JIRSS ("2015")
Vol. 14, No. 2, pp 93-118
DOI: 10.7508/jirss.2015.02.006

A New Modification of the Classical Laplace Distribution

R. Mahmoudvand[1], J. Faradmal[2], N. Abbasi[3], K. Lurz[4].

[1]Department of Statistics, Bu-Ali Sina University, Hamedan, Iran.
[2]Modelling of Noncommunicable Diseases Research Center & Department of Biostatistics, School of Public Health, Hamadan University of Medical Sciences, Hamedan, Iran.
[3]Department of Statistics, Payame Noor University, I. R., Iran.
[4]prognostica GmbH, Würzburg, Germany

Abstract. Several modifications of the Laplace distribution have been introduced and applied in various fields up to this day. In this paper, we introduce a modified symmetric version of the classical Laplace distribution. We provide a comprehensive theoretical description of this distribution. In particular, we derive the formulas for the kth moment, quantiles and several useful alternative representations of the distribution. We derive the maximum likelihood estimators of the parameters and investigate their properties via simulation. Finally, we analyse three real-world datasets to illustrate the usefulness of the modified classical Laplace distribution. The results suggest that further improvement to classical Laplace distribution fitting is possible and the new model provides an attractive alternative to the classical Laplace distribution.
Keywords. Laplace distribution, Maximum likelihood estimation, Jackknife method, Order statistics.

MSC: 62E15; 62E10.

Rahim Mahmoudvand(✉)(r.mahmodvand@gmail.com),
Javad Faradmal(javad.faradmal@umsha.ac.ir),
Narges Abbasi(n_abbasi@pnu.ac.ir),
Kristina Lurz (kristina.lurz@prognostica-consulting.de).

1 Introduction

Probability distributions have an undeniable impact on the quality of the procedures used in a statistical analysis. Therefore, considerable efforts have been made concerning the development of distributions along with relevant statistical methods. However, many cases remain where the real data does not follow any of those models. The famous quote by Box (*Essentially, all models are wrong, but some are useful*, see, for example, Box and Draper, 1987) confirms the usefulness of new developments in distribution modelling.

The Laplace distribution belongs to the oldest distributions in probability theory. Its instances continue to enjoy applications in a variety of disciplines which range from image and speech recognition through ocean engineering to finance. These days, they often are the first choice whenever the distribution of the data reveals heavier than Gaussian tails (Kotz et al. , 2001).

Up to this day, many studies have been published with extensions and applications of the Laplace distribution. Extensions to a skewed model as well as to a multivariate setting can be found, for example, in Kotz et al. (2001) and references therein. Liu and Kozubowski (2015) have studied a class of probability distributions on the positive line, which arise when folding the classical Laplace distribution around the origin. Yu and Moyeed (2001) and Yu and Zhang (2005) have proposed a three-parameter asymmetric Laplace distribution. Cordeiro and Lemonte (2011) have proposed the so-called *beta Laplace* distribution as an extension of the Laplace distribution. A parametric link between the minimisation of the sum of the absolute deviates in regression models and the maximum likelihood theory has been considered, for example, by Koenker and Machado (1999), Greasy and Bottai (2007) and Shi et al. (2014). Song et al. (2014) have proposed a robust estimation procedure for mixture linear regression models by assuming that the error term follows a Laplace distribution. Kozubowski et al. (2013) have considered the application of the multivariate generalised Laplace distribution for the construction of a class of moving average vector processes. Nevertheless, the current forms of the Laplace distribution (both classical and generalised forms) have a sharp peak in the middle, which potentially restricts their usefulness.

In the light of this issue, we present a new probability distribution in this paper that can be derived from the symmetric Laplace distribution and can be used for modelling and analysing real data, when a flat shape in the middle of the distribution can be observed. In order to motivate our proposal, we consider the following example. The OECD[1] Jobs Strategy recommends that governments take measures aimed at increasing

[1]Organisation for Economic Co-operation and Development

working-time flexibility (see OECD, 2004). A common work pattern in Iranian offices is to begin between 7:30 and 8:00 am and end at 3:30 pm. Consequently, the daily working time hour has a flat shape in the middle of distribution.

The paper is organised as follows: In Section 2, we provide a brief description of the standard Laplace distribution and introduce a new modified form of it. Section 3 considers maximum likelihood estimation of the parameters of the new distribution. We compare the fits of two models, classical and modified distribution, to three real-world datasets in Section 4. Finally, we give a summary and conclusion in Section 5.

2 Laplace distribution

A standard form of the Laplace distribution is given by the following probability density function:

$$f(x;\theta,\sigma) = \frac{1}{2\sigma}e^{-\frac{|x-\theta|}{\sigma}} \ , \ \ x,\theta \in \mathbb{R} \text{ and } \sigma > 0. \tag{2.1}$$

Apart from *classical Laplace distribution*, this distribution is also known as *double exponential distribution*. It has mean θ and variance $2\sigma^2$.

The Laplace distribution resembles the normal distribution in several characteristics (e. g. unimodality, symmetry), but is sharper at the peak and has fatter tails. A random variable with a distribution that has a sharp probability density corresponds to low entropy (Viola , 1995). Therefore, this property holds for the classical Laplace distribution and a way to increase its entropy is to decrease its sharpness. Note that the entropy of a random variable X can be interpreted as a representation of either the average amount of uncertainty that exists regarding the value of X or the average information received when X is observed.

In the light of the above discussion, we introduce a new modification of the classical Laplace distribution, in which the middle of the distribution is flat instead of sharp. It is worth mentioning that the uniform distribution on a given interval is the maximum entropy distribution among all continuous distributions (Van Campenhout and Cover , 1981). Thus, we use the uniform along with the classical Laplace distribution to create the new distribution.

2.1 Definition and basic properties of the modified classical Laplace distribution

The modified classical Laplace distribution is a probability distribution on $(-\infty, +\infty)$ and its probability density function is given by

$$f(x;\theta,\sigma) = \begin{cases} \frac{1}{3\sigma}\exp\left\{\frac{x-\theta}{\sigma}\right\}, & x < \theta, \\ \frac{1}{3\sigma}, & \theta \le x < \theta+\sigma, \\ \frac{1}{3\sigma}\exp\left\{1-\frac{x-\theta}{\sigma}\right\}, & x \ge \theta+\sigma, \end{cases} \tag{2.2}$$

where $\theta \in \mathbb{R}$ and $\sigma > 0$. Here, we use the notation $\mathcal{CL}(\theta,\sigma)$ for the classical Laplace distribution and $\mathcal{MCL}(\theta,\sigma)$ for the new modified version, respectively.

Proposition 2.1. *Let $Z \sim \mathcal{MCL}(0,1)$. Then*

$$E\left(Z^k\right) = \frac{1}{3}\left[(-1)^k k! + \sum_{\ell=0}^{k+1}\frac{k!}{\ell!}\right] = \begin{cases} \sum_{\ell=1}^{k+1}\frac{k!}{3\ell!} & \text{if } k \text{ is odd}, \\ \sum_{\ell=1}^{k+1}\frac{k!}{3\ell!} + \frac{2}{3}k! & \text{if } k \text{ is even}. \end{cases} \tag{2.3}$$

Corollary 2.1. Let $X \sim \mathcal{MCL}(\theta,\sigma)$. Then

$$E\left(X^k\right) = \sum_{\ell=0}^{k}\binom{k}{\ell}\sigma^\ell E\left(Z^\ell\right)\theta^{k-\ell}. \tag{2.4}$$

Using Corollary 2.1, we obtain $\text{var}(X) = \frac{79}{36}\sigma^2$. Note also that $\mathcal{MCL}(\theta,\sigma)$ is symmetric around $\theta + \frac{\sigma}{2}$, that means

$$f\left(\theta+\frac{\sigma}{2}-x\right) = f\left(\theta+\frac{\sigma}{2}+x\right),\ x \in \mathbb{R}.$$

Consequently, the mean and median are equal to $\theta + \frac{\sigma}{2}$. In contrast to the classical Laplace distribution where the mode is unique, the mode of the modified classical Laplace distribution is an interval, namely $[\theta, \theta+\sigma]$.

The above results indicate that variance and mean for the modified classical Laplace distribution are greater than the variance and mean of the classical Laplace distribution under equal values for θ and σ. In addition, using Corollary 2.1, the kurtosis for the modified Laplace distribution equals 2.48, whereas the classical Laplace has kurtosis equal to 3.

Proposition 2.2. *Let $X \sim \mathcal{MCL}(\theta,\sigma)$. Then the quantile Q_p can explicitly be written as $Q_p = \sigma q_p + \theta$, where*

$$q_p = \begin{cases} \log(3p), & p < \frac{1}{3}, \\ 3p-1, & \frac{1}{3} \le p < \frac{2}{3}, \\ 1-\log\left(3(1-p)\right), & p \ge \frac{2}{3}. \end{cases} \tag{2.5}$$

Using Proposition 2, random variate generation from $\mathcal{MCL}(\theta,\sigma)$ is straightforward. Since the cumulative distribution function of $\mathcal{MCL}(\theta,\sigma)$ has a closed form expression – and so does its inverse – the inversion method can be applied. We describe a $\mathcal{MCL}(\theta,\sigma)$ generator based on Proposition 2.

- Generate a uniform random variate U on $[0;1]$;
- Compute q_U by equation (2.5);
- Return $y = \sigma q_U + \theta$ as a random variate from $\mathcal{MCL}(\theta,\sigma)$.

Proposition 2.3. *A standard modified classical Laplace random variable Z has the representation*

$$Z \stackrel{d}{=} B\left(Y + \frac{B+1}{2}\right) + \left(1-B^2\right)U, \tag{2.6}$$

where $Y \sim \mathrm{Exp}(1)$, (that is Y has an exponential distribution with parameter 1), U is uniformly distributed on $(0,1)$, B takes the values $\{-1,0,1\}$ with probabilities $1/3$, and the random variables Y, U and B are independent.

Proof. Conditioning on B and using the independence between the random variables B, U and Y, we find that

$$\begin{aligned} F_Z(z) &= P(Z \le z) = \frac{P(Y \ge -z) + P(U \le z) + P(Y+1 \le z)}{3} \\ &= \begin{cases} \frac{1}{3}P(Y \ge -z), & z < 0, \\ \frac{1}{3} + \frac{1}{3}P(U \le z), & 0 \le z < 1, \\ \frac{2}{3} + \frac{1}{3}P(Y+1 \le z), & z \ge 1 \end{cases} \\ &= \begin{cases} \frac{1}{3}e^{z}, & z < 0, \\ \frac{1}{3} + \frac{1}{3}z, & 0 \le z < 1, \\ 1 - \frac{1}{3}e^{-(z-1)}, & z \ge 1. \end{cases} \end{aligned}$$

□

Corollary 2.2. Let $Z \sim \mathcal{MCL}(0,1)$, then

$$Z \stackrel{d}{=} \log\left(U_1^{-B}\right) + \left(1 - B^2\right) U_2 + B\frac{B+1}{2},$$
$$Z \stackrel{d}{=} B\left(Y_1 + \frac{B+1}{2}\right) + \left(1 - B^2\right) e^{-Y_2},$$

where $U_i \stackrel{i.i.d}{\sim} U(0,1)$ and $Y_i \stackrel{i.i.d}{\sim} \text{Exp}(1)$, $i = 1, 2$, and all variables are independent from the random variable B as defined in Proposition 1.

2.2 Comparison with the classical Laplace distribution

Figure 1 provides plots of the classical Laplace and the modified classical Laplace densities. We can see that the modified classical Laplace distribution can be used when the values in the centre of the distribution are uniformly distributed.

The classical Laplace distribution can be thought of as two exponential distributions (with additional location parameters) spliced together back-to-back. Now, if the upper exponential distribution in the standard Laplace distribution is divided into two distributions – a uniform distribution and a truncated exponential distribution –, we obtain the modified classical Laplace distribution by additionally replacing the coefficient 1/2 on either side by 1/3. It is easy to see that $\mathcal{CL}(\theta, \sigma)$ and $\mathcal{MCL}(\theta, \sigma)$ have the Shannon entropies (see Shannon (1948)) $\log(2\sigma e)$ and $\log\left(3\sigma e^{\frac{2}{3}}\right)$, respectively. Consequently, a larger entropy is achieved by means of the modified distribution in comparison to the classical Laplace distribution.

2.3 Order statistics

Let $X_1, X_2, \ldots, X_n \stackrel{i.i.d}{\sim} \mathcal{MCL}(0,1)$, and denote by $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$ the corresponding order statistics. Then the following results are obtained by direct application of the formulas for order statistics of distributions.

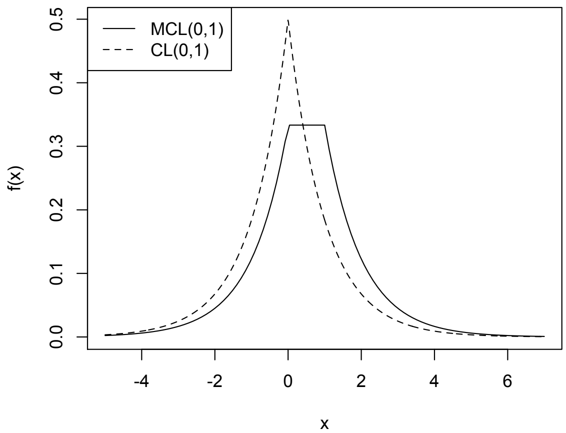

Figure 1: Probability density plot for the modified Laplace distribution.

Proposition 2.4. *The cumulative distribution function (c.d.f) of* $X_{(1)}$ *and* $X_{(n)}$ *are*

$$F_{X_{(1)}}(t) = \begin{cases} 1 - \left(1 - \frac{e^t}{3}\right)^n, & t \leq 0, \\ 1 - \left(1 - \frac{1+t}{3}\right)^n, & 0 < t \leq 1, \\ 1 - \left(\frac{e^{1-t}}{3}\right)^n, & t > 1, \end{cases} \tag{2.7}$$

and

$$F_{X_{(n)}}(t) = \begin{cases} \left(\frac{e^t}{3}\right)^n, & t \leq 0, \\ \left(\frac{1+t}{3}\right)^n, & 0 < t \leq 1, \\ \left(1 - \frac{e^{1-t}}{3}\right)^n, & t > 1. \end{cases} \tag{2.8}$$

Exploiting the relation $F_{X_{(1)}}(t) = 1 - F_{X_{(n)}}(1 - t)$ from Proposition 2.4, we can find a limiting distribution for $X_{(1)}$ and $X_{(n)}$. Note that $X_{(1)} - \log\left(\frac{3}{n}\right)$ has the following c.d.f.:

$$F_{X_{(1)}-\log(3/n)}(t) = \begin{cases} 1-\left(1-\frac{e^t}{n}\right)^n, & t+\log(3/n) \leq 0, \\ 1-\left(1-\frac{1+t+\log(3/n)}{3}\right)^n, & 0 < t+\log(3/n) \leq 1, \\ 1-\left(\frac{e^{1-t}}{n}\right)^n, & t+\log(3/n) > 1. \end{cases} \tag{2.9}$$

Therefore, the limiting distribution of $X_{(1)} - \log\left(\frac{3}{n}\right)$ is given by

$$F_{X_{(1)}-\log(3/n)}(t) \longrightarrow 1-e^{-e^t}, \quad t \in \mathbb{R}. \tag{2.10}$$

Figure 2 shows the empirical probability density (obtained via 10000 simulations from the distribution $\mathcal{MCL}(0,1)$) and the theoretical probability density of $X_{(1)}$ for sample sizes 30, 50, 100 and 500. These plots support the suitability of the limiting distribution as an approximation to the distribution of $X_{(1)}$.

3 Estimation

Let $x_1, \ldots, x_n$ be a sample from $\mathcal{MCL}(\theta,\sigma)$ and $x_{(1)} \leq x_{(2)} \leq \ldots, \leq x_{(n)}$ the corresponding ordered sample. In order to obtain the maximum likelihood estimates of the parameters θ and σ, we will discuss three cases in the following paragraphs:

- σ is known, but θ is unknown;
- θ is known, but σ is unknown;
- Both θ and σ are unknown.

3.1 σ is known

We derive the likelihood function by exploiting the following alternative representation of the density function of the modified Laplace distribution from equation (2.2):

$$f(x;\theta,\sigma) = \frac{1}{3\sigma}\exp\left\{-\frac{1}{2\sigma}\left[|x-\theta|+|x-\sigma-\theta|-\sigma\right]\right\}\ ;\ x \in \mathbb{R}. \tag{3.1}$$

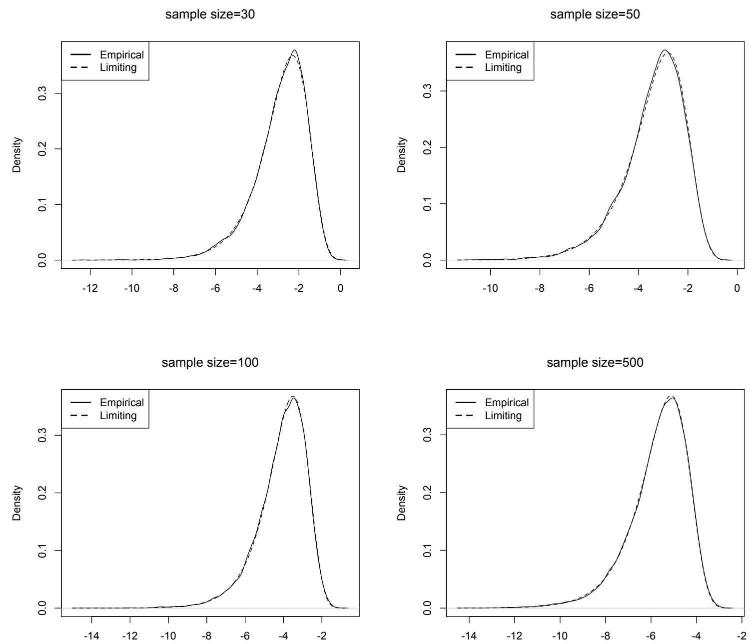

Figure 2: Limiting distribution of $X_{(1)}$ for sample sizes 30, 50, 100 and 500, obtained via 10000 simulations.

Using the representation (3.1), the likelihood function for the modified Laplace distribution can be written as

$$L(\theta) = \left(\frac{1}{3\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma}\left[\sum_{i=1}^{n}|x_i - \theta| + \sum_{i=1}^{n}|\underbrace{x_i - \sigma}_{y_i} - \theta|\right] + \frac{n}{2}\right\}. \tag{3.2}$$

Since σ is known, we can write

$$L(\theta) = \left(\frac{1}{3\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma}\sum_{i=1}^{2n}|y_i - \theta| + \frac{n}{2}\right\}, \tag{3.3}$$

where

$$y_i = \begin{cases} x_i - \sigma, & i = 1, \ldots, n, \\ x_{i-n}, & i = n+1, \ldots, 2n. \end{cases}$$

It is well known that the expression $\sum_{i=1}^{2n} |y_i - \theta|$ is minimised when θ is replaced by the median of the observations $y_1, \ldots, y_{2n}$ (see, e.g., Norton (1984)). Therefore, we have

$$\hat{\theta}_{MLE} = r y_{(n)} + (1-r) y_{(n+1)}, \tag{3.4}$$

where $0 < r < 1$ is an arbitrary constant and $y_{(1)} \leq y_{(2)} \leq \ldots \leq y_{(2n)}$ are the ordered values.

The estimator $\hat{\theta}_{MLE}$ has two properties that are presented in the following proposition.

Proposition 3.1. *Let $\hat{\theta}_{MLE}$ be defined by equation (3.4). Then the following properties hold:*

(i) $\min\left\{x_{(n)} - \sigma, x_{(1)}\right\} \leq \hat{\theta}_{MLE} \leq x_{(\lceil \frac{n+1}{2} \rceil)}$, *where $\lceil v \rceil$ stands for the smallest integer greater or equal to v;*

(ii) $\hat{\theta}_{MLE}$ *is non-increasing with respect to σ.*

Proof. Since $\sigma > 0$, we can consider two extreme cases:

- $x_{(n)} - \sigma \leq x_{(1)}$
- $x_{(i)} - \sigma \leq x_{(i)} \leq x_{(i+1)} - \sigma$ for $i = 1, \ldots, n-1$.

Now, it is easy to see that $\hat{\theta}_{MLE} \in \left[x_{(n)} - \sigma, x_{(1)}\right]$ in the first case and

$$\hat{\theta}_{MLE} \in \begin{cases} \left[x_{(\frac{n}{2})},\ x_{(\frac{n}{2}+1)} - \sigma\right] & \text{for even } n, \\ \\ \left[x_{(\frac{n+1}{2})} - \sigma,\ x_{(\frac{n+1}{2})}\right] & \text{for odd } n, \end{cases}$$

in the second case. This completes the proof of (i). The proof of (ii) is clear by definition of $\hat{\theta}_{MLE}$. □

3.2 θ is known

Without loss of generality, we assume that $\theta = 0$. Thus, the loglikelihood results to

$$\ell(\sigma) = \log L(\sigma) = -n \log(3\sigma) - \frac{1}{2\sigma}\left[\sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |x_i - \sigma|\right] + \frac{n}{2}. \tag{3.5}$$

$\ell(\sigma)$ is continuous everywhere and differentiable except at $\sigma = x_1, \ldots, x_n$. In case the differentiation of $\ell(\sigma)$ exists, we have

$$\begin{aligned}\ell'(\sigma) &= -\frac{n}{\sigma} + \sum_{i=1}^{n} \frac{|x_i|}{2\sigma^2} + \sum_{i=1}^{n} \frac{|x_i - \sigma|}{2\sigma(x_i - \sigma)} + \sum_{i=1}^{n} \frac{|x_i - \sigma|}{2\sigma^2} \\ &= -\frac{n}{\sigma} + \frac{1}{2}\sum_{i=1}^{n} \frac{x_i}{\sigma^2}\left\{\frac{|x_i|}{x_i} + \frac{|x_i - \sigma|}{x_i - \sigma}\right\} = \frac{1}{\sigma}\left\{\sum_{i=1}^{n} \frac{w_i(\sigma)x_i}{\sigma} - n\right\},\end{aligned}$$

where

$$w_i(\sigma) = \begin{cases} -1, & -\infty < x_i < 0, \\ 0, & 0 \le x_i < \sigma, \\ +1, & \sigma < x_i < \infty. \end{cases}$$

Since $w_i(\sigma_1) \ge w_i(\sigma_2)$ for $i = 1, \ldots, n$ when $0 < \sigma_1 < \sigma_2$, we obtain the inequality

$$\left\{\sum_{i=1}^{n} \frac{w_i(\sigma_1)x_i}{\sigma_1} - n\right\} > \left\{\sum_{i=1}^{n} \frac{w_i(\sigma_2)x_i}{\sigma_2} - n\right\}. \tag{3.6}$$

Let us now consider the possible solutions for $\ell'(\sigma) = 0$. If $x_i < 0$ for all i, we have $w_i(\sigma) = -1$ for all i and $\hat{\sigma}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} |x_i|$.

Next, assume $x_i > 0$ for some i. In this case, let $p := \min\{i, x_{(i)} > 0\}$. In addition, let $m \ge p - 1$ be the smallest integer such that either $x_{(m)} < \sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n} < x_{(m+1)}$ or $\sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n} < x_{(m)}$ holds, where $x_{(n+1)} = \infty$ and $x_{(0)} = 0$. We have:

(i) If $x_{(m)} < \sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n} < x_{(m+1)}$, then

$$\ell'(\sigma) \begin{cases} > 0, & 0 < \sigma < \sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n}, \\ = 0, & \sigma = \sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n}, \\ < 0, & \sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n} < \sigma < x_{(m+1)}. \end{cases}$$

In addition, applying $\sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n} < \sigma_1 < x_{(m+1)} < \sigma_2$ to equation (3.6), we get

$$\left\{\sum_{i=1}^{n} \frac{w_i(\sigma_2)x_i}{\sigma_2} - n\right\} < \left\{\sum_{i=1}^{n} \frac{w_i(\sigma_1)x_i}{\sigma_1} - n\right\} < 0,$$

from which follows $\ell'(\sigma) < 0$ for all $\sigma > x_{(m+1)}$. Therefore, we can extend the previous result to

$$\ell'(\sigma) \begin{cases} > 0, & 0 < \sigma < \sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n}, \\ = 0, & \sigma = \sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n}, \\ < 0, & \sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n} < \sigma < \infty, \end{cases}$$

and consequently $\hat{\sigma}_{MLE} = \sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n}$.

(ii) If $\sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n} < x_{(m)}$, then $\ell(\sigma)$ is increasing on $(0, x_{(m)})$ and decreasing on $(x_{(m)}, \infty)$, which implies $\hat{\sigma}_{MLE} = x_{(m)}$.

By applying equality

$$\sum_{i=1}^{n} \frac{w_i(x_{(m+1)})x_i}{n} = \frac{1}{n}\sum_{i=1}^{n} |x_i| - \frac{1}{n}\sum_{i=p}^{m} x_{(i)},$$

the above results can be summarised in Proposition 3.2.

Proposition 3.2. *Let $X_1, \ldots, X_n \overset{i.i.d}{\sim} \mathcal{MCL}(\theta, \sigma)$ under known θ. Then*

$$\hat{\sigma}_{MLE} = \begin{cases} \overline{|Y|}, & \text{if } Y_i < 0 \text{ for all } i, \\ \max\left\{\overline{|Y|} - \frac{1}{n}\sum_{i=p}^{m} Y_{(i)} \ , \ Y_{(m)}\right\}, & \text{if } Y_i > 0 \text{ for some } i, \end{cases}$$

where $Y_i = X_i - \theta,\ i = 1, \ldots, n,\ p = \min\{i, Y_{(i)} > 0\}$ and $\overline{|Y|} = \frac{1}{n}\sum_{i=1}^{n} |Y_i|$. In addition, m can be obtained by

$$m = p - 1 + \sum_{j=p}^{n} \frac{|z_j| + z_j}{2z_j},$$

where $z_p = \overline{|Y|} - Y_{(p)}$ *and* $z_j = \overline{|Y|} - \frac{1}{n}\sum_{i=p}^{j-1} Y_{(i)} - Y_{(j)}$ *for* $j = p+1, \ldots, n$. *The* $\sum_{i=p}^{m} Y_{(m)}$ *are assumed to be equal to zero for* $p > m$.

Proposition 3.2 results in two properties for $\hat{\sigma}_{MLE}$, which are provided in the following proposition.

Proposition 3.3. *Let* $\hat{\sigma}_{MLE}$ *be defined as in Proposition 3.2. Then the following properties hold:*

(i) $\overline{|Y|} - \frac{1}{n}\sum_{i=p}^{n} Y_{(i)} \leq \hat{\sigma}_{MLE} \leq \overline{|Y|}$,

(ii) $\hat{\sigma}_{MLE}$ *is non-increasing with respect to* θ.

Proof. Assertion (i) is clear by definition of $\hat{\sigma}_{MLE}$. With respect to assertion (ii), we suppose that $\theta_1 < \theta_2$ are two arbitrary constants and denote by $\hat{\sigma}_1$ and $\hat{\sigma}_2$ the corresponding MLEs for σ. Using (i), it is sufficient to show that

$$\overline{|x-\theta_1|} - \frac{1}{n}\sum_{i=p}^{n}(x_{(i)} - \theta_1) \geq \overline{|x-\theta_2|}.$$

$\hat{\sigma}_{MLE}$ equals the left hand side of this inequality when it is larger than $x_{(n)} - \theta_1$, and the right hand side for $\overline{|x-\theta_2|} < x_{(n)} - \theta_2$. Therefore, using inequality $\theta_1 < \theta_2$, we have

$$\overline{|x-\theta_1|} - \frac{1}{n}\sum_{i=p}^{n}(x_{(i)} - \theta_1) \geq x_{(n)} - \theta_1 > x_{(n)} - \theta_2 \geq \overline{|x-\theta_2|},$$

which completes the proof of Proposition 3.3. □

3.2.1 Two illustrative examples

Example 3.1. The ordered values corresponding to a random sample of size $n = 10$ taken from $\mathcal{MCL}(0,1)$ are -1.29, -0.81, -0.40, -0.27, 0.37, 0.56, 1.05, 1.23, 1.32, 1.46. Applying Proposition 3.2 we obtain $p = 5$ and $z_5 = 0.506, z_6 = 0.279, z_7 = -0.267, z_8 = -0.552, z_9 = -0.765, z_{10} = -1.037$. Consequently, $m = (5-1) + 2 = 6$, and we have

$$\hat{\sigma}_{MLE} = \max\left\{\overline{|Y|} - \frac{1}{10}\sum_{i=5}^{6} Y_{(i)}\ ,\ Y_{(6)}\right\} = \max\{0.876 - 0.093\ ,\ 0.56\} = 0.783.$$

Example 3.2. Let a sample of size $n = 10$ be taken from $\mathcal{MCL}(0, 1)$ with ordered values -2.08, -1.36, -0.21, 0.08, 0.09, 0.85, 0.86, 1.13, 1.62, 1.76. Applying Proposition 3.2 we obtain $p = 4$ and $z_4 = 0.924, z_5 = 0.906, z_6 = 0.137, z_7 = 0.042, z_8 = -0.314, z_9 = -0.917, z_{10} = -1.219$. Consequently, $m = (4 - 1) + 4 = 7$, and we have

$$\hat{\sigma}_{MLE} = \max\left\{\overline{|Y|} - \frac{1}{10}\sum_{i=4}^{7} Y_{(i)}\ ,\ Y_{(7)}\right\} = \max\{1.004 - 0.188\ ,\ 0.86\} = 0.86.$$

3.3 Both θ and σ are unknown

The case that both parameters θ and σ are unknown constitutes the most realistic real-world situation in comparison to the previous two cases with at least one of the two known. Recall that

$$\ell(\theta, \sigma) = -n\log(3\sigma) - \frac{1}{2\sigma}\left[\sum_{i=1}^{n} |x_i - \theta| + \sum_{i=1}^{n} |x_i - \sigma - \theta|\right] + \frac{n}{2}.$$

Using some results from previous subsections we have

$$\begin{aligned}
\frac{\mathrm{d}\ell(\theta,\sigma)}{\mathrm{d}\theta} &= \frac{1}{\sigma}\sum_{i=1}^{n} w_i(\theta,\sigma),\\
\frac{\mathrm{d}\ell(\theta,\sigma)}{\mathrm{d}\sigma} &= \frac{1}{\sigma}\left\{\sum_{i=1}^{n} \frac{w_i(\theta,\sigma)(x_i-\theta)}{\sigma} - n\right\},
\end{aligned}$$

where

$$w_i(\theta,\sigma) = \begin{cases} -1, & x_i < \theta,\\ 0, & \theta \le x_i < \theta + \sigma,\\ +1, & x_i \ge \theta + \sigma. \end{cases}$$

It is easy to see that $w_i(\theta, \sigma)$ is a non-increasing function in both θ and σ. We can conclude that $\frac{\mathrm{d}\ell(\theta,\sigma)}{\mathrm{d}\theta}$ is a step function varying between $-\frac{n}{\sigma}$ and $\frac{n}{\sigma}$ in θ for a fixed σ. Setting the derivative $\frac{\mathrm{d}\ell(\theta,\sigma)}{\mathrm{d}\sigma}$ equal to zero gives us

$$\sum_{i=1}^{n} \frac{w_i(\theta,\sigma)(x_i-\theta)}{\sigma} - n = 0.$$

The left hand side of this function is continuous and differentiable everywhere as well as decreasing with respect to σ, which guarantees the existence and uniqueness of $\hat{\sigma}_{MLE}$.

Now, note that $\frac{\mathrm{d}\ell(\theta,\sigma)}{\mathrm{d}\theta} = -\frac{n}{\sigma}$ for $\theta > x_{(n)}$, which means that $\ell(\theta,\sigma)$ is decreasing with respect to θ for $\theta > x_{(n)}$. Thus, we can write $\ell(\theta,\sigma) \le \ell(x_{(n)},\sigma)$ for $\theta > x_{(n)}$. In addition, using the previous subsection, we find that $\ell(x_{(n)},\sigma) \le \ell(\hat{\theta}_0,\hat{\sigma}_0)$, where $\hat{\sigma}_0 = \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{\theta}_0|$ and $\hat{\theta}_0 = x_{(n)}$. By making use of the recursive procedure

$$\begin{cases} \hat{\theta}_j = ry_{(n)} + (1-r)y_{(n+1)} \text{ with } y_i = \begin{cases} x_i - \hat{\sigma}_{j-1}, & i = 1,\dots,n, \\ x_{i-n}, & i = n+1,\dots,2n, \end{cases} \\ \hat{\sigma}_j \text{ obtained by Proposition 3.2 with } \theta = \hat{\theta}_j \end{cases}$$

for $j = 1,2,\dots u$, where u is the smallest number with $\frac{\mathrm{d}\ell(\hat{\theta}_u,\sigma)}{\mathrm{d}\sigma} \ge 0$ for $\sigma = \hat{\sigma}_u$ and $\theta = \hat{\theta}_{u+1}$, we can conclude that $\hat{\theta}_{u+1}$ and $\hat{\sigma}_u$ are the maximum likelihood estimations of the parameters θ and σ. Here, $\hat{\sigma}_i$ and $\hat{\theta}_i$ satisfy

$$\begin{aligned} \hat{\sigma}_0 &> \hat{\sigma}_1 > \dots > \hat{\sigma}_u, \\ \hat{\theta}_1 &< \hat{\theta}_2 < \dots < \hat{\theta}_{u+1}. \end{aligned}$$

In order to prove these inequalities, it is enough to show $\hat{\sigma}_0 > \hat{\sigma}_1$. Provided the validity of the inequality $\hat{\sigma}_0 > \hat{\sigma}_1$, Propositions 3.1 and 3.3 can be used to prove $\hat{\sigma}_i > \hat{\sigma}_{i+1}$ for higher i. We intend to show $\min_{x_1,\dots,x_n} \hat{\sigma}_0 > \max_{x_1,\dots,x_n} \hat{\sigma}_1$. From Proposition 3.2 we can conclude $n\hat{\sigma}_1 \le \sum_{i=1}^{n}|x_i - \hat{\theta}_1|$. Furthermore, the definition of $\hat{\sigma}_0$ reveals $\hat{\sigma}_0 \ge \min_{i=1,\dots,n-1}(x_{(i+1)} - x_{(i)})\frac{n-1}{2}$. Therefore, we must show $\min_{i=1,\dots,n-1}(x_{(i+1)} - x_{(i)})\frac{n-1}{2} > \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{\theta}_1|$. For this purpose, let $x_1,\dots,x_n$ be a sample with fixed differences $x_{(i+1)} - x_{(i)}$ for $i = 1,\dots,n-1$. Then we have $\hat{\sigma}_0 = \frac{1}{n}\sum_{i=1}^{n}|x_i - x_{(n)}| = \min_{i=1,\dots,n-1}(x_{(i+1)} - x_{(i)})\frac{n-1}{2}$ and

$$\hat{\theta}_1 \in \begin{cases} \left[x_{(\frac{3n}{4})} - \hat{\sigma}_0,\ x_{(\frac{n}{4}+1)}\right] & \text{if } n \text{ is a multiple of 4}, \\ \left[x_{([\frac{n}{4}]+1)},\ x_{([\frac{3n}{4}]+1)} - \hat{\sigma}_0\right] & \text{if } n \text{ is even, but is not a multiple of 4}, \\ \left[x_{(n-1)} - \hat{\sigma}_0,\ x_{(\frac{n+1}{2})}\right] & \text{if } n \text{ is odd}. \end{cases}$$

Thus, $x_{(1)} \leq \hat{\theta}_1 \leq x_{\left(\frac{n+1}{2}\right)}$, and $\hat{\theta}_1$ is closer to the median of $x_1, \ldots x_n$ than $\hat{\theta}_0$, from which easily follows $\sum_{i=1}^{n} |x_i - \text{Median}\{x_i\}| \leq \sum_{i=1}^{n} |x_i - \hat{\theta}_1| \leq n\hat{\sigma}_0$. Finally, using the inequality $\hat{\sigma}_0 > \hat{\sigma}_1$, we get $\hat{\theta}_1 > \hat{\theta}_2$ by Proposition 3.1. Using $\hat{\theta}_1 > \hat{\theta}_2$ we find $\hat{\sigma}_1 > \hat{\sigma}_2$ by Proposition 3.3 and so on.

4 Empirical Study

4.1 Simulation study

We investigate the behaviour of the maximum likelihood estimators using sample realisations of various $\mathcal{MCL}(\theta, \sigma)$. We consider the following three sets of parameters:

- $\theta = 5,\ \sigma = 0.5,$
- $\theta = 10,\ \sigma = 1,$
- $\theta = 20,\ \sigma = 2.$

For all data generating processes, 1000 datasets of length $N = 10, 30$ and 50 are simulated by means of the statistical software R. Empirical probability density plots for a realisation of these simulations are provided in Figures 3–5.
We use the standard simulation procedure to obtain estimates of the parameters θ and σ. In this way, a set $\left\{(\hat{\theta}^{(j)}, \hat{\sigma}^{(j)}), j = 1, \ldots, 1000\right\}$ of 1000 estimates for the pair (θ, σ) results from the simulation. Then, the mean and the mean squared error (MSE) of $\hat{\theta}$ are estimated by the following formulas:

$$\text{Mean} = \frac{1}{1000} \sum_{j=1}^{1000} \hat{\theta}^{(j)}, \tag{4.1}$$

$$\text{MSE} = \frac{1}{1000} \sum_{j=1}^{1000} \left(\hat{\theta}^{(j)} - \theta\right)^2, \tag{4.2}$$

where $\hat{\theta}^{(j)}$ and θ are replaced by $\hat{\sigma}^{(j)}$ and σ, respectively, when we want to estimate σ. The results in Table 1 show that bias and MSE in estimating the parameters θ and σ decrease with the sample size.

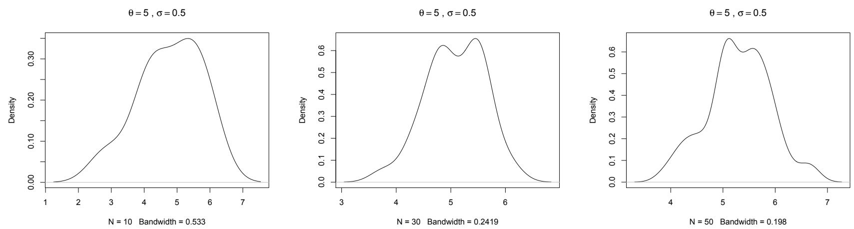

Figure 3: Examples of empirical probability density plots for $\mathcal{MCL}(5, 0.5)$ under samples of size 10, 30 and 50.

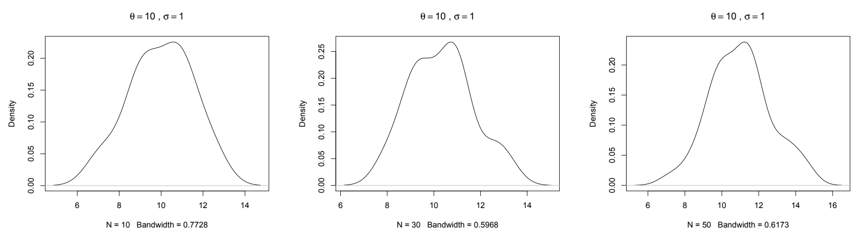

Figure 4: Examples of empirical probability density plots for $\mathcal{MCL}(10, 1)$ under samples of size 10, 30 and 50.

Table 1: MLE properties for $\mathcal{CL}(\theta, \sigma)$ and $\mathcal{MCL}(\theta, \sigma)$ in simulated data.

(θ, σ)	size	$\hat{\theta}_{MLE}$		$\hat{\sigma}_{MLE}$	
		Mean	MSE	Mean	MSE
(5,0.5)	10	5.006	0.052	0.481	0.022
	30	5.000	0.017	0.496	0.007
	50	5.002	0.009	0.494	0.004
(10,1)	10	10.034	0.202	0.951	0.088
	30	9.991	0.066	0.990	0.029
	50	10.003	0.039	0.994	0.017
(20,2)	10	20.070	0.806	1.905	0.354
	30	20.025	0.270	1.989	0.121
	50	19.998	0.159	1.997	0.074

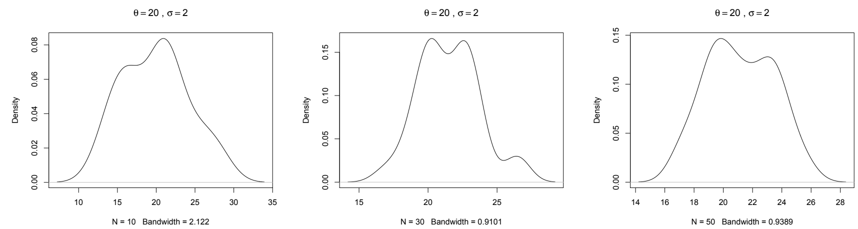

Figure 5: Examples of empirical probability density plots for $\mathcal{MCL}(20, 2)$ under samples of size 10, 30 and 50.

4.2 Real data

Three different datasets are considered to investigate the performance of the modified Laplace distribution on real data. The first one is related to currency exchange rates, the second concerns river flood heights and the last loss ratios in insurance companies. Previous studies examined the possibility of using the Laplace distribution and its extensions on similar datasets. For example, for modelling the exchange rates by means of the Laplace distribution see Kotz et al. (2001), Aryal (2006) and references therein. In addition, Iliopoulos and Balakrishnan (2012) and Bain and Engelhard (1973) studied the flood height data. Modelling loss ratios is according to our knowledge a novel application. It should be mentioned that our aim is not to find the best fitting distribution for the data, but to demonstrate that the modified Laplace distribution is able to provide a better fit than the classical Laplace distribution. We use the Kolmogorov-Smirnov (K-S) test to examine the significance of our results and exploit the Jackknife method (see, for example,Efron (1981)) to estimate the standard error of parameters estimates for $\mathcal{MCL}(\theta, \sigma)$. The standard errors of parameters estimates from $\mathcal{CL}(\theta, \sigma)$ are computed via the R packages `VGAM` (see, for example, Yee (2010)) and `ExtDist` (see Wu et al. (2015)).

Exchange rate data

We consider the annual exchange rates between the United States Dollar (USD) and the Czech Republic Koruna (CZK, 1990-2014), United Kingdom Pound (GBP, 1975-2014), Swedish Krona (SEK, 1960-2014) and Danish Krone (DKK, 1960-2014). The data were obtained from the OECD (2014), available from the website `http://stats.oecd.org`. We examine the fits of $\mathcal{MCL}(\theta, \sigma)$ and $\mathcal{CL}(\theta, \sigma)$ for the logarithm of these data. Table 2 displays ML estimates and their estimated standard errors. Standard errors are low across all parameters, indicating the high precision of the estimates. The probability densities of $\mathcal{MCL}(\theta, \sigma)$ and $\mathcal{CL}(\theta, \sigma)$ using ML estimates as well as the empirical probability density of the data are shown in Figure 6. The plots suggest that further improvement to the classical Laplace distribution fitting can be reached by fitting the modified Laplace distribution $\mathcal{MCL}(\theta, \sigma)$. K-S test statistics and associated p-values are reported in Table 3. They support the graphical indications of a quite reasonable fit by $\mathcal{MCL}(\theta, \sigma)$.

Table 2: Estimated distribution parameters for the exchange rates data.

Currency	$\mathcal{MCL}(\theta, \sigma)$				$\mathcal{CL}(\theta, \sigma)$			
	$\hat{\theta}_{MLE}$		$\hat{\sigma}_{MLE}$		$\hat{\theta}_{MLE}$		$\hat{\sigma}_{MLE}$	
	Estimate	Std.Error	Estimate	Std.Error	Estimate	Std.Error	Estimate	Std.Error
CZK	3.134	0.0560	0.199	0.0290	3.279	0.0497	0.204	0.0408
DKK	1.808	0.0188	0.124	0.0188	1.889	0.0230	0.125	0.0169
SEK	1.710	0.0591	0.193	0.0143	1.847	0.0350	0.199	0.0269
GBP	-0.560	0.0168	0.093	0.0116	-0.493	0.0087	0.097	0.0153

Table 3: Kolmogorov-Smirnov results of fitting $\mathcal{MCL}(\theta, \sigma)$ and $\mathcal{CL}(\theta, \sigma)$ to the exchange rates data.

Quantity	CZK		DKK		SEK		GBP	
	MCL	CL	MCL	CL	MCL	CL	MCL	CL
K-S	0.1399	0.1937	0.1045	0.1437	0.1638	0.2220	0.1074	0.1649
p-value	**0.6612**	0.2687	**0.5852**	0.2059	**0.1044**	0.0089	**0.7057**	0.2029

Flood height data

Bain and Engelhard (1973) considered 33 years (1918-1950) of flood data from two stations on Fox River, Wisconsin. While they modelled the data using a Laplace distribution, we investigate it fitting the modified classical Laplace distribution. Table 4 shows ML estimates and their associated standard errors for both $\mathcal{MCL}(\theta, \sigma)$ and $\mathcal{CL}(\theta, \sigma)$. The resulting densities are displayed in Figure 7 together with the empirical

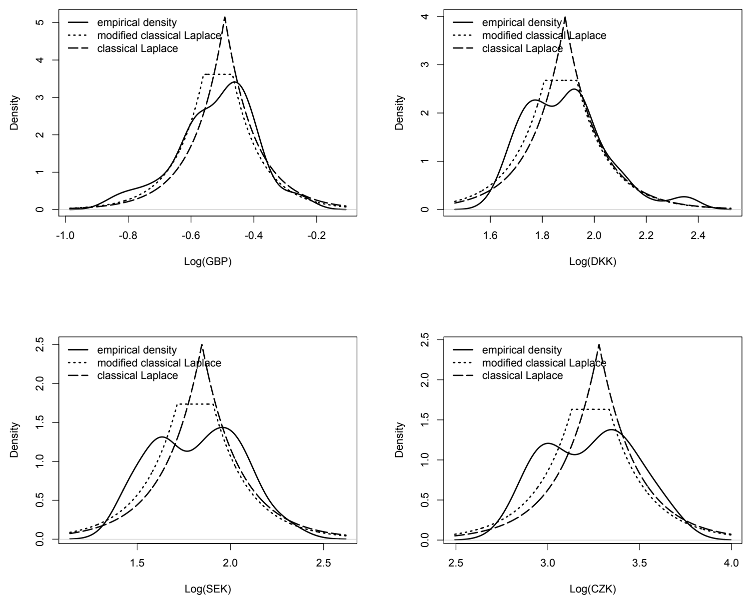

Figure 6: Graphical fitting of $\mathcal{MCL}(\theta, \sigma)$ and $\mathcal{CL}(\theta, \sigma)$ to the exchange rates data.

density of the data, which suggests an improved fit of $\mathcal{MCL}(\theta, \sigma)$ compared to $\mathcal{CL}(\theta, \sigma)$. For these data, the K-S statistic equals 0.1027 and 0.1598 for $\mathcal{MCL}(\theta, \sigma)$ and $\mathcal{CL}(\theta, \sigma)$, respectively. Their associated p-values of 0.877 and 0.3686 indicate an improvement by $\mathcal{MCL}(\theta, \sigma)$.

Table 4: Estimated distribution parameters for flood data.

$\mathcal{MCL}(\theta, \sigma)$				$\mathcal{CL}(\theta, \sigma)$			
$\hat{\theta}_{MLE}$		$\hat{\sigma}_{MLE}$		$\hat{\theta}_{MLE}$		$\hat{\sigma}_{MLE}$	
Estimate	Std.Error	Estimate	Std.Error	Estimate	Std.Error	Estimate	Std.Error
7.923	0.9295	3.200	0.4633	10.131	0.4149	3.361	0.5851

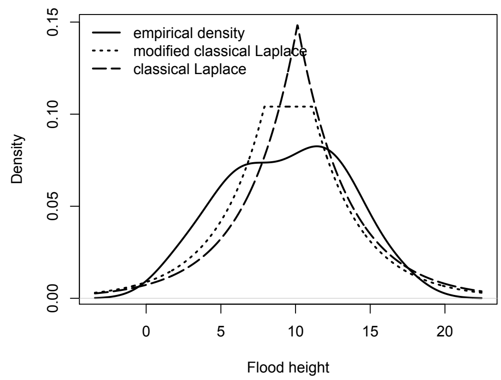

Figure 7: Graphical fitting of $\mathcal{MCL}(\theta, \sigma)$ and $\mathcal{CL}(\theta, \sigma)$ to flood height data.

Loss ratio data

The loss ratio enables insurance companies to determine the overall profitability of the policies they issue. It is computed by dividing the claims paid by an insurer over the premiums earned, usually for the period of a year.

We consider the logarithm of the loss ratio for private cars in motor insurance, based on about 200000 policies and 28721 associated claims in the period Mar 2010 to Mar 2013. It is common to include the gender of the main driver in the actuarial ratemaking. It is also common in practice to start a new period if some changes occur in the observable characteristics of the policies. The policy is then represented as two different lines in the database, and observations are recorded separately for the two periods. Therefore, we decided to divide these data by gender and year.

The maximum likelihood estimates of the parameters for the two distributions $\mathcal{MCL}(\theta, \sigma)$ and $\mathcal{CL}(\theta, \sigma)$ are reported in Table 5. It is evident from the results in Table 5 that the standard errors are low across all parameters, indicating the high precision of the estimates.

The probability densities together with the empirical probability density as displayed in Figure 8 indicate a fairly good fit for the logarithm of the loss ratio, where the graphical evaluation favours $\mathcal{MCL}(\theta, \sigma)$ over $\mathcal{CL}(\theta, \sigma)$ in terms of the fit. Furthermore, the results of the K-S test and associated p-values are reported in Table 6. The p-values returned by the K-S test in this table indicate that the logarithm of the loss ratio for female policy holders might follow $\mathcal{MCL}(\theta, \sigma)$ when employing a significance level of 0.001. Table 6 also shows that using $\mathcal{MCL}(\theta, \sigma)$ results in a significantly improved fit in terms of K-S test statistic and associated p-values.

Table 5: Estimated distribution parameters for the loss ratio data.

Data	$\mathcal{MCL}(\theta,\sigma)$				$\mathcal{CL}(\theta,\sigma)$			
	$\hat{\theta}_{MLE}$		$\hat{\sigma}_{MLE}$		$\hat{\theta}_{MLE}$		$\hat{\sigma}_{MLE}$	
	Estimate	Std.Error	Estimate	Std.Error	Estimate	Std.Error	Estimate	Std.Error
M1	0.3864	0.0244	0.7654	0.0060	0.7566	0.0120	0.8219	0.0101
M2	0.5769	0.0093	0.7026	0.0062	0.9580	0.0105	0.7615	0.0082
M3	0.5815	0.0172	0.7738	0.0081	0.9606	0.0134	0.8305	0.0104
F1	0.3998	0.0431	0.7228	0.0170	0.7651	0.0172	0.7764	0.0165
F2	0.5485	0.0679	0.6883	0.0678	0.9430	0.0147	0.7300	0.0130
F3	0.5003	0.0234	0.7320	0.0240	0.8494	0.0101	0.7867	0.0154

Table 6: Kolmogorov-Smirnov results of fitting $\mathcal{MCL}(\theta, \sigma)$ and $\mathcal{CL}(\theta, \sigma)$ to the loss ratio data.

Sex	Year	length	$\mathcal{MCL}$		$\mathcal{CL}$	
			K-S	p-value	K-S	p-value
Male	2010-2011	6714	0.0258	0.0003	0.0424	7.1e-11
	2011-2012	8629	0.0861	2.1e-16	0.0379	3.2e-11
	2012-2013	6382	0.0316	5.8e-06	0.0436	4.8e-11
Female	2010-2011	2208	0.0250	0.1274	0.0429	0.0006
	2011-2012	3143	0.0303	0.0062	0.0505	2.3e-07
	2012-2013	1645	0.0336	0.0482	0.0456	0.0022

5 Conclusion

We have introduced a new modified Laplace distribution, which replaces the sharp peak of its classical counterpart by a flat segment in the centre of the distribution. The theoretical comparison with the classical Laplace distribution indicates that the classical Laplace distribution is more informative than the modified classical Laplace

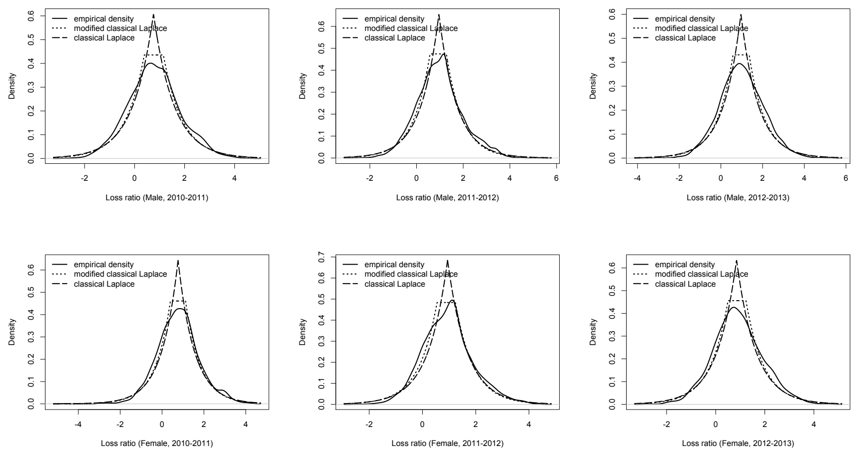

Figure 8: Graphical fitting of $\mathcal{MCL}(\theta, \sigma)$ and $\mathcal{CL}(\theta, \sigma)$ to the loss ratio data.

distribution. A practical analysis of both distributions on real data as part of our study suggests further improvement in classical Laplace distribution fitting by means of the newly proposed distribution.

It may be interesting to extend the results of this paper to cover asymmetry, which is not pursued in the present paper. A suggestion for a modified Laplace distribution of asymmetric shape is provided in the density function

$$f_X(x; \theta, \mu, \sigma, \delta) = \frac{1}{\mu + \sigma + \delta} \begin{cases} e^{-\frac{|x-\theta|}{\delta}}, & x \leq \theta, \\ 1, & \theta < x \leq \theta + \mu, \\ e^{-\frac{|x-\theta-\mu|}{\delta}}, & x > \theta + \mu. \end{cases}$$

The modified Laplace distribution $\mathcal{MCL}(\theta, \sigma)$ introduced in this paper can then be obtained as a special case of the above distribution by setting $\mu = \delta = \sigma$.

Acknowledgement

The authors wish to express their thanks to the anonymous reviewers, whose comments led to an improved version of this paper.

References

Aryal, G.R. (2006), Study of Laplace and Related Probability Distributions and Their Applications. Ph.D Thesis, Department of Mathematics, University of South Florida.

Bain, L. and Engelhard, M. (1973), Interval estimation for the two parameter double exponential distribution. Technometrics 15, 875-887.

Box, George E.P. and Norman R. Draper (1987), Empirical Model-Building and Response Surfaces. John Wiley & Sons, New York.

Cordeiro, G.M. and Lemonte, A.J. (2011), The Beta Laplace Distribution. Statistics & Probability Letters, 81(8), 973–982.

Efron, B. (1981), Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. Biometrika 68, 589-599.

Geraci, M. and Bottai, M. (2007), Quantile regression for longitudinal data using the asymmetric Laplace distribution. Biostatistics 8(1), 140-154.

Wu, W. Godfrey, A.J.R. Govindaraju, K. and Pirikahu, S. ExtDist: Extending the Range of Functions for Probability Distributions. R package version 0.6-3. `http://CRAN.R-project.org/package=ExtDist`

Illipoulos, G. and Balakrishnan, N. (2012), Exact likelihood inference for Laplace distribution based on Type-II censored samples. Journal of Statistical Planning and Inference 142(7):2179-2189.

Koenker, R. and Machado, J.A.F. (1999), Goodness of fit and related inferences processes for quantile regression. Journal of the American Statistical Association, 94(448), 1296–1310.

Kotz, S. Kozubowski, T. and Podgorski, K. (2001), The Laplace Distribution and Generalizations: A Revisit with New Applications. Springer.

Kozubowski, T.J. Podgorski, K. and Rychlik, I. (2013), Multivariate Generalized Laplace Distribution and Related Random Fields. Journal of Multivariate Analysis, 113: 59–72.

Liu, Y. and Kozubowski, T.J. (2015), A folded Laplace distribution. Journal of Statistical Distributions and Applications, DOI 10.1186/s40488-015-003-9.

Norton, R.M. (1984), The Double Exponential Distribution: Using Calculus to Find a Maximum Likelihood Estimator. The American Statistician, 38: 135–136.

OECD: Clocking in and clocking out: Recent trends in working hours (OECD, 2004). Available at
`http://www.oecd.org/dataoecd/42/49/33821328.pdf`.

Shannon, C. E. (1948), A mathematical theory of communication. Bell System Tec. J., 27, 379–432.

Shi, J., Chen, K., and Song, W. (2014), Robust errors-in-variable linear regression via Laplace distribution. Statistics & Probability Letters, 84, 113–120.

Song, W, Yao, W and Xing, Y. (2014), Robust mixture regression model fitting by Laplace distribution. Computational Statistics & Data Analysis, 71 (C), 128–137.

Van Campenhout, J. M. and Cover, T. M. (1981), Maximum entropy and conditional probability. Information Theory, IEEE Transactions on, 27(4), 483–489.

Viola, P. (1995), Alignment by Maximization of Mutual Information. Ph.D. Thesis, Massachusetts Institute of Technology, USA.

Yu, K. and Moyeed, R. A. (2001), Bayesian quantile regression. Statistics and Probability Letters, 54(4), 437-447.

Yu, K. and Zhang, J. (2005), A three-parameter asymmetric Laplace distribution and its extension. Communications in Statistics-Theory and Methods, 34, 1867-1879.

Yee, T. W. (2010), The VGAM package for categorical data analysis. Journal of Statistical Software, 32, 1–34.
`http://www.jstatsoft.org/v32/i10/`.