Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 27 (2024), No. 4, pp. 785–796
DOI : 10.47974/JSMS-1018

The unit log-normal distribution : An alternative for beta and Kumaraswamy distributions

Lucas David Ribeiro-Reis
Department of Economics
Federal University of Alagoas
Santana do Ipanema-AL
Brazil

Abstract

In this paper, through a transformation in the random variable with log-normal distribution, a new distribution was introduced called unit log-normal. Some properties of this new distribution are investigated. It is shown that the unit log-normal distribution belongs to the exponential family and that the maximum likelihood estimators for the two parameters that index the distribution have closed-forms. The accuracy of the maximum likelihood estimators for the two parameters is performed through a simulation study. Two applications to real data are considered, showing the potentiality of the new distribution.

Subject Classification: *(2010) 65C10, 46N30.*

Keywords: *Unit log-normal distribution; Exponential family; Maximum likelihood method; Monte Carlo simulation.*

1. Introduction

It is often common to have to work with data that are in the bounded interval, such as proportion of people in the poverty line, income concentration indexes, etc. that are in the interval (0, 1). The best known distributions and used to analyze such data are the beta and Kumaraswamy (Kumaraswamy, 1980) distributions. These two distributions have two shape parameters, which give flexibility to the distribution. An advantage of the beta distribution over Kumaraswamy distribution is that it belongs to the exponential family and has closed forms for ordinary moments. Already, the advantage of the Kumaraswamy distribution over beta distribution is that the Kumaraswamy distribution has closed forms for

E-mail: econ.lucasdavid@gmail.com

the cumulative distribution function (cdf) and the quantile function, since the cdf of the beta distribution depends on the incomplete beta function.

Recently, several distributions for the unit interval were introduced, being an alternative to the beta and Kumaraswamy distributions. Examples of these new distributions are: log-Lindley (Gómez-Déniz et. al., 2014), unit-inverse Gaussian (Ghitany et. al., 2019), unit Gompertz (Mazucheli et. al., 2019), unit Weibull (Mazucheli et. al., 2020), unit log-logistic (Ribeiro-Reis, 2021), log-weighted exponential (Altun 2021), truncated exponentiated-exponential (Ribeiro-Reis, 2022). All these distributions are bi-parametric.

In this paper, a bi-parametric distribution, obtained through of the transformation of the log-normal distribution, is proposed. This new distribution has some merits. It has several formats for the probability density function (pdf), it belongs to the exponential family of distributions, it has closed forms for the maximum likelihood estimators (MLEs) that index the distribution, the expected information matrix and the standard errors of the MLEs are shown in simple forms. Thus, iterative methods are not needed to obtain the MLEs. One disadvantage is the fact that like the beta distribution, this new distribution doesn't have a closed form for the cdf. The cdf of the proposed distribution depends of the error function, equal to the extended normal distribution proposed by Nwezza and Ugwuowo (2022).

Some properties of the unit log-normal distribution are presented in Section 2. In Section 3, it is shown that this distribution belongs to the exponential family. In Section 4, the maximum likelihood method for estimation of the parameters, their biases, expected information matrix and asymptotic standard errors are investigated. In Sections 5 and 6, simulations study and applications to real data are performed, respectively. Finally, Section 7 concludes the paper.

2. The proposed distribution

The cdf and pdf of a random variable X with log-normal distribution are expressed as

$$F_{\mathrm{LN}}(x;\mu,\sigma^2)=\frac{1}{2}+\frac{1}{2}\mathrm{erf}\left(\frac{\log x-\mu}{\sqrt{2\sigma^2}}\right),\quad x>0$$

and

$$f_{\mathrm{LN}}(x;\mu,\sigma^2)=\frac{1}{x\sqrt{2\pi\sigma^2}}\exp\left[-\frac{1}{2\sigma^2}(\log x-\mu)^2\right],\quad x>0,$$

respectively, where $-\infty<\mu<\infty$, $\sigma^2>0$ and $\mathrm{erf}(z)=\frac{2}{\sqrt{\pi}}\int_0^z \mathrm{e}^{-t^2}dt$, $z\in\mathbb{R}$, is the error function. The error function can still be given as $\mathrm{erf}(z)=2F_{\mathcal{N}}(z\sqrt{2})-1=\mathrm{sign}(z)\,F_{\chi}(2z^2;1)$, where $F_{\mathcal{N}}(\cdot)$ is the cdf of the standard normal distribution, $\mathrm{sign}(\cdot)$ is the signal function and $F_{\chi}(\cdot;v)$ is the cdf of the chi-square distribution with v degree-freedom.

Let now $Y=X/(1+X)$, then the cdf and pdf of Y are

$$F_{\mathrm{ULN}}(x;\mu,\sigma^2)=\frac{1}{2}+\frac{1}{2}\mathrm{erf}\left(\frac{\log[y/(1-y)]-\mu}{\sqrt{2\sigma^2}}\right),\quad 0<y<1$$

and

$$f_{\mathrm{ULN}}(y;\mu,\sigma^2)=\frac{1}{(y-y^2)\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}\left[\log\left(\frac{y}{1-y}\right)-\mu\right]^2\right\},\quad 0<y<1, \tag{1}$$

respectively, where $-\infty<\mu<\infty$ and $\sigma^2>0$. This new distribution is called of unit log-normal and is denoted by $Y\sim\mathrm{ULN}(\mu,\sigma^2)$. For $\mu=0$ and $\sigma^2=1$, the standard ULN distribution is obtained. Figure 1 displays some densities for fixed parameters. Note that the new distribution has different shapes for the density function, which can be symmetric, right symmetric, left symmetric, U-shaped and M-shaped.

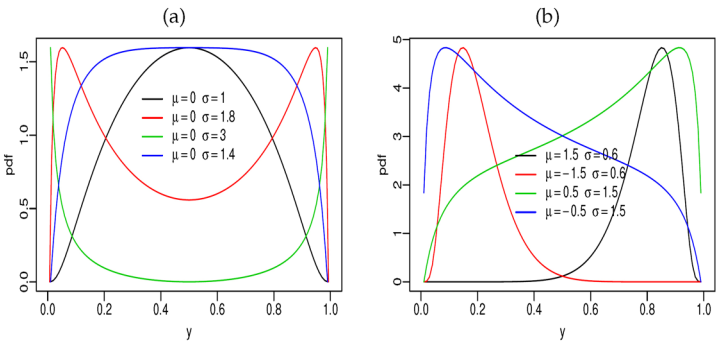

Figure 1

Some ULN pdfs.

By inverting $F_{uln}(x;\mu,\sigma) = u,$ the quantile function of Y is

$$Q_{\text{ULN}}(u;\mu,\sigma^2) = \frac{\exp\{\mu+\sqrt{2\sigma^2}\,\text{erf}^{-1}(2u-1)\}}{1+\exp\{\mu+\sqrt{2\sigma^2}\,\text{erf}^{-1}(2u-1)\}}, \quad 0 < u < 1, \tag{2}$$

where $\text{erf}^{-1}(\cdot)$ is the inverse function of $\text{erf}(\cdot)$. Then,

$$\text{erf}^{-1}(z) = Q_{\mathcal{N}}((1+z)/2)/\sqrt{2} = \text{sign}(z)\sqrt{Q_{\chi}(|z|;1)/2},$$

where $Q_{\mathcal{N}}(\cdot)$ is the quantile function of the standard normal distribution and $Q_{\chi}(\cdot;v)$ is the quantile function of the chi-square distribution with v degree-freedom.

When $u = 1/2,$ the median is obtained. Thus, the median of Y is

$$\text{Md}(Y) = \frac{e^{\mu}}{1+e^{\mu}}.$$

Note that, for $\mu = 0,$ $\text{Md}(Y) = 1/2,$ and then the distribution is symmetric around of $y = 0.5$.

From Equation (2), Y can be easily simulated. Basically, if $V \sim \mathcal{U}(0,1),$ then

$$Y = \frac{\exp\{\mu+\sqrt{2\sigma^2}\,\text{erf}^{-1}(2V-1)\}}{1+\exp\{\mu+\sqrt{2\sigma^2}\,\text{erf}^{-1}(2V-1)\}}, \tag{3}$$

has ULN distribution with density function (1). Another way to generate Y is doing $Y = e^X/(1+e^X),$ where $X \sim \mathcal{N}(\mu,\sigma^2)$ distribution.

Taking the first derivative of the logarithmic of the density function ane equaling to zero

$$\frac{d}{dy}\log f_{\text{ULN}}(y;\mu,\sigma^2) = -\frac{1}{(y-y^2)}\left\{1-2y+\frac{1}{\sigma^2}\left[\log\left(\frac{y}{1-y}\right)-\mu\right]\right\} = 0.$$

So, the mode is the root of

$$1-2y+\frac{1}{\sigma^2}\left[\log\left(\frac{y}{1-y}\right)-\mu\right] = 0.$$

3. Exponetial family

Let the random variable X with pdf $f(x;\theta),$ in which θ is a s-vector of parameters. This random variable belongs to the exponential family s-dimensional if its pdf can be written as

$$f(x;\theta) = h(x)\exp\left[\sum_{j=1}^{s}\eta_j(\theta)t_j(x) - b(\theta)\right], \tag{4}$$

where the functions $\eta_j(\theta)$, $b(\theta)$, $t_j(x)$ and $h(x)$ assume values in subsets of the reals.

Let $w = \log[y/(1-y)]$, note that, the pdf (1) can be written as

$$f_{uln}(y;\mu,\sigma^2) = \frac{1}{(y-y^2)\sqrt{2\pi}}\exp\left\{-\frac{w^2}{2\sigma^2} + \frac{w\mu}{\sigma^2} - \left[\frac{\mu^2}{2\sigma^2} + \frac{1}{2}\log\sigma^2\right]\right\},$$

that belongs to exponential family (4) ($s = 2$), where $\eta_1(\mu,\sigma^2) = \mu/\sigma^2$, $\eta_2(\mu,\sigma^2) = -1/(2\sigma^2)$, $t_1(y) = w$, $t_2(y) = w^2$, $b(\mu,\sigma^2) = \mu^2/(2\sigma^2) + 0.5\log\sigma^2$ and $h(y) = 1/[(y-y^2)\sqrt{2\pi}]$. Thus, by the factorization criterion, $t_1(y)$ and $t_2(y)$ are sufficient statistics for μ and σ^2, respectively. The fact that Y belongs to the bi-parametric exponential family, it is possible to show that

$$\mathbb{E}[t_1(y)] = \mathbb{E}[w] = \mu \quad \text{and} \quad \mathbb{E}[t_2(y)] = \mathbb{E}[w^2] = \mu^2 + \sigma^2.$$

4. Estimation

Consider the random variables $Y_i, \cdots, Y_n \sim \text{ULN}(\mu,\sigma^2)$, and their observed values $y_i, \cdots, y_n$. From Equation (1), the log-likelihood for $(\mu,\sigma^2)^{\text{T}}$ is

$$\mathcal{L}(\mu,\sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \sum_{i=1}^{n}\log(y_i - y_i^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[\log\left(\frac{y_i}{1-y_i}\right) - \mu\right]^2.$$

Differentiating $\mathcal{L}(\mu,\sigma^2)$ with respect to μ and σ^2,

$$\frac{\partial\mathcal{L}(\mu,\sigma^2)}{\partial\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left[\log\left(\frac{y_i}{1-y_i}\right) - \mu\right],$$

$$\frac{\partial\mathcal{L}(\mu,\sigma^2)}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}\left[\log\left(\frac{y_i}{1-y_i}\right) - \mu\right]^2.$$

The MLEs of μ and σ^2, says $\hat{\mu}$ and $\hat{\sigma}^2$, are given equaling the equations above to zero. Then,

$$\hat{\mu} = \overline{w}$$

and

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left[w_i - \overline{w}\right]^2 = \frac{1}{n}\sum_{i=1}^{n}w_i^2 - \overline{w}^2,$$

where $w_i = \log[y_i/(1-y_i)]$ and $\bar{w} = n^{-1}\sum_{i=1}^{n} w_i$ is the mean of the w_i.

Note that $\hat{\mu}$ is a estimator unbiased of μ. From Section 3,

$$\mathbb{E}[\hat{\mu}] = \mathbb{E}[\bar{w}] = \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^{n} w_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[w_i] = \frac{1}{n}n\mu = \mu.$$

From Section 3, $\mathbb{V}(w_i) = \mathbb{E}[w_i^2] - (\mathbb{E}[w_i])^2 = \sigma^2$, then

$$\mathbb{V}(\bar{w}) = \mathbb{V}\left(\frac{1}{n}\sum_{i=1}^{n} w_i\right) = \frac{1}{n^2}\mathbb{V}\left(\sum_{i=1}^{n} w_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{V}(w_i) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}.$$

Hence, $\mathbb{E}[\bar{w}^2] = \mathbb{V}(\bar{w}) + (\mathbb{E}[\bar{w}])^2 = \frac{\sigma^2}{n} + \mu^2$. Then,

$$\mathbb{E}[\hat{\sigma}^2] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[w_i^2] - \mathbb{E}[\bar{w}^2] = \frac{1}{n}n(\mu^2+\sigma^2) - \frac{\sigma^2}{n} - \mu^2 = \sigma^2\left(1-\frac{1}{n}\right).$$

So, $\hat{\sigma}^2$ is a estimator biased for σ^2 and its bias is $B(\hat{\sigma}^2) = 1-1/n$. Note that, when $n \to \infty$, the bias $B(\hat{\sigma}^2) \to 0$, indicating that $\hat{\sigma}^2$ is a consistent estimator of σ^2. Taking the second derivatives

$$\frac{\partial^2 \mathcal{L}(\mu,\sigma^2)}{\partial \mu^2} = -\frac{n}{\sigma^2},$$
$$\frac{\partial^2 \mathcal{L}(\mu,\sigma^2)}{\partial \sigma^4} = \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{i=1}^{n}[w_i-\mu]^2.$$

Once that $\hat{\sigma}^2 > 0$, the second derivatives, evaluated in $\hat{\mu}$ and $\hat{\sigma}^2$

$$\frac{\partial^2 \mathcal{L}(\mu,\sigma^2)}{\partial \mu^2} = -\frac{n}{\hat{\sigma}^2} < 0,$$
$$\frac{\partial^2 \mathcal{L}(\mu,\sigma^2)}{\partial \sigma^4} = \frac{1}{\hat{\sigma}^4}\left\{\frac{n}{2} - \frac{1}{\hat{\sigma}^2}\sum_{i=1}^{n}[w_i-\mu]^2\right\} = -\frac{n}{2\hat{\sigma}^4} < 0,$$

which shows that, of fact, $\hat{\mu}$ and $\hat{\sigma}^2$ are the points that maximize the log-likelihood $\mathcal{L}(\mu,\sigma^2)$. The hessian matrix is

$$J(\mu,\sigma^2) = \begin{bmatrix} J_{\mu\mu} & J_{\mu\sigma^2} \\ J_{\sigma^2\mu} & J_{\sigma^2\sigma^2} \end{bmatrix},$$

where $J_{mn} = \partial^2 \mathcal{L}(\mu,\sigma^2)/\partial\theta_m \partial\theta_n$, with $J_{mn} = J_{nm}$. Only missing $J_{\mu\sigma^2}$, that is

$$J_{\mu\sigma^2} = -\frac{1}{\sigma^4}\sum_{i=1}^{n}[w_i-\mu] = -\frac{1}{\sigma^4}\sum_{i=1}^{n} w_i + \frac{n\mu}{\sigma^4}.$$

The expected information matrix is given by $K(\mu,\sigma^2) = -\mathbb{E}[J(\mu,\sigma^2)]$. Taking the expect value of the second derivatives

$$\begin{aligned}
\mathbb{E}[J_{\mu\mu}] &= -\frac{n}{\sigma^2},\\
\mathbb{E}[J_{\mu\sigma^2}] &= -\frac{1}{\sigma^4}\sum_{i=1}^{n}\mathbb{E}[w_i] + \frac{n\mu}{\sigma^4} = -\frac{n\mu}{\sigma^4} + \frac{n\mu}{\sigma^4} = 0,\\
\mathbb{E}[J_{\sigma^2\sigma^2}] &= \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\left[\sum_{i=1}^{n}\mathbb{E}[w_i^2] - 2\mu\sum_{i=1}^{n}\mathbb{E}[w_i] + n\mu^2\right]\\
&= \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}[n(\mu^2+\sigma^2) - 2n\mu^2 + n\mu^2]\\
&= \frac{n}{2\sigma^4} - \frac{n}{\sigma^4} = -\frac{n}{2\sigma^4}.
\end{aligned}$$

So, the expected information matrix is

$$K(\mu,\sigma^2) = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{bmatrix}.$$

Under the usual regularity conditions of the MLEs, for n large,

$$\begin{pmatrix} \hat{\mu} \\ \hat{\sigma}^2 \end{pmatrix} \overset{a}{\sim} \mathcal{N}\left(\begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}, K(\mu,\sigma^2)^{-1}\right),$$

where $\overset{a}{\sim}$ denotes asymptotic distribution and

$$K(\mu,\sigma^2)^{-1} = \frac{2\sigma^6}{n^2}\begin{bmatrix} \frac{n}{2\sigma^4} & 0 \\ 0 & \frac{n}{\sigma^2} \end{bmatrix} = \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{bmatrix}.$$

Confidence intervals and hypothesis testing can be performed using the asymptotic distribution of the MLEs. Then, the standard errors (SEs) of $\hat{\mu}$ and $\hat{\sigma}^2$ are given by $\text{se}(\hat{\mu}) = \hat{\sigma}/\sqrt{n}$ and $\text{se}(\hat{\sigma}^2) = \hat{\sigma}^2\sqrt{2/n}$, respectively.

5. Simulation study

In this section, a simulation study with $10{,}000$ replications is performed to evaluate some asymptotic properties of the MLEs for the two parameters of the ULN distribution. The observations are generated using Equation (3) with sample sizes $n = \{100, 200, 300, 400\}$, for two scenarios. The true parameters are: $\mu = -2.3$ and $\sigma^2 = 1.2$ for scenario 1 and $\mu = 1.7$ and $\sigma^2 = 0.5$ for scenario 2. The simulations were carried out in the programming language R (R Core Team, 2020)

Table 1 lists the average estimates (AEs) and the mean squared errors (MSEs), for the two scenarios. It is noticed that the MLEs converge to the true parameters and the MSEs decrease when the sample size n increases.

Table 1

Simulations study under scenarios 1 to 2.

n	Par	scenario 1			scenario 2	
		AE	MSE		AE	MSE
100	μ	–2.300239	0.011727		1.699846	0.004886
	σ^2	1.187214	0.029482		0.494672	0.005118
200	μ	–2.300463	0.005926		1.699701	0.002469
	σ^2	1.192631	0.014290		0.496930	0.002481
300	μ	–2.300453	0.003947		1.699708	0.001645
	σ^2	1.195294	0.009606		0.498039	0.001668
400	μ	–2.300175	0.002992		1.699887	0.001247
	σ^2	1.196519	0.007266		0.498550	0.001261

6. Applications

Two applications to real data are considered. The popular beta and Kumaraswamy models were chosen as competitive models. The pdfs of the beta and Kumaraswamy distributions are

$$f\,b(y;\alpha,\beta)=\frac{1}{B(\alpha,\beta)}y^{\alpha-1}(1-y)^{\beta-1},\quad 0<y<1$$

and

$$f\,k(y;\alpha,\beta)=\alpha\beta y^{\alpha-1}(1-y^{a})^{\beta-1},\quad 0<y<1,$$

respectively, where $\alpha,\beta>0$ and $B(\alpha,\beta)=\int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt$ is the beta function.

The choice of the best models are based on the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Hannan-Quinn Information Criterion (HQIC) statistics. The best model is the one with the lowest values of these statistics.

Two data sets were considered, namely:

- (data 1) These data correspond to the Firjan health index (Firjan, 2020) of 853 municipalities in the State of Minas Gerais, Brazil (year

2018). This index is between 0 and 1. The closer to 1, the better the health condition. While values close to 0, it indicates that health conditions are more precarious.

- (data 2) These data refer to the proportion of households with per capita household income below the extreme poverty line in the 27 Brazilian states, in 2014 year (Ipeadata, 2020)..

Table 2 presents the descriptive statistics of the two data sets. Note that, the data 1 has a concentration closer to 1, while the values from data 2 are concentrated closer to 0.

Table 2

Descriptive statistics.

Description	data 1	data 2
Min.	0.3569	0.0109
1st Qu.	0.7073	0.0165
Median	0.7935	0.0380
Mean	0.7726	0.0429
3rd Qu.	0.8535	0.0642
Max.	0.9723	0.1033

The output of the MLEs, with standard errors in parenthesis, and the information criteria for date 1 and date 2 are listed in Tables 3 and 4, respectively. Note that all estimates are quite significant, given their low standard errors.In both applications, the ULN distribution was the one that presented the lowest values of the information criteria.

Table 3

MLEs, SEs and information criteria for data 1.

Distribution	Estimates		AIC	BIC	HQIC
ULN (μ, σ^2)	1.3329	0.4294	–1488.59	–1479.09	–1484.95
	(0.0224)	(0.0208)			
Beta (α, β)	11.2275	3.3156	–1478.25	–1468.75	–1474.61
	(0.5481)	(0.1534)			
Kw (α, β)	7.8381	4.1046	–1487.26	–1477.76	–1483.62
	(0.2660)	(0.2481)			

Table 4

MLEs, SEs and information criteria for data 2.

Distribution	Estimates		AIC	BIC	HQIC
ULN (μ, σ^2)	–3.3606	0.6005	–119.40	–116.81	–118.63
	(0.1491)	(0.1634)			
Beta (α, β)	2.0286	45.2737	–119.23	–116.64	–118.46
	(0.5132)	(12.9075)			
Kw (α, β)	1.4866	91.5259	–118.94	–116.35	–118.17
	(0.2242)	(59.6844)			

The histograms with the estimated densities for date 1 and date 2 are in Figures 2(a) and 2(b), respectively. Note that, for data 1, the ULN pdf has better behavior, corroborating with the information criteria. As for date 2, the result from the graphical analysis is inconclusive. However, the information criteria indicate that the ULN model is more suitable.

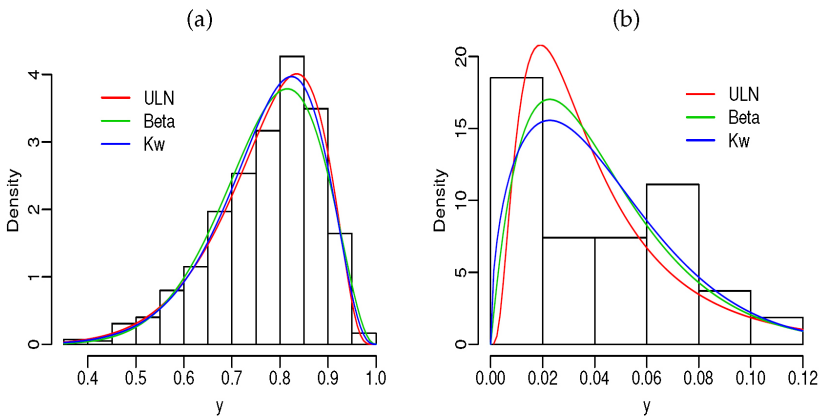

Figure 2

Estimated densities (a) for data 1 and (b) for data 2.

7. Conclusions

In this paper, a new probability distribution for the range (0.1) was introduced. This new distribution has shapes for its density function, which can be left asymmetric, right asymmetric, symmetric, U-shaped

and M-shaped. The two parameters that index the distribution are estimated by maximum likelihood. It was demonstrated that these estimators, as well as the expected information matrix, have a closed form. Thus, iterative methods are not needed to obtain these estimators. It was demonstrated, analytically and numerically via simulations study, the consistency of the maximum likelihood estimators.

It has been shown that the new distribution belongs to the exponential family, which is an advantage, since of the many distributions that have been proposed recently, they do not have such a feature. The fact that the new distribution belongs to the exponential family guarantees that the statistics used to obtain the maximum likelihood estimators are sufficient and complete. And the expected value and variance of these statistics can be obtained easily.

The usefulness of the new model was shown through two applications to real data. The new model performed better than the popular beta and Kumaraswamy models. Due to its properties, it is expected that, in the future, this new distribution will be widely used to analyze data in the interval (0, 1).

References

[1] Altun, E., The log-weighted exponential regression model: alternative to the beta regression model. *Communications in Statistics-Theory and Methods*, 50(10) : 2306–2321 (2021).

[2] Firjan (2020). Federação das Indústrias do Estado do Rio de Janeiro. *Índice FIRJAN de Desenvolvimento Municipal*.

[3] Ghitany, M. E., Mazucheli, J., Menezes, A. F. B., and Alqallaf, F., The unit-inverse Gaussian distribution: A new alternative to two-parameter distributions on the unit interval. *Communications in Statistics-Theory and Methods*, 48(14) : 3423–3438 (2019).

[4] Gómez-Déniz, E., Sordo, M. A., and Calderín-Ojeda, E. (2014). The log–Lindley distribution as an alternative to the beta regression model with applications in insurance. *Insurance: Mathematics and Economics*, 54:49–57.

[5] Ipeadata. Instituto de Pesquisa Econômica Aplicada. *Base de Dados* (2020).

[6] Kumaraswamy, P., A generalized probability density function for doublebounded random processes. *Journal of Hydrology*, 46(1-2) : 79–88 (1980).

[7] Mazucheli, J., Menezes, A. F. B., and Dey, S., Unit-Gompertz distribution with applications. *Statistica*, 79(1) : 25–43 (2019).

[8] Mazucheli, J., Menezes, A. F. B., Fernandes, L. B., de Oliveira, R. P., and Ghitany, M. E., The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates. *Journal of Applied Statistics*, 47(6) : 954–974 (2020).

[9] Nwezza, E. E. and Ugwuowo, F. I., An extended normal distribution for reliability data analysis. *Journal of Statistics and Management Systems*, 25(2) : 369–392 (2022).

[10] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria(2020).

[11] Ribeiro-Reis, L. D., Unit log-logistic distribution and unit log-logistic regression model. *Journal of the Indian Society for Probability and Statistics*, 22(2) : 375–388 (2021).

[12] Ribeiro-Reis, L. D., Truncated exponentiated-exponential distribution: A distribution for unit interval. *Journal of Statistics and Management Systems*, pages 1–12 (2022).

Received June, 2022