Unit-log-symmetric models: Characterization, statistical properties and application to internet access data

Roberto Vila *1, 2, Narayanaswamy Balakrishnan, † 2, Helton Saulo ‡ 1, and Peter Zörnig §1

1 Department of Statistics, University of Brasília, Brasília, Brazil
2 Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada

November 29, 2022

Abstract

We present the unit-log-symmetric model which is based on the bivariate log-symmetric distribution. It is a flexible family of distributions over the interval (0, 1). We study mathematical properties like stochastic representations, symmetry, modality, moments, quantile function, entropy and maximum likelihood estimators, giving special attention to the specific cases of the unit-log-normal, the unit-log-Student-t and the unit-log-Laplace distributions. Simulation results and practical application are also presented.

Keywords. Unit-log-symmetric distribution · Log-symmetric distribution · Bivariate model · Bounded distributions · MCMC.

Mathematics Subject Classification (2010). MSC 60E05 · MSC 62Exx · MSC 62Fxx.

1 Introduction

Usually, distribution theory has focused on unbounded distributions. However, in the recent past, bounded distributions have also become a subject of interest. In many practical applications, for example in medicine, biology, economics and financial sciences arise bounded data. Some of the most common models used in this case are the beta distribution and its generalizations,
the Kumaraswamy and the Topp-Leone distribution, see e.g. Korkmaz (2020), Jodr´a (2020) and Bakouch et al. (2022). In view of constructing flexible models for real world problems, it is desirable to have a whole family of distributions available. For this purpose, transformations have been developed in the statistical literature which construct a bounded distribution for any given distribution over the real line (−∞, ∞) or over (0, ∞), respectively. In Korkmaz (2020, p. 2098), three transformations of these types are described. One of them (see line 6 of p. 2098) is based on the logistic function and transforms any distribution over (−∞, ∞) into a distribution over (0, 1).

In the present paper we study a transformation that contains the latter as a special case (for μ = 0 and σ = 1). However, our transformation is derived from the bivariate log-symmetric distribution which is defined, to our knowledge, for the first time in this paper. We introduce our model, the unit-log-symmetric (ULS) distribution in Section 2. We devote particular attention to the special cases unit-log-normal, unit-log-Student-t and unit-log-Laplace distributions. In Section 3 mathematical properties of our model are studied. Stochastic representations of the ULS are given which are used to simplify the density considerably; see (3.8). The symmetry of the ULS density is established and the above-mentioned special cases are explicitly formulated, and graphically illustrated. In Subsection 3.3 the modality is studied. The unit-log-normal density may be unimodal or bimodal. The unit-log-Student-t density may have a bathtub shape or obey two minimum points. The unit-log-Laplace density may be unimodal shape or obey two minimum points. The quantile function is investigated in Subsection 3.4, the moments in Subsection 3.5 and the Shannon entropy in Subsection 3.6. Finally, the maximum likelihood (ML) estimators are studied in Subsection 3.7. In Section 4 we present Monte Carlo simulation results for evaluating the performance of the estimators and for assessing the empirical distribution of the residuals. The paper ends with applications of the unit-log-normal, unit-log-Student-t and unit-log-Laplace distributions to a real world data set. The three models are fitted to the proportion of world population using Internet (Subsection 5). In Section 6, we provide some concluding remarks.

2 The Unit-Log-Symmetric Distribution

We say that a continuous random variable W with support (0, 1) follows a unit-log-symmetric (ULS) distribution with parameter σρ = σ√2(1 − ρ), denoted by W ∼ ULS(σρ, gc), if its probability density function (PDF) is given by

\[ f_W(w; \sigma_\rho) = \frac{1}{w(1-w)\sigma^2\sqrt{1-\rho^2}Z_{gc}} \int_0^\infty \frac{1}{t} g_c \left( \tilde{t}_w^2 - 2\rho \tilde{t}_w \tilde{t} + \tilde{t}^2 \right) dt, \quad 0 < w < 1, \tag{2.1} \]

\[ t_w = \left( \frac{w}{1-w} \right) t, \quad \tilde{t}_w = \log \left( \frac{t_w}{\eta} \right)^{1/\sigma}, \quad \tilde{t} = \log \left( \frac{t}{\eta} \right)^{1/\sigma}, \quad \eta = \exp(\mu), \]
where $\mu \in \mathbb{R}$, $\sigma > 0$, $\rho \in (-1, 1)$, respectively. Further, $Z_{g_c}$ is a positive constant defined as

$$Z_{g_c} = \int_0^\infty \int_0^\infty \frac{1}{t_1 t_2 \sigma^2 \sqrt{1-\rho^2}} \sigma^2 \sqrt{1-\rho^2} \frac{(t_1^2 - 2\rho \tilde{t}_1 \tilde{t}_2 + \tilde{t}_2^2)}{\sqrt{1-\rho^2}} \, dt_1 dt_2 = \pi \int_0^\infty g_c(u) \, du,$$  \hspace{1cm} (2.2)

and $g_c$ is a scalar function referred as the density generator which may or may not depend on extra parameters considered known (Fang et al., 1990).

**Remark 2.1.** We decided to define the PDF (2.1) as a uniparametric function for two reasons: first, to avoid identification problems (Proposition 3.1) and second, because the shape of this PDF depends only on the parameter $\sigma$ (Subsection 3.3).

For simplicity of presentation, in this paper we focus our study on the density generators of unit-log-normal, unit-log-Student-$t$ and unit-log-Laplace (Table 1).

### Tab. 1: Normalization constants ($Z_{g_c}$) and density generators ($g_c$) for some ULS distributions.

| Distribution            | $Z_{g_c}$   | $g_c$                          | Parameter |
|-------------------------|-------------|--------------------------------|-----------|
| Unit-log-normal         | $2\pi$      | $\exp(-x/2)$                   | $-\infty$ |
| Unit-log-Student-$t$    | $\Gamma(\nu/2)\pi \Gamma((\nu+2)/2)$ | $(1 + \frac{\nu}{\nu})^{-\nu-2}/2$ | $\nu > 0$ |
| Unit-log-Laplace        | $\pi$       | $K_0(\sqrt{2x})$              | $-\infty$ |

Here, in the Table 1, $K_0(u) = \int_0^\infty t^{-1} \exp(-t - \frac{u^2}{4t}) \, dt / 2$, $u > 0$, is the Bessel function of the third kind (for more details on the main properties of $K_0$, see appendix of Kotz et al., 2001).

### 3 Some basic properties of the model

In this section, some mathematical properties of proposed unit-log-symmetric distribution are discussed. For this purpose, the definition of the bivariate log-symmetric distribution is fundamental.

We say that a continuous random vector $(T_1, T_2)$, with $T_1$ and $T_2$ identically distributed, follows a bivariate log-symmetric (BLS) distribution if its joint PDF is given by

$$f_{T_1, T_2}(t_1, t_2; \theta) = \frac{1}{t_1 t_2 \sigma^2 \sqrt{1-\rho^2} Z_{g_c}} \sigma^2 \sqrt{1-\rho^2} g_c \left( \frac{t_1^2 - 2\rho t_1 \tilde{t}_2 + \tilde{t}_2^2}{1-\rho^2} \right), \quad t_1, t_2 > 0,$$

$$\tilde{t}_i = \log \left[ \left( \frac{t_i}{\eta} \right)^{1/\sigma} \right], \quad \eta = \exp(\mu), \quad i = 1, 2,$$

where $\theta = (\eta, \sigma, \rho)$ is the parameter vector with $\mu \in \mathbb{R}$, $\sigma > 0$, $\rho \in (-1, 1)$, and $g_c$ and $Z_{g_c}$ are as in (2.1) and (2.2), respectively. Observe that $Z_{g_c}$ in (2.2) is the normalization constant for $f_{T_1, T_2}$ and that the joint PDF of $(T_1, T_2)$ can be obtained as the PDF of an exponential transformation of an elliptically symmetric vector (Balakrishnan and Lai, 2009, Chapter 13, p. 591).
3.1 Stochastic representation

A simple observation shows that, if \((T_1, T_2) \sim \text{BLS}(\theta, g_c)\) then

\[
\mathbb{P}\left( \frac{T_1}{T_1 + T_2} \leq w \right) = \mathbb{P}\left( \frac{T_1}{T_2} \leq \frac{w}{1-w} \right), \quad 0 < w < 1. \tag{3.1}
\]

And therefore,

\[
f_{\frac{T_1}{T_1+T_2}}(w) = \frac{1}{(1-w)^2} f_{\frac{T_1}{T_2}}\left( \frac{w}{1-w} \right) = \frac{1}{(1-w)^2} \int_0^{\infty} t f_{T_1, T_2}(t_w, t; \theta) \, dt = f_W(w; \sigma). \tag{3.2}
\]

That is, if \((T_1, T_2) \sim \text{BLS}(\theta, g_c)\) the ULS random variable \(W\) admits the stochastic representation:

\[
W = \frac{T_1}{T_1 + T_2}. \tag{3.2}
\]

3.2 Characterizations

From the classic stochastic representation of a bivariate elliptical vector (Balakrishnan and Lai, 2009, Subsection 13.2.3, p. 593) it follows that \((T_1, T_2) \sim \text{BLS}(\theta, g_c)\) has the following representation:

\[
T_1 = \eta \exp(\sigma Z_1), \quad T_2 = \eta \exp(\sigma [\rho Z_1 + \sqrt{1-\rho^2} Z_2]), \tag{3.3}
\]

where \(Z_1 = RDU_1\) and \(Z_2 = R\sqrt{1-D^2}U_2; U_1, U_2, R,\) and \(D\) are mutually independent, \(\rho \in (-1,1), \eta = \exp(\mu), \mathbb{P}(U_i = -1) = \mathbb{P}(U_i = 1) = 1/2, i = 1, 2,\) and the variables \(D\) and \(R\) have PDFs \(f_D(d) = 2/(\pi \sqrt{1-d^2}), \, d \in (0,1),\) and \(f_R(r) = 2rg_c(r^2)/[\int_0^{\infty} g_c(u) \, du], \, r > 0,\) respectively.

Hence, \(T_1/T_2 = \exp(\sigma [(1-\rho)Z_1 - \sqrt{1-\rho^2} Z_2]).\) Using this identity we get

\[
\mathbb{P}\left( \frac{T_1}{T_2} \leq \frac{w}{1-w} \right) = \mathbb{P}(Z_\rho \leq A(w)), \quad 0 < w < 1. \tag{3.6}
\]

By using (3.6), the following result is immediate.
Proposition 3.1. If the distribution of $Z_\rho$ depends (or not) only on the parameters of the density generator $g_c$, then the ULS distribution is identifiable.

The inverse function of $A$, denoted by $A^{-1}$, is expressed as

$$A^{-1}(w) = \frac{1}{1 + \exp(-\sigma_\rho w)}. \quad (3.7)$$

Using this notation, from (3.6) the following result follows immediately.

Proposition 3.2 (Another stochastic representation). Given the distribution of $Z_\rho$, we have $A^{-1}(Z_\rho) \sim \text{ULS}(\sigma_\rho, g_c)$. On the other hand, if $W \sim \text{ULS}(\sigma_\rho, g_c)$ then $A(W)$ and $Z_\rho$ have the same distribution.

Differentiating (3.6) with respect to $w$, the ULS PDF (2.1) of $W$ is characterized as

$$f_W(w; \sigma_\rho) = \frac{1}{w(1-w)\sigma_\rho} f_{Z_\rho}(A(w)), \quad 0 < w < 1, \quad (3.8)$$

where $f_{Z_\rho}$ denotes the PDF of $Z_\rho$.

That is, the distribution of $W$ is completely determined by the distribution of $Z_\rho$.

Proposition 3.3 (Symmetry). The ULS PDF (2.1) is symmetric at the point $w_0 = 1/2$ provided that the PDF of $Z_\rho$ is symmetric at the point $z_0 = 0$. Further, in this case, the median and the mean of a ULS distribution both occur at the point $w_0 = 1/2$.

Proof. A simple algebraic manipulation shows that, $f_W((1/2) - w; \sigma_\rho) = f_W((1/2) + w; \sigma_\rho), \forall 0 < w < 1$, provided that $f_{Z_\rho}(-z) = f_{Z_\rho}(z), \forall z \in \mathbb{R}$. Then the proof follows.

Remark 3.4. Since the PDFs of $Z_1$ and $Z_2$ are even functions (Proposition 1; Item ii, of Saulo et al., 2022), and provided that $Z_1$ and $Z_2$ are independent, we have that $f_{Z_\rho}(-z) = f_{Z_\rho}(z), \forall z \in \mathbb{R}$.

3.2.1 Unit-log-normal

It is widely known that the bivariate log-normal distribution admits a stochastic representation as in (3.3), where $Z_1$ and $Z_2$ are independent and identically distributed (i.i.d.) standard normal random variables (Corollary 1 of Saulo et al., 2022). Consequently, $Z_\rho$ in (3.4) is standard normally distributed. Hence, by using (3.6) and (3.8) we obtain

$$F_W(w; \sigma_\rho) = \Phi(A(w)), \quad f_W(w; \sigma_\rho) = \frac{1}{w(1-w)\sigma_\rho} \phi(A(w)), \quad 0 < w < 1. \quad (3.9)$$

Here, $\phi$ and $\Phi$ denote the PDF and CDF of a standard normal distribution, respectively.
3.2.2 Unit-log-Student-t

It is well-known that the bivariate log-Student-t distribution has a stochastic representation as in (3.3), where \( Z_1 = Z_1^* \sqrt{\nu/Q} \sim t_\nu \) and \( Z_2 = Z_2^* \sqrt{\nu/Q} \sim t_\nu \) (Corollary 2 of Saulo et al., 2022). Here, \( Q \sim \chi^2_\nu \) (chi-square with \( \nu \) degrees of freedom) is independent of \( Z_1^* \) and \( \rho Z_1^* + \sqrt{1 - \rho^2} Z_2^* \); whereas \( Z_1^* \) and \( Z_2^* \) are i.i.d. standard normal random variables.

Since \( \left[ \frac{1}{\sqrt{2(1 - \rho)}} \right] [(1 - \rho) Z_1^* - \sqrt{1 - \rho^2} Z_2^*] \sim N(0, 1) \) we have

\[
Z_\rho = \frac{1}{\sqrt{2(1 - \rho)}} [(1 - \rho) Z_1^* - \sqrt{1 - \rho^2} Z_2^*] \sqrt{\nu} \sim t_\nu. \tag{3.10}
\]

By combining (3.6) and (3.8) with (3.10) we get

\[
F_W(w; \sigma_\rho) = F_\nu(A(w)), \quad f_W(w; \sigma_\rho) = \frac{1}{w(1 - w)\sigma_\rho} f_\nu(A(w)), \quad 0 < w < 1.
\]

Here, \( f_\nu \) and \( F_\nu \) denote the PDF and CDF of a Student-\( t \) distribution with \( \nu \) degrees of freedom, respectively.

3.2.3 Unit-log-Laplace

By using the general algorithm for simulation of symmetric bivariate Laplace variables (Kotz et al., 2001, Subsection 5.1.4, p. 234), it follows that the bivariate log-Laplace distribution has a stochastic representation as in (3.3), where \( Z_1 = Z_1^* \sqrt{B} \sim \text{Laplace}(0, 1/\sqrt{2}) \) and \( Z_2 = Z_2^* \sqrt{B} \sim \text{Laplace}(0, 1/\sqrt{2}) \). Here, \( B \) is a standard exponential variable independent of \( Z_1^* \) and \( \rho Z_1^* + \sqrt{1 - \rho^2} Z_2^* \); while \( Z_1^* \) and \( Z_2^* \) are i.i.d. standard normal random variables.

Since \( \left[ \frac{1}{\sqrt{2(1 - \rho)}} \right] [(1 - \rho) Z_1^* - \sqrt{1 - \rho^2} Z_2^*] \sim N(0, 1) \) we have

\[
Z_\rho = \frac{1}{\sqrt{2(1 - \rho)}} [(1 - \rho) Z_1^* - \sqrt{1 - \rho^2} Z_2^*] \sqrt{B} \sim \text{Laplace}(0, 1/\sqrt{2}). \tag{3.11}
\]

Then, by (3.6), (3.8) and (3.10) we obtain

\[
F_W(w; \sigma_\rho) = F_\ell(A(w)), \quad f_W(w; \sigma_\rho) = \frac{1}{w(1 - w)\sigma_\rho} f_\ell(A(w)), \quad 0 < w < 1, \tag{3.12}
\]

where \( F_\ell \) and \( f_\ell \) denote the CDF and PDF of Laplace distribution with location parameter 0 and scale parameter \( 1/\sqrt{2} \).

Figure 1 displays different shapes of the unit-log-symmetric PDFs for different choices of parameters. From this figure, we observe that the parameter \( \sigma_\rho \) controls the shape of the unit-log-normal, unit-log-Laplace and unit-log-Student-\( t \) densities. We also note that the PDFs of the unit-log-normal, unit-log-Laplace and unit-log-Student-\( t \) models are symmetric. Finally, from Figure 1(d), we see the clear effect the kurtosis parameter \( \nu \) has on the unit-log-Student-\( t \) density.
3.3 Modality

Differentiating $A$ in (3.5) with respect to $w$ yields

$$A'(w) = \frac{1}{w(1-w)\sigma_{\rho}} \quad \text{and} \quad A''(w) = (2w - 1)[A'(w)]^2\sigma_{\rho}. \quad (3.13)$$

Suppose that the derivative of $f_{W}(w;\sigma_{\rho})$ with respect to $w$ exists at each point in its domain. In the case that there is a countable number of points on its domain where $f_{W}$ is not differentiable, by definition, these ones are considered critical points of $f_{W}$. Hence, from (3.8) we have $f_{W}(w;\sigma_{\rho}) = A'(w)f_{Z_{\rho}}(A(w))$ and then

$$f'_{W}(w;\sigma_{\rho}) = A''(w)f_{Z_{\rho}}(A(w)) + [A'(w)]^2f'_{Z_{\rho}}(A(w)).$$
If \( f'_{Z_{\rho}}(z) = -r(z)f_{Z_{\rho}}(z) \), for some real-valued function \( r \), then
\[
f'_W(w; \sigma_{\rho}) = \{A''(w) - [A'(w)]^2r(A(w))\}f_{Z_{\rho}}(A(w)) \quad (3.13)
\]
Hence,
\[
f'_W(w; \sigma_{\rho}) = 0 \iff (2w - 1)\sigma_{\rho} = r(A(w)) \iff [2A^{-1}(A(w)) - 1] \sigma_{\rho} = r(A(w)).
\]
Therefore, by using (3.7), a critical point \( w \) of the ULS PDF (2.1) satisfies the equation:
\[
\left[\frac{2}{1 + \exp(-\sigma_{\rho}A(w))} - 1\right] \sigma_{\rho} = r(A(w)),
\]
or, equivalently,
\[
tanh\left(\frac{\sigma_{\rho}^2}{2} y\right) = \frac{r(\sigma_{\rho} y)}{\sigma_{\rho}}, \quad \text{with } y = A(w)/\sigma_{\rho}. \quad (3.14)
\]

### 3.3.1 Unit-log-normal

In this case, by Subsubsection 3.2.1, we have \( Z_{\rho} \sim N(0, 1) \), and then \( r(z) = z \). Therefore, by (3.14), the equation of critical points of the ULS PDF is given by
\[
tanh\left(\frac{\sigma_{\rho}^2}{2} y\right) = y, \quad \text{with } y = A(w)/\sigma_{\rho},
\]
if and only if
\[
\frac{\sigma_{\rho}^2}{2} y = \arctanh(y). \quad (3.15)
\]
Note that there always exists at least one solution of the above equation. The number of solutions of (3.15) depends on whether \( \sigma_{\rho} \) is larger or smaller than \( \sqrt{2} \). If \( \sigma_{\rho} \leq \sqrt{2} \) then (3.15) has a unique solution, given by \( y_0 = 0 \) (Figure 2), and if \( \sigma_{\rho} > \sqrt{2} \) then (3.15) has two additional non-trivial solutions (Figure 3), \( y_- = y_-(\sigma_{\rho}) < 0 \) and \( y_+ = y_+(\sigma_{\rho}) > 0 \) (which depend on \( \sigma_{\rho} \)), with \( y_+ = -y_- \). That is, \( f_W(w; \sigma_{\rho}) \) has one critical point at \( w_0 = 1/2 \), or three critical points, \( w_- = A^{-1}(\sigma_{\rho}y_-), \quad w_0 = 1/2 \) and \( w_+ = A^{-1}(\sigma_{\rho}y_+) \). Since \( \lim_{w \to 0^+} f_W(w; \sigma_{\rho}) = \lim_{w \to 1^-} f_W(w; \sigma_{\rho}) = 0 \), we obtain the following result.

**Theorem 3.5.** Let \( W \sim ULS(\sigma_{\rho}, g_c) \), with \( g_c(x) = \exp(-x/2) \). The following holds:

- If \( \sigma_{\rho} \leq \sqrt{2} \) then the ULS PDF \( f_W(w; \sigma_{\rho}) \) is unimodal, with mode \( w_0 = 1/2 \). On the other hand, if \( \sigma_{\rho} > \sqrt{2} \) then the ULS PDF \( f_W(w; \sigma_{\rho}) \) is bimodal, with modes
  \[
  w_- = A^{-1}(\sigma_{\rho}y_-) = \frac{1}{1 + \exp(-\sigma_{\rho}^2y_-)} \quad \text{and} \quad w_+ = A^{-1}(\sigma_{\rho}y_+) = \frac{1}{1 + \exp(-\sigma_{\rho}^2y_+)},
  \]
  and minimum point \( w_0 = 1/2 \), so that \( 0 < w_- < w_0 < w_+ < 1 \). Further, the graph of \( f_W(w; \sigma_{\rho}) \) is symmetric with respect to the point \( w_0 = 1/2 \) (Figure 1 a).
3.3.2 Unit-log-Student-\(t\)

In this case, by Subsubsection 3.2.2, we have \(Z_{\rho} \sim t_{\nu}\), and then \(r(z) = (\nu + 1)z/(\nu + z^2)\). Therefore, by (3.14), the equation of critical points of the ULS PDF is given by

\[
\tanh \left( \frac{\sigma_{\rho}^2}{2} y \right) = \frac{(\nu + 1)y}{\nu + \sigma_{\rho}^2 y^2}, \quad \text{with } y = A(w)/\sigma_{\rho},
\]

or, equivalently,

\[
\tanh(z) = \frac{2(\nu + 1)z}{\nu \sigma_{\rho}^2 + 4z^2}, \quad \text{with } z = \sigma_{\rho}^2 y/2. \tag{3.16}
\]

Note that \(w_0 = 1/2\) is a critical point of the ULS PDF because \(z_0 = 0\) satisfies (3.16).

Since \(\lim_{z \to \pm\infty} \tanh(z) = \pm 1\), by an analysis of the graphs of functions \(z \mapsto \tanh(z)\) and \(2(\nu + 1)z/(\nu \sigma_{\rho}^2 + 4z^2)\), we have that the Equation (3.16) has one solution, given by \(z_0 = 0\), or
three solutions, $z_- < z_0 = 0 < z_+ = -z_-$, that depend on the choice of parameters $\nu$ and $\sigma_\rho$ (Figures 4 and 5). That is, $f_W(w; \sigma_\rho)$ has one critical point at $w_0 = 1/2$, or three critical points, $0 < w_- = A^{-1}(2z_-/\sigma_\rho) < w_0 = 1/2 < w_+ = A^{-1}(2z_+/\sigma_\rho) < 1$. Since $\lim_{w_0^+} f_W(w; \sigma_\rho) = \lim_{w_0^-} f_W(w; \sigma_\rho) = \infty$ and $w_0 = 1/2$ is a critical point (and point of symmetry, see Proposition 3.3) of $f_W(w; \sigma_\rho)$, we get the following result.

**Theorem 3.6.** If $W \sim ULS(\sigma_\rho, g_c)$, with $g_c(x) = (1+x/\nu)^{-(\nu+2)/2}$, then the ULS PDF $f_W(w; \sigma_\rho)$ has a bathtub shape with minimum point $w_0 = 1/2$ or is decreasing-increasing-decreasing-increasing with minimum points

$$w_- = A^{-1}(2z_-/\sigma_\rho) = \frac{1}{1 + \exp(-2z_-)} \quad \text{and} \quad w_+ = A^{-1}(2z_+/\sigma_\rho) = \frac{1}{1 + \exp(-2z_+)}$$

and maximum point $w_0 = 1/2$. Further, the graph of the ULS PDF is symmetric with respect to the point $w_0 = 1/2$ (Figure 1 c, d).

### 3.3.3 Unit-log-Laplace

By Subsubsection 3.2.3, $Z_\rho \sim \text{Laplace}(0, 1/\sqrt{2})$, and then $r(z) = \sqrt{2}z/|z|$, $z \neq 0$. The function $f_W(w; \sigma_\rho)$ is not differentiable at $w_0 = 1/2$, then this point is a critical point with $f_W(w_0; \sigma_\rho) = 4/\sigma_\rho$. So the Equation (3.14) of critical points (excluding $w_0$) of the ULS PDF is given by

$$\tanh\left(\frac{\sigma_\rho^2 y}{2}\right) = \frac{\sqrt{2}}{\sigma_\rho} \frac{y}{|y|} \quad \text{with} \quad y = A(w)/\sigma_\rho,$$

which implies that $y_\pm = \pm 2\arctanh(\sqrt{2}/\sigma_\rho)/\sigma_\rho^2$. In other words, $f_W(w; \sigma_\rho)$ has one critical point at $w_0 = 1/2$, or three critical points, $0 < w_- = A^{-1}(\sigma_\rho y_-) < w_0 = 1/2 < w_+ = A^{-1}(\sigma_\rho y_+) < 1$. Since

$$\lim_{w_0^+} f_W(w, \sigma_\rho) = \lim_{w_0^-} f_W(w, \sigma_\rho) = \begin{cases} 0, & \sigma_\rho < \sqrt{2}, \\ 1/\sqrt{2}, & \sigma_\rho = \sqrt{2}, \\ \infty, & \sigma_\rho > \sqrt{2}, \end{cases}$$

and $w_0 = 1/2$ is a critical point of symmetry (Proposition 3.3) of $f_W(w; \sigma_\rho)$, we obtain the following result.

**Theorem 3.7.** Let $W \sim ULS(\sigma_\rho, g_c)$, with $g_c(x) = K_0(\sqrt{2}x)$. The following holds:

If $\sigma_\rho \leq \sqrt{2}$ then the ULS PDF $f_W(w; \sigma_\rho)$ is unimodal, with mode $w_0 = 1/2$. On the other hand, if $\sigma_\rho > \sqrt{2}$ then the ULS PDF $f_W(w; \sigma_\rho)$ is decreasing-increasing-decreasing-increasing with minimum points

$$w_- = A^{-1}(\sigma_\rho y_-) = \frac{1}{1 + \exp(2\arctanh(\sqrt{2}/\sigma_\rho))} \quad \text{and} \quad w_+ = A^{-1}(\sigma_\rho y_+) = \frac{1}{1 + \exp(-2\arctanh(\sqrt{2}/\sigma_\rho))},$$

and maximum point $w_0 = 1/2$, so that $0 < w_- < w_0 < w_+ < 1$. Further, $f_W(w; \sigma_\rho)$ is not differentiable and symmetric at $w_0 = 1/2$ (Figure 1 b).
3.4 Quantile function

From (3.6), the $p$-quantile function $Q_W(p)$ of $W \sim \text{ULS}(\sigma_\rho, g_c)$, for $p \in (0, 1)$ given, satisfies

$$p = F_W(Q_W(p); \sigma_\rho) = \mathbb{P}(Z_\rho \leq A(Q_W(p))),$$

where $Z_\rho$ is as in (3.4). Hence,

$$Q_{Z_\rho}(p) = A(Q_W(p)) \iff Q_W(p) = A^{-1}(Q_{Z_\rho}(p)) = \frac{1}{1 + \exp(-\sigma_\rho Z_\rho(p))}. \quad (3.18)$$

Since the quantiles of Gaussian, Student-$t$, and Laplace variables are well-known, by applying the results of Subsubsections 3.2.1, 3.2.2, and 3.2.3 on the distribution of $Z_\rho$, we have the quantiles Table 2.

| Distribution            | $g_c$                         | $Q_W(p)$                                               |
|-------------------------|-------------------------------|-------------------------------------------------------|
| Unit-log-normal         | $\exp(-x/2)$                  | $[1 + \exp(-\sigma_\rho \Phi^{-1}(p))]^{-1}$          |
| Unit-log-Student-$t$    | $(1 + x)^{-(\nu+2)/2}$        | $[1 + \exp(-\sigma_\rho F^{-1}_\nu(p))]^{-1}$         |
| Unit-log-Laplace        | $K_0(\sqrt{2x})$              | $[1 + (2p)^{-\sigma_\rho/\sqrt{2}}]^{-1}\mathbb{I}_{(0,1/2]}(p) + [1 + (2 - 2p)^{\sigma_\rho/\sqrt{2}}]^{-1}\mathbb{I}_{[1/2,\infty]}(p)$ |

Here, $\Phi^{-1}$ and $F^{-1}_\nu$ denote the respective quantiles of the standard normal and Student-$t$ distributions.

3.5 Moments

Since for $W \sim \text{ULS}(\sigma_\rho, g_c)$, $W = T_1/(T_1 + T_2)$, with $(T_1, T_2) \sim \text{BLS}(\theta, g_c)$, its is clear that $0 \leq \mathbb{E}(W^r) \leq 1$, $r > 0$. Therefore, the positive moments of $W$ always exist.

In general, the real moments of $W$ admit the following representation:

$$\mathbb{E}(W^r) = \mathbb{E}([A^{-1}(Z_\rho)]^r) \overset{(3.7)}{=} \mathbb{E} \left[ \frac{1}{(1 + \exp(-\sigma_\rho Z_\rho))^r} \right], \quad r \in \mathbb{R},$$

where $Z_\rho$ is given in (3.4).

Notice that, by using a binomial expansion, the negative integer moments of $W$ can be written as

$$\mathbb{E}(W^{-n}) = \mathbb{E} \left[ (1 + \exp(-\sigma_\rho Z_\rho))^n \right] = \sum_{k=0}^{n} \binom{n}{k} M_{Z_\rho}(-k\sigma_\rho), \quad n \in \mathbb{N},$$

where $M_{Z_\rho}(t)$ is the moment generating function (MGF) of $Z_\rho$.

Since the MGFs of Gaussian and Laplace variables exist, by using the results of Subsubsections 3.2.1 and 3.2.3 on the distribution of $Z_\rho$, we have the negative moments corresponding to unit-log-normal and unit-log-Laplace (Table 3). On the other hand, since $Z_\rho \sim t_\nu$ as $g(x) = (1 +
\[
x/\nu - (\nu+2)/2 \quad \text{(Subsubsection 3.2.2)}, \text{it is clear that the negative moments corresponding to unit-log-Student-} t \text{ does not exist.}
\]

**Tab. 3:** Density generators \((g_c)\) and negative moments for some ULS distributions.

| Distribution          | \(g_c\)                                                                 | \(\mathbb{E}(W^{-n})\)                                                                 | Restriction |
|-----------------------|--------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|-------------|
| Unit-log-normal       | \(\exp(-x/2)\)                                                          | \(\sum_{k=0}^{n} \binom{n}{k} \exp\left(\frac{1}{2} k^2 \sigma^2_\rho\right)\)          | -           |
| Unit-log-Student-\(t\) | \((1 + \frac{x}{\nu})^{-(\nu+2)/2}\)                                  | \(\#\)                                                                                | -           |
| Unit-log-Laplace      | \(K_0(\sqrt{2x})\)                                                     | \(\sum_{k=0}^{n} \binom{n}{k} (1 - \frac{1}{2} k^2 \sigma^2_\rho)^{-1}\) \(n < \sqrt{2}/\sigma_\rho, \sigma_\rho \leq \sqrt{2}/2\) |             |

### 3.6 Shannon entropy

The entropy of a random variable can be interpreted as the average level of uncertainty inherent to the variable’s possible outcomes. Following Shannon and Weaver (1949), the entropy \(H\) of a random variable \(W \sim ULS(\sigma_\rho, g_c)\), which takes values in the interval \((0, 1)\) and is distributed according to \(f_W(\cdot; \sigma_\rho)\) in (3.8), is defined as

\[
H(W) = -\mathbb{E}[\log f_W(W, \sigma_\rho)].
\]

By using the Formula (3.8) and the representation in (3.3), it is an easy task to see that

\[
H(W) = 2\mu + \log(\sigma_\rho) + \sigma(1 + \rho)\mathbb{E}(Z_1) + \sqrt{1 - \rho^2}\mathbb{E}(Z_2)] + H(Z_\rho),
\]

where \(Z_1\) and \(Z_2\) are given in Subsection 3.2, and \(Z_\rho\) is as in (3.4). Since \(Z_1\) and \(Z_2\) are identically distributed (Proposition 1; Item ii, of Saulo et al., 2022), the above identity is written as

\[
H(W) = 2\mu + \log(\sigma_\rho) + \sigma(1 + \rho + \sqrt{1 - \rho^2})\mathbb{E}(Z_1) + H(Z_\rho).
\]

It is knowledge that the entropies of Gaussian, Student-\(t\), and Laplace variables exist. Then, by using the results of Subsubsections 3.2.1, 3.2.2, and 3.2.3 on the distribution of \(Z_\rho\), we have the entropy Table 4.

**Tab. 4:** Density generators \((g_c)\) and entropies for some ULS distributions.

| Distribution          | \(g_c\) | \(H(W)\)                                                                                     |
|-----------------------|---------|-----------------------------------------------------------------------------------------------|
| Unit-log-normal       | \(\exp(-x/2)\) | \(2\mu + \log(\sigma_\rho) + \frac{1}{2} \log(2\pi) + 1\)                                  |
| Unit-log-Student-\(t\) | \((1 + \frac{x}{\nu})^{-(\nu+2)/2}\) | \(2\mu + \log(\sigma_\rho) + \frac{\nu+1}{2} \left[\Psi\left(\frac{\nu+1}{2}\right) - \Psi\left(\frac{\nu}{2}\right)\right] + \log\left[\sqrt{\pi} B\left(\frac{\nu}{2}, \frac{1}{2}\right)\right]\) |
| Unit-log-Laplace      | \(K_0(\sqrt{2x})\) | \(2\mu + \log(\sigma_\rho) + \log(\sqrt{2\pi})\)                                           |

Here, \(\Psi(x) = \Gamma'(x)/\Gamma(x)\) is the digamma function and \(B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x + y)\) is the beta function, and \(\Gamma(t) = \int_0^\infty x^{t-1} \exp(-x) \, dx, \, t > 0\), is the gamma function.
3.7 Maximum likelihood estimation

Let \( \{W_i : i = 1, \ldots, n\} \) be a univariate random sample of size \( n \) from the ULS(\( \sigma, \rho \)) distribution with PDF as given in (3.8), and let \( w_i \) be the correspondent observations of \( W_i \). Then, the log-likelihood function for \( \sigma, \rho \) is written as

\[
\ell(\sigma, \rho) = -n \log(\sigma) - \sum_{i=1}^{n} \log[w_i(1 - w_i)] + \sum_{i=1}^{n} \log f_{Z_{\rho}}(A(w_i)),
\]

where the random variable \( Z_{\rho} \) is defined in (3.4) and \( A \) is as in (3.5). In the case that a supremum \( \hat{\sigma}, \hat{\rho} \) exists, it must satisfy

\[
\frac{\partial \ell(\sigma, \rho)}{\partial \sigma} \bigg|_{\sigma = \hat{\sigma}, \rho = \hat{\rho}} = 0 \quad \text{and} \quad \frac{\partial^2 \ell(\sigma, \rho)}{\partial \sigma^2} \bigg|_{\sigma = \hat{\sigma}, \rho = \hat{\rho}} < 0,
\]

with

\[
\frac{\partial \ell(\sigma, \rho)}{\partial \sigma} = -\frac{n}{\sigma} - \frac{1}{\sigma^2} \sum_{i=1}^{n} \log \left( \frac{w_i}{1 - w_i} \right) G(A(w_i));
\]

\[
\frac{\partial^2 \ell(\sigma, \rho)}{\partial \sigma^2} = -\frac{1}{\sigma} \left[ 2 \frac{\partial \ell(\sigma, \rho)}{\partial \sigma} + \frac{n}{\sigma} \right] + \frac{1}{\sigma^2} \sum_{i=1}^{n} \log^2 \left( \frac{w_i}{1 - w_i} \right) G'(A(w_i)).
\]

Here, we are adopting the notation:

\[
G(x) = f'_{Z_{\rho}}(x) / f_{Z_{\rho}}(x).
\]

Notice that the likelihood equation \( \frac{\partial \ell(\sigma, \rho)}{\partial \sigma} \bigg|_{\sigma = \hat{\sigma}, \rho = \hat{\rho}} = 0 \) in (3.19) can be written as follows

\[
-\frac{1}{n} \sum_{i=1}^{n} \log \left( \frac{w_i}{1 - w_i} \right) G(A(w_i)) \bigg|_{\sigma = \hat{\sigma}, \rho = \hat{\rho}} = \hat{\sigma}_\rho.
\]

Any nontrivial root \( \hat{\sigma}_\rho \) of the above equation is known as a ML estimate of \( \sigma_\rho \) in the loose sense. When the parameter value provides the absolute maximum of the log-likelihood function, it is called an ML estimate in the strict sense.

3.7.1 Unit-log-normal

By Subsubsection 3.2.1, \( Z_{\rho} \sim N(0, 1) \). Hence, \( G(x) \) in (3.21) and its derivative are given by \( G(x) = -x \) and \( G'(x) = -1 \). Thus, by using (3.22), the ML estimate of \( \sigma_\rho^2 \) in the loose sense is given by

\[
\hat{\sigma}_\rho^2 = \frac{1}{n} \sum_{i=1}^{n} \log^2 \left( \frac{w_i}{1 - w_i} \right).
\]
By using (3.20), it is easy to verify that \( \frac{\partial^2 \ell(\sigma^2)}{\partial \sigma^2} \bigg|_{\sigma^2 = \hat{\sigma}^2} < 0 \). Then, \( \hat{\sigma}^2 \) is an ML estimate in the strict sense.

Let \( \hat{\sigma}^2 \) be the corresponding ML estimator of \( \sigma^2 \). By using Proposition 3.2, we have \( A(W) = Z_\rho \sim N(0, 1) \), for \( W \sim ULS(\sigma_\rho, g_c) \). Then (for \( k = 1, 2, \ldots \))

\[
\mathbb{E}[(\hat{\sigma}^2)^k] = \mathbb{E} \left[ \log^2 \left( \frac{W}{1-W} \right) \right] = \sigma_\rho^{2k} \mathbb{E} [A^{2k}(W)] = \sigma_\rho^{2k} \mathbb{E} (Z_\rho^{2k}) = \sigma_\rho^{2k} \frac{(2k)!}{2^k k!}.
\]

(3.23)

**Proposition 3.8.** Let \( W_1, \ldots, W_n \) be independent and identically distributed random variables with the unit-log-normal distribution (3.9), where \( \sigma_\rho \) is unknown. Then the ML estimator \( \hat{\sigma}^2 \) of \( \sigma_\rho \) is

1. unbiased;
2. consistent;
3. asymptotically normal; that is, \( \sqrt{n}(\hat{\sigma}^2 - \sigma^2) \) converges in distribution to a normal distribution with mean zero and variance \( 2\sigma^2 \);
4. efficient.

**Proof.** By taking \( n = 1 \) in (3.23), the unbiasedness of \( \hat{\sigma}^2 \) follows.

By the Strong Law of Large Numbers and the Central Limit Theorem, it follows that \( \hat{\sigma}^2 \) is consistent and asymptotically normal.

Finally, since the variance of \( \hat{\sigma}^2 \) coincides with the Cramér-Rao lower bound \( [nI(\sigma_\rho)]^{-1} \), where \( I(\sigma_\rho) = -\mathbb{E}[\partial^2 \log f_W(w; \sigma_\rho)/\partial (\sigma_\rho^2)] = 1/(2\sigma^2) \) is the Fisher information in one observation from \( f_W(w; \sigma_\rho) \), the efficiency of \( \hat{\sigma}^2 \) follows.

\( \square \)

### 3.7.2 Unit-log-Student-\( t \)

By Subsubsection 3.2.2, \( Z_\rho \sim t_\nu \). Then, \( G(x) \) in (3.21) and its derivative are given by \( G(x) = -(\nu + 1)x/(\nu + x^2) \) and \( G'(x) = -(\nu + 1)(\nu - x^2)/(\nu + x^2)^2 \). Therefore, by using (3.22), the ML estimate of \( \sigma_\rho^2 \) in the loose sense satisfies

\[
\frac{\nu + 1}{n} \sum_{i=1}^{n} \frac{\log^2 \left( \frac{w_i}{1-w_i} \right)}{\nu \hat{\sigma}^2 + \log^2 \left( \frac{w_i}{1-w_i} \right)} = 1.
\]

(3.24)

**Proposition 3.9.** Any ML estimate of \( \sigma_\rho^2 \) in the loose sense is an ML estimate in the strict sense.

**Proof.** Suppose that \( \hat{\sigma}^2 \) is an ML estimate of \( \sigma_\rho^2 \) in the loose sense. From (3.20) it follows that \( \hat{\sigma}^2 \) is an ML estimate in the strict sense whenever

\[
\frac{\nu + 1}{n} \sum_{i=1}^{n} \log^2 \left( \frac{w_i}{1-w_i} \right) \frac{\nu \hat{\sigma}^2 - \log^2 \left( \frac{w_i}{1-w_i} \right)}{[\nu \hat{\sigma}^2 + \log^2 \left( \frac{w_i}{1-w_i} \right)]^2} > -1.
\]
By using the partial fraction expansion
\[
\frac{x - y}{(x + y)^2} = \frac{2x}{(x + y)^2} - \frac{1}{x + y},
\]
the above inequality is written as
\[
\frac{2\nu(\nu + 1)\hat{\sigma}^2_\rho}{n} \sum_{i=1}^{n} \log^2 \left( \frac{w_i}{1 - w_i} \right) - \frac{\nu + 1}{n} \sum_{i=1}^{n} \nu \hat{\sigma}^2_\rho + \log^2 \left( \frac{w_i}{1 - w_i} \right) > -1.
\]
Since \( \hat{\sigma}^2_\rho \) is an ML estimate in the loose sense, it satisfies (3.24). Consequently, the above inequality is true. This completes the proof of the result.

### 3.7.3 Unit-log-Laplace

In this case, note that
\[
f_W(w; \sigma_\rho) = \exp\left[\theta T(w) - \psi(\theta)\right] h(w), \quad W \sim \text{ULS}(\sigma_\rho, g_c), \quad \text{with} \quad \theta = 1/\sigma_\rho,
\]

\[
T(w) = -\sqrt{2} \log \left( \frac{w}{1 - w} \right), \quad \psi(\theta) = -\log(\theta) \quad \text{and} \quad h(w) = \frac{1}{\sqrt{2w(1 - w)}}.
\]

Then, \( f_W \) belongs to the one-parameter exponential family with \( \Theta = \{\theta : \theta > 0\} \) parameter space.

**Proposition 3.10.**

1. For \( \theta \in \Theta^o \), the interior of \( \Theta \), all moments of \( T(W) \) exist, and \( \psi(\theta) \) is infinitely differentiable at any such \( \theta \). Furthermore,
   \[
   \mathbb{E} \left( \sqrt{2} \log \left( \frac{W}{1 - W} \right) \right) = \sigma_\rho \quad \text{and} \quad \text{Var} \left( \sqrt{2} \log \left( \frac{W}{1 - W} \right) \right) = \sigma^2_\rho.
   \]

2. Given an i.i.d. sample of size \( n \) from \( f_W(w; \sigma_\rho) \),
   \[
   -\sqrt{2} \sum_{i=1}^{n} \log \left( \frac{W_i}{1 - W_i} \right)
   \]
   is minimally sufficient.

3. The Fisher information function exists, is finite at all \( \theta \in \Theta^o \) and equals \( I(\theta) = [I(\sigma_\rho)]^{-1} = \sigma^2_\rho \).

**Proof.** The proof follows by direct application of Proposition 16.1 of DasGupta (2008).

**Proposition 3.11.** Let the true \( \theta = \theta_0 \in \Theta^o \). Then, for all large \( n \), with probability 1, a unique ML estimator of \( \sigma_\rho \) exists, is consistent, and is asymptotically normal.
Proof. Since \(\psi''(\theta) = \sigma^2 \rho > 0\), by Theorem 16.1 of DasGupta (2008), a unique ML estimator of \(\theta = 1/\sigma\rho\) exists, is consistent, and is asymptotically normal. Hence, by invariance of the ML estimator, by the continuous application theorem for convergence in probability and by the Delta method, the proof of the result follows.

Further, in the unit-log-Laplace case, the likelihood function for \(\sigma\rho\) is written as

\[
L(\sigma\rho) = \prod_{i=1}^{n} \frac{1}{w_i(1-w_i)\sigma_\rho} f_\ell(A(w_i)),
\]

where \(A\) is given in (3.5) and \(f_\ell\) is the PDF of Laplace distribution with location parameter 0 and scale parameter \(1/\sqrt{2}\). Since \(f_\ell(x)\) is not differentiable at \(x = 0\) we cannot use the log-likelihood function differentiation method as in the previous two cases. In what follows we maximize \(L(\sigma\rho)\).

Indeed, an algebraic manipulation gives

\[
L(\sigma\rho) = \frac{1}{2^{n/2}} \left[ \prod_{i=1}^{n} \frac{1}{w_i(1-w_i)} \right] \exp \left( -\frac{\sqrt{2}}{\sigma\rho} \sum_{i=1}^{n} \log \left( \frac{w_i}{1-w_i} \right) \right).
\]

Since the function \(\sigma\rho \mapsto \exp(-t/\sigma\rho)/\sigma^n\rho\), \(t > 0\), reaches a maximum at \(\sigma\rho = t/n\), the above right-hand expression is at most

\[
\exp(-n) \left[ \prod_{i=1}^{n} \frac{1}{w_i(1-w_i)} \right] \left/ \left[ \frac{1}{n} \sum_{i=1}^{n} \log \left( \frac{w_i}{1-w_i} \right) \right] \right|^n.
\]

Thus, the ML estimate of \(\sigma^2\) is

\[
\hat{\sigma}_\rho^2 = \frac{\sqrt{2}}{n} \sum_{i=1}^{n} \log \left( \frac{w_i}{1-w_i} \right).
\]

This corroborates the uniqueness of the ML estimator evidenced in Proposition 3.11. Moreover, by Proposition 3.10, Item 1, the ML estimator of \(\sigma^2\rho\) is unbiased and consistent.

4 Simulation results

In this section, we present Monte Carlo simulation studies for the unit-log-normal model (the results of the unit-log-Student-t and unit-log-Laplace are omitted here due to space limitations), considering different scenarios of parameters and sample sizes. The first part of the study presents the evaluation of the ML estimates, while the second evaluates the empirical distribution of the generalized Cox-Snell (GCS) and randomized quantile (RQ) residuals, which are given, respectively, by

\[
r_i^{\text{GCS}} = -\log[1 - \hat{F}_W(w_i; \hat{\sigma}_\rho)] \quad \text{and} \quad r_i^{\text{RQ}} = \Phi^{-1}(\hat{F}_W(w_i; \hat{\sigma}_\rho)), \quad i = 1, \ldots, n,
\]
where $\hat{F}_W$ is the fitted unit-log-normal or unit-log-Student-$t$ CDF, and $\hat{\sigma}_\rho$ is the ML estimate of $\sigma_\rho$. When the models are correctly specified, the GCS is asymptotically standard exponentially distributed, while the RQ is asymptotically standard normal distributed. The study of the distributions of these residuals is important because they will be used to assess goodness of fit.

Both studies consider simulated data generated from each one of the unit-log-normal and unit-log-Student-$t$ models according to the following stochastic relation

$$W = A^{-1}(F_W^{-1}(U)) = \frac{1}{1 + \exp(-\sigma_\rho F_W^{-1}(U))},$$

where $F_W$ is the unit-log-normal or unit-log-Student-$t$ CDF and $U \sim U[0,1]$. The Monte Carlo simulation experiments were performed using the R environment; see http://www.r-project.org.

4.1 Maximum likelihood estimation

The simulation scenario considers the following settings: sample size $n \in \{50, 100, 150, 250, 600\}$ and true parameter $\sigma_\rho \in \{0.15, 0.25, 0.50, 1.0, 1.5\}$, with 500 Monte Carlo replications for each sample size. To study the ML estimators, we use compute the bias and root mean square error (RMSE), which are computed as

$$\hat{\text{Bias}}(\hat{\theta}) = \frac{1}{M} \sum_{i=1}^{M} \hat{\theta}^{(i)} - \theta, \quad \hat{\text{RMSE}}(\hat{\theta}) = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (\hat{\theta}^{(i)} - \theta)^2},$$

where $\theta$ and $\hat{\theta}^{(i)}$ are the true parameter value and its respective $i$-th ML estimate, and $M$ is the number of Monte Carlo replicas.

The ML estimation results are shown in Figure 6 wherein the empirical bias and RMSE are all reported. It is possible to see that the simulations produced the expected outcomes. We observe that the bias and RMSE to approach zero as $n$ grows.

Fig. 6: Monte Carlo simulation results for the unit-log-normal model.
4.2 Empirical distribution of residuals

In this subsection, we show the performance of GCS and RQ residuals. We report the empirical means of the following descriptive statistics: mean, median, standard deviation (Sd), coefficient of skewness and coefficient of kurtosis, whose values are expected to be 1, 0.69, 1, 2 and 6, respectively, for the GCS residual, and 0, 0, 1, 0 and 0, respectively, for the RQ residual. From Figures 7 and 8, we note that the considered residuals conform well with their reference distributions, and we can therefore use the GCS and RQ residuals to verify the fit of the proposed models.

Fig. 7: Monte Carlo simulation results of the RQ residuals for the unit-log-normal model.

5 Application to real data

In this section a real-world data set, corresponding to share of the population using internet, is analyzed. Figure 9 shows the share of the population that is accessing the internet for 201 countries of the world in 2015. Here, the internet can be accessed through a computer, mobile phone, games machine, digital TV, etc; see https://ourworldindata.org/internet. From Figure 9, we observe high rates of people online in richer countries and much lower rates in the developing world. For example, three-quarters (74.55%) of people in the US were online, whereas in India 14.9% used the internet.
Fig. 8: Monte Carlo simulation results of the GCS residuals for the unit-log-normal model.

In terms of stochastic representation, the share of the population using internet, represented by the random variable \( W \), can be written as

\[
W = \frac{T_1}{T_1 + T_2}
\]  

(5.1)

where \( T_1 \in \mathbb{R}^+ \) and \( T_2 \in \mathbb{R}^+ \) are two random variables following log-symmetric distributions representing the number of people with internet access and the number of people without internet access, respectively. That is, the sum \( T_1 + T_2 \) represents the total population, and the ratio (5.1) has support in the unit interval \((0, 1)\) with PDF (2.1).

Table 5 provides descriptive statistics for the share of the population using the internet, including the mean, median, standard deviation (SD), coefficient of variation (CV), skewness (CS), (excess) kurtosis (CK), and minimum and maximum values. From this table, we note that the mean is almost equal to the median. In addition, the CV (dispersion around the mean) is less than 100%, indicating a low dispersion of data around the mean. Finally, we note the data show skewness near zero and low degree of kurtosis.

**Tab. 5:** Summary statistics for the internet data.

| Mean    | Median | SD    | CV       | CS       | CK       | minimum | maximum | size |
|---------|--------|-------|----------|----------|----------|---------|---------|------|
| 0.480   | 0.497  | 0.289 | 60.276\% | 0.008    | -1.316   | 0.011   | 0.983   | 201  |

Table 6 presents the ML estimates and SEs for the unit-log-normal, unit-log-Student-\( t \) and
unit-log-Laplace model parameters. This table also reports the log-likelihood value, and Akaike (AIC) and Bayesian information (BIC) criteria. For comparative purpose, the results of the beta model (Ferrari and Cribari-Neto, 2004) are provided as well. The results of Table 6 reveal that the proposed unit-log-normal and unit-log-Student-$t$ models provide better adjustments than the beta model based on the values of log-likelihood, AIC and BIC. In general, the unit-log-normal model provides the best fit.

**Fig. 9:** Share of the population using the internet, 2015.

**Table 6:** ML estimates (with SE in parentheses) and model selection measures for fit to the internet data.

|                | Unit-log-normal | Unit-log-Student-$t$ | Unit-log-Laplace | Beta     |
|----------------|-----------------|----------------------|------------------|----------|
| $\sigma_p$ (or $\mu$) | 1.6645          | 1.5390               | 1.9794           | -0.7386  |
|                | (0.0830)        | (0.0860)             | (0.2341)         | (0.0397) |
| $\nu$ (or $\phi$) | 10              |                      | 2.1860           | 2.1860   |
|                |                 |                      | (0.1849)         | (0.1849) |
| Log-lik.       | 3.2534          | 1.2758               | -21.9354         | 1.1450   |
| AIC            | -4.5068         | -0.5516              | 45.8709          | 1.7098   |
| BIC            | -1.2035         | 2.7517               | 49.1742          | 8.3164   |

Figure 10 shows the QQ plots with simulated envelope of the GCS and RQ residuals for the
unit-log-normal, unit-log-Student-t, unit-log-Laplace and beta models considered in Table 6. We observe that the unit-log-normal model provides better fit than the unit-log-Student-t, unit-log-Laplace and beta models. Figure 11 displays the histogram and superimposed fitted PDFs, and fitted CDFs (empirical CDF in gray).

Fig. 10: QQ plot and its envelope for the GCS (top) and RQ (bottom) residuals in the indicated model for the internet data.

Fig. 11: Histogram of the internet data with fitted PDFs (left) and fitted CDFs (right).
6 Conclusions

In the present paper we proposed a family of flexible models over the interval \((0, 1)\). By adequately defining the density generator we can transform any distribution over the real line into a bounded distribution over the interval \((0, 1)\). In this way we have transformed the normal, the Student-\(t\) and the Laplace distribution into the unit-log-normal, into the unit-log-Student-\(t\) and into the unit-log-Laplace distribution, respectively. These three cases have been intensively studied above. However, many other unit-log-symmetric distributions can be investigated, for example, the unit-log Kotz type, the unit-log Pearson Type VII, the unit-log hyperbolic, the unit-log slash, the unit-log-logistic and the unit-log-power exponential distributions. These may be useful alternatives to “conventional” models like beta generalizations, the Kumaraswamy and the Topp-Leone distribution, for example (see the introduction). It is clear that by suitably choosing of location and scale parameters, distributions over any bounded support can be obtained.

Acknowledgements This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) (Finance Code 001).

Disclosure statement There are no conflicts of interest to disclose.

References

Balakrishnan, N., Lai, C-D., Continuous Bivariate Distributions, Springer-Verlag, New York, 2009.

Bakouch, H. S., Hussain, T., Chesneau, C. Jónas, T., “A notable bounded probability distribution for environmental and lifetime data”, Earth Science Informatics, 15, pp. 1607-1620, 2022.

DasGupta, A., Asymptotic Theory of Statistics and Probability, Springer Texts in Statistics, Springer New York, 2008.

Davison, A. C., Statistical Models (First ed.). Cambridge Series in Statistical and Probabilistic 408 Mathematics, Cambridge University Press, 2008.

Fang, K. T., Kotz, S., Ng, K. W., Symmetric multivariate and related distributions, Chapman and Hall, London, 1990.

Ferrari, S., Cribari-Neto, F., “Beta regression for modelling rates and proportions”, Journal of Applied Statistics, 31, pp. 799-815, 2004.

Jodrá, P., “A bounded distribution derive from the shifted Gompertz law”, Journal of King Saud University-Science, 32, pp. 523-536, 2020.

Korkmaz, M. Ç., “A new heavy-tailed distribution defined on the bounded interval: the logit slash distribution and its application”, Journal of Applied Statistics, 47, pp. 2097-2119, 2020.
Kotz, S., Kozubowski, T. J., Podgórski, K., *The Laplace Distribution and Generalizations*, 2001.

Saulo, H., Vila, R., Cordeiro, S. S., Leiva, V., “Bivariate symmetric Heckman models and their characterization”, *Journal of Multivariate Analysis*, 105097, 2022.

Shannon, C. E., Weaver, W., “The Mathematical Theory of Communication”, Univ of Illinois Press, 1949.