

# Logistic or not Logistic?

James S. Allison<sup>1</sup>  | Bruno Ebner<sup>2</sup> | Marius Smuts<sup>1</sup>

<sup>1</sup>School of Mathematical and Statistical Sciences, North-West University, 1 Hoffman Street, Potchefstroom, North West, South Africa

<sup>2</sup>Institute of Stochastics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

## Correspondence

James S. Allison, School of Mathematical and Statistical Sciences, North-West University, 1 Hoffman Street, Potchefstroom, North West, South Africa.  
Email: [james.allison@nwu.ac.za](mailto:james.allison@nwu.ac.za)

We propose a new class of goodness-of-fit tests for the logistic distribution based on a characterization related to the density approach in the context of Stein's method. This characterization-based test is a first of its kind for the logistic distribution. The asymptotic null distribution of the test statistic is derived and it is shown that the test is consistent against fixed alternatives. The finite sample power performance of the newly proposed class of tests is compared to various existing tests by means of a Monte Carlo study. It is found that this new class of tests are especially powerful when the alternative distributions are heavy tailed, like Student's  $t$  and Cauchy, or for skew alternatives such as the log-normal, gamma and chi-square distributions.

## KEYWORDS

density approach, empirical characteristic function, goodness-of-fit, logistic distribution, Stein's method

## 1 | INTRODUCTION

The logistic distribution apparently found its origin in the mid-19th century in the writings of (Verhulst, 1838; Verhulst, 1845). Since then it has been used in many different areas such as logistic regression, logit models and neural networks. The logistic law has become a popular choice of model in reliability theory and survival analysis (see e.g., Kannisto, 1999) and lately in finance (Ahmad, 2018). The United States Chess Federation and FIDE have recently changed its formula for calculating chess ratings of players by using the more heavy tailed logistic distribution instead of the lighter tailed normal distribution (Aldous, 2017; Elo, 1978). For a detail account on the his-

-----  
This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Statistica Neerlandica* published by John Wiley & Sons Ltd on behalf of Netherlands Society for Statistics and Operations Research.

tory and application of the logistic distribution, the interested reader is referred to Johnson, Kotz, and Balakrishnan (1995).

In the literature some goodness-of-fit tests for assessing whether the observed data are realizations from the logistic distribution have been developed and studied. These include tests based on the empirical distribution function (Stephens, 1979), normalized spacings (Lockhart, O'Reilly, & Stephens, 1986), chi-squared-type statistics (Aguirre & Nikulin, 1994), orthogonal expansions (Cuadras & Lahlou, 2000), empirical characteristic and moment generating functions (Meintanis, 2004; Epps, 2005), and the Gini index (Alizadeh Noughabi, 2017).

Balakrishnan (1991) provides an excellent discussion on the logistic distribution including some of the goodness-of-fit tests mentioned above. Nikitin and Ragozin (2019) recently proposed a test based on a characterisation of the logistic distribution involving independent shifts. In this paper the authors remarked that “no goodness-of-fit tests of the composite hypothesis to the logistic family based on characterizations are yet known.” Although, as mentioned, some tests exist for the logistic distribution, they are few in number compared to those for other distributions such as the normal, exponential and the Rayleigh distribution. In this paper we propose a new class of tests for the logistic distribution based on a new characterisation filling the gap reported by Nikitin and Ragozin (2019). To be precise, we write shorthand  $L(\mu, \sigma)$ ,  $\mu \in \mathbb{R}$ ,  $\sigma > 0$ , for the logistic distribution with location parameter  $\mu$  and scale parameter  $\sigma$  if the density is defined by

$$f(x, \mu, \sigma) = \frac{1}{\sigma} \frac{\exp\left(-\frac{x-\mu}{\sigma}\right)}{\left(1 + \exp\left(-\frac{x-\mu}{\sigma}\right)\right)^2} = \frac{1}{4\sigma} \left(\operatorname{sech}\left(-\frac{x-\mu}{2\sigma}\right)\right)^2, x \in \mathbb{R}, \quad (1)$$

where  $\operatorname{sech}(\cdot) = (\cosh(\cdot))^{-1}$  is the hyperbolic secant. Note that  $X \sim L(\mu, \sigma)$  if, and only if,  $\frac{X-\mu}{\sigma} \sim L(0, 1)$  and hence the logistic distribution belongs to the location-scale family of distributions, for a detailed discussion see Johnson et al. (1995), chapter 23. In the following we denote the family of logistic distributions by  $\mathcal{L} := \{L(\mu, \sigma) : \mu \in \mathbb{R}, \sigma > 0\}$ , a family of distributions which is closed under translation and rescaling. Let  $X, X_1, X_2, \dots$  be real-valued independent and identically distributed (iid.) random variables with distribution  $\mathbb{P}^X$  defined on an underlying probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . We test the composite hypothesis

$$H_0 : \mathbb{P}^X \in \mathcal{L}, \quad (2)$$

against general alternatives based on the sample  $X_1, \dots, X_n$ .

The novel procedure is based on the following new characterization of the standard Logistic distribution, which is related to the density method in the broad theory of Stein's method for distributional approximation, see for example, Chen, Goldstein, and Shao (2011); Ley and Swan (2013), and the Stein–Tikhomirov approach, see Arras, Mijoule, Poly, and Swan (2017).

**Theorem 1.** *Let  $X$  be a random variable with absolutely continuous density  $p$ . Then  $X$  follows a standard logistic distribution  $L(0, 1)$  if, and only if*

$$\mathbb{E} \left[ f_t'(X) - \frac{1 - \exp(-X)}{1 + \exp(-X)} f_t(X) \right] = 0, \quad (3)$$

*holds for all  $t \in \mathbb{R}$ , where  $f_t(x) = \exp(itx)$  and  $i$  is the imaginary unit.*

*Proof.* For  $X \sim L(0, 1)$  direct calculation shows the assertion. Let  $X$  be a random variable with absolutely continuous density function  $p$  such that,

$$\mathbb{E} \left[ \left( it - \frac{1 - \exp(-X)}{1 + \exp(-X)} \right) \exp(itX) \right] = 0,$$

holds for all  $t \in \mathbb{R}$ . Note that since  $-it\mathbb{E}[\exp(itX)]$  is the Fourier–Stieltjes transform of the derivative of  $p$  we have

$$0 = \mathbb{E} \left[ \left( it - \frac{1 - \exp(-X)}{1 + \exp(-X)} \right) \exp(itX) \right] = \int_{-\infty}^{\infty} \left( -p'(x) - \frac{1 - \exp(-X)}{1 + \exp(-X)} p(x) \right) \exp(itx) \, dx,$$

for all  $t \in \mathbb{R}$ . By standard properties of the Fourier–Stieltjes transform, we hence note that  $p$  must satisfy the ordinary differential equation

$$p'(x) + \frac{1 - \exp(-x)}{1 + \exp(-x)} p(x) = 0,$$

for all  $x \in \mathbb{R}$ . By separation of variables it is straightforward to see, that the only solution satisfying  $\int_{-\infty}^{\infty} p(x) dx = 1$  is  $p(x) = f(x, 0, 1)$ ,  $x \in \mathbb{R}$ , and  $X \sim L(0, 1)$  follows. ■

To model the standardization assumption, we consider the so-called scaled residuals  $Y_{n,1}, \dots, Y_{n,n}$ , given by

$$Y_{n,j} = \frac{X_j - \hat{\mu}_n}{\hat{\sigma}_n}, \quad j = 1, \dots, n.$$

Here,  $\hat{\mu}_n = \hat{\mu}_n(X_1, \dots, X_n)$  and  $\hat{\sigma}_n = \hat{\sigma}_n(X_1, \dots, X_n)$  denote consistent estimators of  $\mu \in \mathbb{R}$  and  $\sigma > 0$  such that

$$\hat{\mu}_n(bX_1 + c, \dots, bX_n + c) = b\hat{\mu}_n(X_1, \dots, X_n) + c, \tag{4}$$

$$\hat{\sigma}_n(bX_1 + c, \dots, bX_n + c) = b\hat{\sigma}_n(X_1, \dots, X_n), \tag{5}$$

holds for each  $b > 0$  and  $c \in \mathbb{R}$ . By (4) and (5) it is easy to see that  $Y_{n,j}, j = 1, \dots, n$ , do not depend on the location nor the scale parameter, so we assume  $\mu = 0$  and  $\sigma = 1$  in the following. The test statistic

$$T_n = n \int_{-\infty}^{\infty} \left| \frac{1}{n} \sum_{j=1}^n \left( it - \frac{1 - \exp(-Y_{n,j})}{1 + \exp(-Y_{n,j})} \right) \exp(itY_{n,j}) \right|^2 \omega(t) \, dt,$$

is the weighted  $L^2$ -distance from (3) to the 0-function. Here,  $\omega(\cdot)$  denotes a symmetric, positive weight function satisfying  $\int_{-\infty}^{\infty} \omega(t) \, dt < \infty$ , that guaranties that the considered integrals are finite. Since under the hypothesis (2)  $T_n$  should be close to 0, we reject  $H_0$  for large values of  $T_n$ .

Note that  $T_n$  is in the structural spirit of section 5.4.2 in Anastasiou et al. (2023). It only depends on the scaled residuals  $Y_{n,j}, j = 1, \dots, n$ , and as a consequence it is invariant due to affine transformations of the data, that is, w.r.t. transformations of the form  $x \mapsto bx + c$ ,  $b > 0, c \in \mathbb{R}$ . This is indeed a desirable property, since the family  $\mathcal{L}$  is closed under affine transformations.

Direct calculations show with  $\omega(t) = \omega_a(t) = \exp(-at^2)$ ,  $t \in \mathbb{R}$ ,  $a > 0$ , that the integration-free and numerically stable version is

$$T_{n,a} = \frac{1}{4a^2 n} \sqrt{\frac{\pi}{a}} \sum_{j,k=1}^n \frac{\exp(-(Y_{n,j,k}^+)^2/4a)}{(\exp(Y_{n,j}) + 1)(\exp(Y_{n,k}) + 1)} \left[ (4a^2 + 2a + (Y_{n,j,k}^-)^2) \exp(Y_{n,j,k}^+ + Y_{n,j,k}/a) \right. \\ \left. - \exp(Y_{n,j,k}/a) \left( (4a^2 + 2a(2Y_{n,j,k}^- - 1) + (Y_{n,j,k}^-)^2) \exp(Y_{n,j}) \right. \right. \\ \left. \left. + (4a^2 - 2a(2Y_{n,j,k}^- + 1) + (Y_{n,j,k}^-)^2) \exp(Y_{n,k}) \right) \right. \\ \left. + (4a^2 + 2a - (Y_{n,j,k}^-)^2) \exp(Y_{n,j,k}/a) \right],$$

with  $Y_{n,j,k}^+ = Y_{n,j} + Y_{n,k}$ ,  $Y_{n,j,k}^- = Y_{n,j} - Y_{n,k}$ , and  $Y_{n,j,k} = Y_{n,j} \cdot Y_{n,k}$  for  $j, k = 1, \dots, n$ . Here,  $a$  is a so-called *tuning parameter*, which allows some flexibility in the choice of the right test statistic  $T_{n,a}$ . A good choice of  $a$  is suggested in Section 3.

The rest of the paper is organised as follows. In Section 2 the asymptotic behavior of the new test is investigated under the null and alternative distribution, respectively. The results of a Monte Carlo study is presented in Section 3, while all the tests are applied to a real-world dataset in Section 4. The paper concludes in Section 5 with some concluding remarks and an outlook for future research.

## 2 | LIMIT DISTRIBUTION UNDER THE NULL HYPOTHESIS AND CONSISTENCY

In what follows let  $X_1, X_2, \dots$  be iid. random variables, and in view of affine invariance of  $T_n$  we assume w.l.o.g.  $X_1 \sim L(0, 1)$ . A suitable setup for deriving asymptotic theory is the Hilbert space  $\mathbb{H}$  of measurable, square integrable functions  $\mathbb{H} = L^2(\mathbb{R}, \mathcal{B}, \omega(t)dt)$ , where  $\mathcal{B}$  is the Borel- $\sigma$ -field of  $\mathbb{R}$ . Notice that the functions figuring within the integral in the definition of  $T_n$  are  $(\mathcal{A} \otimes \mathcal{B})$ -measurable random elements of  $\mathbb{H}$ . We denote by

$$\|f\|_{\mathbb{H}} = \left( \int_{-\infty}^{\infty} |f(t)|^2 \omega(t) dt \right)^{1/2}, \quad \langle f, g \rangle_{\mathbb{H}} = \int_{-\infty}^{\infty} f(t)g(t) \omega(t) dt,$$

the usual norm and inner product in  $\mathbb{H}$ . In the following, we assume that the estimators  $\hat{\mu}_n$  and  $\hat{\sigma}_n$  allow linear representations

$$\sqrt{n}\hat{\mu}_n = \frac{1}{\sqrt{n}} \sum_{j=1}^n \psi_1(X_j) + o_{\mathbb{P}}(1), \\ \sqrt{n}(\hat{\sigma}_n - 1) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \psi_2(X_j) + o_{\mathbb{P}}(1),$$

where  $o_{\mathbb{P}}(1)$  denotes a term that converges to 0 in probability, and  $\psi_1$  and  $\psi_2$  are measurable functions with

$$\mathbb{E}[\psi_1(X_1)] = \mathbb{E}[\psi_2(X_1)] = 0, \quad \text{and} \quad \mathbb{E}[\psi_1^2(X_1)] < \infty, \mathbb{E}[\psi_2^2(X_1)] < \infty.$$

The interested reader finds formulas for the functions  $\psi_1$  and  $\psi_2$  in Appendix A for maximum-likelihood and moment estimators. By the symmetry of the weight function  $\omega(\cdot)$  straightforward calculations show

$$T_n = \int_{-\infty}^{\infty} Z_n^2(t) \omega(t) dt,$$

where

$$Z_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \kappa(t, Y_{n,j}), \quad t \in \mathbb{R},$$

and

$$\begin{aligned} \kappa(t, x) = & (1 + \exp(-x))^{-1} [((1 - t) \cos(xt) - (t + 1) \sin(xt)) \exp(-x) \\ & - (t + 1) \cos(xt) - (t - 1) \sin(xt)], \quad t, x \in \mathbb{R}. \end{aligned}$$

Clearly,  $Z_n(t)$  is a sum of dependent random variables. In order to find an asymptotic equivalent stochastic process we use a first-order multivariate Taylor expansion and consider with

$$\begin{aligned} h(t, x) = & (1 + \exp(-x))^{-2} [((t + 1) \cos(xt) - (t - 1) \sin(xt)) t \exp(-2x) \\ & + 2(t^2 + 1)(\cos(xt) - \sin(xt)) \exp(-x) + ((t - 1) \cos(xt) - (t + 1) \sin(xt)) t], \quad t, x \in \mathbb{R}, \end{aligned}$$

the auxiliary process

$$Z_n^*(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \kappa(t, X_j) + \hat{\mu}_n h(t, X_j) + (\hat{\sigma}_n - 1) X_j h(t, X_j), \quad t \in \mathbb{R}.$$

In view of (6) and (6) we define the second auxiliary process

$$Z_n^{**}(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \kappa(t, X_j) + \psi_1(X_j) \mathbb{E}[h(t, X_1)] + \psi_2(X_j) \mathbb{E}[X_1 h(t, X_1)], \quad t \in \mathbb{R},$$

which is a sum of centered iid. random variables. Note that using

$$\begin{aligned} \mathbb{E}(\exp(-2X_1)/(1 + \exp(-X_1))^2) &= 1/3, \quad \mathbb{E}(|X_1| \exp(-X_1)/(1 + \exp(-X_1))^2) = \log(2)/3 - 1/12, \\ \mathbb{E}(\exp(-X_1)/(1 + \exp(-X_1))^2) &= 1/6, \text{ and} \\ \mathbb{E}(|X_1| \exp(-2X_1)/(1 + \exp(-X_1))^2) &= 2 \log(2)/3 + 1/12, \end{aligned}$$

we have by straightforward calculations  $\mathbb{E}[|h(t, X_1)|] < \infty$  and  $\mathbb{E}[|X_1 h(t, X_1)|] < \infty$ . In the following, we denote by  $\xrightarrow{D}$  weak convergence (or alternatively convergence in distribution), whenever random elements (or random variables) are considered, and in the same manner by  $\xrightarrow{P}$  convergence in probability.

**Theorem 2.** *Under the standing assumptions, we have*

$$Z_n \xrightarrow{D} Z, \quad \text{as } n \rightarrow \infty,$$

in  $\mathbb{H}$ , where  $Z$  is a centered Gaussian process having covariance kernel

$$\begin{aligned} K(s, t) = & \mathbb{E}[\kappa(s, X_1)\kappa(t, X_1)] + \mathbb{E}[h(s, X_1)]\mathbb{E}[\psi_1(X_1)\kappa(t, X_1)] + \mathbb{E}[h(t, X_1)]\mathbb{E}[\psi_1(X_1)\kappa(s, X_1)] \\ & + \mathbb{E}[X_1 h(s, X_1)]\mathbb{E}[\psi_2(X_1)\kappa(t, X_1)] + \mathbb{E}[X_1 h(t, X_1)]\mathbb{E}[\psi_2(X_1)\kappa(s, X_1)] \\ & + \mathbb{E}[\psi_1^2(X_1)]\mathbb{E}[h(s, X_1)]\mathbb{E}[h(t, X_1)] + \mathbb{E}[\psi_2^2(X_1)]\mathbb{E}[X_1 h(s, X_1)]\mathbb{E}[X_1 h(t, X_1)] \\ & + \mathbb{E}[\psi_1(X_1)\psi_2(X_1)](\mathbb{E}[h(s, X_1)]\mathbb{E}[X_1 h(t, X_1)] + \mathbb{E}[X_1 h(s, X_1)]\mathbb{E}[h(t, X_1)]), \quad s, t \in \mathbb{R}. \end{aligned}$$

Furthermore, we have  $T_n \xrightarrow{D} \|Z\|_{\mathbb{H}}^2$ , as  $n \rightarrow \infty$ .

*Proof.* In a first step, we note that after some algebra using a multivariate Taylor expansion around  $(\mu, \sigma) = (0, 1)$  we have

$$\|Z_n - Z_n^*\|_{\mathbb{H}} \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \rightarrow \infty.$$

Furthermore using the linear representations in (6) and (6) and the law of large numbers in Hilbert spaces, it follows that

$$\|Z_n^* - Z_n^{**}\|_{\mathbb{H}} \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \rightarrow \infty,$$

and by the triangular inequality, we see that  $Z_n$  has the same limiting distribution as  $Z_n^{**}$ . Since by the central limit theorem in Hilbert spaces  $Z_n^{**} \xrightarrow{D} Z$  in  $\mathbb{H}$ , where  $Z$  is the stated Gaussian limit process with covariance kernel  $K(s, t) = \mathbb{E}[Z_1^{**}(s)Z_1^{**}(t)]$  which gives the stated formula after a short calculation. The next statement is a direct consequence of the continuous mapping theorem. ■

In the rest of this section, we assume that the underlying distribution is a fixed alternative to  $H_0$  and that the distribution is absolutely continuous, as well as in view of affine invariance of the test statistic, we assume  $\mathbb{E}[X] = 0$  and  $\mathbb{E}[X^2] < \infty$ . Furthermore, we assume that

$$(\hat{\mu}_n, \hat{\sigma}_n) \xrightarrow{\mathbb{P}} (0, 1), \quad \text{as } n \rightarrow \infty. \quad (6)$$

**Theorem 3.** *Under the standing assumptions, we have as  $n \rightarrow \infty$ ,*

$$\frac{T_n}{n} \xrightarrow{\mathbb{P}} \int_{-\infty}^{\infty} \left| \mathbb{E} \left[ \left( it - \frac{1 - \exp(-X)}{1 + \exp(-X)} \right) \exp(itX) \right] \right|^2 \omega(t) dt = \Delta. \quad (7)$$

The proof of Theorem 3 follows the lines of the proof of theorem 3.1 in Ebner, Eid, and Klar (2022) and since it does not provide further insights it is omitted. Notice that by the characterization of the logistic law in Theorem 1, we have  $\Delta = 0$  if and only if  $X \sim L(0, 1)$ . This implies

that  $T_n \xrightarrow{\mathbb{P}} \infty$ , as  $n \rightarrow \infty$ , for any alternative with existing second moment. Thus we conclude that the test based on  $T_n$  is consistent against each such alternative.

Note that, if the condition in (6) does not hold, that is, the estimators converge to some other stochastic limit (or diverge), we conjecture that the testing procedure will still be consistent, since the limit on the right-hand side of (7) will be even larger.

### 3 | SIMULATION RESULTS

In this section the finite sample performance of the newly proposed test  $T_{n,a}$  is compared to various existing tests for the logistic distribution by means of a Monte Carlo study. We consider the traditional tests (based on the empirical distribution function) of Kolmogorov–Smirnov ( $KS_n$ ), Cramér–von Mises ( $CM_n$ ), Anderson–Darling ( $AD_n$ ), and Watson ( $WA_n$ ), a test proposed by Alizadeh Noughabi (2017) based on an estimate of the Gini index ( $G_n$ ), as well as a test by Meintanis (2004), based on the empirical characteristic function, with calculable form

$$R_{n,v} = \frac{4v^2\pi^2}{n} \sum_{j,k=1}^n \frac{\sinh(Y_j + Y_k)}{(Y_j + Y_k) [4v^2\pi^2 + (Y_j + Y_k)^2]} - 4\pi^2 \sum_{j=1}^n S(v, Y_j) + n \left[ \frac{2v\pi^2}{3} + 2 \sum_{k=1}^{v-1} \frac{v-k}{k^2} \right],$$

where

$$S(v, x) = \sum_{k=1}^v \frac{(2k-1) [\{x^2 + (2k-1)^2\pi^2\} \cosh(x) - 2x \sinh(x)]}{[x^2 + (2k-1)^2\pi^2]^2}.$$

We also include a new test ( $S_n$ ) constructed similarly to that of  $T_{n,a}$ , but setting  $f_t(x) = \exp(tx)$ ,  $t \in (-1, 1)$ , in theorem 1.1. This new “moment generating” function based test is then given by the  $L^2$ -statistic

$$S_n = \frac{n}{2} \int_{-1}^1 \left| \frac{1}{n} \sum_{j=1}^n \left( t - \frac{1 - \exp(-Y_{nj})}{1 + \exp(-Y_{nj})} \right) \exp(tY_{nj}) \right|^2 dt.$$

Direct calculations lead with  $Y_{nj,k}^+ = Y_{nj} + Y_{nk}$  to the numerical stable version

$$S_n = \frac{1}{n} \sum_{j,k=1}^n \frac{(e^{Y_{nj}} + e^{Y_{nk}} + e^{Y_{nj,k}^+} + 1)^{-1}}{(Y_{nj,k}^+)^3} \left( e^{2Y_{nj,k}^+} - e^{-Y_{nj,k}^+} + (1 - Y_{nj,k}^+) (e^{2Y_{nj} + Y_{nk}} + e^{Y_{nj} + 2Y_{nk}}) \right. \\ \left. + e^{Y_{nj,k}^+} (2(Y_{nj,k}^+)^2 - 2Y_{nj,k}^+ + 1) - (1 + Y_{nj,k}^+) (e^{-Y_{nj}} + e^{-Y_{nk}}) - 2(Y_{nj,k}^+)^2 - 2Y_{nj,k}^+ - 1 \right).$$

A significance level of  $\alpha = 0.05$  was used throughout the study and empirical critical values were obtained from 100,000 independent Monte Carlo replications, with unknown parameters estimated by method of moments (maximum likelihood estimation yielded similar results and are available from the authors, therefore we only display results based on method of moments). The critical values for  $T_{n,a}$  are given in Table 1 for different values of  $\alpha$ ,  $n$  and  $a$ . The power estimates were calculated for sample sizes  $n = 20$  and  $n = 50$  using 10,000 independent Monte Carlo simulations. The alternative distributions considered were the Normal (N), Student’s  $t$  ( $t$ ), Cauchy

TABLE 1 Critical values for the new test  $T_{n,\alpha}$ .

	$n = 20$			$n = 50$		
	$T_{n,3}$	$T_{n,4}$	$T_{n,5}$	$T_{n,3}$	$T_{n,4}$	$T_{n,5}$
$\alpha = 0.01$	1.011	0.701	0.525	1.091	0.759	0.580
$\alpha = 0.05$	0.684	0.459	0.339	0.714	0.487	0.363
$\alpha = 0.10$	0.531	0.350	0.254	0.555	0.374	0.276

TABLE 2 Estimated powers for alternative distributions for  $n = 20$ .

Alternative	$T_{n,3}$	$T_{n,4}$	$T_{n,5}$	$S_n$	$R_{n,1}$	$R_{n,2}$	$R_{n,3}$	$KS_n$	$CM_n$	$AD_n$	$WA_n$	$G_n$
$L(0,1)$	5	5	5	5	5	5	5	5	5	5	5	5
$N(0,1)$	3	3	2	1	6	7	9	4	4	4	4	5
$t_2$	<b>37</b>	<b>37</b>	<b>37</b>	<b>38</b>	31	27	23	33	<b>37</b>	<b>37</b>	36	33
$t_5$	<b>9</b>	<b>9</b>	<b>9</b>	<b>10</b>	<b>9</b>	8	7	7	8	8	8	8
$t_{10}$	4	4	4	4	5	<b>6</b>	<b>6</b>	5	4	5	5	5
$C(0,1)$	76	75	74	69	62	58	53	76	<b>79</b>	<b>79</b>	<b>79</b>	75
$LP(0,1)$	<b>13</b>	<b>13</b>	<b>13</b>	<b>13</b>	8	6	5	12	<b>13</b>	<b>13</b>	<b>13</b>	10
$LN(1)$	<b>87</b>	<b>87</b>	<b>87</b>	76	48	39	33	75	85	<b>87</b>	80	28
$LN(1.5)$	<b>98</b>	<b>98</b>	<b>98</b>	94	70	61	53	94	97	<b>98</b>	97	52
$LN(2)$	99	99	99	98	81	73	64	98	<b>100</b>	<b>100</b>	99	72
$\Gamma(1)$	<b>70</b>	<b>70</b>	69	52	26	20	16	53	67	<b>71</b>	61	17
$\Gamma(2)$	<b>41</b>	<b>41</b>	40	28	17	14	12	28	36	38	31	15
$\Gamma(3)$	<b>26</b>	<b>26</b>	<b>26</b>	17	12	11	11	19	23	24	20	13
$U(\sqrt{3}, -\sqrt{3})$	16	8	5	0	48	<b>55</b>	<b>58</b>	13	21	28	27	30
$B(2,2)$	5	3	2	0	19	<b>24</b>	<b>28</b>	6	9	11	12	14
$B(3,5)$	13	11	10	4	11	13	<b>14</b>	11	13	<b>14</b>	13	13
$\chi^2_2$	<b>71</b>	<b>71</b>	70	52	27	21	17	53	67	<b>71</b>	61	16
$\chi^2_5$	<b>32</b>	<b>32</b>	31	21	15	13	12	23	27	29	24	14
$\chi^2_{10}$	<b>15</b>	<b>15</b>	<b>15</b>	10	10	10	10	12	14	<b>15</b>	12	11
$\chi^2_{15}$	<b>12</b>	<b>11</b>	<b>11</b>	7	8	9	10	9	10	<b>11</b>	10	9

Note: Bold indicates the two highest powers for each alternative.

(C), Laplace (LP), log-normal (LN), gamma ( $\Gamma$ ), uniform ( $U$ ), beta ( $B$ ), and chi-square ( $\chi^2$ ) distributions. Tables 2 and 3 contain these power estimates. Apart from the mentioned alternative distributions, we also considered some local alternatives. Table 4 contains the local power estimates for  $n = 20$  (top row) and  $n = 50$  (bottom row), where we simulated data from a mixture of the logistic and Cauchy distribution, that is, we sample from a logistic distribution with probability  $1 - p$  and from a Cauchy distribution with probability  $p$ . Similarly, Table 5 contains the local power estimates where we simulated data from a mixture of the logistic and log-normal distribution. The tables contain the percentage of times that the null hypothesis in (1) is rejected, rounded to the nearest integer. All calculations were performed in R (R Core Team, 2020).



TABLE 3 Estimated powers for alternative distributions for  $n = 50$ .

Alternative	$T_{n,3}$	$T_{n,4}$	$T_{n,5}$	$S_n$	$R_{n,1}$	$R_{n,2}$	$R_{n,3}$	$KS_n$	$CM_n$	$AD_n$	$WA_n$	$G_n$
$L(0,1)$	5	5	5	5	5	5	5	5	5	5	5	5
$N(0,1)$	5	3	3	0	4	7	<b>10</b>	5	6	6	7	<b>9</b>
$t_2$	65	64	64	59	57	54	51	60	66	<b>67</b>	<b>67</b>	65
$t_5$	11	12	12	<b>15</b>	<b>14</b>	13	12	9	11	11	10	11
$t_{10}$	5	4	4	4	5	<b>6</b>	<b>6</b>	5	5	5	<b>6</b>	5
$C(0, 1)$	98	98	97	91	91	90	88	98	<b>99</b>	<b>99</b>	<b>99</b>	98
$LP(0, 1)$	19	18	18	15	13	11	9	20	<b>23</b>	22	<b>24</b>	19
$LN(1)$	<b>100</b>	<b>100</b>	<b>100</b>	96	83	76	70	99	<b>100</b>	<b>100</b>	<b>100</b>	48
$LN(1.5)$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	96	93	90	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	87
$LN(2)$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	99	98	96	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	98
$\Gamma(1)$	<b>99</b>	<b>99</b>	<b>99</b>	73	48	39	31	94	<b>99</b>	<b>99</b>	97	21
$\Gamma(2)$	<b>87</b>	<b>87</b>	<b>87</b>	39	27	22	19	67	81	86	73	28
$\Gamma(3)$	68	<b>69</b>	<b>69</b>	24	18	17	14	48	59	65	50	30
$U(\sqrt{3}, -\sqrt{3})$	78	66	51	0	93	<b>97</b>	<b>98</b>	44	68	84	77	74
$B(2,2)$	29	19	11	0	47	<b>64</b>	<b>71</b>	16	26	35	34	42
$B(3,5)$	<b>46</b>	43	40	1	12	18	21	30	39	<b>45</b>	37	40
$\chi^2_2$	<b>99</b>	<b>99</b>	<b>99</b>	75	49	40	33	95	<b>99</b>	<b>99</b>	97	20
$\chi^2_5$	78	<b>79</b>	<b>79</b>	30	22	19	16	57	70	75	61	30
$\chi^2_{10}$	<b>46</b>	<b>46</b>	<b>46</b>	13	12	13	13	31	38	41	32	29
$\chi^2_{15}$	<b>31</b>	<b>31</b>	30	8	8	10	11	22	26	28	22	24

Note: Bold indicates the two highest powers for each alternative.

The newly proposed tests are especially powerful when the alternative distributions are heavy tailed, like the Student's  $t$  and Cauchy, or for skew alternatives such as the log-normal, gamma and chi-square distributions. The test of Meintanis produces the highest estimated powers when the alternatives have lighter tails or have bounded support such as the uniform and Beta distributions. When comparing the traditional tests, it is clear that the Anderson–Darling test has superior estimated powers.

For the mixture of the logistic and Cauchy distribution, the newly proposed tests ( $T_{n,a}$  and  $S_n$ ) as well as the test of Meintanis have the highest estimated powers for small values of the mixing parameter  $p$  (i.e., closer to the null distribution). The more traditional tests have slightly higher estimated powers for increasing values of the mixing parameter. This trend is similar when considering the mixture of the logistic and log-normal distributions.

Overall, the newly proposed test  $T_{n,a}$  performs favorably relative to the existing tests and to a lesser extend the test  $S_n$ , which is based on the moment generating function. For practical implementation of the test, we advise choosing the tuning parameter as  $a = 3$  as this choice produced high estimated powers for most alternatives considered. Alternatively, one can use the methods described in Allison and Santana (2015) or Tenreiro (2019) to choose this parameter data dependently.

TABLE 4 Estimated local powers for mixture with Cauchy distribution for  $n = 20$  (top row) and  $n = 50$  (bottom row).

Mixing proportion ( $p$ )	$T_{n,3}$	$T_{n,4}$	$T_{n,5}$	$S_n$	$R_{n,1}$	$R_{n,2}$	$R_{n,3}$	$KS_n$	$CM_n$	$AD_n$	$WA_n$	$G_n$
0	5	5	5	5	5	5	5	5	5	5	5	5
	5	5	5	5	5	5	5	5	5	5	5	5
0.05	<b>12</b>	<b>12</b>	<b>12</b>	<b>12</b>	<b>12</b>	11	11	11	<b>12</b>	<b>12</b>	11	11
	20	20	20	<b>22</b>	<b>22</b>	<b>22</b>	21	18	19	19	18	18
0.1	17	<b>18</b>	<b>18</b>	<b>19</b>	17	16	15	16	17	17	16	16
	30	31	31	<b>35</b>	<b>34</b>	33	32	28	30	31	30	31
0.15	<b>24</b>	<b>24</b>	<b>24</b>	<b>25</b>	23	21	20	21	23	23	22	22
	41	42	43	<b>46</b>	<b>45</b>	44	43	38	40	41	40	43
0.2	28	<b>29</b>	<b>29</b>	<b>31</b>	28	26	24	25	27	27	26	27
	51	52	52	<b>55</b>	<b>54</b>	53	51	48	51	51	50	51
0.3	<b>38</b>	<b>38</b>	<b>38</b>	<b>39</b>	36	33	31	34	36	37	35	37
	66	66	66	<b>68</b>	<b>67</b>	65	64	62	65	66	65	65
0.4	46	<b>47</b>	<b>47</b>	<b>47</b>	42	39	36	42	45	45	44	43
	<b>77</b>	<b>77</b>	<b>77</b>	76	76	74	72	73	76	<b>77</b>	76	76
0.5	<b>55</b>	<b>55</b>	<b>55</b>	54	49	46	42	51	54	54	53	51
	84	84	84	82	82	80	78	82	84	<b>85</b>	<b>85</b>	83
0.6	<b>60</b>	<b>60</b>	<b>60</b>	58	52	48	45	57	<b>60</b>	<b>60</b>	59	57
	89	89	89	86	85	84	82	87	<b>90</b>	<b>90</b>	<b>90</b>	89
0.7	65	65	65	63	56	52	48	63	<b>66</b>	<b>66</b>	65	62
	<b>93</b>	92	92	88	87	86	84	91	<b>93</b>	<b>93</b>	<b>94</b>	92
0.8	70	69	69	65	58	54	50	68	<b>71</b>	<b>71</b>	<b>71</b>	67
	<b>96</b>	95	95	90	89	88	86	95	<b>96</b>	<b>96</b>	<b>96</b>	95
0.9	73	73	72	68	61	57	52	72	<b>76</b>	<b>75</b>	<b>75</b>	71
	97	97	96	91	91	90	88	96	<b>98</b>	<b>98</b>	<b>98</b>	97
1	77	76	75	70	63	59	54	77	<b>80</b>	<b>80</b>	<b>80</b>	75
	98	98	98	92	91	90	88	98	<b>99</b>	<b>99</b>	<b>99</b>	98

Note: Bold indicates the two highest powers for each alternative.

## 4 | PRACTICAL APPLICATION

In this section all the tests considered in the Monte Carlo study are applied to the “Bladder cancer” dataset. This dataset contains the monthly remission times of 128 patients, denoted by  $x_1, \dots, x_{128}$ , who were diagnosed with bladder cancer and can be found in Lee and Wang (2003). The data were also studied and analyzed by Noughabi (2022) and Al-Shomrani, Shawky, Arif, and Aslam (2016). We are interested in testing whether the log of the remission times,  $w_j = \log(x_j), j = 1, \dots, 128$ , follow a logistic distribution. The method of moment estimates of  $\mu$  and  $\sigma$  are  $\hat{\mu}_n = \hat{\mu}(w_1, \dots, w_{128}) = 1.753$  and  $\hat{\sigma}_n = \hat{\sigma}(w_1, \dots, w_{128}) = 0.592$ , respectively. Figure 1 represents the probability plot of  $G^{-1}\left(\frac{k}{n+1}\right)$  versus  $y_{(k)}$ , where  $G^{-1}(\cdot)$  denotes the quantile function of

TABLE 5 Estimated local powers for mixture with Log-Normal distribution for  $n = 20$  (top row) and  $n = 50$  (bottom row).

Mixing proportion ( $p$ )	$T_{n,3}$	$T_{n,4}$	$T_{n,5}$	$S_n$	$R_{n,1}$	$R_{n,2}$	$R_{n,3}$	$KS_n$	$CM_n$	$AD_n$	$WA_n$	$G_n$
0	5	5	5	5	5	5	5	5	5	5	5	5
	5	5	5	5	5	5	5	5	5	5	5	5
0.05	7	7	7	7	7	6	6	6	7	7	7	5
	8	8	8	9	9	8	8	7	7	8	7	6
0.1	8	8	8	9	8	7	6	7	8	8	7	6
	10	11	11	13	12	11	11	10	11	11	11	8
0.15	10	10	10	11	9	8	7	8	9	9	9	7
	13	14	14	16	15	14	13	13	14	14	13	10
0.2	11	12	12	12	10	9	8	11	11	11	11	8
	17	17	17	20	19	18	16	17	18	18	18	13
0.3	15	15	15	16	13	11	9	15	15	15	15	10
	23	24	24	25	24	23	21	26	27	27	28	19
0.4	19	19	19	20	14	12	10	19	20	19	19	13
	33	33	33	32	31	28	26	38	41	39	42	27
0.5	25	25	25	25	18	15	12	26	28	26	27	16
	43	42	42	39	36	33	31	52	54	52	56	36
0.6	31	31	31	29	21	17	14	32	35	34	35	20
	57	55	53	45	42	39	36	65	70	67	72	45
0.7	41	40	40	36	26	21	17	41	46	44	46	25
	71	69	67	53	49	45	42	78	83	81	84	51
0.8	51	50	49	44	30	25	20	51	57	56	56	28
	85	83	82	63	58	54	50	88	92	92	93	54
0.9	66	66	65	56	37	30	26	61	69	69	67	28
	95	94	94	76	68	63	58	96	98	98	98	53
1	87	87	87	77	48	39	33	75	85	87	80	29
	100	100	100	96	83	76	70	99	100	100	100	48

Note: Bold indicates the two highest powers for each alternative.

the standard logistic distribution and  $y_{(k)} = \frac{w_{(k)} - \hat{\mu}_n}{\hat{\sigma}_n}, k = 1, \dots, 128$ . This probability plot suggests that the underlying distribution of the data might be the logistic distribution. Table 6 contains the test statistic values as well as the corresponding estimated  $p$ -values (calculated based on 10,000 samples of size 128 simulated from the standard logistic distribution) for the eight tests for testing the goodness-of-fit for the logistic distribution. All the tests do not reject the null hypothesis that the log of the remission times is logistic distributed. These findings are in agreement with that of Al-Shomrani et al. (2016), where they concluded that the remission times follows a log-logistic distribution.

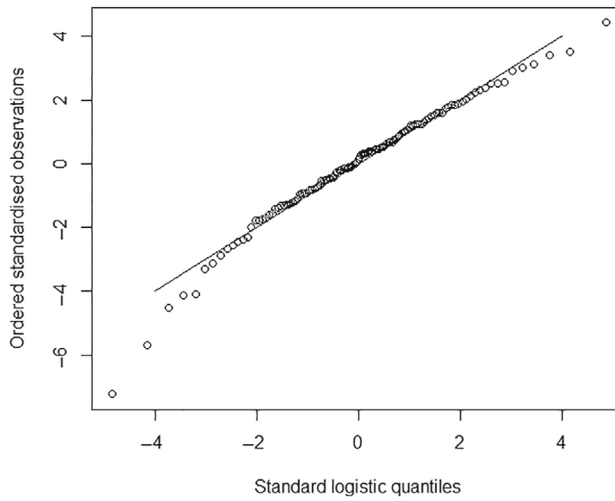


FIGURE 1 Probability plot of the cancer dataset.

TABLE 6 Estimated  $p$ -values for the cancer dataset.

Test	Test statistic value	$p$ -Value
$T_{n,3}$	0.500	0.171
$S_n$	19.75	0.329
$R_{n,1}$	169.4	0.602
$KS_n$	0.061	0.404
$CM_n$	0.072	0.401
$AD_n$	0.440	0.421
$WA_n$	0.043	0.680
$G_n$	0.325	0.147

## 5 | CONCLUSION AND OPEN QUESTIONS

We have shown a new characterisation of the logistic law and proposed a weighted affine invariant  $L^2$ -type test. Monte Carlo results show that it is competitive to the state-of-the-art procedures. Asymptotic properties have been derived, including the limit null distribution and consistency against a large class of alternatives. We conclude the paper by pointing out open questions for further research.

Following the methodology in Baringhaus, Ebner, and Henze (2017), we have under alternatives satisfying a weak moment condition  $\sqrt{n}(T_n/n - \Delta) \xrightarrow{D} N(0, \tau^2)$ , as  $n \rightarrow \infty$  where  $\tau^2 > 0$  is a specified variance, for details, see theorem 1 in Baringhaus et al. (2017). Since the calculations are too involved to get further insights, we leave the derivation of formulas open for further research. Note that such results can lead to confidence intervals for  $\Delta$  or approximations of the power function, for examples of such results see Dörr, Ebner, and Henze (2021) and Ebner, Henze, and Strieder (2022) in the multivariate normality setting. Note that Theorem 1 can also be considered

for other (so-called) test functions  $f_t$ , like, that is, the moment generating function  $f_t(x) = \exp(tx)$ . This will (again) lead to a characterization of the  $L(0, 1)$  law, and to similar but different families of test statistics, which then should be compared to the presented method.

Due to the increasing popularity of the log-logistic distribution in survival analysis, another avenue for future research is to adapt our test for scenarios where censoring is present. One possibility is to estimate the expected value in (3) by estimating the law of the survival times by the well-known Kaplan–Meier estimate. Some work on this has been done in the case of testing for exponentiality (see, e.g., Cuparić & Milošević, 2022 and Bothma, Allison, Cockeran, & Visagie, 2021).

## ACKNOWLEDGMENTS

The authors thank an anonymous referee for helpful comments that led to an improvement of the paper.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

James S. Allison  <https://orcid.org/0000-0002-9123-6793>

## REFERENCES

- Aguirre, N., & Nikulin, M. (1994). Chi-squared goodness-of-fit test for the family of logistic distributions. *Kybernetika*, 30(3), 214–222.
- Ahmad, M. (2018). Analysis of financial data by generalized logistic distribution. *International Journal of Statistics and Applied Mathematics*, 3(5), 69–72.
- Aldous, D. (2017). Elo ratings and the sports model: A neglected topic in applied probability? *Statistical Science*, 32(4), 616–629.
- Alizadeh Noughabi, H. (2017). Gini index based goodness-of-fit test for the logistic distribution. *Communications in Statistics-Theory and Methods*, 46(14), 7114–7124.
- Allison, J., & Santana, L. (2015). On a data-dependent choice of the tuning parameter appearing in certain goodness-of-fit tests. *Journal of Statistical Computation and Simulation*, 85(16), 3276–3288.
- Al-Shomrani, A. A., Shawky, A., Arif, O. H., & Aslam, M. (2016). Log-logistic distribution for survival data analysis using mcmc. *Springerplus*, 5(1), 1–16.
- Anastasiou, A., Barp, A., Briol, F.-X., Ebner, B., Gaunt, R. E., Ghaderinezhad, F., ... Swan, Y. (2023). Stein's method meets computational statistics: A review of some recent developments. *Statistical Science*, 38(1), 120–139.
- Arras, B., Mijoule, G., Poly, G., & Swan, Y. (2017). A new approach to the stein-tikhomirov method: With applications to the second wiener chaos and Dickman convergence. *arXiv preprint, arXiv:1605.06819*.
- Balakrishnan, N. (1991). *Handbook of the logistic distribution*. Boca Raton: CRC Press.
- Baringhaus, L., Ebner, B., & Henze, N. (2017). The limit distribution of weighted  $L^2$ -goodness-of-fit statistics under fixed alternatives, with applications. *Annals of the Institute of Statistical Mathematics*, 69(5), 969–995.
- Betsch, S., & Ebner, B. (2020). Testing normality via a distributional fixed point property in the stein characterization. *Test*, 29(1), 105–138.
- Bickel, P. J., & Doksum, K. A. (2015). *Mathematical statistics: Basic ideas and selected topics* (Vol. 1, 2nd ed.). Boca Raton, Fla: CRC Press.
- Bothma, E., Allison, J., Cockeran, M., & Visagie, J. (2021). Characteristic function and Laplace transform-based test for exponentiality in the presence of random right censoring. *Stat*, 10(1), e394.
- Chen, L. H. Y., Goldstein, L., & Shao, Q.-M. (2011). *Normal approximation by steins method*. Berlin, Germany: Springer.

- Cuadras, C. M., & Lahlou, Y. (2000). Some orthogonal expansions for the logistic distribution. *Communications in Statistics-Theory and Methods*, 29(12), 2643–2663.
- Cuparić, M., & Milošević, B. (2022). New characterization-based exponentiality tests for randomly censored data. *Test*, 31(2), 461–487.
- Dörr, P., Ebner, B., & Henze, N. (2021). Testing multivariate normality by zeros of the harmonic oscillator in characteristic function spaces. *Scandinavian Journal of Statistics*, 48(2), 456–501.
- Ebner, B., Eid, L., & Klar, B. (2022). Cauchy or not cauchy? New goodness-of-fit tests for the cauchy distribution. *Statistical Papers*. <https://doi.org/10.1007/s00362-022-01382-0>
- Ebner, B., Henze, N., & Strieder, D. (2022). Testing normality in any dimension by fourier methods in a multivariate stein equation. *Canadian Journal of Statistics*, 50(3), 992–1033.
- Elo, A. (1978). *The rating of chessplayers, past and present*. New York: Arco Pub. Retrieved from <https://books.google.de/books?id=8pMnAQAAAJ>
- Epps, T. W. (2005). Tests for location-scale families based on the empirical characteristic function. *Metrika*, 62(1), 99–114.
- Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous univariate distributions* (Vol. 2). John Wiley & Sons.
- Kannisto, V. (1999). *Trends in the mortality of the oldest-old*. In J. Alho (Ed.), *Statistics, registers and science* (pp. 177–194). Odense: Odense University Press.
- Lee, E. T., & Wang, J. (2003). *Statistical methods for survival data analysis* (Vol. 476). New York: John Wiley & Sons.
- Ley, C., & Swan, Y. (2013). Stein's density approach and information inequalities. *Electronic Communications in Probability*, 18, 1–14.
- Lockhart, R. A., O'Reilly, F., & Stephens, M. A. (1986). Tests of fit based on normalized spacings. *Journal of the Royal Statistical Society: Series B (Methodological)*, 48(3), 344–352.
- Meintanis, S. G. (2004). Goodness-of-fit tests for the logistic distribution based on empirical transforms. *Sankhyā: The Indian journal of. Statistics*, 66(2), 306–326.
- Millard, S. P. (2013). *Envstats: An R package for environmental statistics*. New York, NY: Springer. Retrieved from <http://www.springer.com>
- Nikitin, Y. Y., & Ragozin, I. (2019). Goodness-of-fit tests based on a characterization of logistic distribution. *Vestnik St. Petersburg University, Mathematics*, 52(2), 169–177.
- Noughabi, H. A. (2022). A new goodness-of-fit test for the logistic distribution. *Sankhyā B: The Indian journal of Statistics*, 84, 303–319.
- R Core Team. (2020). *R: A language and environment for statistical computing [Computer software manual]*. Vienna, Austria: R Core Team Retrieved from <http://www.R-project.org/>
- Stephens, M. A. (1979). Tests of fit for the logistic distribution based on the empirical distribution function. *Biometrika*, 66(3), 591–595.
- Tenreiro, C. (2019). On the automatic selection of the tuning parameter appearing in certain families of goodness-of-fit tests. *Journal of Statistical Computation and Simulation*, 89(10), 1780–1797.
- Verhulst, P.-F. (1838). Notice sur la loi que la population suit dans son accroissement. *Correspondence Mathématique et Physique*, 10, 113–126.
- Verhulst, P.-F. (1845). Recherches mathématiques sur la loi d'accroissement de la population. *Nouveaux mémoires de l'Académie Royale Des Sciences et Belles-Lettres de Bruxelles*, 18, 14–54.

**How to cite this article:** Allison, J. S., Ebner, B., & Smuts, M. (2023). Logistic or not Logistic?. *Statistica Neerlandica*, 1–15. <https://doi.org/10.1111/stan.12292>

## APPENDIX A. ASYMPTOTIC REPRESENTATION OF ESTIMATORS

In this section we derive explicit formulae for the linear representations of the estimators in (6) and (6), for comparison we refer to Meintanis (2004), p. 313.

### A.1 Maximum-likelihood estimators

The maximum-likelihood estimators  $\hat{\mu}_n^{ML}$  and  $\hat{\sigma}_n^{ML}$  of the parameters  $\mu$  and  $\sigma$  in (1) satisfy the equations, see displays (23.35) and (23.36) in Johnson et al. (1995),

$$\sum_{j=1}^n \left[ 1 + \exp\left(\frac{X_j - \hat{\mu}_n^{ML}}{\hat{\sigma}_n^{ML}}\right) \right] = \frac{n}{2},$$

$$\sum_{j=1}^n \frac{X_j - \hat{\mu}_n^{ML}}{\hat{\sigma}_n^{ML}} \left[ \frac{1 - \exp\left(\frac{X_j - \hat{\mu}_n^{ML}}{\hat{\sigma}_n^{ML}}\right)}{1 + \exp\left(\frac{X_j - \hat{\mu}_n^{ML}}{\hat{\sigma}_n^{ML}}\right)} \right] = n.$$

An implementation is found in the R-package `EnvStats`, see Millard (2013). Direct calculations show that the score vector of  $X \sim L(\mu, \sigma)$ ,  $\mu \in \mathbb{R}$ ,  $\sigma > 0$ , is

$$U_{(\mu,\sigma)}(X) = \frac{1}{\sigma^2} \tanh\left(\frac{X - \mu}{2\sigma}\right) (\sigma, (X - \mu))^T + (0, -\sigma^{-1})^T,$$

where  $x^T$  stands for the transpose of a vector  $x$ . The Fisher information matrix is

$$I_{(\mu,\sigma)} = \sigma^{-2} \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{\pi^2+3}{9} \end{pmatrix},$$

which is easily inverted due to the diagonal form. By Bickel and Doksum (2015), section 6.2.1, we hence have for  $\mu = 0$  and  $\sigma = 1$  the asymptotic expansions

$$\begin{aligned} \sqrt{n}\hat{\mu}_n^{ML} &= \frac{3}{\sqrt{n}} \sum_{j=1}^n \tanh(X_j/2) + o_{\mathbb{P}}(1) \text{ and } \sqrt{n}(\hat{\sigma}_n^{ML} - 1) \\ &= \frac{9}{(\pi^2 + 3)\sqrt{n}} \sum_{j=1}^n (X_j \tanh(X_j/2) - 1) + o_{\mathbb{P}}(1). \end{aligned}$$

### A.2 Moment estimators

Since for  $X \sim L(\mu, \sigma)$ ,  $\mu \in \mathbb{R}$ ,  $\sigma > 0$ , we have  $\mathbb{E}[X] = \mu$  and  $\mathbb{V}[X] = \pi^2\sigma^2/3$  the moment estimators  $\hat{\mu}_n^{ME}$  and  $\hat{\sigma}_n^{ME}$  are given by  $\hat{\mu}_n^{ME} = \frac{1}{n} \sum_{j=1}^n X_j = \bar{X}_n$  and  $\hat{\sigma}_n^{ME} = \frac{\sqrt{3}}{\pi} S_n$ , where  $S_n^2 = \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X}_n)^2$ . An implementation is found in the R-package `EnvStats`, see Millard (2013). By the same arguments as in Betsch and Ebner (2020), p. 113, we have

$$\sqrt{n}\hat{\mu}_n^{ME} = \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j \quad \text{and} \quad \sqrt{n}(\hat{\sigma}_n^{ME} - 1) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \frac{1}{2} \left( \frac{3}{\pi^2} X_j^2 - 1 \right) + o_{\mathbb{P}}(1).$$

Note that the unbiased moment estimators use  $\tilde{S}_n^2 = \frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X}_n)^2$  instead of  $S_n^2$ .