On a Family of Discrete Log-Symmetric Distributions

Helton Saulo¹,² · Roberto Vila¹ · Leonardo Paiva¹ · N. Balakrishnan² · Marcelo Bourguignon³

Accepted: 6 May 2021 / Published online: 27 May 2021
© Grace Scientific Publishing 2021

Abstract
The use of continuous probability distributions has been widespread in problems with purely discrete nature. In general, such distributions are not appropriate in this scenario. In this paper, we introduce a class of discrete and asymmetric distributions based on the family of continuous log-symmetric distributions. Some properties are discussed as well as estimation by the maximum likelihood method. A Monte Carlo simulation study is carried out to evaluate the performance of the estimators, and censored and uncensored data sets are used to illustrate the proposed methodology.

Keywords Discrete distributions · Maximum likelihood methods · Monte Carlo simulation · R software

1 Introduction
Continuous log-symmetric distributions are of particular interest for describing strictly positive and asymmetric data with the possibility of outlier observations; see, for example, [1, 3, 4, 9, 10, 13–15], for some discussions and applications of log-symmetric models. A continuous random variable \( Y \) follows a log-symmetric distribution if its probability density function (PDF) is given by

\[
f_Y(y|\theta) = \frac{(Z_g)^{-1}}{\sqrt{\phi} y} g\left[a_\phi^2(y)\right], \quad a_\phi(y) = \log\left(\frac{y}{\lambda}\right)^{1/\sqrt{\phi}}, \quad y > 0;
\]

where \( \theta = (\lambda, \phi) \), \( \lambda > 0 \) is a scale parameter and also the median of \( Y \), \( \phi > 0 \) is a shape parameter associated with the skewness or relative dispersion, \( Z_g = \int_{-\infty}^{\infty} g(w^2) \, dw \) is the

Marcelo Bourguignon
m.p.bourguignon@gmail.com

¹ Department of Statistics, University of Brasília, Brasília, Brazil
² Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada
³ Department of Statistics, Federal University of Rio Grande do Norte, Natal, Brazil
partition function, and the function \( g \) is a density generating kernel such that \( g(u) > 0 \) for \( u > 0 \). The function \( g \) is associated with an additional parameter \( \xi \) (or vector \( \xi \)). We use the notation \( Y \sim \text{LS}(\theta, g) \). Note that if \( g(u) \) in (1) is \( \exp(-u/2); [1 + (u/\xi)]^{-1} \), \( \xi > 0; \exp[-u^{1/(1+\xi)}/2], -1 < \xi \leq 1; \sqrt{\xi_2} \exp(-\xi_2 u/2) + [(1 - \xi_1)/\xi_1] \exp(-u/2), 0 < \xi_1, \xi_2 < 1; \cosh(u^{1/2}) \exp[-(2/\xi^2) \sinh^2(u^{1/2})], \xi > 0; \) or \( \cosh(u^{1/2}) [\xi_2 \xi_1^2 + 4 \sinh^2(u^{1/2})]^{-1/2} \), \( \xi_1, \xi_2 > 0; \) we have the log-normal, log-Student-\( \nu \), log-power-exponential, log-contaminated-normal, extended Birnbaum-Saunders or extended Birnbaum-Saunders-\( t \) distributions, respectively; see [13, 15]. If \( Y \sim \text{LS}(\theta, g) \), then the associated cumulative distribution function (CDF) is given by \( F_Y(y|\theta) = G[a_g(y)] \), where the function \( G : (-\infty, +\infty) \to [0, 1] \) is defined as

\[
G(r) = (Z_g)^{-1} \int_{-\infty}^{r} g(z^2) \, dz, \quad -\infty < r < +\infty. \tag{2}
\]

This mapping is easily seen to have the following properties:

(a) \( G(0) = 0.5, G(+\infty) = \lim_{r \to +\infty} G(r) = 1, G(-\infty) = \lim_{r \to -\infty} G(r) = 0; \)

(b) \( G(\cdot) \) is a continuous function and that \( G(\cdot) \) is strictly monotonically increasing.

Hence \( G(\cdot) \) has an inverse function, denoted by \( G^{-1}(\cdot); \)

(c) \ From Items (a) and (b), \( G(\cdot) \) is a CDF; and

(d) \( G^{-1}(1-p) = -G^{-1}(p) \) for \( p \in (0, 1) \) given.

Despite the huge use of log-symmetric distributions—its most famous member is the log-normal model—they are not appropriate in purely discrete contexts. For example, to model the number of cycles before failure of a equipment or the number of weeks to cure a patient, among others; see [16]. Moreover, despite useful, continuous log-symmetric models do not include the zero. In this paper, we define a discrete random variable associated to \( Y \) in (1) as \( X = [Y] \), where \([y]\) denotes the largest integer less than or equal to \( y \). In other words, we propose a class of discrete log-symmetric distributions. The proposed class incorporates every distribution belonging to the log-symmetric family, and it is useful for asymmetric and non-negative discrete data. Furthermore, the proposed class of discrete log-symmetric distributions includes as special cases at least ten new models.

The rest of the paper proceeds as follows. In Sect. 2, we introduce the class of discrete log-symmetric models. In Sect. 3, we discuss some mathematical properties. In Sect. 4, estimation of the model parameters are approached via the maximum likelihood method for the censored and uncensored cases. In Sect. 5, we carry out a simulation study to evaluate the performance of the estimators taking into account different censoring proportions. In Sect. 6, we illustrate the proposed methodology with two real data sets. Finally, in Sect. 7, we make some concluding remarks and discuss future work.
2 Discrete Log-Symmetric Distributions

We say that a discrete random variable \( X \), taking values in the set \( \{0, 1, \ldots\} \), follows a discrete log-symmetric distribution with parameter vector \( \theta = (\lambda, \phi) \), where \( \lambda > 0, \phi > 0 \), denoted by \( X \sim LS_d(\theta, g) \), if its probability mass function (PMF) is given by

\[
p(x|\theta) = G[a_\theta(x+1)] - G[a_\theta(x)], \quad x = 0, 1, \ldots,
\]

(3)

where \( a_\theta(\cdot) \) and \( G(\cdot) \) are as in (1) and (2), respectively. Note that from (2) and (3), we can obtain the following discrete log-symmetric distributions: log-normal, log-Student-\( t \), log-power-exponential, log-contaminated-normal, extended Birnbaum-Saunders and extended Birnbaum-Saunders-\( t \) distributions. Note also that \( G[a_\theta(0)] = G(-\infty) = 0 \) and \( G[a_\theta(+\infty)] = G(+\infty) = 1 \). Given the density generating kernel \( g \), defined below (1), the parameters \( \lambda \) and \( \phi \) completely determine the PMF (3) at \( x = 0 \). Since \( G(\cdot) \) and \( a_\theta(\cdot) \) are strictly increasing functions, and it is clear that \( p(x|\theta) \) is a PMF.

The CDF, reliability function (RF) and hazard rate (HR) of the \( LS_d \) distribution, respectively, are given by

\[
F(x|\theta) = 1 - R(x|\theta) = G[a_\theta([x] + 1)], \quad x \geq 0;
\]

\[
H(x|\theta) = \frac{p(x|\theta)}{p(x|\theta) + R(x|\theta)} = \frac{G[a_\theta(x+1)] - G[a_\theta(x)]}{1 - G[a_\theta(x)]}, \quad x = 0, 1, \ldots.
\]

3 Mathematical Properties

This section, if not explicitly mentioned otherwise, consists of mathematical properties valid for any discrete random variable \( X \) with support \( \{0, 1, \ldots\} \).

Let \( (b_n) \) be a sequence of real numbers. For technical reasons in the next result we use the convention \( \prod_{y=0}^{n-1} b_y = 1 \). The next result provides a characterization of the PMF and RF of a discrete distribution in terms of the HR.

**Proposition 1** If \( X \) is a discrete random variable then, for each \( x = 0, 1, 2, \ldots \),

(a) \( p(x|\theta) = \frac{H(x|\theta)}{1-H(x|\theta)} \prod_{y=0}^{x-1} \left[ 1 - H(y|\theta) \right] \);

(b) \( R(x|\theta) = \prod_{y=0}^{x-1} \left[ 1 - H(y|\theta) \right] \);
where \( H(\cdot | \theta) \) is the HR.

**Proof** By using the identity

\[
p(x|\theta) = R(x|\theta) - R(x+1|\theta) = \left[p(x|\theta) + R(x|\theta)\right] H(x|\theta), \quad x = 0, 1, \ldots,
\]

we have

\[
1 = H(x|\theta) + \frac{R(x|\theta)H(x|\theta)}{p(x|\theta)}, \quad x = 0, 1, 2, \ldots.
\]

Since \( p(x|\theta)/H(x|\theta) = p(x|\theta) + R(x|\theta) = R(x-1|\theta) \), it follows that

\[
\frac{R(x|\theta)}{R(x-1|\theta)} = 1 - H(x|\theta), \quad x = 0, 1, 2, \ldots.
\]

Exchanging \( x \) for \( y \) in the above identity and then multiplying from \( y = 0 \) to \( y = x-1 \), we get

\[
R(x|\theta) = \prod_{y=0}^{x-1} \frac{R(y|\theta)}{R(y-1|\theta)} = \prod_{y=0}^{x-1} \left[1 - H(y|\theta)\right], \quad x = 0, 1, 2, \ldots,
\]

verifying the identity for \( R(x|\theta) \). On the other hand, combining the above identity with the definition of HR, the identity for \( p(x|\theta) \) follows. \( \square \)

### 3.1 Moments and Variance

**Theorem 1** If \( X \) is a discrete random variable possessing all the higher-order moments, then

(a) \( \mathbb{E}(X^r) = \sum_{x=0}^{\infty} [(x + 1)^r - x^r] R(x|\theta); \)

(b) \( \mathbb{E}(X^r) = \sum_{x=0}^{\infty} \sum_{k=0}^{r} \sum_{i=0}^{r-k} \binom{r-k}{i} x^{k+i} R(x|\theta); \)

(c) \( \text{Var}(X) = 2 \sum_{x=0}^{\infty} xR(x|\theta) + \sum_{x=0}^{\infty} R(x|\theta) \left[1 - \sum_{x=0}^{\infty} R(x|\theta)\right]; \)

where \( R(\cdot | \theta) \) is the RF.

**Proof** In order to prove Item (a), using the telescopic series \( \sum_{x=0}^{i-1} [(x + 1)^r - x^r] = i^r \), it follows that

\[
\mathbb{E}(X^r) = \sum_{i=0}^{\infty} \sum_{x=0}^{\infty} 1_{\{x < i\}} [(x + 1)^r - x^r] p(i|\theta) = \sum_{x=0}^{\infty} [(x + 1)^r - x^r] \sum_{i=0}^{\infty} 1_{\{i > x\}} p(i|\theta),
\]

where in the second equality we exchange the orders of the summations because the following series

\[
\sum_{x=0}^{\infty} 1_{\{x < i\}} [(x + 1)^r - x^r] p(i|\theta) = \sum_{x=0}^{\infty} 1_{\{x < i\}} \cdot [(x + 1)^r - x^r] p(i|\theta) = i^r p(i|\theta) =: M_i
\]
is finite for each $i = 0, 1, \ldots$, and, by hypothesis, the expectation $\mathbb{E}(X^r) = \sum_{i=0}^\infty M_i$ always exists. This proves the first item. The second item follows by combining the expression for $\mathbb{E}X^r$ given in the first item with the polynomial identity $a^n - b^n = (a - b) \sum_{k=0}^r a^{r-k} b^k$ and with the binomial expansion. The proof of the third item immediately follows from Item (a). Thus, the proof is complete. 

\[ \square \]

3.2 The $p$-Quantile

Theorem 2 Let $X = \lfloor Y \rfloor$ be a discrete random variable obtained from a positive continuous random variable $Y$ with CDF $F_Y(\cdot|\theta)$. Given $p \in (0, 1)$, let $Q_p = F_Y^{-1}(p|\theta)$ be the $p$-quantile for $Y$. The following statements are valid:

(a) If $Q_p > 0$ is a natural number, then $Q_p - 1$ is the $p$-quantile for $X$;
(b) If $Q_p > 0$ is not a natural number, then all $y \in ([Q_p], [Q_p] + 1)$ is a $p$-quantile for $X$.

Proof Given $p \in (0, 1)$, assume that $Q_p = F_Y^{-1}(p|\theta)$. By using the relations, for all $x > 0$,

$$F_X(x^-|\theta) = F_Y([x]|\theta);$$  \hfill (4)

$$[x] \leq x < [x] + 1;$$  \hfill (5)

where $F_X(x^-|\theta) = \lim_{\delta \to 0} F_X(x - \delta|\theta)$ for all $\delta > 0$, we have

$$F_X([Q_p]^-|\theta) \leq F_X(Q_p - 1|\theta) = F_X(Q_p^-|\theta) = F_Y(Q_p|\theta) = p,$$

whenever $Q_p > 0$ is a natural number. Then, by definition of $p$-quantile for a discrete random variable, the statement in Item (a) follows.

Already, when $Q_p > 0$ is not a natural number, from (4) and (5) we have

$$F_X(Q_p^-|\theta) = F_Y([Q_p]^-|\theta) \leq F_Y(Q_p|\theta) = p;$$

$$F_X(Q_p|\theta) = F_X([Q_p] + 1^-|\theta) = F_Y([Q_p] + 1|\theta) \geq F_Y(Q_p|\theta) = p.$$

Therefore, $F_X(Q_p^-|\theta) \leq p \leq F_X(Q_p|\theta)$. Hence, the $p$-quantile for $X$ can be represented by any value in the interval $([Q_p], [Q_p] + 1)$, and the proof of Item (b) follows. This completes the proof. \[ \square \]

The following two results are applied exclusively to random variables with discrete log-symmetric distribution.
Proposition 2 Let $X$ be a random variable with $\text{LS}_d$ distribution. The following statements hold:

(a) If $\lambda$ is a natural number, then $\lambda - 1$ is the median for $X$;
(b) If $\lambda$ is not a natural number, then the median of the distribution of $X$ can be represented by any value in the set $[\lfloor \lambda \rfloor, \lfloor \lambda \rfloor + 1)$.

**Proof** Let $X \sim \text{LS}_d(\theta, g)$. Since $G(\cdot)$ and $a_\theta(\cdot)$ are strictly increasing functions, the function $G[a_\theta(\cdot)]$ is a strictly increasing CDF corresponding to some continuous random variable $Y$ with log-symmetric distribution $\text{LS}(\theta, g)$. Furthermore, note that the median $Q_{0.5}$ for $Y$ can be written as

$$Q_{0.5} = (G \circ a_\theta)^{-1}(0.5) = a_\theta^{-1}[G^{-1}(0.5)] = \lambda \exp \left[ \sqrt{\phi} G^{-1}(0.5) \right] = \lambda,$$

where in the last equality we use that $G^{-1}(0.5) = 0$; see Item (a) below Item (2) in Sect. 1. Then, by Theorem 2, the proof of Items (a) and (b) follows. \hfill $\Box$

Let $X \sim \text{LS}_d(\theta, g)$. For given $p \in (0, 1)$, let $Q_d; p$ be a $p$-quantile for $X$. Let us define

- **Dispersion:** $\zeta = Q_{d;0.75} - Q_{d;0.25}, \quad 0 < \zeta < \infty$;
- **Relative dispersion:** $\varpi = \frac{\zeta}{\zeta + 2Q_{d;0.25}}, \quad 0 < \varpi < 1$;
- **Skewness:** $\lambda(p) = \frac{Q_{d;p} + Q_{d;1-p} - 2Q_{d;0.5}}{Q_{d;1-p} + Q_{d;p}}, \quad 0 < \lambda(p) < 1, \quad 0 < p < 0.5$;
- **Kurtosis:** $\kappa = \frac{Q_{d;7/8} - Q_{d;5/8} + Q_{d;3/8} - Q_{d;1/8}}{Q_{d;6/8} - Q_{d;2/8}}, \quad 0 \leq \kappa < \infty$.

The relative dispersion, skewness and kurtosis have appeared in [2, 6, 17], respectively.

Proposition 3 Given $p \in (0, 1)$, let $X \sim \text{LS}_d(\theta, g)$ and let $Q_p$ be the $p$-quantile of the corresponding continuous log-symmetric distribution. If $Q_p$ is a natural number, then

(a) $\zeta = 2\lambda \sinh \left[ \sqrt{\phi} G^{-1}(0.75) \right]$;
(b) $\varpi = \left\{ \cotanh \left[ \sqrt{\phi} G^{-1}(0.75) \right] - \coth \left[ \sqrt{\phi} G^{-1}(0.75) \right] \right\}^{-1}$;
(c) $\lambda(p) = \lambda$;
(d) $\kappa = \frac{\sin \left[ \sqrt{\phi} G^{-1}(7/8) \right] - \sin \left[ \sqrt{\phi} G^{-1}(5/8) \right]}{\sin \left[ \sqrt{\phi} G^{-1}(6/8) \right]}$;

where $G(\cdot)$ was defined in (2).
Proof Since $Q_p$ is a natural number, by Theorem 2, $Q_{d:p} = Q_p − 1$ is a $p$-quantile for $X$, where $Q_p = \lambda \exp \left[ \sqrt{\phi} G^{-1}(p) \right]$. By using the identity $G^{-1}(1−p) = −G^{-1}(p)$ (see Item (d) below Item (2) in Sect. 1), from Proposition 2 and a simple algebraic computation, the proof of statements in Items (a–d) follows. □

3.3 Modality Properties

The next result shows that the shape of a discrete log-symmetric distribution depends on choice of density generating kernel and on the distance between the modes of the corresponding continuous log-symmetric distribution.

**Theorem 3** Let $g$ be a density generating kernel so that the corresponding continuous log-symmetric distribution, of a random variable $Y$, is bimodal. Then the discrete log-symmetric distribution of $X = \lfloor Y \rfloor$ has the following shapes:

(a) It is bimodal, whenever the distance between the modes is big enough;
(b) It is unimodal, whenever the distance between the modes is small enough.

**Proof** Since the proof of Item (b) follows the same analysis and steps as the first item, we are concerned with proving only Item (a).

Let $f_Y(y|\theta)$, $\theta > 0$, be the bimodal PDF of the continuous random variable $Y \sim \text{LS}(\theta, g)$, where the distance between their modes, denoted by $y_0 > 0$ and $y_\epsilon = y_0 + \epsilon$, is big enough ($\epsilon > 6$). From bimodality property, there is $y^* \in (y_0, y_\epsilon)$ such that the following inequalities hold:

$$f_Y(y|\theta) \geq f_Y(y-1|\theta) \quad \text{for all } y \leq y_0 \text{ and } y^* \leq y \leq y_\epsilon;$$  \hfill (6)

$$f_Y(y|\theta) \geq f_Y(y+1|\theta) \quad \text{for all } y^* \leq y \leq y_0 \text{ and } y \geq y_\epsilon. \hfill (7)$$

If $x$ is a natural number such that $x \leq \lfloor y_0 \rfloor − 1$ and $\lfloor y^* \rfloor + 1 \leq x \leq \lfloor y_\epsilon \rfloor − 1$, from above inequalities, we have

$$p(x|\theta) = \sum_{y=x}^{x+1} f_Y(y|\theta) \geq \sum_{y=x}^{x+1} f_Y(y-1|\theta) = p(x-1|\theta).$$

Already, if $x$ is a natural number such that $\lfloor y^* \rfloor - 1 \geq x \geq \lfloor y_0 \rfloor + 1$ and $x \geq \lfloor y_\epsilon \rfloor + 1$, we have

$$p(x|\theta) = \sum_{y=x}^{x+1} f_Y(y|\theta) \geq \sum_{y=x}^{x+1} f_Y(y+1|\theta) = p(x+1|\theta).$$

In other words, we have the following...
\[ p(0|\theta) \leq p(1|\theta) \leq \cdots \leq p([y_0] - 1|\theta); \]
\[ p([y_0] + 1|\theta) \geq \cdots \geq p([y_\ast] - 2|\theta) \geq p([y_\ast] - 1|\theta); \]  
(8) 
\[ p([y_\ast] + 1|\theta) \leq p([y_\ast] + 2|\theta) \leq \cdots \leq p([y_\ast] - 1|\theta); \]
\[ p([y_\ast] + 1|\theta) \geq p([y_\ast] + 2|\theta) \geq \cdots . \]  
(9)

From monotonicities (8) and (9), one can guarantee the bimodality property of the discrete log-symmetric distribution. By using (8), we show how to obtain only one of the modes, since the other one can be obtained following a similar path. Indeed, by (8) it remains to relate the probabilities \( p([y_0] - 1|\theta), p([y_0]|\theta) \) and \( p([y_0] + 1|\theta) \) to find the first mode, denoted by \( x_0 \), of the discrete log-symmetric distribution. A simple observation shows that \( x_0 \) is given by

\[
x_0 = \begin{cases} 
[y_0] + 1, & \text{if } p([y_0] - 1|\theta) \leq p([y_0]|\theta) < p([y_0] + 1|\theta), \\
[y_0], & \text{if } p([y_0] - 1|\theta) < p([y_0]|\theta) > p([y_0] + 1|\theta), \\
[y_0] - 1, & \text{if } p([y_0] - 1|\theta) > p([y_0]|\theta) \geq p([y_0] + 1|\theta), \\
[y_0] - 1 \text{ and } [y_0], & \text{if } p([y_0] - 1|\theta) = p([y_0]|\theta) > p([y_0] + 1|\theta), \\
[y_0] \text{ and } [y_0] + 1, & \text{if } p([y_0] - 1|\theta) < p([y_0]|\theta) = p([y_0] + 1|\theta), \\
[y_0], & \text{if } p([y_0] - 1|\theta) = p([y_0] + 1|\theta) < p([y_0]|\theta). 
\end{cases}
\]

Notice that, any other possible relation between \( p([y_0] - 1|\theta), p([y_0]|\theta) \) and \( p([y_0] + 1|\theta) \) contradicts the fact that \( y_0 \) is a mode of the continuous log-symmetric distribution. As mentioned above, using (9), the second mode of the discrete log-symmetric distribution is obtained in an analogous way. So we completed the proof.

\[\square\]

As an immediate consequence of Theorem 3, the following result follows.

**Corollary 1** Let \( g \) be a density generating kernel so that the corresponding continuous log-symmetric distribution, of a random variable \( Y \), is unimodal. Then the discrete log-symmetric distribution of \( X = [Y] \) is also unimodal.

### 4 Maximum Likelihood Estimation

#### 4.1 Uncensored Data

Let \((X_1, \ldots, X_n)\) be a random sample of size \( n \) from a random variable \( X \) with PMF given by (3) and \( x = (x_1, \ldots, x_n) \) their observations (data). Then, the log-likelihood function for a parameter vector \( \theta = (\lambda, \phi) \) is given by
\[
\ell(\theta) = \ell(\theta|x) = \sum_{i=1}^{n} \log[p(x_i|\theta)] = \sum_{i=1}^{n} \log \left\{ G[a_\theta(x_i + 1)] - G[a_\theta(x_i)] \right\}. \quad (10)
\]

The roots of the system formed by the partial derivatives of the log-likelihood function \( \ell(\theta) \) with respect to \( \lambda \) and \( \phi \) are the estimates of these parameters, respectively. Thus, we must solve the following system of equations:

\[
\frac{\partial \ell(\theta)}{\partial \theta} = (Z_g)^{-1} \sum_{i=1}^{n} \sum_{j=0}^{1} (-1)^{j+1} \frac{\partial a_\theta(x_i + j)}{\partial \theta} \frac{g[a_\theta^2(x_i + j)]}{p(x_i|\theta)} = 0, \quad \theta \in \{ \lambda, \phi \},
\]

where \( Z_g = \int_{-\infty}^{\infty} g(w^2)dw \) and,

\[
\frac{\partial a_\theta(x_i)}{\partial \lambda} = -\left( \lambda \phi^{1/2} \right)^{-1}, \quad \frac{\partial a_\theta(x_i)}{\partial \phi} = \log \left( \frac{x_i}{\lambda} \right)^{1/(2\phi^{3/2})}. \quad (11)
\]

Note that they must be solved by an iterative procedure for nonlinear optimization, such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton method; see Mittelhammer et al. [5, p. 199].

Inference for \( \theta \) of the \( \text{LS}_d \) model can be based on the asymptotic distribution of the maximum likelihood estimator \( \hat{\theta} \). Under classic regularity conditions, this estimator is bivariate normal distributed with mean \( \theta \) and covariance matrix \( \Sigma_\theta \), namely,

\[
\sqrt{n} (\hat{\theta} - \theta) \xrightarrow{D} N_2(0, \Sigma_\theta = (J(\theta))^{-1}),
\]

as \( n \to \infty \), where \( \xrightarrow{D} \) means “convergence in distribution”, \( J(\theta) \) is the expected Fisher information matrix, and \( J(\theta) = \lim_{n \to \infty} (1/n) [I(\theta)] \). Observe that \( [J(\theta)]^{-1} \) is a consistent estimator of the asymptotic covariance matrix of \( \hat{\theta} \). Observe also that one may use the Hessian matrix to obtain the observed version of the expected Fisher information matrix.

The Hessian matrix of \( \ell(\theta) \) is given by

\[
[\hat{\ell}_{\theta\theta'}(\theta)]_{2 \times 2} = \begin{bmatrix}
\frac{\partial^2 \ell(\theta)}{\partial \lambda^2} & \frac{\partial^2 \ell(\theta)}{\partial \lambda \partial \phi} \\
\frac{\partial^2 \ell(\theta)}{\partial \phi \partial \lambda} & \frac{\partial^2 \ell(\theta)}{\partial \phi^2}
\end{bmatrix},
\]

where its elements, for each \( \theta, \theta' \in \{ \lambda, \phi \} \), are

\[
\hat{\ell}_{\theta\theta'}(\theta) = (Z_g)^{-1} \sum_{i=1}^{n} \sum_{j=0}^{1} (-1)^{j+1} \left[ \frac{\partial^2 a_\theta(x_i + j)}{\partial \theta \partial \theta'} + \Theta_j(x_i) - \Omega_j(x_i) \right] g[a_\theta^2(x_i + j)] p(x_i|\theta) \frac{g[a_\theta^2(x_i + j)]}{p(x_i|\theta)}.
\]

Here we adopt the following notation:

\[
\Theta_j(x_i) = 2a_\theta(x_i + j) g'[a_\theta^2(x_i + j)] \frac{\partial a_\theta(x_i + j)}{\partial \theta} \frac{\partial a_\theta(x_i + j)}{\partial \theta'}; \quad (12)
\]
\[ \Omega_j(x_i) = (Z_g)^{-1} \frac{\partial a_\theta(x_i + j)}{\partial \theta} \sum_{k=0}^{1} (-1)^{k+1} \frac{\partial a_\theta(x_i + k)}{\partial \theta'} g \left[ a^2(x_i + k) \right] \frac{p(x_i | \theta)}{p(x_i | \theta)}; \]  

whenever the density generating kernel \( g \) be differentiable. The above second-order partial derivatives of \( a_\theta(\cdot) \), with respect to the parameters, are given by

\[
\frac{\partial^2 a_\theta(x_i)}{\partial \lambda^2} = \left( \lambda^2 \phi^2 / 2 \right); \quad \frac{\partial^2 a_\theta(x_i)}{\partial \phi^2} = \log \left( \frac{x_i}{\lambda} \right)^{-3/(4 \phi^2)}; \quad \frac{\partial a_\theta(x_i)}{\partial \lambda} = (2 \lambda \phi^3 / 2)^{-1}; \quad \frac{\partial a_\theta(x_i)}{\partial \phi} = (2 \lambda \phi^3 / 2)^{-1}.
\]

Under certain regularity conditions, the Fisher information matrix

\[ [I_{\theta \theta'}(\Theta)]_{2 \times 2} = - \begin{bmatrix} \mathbb{E} \frac{\partial^2 p(X | \theta)}{\partial \lambda^2} & \mathbb{E} \frac{\partial^2 p(X | \theta)}{\partial \phi \partial \lambda} \\ \mathbb{E} \frac{\partial^2 p(X | \theta)}{\partial \phi^2} & \mathbb{E} \frac{\partial^2 p(X | \theta)}{\partial \lambda^2} \end{bmatrix}, \quad X \sim L_S d(\Theta, g), \]

has elements of the following form

\[ I_{\theta \theta'}(\Theta) = (Z_g)^{-1} \sum_{x=0}^{\infty} \sum_{j=0}^{1} (-1)^j \left[ \frac{\partial^2 a_\theta(x + j)}{\partial \theta \partial \theta'} + \Theta_j(x) - \Omega_j(x) \right] g \left[ a^2(x + j) \right], \]

for each \( \theta, \theta' \in \{ \lambda, \phi \} \), where \( \Theta_j(\cdot) \) and \( \Omega_j(\cdot) \) are given in (12) and (13), respectively, and whenever the above series converges absolutely.

The extra parameter \( \xi \) (or parameter vector \( \xi \)) associated with \( g \) is selected by using the profile log-likelihood function. For instance, in the case of the discrete log-Student-\( t \) distribution, two steps are required:

(i) Let \( \hat{\xi}_k = k \) and for each \( k = 1, \ldots, 100 \) compute the \( k \)-th maximum likelihood estimate of \( \Theta_k = (\hat{\lambda}_k, \hat{\phi}_k)^T \), \( \hat{\Theta}_k = (\hat{\lambda}_k, \hat{\phi}_k)^T \) say. Compute also the \( k \)-th log-likelihood function value \( \ell_k(\hat{\Theta}_k) \);

(ii) The final estimate of \( \hat{\xi} \), \( \hat{\xi} = \hat{\xi}_k \) say, is the one which maximizes the log-likelihood function, that is, \( \hat{\xi} \in \{ \text{argmax}_{\hat{\xi}_k} \ell_k(\hat{\Theta}_k) \} \), and the estimate of \( \Theta \) is \( \hat{\Theta}_k = (\hat{\lambda}_k, \hat{\phi}_k)^T \).

### 4.2 Censored Data

Let \( Y_i \sim L_S(\Theta, g) \) be the failure time of the \( i \)-th individual and let \( \delta_i \) indicate whether the \( i \)-th individual is censored or not. Let us define \( d_k = \text{“number of failures at time } t_k \”, \ q_k = \text{“number censored at time } t_k \” \) and \( N_k = \sum_{i=k}^{\infty} (d_i + q_i) \). Note that \( N_k - d_k \) represents the number survived just before time \( t_k + 1 \). That is, in each given time \( t_k \), there are \( d_k \) failures and \( N_k - d_k \) survivors.

Since the data are discrete observing \( \{(Y_i, \delta_i)\} \) is equivalent to observing \( \{(d_k, q_k)\} \), the likelihood function for the random censoring is given by

\[ g_c \left( Y_i | \Theta, g \right) = \begin{cases} p(Y_i | \Theta, g) & \text{if } \delta_i = 1, \\
0 & \text{otherwise.} \end{cases} \]
\[ L^R(\theta) = \prod_{i=1}^{n} \left[ f_Y(y_i|\theta) \right]^{\delta_i} \left[ 1 - F_Y(y_i|\theta) \right]^{1 - \delta_i} = \prod_{k=1}^{\infty} \left[ p(x_k|\theta) \right]^{d_k} \left[ p(x_k|\theta) + R(x_k|\theta) \right]^{q_k}. \]

This type of censoring has as special case type I and II censoring. The corresponding log-likelihood is

\[ \ell^R(\theta) = \sum_{k=1}^{\infty} \left\{ d_k \log[p(x_k|\theta)] + q_k \log \left[ p(x_k|\theta) + R(x_k|\theta) \right] \right\} \]

\[ = \sum_{k=1}^{\infty} d_k \log \left\{ G[a_\theta(x_k + 1)] - G[a_\theta(x_k)] \right\} + \sum_{k=1}^{\infty} q_k \log \left\{ 1 - G[a_\theta(x_k)] \right\} \]

\[ \tag{15} \]

**Remark 1** By Proposition 1, the log-likelihood (15) can be rewritten in terms of HR as

\[ \ell^R(\theta) = \sum_{k=1}^{\infty} (d_k + q_k) \left( \frac{d_k}{d_k + q_k} \log [H(x_k|\theta)] - \log [1 - H(x_k|\theta)] + \sum_{j=0}^{x_k-1} \log [1 - H(y)|\theta] \right), \]

whenever the above series converges absolutely.

Differentiating in (15), a straightforward computation shows that

\[ \frac{\partial \ell^R(\theta)}{\partial \theta} = (Z_g)^{-1} \sum_{k=1}^{\infty} d_k \sum_{j=0}^{1} (-1)^{j+1} \frac{\partial a_\theta(x_k + j)}{\partial \theta} \frac{g[a_\theta^2(x_k + j)]}{p(x_k|\theta)} \]

\[ - (Z_g)^{-1} \sum_{k=1}^{\infty} q_k \frac{\partial a_\theta^2(x_k)}{\partial \theta} \frac{g[a_\theta^2(x_k)]}{p(x_k|\theta) + R(x_k|\theta)}, \quad \theta \in \{\lambda, \phi\}, \]

where \( Z_g = \int_{-\infty}^{\infty} g(w^2) \, dw \). The mixed partial derivatives of \( \ell^R(\theta) \) are given by

\[ \frac{\partial^2 \ell^R(\theta)}{\partial \theta \partial \theta'} = (Z_g)^{-1} \sum_{k=1}^{\infty} d_k \sum_{j=0}^{1} (-1)^{j+1} \left[ \frac{\partial^2 a_\theta(x_k + j)}{\partial \theta \partial \theta'} + \Theta_j(x_k) - \Omega_j(x_k) \right] \frac{g[a_\theta^2(x_k + j)]}{p(x_k|\theta)} \]

\[ - (Z_g)^{-1} \sum_{k=1}^{\infty} q_k \left[ \frac{\partial^2 a_\theta(x_k)}{\partial \theta \partial \theta'} + \Theta_0(x_k) + \hat{\Omega}(x_k) \right] \frac{g[a_\theta^2(x_k)]}{p(x_k|\theta) + R(x_k|\theta)}, \quad \theta, \theta' \in \{\lambda, \phi\}, \]

where

\[ \hat{\Omega}(x_k) = (Z_g)^{-1} \frac{g[a_\theta^2(x_k)]}{p(x_k|\theta) + R(x_k|\theta)} \frac{\partial a_\theta(x_k)}{\partial \theta} \frac{\partial a_\theta(x_k)}{\partial \theta'}, \]

and \( \Theta_j(\cdot), \Omega_j(\cdot) \) are as in (12) and (13), respectively. The first and second derivatives of function \( a_\theta(\cdot) \) are given in (11) and (14), respectively.
5 Monte Carlo Simulation Study

A Monte Carlo simulation study was carried out to evaluate the performance of the maximum likelihood estimators for the LS\(_d\) models, particularly the log-normal, log-Student-\(t\), log-contaminated-normal, log-power-exponential, extended Birnbaum-Saunders and extended Birnbaum-Saunders-\(t\) cases. Note that when \(\phi = 4\) (fixed) the Birnbaum-Saunders and Birnbaum-Saunders-\(t\) are obtained. All numerical evaluations were done in the \(\mathbb{R}\) software; see [8].

The simulation scenario considers: sample size \(n \in \{40, 120, 400\}\), values of true parameters \(\phi \in \{1, 4, 8\}\), \(\lambda \in \{2.00\}\), censoring proportions \(\{0\%, 10\%, 30\%\}\), and 1, 000 Monte Carlo replications for each sample size. The values of the true extra parameters are presented in the caption of each table.

The maximum likelihood estimation results are presented in Tables 1, 2, 3, 4, 5 and 6. The following sample statistics for the maximum likelihood estimates are reported: empirical mean, bias, and mean squared error (MSE). A look at the results in Tables 1, 2, 3, 4, 5 and 6 allows us to conclude that, as the sample size increases, the bias and MSE of all the estimators decrease, indicating that they are asymptotically unbiased, as expected. Moreover, as the censoring proportion increases, the performances of the estimators of \(\phi\) and \(\lambda\), deteriorate. Generally, all of these results show the good performance of the proposed model.

6 Illustrative Examples

The LS\(_d\) models are now used to analyze two real-world data sets. It is considered the following discrete LS\(_d\) models: log-normal (LN), log-Student-\(t\) (L-\(t\)), log-contaminated-normal (LCN), log-power-exponential (LPE), Birnbaum-Saunders (BS), extended Birnbaum-Saunders (EBS), Birnbaum-Saunders-\(t\) (BS-\(t\)), and extended Birnbaum-Saunders-\(t\) (EBS-\(t\)).

Example 6.1 The first data set corresponds to the number of times that a DEC-20 computer broke down in each of 128 consecutive weeks of operation. This computer has operated at the Open University during the 1980s; see Table 7 and [12]. Descriptive statistics for the computer breaks data set are the following: 128 (sample size), 0 (minimum), 22 (maximum), 3 (median), 4.016 (mean), 3.808 (standard deviation), 94.839 (coefficient of variation), 1.732 (coefficient of skewness) and 3.995 (coefficient of kurtosis). From these results, we observe the positive skewness and a high degree of kurtosis. Figure 1(left) shows the histogram for the computer breaks data, from where it is confirmed the positive skewness. Moreover, Fig. 1(right) shows the usual and adjusted boxplots, and we note that some potential outliers are not in fact outliers when the adjusted boxplot is observed.

Table 8 presents the maximum likelihood estimates, computed by the BFGS method, and standard errors (SEs) for the LS\(_d\) models parameters. Moreover, the \(p\)
values of the $\chi^2$ and Cramer-Von Mises (CVM) statistics, and the the Akaike (AIC) and Bayesian information (BIC) criteria, are also reported. For comparison, the results of the discrete Burr, Poisson and binomial distributions, are given as well; see [7]. The results of Table 8 reveal that the discrete BS model provides the best adjustment compared to other models based on the values of AIC and BIC.

Example 6.2 The second data set refers to the number of physiotherapy sessions until a patient’s chronic back pain is reduced or alleviated; see Table 9. The patients were submitted to electric currents and the study was developed by the School of Physiotherapy Clinics of City University of Sao Paulo (UNICID), Sao Paulo, Brazil; see [11]. Observations were considered censored to the right when patients did not report pain reduction or relief after 12 treatment sessions, or if they had been lost to follow-up. Such as in [16], the variable of interest is defined as $T = X - 1$, $t = 0, 1, 2, 3, \ldots$, where $t = 0$ denotes a patient who presented pain relief in the first session performed. Descriptive statistics for the pain relief data are the following: 100(sample size), 0(minimum), 11(maximum), 0(median), 0.98(mean), 1.933(standard deviation), 197.258(coefficient of variation), 2.802(coefficient of skewness) and 9.015(coefficient of kurtosis). These statistics values indicate the positive skewness

| \( n \) | Cen. | \( \phi = 1 \) | \( \phi = 4 \) | \( \phi = 8 \) |
|--------|------|----------------|----------------|----------------|
|        |      | Mean | Bias | MSE  | Mean | Bias | MSE  | Mean | Bias | MSE  |
| 40     | 0%   | 1.0119 | 0.0119 | 0.0851 | 4.0752 | 0.0752 | 1.7247 | 8.1388 | 0.1388 | 7.5728 |
|        | \( \hat{\phi} \) | 2.0262 | 0.0262 | 0.1166 | 2.1126 | 0.1126 | 0.5629 | 2.2586 | 0.2586 | 1.4043 |
| 120    | \( \hat{\phi} \) | 1.0059 | 0.0059 | 0.0260 | 4.0413 | 0.0413 | 0.5237 | 8.0768 | 0.0768 | 2.1683 |
|        | \( \hat{\lambda} \) | 2.0086 | 0.0086 | 0.0394 | 2.0317 | 0.0317 | 0.1759 | 2.0750 | 0.0750 | 0.3897 |
| 400    | \( \hat{\phi} \) | 1.0041 | 0.0041 | 0.0077 | 4.0184 | 0.0184 | 0.1522 | 8.0404 | 0.0404 | 0.6641 |
|        | \( \hat{\lambda} \) | 2.0021 | 0.0021 | 0.0110 | 2.0086 | 0.0086 | 0.0464 | 2.0198 | 0.0198 | 0.1020 |
| 40     | 10%  | 1.0189 | 0.0189 | 0.0998 | 4.1507 | 0.1507 | 2.2755 | 8.3002 | 0.3002 | 10.0324 |
|        | \( \hat{\phi} \) | 2.0301 | 0.0301 | 0.1198 | 2.1154 | 0.1154 | 0.5850 | 2.2528 | 0.2528 | 1.4323 |
| 120    | \( \hat{\phi} \) | 1.0114 | 0.0114 | 0.0290 | 4.0519 | 0.0519 | 0.5978 | 8.1162 | 0.1162 | 2.5822 |
|        | \( \hat{\lambda} \) | 2.0099 | 0.0099 | 0.0396 | 2.0416 | 0.0416 | 0.1795 | 2.0842 | 0.0842 | 0.4037 |
| 400    | \( \hat{\phi} \) | 1.0015 | 0.0015 | 0.0090 | 4.0262 | 0.0262 | 0.1817 | 8.0675 | 0.0675 | 0.8152 |
|        | \( \hat{\lambda} \) | 2.0012 | 0.0012 | 0.0109 | 2.0014 | 0.0014 | 0.0466 | 2.0073 | 0.0073 | 0.0984 |
| 40     | 30%  | 1.0562 | 0.0562 | 0.1839 | 4.4113 | 0.4113 | 5.7506 | 8.9965 | 0.9965 | 28.9825 |
|        | \( \hat{\phi} \) | 2.0435 | 0.0435 | 0.1407 | 2.1396 | 0.1396 | 0.6858 | 2.2925 | 0.2925 | 1.7228 |
| 120    | \( \hat{\phi} \) | 1.0218 | 0.0218 | 0.0458 | 4.1116 | 0.1116 | 1.1065 | 8.2539 | 0.2539 | 5.0762 |
|        | \( \hat{\lambda} \) | 2.0124 | 0.0124 | 0.0438 | 2.0450 | 0.0450 | 0.1915 | 2.0868 | 0.0868 | 0.4270 |
| 400    | \( \hat{\phi} \) | 1.0093 | 0.0093 | 0.0133 | 4.0840 | 0.0840 | 0.3158 | 8.1961 | 0.1961 | 1.5102 |
|        | \( \hat{\lambda} \) | 2.0046 | 0.0046 | 0.0119 | 2.0076 | 0.0076 | 0.0501 | 2.0131 | 0.0131 | 0.1043 |

Table 1 Empirical values of mean, bias and MSE from simulated discrete log-normal data for the indicated maximum likelihood estimators

\( n \) = sample size, \( \text{Cen.} \) = censoring point, \( \phi \) = discrete log-normal parameter, \( \hat{\phi} \) and \( \hat{\lambda} \) = maximum likelihood estimators, \( \text{Mean} \) = mean, \( \text{Bias} \) = bias, \( \text{MSE} \) = mean squared error.
and a high degree of kurtosis. Figure 2 shows the histogram and the fitted survival function by the Kaplan–Meier (KM) method.

The maximum likelihood estimates of the discrete log-symmetric distribution parameters, along with AIC and BIC criteria are reported in Table 10. We have also provided the results of the following discrete distributions: Weibull, gamma and Lindley; see [7]. We note that the log-Student-\(t\) model provides better adjustment compared to the other models based on the values of AIC and BIC. Table 11 presents the fitted survival functions obtained by the KM and the discrete log-symmetric models. These results suggest that (extended) Birnbaum-Saunders and log-normal models yield the best fits to the pain relief data (Fig. 3).
7 Concluding Remarks

We have proposed a new class of distributions to deal with cases where the data are discrete, asymmetric and nonnegative. The proposed approach is a discrete version of the family of continuous log-symmetric distributions. We have considered estimation about the model parameters based on the maximum likelihood method with censored and uncensored data. A Monte Carlo simulation study was carried out to evaluate the behavior of the maximum likelihood estimators. We have applied the proposed models to two real-world data sets. In general, the results have shown that

---

Table 3  Empirical values of mean, bias and MSE from simulated discrete log-contaminated-normal data for the indicated maximum likelihood estimators with $\xi = (0.5, 0.5)^T$

| n   | Cen. | $\phi = 1$       | $\phi = 4$       | $\phi = 8$       |
|-----|------|------------------|------------------|------------------|
|     |      | Mean  | Bias  | MSE   | Mean  | Bias  | MSE   | Mean  | Bias  | MSE   |
| 40  | 0%   | $\hat{\phi}$ | 1.0057 | 0.0057 | 0.0947 | 4.0459 | 0.0459 | 1.8954 | 8.0944 | 0.0944 | 7.9832 |
|     |      | $\hat{\lambda}$ | 2.0437 | 0.0437 | 0.1706 | 2.1750 | 0.1750 | 0.8474 | 2.3687 | 0.3687 | 2.1616 |
| 120 | 0%   | $\hat{\phi}$ | 1.0024 | 0.0024 | 0.0300 | 4.0207 | 0.0207 | 0.6042 | 8.0165 | 0.0165 | 2.5295 |
|     |      | $\hat{\lambda}$ | 2.0161 | 0.0161 | 0.0545 | 2.0589 | 0.0589 | 0.2503 | 2.1297 | 0.1297 | 0.5661 |
| 400 | 0%   | $\hat{\phi}$ | 1.0020 | 0.0020 | 0.0085 | 4.0061 | 0.0061 | 0.1779 | 7.9912 | -0.0088 | 0.8225 |
|     |      | $\hat{\lambda}$ | 2.0024 | 0.0024 | 0.0165 | 2.0161 | 0.0161 | 0.0745 | 2.0410 | 0.0410 | 0.1751 |
| 40  | 10%  | $\hat{\phi}$ | 1.0336 | 0.0336 | 0.1484 | 4.1971 | 0.1971 | 3.1525 | 8.4153 | 0.4153 | 12.8221 |
|     |      | $\hat{\lambda}$ | 2.0562 | 0.0562 | 0.1819 | 2.1952 | 0.1952 | 0.9257 | 2.3944 | 0.3944 | 2.3858 |
| 120 | 10%  | $\hat{\phi}$ | 1.0154 | 0.0154 | 0.0396 | 4.0676 | 0.0676 | 0.7806 | 8.1552 | 0.1552 | 3.3793 |
|     |      | $\hat{\lambda}$ | 2.0182 | 0.0182 | 0.0588 | 2.0688 | 0.0688 | 0.2649 | 2.1338 | 0.1338 | 0.5975 |
| 400 | 10%  | $\hat{\phi}$ | 1.0067 | 0.0067 | 0.0103 | 4.0278 | 0.0278 | 0.2113 | 8.0581 | 0.0581 | 0.9135 |
|     |      | $\hat{\lambda}$ | 1.9988 | -0.0012 | 0.0163 | 2.0077 | 0.0077 | 0.0735 | 2.0222 | 0.0222 | 0.1586 |
| 40  | 30%  | $\hat{\phi}$ | 1.0644 | 0.0644 | 0.2448 | 4.4774 | 0.4774 | 7.4371 | 9.1280 | 1.1280 | 33.9336 |
|     |      | $\hat{\lambda}$ | 2.0635 | 0.0635 | 0.2088 | 2.2106 | 0.2106 | 1.0845 | 2.4422 | 0.4422 | 3.1094 |
| 120 | 30%  | $\hat{\phi}$ | 1.0310 | 0.0310 | 0.0591 | 4.1681 | 0.1681 | 1.3522 | 8.4147 | 0.4147 | 6.3718 |
|     |      | $\hat{\lambda}$ | 2.0237 | 0.0237 | 0.0635 | 2.0767 | 0.0767 | 0.2895 | 2.1436 | 0.1436 | 0.6536 |
| 400 | 30%  | $\hat{\phi}$ | 1.0126 | 0.0126 | 0.0161 | 4.0609 | 0.0609 | 0.3648 | 8.1301 | 0.1301 | 1.6252 |
|     |      | $\hat{\lambda}$ | 2.0005 | 0.0005 | 0.0176 | 2.0102 | 0.0102 | 0.0771 | 2.0237 | 0.0237 | 0.1628 |
the proposed discrete family proved to be an useful model for discrete data. As part of future research, it is of interest to discuss regression models as well as multivariate extensions. Moreover, time series models based on the proposed class may be of interest. Work on these issues is currently in progress and we hope to report some findings in future papers.

Table 4 Empirical values of mean, bias and MSE from simulated discrete log-power-exponential data for the indicated maximum likelihood estimators with $\xi = -0.5$

| n    | Cen. | $\phi = 1$ |       | $\phi = 4$ |       | $\phi = 8$ |       |
|------|------|------------|-------|------------|-------|------------|-------|
|      |      | Mean | Bias | MSE | Mean | Bias | MSE | Mean | Bias | MSE |
| 40   | 0%   | 0.9783 | -0.0217 | 0.0470 | 3.9497 | -0.0503 | 1.1520 | 7.9186 | -0.0814 | 5.6648 |
|      |      | $\hat{\lambda}$ | 2.0072 | 0.0072 | 0.0481 | 2.0493 | 0.0493 | 0.2856 | 2.1332 | 0.1332 | 0.6806 |
| 120  |      | 0.9970 | -0.0030 | 0.0134 | 4.0014 | 0.0014 | 0.3399 | 7.9966 | -0.0034 | 1.5695 |
|      |      | $\hat{\lambda}$ | 2.0032 | 0.0032 | 0.0153 | 2.0185 | 0.0185 | 0.0863 | 2.0482 | 0.0482 | 0.1996 |
| 400  |      | 0.9969 | -0.0031 | 0.0042 | 3.9906 | -0.0094 | 0.1019 | 7.9864 | -0.0136 | 0.4742 |
|      |      | $\hat{\lambda}$ | 1.9998 | -0.0002 | 0.0047 | 2.0036 | 0.0036 | 0.0282 | 2.0098 | 0.0098 | 0.0640 |
| 40   | 10%  | 0.9941 | -0.0059 | 0.0612 | 4.0123 | 0.0123 | 1.5967 | 8.0169 | 0.0169 | 8.0172 |
|      |      | $\hat{\lambda}$ | 2.0046 | 0.0046 | 0.0510 | 2.0472 | 0.0472 | 0.2900 | 2.1398 | 0.1398 | 0.6989 |
| 120  |      | 0.9999 | -0.0001 | 0.0167 | 4.0151 | 0.0151 | 0.4585 | 8.0260 | 0.0260 | 2.0633 |
|      |      | $\hat{\lambda}$ | 2.0010 | 0.0010 | 0.0165 | 2.0099 | 0.0099 | 0.0866 | 2.0349 | 0.0349 | 0.1968 |
| 400  |      | 1.0020 | 0.0020 | 0.0053 | 4.0178 | 0.0178 | 0.1283 | 8.0441 | 0.0441 | 0.6083 |
|      |      | $\hat{\lambda}$ | 2.0010 | 0.0010 | 0.0050 | 2.0044 | 0.0044 | 0.0298 | 2.0084 | 0.0084 | 0.0661 |
| 40   | 30%  | 1.0093 | 0.0093 | 0.0890 | 4.1602 | 0.1602 | 3.3402 | 8.2982 | 0.2982 | 15.8037 |
|      |      | $\hat{\lambda}$ | 2.0111 | 0.0111 | 0.0646 | 2.0512 | 0.0512 | 0.3185 | 2.1408 | 0.1408 | 0.7509 |
| 120  |      | 1.0040 | 0.0040 | 0.0250 | 4.0485 | 0.0485 | 0.7736 | 8.1109 | 0.1109 | 3.7337 |
|      |      | $\hat{\lambda}$ | 2.0019 | 0.0019 | 0.0191 | 2.0104 | 0.0104 | 0.0927 | 2.0350 | 0.0350 | 0.2090 |
| 400  |      | 1.0026 | 0.0026 | 0.0079 | 4.0215 | 0.0215 | 0.1934 | 8.0665 | 0.0665 | 0.9711 |
|      |      | $\hat{\lambda}$ | 2.0010 | 0.0010 | 0.0056 | 2.0045 | 0.0045 | 0.0318 | 2.0093 | 0.0093 | 0.0694 |
Table 5 Empirical values of mean, bias and MSE from simulated discrete extended Birnbaum-Saunders data for the indicated maximum likelihood estimators with $\xi = 0.5$

| $n$  | Cen. | $\phi = 1$ | | $\phi = 4$ | | $\phi = 8$ | |
|------|------|------------|---|------------|---|------------|
|      |      | Mean | Bias | MSE | Mean | Bias | MSE | Mean | Bias | MSE |
| 40   | 0%   | 0.8741 | -0.1259 | 0.1721 | 3.9719 | -0.0281 | 0.9531 | 7.9871 | -0.0129 | 4.3427 |
|      |      | $\hat{\phi}$ | 2.0126 | 0.0126 | 0.0075 | 2.0117 | 0.0117 | 0.0275 | 2.0152 | 0.0152 | 0.0542 |
|      |      | $\hat{\lambda}$ | 2.0049 | 0.0049 | 0.0029 | 2.0032 | 0.0032 | 0.0090 | 2.0037 | 0.0037 | 0.0185 |
| 120  |      | 0.9747 | -0.0253 | 0.0353 | 4.0090 | 0.0090 | 0.3178 | 8.0289 | 0.0289 | 1.3258 |
|      |      | $\hat{\phi}$ | 0.9913 | -0.0087 | 0.0117 | 4.0042 | 0.0042 | 0.0978 | 8.0266 | 0.0266 | 0.4048 |
|      |      | $\hat{\lambda}$ | 2.0020 | 0.0020 | 0.0009 | 2.0011 | 0.0011 | 0.0026 | 2.0004 | 0.0004 | 0.0051 |
| 400  |      | 0.8825 | -0.1175 | 0.1945 | 4.0107 | 0.0107 | 1.2217 | 8.1117 | 0.1117 | 5.3717 |
|      |      | $\hat{\phi}$ | 2.0128 | 0.0128 | 0.0076 | 2.0120 | 0.0120 | 0.0289 | 2.0172 | 0.0172 | 0.0565 |
|      |      | $\hat{\lambda}$ | 2.0047 | 0.0047 | 0.0028 | 2.0043 | 0.0043 | 0.0096 | 2.0031 | 0.0031 | 0.0187 |
| 40   | 10%  | 0.9699 | -0.0301 | 0.0460 | 4.0048 | 0.0048 | 0.3449 | 8.0872 | 0.0872 | 1.6128 |
|      |      | $\hat{\phi}$ | 0.9907 | -0.0093 | 0.0134 | 4.0000 | 0.0000 | 0.1167 | 8.0092 | 0.0092 | 0.4932 |
|      |      | $\hat{\lambda}$ | 2.0005 | 0.0005 | 0.0009 | 1.9997 | -0.0003 | 0.0026 | 2.0005 | 0.0005 | 0.0050 |
| 40   | 30%  | 0.7674 | -0.2326 | 0.3577 | 4.0841 | 0.0841 | 2.0748 | 8.3517 | 0.3517 | 8.9625 |
|      |      | $\hat{\phi}$ | 2.0152 | 0.0152 | 0.0067 | 2.0178 | 0.0178 | 0.0343 | 2.0280 | 0.0280 | 0.0692 |
|      |      | $\hat{\lambda}$ | 2.0052 | 0.0052 | 0.0028 | 2.0057 | 0.0057 | 0.0106 | 2.0055 | 0.0055 | 0.0210 |
| 120  |      | 0.9263 | -0.0737 | 0.0967 | 4.0323 | 0.0323 | 0.5129 | 8.1723 | 0.1723 | 2.5286 |
|      |      | $\hat{\phi}$ | 0.9852 | -0.0148 | 0.0212 | 4.0188 | 0.0188 | 0.1660 | 8.0638 | 0.0638 | 0.7290 |
|      |      | $\hat{\lambda}$ | 2.0009 | 0.0009 | 0.0009 | 2.0011 | 0.0011 | 0.0028 | 2.0028 | 0.0028 | 0.0056 |
Table 6  Empirical values of mean, bias and MSE from simulated discrete extended Birnbaum-Saunders-$t$ data for the indicated maximum likelihood estimators with $\xi = (0.5, 4)^{1}$

| $n$ | Cen. | $\phi = 1$ | $\phi = 4$ | $\phi = 8$ |
|-----|------|------------|------------|------------|
|     |      | Mean | Bias  | MSE  | Mean | Bias  | MSE  | Mean | Bias  | MSE  |
| 40  | 0%   | 0.9871 | -0.0129 | 0.1511 | 3.9863 | -0.0137 | 1.4404 | 8.0258 | 0.0258 | 6.0994 |
|     |      | 2.0028 | 0.0028 | 0.0114 | 2.0100 | 0.0100 | 0.0387 | 2.0167 | 0.0167 | 0.0756 |
| 120 | 0.5% | 0.9998 | -0.0002 | 0.0457 | 4.0319 | 0.0319 | 0.4972 | 8.0689 | 0.0689 | 2.1333 |
|     |      | 2.0026 | 0.0026 | 0.0037 | 2.0061 | 0.0061 | 0.0128 | 2.0086 | 0.0086 | 0.0248 |
| 400 | 0.5% | 0.9984 | -0.0016 | 0.0141 | 3.9996 | -0.0004 | 0.1569 | 8.0282 | 0.0282 | 0.6348 |
|     |      | 2.0018 | 0.0018 | 0.0010 | 2.0031 | 0.0031 | 0.0038 | 2.0035 | 0.0035 | 0.0074 |
| 40  | 10%  | 1.0002 | 0.0002 | 0.1663 | 4.1093 | 0.1093 | 1.8263 | 8.2395 | 0.2395 | 7.9333 |
|     |      | 2.0009 | 0.0009 | 0.0109 | 2.0029 | 0.0029 | 0.0387 | 2.0095 | 0.0095 | 0.0760 |
| 120 | 10%  | 1.0055 | 0.0055 | 0.0512 | 4.0492 | 0.0492 | 0.5522 | 8.1287 | 0.1287 | 2.3725 |
|     |      | 2.0027 | 0.0027 | 0.0037 | 2.0044 | 0.0044 | 0.0129 | 2.0080 | 0.0080 | 0.0245 |
| 400 | 10%  | 1.0003 | 0.0003 | 0.0157 | 3.9944 | -0.0056 | 0.1649 | 8.0085 | 0.0085 | 0.6848 |
|     |      | 2.0011 | 0.0011 | 0.0012 | 2.0020 | 0.0020 | 0.0042 | 2.0028 | 0.0028 | 0.0078 |
| 40  | 30%  | 0.9829 | -0.0171 | 0.2728 | 4.1695 | 0.1695 | 2.7474 | 8.4158 | 0.4158 | 12.0941 |
|     |      | 2.0027 | 0.0027 | 0.0114 | 2.0055 | 0.0055 | 0.0438 | 2.0151 | 0.0151 | 0.0882 |
| 120 | 30%  | 1.0115 | 0.0115 | 0.0708 | 4.1024 | 0.1024 | 0.7985 | 8.2566 | 0.2566 | 3.5320 |
|     |      | 2.0029 | 0.0029 | 0.0038 | 2.0065 | 0.0065 | 0.0140 | 2.0119 | 0.0119 | 0.0273 |
| 400 | 30%  | 1.0009 | 0.0009 | 0.0238 | 4.0094 | 0.0094 | 0.2337 | 8.0494 | 0.0494 | 0.9990 |
|     |      | 2.0014 | 0.0014 | 0.0012 | 2.0037 | 0.0037 | 0.0047 | 2.0049 | 0.0049 | 0.0087 |

Table 7  Computer breaks data

| x   | 0  | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 16 | 17 | 22 |
|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
|     | 15 | 19 | 23 | 14 | 15 | 10 | 8  | 4  | 6  | 2  | 3  | 3  | 2  | 1  | 1  | 1  | 1  |
Fig. 1 Histogram (left) and boxplots (right) for the computer breaks data
| Discrete distribution          | Estimates   | SE   | \( p \) value     | AIC     | BIC     |
|--------------------------------|-------------|------|-------------------|---------|---------|
| Log-normal                    | \( \lambda = 3.2280 \) | 0.2526 | 0.7841 | 0.6959 | 643.5141 | 652.0702 |
|                               | \( \phi = 0.7541 \) | 0.1048 | \( \chi^2 \)    | CVM     |         |
| Log-Student-\( t \)          | \( \lambda = 3.2653 \) | 0.2574 | 0.8306 | 0.7762 | 644.2248 | 652.7809 |
|                               | \( \phi = 0.7065 \) | 0.1026 | \( \chi^2 \)    | CVM     |         |
| Log-Contaminated-Normal       | \( \lambda = 3.2283 \) | 0.2526 | 0.7906 | 0.6967 | 645.5205 | 656.9287 |
|                               | \( \phi = 0.6858 \) | 0.0953 | \( \chi^2 \)    | CVM     |         |
|                               | \( \xi = 0.9 \) |         |         |         |         |         |
| Log-Power-Exponential         | \( \lambda = 3.1624 \) | 0.2770 | 0.7096 | 0.5151 | 642.8785 | 651.4346 |
|                               | \( \phi = 1.0176 \) | 0.0555 | \( \chi^2 \)    | CVM     |         |
|                               | \( \xi = -0.2 \) |         |         |         |         |         |
| Birnbaum-Saunders             | \( \lambda = 3.1704 \) | 0.0589 | 0.6007 | 0.5283 | 640.6061 | 646.3102 |
|                               | \( \xi = 0.9 \) |         |         |         |         |         |
| Extended Birnbaum-Saunders    | \( \lambda = 3.1436 \) | 0.2090 | 0.6517 | 0.4700 | 642.4045 | 650.9605 |
|                               | \( \phi = 2.9392 \) | 0.0438 | \( \chi^2 \)    | CVM     |         |
|                               | \( \xi = 1.1 \) |         |         |         |         |         |
| Birnbaum-Saunders-\( t \)    | \( \lambda = 3.1803 \) | 0.3008 | 0.7966 | 0.5799 | 642.9026 | 651.4587 |
|                               | \( \xi = 0.9 \) |         |         |         |         |         |
|                               | \( \xi = 20.0 \) |         |         |         |         |         |
| Extended Birnbaum-Saunders-\( t \) | \( \lambda = 3.1406 \) | 0.2193 | 0.6492 | 0.4364 | 644.5534 | 655.9615 |
|                               | \( \phi = 2.0820 \) | 0.0818 | \( \chi^2 \)    | CVM     |         |
|                               | \( \xi = 1.3 \) |         |         |         |         |         |
|                               | \( \xi = 20.0 \) |         |         |         |         |         |
| Burr                          | \( \alpha = 3.3087 \) | 0.6807 | 0.0000 | 0.0058 | 689.1619 | 694.8666 |
|                               | \( \beta = 0.2436 \) | 0.0526 | \( \chi^2 \)    | CVM     |         |
| Poisson                       | \( \lambda = 4.0156 \) | 0.1771 | 0.0000 | 0.0004 | 771.9487 | 774.8007 |
| Binomial                      | \( \theta = 0.0314 \) | 0.0014 | 0.0000 | 0.0003 | 782.9518 | 785.8038 |
Table 9  Number of sessions until a patient’s chronic back pain is reduced or alleviated

| Sessions | T | No at risk | No of events | Censoring indicator |
|----------|---|------------|--------------|---------------------|
| 1        | 0 | 100        | 64           | 0                   |
| 2        | 1 | 36         | 16           | 0                   |
| 3        | 2 | 20         | 5            | 1                   |
| 4        | 3 | 14         | 4            | 0                   |
| 5        | 4 | 10         | 4            | 0                   |
| 6        | 5 | 6          | 3            | 0                   |
| 7        | 6 | 3          | 0            | 0                   |
| 8        | 7 | 3          | 1            | 0                   |
| 9        | 8 | 2          | 0            | 0                   |
| 10       | 9 | 2          | 1            | 0                   |
| 11       | 10| 1          | 0            | 0                   |
| 12       | 11| 1          | 0            | 1                   |

Fig. 2  Histogram (left) and KM (right) for the pain relief data
Fig. 3: Estimation of the survival function using the KM (solid), discrete log-symmetric distributions, discrete distributions Weibull, Gamma and Lindley (dashed) with the pain relief data.
Table 10 Maximum likelihood estimates and model selection measures for fit to the pain relief data

| Discrete distribution | Estimates       | SE   | AIC     | BIC     |
|-----------------------|-----------------|------|---------|---------|
| Log-normal            | $\hat{\lambda} = 2.3229$ | 0.1329 | 540.0872 | 549.1191 |
|                       | $\hat{\phi} = 0.462$   | 0.0571 |          |         |
| Log-Student-t         | $\hat{\lambda} = 1.8745$ | 0.1250 | 513.1153 | 522.1472 |
|                       | $\hat{\phi} = 0.122$   | 0.0538 |          |         |
|                       | $\xi = 2$          |      |          |         |
| Log-Power-Exponential | $\hat{\lambda} = 2.0046$ | 0.0999 | 528.1035 | 537.1354 |
|                       | $\hat{\phi} = 0.1713$ | 0.0259 |          |         |
|                       | $\xi = 0.5$        |      |          |         |
| Log-Contaminated-Normal | $\hat{\lambda} = 1.8654$ | 0.1329 | 513.3146 | 525.3571 |
|                       | $\hat{\phi} = 0.1018$ | 0.0519 |          |         |
|                       | $\xi = 0.37:0.10$   |      |          |         |
| Birnbaum-Saunders     | $\hat{\lambda} = 2.4767$ | 0.8043 | 543.6507 | 549.6719 |
|                       | $\xi = 0.7$        |      |          |         |
| Extended Birnbaum-Saunders | $\hat{\lambda} = 2.3263$ | 0.1560 | 540.2164 | 549.2483 |
|                       | $\hat{\phi} = 184.8547$ | 0.0701 |          |         |
|                       | $\xi = 0.1$        |      |          |         |
| Birnbaum-Saunders-t   | $\hat{\lambda} = 1.8966$ | 0.2047 | 515.0695 | 524.1014 |
|                       | $\xi = 0.4:2.0$    |      |          |         |
| Extended Birnbaum-Saunders-t | $\hat{\lambda} = 1.8751$ | 0.1477 | 515.1634 | 527.2060 |
|                       | $\hat{\phi} = 49.0515$ | 0.1995 |          |         |
|                       | $\xi = 0.1:2.0$    |      |          |         |
| Weibull               | $\hat{\alpha} = 3.3464$ | 0.2387 | 605.6825 | 611.7038 |
|                       | $\hat{\beta} = 1.2383$ | 0.0739 |          |         |
| Gamma                 | $\hat{\alpha} = 1.9100$ | 0.2150 | 588.5916 | 594.6129 |
|                       | $\hat{\beta} = 1.6114$ | 0.2081 |          |         |
| Lindley               | $\hat{\theta} = 0.5421$ | 0.0331 | 599.2233 | 602.2339 |
Table 11  Estimates of the survival function via KM, discrete log-symmetric distributions and discrete distributions Weibull, Gamma and Lindley

| x  | KM   | LN  | L-t | LPE | LCN  | BS   | EBS  | BS-t | EBS-t | Weibull | Gamma | Lindley |
|----|------|-----|-----|-----|------|------|------|------|-------|---------|-------|---------|
| 0  | 1.00 | 0.89 | 0.89 | 0.87 | 0.88 | 0.91 | 0.89 | 0.87 | 0.89 | 0.87 | 0.87 | 0.87 |
| 1  | 0.43 | 0.38 | 0.43 | 0.43 | 0.43 | 0.43 | 0.43 | 0.43 | 0.43 | 0.43 | 0.43 | 0.43 |
| 2  | 0.24 | 0.16 | 0.24 | 0.24 | 0.24 | 0.24 | 0.24 | 0.24 | 0.24 | 0.24 | 0.24 | 0.24 |
| 3  | 0.19 | 0.12 | 0.19 | 0.19 | 0.19 | 0.19 | 0.19 | 0.19 | 0.19 | 0.19 | 0.19 | 0.19 |
| 4  | 0.16 | 0.13 | 0.16 | 0.16 | 0.16 | 0.16 | 0.16 | 0.16 | 0.16 | 0.16 | 0.16 | 0.16 |
| 5  | 0.13 | 0.11 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 |
| 6  | 0.09 | 0.07 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 |
| 7  | 0.07 | 0.06 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 |
| 8  | 0.06 | 0.05 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 |
| 9  | 0.05 | 0.04 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |
| 10 | 0.03 | 0.02 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| 12 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
Declarations

Conflict of Interest The authors declare no potential conflict of interests.

References

1. Balakrishnan N, Saulo H, Bourguignon M, Zhu X (2017) On moment-type estimators for a class of log-symmetric distributions. Comput Stat 32(4):1339–1355
2. Hinkley DV (1975) On power transformations to symmetry. Biometrika 62:101–111
3. Jones MC (2008) On reciprocal symmetry. J Stat Plann Infer 138:3039–3043
4. Medeiros FMC, Ferrari SLP (2017) Small-sample testing inference in symmetric and log-symmetric linear regression models. Statistica Neerlandica 71:200–224
5. Mittelhammer RC, Judge GG, Miller DJ (2000) Econometric foundations. Cambridge University Press, New York
6. Moors JJA (1988) A quantile alternative for kurtosis. Statistician 37:25–32
7. Nadarajah S (2015) A statistical analysis of Iraq body counts. Qual Quant 49:21–37
8. R Core Team (2016) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
9. Saulo H, Leão J (2017) On log-symmetric duration models applied to high frequency financial data. Econ Bull 37:1089–1097
10. Saulo H, Leão J, Nobre L, Balakrishnan N (2021) A class of asymmetric regression models for left-censored data. Braz J Prob Stat 35:62–84
11. Silva JF, Liebano RE, Corrêa JB, Matsushita RY, Nakano EY (2017) Analysis of the time to relieving pain in patients with chronic non-specific low back pain via Cox proportional hazard model. Ciência e Natura 39:233–243
12. Trenkler, D. (1995). A handbook of small data sets: Hand, d.j., daly, f., lunn, a.d., mcconway, k.j. & ostrowski, e. (1994): Chapman & hall, london. Comput Stat Data Anal 19(1):101–101
13. Vanegas LH, Paula GA (2016a) An extension of log-symmetric regression models: R codes and applications. J Stat Simul Comput 86:1709–1735
14. Vanegas LH, Paula GA (2016b) ssym: Fitting semi-parametric log-symmetric regression models. R package version 1(5):7
15. Ventura M, Saulo H, Leiva V, Monsueto S (2019) Log-symmetric regression models: information criteria and application to movie business and industry data with economic implications. Appl Stochast Models Bus Ind 35(4):963–977
16. Vila R, Nakano EY, Saulo H (2019) Theoretical results on the discrete Weibull distribution of Nakagawa and Osaki. Statistics 53(2):339–363
17. Zwilling D, Kokoska S (2000) Standard probability and statistical tables and formula. Chapman & Hall, Boca Raton

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.