also, let $W_1, W_2 \sim Exp(1)$ be independent of N. Then*

$$Y \stackrel{d}{=} \lambda\left(\frac{W_1}{\beta + sign(\delta)\alpha} + NW_2\right) \sim LM - AL(\beta,\lambda),$$

where in all of these $\alpha^2 - \beta^2 = 1$ and $\delta = \frac{\lambda}{sign(\delta)\alpha+\beta}$.

Proposition 3.7. *Let $Y \sim LM-AL(\beta,\lambda)$. Then, the skewness and kurtosis of Y can be obtained as follows*

$$\gamma_y = \frac{\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3}{(\mu_2-\mu_1^2)^{1.5}}, \qquad \kappa_y = \frac{\mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4}{(\mu_2-\mu_1^2)^2} - 3$$

where

$$\begin{aligned}
\mu_1 &= E(Y) = \lambda(\beta + sign(\delta)\alpha),\\
\mu_2 &= E(Y^2) = 4\lambda^2(\alpha^2+\beta^2),\\
\mu_3 &= E(Y^3) = 12\lambda^3(\alpha^2+\beta^2)(\beta+sign(\delta)\alpha),\\
\mu_4 &= E(Y^4) = 24\lambda^4(3\alpha^4+10\alpha^2\beta^2+3\beta^4).
\end{aligned}$$

To show the flexibility and supremacy of the LM-AL model in covering the skewness and kurtosis of the data, we computed the maximum range of skewness and kurtosis of the LM-AL distribution. The obtained findings indicate that the range of skewness (-2,2) and kurtosis (5,9) for the LM-AL is wider than the other skewed distributions introduced in the literature including skew-t normal (Gómez, Venegas, and Bolfarine (2007)).

Remark 3.8. For $\beta = 0$, $Y \sim LM-AL(\beta,\lambda)$ reduces to the location mixture of Laplace distribution, denoted by $LM-L(\lambda)$, if and only if

$$Y \stackrel{d}{=} \begin{cases} sign(-\delta)X_1, & with\ probability\ \frac{1}{4},\\ sign(\delta)X_2, & with\ probability\ \frac{3}{4}, \end{cases}$$

where $sign(-\delta)X_1 \stackrel{d}{=} sign(\delta)X_2 \sim Exp(\lambda)$.

4. Parameter estimation with ECM algorithm

In this section, we propose an ECM algorithm for determining the maximum likelihood estimate $\Theta = (\beta,\lambda)$. The ECM algorithm replaces a complicated M-step of the EM algorithm with several computationally simpler conditional maximization (CM) steps. A CM-step might be in closed form or it might itself require iteration, but because the CM maximizations are over smaller dimensional spaces, often they are faster, simpler, and more stable than the corresponding full maximization required in the M-step of the EM algorithm, especially when iteration is required (McLachlan et al., 2004). In the following theorem, we present the conditional distribution of *W*, given *Y* = y, which are required in the E-step of the ECM algorithm.

Theorem 4.1. *Let $Y = \delta W + \lambda X$, where $W \sim Exp(1)$, and $X \sim AL(\beta)$, is independently of W. Then, the conditional PDF of W, given Y = y, for $y \in \mathbb{R}$, is as follows:*

(a) *For $\delta y < 0$,*

$$W|(Y=y) \sim Exp\left(\frac{1}{2}\right);$$

(b) *For* $\delta y \geq 0$,

$$f_{W|Y=y}(w;\beta,\lambda) = c \begin{cases} e^{-\frac{sign(\delta)\alpha-\beta}{\lambda}(y+2\beta\lambda w)}, & 0 \leq w < \frac{\beta+sign(\delta)\alpha}{\lambda}y, \\ e^{\frac{sign(\delta)\alpha+\beta}{\lambda}(y+2\lambda w(\beta-sign(\delta)\alpha))}, & w \geq \frac{\beta+sign(\delta)\alpha}{\lambda}y, \end{cases}$$

where $c = \frac{2sign(\delta)\beta}{(\alpha+sign(\delta)\beta)e^{-(\frac{sign(\delta)\alpha-\beta}{\lambda})y}-\alpha e^{-(\frac{sign(\delta)\alpha+\beta}{\lambda})y}}$ *is the normalizing constant.*

Now, let $Y_1, Y_2, ..., Y_n$ be a random of size n from the LM-AL distribution. Then, we use the following hierarchical representation for Y_i:

(8) $$Y_i|(W_i = w_i) \sim AL(\delta w_i, \lambda, \beta) \quad and \quad W_i \sim Exp(1), \quad i = 1,2,...,n,$$

where $\delta = \frac{\lambda}{sign(\delta)\alpha+\beta}$. Based on the hierarchical representation in (8), we consider random vector $\boldsymbol{x} = (\boldsymbol{y}, \boldsymbol{w})^T$, where $\boldsymbol{y} = (y_1, y_2, ..., y_n)^T$ is a random variable associated with the observed data and $\boldsymbol{w} = (w_1, ..., w_n)^T$ is the latent random vector. The complete data log-likelihood is then given by

(9)
$$l_c(\Theta|\boldsymbol{y},\boldsymbol{w}) = -n\log(2\alpha\lambda) + \frac{\beta}{\lambda}\sum_{i=1}^{n} y_i - \left(\frac{2\beta}{sign(\delta)\alpha+\beta}\right)\sum_{i=1}^{n} w_i - \frac{\alpha}{\lambda}\sum_{i=1}^{n}\left|y_i - \frac{\lambda}{sign(\delta)\alpha+\beta}w_i\right|.$$

As w_i are unobservable, we need to compute the conditional expectation of the log-likelihood function of complete data and then replace $g(w_i)$ with its conditional expectation, given y_i. Therefore, the ECM algorithm for ML estimation of the LM-AL distribution can be summarized through the following two iterative steps.

• E-step: At the iteration k, we compute the expected-log-likelihood function as

$$Q(\Theta|\hat{\Theta}^{(k)}) = -n\log(2\alpha\lambda) + \frac{\beta}{\lambda}\sum_{i=1}^{n} y_i - \left(\frac{2\beta}{sign(\delta)\alpha+\beta}\right)\sum_{i=1}^{n}\psi_{1i}^{(k)} - \frac{\alpha}{\lambda}\sum_{i=1}^{n}\psi_{2i}^{(k)},$$

where $\psi_{1i}^{(k)} = E(W_i \mid y_i, \hat{\theta}^{(k)})$ and $\psi_{2i}^{(k)} = E(\left|y_i - \frac{\lambda}{sign(\delta)\alpha+\beta}W_i\right| \mid y_i, \hat{\theta}^{(k)})$.

• CM-step: Maximizing $Q(\Theta|\hat{\Theta}^{(k)})$ over β and λ is done as follows:

$$\hat{\lambda}^{(k+1)} = \frac{\hat{\alpha}^{(k)}\sum_{i=1}^{n}\psi_{2i}^{(k)} - \hat{\beta}^{(k)}\sum_{i=1}^{n} y_i}{n},$$

ana then by fixing $\hat{\lambda}^{(k)} = \hat{\lambda}^{(k+1)}$, we have

$$\hat{\beta}^{(k+1)} = arg\max_{\beta}\ Q(\Theta|\hat{\Theta}^{(k)}).$$

The above procedure is iterated until convergence of $l(\hat{\Theta}^{(k+1)}) - l(\hat{\Theta}^{(k)}) < \upsilon$, where the tolerance υ is considered as 10^{-5}.

In order to provide a detailed guidance of the implementation method, a pseudocode of the above ECM algorithm is summarized in Algorithm 1.

The performance of the estimation method proposed in this section is examined through a simulation study in Section 6.

Algorithm 1 Implementation procedure of the ECM algorithm for fitting the LM-AL.

procedure LM-AL($\mathbf{y}$, $\mathbf{x}$)
  inputs: $\{y_i, \mathbf{x}_i\}_{i=1}^{n}$ - the set of input data;
  $k_{\max}$ = 10000 - the maximum number of iteration; $\upsilon = 10^{-5}$- the prespecified tolerance.
  initialize: Set the algorithm counter to k=0.
  Obtain the initial value of the parameter vector of $\theta^{(0)}$.
  Compute the initial log-likelihood as $\ell_{obs}^{(k)} = \ell(\hat{\theta}^{(k)} \mid \mathbf{y})$.
  $\mathbf{E-step}$:
  for $k = 1$ to k_{max} **do**
    for $i = 0$ to n **do**
      $\psi_{1i}^{(k)} = E(W_i \mid y_i, \hat{\theta}^{(k)})$ and $\psi_{2i}^{(k)} = E\left(\left|y_i - \frac{\lambda}{sign(\delta)\alpha+\beta} W_i\right| \mid y_i, \hat{\theta}^{(k)}\right)$.
      Compute the conditional expectation, at the $k-th$ iteration, as

$$Q(\Theta|\hat{\Theta}^{(k)}) = -n\log(2\alpha\lambda) + \frac{\beta}{\lambda}\sum_{i=1}^{n} y_i - \left(\frac{2\beta}{sign(\delta)\alpha+\beta}\right)\sum_{i=1}^{n}\psi_{1i}^{(k)} - \frac{\alpha}{\lambda}\sum_{i=1}^{n}\psi_{2i}^{(k)}.$$

    end for
    $\mathbf{CM-steps}$:
    $\hat{\lambda}^{(k+1)} = \frac{\hat{\alpha}^{(k)}\sum_{i=1}^{n}\psi_{2i}^{(k)} - \hat{\beta}^{(k)}\sum_{i=1}^{n} y_i}{n}$;
    Fix $\lambda^{(k)} = \hat{\lambda}^{(k+1)}$ and update $\hat{\beta}^{(k)}$ by maximizing $Q(\Theta|\hat{\Theta}^{(k)})$ over β,

$$\hat{\beta}^{(k+1)} = arg\max_{\beta}\ Q(\Theta|\hat{\Theta}^{(k)});$$

    Update k= (k+1) and $\Theta^{(k)} = \Theta^{(k+1)}$.
  end for
  Repeat the above steps until convergence of $l(\hat{\Theta}^{(k+1)}) - l(\hat{\Theta}^{(k)}) < \upsilon$.
end procedure

5. Application of the LM-AL family in survival analysis

Survival analysis is one of the most important and widely used concepts defined in the statistical literature that is used in many fields such as medical sciences, finance and reliability. Accelerated failure time models (AFT) are an important family in survival analysis. Let T is a positive random variable representing time of the occurrence of an event. The AFT model can be defined based on the relationship between survival function and a set of covariates as follows

$$S(t|\mathbf{X}) = S_0(te^{-\mathbf{X}^T\gamma}),$$

where $S_0(t)$ is the baseline survival and $\gamma = (\gamma_1, \gamma_2, ..., \gamma_p)^T$ and $\mathbf{X}^T = (X_1, X_2, ..., X_p)$ are a vector of regression coefficient and a column vector of the covariates, respectively. The relationship between covariates and the survival time can be defined as a linear relation between the logarithm of survival time and a set of covariates, that is

$$Y = \ln(T) = \mathbf{X}^T\gamma + \sigma\varepsilon,$$

where $\sigma > 0$ is an unknown scale parameter and the error term ε is a random variable that follows a certain parametric distribution. There are several types of parametric models that conform to the AFT model, such as exponential, Weibull, log-normal, log-logistic, gamma and generalized gamma models.

In this section, we choose a model based on the LM-AL distribution, when $\delta > 0$. Because survival analysis focuses on techniques based on positive random variables, we consider the case when the LM-AL distribution introduced in Section 3 is truncated on the left at zero.

Proposition 5.1. *Let $Y \sim LM-AL(\beta,\lambda)$ with PDF defined in (7), then the random variable $T = Y|Y > 0$ is said to be distributed as truncated location mixture of AL distribution at zero, if its PDF has the following form:*

$$f(t;\alpha,\lambda) = \frac{(\alpha+\beta)e^{-\frac{(\alpha-\beta)}{\lambda}t} - \alpha e^{-\frac{(\alpha+\beta)}{\lambda}t}}{\lambda\beta(3\alpha+\beta)}, \quad t > 0. \tag{10}$$

Remark 5.2. The PDF defined in (10) belongs to the family of weighted exponential distribution in terms of the survival function, i.e.,

$$T \sim WES(\lambda, \frac{\lambda}{\alpha-\beta}, 2\beta, \frac{\beta}{\alpha}).$$

5.1. **Parameter estimation of WES AFT model with MCECM algorithm.** The EM algorithm is an iterative method to find maximum likelihood estimates of parameters through the recursive implementation of the E- step (expectation) and the M-step (maximization). The advantage of the EM algorithm is to exploit the simplicity of the likelihood function. But sometimes, due to the complexity in calculating the Q function, it is suggested that the Q function be estimated by the Monte Carlo (MC) method, to avoid the analytical complexity in the E-step. Let $(\boldsymbol{T},\boldsymbol{Z})$ be an independent random sample of size n from a $WES(\lambda, \frac{\lambda}{\alpha-\beta}, 2\beta, \frac{\beta}{\alpha})$ distribution with PDF in (10), where $\boldsymbol{T} = (t_1,t_2,...,t_m)^T$ denote the observed survival times (uncensored) and $\boldsymbol{Z} = (z_{m+1},...,z_n)^T$ is the missing (censored) part with $z_i > R_i$. Then, the complete data log-likelihood for the parameter vector $\boldsymbol{\Psi} = (\Theta^T,\gamma^T)^T$, such that $\Theta = (\beta,\sigma,\lambda)^T$

and $\gamma = (\gamma_1, \gamma_2, ..., \gamma_p)^T$, can be written as

(11)

$$\begin{aligned}
Q(\Psi|\Psi^{(k)}) &= E_{\Psi^{(k)}}(\log L_c(\Psi)|t,z) \\
&\propto -n\log(\sigma\lambda\beta(\beta+3\alpha)) - \frac{1}{\sigma}\sum_{i=1}^{n} X_i^T\gamma \\
&+ \frac{1}{\sigma}\left(\sum_{i=1}^{m}\log(t_i) + \sum_{i=m+1}^{n}\int_{R_i}^{\infty}\log(z_i)p(z_i|\Psi^{(k)})dz_i\right) \\
&+ \sum_{i=1}^{m}\log((\alpha+\beta)e^{-\frac{(\alpha-\beta)e^{\frac{\log(t_i)-X_i^T\gamma}{\sigma}}}{\lambda}} - \alpha e^{-\frac{(\alpha+\beta)e^{\frac{\log(t_i)-X_i^T\gamma}{\sigma}}}{\lambda}}) \\
&= -n\log(\sigma\lambda\beta(\beta+3\alpha)) - \frac{1}{\sigma}\sum_{i=1}^{n} X_i^T\gamma + \frac{1}{\sigma}\left(\sum_{i=1}^{m}\log(t_i) + \frac{1}{M}\sum_{j=1}^{M}\sum_{i=m+1}^{n}\log(z_{i,j})\right) \\
&+ \sum_{i=1}^{m}\log((\alpha+\beta)e^{-\frac{(\alpha-\beta)e^{\frac{\log(t_i)-X_i^T\gamma}{\sigma}}}{\lambda}} - \alpha e^{-\frac{(\alpha+\beta)e^{\frac{\log(t_i)-X_i^T\gamma}{\sigma}}}{\lambda}}),
\end{aligned}$$

where $\boldsymbol{z}^{(j)} = (z_{m+1,j}, ..., z_{n,j})$, for $j = 1,..,M$ are generated from $p(z|\Psi^{(k)})$. By derivation of (11) with respect to the parameters of model and then equating the results to zero, we get the systems of equations which cannot be solved directly. Therefore, the maximum likelihood estimator of the parameters have not an explicit form and must be solved numerically using an iterative procedure. Hence, the M-step can be summarized through the following two steps:

$$\gamma^{(k+1)} = arg\max_{\gamma}\ Q(\Theta,\gamma|\Theta^{(k)},\gamma^{(k)}),$$

$$\Theta^{(k+1)} = arg\max_{\Theta}\ Q(\Theta,\gamma|\Theta^{(k)},\gamma^{(k+1)}).$$

6. **Empirical evaluation**

In this section, a Monte Carlo simulation study is carried out for evaluating the performance of the estimation algorithm under two scenarios. In addition, two real data sets are analyzed to demonstrate the suitability of the proposed distribution in data analysis.

6.1. **Simulation study.** In this section, a simulation study is done for evaluating the performance of the estimation method based on the ECM algorithm proposed in Section 4. For this purpose, we use the following simulation scheme for two different levels of skewness in the LM-AL distribution. Firstly, we generate a random sample with sample size n =50, 200, and 500 from LM-AL distribution with parameters (1) $\beta = 0.06$ and $\lambda = 1$; and (2) $\beta = 1.73, \lambda = 0.1$, corresponding to the values of the skewness $\gamma_y = 0.50$ (low), and $\gamma_y = 1.99$ (high), respectively. For this combination of parameter values and sample sizes, we generated N=5000 independent data sets. In each iteration, the ML estimates are determined by using the ECM algorithm. Then, for evaluating the performance of the ML estimates obtained, we computed absolute bias (*AB*), variance, and root-mean-square error (*RMSE*). For each parameter θ, *AB* and *RMSE* are defined as

$$AB = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{\theta}_i - \theta\right| \quad and \quad RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta)^2}\,,$$

where $\hat{\theta}_i$ are the parameter estimates of β and λ obtained from the i-th replication. The corresponding results of this simulation are presented in Table 1, while Figure 2 shows the histogram of the simulated data with sample size $n = 500$ and the PDF of LM-AL distribution for the estimated values.

TABLE 1. Simulation results for the LM-AL distribution.

Par.		$\beta = 0.06, \lambda = 1$		$\beta = 1.73, \lambda = 0.1$	
	n	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\lambda}$
AB	50	0.0146	0.0135	0.4768	0.0109
	200	0.0038	0.0053	0.0925	0.0026
	500	0.0017	0.0027	0.0333	0.0011
Var	50	0.0010	0.0152	0.7867	0.0008
	200	0.0001	0.0035	0.1114	0.0002
	500	0.0000	0.0014	0.0294	0.0001
RMSE	50	0.0353	0.1241	1.0069	0.0310
	200	0.0093	0.0597	0.3463	0.0149
	500	0.0047	0.0374	0.1746	0.0090

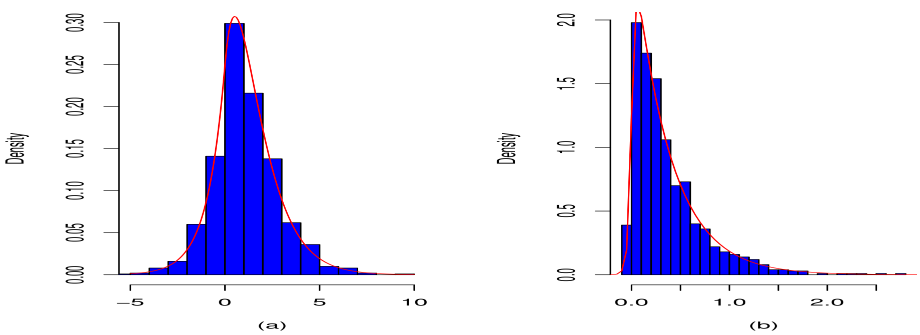

FIGURE 2. The PDF of LM-AL distribution for the estimated values based on tha different levels of skewness $\gamma_y = 0.50$ (a) and $\gamma_y = 1.99$ (b).

TABLE 2. Simulation results for the WES distribution.

Model		WES(1,3.7,3.4,1.16)					WES(0.7,1.8,2.2,0.74)				
Par.	n	γ_0	γ_1	γ_2	β	λ	γ_0	γ_1	γ_2	β	λ
AB	100	0.0248	0.0824	0.1087	0.4530	0.2426	0.0452	0.0660	0.1342	0.2361	0.1813
	500	0.0163	0.0651	0.1049	0.1933	0.1549	0.0131	0.0763	0.1070	0.0952	0.0855
	1000	0.0088	0.0147	0.0706	0.1291	0.0511	0.0082	0.0411	0.0952	0.0281	0.0323
Var	100	0.1264	0.0632	0.0166	0.3566	0.1905	0.1211	0.0569	0.0172	0.1814	0.0704
	500	0.0251	0.01331	0.0032	0.1052	0.0869	0.0271	0.0137	0.0032	0.0758	0.0132
	1000	0.0123	0.0067	0.0016	0.0618	0.05621	0.0022	0.0071	0.0023	0.0068	0.0086
RMSE	100	0.3556	0.2645	0.1688	0.7164	0.4395	0.2571	0.3372	0.1973	0.4258	0.3011
	500	0.1093	0.1077	0.0996	0.1596	0.1331	0.0692	0.0829	0.0668	0.17034	0.0838
	1000	0.0712	0.0781	0.0542	0.0873	0.07199	0.0318	0.0573	0.0218	0.0861	0.0571

In the second scenario of the simulation study, we evaluate the performance of the parameter estimation method in accelerated failure time models. For this purpose, we consider the following two scenarios. Time-to-event data are generated from the following model:

$$\log(T) = \gamma_0 + \gamma_1 X_1 + \gamma_2 X_2 + \varepsilon,$$

where $(\gamma_0, \gamma_1, \gamma_2) = (0.1, 0.7, 0.8)$ and covariates X_1 and X_2 are independent random variables, which are simulated from Bernoulli distribution with success probability 0.75 and standard normal distribution, respectively. We censor the observations T_i that are greater than 14. This mechanism generates samples with 10%–30% censored observations. For the distribution of the error terms ε, we suppose that ε is a random variable such that $\log(\varepsilon) \sim WES$ with PDF as in (10) with parameters; (1) $\beta = 1.1$ and $\lambda = 0.7$, and (2) $\beta = 1.7$ and $\lambda = 1$. Table 2 presents the corresponding results of the WES AFT model. We observe the following from these two simulation studies:

1) As the sample size increases, the mean of the estimated value approaches the true value of the parameter;
2) AB and RMSE both decrease with increasing sample size.

6.2. **Data analysis.** In this subsection, we analyze two real data sets using the developed model. We then use Akaike information criteria (AIC) and Bayesian information criteria (BIC) to compare the fitted models.

6.2.1. ***Example 1: Bladder Cancer Data.*** As the first example, we consider the set of data reported by Shanker et al. (2015). This data set consists of remission times (in months) of a random sample of 128 bladder cancer patients. To evaluate the performance of the proposed model, we fitted several models to this set of data. Table 3 presents the obtained results, including the MLEs of model parameters and Akaike information criterion (AIC) and Bayesian information criteria (BIC) values for the models. The graphical fitness of the WES model is displayed in Figure 3. It can be seen from Table 3, that the WES model has lower AIC and BIC values than all other models. Thus, the WES model provides a better fit for this bladder cancer data.

TABLE 3. Bladder Cancer data: Maximum likelihood estimates and AIC and BIC values.

Model	Lindley	Log-normal	Weibull	Gamma	WE	WES
β	-	-	1.0514	1.1851	11.5099	1.0545
λ	0.1991	-	0.1062	0.1287	0.1173	3.3800
μ	-	1.7423	-	-	-	-
σ	-	1.0646	-	-	-	-
AIC	835.7925	829.3131	827.7849	826.1691	825.4546	823.9858
BIC	838.6445	835.0171	833.489	831.8732	831.1586	829.6898

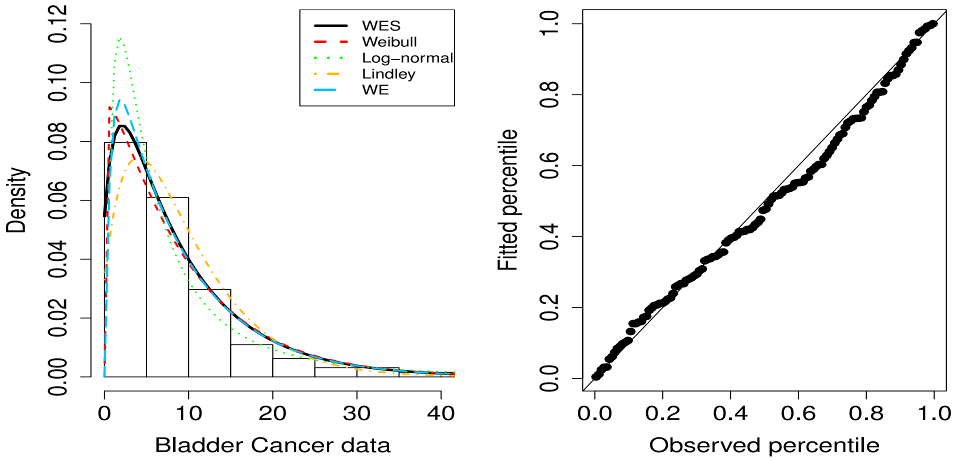

FIGURE 3. Histogram of the Bladder Cancer data and fitted WES, WE, Weibull, Log-normal and Lindley models (left figure) and P-P plot based on the fitted WES model (right figure).

6.2.2. ***Example 2: Monoclonal Gammopathy Data***. As the second example, we consider the monoclonal gammopathy of undetermined significance (MGUS) data, available in the survival package of the R software. This data set contains information on the survival time and a few known predictors of progression of 1384 patients in a study collected at Mayo Clinic between 1994 and 1999. The survival times (futime) indicate the time until death or last contact in months, together with an indicator variable (death) associated with the status of the patient at the end of the study (0,1 for censored and dead). This data set has been analyzed by some other authors, including Castaneda-Avila et al. (2021) and Epstein et al. (2019). We now fit the WES model and compare the results with these of some other prominent lifetime models. Specifically, we fit an AFT model with five covariates “age” (in years),“sex” (Female=1, Male=2), “hgb” (hemoglobin), “creat” (creatinine), “mspike” (size of the monoclonal serum splike) as well as an intercept. Table 4 presents the ML estimates of the parameters of each model with the corresponding AIC and BIC values. We observe from Table 4 that the estimates of the regression parameters are all quite close for different models. However, the AIC and BIC values reveal that the proposed model performs better than other models considered.

TABLE 4. MGUS data: Maximum likelihood estimates and AIC and BIC values.

Model	Log-normal	Log-logistic	WE	Gamma	Weibull	WES
Intercept	5.2362	6.1036	6.5123	6.5306	6.9845	7.1053
Age	-0.0460	-0.0468	-0.0482	-0.0479	-0.0476	-0.0474
Sex	-0.5402	-0.4743	-0.4474	-0.4108	-0.4025	-0.3865
Hgb	0.2166	0.1658	0.1438	0.1241	0.1189	0.1099
Creat	-0.1340	-0.1702	-0.0677	-0.0463	-0.0490	-0.0507
Mspike	0.0169	-0.0147	-0.0262	-0.0376	-0.0356	-0.0394
β	-	-	0.0100	1.0105	1.1167	0.1417
λ	-	-	1.3921	0.1280	0.2352	0.7611
σ	1.4877	0.7786	0.7010	0.8786	0.9560	1.1298
AIC	10928.73	10837.01	10748.87	10697.09	10694.54	10680.55
BIC	10970.32	10878.6	10795.66	10738.95	10736.13	10727.34

From Table 4, the AFT regression model for this data set is given by

$$\begin{aligned} log(T_j) =& 7.1053 - 0.0474 \times Age_j - 0.3865 \times Sex_j + 0.1099 \times Hgb_j \\ & - 0.0507 \times Creat_j - 0.0394 \times Mspike_j + 1.1298 \times \varepsilon_j, \end{aligned}$$

where T_j represents the survival time for j-th individual and ε_j is an error variable such that $log(\varepsilon_j) \sim WES(0.761, 0.877, 0.283, 0.14)$ with PDF as in (10). The coefficient $\hat{\gamma}$ in Table 4 expresses the impact of a unit change in predictors on the survival time, provided that all other variables are held constant so that one unit change in predictor X_j changes the survival time by $exp(\hat{\gamma}_j)$. For example, from Table 4, it can be seen that a unit increase in patient age reduces survival time by 4.63%. Also, the survival time for males is 32.06% shorter than for females, and an increase in the hemoglobin level increases patient survival time by 11.61%.

7. Concluding remarks

In this paper, we have introduced a new family of skewed distributions based on the location mixture of asymmetric Laplace (LM-AL) distribution. The proposed model is long-tailed and has a wider range of skewness and kurtosis than many other known distributions. This distribution is especially useful for modeling unimodal and long-tailed data. The usefulness and practicality of the proposed model and the developed estimation method are demonstrated through detailed Monte Carlo simulations as well as two real data sets.

Appendix A

We provide here the proofs of some theorems and also details about the derivation of the PDF of the LM-AL distribution.

A.1. Proof of Proposition 2.2.

Proof. As the support of the variable Y is all real numbers, it suffices to show that $Pr(X_2 - \tau X_1 \leq 0) \neq \frac{1}{2}$. We specifically have

$$Pr(X_2 - \tau X_1 \leq 0) = \int_0^\infty Pr(X_2 - \tau X_1 \leq 0 | X_1 = x)\ f_{X_1}(x)\ dx = \int_0^\infty F_{X_2}(\tau x)\ f_{X_1}(x)\ dx,$$

where F_X and f_X are the cumulative distribution function (CDF) and probability density function (PDF) of the exponential random variable with mean λ, respectively. Upon substituting $f_X(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}}$ and $F_X(\tau x) = 1 - e^{-\frac{\tau x}{\lambda}}$ in the above expression and then carrying out the integration, we find

$$Pr(X_2 - \tau X_1 \leq 0) = \frac{\tau}{1+\tau},$$

It is then evident that if $\tau \neq 1$, then $Pr(X_2 - \tau X_1 \leq 0) \neq \frac{1}{2}$. Also, the PDF of Y is readily obtained as

$$F_Y(y) = Pr(X_2 - \tau X_1 \leq y) = \int_0^\infty Pr(X_2 - \tau X_1 \leq y | X_1 = x)\ f_{X_1}(x)\ dx = \int_0^\infty F_{X_2}(y + \tau x)\ f_{X_1}(x)\ dx,$$

from which we then obtain the following:

1) If $y < 0$, then $x \geq \frac{-y}{\tau}$, and

$$f_Y(y) = \int_0^\infty f_{X_2}(y + \tau x) f_{X_1}(x)\ dx = \frac{e^{-\frac{y}{\lambda}}}{\lambda^2} \int_{-\frac{y}{\tau}}^\infty e^{-\frac{(\tau+1)}{\lambda}x}\ dx = \frac{e^{\frac{y}{\lambda\tau}}}{\lambda(\tau+1)};$$

2) If $y \geq 0$, then $x \geq 0$, and

$$f_Y(y) = \int_0^\infty f_{X_2}(y + \tau x) f_{X_1}(x)\ dx = \frac{e^{-\frac{y}{\lambda}}}{\lambda^2} \int_0^\infty e^{-\frac{(\tau+1)}{\lambda}x}\ dx = \frac{e^{-\frac{y}{\lambda}}}{\lambda(\tau+1)}.$$

Hence the theorem holds. □

A.2. Proof of Theorem 3.3**.**

Proof. Based on the definition of conditional probability, we have

$$F_Y(y;\lambda_1,\lambda_2,\theta,\gamma) = Pr(X_2 \le y | 0 < X_1 \le \theta X_2) = \frac{Pr(X_2 \le y, 0 < X_1 \le \theta X_2)}{Pr(0 < X_1 \le \theta X_2)},$$

It can then be easily shown that

$$f_Y(y;\lambda_1,\lambda_2,\theta,\gamma) = f_{X_2}(y;\lambda_2)\frac{Pr(0 < X_1 \le \theta X_2 | X_2 = y)}{Pr(0 < X_1 \le \theta X_2)}, \tag{12}$$

As $\gamma \ge 0$, we can write

$$Pr(0 < X_1 \le \theta X_2 | X_2 = y) = 1 - \frac{1}{1+\gamma} e^{-\frac{\theta y}{\lambda}}$$

and

$$Pr(0<X_1 \le \theta X_2) = \int_0^\infty Pr(0<X_1 \le \theta X_2|X_2=y)\, f_{X_2}(y;\lambda_2)\, dy = \frac{\lambda_1\gamma + \theta(1+\gamma)\lambda_2}{(\lambda_1+\lambda_2\theta)(1+\gamma)}.$$

By using the above expressions in (12), the proof gets completed. □

A.3. PDF of LM-AL distribution in (6)**.**

Proof. As the proofs for the cases $\delta \ge 0$ and $\delta < 0$ are the same, we only consider here the case $\delta \ge 0$. According to the definition of the PDF of LM-AL in (5), we have:

(1) For y¡0,

$$\begin{aligned} f_Y(y;\beta,\lambda) =& \frac{1}{2\alpha\lambda}\int_0^\infty e^{-\frac{\alpha}{\lambda}|y-\delta w| + \frac{\beta}{\lambda}(y-\delta w)} e^{-w} dw \\ =& \frac{e^{\frac{\alpha+\beta}{\lambda}y}}{2\alpha\lambda}\int_0^\infty e^{-(\frac{(\alpha+\beta)\delta+\lambda}{\lambda})w} dw = \frac{e^{\frac{\beta+\alpha}{\lambda}y}}{2\alpha(\alpha\delta+\beta\delta+\lambda)}, \end{aligned}$$

(2) For $y \ge 0$,

$$\begin{aligned} f_Y(y;\beta,\lambda) =& \frac{1}{2\alpha\lambda}\int_0^\infty e^{-\frac{\alpha}{\lambda}|y-\delta w| + \frac{\beta}{\lambda}(y-\delta w)} e^{-w} dw \\ =& \frac{1}{2\alpha\lambda}\left(\int_0^{\frac{y}{\delta}} e^{\frac{\beta-\alpha}{\lambda}(y-\delta w)} e^{-w} dw + \int_{\frac{y}{\delta}}^\infty e^{\frac{\alpha+\beta}{\lambda}(y-\delta w)} e^{-w} dw\right) \\ =& \frac{(\beta\delta+\alpha\delta+\lambda)\, e^{-\frac{(\alpha-\beta)}{\lambda}y} - 2\,\alpha\,\delta\, e^{-\frac{y}{\delta}}}{2\,\alpha\,(\beta\delta+\alpha\delta+\lambda)\,(\beta\delta-\alpha\delta+\lambda)}, \end{aligned}$$

Upon substituting $\delta = \frac{\lambda}{\alpha+\beta}$ in the above equations, the required result is obtained. □

A.3. Proof of Proposition 3.6.

Proof. Consider a situation in which $\delta > 0$. Let $W_1 \sim Exp(\alpha - \beta)$, and $\alpha^2 - \beta^2 = 1$. Then by definition $\delta = \frac{\lambda}{\alpha+\beta}$ it can be easily shown that $\lambda W_1 \overset{d}{=} \delta W$, where $W \sim Exp(1)$. Now

(I) Let $W_2, W_3 \sim Exp(\alpha + \beta)$. Then, according to the Mean-mixture representation of the AL distribution in subsection 2.1, we can write

$$Y \overset{d}{=} \lambda(W_1 + W_2 - (\alpha - \beta)^2 W_3) = \delta W + \lambda X \sim LM - AL(\beta, \lambda),$$

where $X \overset{d}{=} W_2 - (\alpha - \beta)^2 W_3 \sim AL(\beta)$.

(II) According to the Mean-variance mixture representation of the AL distribution in subsection 2.1, we have

$$Y \overset{d}{=} \lambda(W_1 + \beta W_2 + \sqrt{W_2} Z) = \delta W + \lambda X \sim LM - AL(\beta, \lambda),$$

where $X \overset{d}{=} \beta W_2 + \sqrt{W_2} Z \sim AL(\beta)$.

(III) According to the Variance mixture epresentation of the AL distribution in subsection 2.1, we have

$$Y \overset{d}{=} \lambda \left(\frac{W_1}{\beta + \alpha} + N W_2 \right) = \delta W + \lambda X \sim LM - AL(\beta, \lambda),$$

where $X = N W_2 \sim AL(\beta)$.

□

A.4. Proof of Theorem 4.1**.** Use of Bayes formula readily yields the following:

1) For $y < 0$,

$$f(w|y) = \frac{f(y|w)f(w)}{Pr(y<0)} = \frac{4\alpha\lambda\ e^{-\frac{\alpha}{\lambda}|y-\delta w| + \frac{\beta}{\lambda}(y-\delta w)} e^{-w}}{2\ \alpha\lambda\ e^{\frac{\alpha+\beta}{\lambda}y}} = \frac{2\ e^{\frac{\alpha+\beta}{\lambda}(y-\delta w)-w}}{e^{\frac{\alpha+\beta}{\lambda}y}} = 2e^{-2w},$$

2) For $y \geq 0$, if $w < \frac{y}{\delta}$, then

$$f(w|y) = c\ f(y|w)f(w) = c\ \frac{e^{\frac{\beta-\alpha}{\lambda}(y-\delta w)-w}}{2\alpha\lambda},$$

and if $w > \frac{y}{\delta}$,

$$f(w|y) = c\ f(y|w)f(w) = c\ \frac{c\ e^{\frac{\beta+\alpha}{\lambda}(y-\delta w)-w}}{2\alpha\lambda},$$

where $c = \frac{1}{Pr(y \geq 0)}$ and $\delta = \frac{\lambda}{\alpha+\beta}$.

8. Aknowledgement

We would like to thank the reviewers for their thoughtful comments and efforts towards improving our manuscript.

References

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

[2] Castaneda-Avila, MA., Ulbricht, CM., & Epstein, MM. (2021). Risk factors for monoclonal gammopathy of undetermined significance: a systematic review. *Annals of Hematology,* **100**, 855–863. https://doi.org/10.1007/s00277-021-04400-7

[3] Epstein, MM., Saphirak, C., Zhou, Y., Leblanc, C., Rosmarin, AG., Ash, A., Singh, S., Fisher, K., Birmann, B., & Gurwitz, JH. (2019). Identifying monoclonal gammopathy of undetermined significance in electronic health data, **29**, 69-76. https://doi: 10.1002/pds.4912

[4] Fernandez, C., & Steel, MFJ. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association,* **93**, 359-371. https://doi.org/10.2307/2669632

[5] Gómez, HW., Venegas, O., & Bolfarine, H. (2007). Skew-symmetric distributions generated by the distribution function of the normal distribution. *Environmetrics,* **18**, 395–407. https://doi.org/10.1002/env.817

[6] Gupta, RD., & Kundo, DA. (2009). New class of weighted exponential distribution. *Statistics,* **43**, 621-634. https://doi.org/10.1080/02331880802605346

[7] Henze, N. (1986). A probabilistic representation of the skew-normal distribution. *Scandinavian Journal of Statistics*, **13**, 271-275.

[8] Kotz, S., Kozubowski, TJ., & Podgórski, K. (2001). The Laplace distribution and Generalizations: *A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Birkhäuser, Boston. https://doi.org/10.1007/978-1-4612-0173-1_5

[9] Krupskii, P., Joe, H., Lee, D., & Genton MG. (2018). Extreme-value limit of the convolution of exponential and multivariate normal distributions: Link to the Hüsler–Reiß distribution. *Multivariate Analysis*, **163**, 80-95. https://doi.org/10.1016/j.jmva.2017.10.006

[10] McLachlan, GJ., Krishnan, T., & Ng, SK. (2004). The EM algorithm. Technical report, Humboldt-Universit¨at Berlin, Center for Applied Statistics and Economics (CASE).

[11] Naderi, M., Arabpour, A., & Jamalizadeh, A. (2018). Multivariate normal mean-variance mixture distribution based on Lindley distribution, *Comm. Statist. Simulation Comput.* **48**, 1179–1192. https://doi.org/10.1080/03610918.2017.1307400

[12] Naderi, M., Mirfarah, E., Wang, WL., & Lin, TI. (2023). Robust mixture regression modeling based on the normal mean-variance mixture distributions. *Computational Statistics* & *Data Analysis*, **180**, p.107661. https://doi.org/10.1016/j.csda.2022.107661

[13] Negarestani, H., Jamalizadeh, A., Shafiei, S., & Balakrishnan, N. (2018). Mean mixtures of normal distributions: properties, inference and application. *Metrika*, **82**, 501–528. https://doi.org/10.1007/s00184-018-0692-x

[14] Pourmousa, R., Jamalizadeh, A., & Rezapour, M. (2015). Multivariate normal mean variance mixture distribution based on Birnbaum-Saunders distribution. *Journal of Statistical Computation and Simulation,* **85**, 2736–2749. https://doi.org/10.1080/00949655.2014.937435

[15] Shanker, R., Fesshaye, H., & Selvaraj, S. (2015). On modeling of lifetimes data using exponential and Lindley distributions. *Biometrics* & *Biostatistics International Journal,* **2(5)**, 140-147. https://doi.org/10.15406/bbij.2015.02.00042

NARJES GILANI
ORCID NUMBER: 0009-0008-7592-5485
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND COMPUTER
SHAHID BAHONAR UNIVERSITY OF KERMAN
KERMAN, IRAN
Email address: narges_gilani@math.uk.ac.ir

REZA POURMOUSA
ORCID NUMBER: 0000-0001-6412-7260
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND COMPUTER
SHAHID BAHONAR UNIVERSITY OF KERMAN
KERMAN, IRAN
Email address: pourm@uk.ac.ir