ECONSTOR

Make Your Publications Visible.

A Service of



Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Kreer, Markus; Kızılersü, Ayşe; Guscott, Jake; Schmitz, Lukas Christopher; Thomas, Anthony W.

Article — Published Version Maximum likelihood estimation for left-truncated loglogistic distributions with a given truncation point

Statistical Papers

Provided in Cooperation with: Springer Nature

Suggested Citation: Kreer, Markus; Kızılersü, Ayşe; Guscott, Jake; Schmitz, Lukas Christopher; Thomas, Anthony W. (2024) : Maximum likelihood estimation for left-truncated log-logistic distributions with a given truncation point, Statistical Papers, ISSN 1613-9798, Springer, Berlin, Heidelberg, Vol. 65, Iss. 9, pp. 5409-5445, https://doi.org/10.1007/s00362-024-01603-8

This Version is available at: https://hdl.handle.net/10419/314987

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.



http://creativecommons.org/licenses/by/4.0/

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.





REGULAR ARTICLE



Maximum likelihood estimation for left-truncated log-logistic distributions with a given truncation point

Markus Kreer¹ · Ayşe Kızılersü² · Jake Guscott² · Lukas Christopher Schmitz³ · Anthony W. Thomas²

Received: 27 October 2022 / Revised: 29 July 2024 / Published online: 10 September 2024 © The Author(s) 2024

Abstract

For a sample X_1, X_2, \ldots, X_N of independent identically distributed copies of a loglogistically distributed random variable X the maximum likelihood estimation is analysed in detail if a left-truncation point $x_L > 0$ is introduced. Due to scaling properties it is sufficient to investigate the case $x_L = 1$. Here the corresponding maximum likelihood equations for a normalised sample (i.e. a sample divided by x_I) do not always possess a solution. A simple criterion guarantees the existence of a solution: Let $\mathbb{E}(\cdot)$ denote the expectation induced by the normalised sample and denote by $\beta_0 = \mathbb{E}(\ln X)^{-1}$, the inverse value of expectation of the logarithm of the sampled random variable X (which is greater than $x_L = 1$). If this value β_0 is bigger than a certain positive number β_C then a solution of the maximum likelihood equation exists. Here the number β_C is the unique solution of a moment equation, $\mathbb{E}(X^{-\beta_C}) = \frac{1}{2}$. In the case of existence a profile likelihood function can be constructed and the optimisation problem is reduced to one dimension leading to a robust numerical algorithm. When the maximum likelihood equations do not admit a solution for certain data samples, it is shown that the Pareto distribution is the L^1 -limit of the degenerated left-truncated log-logistic distribution, where $L^1(\mathbb{R}^+)$ is the usual Banach space of functions whose absolute value is Lebesgue-integrable. A large sample analysis showing consistency and asymptotic normality complements our analysis. Finally, two applications to real world data are presented.

Keywords Log-logistic distibution \cdot Truncated distributions \cdot Maximum likelihood estimation

Ayşe Kızılersü ayse.kizilersu@adelaide.edu.au

¹ Feldbergschule, Oberhöchstadter Str. 20, 61440 Oberursel, Taunus, Germany

² CSSM, Department of Physics, University of Adelaide, Adelaide 5005, Australia

³ Institut f
ür Mathematik, Johannes Gutenberg-Universit
ät Mainz, Staudingerweg 9, 55128 Mainz, Germany

1 Preliminaries

The log-logistic distribution, also known as the Fisk distribution, has been popular in the econometric community since the early 1960s because of its better description of income distributions Fisk (1961), compared with the Pareto distribution. On the other hand, hydrologists in the late 1980s suggested that log-logistic distributions were useful for modelling Canadian precipitation data Shoukri et al. (1988), or flood frequencies in Scotland for annual flood maxima Ahmad et al. (1988). The log-logistic distribution is related to the logistic distribution by a logarithmic transform. Practitioners make use of log-logistic distributions because of their easy calculability and closed form expressions for both cumulative probability distribution (cdf) and probability density function (pdf) Reath et al. (2018), He et al. (2020). For x > 0 the pdf and cdf for the log-logistic distributions are given by

$$f(x|\alpha,\beta) = \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1} \frac{1}{\left[1 + (x/\alpha)^{\beta}\right]^2},$$
$$F(x|\alpha,\beta) = \frac{1}{1 + (x/\alpha)^{-\beta}},$$

where $\alpha > 0$ is the scale parameter and $\beta > 0$ the shape parameter. We can recover the Pareto distribution for the tail by expanding the cdf for "large" arguments, $x \gg \alpha$, up to the first order in x/α , $F(x) \simeq 1 - (x/\alpha)^{-\beta}$.

In the analysis of annual flood maxima, the log-logistic distribution in Ahmad et al. (1988) was modified by the introduction of a threshold parameter for practical reasons: a flood maximum in a rainy country like Scotland should always be above a certain threshold level. In this paper we pursue a different approach: we keep the two-parameter distribution but introduce a fixed left-truncation point $x_L > 0$ instead. Thus, we have for $x \ge x_L$, the left-truncated log-logistic pdf and cdf respectively as [see also Kendall and Stuart (1979), Cohen (1991), Guscott (2018)]

$$f_{LT}(x|\alpha,\beta;x_L) = \left(1 + \left(\frac{x_L}{\alpha}\right)^{\beta}\right) \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1} \frac{1}{\left[1 + \left(\frac{x}{\alpha}\right)^{\beta}\right]^2},\tag{1}$$

$$F_{LT}(x|\alpha,\beta;x_L) = \frac{\left(\frac{x}{\alpha}\right)^{\beta} - \left(\frac{x_L}{\alpha}\right)^{\beta}}{1 + \left(\frac{x}{\alpha}\right)^{\beta}},\tag{2}$$

where the subscript "LT" stands for "left-truncated". If a random variable X is log-logistically distributed with positive parameters α , β and left-truncation x_L we can denote $X \sim LL(\alpha, \beta; x_L)$. Using Eq. (2) we generate a random variable $X \sim LL(\alpha, \beta; x_L)$ from a uniformly distributed random variable U in the interval (0, 1), namely

$$X = \left(\frac{\alpha^{\beta}U + x_{L}^{\beta}}{1 - U}\right)^{1/\beta} = \alpha \left(\frac{U + \eta}{1 - U}\right)^{1/\beta},\tag{3}$$

🖉 Springer

where $\eta = (x_L/\alpha)^{\beta}$.

When probability distributions are truncated, some interesting effects can happen. Castillo in del Castillo (1994) demonstrated that if a normal distribution is truncated, there exist finite random samples for which the regular maximum likelihood equations (MLE) do not possess a solution. Instead, a new maximum likelihood estimator as a limit case is obtained, leading to a degenerate one-parameter distribution, namely the exponential distribution"to fit" the sample data appropriately. Similar effects were observed in the analysis of left-truncated Weibull distributions Kreer et al. (2015) Kizilersu et al. (2016), where for some samples the MLE do not possess a solution. Indeed, in Guscott (2018) it is found numerically that for a certain percentage of random samples the MLE for the left-truncated log-logistic distribution do not admit a solution ((see also Table 1) and compare this to Table 1 in Kizilersu et al. (2016) for similar results concerning the truncated Weibull distribution).

These observations are the motivations for this paper, because they clearly need careful mathematical analysis of the left-truncated log-logistic distribution. To the best of our knowledge no study has proven the existence of the maximum likelihood estimator for a random sample drawn from the left-truncated log-logistic distribution. Therefore, any numerical studies without this proof, assuming the existence of a solution, will end up in the worst case either not converging at all, or converging to a degenerate solution in which the parameter estimates can take values like zero or infinity.

Our paper is structured as follows: Sect. 2 contains the main theorems for the existence of a non-trivial solution of the maximum likelihood equations for the left-truncated log-logistic distribution. It also examines the properties of a suitable profile likelihood function and its asymptotic properties such as "consistency" and "asymptotic normality". These results will be relevant for the efficient numerical implementation of an algorithm solving the maximum likelihood equations. As an illustration, in Sect. 3 we apply our technique to cancer data and German and Canadian precipitation data. All proofs are given in Sect. 4.

2 Mathematical results

2.1 Scaling property

We first provide a lemma dealing with a scaling property.

Lemma 1 (Scaling Property) Let $X \sim LL(\alpha, \beta; x_L)$ be a left-truncated log-logistic random variable. Then for any k > 0 we have $kX \sim LL(k\alpha, \beta; kx_L)$.

Proof By assumption $X > x_L$ with $X \sim LL(\alpha, \beta; x_L)$. Thus $kX > kx_L$ and

$$\operatorname{prob}(kX < x \mid kX > kx_L) = \operatorname{prob}(X < x/k|X > x_L)$$
$$= \frac{\left(\frac{x/k}{\alpha}\right)^{\beta} - \left(\frac{x_L}{\alpha}\right)^{\beta}}{1 + \left(\frac{x/k}{\alpha}\right)^{\beta}} = \frac{\left(\frac{x}{k\alpha}\right)^{\beta} - \left(\frac{kx_L}{k\alpha}\right)^{\beta}}{1 + \left(\frac{x}{k\alpha}\right)^{\beta}},$$

🖄 Springer

and the proof is finished.

As a consequence of this scaling property, for simplicity we will assume the lefttruncation point $x_L = 1$ later on. In other words, we rescale the independent identically distributed (i.i.d.) sample X_1, \ldots, X_N with $k = 1/x_L$, leading to $X_1/x_L, \ldots, X_N/x_L$ (which is truncated at 1). At any stage we can go back to the original sample and original parameters by using a rescaling factor $k = x_L$.

2.2 The first-order maximum likelihood equations

We introduce a new parameter $\lambda \equiv \alpha^{\beta}$, and rewrite Eq. (1) as

$$f_{LT}(x|\lambda,\beta) = \left(1 + \frac{x_L^{\beta}}{\lambda}\right) \frac{\beta}{\lambda} x^{\beta-1} \frac{1}{\left[1 + \frac{x^{\beta}}{\lambda}\right]^2}.$$
 (4)

Note that with this notation we have $\eta = x_L^{\beta}/\lambda$.

Using Eq. (4) and a left-truncated log-logistic sample, denoted by $X_1, X_2, ..., X_N$, where all observations are $X_1, X_2, ..., X_N > x_L > 0$, the log-likelihood function is given by

$$\ln L_{LT}(\{X_i\}|\lambda,\beta;x_L) = N \ln \left(1 + \frac{x_L^{\beta}}{\lambda}\right) + N \ln \beta - N \ln \lambda + (\beta - 1) \sum_{i=1}^{N} \ln X_i$$
$$-2 \sum_{i=1}^{N} \ln \left[1 + \frac{X_i^{\beta}}{\lambda}\right].$$
(5)

The maximum likelihood equations are obtained by differentiating Eq. (5) with respect to λ and β and setting the derivatives equal to zero. From $\frac{\partial}{\partial \lambda} \ln L_{LT} = 0$ we obtain for $(\lambda, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+$

$$0 = \frac{2}{N} \sum_{i=1}^{N} \frac{\frac{X_i^{\beta}}{\lambda}}{1 + \frac{X_i^{\beta}}{\lambda}} - 1 - \frac{\frac{x_L^{\beta}}{\lambda}}{1 + \frac{x_L^{\beta}}{\lambda}}$$
(6)

and from $\frac{\partial}{\partial\beta} \ln L_{LT} = 0$ we obtain for $(\lambda, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+$

$$0 = \frac{\frac{x_L^{\beta}}{\lambda} \ln x_L}{1 + \frac{x_L^{\beta}}{\lambda}} + \frac{1}{\beta} + \frac{1}{N} \sum_{i=1}^{N} \ln X_i - \frac{2}{N} \sum_{i=1}^{N} \frac{\frac{X_i^{\beta}}{\lambda}}{1 + \frac{X_i^{\beta}}{\lambda}} \ln X_i.$$
(7)

A solution of these MLE, Eqs. (6)–(7), (if it exists) will be denoted by $(\hat{\lambda}, \hat{\beta})$. For finite, untruncated samples, Gupta et al. (1999) (respectively Antle et al. (1970) for

the logistic distribution) has shown that the MLE have a unique solution for the loglogistic distribution (hence for the logistic distribution). These proofs fail when a left-truncation point $x_L > 0$ is introduced.

For simplicity we will take $x_L = 1$ from now on without loss of generality. For convenience we also define the quantity $S = \sum_{i=1}^{N} \ln X_i$ and introduce a new objective function as in Gupta et al. (1999)

$$\varphi(\lambda,\beta) = \ln L_{LT}(\{X_i\}|\lambda,\beta; x_L = 1) + S.$$
(8)

The extrema of both the function $\varphi(\cdot, \cdot)$ and the log-likelihood function are the same (and so are their MLE), because they only differ by a constant number, *S*.

2.3 Our key theorems

Our first theorem states the existence of a maximum for the objective function under certain conditions.

Theorem 1 (Existence) Consider the i.i.d. left-truncated sample $X_1, \ldots, X_N > 1$, for which at least two observations are different, and define the objective function $\varphi(\cdot, \cdot)$: $\mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ by

$$\varphi(\lambda,\beta) = N \ln\left(1+\frac{1}{\lambda}\right) + N \ln\beta - N \ln\lambda + \beta S - 2\sum_{i=1}^{N} \ln\left[1+\frac{X_i^{\beta}}{\lambda}\right].$$
 (9)

Define $\beta_C > 0$ as the unique solution of

$$\frac{1}{2} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{X_i^{\beta_C}}$$
(10)

and $\beta_0 > 0$ by

$$\frac{1}{\beta_0} = \frac{1}{N} \sum_{i=1}^N \ln X_i.$$
 (11)

Then the following holds true

- (1) For $\beta_0 > \beta_C$ the objective function $\varphi(\cdot, \cdot)$ possesses a global maximum $(\hat{\lambda}, \hat{\beta}) \in \mathbb{R}^+ \times \mathbb{R}^+$.
- (2) For $\beta_0 \leq \beta_C$ the objective function $\varphi(\cdot, \cdot)$ possesses a (local) maximum $(\hat{\lambda}, \hat{\beta}) = (0, \beta_0)$ on the boundary.

The proof is given in Sect. 4.

The next result is important for the numerical computation of the maxima, leads to a profile likelihood function, and provides a curve for the loci¹ of critical points.

¹ Locus: A set of points that satisfy or are determined by some specific condition.

Theorem 2 (Loci of critical points)

Consider the i.i.d. left-truncated sample $X_1, \ldots, X_N > 1$ for which at least two observations are different. Then the following holds true:

(1) For fixed $\beta > \beta_C$ the equation

$$\left. \frac{\partial}{\partial \lambda} \varphi(\lambda, \beta) \right|_{\lambda = \hat{\lambda}} = 0$$

has exactly one positive solution $\hat{\lambda}$ which depends on the fixed parameter β . Furthermore we have the following inequalities

$$\begin{split} &\frac{\partial}{\partial\lambda}\varphi(\lambda,\beta)>0 \quad for \quad 0<\lambda<\hat{\lambda}\\ &\frac{\partial}{\partial\lambda}\varphi(\lambda,\beta)<0 \quad for \quad \lambda>\hat{\lambda} \end{split}$$

and thus

$$\left. \frac{\partial^2}{\partial \lambda^2} \varphi(\lambda, \beta) \right|_{\lambda = \hat{\lambda}} < 0.$$
(12)

(2) For fixed $0 < \beta \leq \beta_C$ the equation

$$\left. \frac{\partial}{\partial \lambda} \varphi(\lambda, \beta) \right|_{\lambda = \hat{\lambda}} = 0,$$

has the only solution $\hat{\lambda} = 0$.

Hence the non-negative continous function $\Lambda(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+_0$ *, defined by*

$$\Lambda(\beta) = \begin{cases} 0, & 0 < \beta \le \beta_C \\ \hat{\lambda}, & \beta > \beta_C \end{cases}$$
(13)

is the locus of all possible critical points of the objective function $\varphi(\cdot, \cdot)$.

The proof is given in Sect. 4.

By inserting the function $\Lambda(\beta)$, as constructed in Eq. (13), into the original objective function Eq. (9) we obtain the "profile likelihood function"

$$\tilde{\varphi}(\beta) = \varphi(\Lambda(\beta), \beta) = \begin{cases} N \ln \beta - \beta S, & \beta \in (0, \beta_C] \\ N \ln \left(1 + \frac{1}{\Lambda(\beta)}\right) + N \ln \beta - N \ln \Lambda(\beta) \\ +\beta S - 2 \sum_{i=1}^{N} \ln \left(1 + \frac{X_i^{\beta}}{\Lambda(\beta)}\right), & \beta \in (\beta_C, \infty) \end{cases}$$
(14)

Deringer





Therefore, we have reduced the two-dimensional maximisation problem for $\varphi(\cdot, \cdot)$ on a finite region to a one-dimensional problem for $\tilde{\varphi}(\cdot)$ on a finite interval. Figure 1 illustrates the problem: the critical points (and therefore our maximum guaranteed by Theorem 1) are located inside the purple rectangle region and must be located on the red curve.² By the condition given by Eq. (12), critical points can only be maxima or saddle points. This simplification is important because it improves the speed of finding the critical points in our numerical studies dramatically. Unfortunately, we were not able to prove uniqueness for the critical points, as the necessary concavity arguments for the Hessian matrix seem intractable. Numerically, however, we always found exactly one critical point for millions of samples generated and this point was always a maximum.

A corollary below summarises the the two theorems discussed above.

Corollary 1 Under the assumptions of Theorems 1 and 2 the following holds true:

(1) If and only if $\beta_0 > \beta_C$ the MLE, Eqs. (6)–(7), possess a solution $(\hat{\lambda}, \hat{\beta}) \neq (0, 0)$ (which might not be unique). Critical points are obtained by constructing the function $\Lambda(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+_0$ as defined for all $\beta > 0$ in Theorem 2, Eq. (13), and then solving Eq. (7), which now reads

$$0 = \frac{N}{\beta} + \sum_{i=1}^{N} \ln X_i - 2 \sum_{i=1}^{N} \frac{X_i^{\beta} \ln X_i}{\Lambda(\beta) + X_i^{\beta}},$$
(15)

for the desired critical point $\hat{\beta} > \beta_C$.

(2) If and only if $\beta_0 \leq \beta_C$ the MLE, Eqs. (6)–(7), do not possess a solution. The likelihood attains a maximum on the boundary $(\hat{\lambda}, \hat{\beta}) = (0, \beta_0)$ with a Pareto

 $^{^2 \}beta_1 < \beta_2$ are the finite boundaries of our region in which we prove the existence of a solution in Sect. 4, see proof of Theorem 1, Fig. 6.

probability distribution given by

$$g(x|\beta_0) = \frac{\beta_0}{x^{1+\beta_0}}.$$
 (16)

The proof is given in Sect. 4.

Our final theorem demonstrates the condition that at least two observations in our sample be different is not only sufficient but also necessary. We see that otherwise no finite maximum of the MLE exists.

Theorem 3 (Necessity of at least two different observations) *Consider the i.i.d. left-truncated sample* $X_1 = X_2 = \cdots = X_N > 1$ *for which all observations are equal to* X_1 , say. Then the following holds true:

- (1) $\beta_0 = 1/\ln X_1$ and $\beta_C = (\ln 2)/\ln X_1$
- (2) For fixed $\beta > \beta_C$ the function

$$\Lambda(\beta) = X_1^\beta - 2$$

satisfies the ML equation for λ , Eq. (6).

(3) The profile likelihood function Eq. (14) reading here as

$$\tilde{\varphi}(\beta) = \varphi(\Lambda(\beta), \beta) = \begin{cases} N \ln \beta - N\beta \ln X_1, & \beta \in (0, \beta_C] \\ -N \ln \left(X_1^{\beta} - 1\right) + N \ln \beta \\ +N\beta \ln X_1 - 2N \ln 2, & \beta \in (\beta_C, \infty) \end{cases}$$

is continous and strictly monotonically increasing in β . It therefore does not possess a finite maximum.

The proof is given in Sect. 4.

Theorem 3, which can be also formulated in the same fashion for the untruncated log-logistic MLE, might explain the phenomenon observed in Shoukri et al. (1988, Table 1, p. 231): where the percentage of failures of their maximisation algorithm for the likelihood is given for small samples. Indeed, when a sample consists of N = 15 data points which are all "nearly equal", a direct optimisation by numerical methods is bound to fail if the parameter region chosen is too "small". The profile likelihood will be increasing in one direction towards the boundaries of the "small" parameter region and the Newton–Raphson solver of the MLE fails to converge. That the MLE problem in this case possesses a solution was established in Gupta et al. (1999).

2.4 Asymptotic analysis and confidence regions

Having established the existence of point estimators $(\hat{\lambda}, \hat{\beta})$ in the previous subsection under certain conditions, namely Eqs. (10) and (11), we now investigate the largesample properties of the MLE where the sample size tends to infinity. Our aim is to demonstrate consistency and asymptotic normality of the left-truncated log-logistic distribution as $N \to \infty$ using standard arguments provided by Newey and McFadden (1994).

2.4.1 Consistency

We first demonstrate that the left-truncated log-logistic probability density function is identifiable, i.e. for a compact set $K \subset \mathbb{R}^+ \times \mathbb{R}^+$ and two distinct $(\lambda_1, \beta_1) \neq$ $(\lambda_2, \beta_2) \in K$ we have according to Eq. (4) that $f_{LT}(x|\lambda_1, \beta_1) \neq f_{LT}(x|\lambda_2, \beta_2)$ for x > 0 with probability 1. To see this note that from $f_{LT}(x|\lambda_1, \beta_1) = f_{LT}(x|\lambda_2, \beta_2)$ we obtain with $x_L = 1$ and assuming without loss of generality $\beta_2 \geq \beta_1$ after some simple algebraic manipulations

$$\Delta(x) = x^{2\beta_2 + \beta_1} - C \frac{\lambda_2}{\lambda_1} x^{2\beta_1 + \beta_2} + 2(1 - C)\lambda_2 x^{\beta_1 + \beta_2} -C\lambda_1 \lambda_2 x^{\beta_2} + \lambda_2^2 x^{\beta_1} = 0,$$
(17)

where the constant $C = (1 + \lambda_2^{-1})\beta_2/[(1 + \lambda_1^{-1})\beta_1]$. The left-hand side of Eq. (17) defines a so called generalised polynomial $\Delta : \mathbb{R}^+ \to \mathbb{R}$ with real powers listed in descending order. Corollary 3.2 from Jameson (2006) asserts the existence of at most 5 positive zeros in our case because the number of terms in descending order is 5. Thus identifiability is proven.

Next we investigate $\mathbb{E}(\ln f_{LT}(x|\lambda, \beta))$ where the expectation $\mathbb{E}(\cdot)$ is defined with respect to the left-truncated log-logistic measure and some integrable function $g(\cdot)$,

$$\mathbb{E}(g(X)) = \int_{x_L}^{\infty} g(x) \left[1 + \left(\frac{x_L}{\alpha}\right)^{\beta} \right] \frac{1}{\left[1 + \left(\frac{x}{\alpha}\right)^{\beta} \right]^2} d\left(\frac{x}{\alpha}\right)^{\beta}, \quad (18)$$

whereby $x_L = 1$ and $\lambda = \alpha^{\beta}$. Use the following upper bound

$$|\ln f_{LT}(x|\lambda,\beta)| = |\ln (\lambda+1) + \ln \beta + (\beta-1)\ln x - 2\ln (\lambda+x^{\beta})|$$

$$\leq \ln (\lambda+1) + \ln \beta + \ln x^{\beta} + 2\ln (\lambda+x^{\beta}), \qquad (19)$$

substitute $y = x^{\beta}$ in Eq. (19) and also in the measure defined in Eq. (18) and because the integrals $\int_{1}^{\infty} dy(\ln y)/(\lambda + y)^{2} < \infty$, $\int_{1}^{\infty} dy(\ln (\lambda + y))/(\lambda + y)^{2} dy < \infty$ we have finally established that for all $(\lambda, \beta) \in K$ the following expectation is finite, $\mathbb{E}(|\ln f_{LT}(x|\lambda, \beta)|) < \infty$. We have therefore established the preconditions of Lemma 2.2 by Newey and McFadden (1994) that guarantees that $\mathbb{E}(\ln f_{LT}(X|\lambda, \beta))$ has a unique maximum. Thus the usual "information inequality" holds true for the lefttruncated log-logistic distribution. Recall next the log-likelihood function Eq. (5) for our i.i.d. sample X_1, \ldots, X_N left-truncated at 1 and suppressing the parameter x_L for convenience,

$$\frac{1}{N}\ln L_{LT}(\{X_1,\ldots,X_N\}|\lambda,\beta) = \ln(\lambda+1) + \ln\beta + (\beta-1)\frac{1}{N}\sum_{i=1}^N\ln X_i$$
$$-\frac{2}{N}\sum_{i=1}^N\ln(\lambda+X_i^\beta)$$

In this last equality, all terms on the right-hand side are continous functions in *X* for each $(\lambda, \beta) \in K$ with probability 1 and all the terms, respectively their absolute values can be bounded above by some functions for which the expectations $\mathbb{E}(\cdot)$ defined above in Eq. (18) are finite, namely $\mathbb{E}(\ln X) < \infty$ and $\mathbb{E}(\ln (\lambda + X^{\beta})) < \infty$ by similar arguments demonstrating the finiteness of the expectation in Eq. (19). Thus Lemma 2.4 by Newey and McFadden (1994) can be applied, yielding the following crucial result

$$\sup_{(\lambda,\beta)\in K} \left\| \frac{1}{N} \ln L_{LT}(\{X_1,\ldots,X_N\}|\lambda,\beta) - \{\ln(\lambda+1) + \ln\beta + (\beta-1)\mathbb{E}(\ln X) - 2\mathbb{E}(\ln(\lambda+X^\beta))\} \right\| \xrightarrow{p} 0$$
(20)

This convergence result Eq. (20) along with uniqueness established in the "information inequality" allows us to apply the general Theorem 2.1 by Newey and McFadden (1994) giving the desired weak consistency, i.e. as $N \to \infty$ we have $(\hat{\lambda}_N, \hat{\beta}_N) \xrightarrow{p} (\lambda, \beta)$ where the subscript *N* on the hatted quantities indicates the sample size and the arrow with superscript *p* indicates convergence in probability.

2.4.2 A remark on our existence criterion for $N \rightarrow \infty$

As the sample size grows, $N \to \infty$, our criterion for the existence of a ML estimator inside the parameter region in Theorem 1, namely Eqs. (10) and (11), converges by the law of large numbers in probability to the following limits

$$\frac{1}{N}\sum_{i=1}^{N}\ln X_{i} \xrightarrow{p} \mathbb{E}\left(\ln X\right), \quad \frac{1}{N}\sum_{i=1}^{N}X_{i}^{-\beta_{C}} \xrightarrow{p} \mathbb{E}\left(X^{-\beta_{C}}\right)$$

and our conditions for the existence of the point estimators become $\beta_0 > \beta_C$ with

$$\frac{1}{\beta_0} = \mathbb{E} \left(\ln X \right), \quad \frac{1}{2} = \mathbb{E} \left(X^{-\beta_C} \right)$$



Fig. 2 $\rho(\lambda)$ as solution for Eq. (22) for various λ



Fig. 3 The ratio β_C/β_0

Evaluating these expressions with the expectation given in Eq. (18) we obtain introducing a new variable $\rho = \beta_C / \beta$

$$\frac{\beta}{\beta_0} = \mathbb{E}\left(\ln X^{\beta}\right) = \frac{(1+\lambda)\ln\left(1+\lambda\right)}{\lambda} \tag{21}$$

$$\frac{1}{2} = \mathbb{E}\left([X^{\beta}]^{-\rho} \right) = (1+\lambda) \int_{1}^{\infty} y^{-\rho} \frac{1}{(\lambda+y)^{2}} dy$$
(22)

From Eq. (22) we can construct numerically a function $\rho(\lambda) : \mathbb{R}^+ \to \mathbb{R}^+$ by solving this equation for various values of λ . The graph is depicted in Fig.2. Now we

p\N	30	50	100	200	
0%	0%	0%	0%	0%	
25%	0.6%	0%	0%	0%	
50%	9%	5.3%	1.4%	0%	
75%	25.1%	22.4%	17%	10%	

Table 1 Frequency of "bad samples" not admitting a log-logistic MLE solution

rewrite Eq. (21) as follows

$$\frac{\beta_C}{\beta_0} = \rho(\lambda) \frac{(1+\lambda)\ln(1+\lambda)}{\lambda}$$
(23)

As we see from the graph in Fig. 3 the left-hand side of Eq. (23) is always strictly less than 1 for $\lambda > 0$ and thus our criterion for a true maximum of the maximum likelihood function is satisfied in the limit $N \rightarrow \infty$. For completeness we depict the statistics of "bad samples" not possessing a solution as provided by the criterion in Theorem 1 in Table 1 where *p* denotes the "truncation percentage".

2.4.3 Asymptotic normality

The consistency of the ML estimator serves as a precondition for the asymptotic normality upon which we will focus in this subsection. We return to the original variables $\theta = (\alpha, \beta)$, where $\alpha^{\beta} = \lambda$ and $x_L \ge 0$, because for these variables the Fisher information matrix for complete (untruncated) samples is diagonal [see also Reath et al. (2018)]. Let us also introduce for convenience of notation $\nabla_{\theta} = \left(\frac{\partial}{\partial \alpha}, \frac{\partial}{\partial \beta}\right)$ the gradient vector and $\nabla_{\theta\theta}$ the Hessian matrix of second derivatives. In this notation Theorem 3.3 by Newey and McFadden (1994) will provide the desired asymptotic normality result: key conditions are a sample of i.i.d. random variables and asymptotic consistency as well as additional five technical conditions, namely

- (i) the true parameters $\theta = (\alpha, \beta) \in interior(K)$ the compact parameter region,
- (ii) the pdf $f(x|\theta) = f(x|\alpha, \beta)$ is twice continuously differentiable and $f(x|\theta) > 0$ in a neighbourhood \mathcal{N} of the true parameters $\theta = (\alpha, \beta)$,
- (iii) $\int \sup_{\theta \in \mathcal{N}} ||\nabla_{\theta} f(x|\theta)|| dx < \infty, \int \sup_{\theta \in \mathcal{N}} ||\nabla_{\theta\theta} f(x|\theta)|| dx < \infty,$
- (iv) $\mathcal{I} = \mathbb{E}((\nabla_{\theta} \ln f(x|\theta) \nabla_{\theta} \ln f(x|\theta)))$ exists and is nonsingular,
- (v) $\mathbb{E}(\sup_{\theta \in \mathcal{N}} ||\nabla_{\theta \theta} \ln f(x|\theta)||) < \infty$ where $|| \cdot ||$ denotes the Euclidean norm.

We quickly check these conditions: (i) and (ii) are obviously satisfied because we only consider true log-logistic distributions with $\theta = (\alpha, \beta) \neq (0, 0)$. It is straight forward to verify the dominance conditions for (iii) and (v) using a practical criterion from Pratt (1960). For example (v) follows by noting that $\nabla_{\theta\theta} \ln f(x|\theta)$ is bounded, uniformly in bounded α and β and β bounded away from zero. For conditions (iv) we use the clever argument of Newey and McFadden (1994) namely the following

identity valid under the "usual" regularity conditions from above

 $\mathcal{I} = -\mathbb{E}\left(\nabla_{\theta,\theta} \ln f(x|\theta)\right) = \mathbb{E}\left(\nabla_{\theta} \ln f(x|\theta)\nabla_{\theta} \ln f(x|\theta)\right).$

From this we conclude that $\theta \mathcal{I} \theta^T > 0$ because $\theta \cdot \nabla_{\theta} \ln f(x|\theta)$ is a non-vanishing function in the argument x on an interval and thus we have strict positivity for all $\theta \neq 0$.

Thus all the necessary conditions of Theorem 3.3 Newey and McFadden (1994) are indeed satisfied and as the sample size gets larger, i.e. $N \to \infty$, the joint ML estimators $(\hat{\alpha}, \hat{\beta})$ converge in distribution to a multivariate normal distribution with mean (α, β) , the true parameters,³ and variance-covariance matrix whose inverse is proportional to the Fisher information matrix

$$\begin{bmatrix} \operatorname{var}(\hat{\alpha}) & \operatorname{cov}(\hat{\alpha}, \hat{\beta}) \\ \operatorname{cov}(\hat{\beta}, \hat{\alpha}) & \operatorname{var}(\hat{\beta}) \end{bmatrix} \approx \frac{1}{N} \begin{bmatrix} \mathcal{I}_{\alpha, \alpha} & \mathcal{I}_{\alpha, \beta} \\ \mathcal{I}_{\beta, \alpha} & \mathcal{I}_{\beta, \beta} \end{bmatrix}^{-1},$$
(24)

where the elements of the Fisher information matrix are given by:

$$\begin{split} \mathcal{I}_{\alpha,\alpha} &= -\mathbb{E}\left(\frac{\partial^2 \varphi}{\partial \alpha^2}\right), \\ \mathcal{I}_{\alpha,\beta} &= -\mathbb{E}\left(\frac{\partial^2 \varphi}{\partial \alpha \partial \beta}\right) = -\mathbb{E}\left(\frac{\partial^2 \varphi}{\partial \beta \partial \alpha}\right) = \mathcal{I}_{\beta,\alpha}, \\ \mathcal{I}_{\beta,\beta} &= -\mathbb{E}\left(\frac{\partial^2 \varphi}{\partial \beta^2}\right). \end{split}$$

with the expectations $\mathbb{E}(\cdot)$ defined above in Eq. (18).

For given left-truncation values $x_L > 0$, we have already introduced parameter, $\eta = (x_L/\alpha)^{\beta}$, after some computations (for details see Appendix A) the desired results for expectations can be given

$$\mathbb{E}\left(\frac{\partial^{2}\varphi}{\partial\alpha^{2}}\right) = N\frac{\beta}{\alpha^{2}}\left\{-\frac{\beta}{3(1+\eta)^{2}}\right\}$$
$$\mathbb{E}\left(\frac{\partial^{2}\varphi}{\partial\alpha\partial\beta}\right) = -\frac{N}{\alpha}\left\{\frac{\eta\left[1+\eta-\eta\left(3+\eta\right)\ln\left(\frac{1+\eta\right)}{\eta}+3\ln(\eta)\right]-(1+3\eta)\ln(1+\eta)}{3(1+\eta)^{2}}\right\}$$
$$= \mathbb{E}\left(\frac{\partial^{2}\varphi}{\partial\beta\partial\alpha}\right)$$
$$\mathbb{E}\left(\frac{\partial^{2}\varphi}{\partial\beta^{2}}\right) = -\frac{N}{3\beta^{2}}\left\{1+\frac{\pi^{2}}{3}(1+\eta)+2(1+\eta)\operatorname{PolyLog}[2,-\eta]-2\frac{\eta}{(1+\eta)}\ln(\eta)\right\}$$

³ $\hat{\theta}$ is asymptotically normal if $\sqrt{N}(\hat{\theta} - \theta) \xrightarrow{d} N(0, I^{-1}(\theta))$ where I^{-1} is called asymptotic variancecovariance of the $\hat{\theta}$. Asymptotic normality imply that the estimator not only converges to the unknown parameters, but it converges fast enough, at a rate $1/\sqrt{N}$ (Newey and McFadden 1994).

+2 ln
$$\eta$$
 ln(1 + η) - $\frac{\left[(1 + \eta)^3 - 1\right]}{(1 + \eta)^2}$ ln²(η) }. (25)

In particular, we obtain for the untruncated case, i.e. $x_L = \eta = 0$, the well-known results of Reath et al. (2018)

$$\mathbb{E}\left(\frac{\partial^2 \varphi}{\partial \alpha^2}\right) = -N \frac{\beta^2}{3\alpha^2},$$
$$\mathbb{E}\left(\frac{\partial^2 \varphi}{\partial \beta \partial \alpha}\right) = 0,$$
$$\mathbb{E}\left(\frac{\partial^2 \varphi}{\partial \beta^2}\right) = -\frac{N}{\beta^2}\left(1 + \frac{(\pi^2 - 6)}{9}\right).$$

For large *N* the asymptotic $(1 - \xi)100$ percent elliptical confidence region for the parameters α and β , the celebrated Wald test, is given by [see e.g. Theorem 6.3.4 Bickel and Doksum (2015), Wingo (1989)]

$$\hat{\mathcal{I}}_{\alpha,\alpha}(\alpha-\hat{\alpha})^2 + 2\hat{\mathcal{I}}_{\alpha,\beta}(\alpha-\hat{\alpha})(\beta-\hat{\beta}) + \hat{\mathcal{I}}_{\beta,\beta}(\beta-\hat{\beta})^2 \leq \chi^2(1-\xi;2), \quad (26)$$

where the hat-symbol "~" indicates that the Fisher matrix is evaluated at the estimated parameters $(\hat{\alpha}, \hat{\beta})$ and χ^2 is the Chi-square distribution with 2 degrees of freedom because the two parameters α and β are estimated. We shall later choose $\xi = 0.95$ and thus $\chi^2(0.05; 2) \approx 5.991$. Inequality Eq.(26) defines the general confidence region⁴ related to our point estimate $(\hat{\alpha}, \hat{\beta})$. When the off-diagonal elements of the Fisher matrix vanish, $\mathcal{I}_{\alpha,\beta} = \mathcal{I}_{\beta,\alpha} = 0$, we can use the diagonal elements for the confidence intervals for α and β respectively. For the 95%-confidence interval of a standard Gaussian we have a multiplication factor 1.96 so that we find in this case for the confidence intervals from Eq. (24)

$$\hat{\alpha} \pm 1.96\sqrt{\operatorname{var}(\hat{\alpha})} \quad \text{and} \quad \hat{\beta} \pm 1.96\sqrt{\operatorname{var}(\hat{\beta})}.$$
 (27)

In this case, the confidence region is an ellipse whose axes are in x- respectively y-direction.

3 Applications to real world data

3.1 Bladder cancer data set from Lee and Wang (2003)

We apply our methods to the bladder cancer data set from Lee and Wang (2003), Table 9.3, p. 231, which has been studied recently in Al-Shomrani et al. (2016) under the hypothesis that the data are distributed according to a log-logistic distribution.

⁴ We choose α as x-axis and β as y-axis.

<i>x</i> _{<i>L</i>} [m]	Ν	<i>α̂</i> [m]	\hat{eta}	ln L
0	128(137)	$5.97 \pm 0.09 (5.98 \pm 0.09)$	$1.695 \pm 0.022 (1.700 \pm 0.022)$	-410.89(-439.57)
0.25	126(135)	$6.11 \pm 0.09 (6.10 \pm 0.09)$	$1.782 \pm 0.022 (1.779 \pm 0.022)$	-402.20(-430.84)
1	120(128)	$6.32\pm0.10(6.37\pm0.10)$	$1.877 \pm 0.023 (1.891 \pm 0.022)$	-379.28(-404.74)
6	64(68)	$8.63 \pm 0.30 (9.01 \pm 0.31)$	$2.239 \pm 0.056 (2.291 \pm 0.058)$	-206.00(-219.49)

Table 2 Remission times for log-logistic distribution for reduced and original bladder cancer data set

The original Table 9.3 contains a set of 137 remission times in months from cancer patients whereas in Al-Shomrani et al. (2016), only 128 data points from this table are considered because nine observations (indicated by a "+" in the original Table 9.3) have been removed. We shall consider the original data set of Lee and Wang (2003) as well as the reduced data subset of Al-Shomrani et al. (2016):

 $\begin{array}{l} 0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02,\\ 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26,\\ 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82,\\ 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90,\\ 2.69, 4.18, 5.34, 7.59, 0.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62,\\ 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25,\\ 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93,\\ 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 12.07, 3.25, 4.50, 6.25,\\ 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 21.73, 2.07,\\ 3.36, 6.93, 8.65, 12.63, 22.69[, 4.65+, 0.87+, 24.80+, 10.86+, 4.70+, 3.02+,\\ 19.36+, 8.60+, 3.33+]. \end{array}$

Using our method, we estimate the parameters of the log-logistic distribution. This was done for both the complete and truncated data set with various truncation points for the original and the reduced bladder cancer set. Our results are displayed in Table 2 (the values in round brackets are for the original data set). Note that for the purpose of comparison we use the scale parameter $\hat{\alpha} = \hat{\lambda}^{1/\hat{\beta}}$. The confidence intervals for the parameters are obtained from Eq. (27) and because the off-diagonal elements of the Fisher information matrix become relevant for increasing truncation we also provide the confidence ellipses according to Eq. (26) in Fig. 4 for the reduced data set of Al-Shomrani et al. (2016) only.

We may compare our result without truncation with Al-Shomrani et al. (2016), which found $\hat{\alpha} = 6.08982$ months and $\hat{\beta} = 1.725158$, with a log-likelihood function value of -411.4574. These authors were using "LLmodel.Optim() function in R with Newton–Raphson options [...] as an iterative process for maximising the log-likelihood function". Our likelihood value being -410.89 is better and will lead to a better estimation of the parameters $\hat{\alpha}$ and $\hat{\beta}$.



Fig. 4 Error ellipses for bladder cancer data (Lee and Wang 2003) with various left-truncations

<i>x</i> _{<i>L</i>} [mm]	Ν	<i>α̂</i> [mm]	\hat{eta}	ln L
0	141	564.0 ± 1.2	11.8 ± 0.1	-825.30
300	141	563.9 ± 1.2	11.8 ± 0.1	-825.22
400	137	565.9 ± 1.2	12.1 ± 0.1	-791.86
500	113	574.5 ± 1.5	13.6 ± 0.2	-616.60
	x _L [mm] 0 300 400 500	x_L [mm] N 0 141 300 141 400 137 500 113	x_L [mm] N $\hat{\alpha}$ [mm] 0 141 564.0 ± 1.2 300 141 563.9 ± 1.2 400 137 565.9 ± 1.2 500 113 574.5 ± 1.5	x_L [mm] N $\hat{\alpha}$ [mm] $\hat{\beta}$ 0 141 564.0 ± 1.2 11.8 ± 0.1 300 141 563.9 ± 1.2 11.8 ± 0.1 400 137 565.9 ± 1.2 12.1 ± 0.1 500 113 574.5 ± 1.5 13.6 ± 0.2

3.2 Annual precipitation for Berlin and Toronto

In the literature the log-logistic distributions is often used to describe precipitation data, for example Shoukri et al. (1988) and Ashkar and Mahdi (2006). For reasons of homogeneity, Shoukri et al. (1988) focused on precipitation data of various Canadian cities. We have chosen the annual precipitation for the city-state of Berlin, Germany as an example, whose data are likely to be i.i.d. because of its small area. The data are available from the German met office, Deutscher Wetterdienst (2022). The annual precipitation data consist of 141 data points starting in 1881 and extending to 2021. Also, to compare our method with Shoukri et al. (1988), we analyse the annual precipitation data for Toronto from 1 July to 30 June annually, as given in Toronto Weather Statistics (2022), from 1937 until 2021. For Berlin our results are given in Table 3 and for Toronto in Table 4. Note that in our analysis the shape parameters, $\hat{\beta}$, are similar for both cities as can be seen nicely in Fig. 5. We compare our results with Shoukri et al. (1988), where for Toronto $\hat{\alpha} = 789.8$ mm and $\hat{\beta} = 14.4$ were given. Whereas the estimates of the scale parameter $\hat{\alpha}$ are in good agreement, the values of $\hat{\beta}$ are different, but of similar order of magnitude.

Table 4 Toronto precipitationdata for log-logistic distribution	<i>x</i> _{<i>L</i>} [mm]	Ν	<i>α̂</i> [mm]	β	ln L
	0	85	771.5 ± 2.7	11.6 ± 0.2	-526.58
	300	85	771.5 ± 2.7	11.6 ± 0.2	-526.58
	400	84	774.1 ± 2.5	12.3 ± 0.2	-514.65
	500	83	775.9 ± 2.5	12.7 ± 0.2	-504.75



Fig. 5 Error ellipses for precipitation data for Berlin and Toronto

3.3 Outlook

The log-logistic distribution is heavier tailed than the log-normal distribution and enjoys a closeness to the Pareto distribution. For modelling the distribution of wealth and income, a left-truncated version of the log-logistic distribution Fisk (1961) is a sensible approach, because in many economies, very low incomes (below a certain threshold) are often earned in a shadow economy. Therefore, they are not captured in official government statistics. We have shown that the log-logistic hypothesis can also explain certain cancer survival data very well and accurately. We can also confirm that in general the log-logistic distribution describes precipitation data regardless of their region fairly well. It is noteworthy that even in recent daily precipitation value if the missing data were replaced by zero. While the effect of this might be negligible for higher annual precipitation it may have a more significant impact for lower annual precipitation. A remedy for the statistical analysis could be the introduction of a lefttruncation point $x_L > 0$ to overcome this problem.

How would one chose a left-truncation point $x_L > 0$ in practical applications? Whenever there is a gap in the observed data between 0 and some threshold value this threshold value could be an obvious candidate for the left-truncation point. In the case of income distribution this could be the value of the tax allowance because below this value the statistical data will be most likely erroneous or incomplete. For the medical example of a patient dying a few days after the fatal diagnosis the reason for the patient's death might not be the illness itself but maybe suicide or a fatal accident. Here a left truncation point bigger than 0.5 months might serve in order to single out these events. For the case of precipitation one could take as a rule of thumb the value of 0.1 mm/h serving as a measurement threshold (see Schnepper et al. (2023)). Hence for a 24 h day one might take 2.4 mm as a value for x_L , which is quite large compared to the typical annual precipitation of 75 mm for the dry regions of Saudi Arabia.

Finally in the case of mixed distributions where one mixture component is light tailed and the other is heavier tailed, such as log-logistic, our method will allow us to estimate the parameters of the log-logistic component with a good degree of accuracy. Thus, one has a tool at hand to decide whether or not a two-component mixture contains a log-logistic component by employing an Expectation-Maximization algorithm for determining the mixture component with better initial parameters to ensure quick and accurate convergence.

4 Proofs for theorems from Sect. 2

4.1 Proof of Theorem 1

We first begin with a lemma dealing with "small" λ for the objective function given in Eq. (9). Keep in mind that we have chosen $x_L = 1$.

Lemma 2 For any $\beta > 0$, there exists some $\lambda_0 > 0$ such that for $\lambda \in [0, \lambda_0)$ we have

$$\varphi(\lambda,\beta) = \varphi(0,\beta) + N\left\{-\frac{2}{N}\sum_{i=1}^{N}\frac{1}{X_i^{\beta}} + 1\right\}\lambda + O(\lambda^2),\tag{28}$$

where

$$\varphi(0,\beta) = \lim_{\lambda \to 0+} \varphi(\lambda,\beta) = \sum_{i=1}^{N} \left[-\beta \ln X_i + \ln \beta \right].$$
⁽²⁹⁾

Proof For fixed $\beta > 0$ we obtain by differentiating Eq. (9) with respect to λ

$$\begin{split} \frac{\partial}{\partial\lambda}\,\varphi(\lambda,\beta) &= -\frac{N}{\lambda}\frac{\frac{1}{\lambda}}{1+\frac{1}{\lambda}} - \frac{N}{\lambda} + \frac{2}{\lambda}\sum_{i=1}^{N}\frac{\frac{X_{i}^{\rho}}{\lambda}}{1+\frac{X_{i}^{\beta}}{\lambda}} \\ &= \frac{N}{\lambda}\left\{\frac{2}{N}\sum_{i=1}^{N}\frac{1}{1+\frac{\lambda}{X_{i}^{\beta}}} - 1 - \frac{1}{1+\lambda}\right\} \end{split}$$

🖉 Springer

$$= \frac{N}{\lambda} \left\{ -\frac{2}{N} \sum_{i=1}^{N} \frac{1}{X_i^{\beta}} \lambda + \lambda + O(\lambda^2) \right\},\tag{30}$$

where for $\lambda \in (0, \lambda_0)$ with $0 < \lambda_0 < 1$ the geometric series is absolutely convergent. Thus integrating Eq. (30) over λ' from 0 to λ we get

$$\varphi(\lambda,\beta) - \varphi(0,\beta) = \lim_{\varepsilon \to 0} \int_{\varepsilon}^{\lambda} N\left\{-\frac{2}{N}\sum_{i=1}^{N}\frac{1}{X_{i}^{\beta}} + 1 + O(\lambda')\right\}d\lambda'$$
$$= N\lambda\left\{-\frac{2}{N}\sum_{i=1}^{N}\frac{1}{X_{i}^{\beta}} + 1\right\} + O(\lambda^{2}).$$

The computation of $\varphi(0, \beta) = \lim_{\varepsilon \to 0} \varphi(\varepsilon, \beta)$ is straight forward.

In the next lemma we derive a result analogous to a result in Gupta et al. (1999) about the behaviour of the objective function $\varphi(\lambda, \beta)$ in a certain rectangle in the first quadrant of the (λ, β) -coordinate plane. The situation is depicted in Fig.6.

Lemma 3 Consider an i.i.d. left-truncated sample $X_1, ..., X_N$ all bigger than one and for which at least two observations are different. For any M > 0 we have the following inequalities valid for $\varphi(\lambda, \beta)$ defined in Eq. (9):

- (1) There exists a $\beta_1 > 0$ such that for $\beta \in (0, \beta_1)$ we have $\varphi(\lambda, \beta) < -M$, independent of λ .
- (2) There exists a $\beta_2 > \beta_1$ such that for $\beta > \beta_2$ we have $\varphi(\lambda, \beta) < -M$, independent of λ .
- (3) For $\beta \in [\beta_1, \beta_2]$ there exists a $\lambda_1 > 0$ such that for $\lambda > \lambda_1$ we also have $\varphi(\lambda, \beta) < -M$.

Proof

(1): We write Eq. (9) as

$$\varphi(\lambda,\beta) = N \ln\left(1+\frac{1}{\lambda}\right) + N \ln\beta - N \ln\lambda + \beta S - 2\sum_{i=1}^{N} \ln\left[1+\frac{X_{i}^{\beta}}{\lambda}\right]$$
$$= \sum_{i=1}^{N} \ln\frac{1+\frac{1}{\lambda}}{1+\frac{X_{i}^{\beta}}{\lambda}} + N \ln\beta + \beta S + \sum_{i=1}^{N} \ln\frac{\frac{1}{\lambda}}{1+\frac{X_{i}^{\beta}}{\lambda}}$$
$$= \sum_{i=1}^{N} \ln\frac{\lambda+1}{\lambda+X_{i}^{\beta}} + N \ln\beta + \beta S + \sum_{i=1}^{N} \ln\frac{1}{\lambda+X_{i}^{\beta}}$$
$$< N \ln\beta + \beta S$$
(31)

because $\lambda + X_i^{\beta} > \lambda + 1 > 1$. From the bound Eq. (54) we see that $\beta_1 > 0$ can be chosen independent of $\lambda \ge 0$ and the first statement is proven.



(3): For $\lambda > 1$ and $\beta \in [\beta_1, \beta_2]$ we have

$$\varphi(\lambda,\beta) \le N \ln (1+1) + N \ln \beta_2 - N \ln \lambda + \beta_2 S \tag{32}$$

because the last term in Eq. (9) is always negative and the argument of the logarithm is always bigger than 1. From inequality Eq. (56) we see the existence of a sufficiently large $\lambda_1 > 1$ such that the third statement follows.

Now we have all the ingredients for the proof of Theorem 1.

Proof of Theorem 1 We start the proof by looking at Fig. 6.

For statement (1) apply Lemma 3 to conclude that outside the shaded area our objective function $\varphi(\lambda, \beta) < -M$ for some arbitrarily large M > 0, i.e. our objective function gets arbitrarily small. Thus a global maximum of the twice differentiable objective function $\varphi(\cdot, \cdot)$ is either inside the shaded region $[0, \lambda_1] \times [\beta_1, \beta_2]$ or on the boundary with $\lambda = 0$, somewhere in the closed interval $[\beta_1, \beta_2]$. One easily computes the only critical point and thus a candidate for a maximum of the objective function on this boundary-line to be β_0 , as given in Eq. (11). By the condition $\beta_0 > \beta_C$ and Lemma 2 this can be excluded: there exists some small $\lambda_0 > 0$ such that for $\lambda \in (0, \lambda_0)$

$$\varphi(\lambda, \beta_0) > \varphi(0, \beta_0)$$

and statement (1) is proven.

For statement (2) we use the following argument: by continuity there exists an $\epsilon > 0$ (first line) and by Lemma 2 a positive λ_0 (second line) such that

$$\varphi(0, \beta_0) = \max \left\{ \varphi(0, \beta) | \beta \in [\beta_0 - \epsilon, \beta_0 + \epsilon] \right\}$$

$$\geq \max \left\{ \varphi(\lambda, \beta) | \beta \in [\beta_0 - \epsilon, \beta_0 + \epsilon], \lambda \in [0, \lambda_0] \right\}$$

and thus β_0 is local maximum.

4.2 Proof of Theorem 2

For $\lambda > 0$ we start from the rewritten Eq. (6)

$$0 = \frac{2}{N} \sum_{i=1}^{N} \frac{X_i^{\beta}}{\lambda + X_i^{\beta}} - 1 - \frac{1}{\lambda + 1}$$
(33)

and investigate this equation Eq. (33) in greater detail. We firstly note the following:

Lemma 4 Consider a left-truncated log-logistic sample X_1, \ldots, X_N all bigger than 1. Define the function $q : \mathbb{R}_+ \to (0, 1)$ by

$$q\left(\beta|(X_1,\ldots,X_N)\right) \equiv \frac{1}{N}\sum_{i=1}^N \left(\frac{1}{X_i}\right)^\beta.$$

Then the following holds true:

- (1) $q(\cdot)$ is a monotonically decreasing function.
- (2) The equation $q(\beta|(X_1, ..., X_N)) = \frac{1}{2}$ has a unique positive solution β_C .
- (3) For $\beta \in (0, \beta_C]$ Eq. (33) only possesses the trivial solution $\hat{\lambda} = \Lambda(\beta) = 0$ and for $\beta > \beta_C$ in addition exactly one positive solution $\hat{\lambda} = \Lambda(\beta) > 0$.
- (4) For any fixed $\beta > 0$ all the roots $\hat{\lambda} = \hat{\lambda}(\beta)$ of the MLE Eq. (33) depend continuously on the sample (X_1, \ldots, X_N) .
- (5) For a fixed sample $(X_1, ..., X_N)$ all the roots $\hat{\lambda}$ of the MLE Eq. (33) depend continuously on $\beta \in (0, +\infty)$. In addition, we have $\lim_{\beta \to \beta_C +} \Lambda(\beta) = 0$, where $\Lambda(\beta)$ is defined in

part (3) of the lemma.

Proof

(1) Note that $0 < 1/X_i < 1...$ and therefore $(1/X_i)^{\beta}$ is monotonically decreasing and the first statement follows.

(2) Next note that because of

$$\lim_{\beta \to 0+} q \left(\beta | (X_1, \dots, X_N) \right) = 1,$$
$$\lim_{\beta \to \infty} q \left(\beta | (X_1, \dots, X_N) \right) = 0,$$

and the monotonicity of $q(\cdot)$ the second statement follows immediately.

(3) To prove the third statement we start from the likelihood equation, Eq. (33), defining for fixed $\beta > 0$ for convenience $Y_i = X_i^{\beta}$,

$$1 + \frac{1}{\lambda + 1} = \frac{2}{N} \sum_{i=1}^{N} \frac{Y_i}{\lambda + Y_i}.$$
(34)

🖉 Springer

Next define by the right-hand side of Eq. (34), the function

$$k(\lambda) = \frac{2}{N} \sum_{i=1}^{N} \frac{Y_i}{\lambda + Y_i}$$
(35)

which is monotonically decreasing in $\lambda \in [0, \infty)$ because

$$k'(\lambda) = -\frac{2}{N} \sum_{i=1}^{N} \frac{Y_i}{(\lambda + Y_i)^2} < 0.$$
(36)

Afterwards define a function by the left-hand side of Eq. (34) as

$$\ell(\lambda) = 1 + \frac{1}{\lambda + 1} \tag{37}$$

which is also monotonically decreasing in $\lambda \in [0, \infty)$ because

$$\ell'(\lambda) = -\frac{1}{(\lambda+1)^2} < 0.$$
(38)

Note that we have

$$\lim_{\lambda \to 0+} k(\lambda) = 2,$$
$$\lim_{\lambda \to +\infty} k(\lambda) = 0,$$

and also

$$\lim_{\lambda \to 0+} \ell(\lambda) = 2,$$
$$\lim_{\lambda \to +\infty} \ell(\lambda) = 1.$$

Thus the trivial solution $\hat{\lambda} = 0$ is always a solution. Next, we consider three cases:

Case I $k'(0) < \ell'(0)$:

In this case according to Eqs. (36) and (38) the function $k(\cdot)$ falls faster than the function $\ell(\cdot)$, at least for argument λ near 0. If this is true for all $\lambda > 0$, i.e. $k'(\lambda) < \ell'(\lambda)$, then clearly there will not be any intersection and $\hat{\lambda} = 0$ is the only solution of Eq. (33). If the latter is not true, then the situation will be as depicted in Fig.7:

In this case $k(\cdot)$ initially falls quicker than $\ell(\cdot)$ and then slows down to intersect at a point $\lambda_1 > 0$ and possibly a second time at a point $\lambda_2 > \lambda_1$. Thus we have for $\lambda_1 > 0$ that the function graph of $k(\cdot)$ crosses the function graph of $\ell(\cdot)$ from below, meaning that the slopes satisfy the following inequality

$$k'(\lambda_1) > \ell'(\lambda_1).$$



Fig. 7 Case I: The functions $k(\lambda)$ (golden) and $\ell(\lambda)$ (purple)

By the intermediate value theorem there will exist a positive number $\lambda^* \in (0,\lambda_1)$ such that

$$k'(\lambda^*) = \ell'(\lambda^*). \tag{39}$$

We shall show that this will lead to a contradiction as follows. We write Eq. (39) in detail as

$$-\frac{2}{N}\sum_{i=1}^{N}\frac{Y_i}{(\lambda^*+Y_i)^2} = -\frac{1}{(\lambda^*+1)^2}$$

or rearranging

$$\frac{1}{N}\sum_{i=1}^{N}Y_i\frac{(\lambda^*+1)^2}{(\lambda^*+Y_i)^2} = \frac{1}{2}.$$
(40)

Define the function $h(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$h(\lambda) = \frac{1}{N} \sum_{i=1}^{N} Y_i \frac{(\lambda+1)^2}{(\lambda+Y_i)^2}.$$
 (41)

D Springer

Thus Eq. (40) is equivalent to $h(\lambda^*) = \frac{1}{2}$. Note that

$$h'(\lambda) = \frac{1}{N} \sum_{i=1}^{N} Y_i \frac{2(\lambda+1)(\lambda+Y_i) - 2(\lambda+1)^2}{(\lambda+Y_i)^3} > 0$$
(42)

because all terms are positive. In particular, the numerator is positive because $Y_i > 1$. Note that we have the following limit

$$\lim_{\lambda \to \infty} h(\lambda) = \frac{1}{N} \sum_{i=1}^{N} Y_i > 1,$$
$$\lim_{\lambda \to 0+} h(\lambda) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{Y_i} < 1.$$

Hence for the equation $h(\lambda) = \frac{1}{2}$ to have a solution in $(0, \lambda_1)$ we need to require that the second limit is smaller than $\frac{1}{2}$, i.e.

$$\frac{1}{N}\sum_{i=1}^N\frac{1}{Y_i}<\frac{1}{2}\iff k'(0)>\ell'(0),$$

which is a contradiction. Thus, in Case I we have as the only real solution $\hat{\lambda} = \lambda_0 = 0$. Note that Case I corresponds to $\beta < \beta_C$.

Case II: $k'(0) > \ell'(0)$:

Here the function $k(\cdot)$ decreases slower than the function $\ell(\cdot)$ for λ near 0 and only here can the graph of $k(\cdot)$ intersect the graph of $\ell(\cdot)$ from above at $\lambda_1 > 0$ before tending to zero as λ gets larger, as depicted in Fig. 8.

We need to show now that this intersection is unique. We note that at λ_1 the graph of $k(\cdot)$ crosses the graph of $\ell(\cdot)$ from above, i.e.

$$k'(\lambda_1) < \ell'(\lambda_1). \tag{43}$$

However, if there were a second crossing point, $\lambda_2 > \lambda_1$, the graph of $k(\cdot)$ would cross the graph of $\ell(\cdot)$ from below, i.e.

$$k'(\lambda_2) > \ell'(\lambda_2).$$

As in Case I we conclude by the intermediate value theorem that there exists a $\lambda^* \in (\lambda_1, \lambda_2)$ such that

$$k'(\lambda^*) = \ell'(\lambda^*)$$



Fig. 8 Case II: the functions $k(\lambda)$ (golden) and $\ell(\lambda)$ (purple)

which is equivalent to

$$\frac{1}{N}\sum_{i=1}^{N}Y_{i}\frac{(\lambda^{*}+1)^{2}}{(\lambda^{*}+Y_{i})^{2}} = \frac{1}{2}$$
(44)

and this in turn is similar to Eq. (40). We use the function $h(\cdot)$, as defined in Eq. (41), to obtain the following limits:

$$\lim_{\lambda \to +\infty} h(\lambda) = \frac{1}{N} \sum_{i=1}^{N} Y_i > 1,$$
$$\lim_{\lambda \to \lambda_1} h(\lambda) > \frac{1}{2}$$

showing that Eq. (44) has no solution. The last limit follows from Eq. (43) and the definition of $h(\cdot)$ from Eq. (44)

$$-\frac{2}{N}\sum_{i=1}^{N}\frac{Y_i}{(\lambda_1+Y_i)^2} < -\frac{1}{(\lambda_1+1)^2} \iff k'(\lambda_1) < \ell'(\lambda_1).$$

Thus, in this case there exists only one positive root λ_1 of Eq. (33). Case II corresponds to $\beta > \beta_C$.

☑ Springer

Case III: $k'(0) = \ell'(0)$:

Here, one needs to look at the next higher derivatives, i.e. $k''(0) > \ell''(0)$, which reads here as $1 < \frac{2}{N} \sum_{i=1}^{N} \frac{1}{Y_i^2}$ or

$$\frac{1}{2} < \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{Y_i} \right)^2 = q(2\beta).$$

This cannot be true if

$$\frac{1}{2} = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{Y_i} \right) = q(\beta),$$

because of the properties of the above defined function $q(\cdot)$. Note that Case III corresponds to $\beta = \beta_C$. Thus, the third statement has been proven.

(4) For the fourth statement rewrite the MLE Eq. (33) as in Eq. (34) and multiply both sides of the equation with $(\lambda + 1) \prod_{i=1}^{N} (\lambda + Y_i)$. Thus, after rearranging one obtains a polynomial of (N + 1)-th order in the variable λ

$$\lambda^{N+1} + b_N(Y_1, ..., Y_N)\lambda^N + ... + b_0(Y_1, ..., Y_N) = 0,$$
(45)

where the coefficients $b_i(Y_1, ..., Y_N)$ are certain continuous polynomial functions in $(Y_1, ..., Y_N)$ obtained in the process of rearranging terms and sorting by the powers of λ . From Appendix A in Ostrowski (1960) or Harris and Martin (1987), the roots $\hat{\lambda}$ of some algebraic equation (i.e. a polynomial equation) depend continuously on the coefficients and the statement is proven.

(5) The last statement is shown in a similar manner to the previous statement and because all X_i^{β} are continuous functions in β respectively, the desired statement is proven because for $\beta \leq \beta_C$ the only real root is $\hat{\lambda} = 0$.

The next lemma provides some additional information about the function $\Lambda(\cdot)$ defined implicitly by Eq. (33) via Lemma 4.(3).

Lemma 5 Consider a left-truncated log-logistic sample $X_1, ..., X_N$ all bigger than 1. Then the non-zero solution of the MLE Eq. (33) $\Lambda(\cdot) : (\beta_C, \infty) \to \mathbb{R}^+$ has the following properties:

(1) For fixed $\beta \Lambda(\beta)$ is a root of Eq. (33) with multiplicity 1.

(2) $\Lambda(\cdot)$ is an analytic function in the range of (β_C, ∞) .

Proof Recall that $x_L = 1$. From Eq. (5) respectively Eq. (9) we obtain for fixed $\beta \in (\beta_C, +\infty)$ by direct computation

$$\frac{\partial}{\partial \lambda} \varphi(\lambda, \beta) \Big|_{\Lambda(\beta)} = \frac{N}{\Lambda(\beta)} \left[\frac{2}{N} \sum_{i=1}^{N} \frac{X_i^{\beta}}{\lambda + X_i^{\beta}} - 1 - \frac{1}{\lambda + 1} \right]_{\Lambda(\beta)} = 0,$$

🖉 Springer

$$\frac{\partial^2}{\partial\lambda^2} \varphi(\lambda,\beta) \bigg|_{\Lambda(\beta)} = -\frac{N}{\Lambda(\beta)} \left[\frac{2}{N} \sum_{i=1}^N \frac{X_i^\beta}{(\lambda+X_i^\beta)^2} - \frac{1}{(\lambda+1)^2} \right]_{\Lambda(\beta)},$$

where the first derivative is equal to zero for $\lambda = \Lambda(\beta)$ by construction. Next we define a polynomial function $p(\cdot) : \mathbb{R}_0^+ \to \mathbb{R}$ by

$$p(\lambda) \equiv \left[\frac{2}{N}\sum_{i=1}^{N}\frac{X_{i}^{\beta}}{\lambda+X_{i}^{\beta}} - \frac{\lambda+2}{\lambda+1}\right](\lambda+1)\prod_{j=1}^{N}(\lambda+X_{j}^{\beta})$$
$$= \frac{2}{N}\sum_{i=1}^{N}X_{i}^{\beta}(\lambda+1)\prod_{\substack{1\leq j\leq N\\j\neq i}}(\lambda+X_{j}^{\beta}) - (\lambda+2)\prod_{j=1}^{N}(\lambda+X_{j}^{\beta}).$$
(46)

The first claim of the lemma states that $\Lambda(\beta)$ is a root of $p(\cdot)$, Eq. (46), but not of $p'(\cdot)$. By construction we have indeed $p(\Lambda(\beta)) = 0$. We show that the assumption $p'(\Lambda(\beta)) = 0$ leads to a contradiction. Compute the derivative in Eq. (46)

$$\begin{split} p'(\lambda) &= \frac{2}{N} \sum_{i=1}^{N} X_i^{\beta} \prod_{\substack{1 \le j \le N \\ j \ne i}} (\lambda + X_j^{\beta}) + \frac{2}{N} \sum_{i=1}^{N} X_i^{\beta} (\lambda + 1) \sum_{k=1}^{N} \prod_{\substack{1 \le j \le N \\ j \ne i, j \ne k}} (\lambda + X_j^{\beta}) \\ &- \prod_{j=1}^{N} (\lambda + X_j^{\beta}) - (\lambda + 2) \sum_{k=1}^{N} \prod_{\substack{1 \le j \le N \\ j \ne k}} (\lambda + X_j^{\beta}). \end{split}$$

Then write down $p'(\Lambda(\beta)) / \prod_{j} (\Lambda(\beta) + X_{j}^{\beta})$, which by our assumption $p'(\Lambda(\beta)) = 0$ vanishes

$$0 = \frac{2}{N} \sum_{i=1}^{N} \frac{X_i^{\beta}}{\Lambda(\beta) + X_i^{\beta}} + \frac{2}{N} (\Lambda(\beta) + 1) \sum_{i=1}^{N} \frac{X_i^{\beta}}{\Lambda(\beta) + X_i^{\beta}} \sum_{k=1}^{N} \frac{1}{\Lambda(\beta) + X_k^{\beta}}$$
$$-1 - (\Lambda(\beta) + 2) \sum_{k=1}^{N} \frac{1}{\Lambda(\beta) + X_k^{\beta}}$$
$$= \frac{2}{N} \sum_{i=1}^{N} \frac{X_i^{\beta}}{\Lambda(\beta) + X_i^{\beta}} \left[1 + (\Lambda(\beta) + 1) \sum_{k=1}^{N} \frac{1}{\Lambda(\beta) + X_k^{\beta}} \right]$$
$$-1 - (\Lambda(\beta) + 2) \sum_{k=1}^{N} \frac{1}{\Lambda(\beta) + X_k^{\beta}}$$
$$= \frac{\Lambda(\beta) + 2}{\Lambda(\beta) + 1} \left[1 + (\Lambda(\beta) + 1) \sum_{k=1}^{N} \frac{1}{\Lambda(\beta) + X_k^{\beta}} \right]$$

$$-1 - (\Lambda(\beta) + 2) \sum_{k=1}^{N} \frac{1}{\Lambda(\beta) + X_k^{\beta}}$$
$$= \frac{\Lambda(\beta) + 2}{\Lambda(\beta) + 1} - 1 = \frac{1}{\Lambda(\beta) + 1}$$

which is a contradiction because $\Lambda(\beta) > 0$ and thus the first claim is proven. Hence, for $\beta \in (\beta_C, +\infty)$ fixed and any critical point $\Lambda(\beta)$ of $\varphi(\cdot, \beta)$ we have always

$$\frac{\partial^2}{\partial \lambda^2} \varphi(\lambda, \beta) \bigg|_{\Lambda(\beta)} \neq 0.$$

We know that for a maximum the second derivative is negative, so we always have

$$\frac{\partial^2}{\partial \lambda^2} \varphi(\lambda,\beta) \bigg|_{\Lambda(\beta)} < 0$$

from which we conclude that for all $\beta > \beta_C$

$$\left[\frac{2}{N}\sum_{i=1}^{N}\frac{X_{i}^{\beta}}{(\lambda+X_{i}^{\beta})^{2}}-\frac{1}{(\lambda+1)^{2}}\right]_{\Lambda(\beta)}>0$$
(47)

For the second claim we note that by its definition as root of multiplicity one the function $\Lambda(\cdot)$ is analytic for arguments $\beta > \beta_C$ by simple application of the implicit function theorem (e.g. Ostrowski (1960)) and the proof is finished.

Now we have all the ingredients for the proof of Theorem 2.

Proof of Theorem 2 Lemma 4.(3) already contains the main statements concerning the unique construction of the function $\Lambda(\cdot) : \mathbb{R}_0^+ \to \mathbb{R}_0^+$. The behaviour of the derivatives $\frac{\partial}{\partial \lambda}\varphi(\lambda,\beta)$ for $\lambda \neq \Lambda(\beta)$ follows from the easily obtained asymptotic for $\lambda \to \infty$

$$\frac{\partial}{\partial \lambda} \varphi(\lambda, \beta) = -\frac{N}{\lambda} \left\{ 1 + O\left(\frac{1}{\lambda}\right) \right\}$$

and Lemma 2 for small $\lambda \ge 0$ and the fact that $\frac{\partial}{\partial \lambda} \varphi(\lambda, \beta)$ possesses only one zero.

The next lemma treats the situation in which the log-logistic distribution degenerates when $\lambda \to 0^+$. We formulate it for the convenience of general $x_L > 0$ and work within the Banach space $L^1(\mathbb{R}^+)$ for which the integrals of their absolute values are Lebesgue-integrable.

Lemma 6 Consider the left-truncated log-logistic distribution given in Eq. (1) with $x_L > 0$. In the limit $\lambda \to 0^+$ respectively $\alpha \to 0^+$ the following holds true:

=

(1) There is L^1 -convergence in the sense of Lebesgue of

$$f_{LT}(\cdot|\alpha,\beta) \xrightarrow{L^1} g(\cdot|\beta) \ as \ \alpha \to 0^+,$$

where

$$g(x|\beta) = \left(\frac{\beta}{x_L}\right) \left(\frac{x}{x_L}\right)^{-(1+\beta)}$$
(48)

is the Pareto probability distribution density.

(2) While the first ML equation for α respectively λ , Eq. (6), becomes redundant, the second ML equation for β , Eq. (7) simplifies to

$$\frac{1}{\beta} = \frac{1}{N} \sum_{i=1}^{N} \ln\left(\frac{X_i}{x_L}\right). \tag{49}$$

Proof (1) To prove statement (1) we first show point-wise convergence for some $x > x_L$ fixed, namely if we choose any arbitrary positive sequence $(\alpha_n)_{n \in \mathbb{N}} \to 0^+$ as $n \to \infty$:

$$\begin{bmatrix} 1 + \left(\frac{x_L}{\alpha_n}\right)^{\beta} \end{bmatrix} \frac{\beta}{\alpha_n} \left(\frac{x}{\alpha_n}\right)^{\beta-1} \frac{1}{\left[1 + \left(\frac{x}{\alpha_n}\right)^{\beta}\right]^2} \\ = \left(\frac{x_L}{\alpha_n}\right)^{\beta} \frac{\beta}{\alpha_n} \left(\frac{x}{\alpha_n}\right)^{\beta-1} \left(\frac{x}{\alpha_n}\right)^{-2\beta} + \mathcal{O}(\alpha_n^{2\beta}) \\ \xrightarrow{n \to \infty} \left(\frac{\beta}{x_L}\right) \left(\frac{x}{x_L}\right)^{-\beta-1} = g(x|\beta).$$

Next we need to show that the functions $f_{LT}(\cdot | \alpha_n, \beta)$ can be bounded by a positive integrable function. For this we choose the following bound:

$$0 < f_{LT}(x|\alpha_n,\beta) < c \cdot g(x|\beta),$$

where the constant c > 0 needs to be chosen in a suitable way. To this end we start from the following

$$\left[1 + \left(\frac{x_L}{\alpha_n}\right)^{\beta}\right] \frac{\beta}{\alpha_n} \left(\frac{x}{\alpha_n}\right)^{\beta-1} \frac{1}{\left[1 + \left(\frac{x}{\alpha_n}\right)^{\beta}\right]^2} < c \left(\frac{\beta}{x_L}\right) \left(\frac{x}{x_L}\right)^{-(1+\beta)}$$

🖄 Springer

which can be reduced to

$$\frac{\alpha_n^{\beta} + x_L^{\beta}}{x_L^{\beta}} \left(\frac{x^{\beta}}{\alpha_n^{\beta} + x^{\beta}} \right)^2 < c.$$

Now there exists an $n_0 > 0$ such that $(\alpha_n^{\beta} + x_L^{\beta})/x_L^{\beta} \le 2$ for $n > n_0$. Because $x^{\beta}/(\alpha_n^{\beta} + x^{\beta}) < 1$ we may choose

$$c = \max\left\{2, \max_{0 \le n \le n_0} \left(1 + \frac{\alpha_n^\beta}{x_L^\beta}\right)\right\}.$$

The desired convergence result in the L^1 -Lebesgue sense follows from Lebesgue's dominated convergence theorem.

(2) To prove statement (2) note that in Eq. (7) as $\alpha_n \to 0^+$ the leading term on the right-hand side is $1/\beta + \ln (x_L/\alpha_n)$ whereas the leading term on the left-hand side is $\sum_{i=1}^{N} \ln (X_i/\alpha_n) \cdot 1$ and rearranging terms and taking the limit yields the desired result.

4.3 Proof of Corollary 1

Proof Part (1) is an immediate consequence of our theorems. The proof of statement (2) is provided by the Lemma 6. \Box

4.4 Proof of Theorem 3

Proof (1): The expressions follow immediately from Theorem 1. Note that $X_1 > 1$ throughout.

(2): Because $\lambda = \Lambda(\beta) > 0$ for $\beta > \beta_C$ we can write Eq. (6) as

$$0 = 2\frac{X_1^{\beta}}{\lambda + X_1^{\beta}} - 1 - \frac{1}{\lambda + 1}.$$

Now inserting $\lambda = \Lambda(\beta) = X_1^{\beta} - 2$ on the right-hand side yields

$$2\frac{X_1^{\beta}}{2(X_1^{\beta}-1)} - 1 - \frac{1}{X_1^{\beta}-1} = \frac{X_1^{\beta}}{X_1^{\beta}-1} - \frac{X_1^{\beta}-1}{X_1^{\beta}-1} - \frac{1}{X_1^{\beta}-1} = 0.$$

(3): We see continuity by straight-forward computation:

$$\lim_{\beta \to \beta_C -} \tilde{\varphi}(\beta) = \lim_{\beta \to \beta_C +} \tilde{\varphi}(\beta) = N \ln \beta_C - N \ln 2.$$

Deringer

For monotonicity we differentiate the profile likelihood with respect to β . For $\beta \in (0, \beta_C)$ we obtain

$$\tilde{\varphi}'(\beta) = N\left(\frac{1}{\beta} - \ln X_1\right) > 0$$

because $\beta < \beta_C = \ln 2 / \ln X_1 < 1 / \ln X_1$. Finally for $\beta \in (\beta_C, +\infty)$ after some easy computation we obtain:

$$\tilde{\varphi}'(\beta) = N\left(\frac{1}{\beta} - \frac{\ln X_1}{X_1^{\beta} - 1}\right) = \left(\frac{1}{\beta} - \frac{1}{\beta C}\frac{\ln 2}{X_1^{\beta} - 1}\right) > 0,$$

where the last inequality follows from the inequality

$$\beta_C(X_1^\beta - 1) > \beta \ln 2.$$

To see this define two functions $F(\beta) = \beta_C (X_1^{\beta} - 1)$ and $G(\beta) = \beta \ln 2$. Obviously we have $F(\beta_C) = \beta_C > G(\beta_C) = \beta_C \ln 2$. With the derivatives $F'(\beta) = \beta_C X_1^{\beta} \ln X_1 = X_1^{\beta} \ln 2 > G'(\beta) = \ln 2$ we obtain the desired result as

$$F(\beta) - G(\beta) = F(\beta_C) - G(\beta_C) + \int_{\beta_C}^{\beta} \left(F'(b) - G'(b) \right) db > 0$$

because all terms on the right-hand side are positive.

Appendix A: Derivatives of log-likelihood function

We start from the objective function rather than the log-likelihood function, Eq. (8) where we redefined λ as $\lambda \equiv \alpha^{\beta}$ and allow the given truncation parameter $x_L \ge 0$

$$\varphi(\alpha, \beta; x_L) := N \ln\left(1 + \left(\frac{x_L}{\alpha}\right)^{\beta}\right) + N \ln\beta - N\beta \ln\alpha$$
$$+\beta \sum_{i=1}^N \ln X_i - 2\sum_{i=1}^N \ln\left(1 + \left(\frac{X_i}{\alpha}\right)^{\beta}\right).$$

The first derivatives read as

$$\frac{\partial \varphi}{\partial \alpha} = N \frac{\beta}{\alpha} \left\{ -\frac{\left(\frac{x_L}{\alpha}\right)^{\beta}}{1 + \left(\frac{x_L}{\alpha}\right)^{\beta}} - 1 + \frac{2}{N} \sum_{i=1}^{N} \frac{\left(\frac{X_i}{\alpha}\right)^{\beta}}{1 + \left(\frac{X_i}{\alpha}\right)^{\beta}} \right\},\,$$

Deringer

$$\frac{\partial \varphi}{\partial \beta} = \frac{N\left(\frac{x_L}{\alpha}\right)^{\beta} \ln \frac{x_L}{\alpha}}{1 + \left(\frac{x_L}{\alpha}\right)^{\beta}} + \frac{N}{\beta} - N \ln \alpha + \sum_{i=1}^{N} \ln X_i - 2 \sum_{i=1}^{N} \frac{\left(\frac{X_i}{\alpha}\right)^{\beta} \ln \frac{x_L}{\alpha}}{1 + \left(\frac{X_i}{\alpha}\right)^{\beta}},$$

and the second derivatives are

$$\begin{split} \frac{\partial^2 \varphi}{\partial \alpha^2} &= N \frac{\beta}{\alpha^2} \left\{ \frac{\left(\frac{x_L}{\alpha}\right)^{\beta} \left[1 + \beta + \left(\frac{x_L}{\alpha}\right)^{\beta}\right]}{\left[1 + \left(\frac{x_L}{\alpha}\right)^{\beta}\right]^2} + 1 - \frac{2}{N} \sum_{i=1}^{N} \frac{\left(\frac{x_i}{\alpha}\right)^{\beta} \left[1 + \beta + \left(\frac{x_i}{\alpha}\right)^{\beta}\right]}{\left[1 + \left(\frac{x_i}{\alpha}\right)^{\beta}\right]^2} \right\},\\ \frac{\partial^2 \varphi}{\partial \alpha \partial \beta} &= -\frac{N}{\alpha} \left\{ \frac{\left(\frac{x_L}{\alpha}\right)^{\beta} \left[1 + \left(\frac{x_L}{\alpha}\right)^{\beta} + \ln\left(\frac{x_L}{\alpha}\right)^{\beta}\right]}{\left[1 + \left(\frac{x_L}{\alpha}\right)^{\beta}\right]^2} + 1 \right.\\ &\left. - \frac{2}{N} \sum_{i=1}^{N} \frac{\left(\frac{x_i}{\alpha}\right)^{\beta} \left[1 + \left(\frac{x_i}{\alpha}\right)^{\beta} + \ln\left(\frac{x_i}{\alpha}\right)^{\beta}\right]}{\left[1 + \left(\frac{x_i}{\alpha}\right)^{\beta}\right]^2} \right\} = \frac{\partial^2 \varphi}{\partial \beta \partial \alpha},\\ &\left. \frac{\partial^2 \varphi}{\partial \beta^2} = -\frac{N}{\beta^2} \left\{ - \frac{\left(\frac{x_L}{\alpha}\right)^{\beta} \ln^2 \left(\frac{x_L}{\alpha}\right)^{\beta}}{\left[1 + \left(\frac{x_L}{\alpha}\right)^{\beta}\right]^2} + 1 + \frac{2}{N} \sum_{i=i}^{N} \frac{\left(\frac{x_i}{\alpha}\right)^{\beta} \ln^2 \left(\frac{x_i}{\alpha}\right)^{\beta}}{\left[1 + \left(\frac{x_i}{\alpha}\right)^{\beta}\right]^2} \right\}. \end{split}$$

For simplicity we introduce a new random variable $Y := (X/\alpha)^{\beta}$ and set $Y_i := (X_i/\alpha)^{\beta}$ and likewise $\eta := (x_L/\alpha)^{\beta}$. This yields the following expectations:

$$\mathbb{E}\left(\frac{\partial^2 \varphi}{\partial \alpha^2}\right) = N \frac{\beta}{\alpha^2} \left\{ \frac{\eta \left(1 + \beta + \eta\right)}{\left(1 + \eta\right)^2} + 1 - 2\mathbb{E}\left(\frac{Y \left(1 + \beta + Y\right)}{\left(1 + Y\right)^2}\right) \right\},\$$
$$\mathbb{E}\left(\frac{\partial^2 \varphi}{\partial \beta \partial \alpha}\right) = -\frac{N}{\alpha} \left\{ \frac{\eta \left(1 + \eta + \ln \eta\right)}{\left(1 + \eta\right)^2} + 1 - 2\mathbb{E}\left(\frac{Y \left(1 + Y + \ln Y\right)}{\left(1 + Y\right)^2}\right) \right\},\$$
$$\mathbb{E}\left(\frac{\partial^2 \varphi}{\partial \beta^2}\right) = -\frac{N}{\beta^2} \left\{ -\frac{\eta \ln^2 \eta}{\left(1 + \eta\right)^2} + 1 + 2\mathbb{E}\left(\frac{Y \ln^2 Y}{\left(1 + Y\right)^2}\right) \right\},\$$

where

$$\mathbb{E}(f(Y)) = \int_{y_L}^{\infty} f(y) \left[1+\eta\right] \frac{1}{\left[1+y\right]^2} dy.$$

$$\mathbb{E}\left(\frac{Y\left(1+\beta+Y\right)}{\left(1+Y\right)^2}\right) = \frac{(3+\beta)+3\left(3+\beta\right)\eta+6\eta^2}{6\left(1+\eta\right)^2}$$

$$\mathbb{E}\left(\frac{Y\left(1+Y+\ln Y\right)}{\left(1+Y\right)^2}\right)$$
(50)

$$= \frac{1}{6(1+\eta)^2} \left\{ (1+\eta) (3+5\eta) + (1+\eta)^3 \ln(1+\eta) - \eta^2 (3+\eta) \ln(\eta) \right\}$$
$$\mathbb{E} \left(\frac{Y \ln^2 Y}{(1+Y)^2} \right)$$
$$= \frac{1}{18} \left\{ -6 + (1+\eta) \pi^2 - \frac{3\eta \ln(\eta) [2(1+\eta) + \eta (3+\eta) (\ln(\eta))]}{(1+\eta)^2} + 6(1+\eta) \ln(\eta) \ln(1+\eta) + 6(1+\eta) \operatorname{PolyLog}[2,-\eta] \right\}$$
(51)

where

PolyLog[2,
$$-\eta$$
] = Li₂[$-\eta$] = $\sum_{k=1}^{\infty} \frac{(-\eta)^k}{k^n}$
= S_{1,1}($-\eta$) = $-\int_0^1 \ln(1+\eta t) \frac{dt}{t}$ (52)

The PolyLog function can also be described in terms the Neilsen generalised polylogarithm function.⁵ Making use of Eq. (51), we now give the final expressions for the Eq. (50)

$$\mathbb{E}\left(\frac{\partial^{2}\varphi}{\partial\alpha^{2}}\right) = N\frac{\beta}{\alpha^{2}}\left\{-\frac{\beta}{3(1+\eta)^{2}}\right\}$$

$$\mathbb{E}\left(\frac{\partial^{2}\varphi}{\partial\alpha\partial\beta}\right) = -\frac{N}{\alpha}\left\{\frac{\eta\left[1+\eta-\eta\left(3+\eta\right)\ln\frac{(1+\eta)}{\eta}+3\ln(\eta)\right]-(1+3\eta)\ln(1+\eta)}{3(1+\eta)^{2}}\right\}$$

$$= \mathbb{E}\left(\frac{\partial^{2}\varphi}{\partial\beta\partial\alpha}\right)$$

$$\mathbb{E}\left(\frac{\partial^{2}\varphi}{\partial\beta^{2}}\right) = -\frac{N}{3\beta^{2}}\left\{1+\frac{\pi^{2}}{3}(1+\eta)+2(1+\eta)\operatorname{PolyLog}[2,-\eta]-2\frac{\eta}{(1+\eta)}\ln(\eta)$$

$$+2\ln(\eta)\ln(1+\eta)-\frac{\left[(1+\eta)^{3}-1\right]}{(1+\eta)^{2}}\ln^{2}(\eta)\right\}$$
(53)

Appendix B: Proof of statement (2) in Lemma 3

We follow the proof of Gupta et al. (1999) and adapt their notation and define the quantities $a = -\ln \lambda$, $s_i = \ln X_i$ and $S = \sum_{i=1}^{N} \ln X_i = \sum_{i=1}^{N} s_i$. For simplicity we now write $\varphi(a, \beta)$ rather than $\varphi(\lambda, \beta)$ and use the fact that all s_i are positive because of the left-truncation equaling one. First, however, we start with two useful propositions:

⁵ PolyLog[*n*, *z*] = Li_{*n*}(*z*) = $\sum_{k=1}^{\infty} \frac{z^k}{k^n}$, the Neilsen generalised polylogarithm function $S_{n,p}(z) = \frac{(-1)^{n+p+1}}{(n-1)!p!} \int_0^1 \ln^{n-1}(t) \ln^p (1-zt) \frac{dt}{t}, S_{n-1,1}(z) = \text{Li}_n(z).$

Proposition 1 For any m > 0 and $c \in \mathbb{R}$ let $A \in \mathbb{R}$ satisfy the following inequalities

(i) $A \le mc$ (ii) $A \le -mc$

Then A satisfies the following inequality:

 $A \leq -m|c|.$

Proof We have $A \le 0$ by adding up inequalities (i) and (ii). Consider first the case when c > 0 in which case inequality (ii) is tighter than inequality (i), thus $A \le -mc = -m|c|$. Likewise, when c < 0 the inequality (i) is tighter than inequality (ii), thus $A \le mc = m(-|c|)$. The case c = 0 is trivial.

Proposition 2 For $z_i \in \mathbb{R}$ we have the following inequalities

$$\frac{1}{N}\sum_{i=1}^{N}z_i \le \ln\left(1 + \exp\left[\frac{1}{N}\sum_{i=1}^{N}z_i\right]\right) \le \frac{1}{N}\sum_{i=1}^{N}\ln\left(1 + e^{z_i}\right).$$

Proof Note that the function $\psi(z) = \ln (1 + \exp [z])$ is convex because $\psi''(z) = e^{z}/(1 + e^{z})^{2} > 0$ for all $z \in \mathbb{R}$. The right-hand side inequality is Jensen's inequality (Hardy et al. 1952, p. 74, Eqs. (3.8.1) and (3.8.2)), whereas the left-hand side follows from monotonicity of the logarithm.

Proof of statement (2) in Lemma 3 In our new notation the objective function Eq. (9) reads as

$$\varphi(a,\beta) = N \ln(1+e^{a}) + N \ln\beta + Na + \beta S - 2\sum_{i=1}^{N} \ln[1+e^{a+\beta s_{i}}].$$
 (9)

We shall derive various inequalities from this equation. Because the last term is negative, we immediately find the following inequality

$$\varphi(a,\beta) \le N \ln\left(1+e^a\right) + N \ln\beta + Na + \beta S.$$
(54)

Using the entire inequality chain from Proposition 2 we obtain another inequality from Eq. (9)

$$\varphi(a,\beta) \le N \ln\left(1+e^a\right) + N \ln\beta - Na - \beta S.$$
(55)

Applying Proposition 1 to the inequalities Eqs. (54) and (55) we get

$$\varphi(a,\beta) \le N \ln\left(1+e^a\right) + N \ln\beta - \beta N \left|\frac{a}{\beta} + \frac{S}{N}\right|.$$
(56)

A final upper bound is obtained from Eq. (9) using Proposition 2 for any i = 1, 2, ..., N

$$\varphi(a,\beta) \leq N \ln(1+e^{a}) + N \ln\beta - \sum_{i=1}^{N} \ln\left[1+e^{a+\beta s_{i}}\right]$$
$$\leq N \ln(1+e^{a}) + N \ln\beta - \beta\left(\frac{a}{\beta}+s_{i}\right).$$
(57)

Case I: First consider $a \ge 0$. We can bound the positive quantity $\ln (1 + e^a)$ from above by $\ln 2 + a$ and thus inequality Eq. (55) altogether yields

$$\varphi(a,\beta) \le N \ln 2 + N \ln \beta - \beta S,\tag{58}$$

the desired upper bound independent of a.

Case II: Consider now a < 0 We can bound the positive quantity $\ln (1 + e^a)$ from above by $\ln 2$. In this part we shall use inequality Eq. (56) reading as

$$\varphi(a,\beta) \le N \ln 2 + N \ln \beta - \beta N \left| \frac{a}{\beta} + \frac{S}{N} \right|$$
(59)

and inequality Eq. (57) reading as

$$\varphi(a,\beta) \le N \ln 2 + N \ln \beta - \beta \left(\frac{a}{\beta} + s_i\right).$$
 (60)

Because not all s_i are equal, there exists an index i_0 with

$$s_{i_0} - \frac{S}{N} = \varepsilon_0 > 0.$$

We have $(-\infty, s_{i_0} - \varepsilon_0) \cup (S/N + \varepsilon_0, +\infty) = \mathbb{R}$. Again we consider two subcases: If we have $-a/\beta \in (S/N + \varepsilon_0, +\infty)$ then

$$-\frac{a}{\beta} > \frac{S}{N} + \varepsilon_0 \tag{61}$$

from which we obtain

$$\left|\frac{a}{\beta} + \frac{S}{N}\right| = -\left(\frac{a}{\beta} + \frac{S}{N}\right) > \varepsilon_0.$$
(62)

Now we use this lower bound Eq. (62) in conjunction with inequality Eq. (59) to obtain

$$\varphi(a,\beta) \le N \ln 2 + N \ln \beta - \beta N \varepsilon_0. \tag{63}$$

🖄 Springer

Likewise, if we have $-a/\beta \in (-\infty, s_{i_0} - \varepsilon_0)$, then

$$-\frac{a}{\beta} < s_{i_0} - \varepsilon_0 \tag{64}$$

from which we obtain

$$\varepsilon_0 < s_{i_0} + \frac{a}{\beta}.\tag{65}$$

Now we use this lower bound Eq. (65) in conjunction with inequality Eq. (60)

$$\varphi(a,\beta) \le N \ln 2 + N \ln \beta - \beta \varepsilon_0. \tag{66}$$

Hence for any a < 0 we obtain from Eqs. (63) and (66) the desired bound independent of a

$$\varphi(a,\beta) \le N \ln 2 + N \ln \beta - \beta \varepsilon_0 \tag{67}$$

since $N \ge 2$.

From Case I Eq. (58) and Case II Eq. (67) we obtain altogether

$$\varphi(a,\beta) \le N \ln 2 + N \ln \beta - \beta \min\{S,\varepsilon_0\}$$
(68)

and thus the existence of a sufficiently large $\beta_2 > \beta_1 > 0$, independent of *a* such that $\varphi(a, \beta) < -M$ for $\beta > \beta_2$ and all $a \in \mathbb{R}$. This concludes the proof.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

- Ahmad MI, Sinclair CD, Werritty A (1988) Log-logistic flood frequency analysis. J Hydrol 98(3–4):205– 224. https://doi.org/10.1016/0022-1694(88)90015-7
- Al-Shomrani AA, Shawky AI, Arif OH, Aslam M (2016) Log-logistic distribution for survival data analysis using MCMC. SpringerPlus 5(1):1774–1790. https://doi.org/10.1186/s40064-016-3476-7
- Antle CE, Klimko L, Harkness W (1970) Confidence intervals for the parameters of the logistic distribution. Biometrika 57(2):397–402. https://doi.org/10.2307/2334848

Ashkar F, Mahdi S (2006) Fitting the log-logistic distribution by generalized moments. J Hydrol 328(3– 4):694–703. https://doi.org/10.1016/j.jhydrol.2006.01.014 Bickel PJ, Doksum KA (2015) Mathematical statistics-basic ideas and selected topics. Chapman and Hall/CRC, New York. https://doi.org/10.1201/9781315369266

 $Cohen\,C\,(1991)\,Truncated\,and\,censored\,samples \\ -- theory\,and\,applications,\,1st\,edn.\,CRC\,Press,\,Boca\,Raton$

del Castillo J (1994) The singly truncated normal distribution: A non-steep exponential family. Ann Inst Stat Math 45(1):57–66. https://doi.org/10.1007/BF00773592

- Deutscher Wetterdienst (2022) Climate data base of the DWD Deutscher Wetterdienst (German Meteorology Office), Offenbach Germany. http://opendata.dwd.de/climate_environment/CDC/regional_averages_ DE/annual/precipitation/regional_averages_rr_year.txt
- Fisk PR (1961) The graduation of income distributions. Econometrica 29(2):171–185. https://doi.org/10. 2307/1909287
- Gupta RC, Akman O, Levin S (1999) A study of log-logistic model in survival analysis. Biom J 41(4):431– 443. https://doi.org/10.1002/(SICI)1521-4036(199907)41:4<431::AID-BIMJ431>3.0.CO;2-U
- Guscott JC (2018) Reliable statistical methods and their applications for testing incomplete multidisciplinary data. MPhil Thesis, Physics Department, University of Adelaide. https://hdl.handle.net/2440/127169
- Hardy G, Littlewood JE, Polya G (1952) Inequalities, 2nd edn. Cambridge University Press, Cambridge (Reprinted 1994)
- Harris G, Martin C (1987) Shorter notes: the roots of a polynomial vary continuously as a function of the coefficients. Proc Am Math Soc 100(2):390–392. https://doi.org/10.2307/2045978
- He X, Chen W, Qian W (2020) Maximum likelihood estimators of the parameters of the log-logistic distribution. Stat Pap 61(5):1875–1892. https://doi.org/10.1007/s00362-018-1011-3
- Jameson GJO (2006) Counting zeros of generalized polynomials: Descartes' rule of signs and Laguerre's extensions. Math Gaz 90(581):223–234. https://doi.org/10.1017/S0025557200179628
- Kendall M, Stuart A (1979) The advanced theory of statistics II—inference and relationship, 4th revised edn. Griffin, London
- Kizilersu A, Kreer M, Thomas AW (2016) Goodness-of-fit testing for left-truncated two-parameter Weibull distributions with known truncation point. Austrian J Stat 45(3):15–42. https://doi.org/10.17713/ajs. v45i3.106
- Kreer M, Kizilersu A, Thomas AW, dos Reis AE (2015) Goodness-of-fit tests and applications for lefttruncated Weibull distributions to non-life insurance. Eur Actuar J 5(1):139–163. https://doi.org/10. 1007/s13385-015-0105-8
- Lee ET, Wang JW (2003) Statistical methods for survival data analysis, 3rd edn. Wiley, Hoboken
- Newey WK, McFadden D (1994) Large sample estimation and hypothesis testing. In: Handbook of econometrics, vol 4. Elsevier, Amsterdam, pp 2111–2245. https://doi.org/10.1016/S1573-4412(05)80005-4
- Ostrowski AM (1960) Solution of equations and systems of equations. Academic Press, New York
- Pratt JW (1960) On interchanging limits and integrals. Ann Math Stat 31(1):74–77. https://doi.org/10.1214/ aoms/1177705988
- Reath J, Dong J, Wang M (2018) Improved parameter estimation of the log-logistic distribution with applications. Comput Stat 33(1):339–356. https://doi.org/10.1007/s00180-017-0738-y
- Schnepper T, Groh J, Gerke HH, Reichert B, Pütz T (2023) Evaluation of precipitation measurement methods using data from a precision lysimeter network. Hydrol Earth Syst Sci 27:3265–3292. https://doi.org/ 10.5194/hess-27-3265-2023
- Shao Q (2004) Notes on maximum likelihood estimation for the three-parameter Burr XII distribution. Comput Stat Data Anal 45(3):675–687. https://doi.org/10.1016/S0167-9473(02)00367-5
- Shoukri MM, Mian IUH, Tracy DS (1988) Sampling properties of estimators of the log-logistic distribution with application to Canadian precipitation data. Can J Stat 16(3):223–236. https://doi.org/10.2307/ 3314729
- Toronto Weather Statistics (2022) Total precipitation—annual data Jul 1–Jun 30 for Toronto. http://toronto. weatherstats.ca/charts/precipitation-wyearly.html
- Wingo DR (1989) The left-truncated Weibull distribution: theory and computation. Stat Pap 30:39–48. https://doi.org/10.1007/BF02924307

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