STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–13.
Published online in International Academic Press (www.IAPress.org)

On Testing the Adequacy of the Logistic Model Based on Negative Cumulative Extropy

Hadi Alizadeh Noughabi [1] and Mohammad Shafaei Noughabi [2,*]

[1] *Department of Statistics, University of Birjand, Birjand, Iran*
[2] *Department of Mathematics and Statistics, University of Gonabad, Gonabad, Iran*

Abstract The logistic distribution has been used for various growth models, and is used in a certain type of regression, known appropriately as logistic regression. In this article, we propose a new goodness-of-fit test for the logistic distribution based on the negative cumulative residual extropy introduced by Tahmasebi and Toomaj (2022). The mean, variance and the other properties of the test statistic is presented. Percentage points of the test statistic are obtained and then power of the test against different alternatives are reported. The results of a simulation study show the test is competitive in terms of power. The proposed statistic is easy to compute and a real data set is used to illustrate the application of the proposed test.

Keywords Logistic distribution, Negative cumulative extropy, Goodness-of-fit tests, Monte Carlo simulation, Critical value, Test power.

AMS 2010 subject classifications 62G10; 62B10

DOI: 10.19139/soic-2310-5070-1938

1. Introduction

Let X be a non-negative absolutely continuous random variable with the cumulative distribution function $F(x)$ and the reliability function $\bar{F}(x) = P(X > x)$. Rao et al. (2004) introduced the cumulative residual entropy (CRE) as follows:

$$CRE(F) = -\int_0^\infty \bar{F}(x)\,\log \bar{F}(x)\,dx\,.$$

The CRE is a measure of uncertainly in the random variable X. This measure has some advantages respect to Shannon entropy (1948). For example, it is always non-negative, can be easily computed from sample data and etc. Some results and extensions regarding CRE have been studied by many authors including Rao (2005), Asadi, and Zohrevand (2007), Di Crescenzo and Longobardi (2006, 2009a, b), Drissi et al. (2008), Navarro et al. (2010), Kapodistria and Psarrakos (2012), Kayal (2016), Psarrakos and Toomaj (2017), Navarro and Psarrakos (2017). Also, Toomaj et al. (2017) investigated the CRE of coherent and mixed systems when the component lifetimes are identically distributed.
Recently, an alternative measure of uncertainty, termed by extropy, was proposed by Lad et al. (2015). For an absolutely continuous non-negative random variable X with probability density function $f(x)$, the extropy of X is defined as

$$J(X) = -\frac{1}{2}\int_0^\infty f^2(x)\,dx. \tag{1}$$

*Correspondence to: Mohammad Shafaei Noughabi (Email: mohamad.shafaee@gmail.com). Department of Mathematics and Statistics, University of Gonabad, Gonabad, Iran.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

The properties of this measure such as the maximum extropy distribution and statistical applications were presented in Lad et al. (2015). Also, fruitful results can be found in Qiu (2017) and Qiu and Jia (2018b) related with extropy and residual extropy properties of order statistics and record values. Furthermore, Qiu et al. (2019) obtained some results on extropy properties of mixed systems. The problem of estimation of extropy has been considered by Alizadeh and Jarrahiferiz (2019).

In analogy with (1), Jahanshahi et al. (2020) proposed a cumulative residual extropy (CREX), defined as

$$J^*(X) = -\frac{1}{2}\int_0^\infty \bar{F}^2(x)\,dx\,. \tag{2}$$

It is clear that $J^*(X)$ is always non-positive. Moreover, Abdul Sathar and Dhanya Nair (2021) introduced a dynamic cumulative residual extropy for the residual random variable $X_t = (X - t|X > t)$ and studied its various properties.

Tahmasebi and Toomaj (2022) proposed an alternative measure analogous to (2) which is useful for measuring uncertainty to the past. They defined the negative cumulative extropy (NCEX) as follows and then investigated some important properties of it.

$$NCEX(X) = \frac{1}{2}\int_0^\infty \big[1 - F^2(x)\big]\,dx\,.$$

Tahmasebi and Toomaj (2022) obtained the value of NCEX for some well-known distributions, for example, NCEX of a uniform distribution on the interval $[0, 1]$ is 1/3.

Also, they applied the NCEX measure for the coherent systems lifetime with identically distributed components. Recently, Alizadeh (2023) used the negative cumulative extropy and constructed a test for uniformity. A theorem of Alizadeh (2023) states that in the class of continuous distributions f, concentrated on $[0, 1]$, it holds

$$0 \leq NCEX(f) \leq 1/2$$

and the value of $NCEX(f) = 1/3$, being uniquely attained by the $U(0, 1)$ density. Based on this property, Alizadeh (2023) proposed the following test statistic for testing uniformity.

$$T_n = \frac{1}{2}\sum_{i=1}^{n-1}\left[1 - \left(\frac{i}{n}\right)^2\right]\big(X_{(i+1)} - X_{(i)}\big).$$

He obtained the percentage points of the test statistic and power of test by simulation and showed that the test based on NCEX has a good performance respect to the other competing tests. Moreover, recently some goodness-of-fit tests for different distributions are suggested by some authors see for example Alizadeh (2021a,b), Alizadeh and Shafaei (2023, 2024).

The logistic distribution has been used for various growth models, and is used in a certain type of regression, known appropriately as logistic regression. There are many applications of the logistic distribution in literature. Early applications back to Verhulst (1838, 1845). This distribution has been used in the study of population growth (Pearl and Reed, 1920), in bioassay studies (Finney, 1947, 1952) and in the analysis of survival distributions (Plackett, 1959). Other applications of this distribution can be found in Balakrishnan (1992). Therefore, in practice, it is important to test whether the underlying distribution has a logistic form. Some researchers investigated different properties of the logistic distribution and then proposed some goodness-of-fit tests for this distribution. One can see for example, Meintanis (2004), Abd-Elfattah (2007), Al-Subh et al. (2012), Al-Shomrani et al. (2016), Alizadeh (2017), Nikitin and Ragozin (2019, 2020), Esmaeili et al. (2020), Heydari et al. (2021), Alizadeh (2022), Opone et al. (2023), and Eltehiwy (2023).

This article is organized as follows. Section 2 describes the logistic distribution and the procedure for estimating the parameters of this model. Also, a test statistic for testing a hypothesis that the sample comes from a logistic distribution based on the negative cumulative extropy is proposed. Properties and some theoretical aspects of the proposed test statistic are discussed. In Section 3 the percentage points of our test statistic are obtained for different sample sizes by a Monte Carlo experiment. Moreover, the results of the power comparison of the proposed test with some known competing tests under various alternatives are presented. The following section contains an illustrative example.

2. The Logistic Distribution and Test Statistic

In this section, we express some properties of the logistic distribution and then construct a goodness-of-fit test statistic for this distribution.

2.1. *The logistic distribution*

The probability density function of the logistic distribution has the following form.

$$f_0(x;\mu,\sigma) = \frac{1}{\sigma}\frac{\exp\{-(x-\mu)/\sigma\}}{\left(1+\exp\{-(x-\mu)/\sigma\}\right)^2}, \qquad -\infty < x < \infty,\ \ \mu \in R,\ \ \sigma > 0$$

where μ and σ are the location and scale parameters, respectively. The cumulative distribution function can be obtained as

$$F_0(x;\mu,\sigma) = \frac{1}{1+\exp\{-(x-\mu)/\sigma\}}.$$

The mean and variance of the distribution are

$$E(X) = \mu \quad ; \quad Var(X) = \frac{\pi^2\sigma^2}{3}.$$

If $Z = (X-\mu)/\sigma$, then Z is called the standard logistic random variable with the following density.

$$f_0(z) = e^{-z}\left(1+e^{-z}\right)^{-2}, \qquad -\infty < z < \infty.$$

The maximum likelihood estimators (MLEs) of the unknown parameters μ and σ are used for computing the test statistic and therefore we express that how they can be obtained.
It is well-known that the MLEs of the parameters cannot be obtained explicitly. Therefore, we use the approximate maximum likelihood estimators (AMLEs) suggested by Balakrishnan and Cohen (1990), which are simple explicit estimators. Through a simulation study, Balakrishnan (1992) showed that these estimators are nearly as efficient as the MLEs. For the complete samples, the calculation of the AMLEs is described as follows. Also, for the doubly censored case these estimators can be found in Balakrishnan (1992). Doubly censored survival data arise in studies where both the time of the originating event and the failure event are either right- or interval-censored. This type of samplings scheme can result when the originating event is not directly observable but is detected via periodic screening. For example, it occurs in the study of HIV (human immunodeficiency virus) infected hemophiliacs whose stored blood samples are screened for evidence of infection with HIV, the virus that causes AIDS. In these studies, the induction period between infection with HIV and the failure event, onset of AIDS, may be doubly censored. This sampling scheme is also seen in transmission studies where the susceptible partners of HIV-positive individuals, whose times of infection are known to within a time interval, are periodically monitored for infection. For more details about the doubly censored data see De Gruttola and Lagakos (1989) and Kim et al. (1993). Also, a complete sample means that the sample has not been censored.
Suppose that $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ are the ordered sample from a logistic distribution. Define

$$p_i = \frac{i}{n+1},\ \ q_i = 1-p_i,\ \ \delta_i = p_i q_i, \qquad \gamma_i = p_i - \delta_i \ln\left(\frac{p_i}{q_i}\right), \qquad i = 1,2,...,n.$$

Further, let

$$z = 2\sum_{i=1}^{n}\delta_i,\ \ B = \frac{2}{z}\sum_{i=1}^{n}\delta_i X_{(i)},\ \ D = \sum_{i=1}^{n}(2\gamma_i - 1)X_{(i)}, \quad and \ \ E = 2\sum_{i=1}^{n}\delta_i X_{(i)}^2 - zB^2.$$

Then the approximate ML estimators of μ and σ can be obtained as

$$\hat{\mu} = B \quad ; \quad \hat{\sigma} = \frac{D + \left(D^2 + 4nE\right)^{1/2}}{2n}.$$

These estimators will be used to computation of the proposed test statistic.

2.2. *The test statistic*

Given a random sample $X_1, ..., X_n$ from a continuous probability distribution F with a density function $f(x)$, the hypothesis of interest is

$$H_0: \ f(x) = f_0(x;\mu,\sigma) = \frac{1}{\sigma}\frac{\exp\{-(x-\mu)/\sigma\}}{\left(1+\exp\{-(x-\mu)/\sigma\}\right)^2}\,, \qquad for\ some\ (\mu,\sigma) \in \Theta,$$

where μ and σ are specified or unspecified and $\Theta = R \times R^{+}$. The alternative to H_0 is

$$H_1: \ f(x) \neq f_0(x;\mu,\sigma)\,, \qquad for\ any\ (\mu,\sigma).$$

Without loss of any generality, one can reduce the above problem of goodness-of-fit, to testing the hypothesis of uniformity on the unit interval, by means of the probability integral transformation $U = F_0(X)$. Therefore if $U_i = F_0(X_i)\ ,\ i = 1, 2, ..., n$ be the transformed sample, the problem becomes the following testing uniformity.

$$H_0: \ f(u) = 1, \quad 0 < u < 1$$

against

$$H_1: \ f(u) \neq 1, \quad 0 < u < 1.$$

Now, we use the test proposed by Alizadeh (2023) for uniformity. Therefore, the proposed test statistic can be constructed as follows.

Let $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ denote the order statistics of the sample. The NCEX of a non-negative random variable X with the absolutely continuous density $f(x)$ and the cumulative distribution function $F(x)$ is given by

$$NCEX(F) = NCEX(f) = \frac{1}{2}\int_0^\infty \left[1 - F^2(x)\right] dx \ = \int_0^1 \frac{\phi(u)}{f(F^{-1}(u))} du\,,$$

where $\phi(u) = \frac{1-u^2}{2}\ ,\ 0 < u < 1$.

Alizadeh (2023) showed that for an f concentrated on $[0,1]$ one always has $0 \leq NCEX(f) \leq 1/2$, and for the $U(0,1)$ density the value of $NCEX(f)$ is 1/3, and this value being uniquely attained by the uniform distribution. Based on this property, we construct our test of H_0. A consistent test of the hypothesis of uniformity is then given by

$$T_n = \ NCEX(F_n),$$

where $NCEX(F_n)$ is the sample estimate of $NCEX(F)$ given by (Tahmasebi and Toomaj, 2020)

$$NCEX(F_n) = \frac{1}{2}\int \left[1 - F_n^2(x)\right]\, dx,$$

where F_n is the empirical distribution function defined by

$$F_n(x) = \sum_{i=1}^{n-1} \frac{i}{n} I_{(X_{(i)}, X_{(i+1)}]}(x)\,,\ \ x \in R$$

and I_A is the indicator function of A. Therefore, the proposed test statistic can be written as

$$\begin{aligned} T_n = NCEX(F_n) &= \tfrac{1}{2}\sum_{i=1}^{n-1}\int_{X_{(i)}}^{X_{(i+1)}}\left[1-\left(\tfrac{i}{n}\right)^2\right]dx = \tfrac{1}{2}\sum_{i=1}^{n-1}\left[1-\left(\tfrac{i}{n}\right)^2\right]\left(U_{(i+1)} - U_{(i)}\right) \\ &= \tfrac{1}{2}\sum_{i=1}^{n-1} W_i\left(U_{(i+1)} - U_{(i)}\right) = \tfrac{1}{2}\sum_{i=1}^{n-1} W_i Z_i\,, \end{aligned}$$

where $W_i = 1 - \left(\frac{i}{n}\right)^2$, $U_i = F_0(X_i;\hat{\mu},\hat{\sigma})\ ,\ i = 1, 2, ..., n$ and $Z_i = U_{(i+1)} - U_{(i)},\ i = 1, ..., n-1,$ denote the transformed sample spacings. Also, $\hat{\mu}$ and $\hat{\sigma}$ are the approximated maximum likelihood estimates of the parameters

μ and σ, respectively.

$$\hat{\mu} = B \quad and \quad \hat{\sigma} = \frac{D + \left(D^2 + 4nE\right)^{1/2}}{2n}.$$

Under the null hypothesis H_0, T_n converges in probability to 1/3 as $n \to \infty$ and under an alternative distribution on $[0, 1]$ with absolutely continuous density f, T_n converges in probability to a number smaller or larger than 1/3 as $n \to \infty$.

Guided by these properties, given any significance level α, and any finite sample size n, our test procedure is then defined by the critical region

$$T_n \leq T^*_{\alpha/2} \quad or \quad T_n \geq T^*_{1-\alpha/2} \tag{3}$$

where $T^*_{\alpha/2}$ and $T^*_{1-\alpha/2}$ are set so that the test has the desired level α for given n. For specific α and n, the T^*_α can be obtained by Monte Carlo methods.

We show that the test statistic is non-negative, i.e., $T_n \geq 0$, and also the test based on T_n is consistent.

Remark 1. Clearly, the proposed test statistic is invariant to transformations of location-scale and also the parameter space is transitive. Therefore, the distribution of the proposed test statistic T_n does not depend on the unknown parameters μ and σ. We will use this property to obtain the critical values of the test statistic.

Theorem 1. Let $X_1, ..., X_n$ be a random sample from an unknown continuous distribution F with probability density function $f(x)$. Then, we have

$$0 \leq T_n \leq 1/2.$$

Proof. Set $g(p) = \frac{1}{2}(1 - p^2)\,,\ 0 < p < 1$. It is easy to show that the maximum $g(p)$ is 1/2. Then, we can write

$$\begin{aligned} T_n = NCEX(F_n) &= \tfrac{1}{2} \sum_{i=1}^{n-1} \left[1 - \left(\tfrac{i}{n}\right)^2\right] \left(U_{(i+1)} - U_{(i)}\right) \\ &\leq \tfrac{1}{2} \sum_{i=1}^{n-1} \left(U_{(i+1)} - U_{(i)}\right) = \tfrac{1}{2} \left(U_{(n)} - U_{(1)}\right) \leq \tfrac{1}{2}. \end{aligned}$$

Also, it is clear that $NCEX(F_n) \geq 0$. Therefore, $0 \leq T_n \leq 1/2$.

Theorem 2. The test based on T_n is consistent.

Proof. We know that as $n \to \infty$, the maximum likelihood estimators $(\hat{\mu}, \hat{\sigma})$ trend to (μ, σ). Based on the Glivenko-Cantelli theorem asserts that

$$\sup_x |F_n(x) - F(x)| \to 0 \qquad as\ \ n \to \infty.$$

Tahmasebi and Toomaj (2022) showed that the empirical NCEX converges to the NCEX of X, i.e.,

$$NCEX(F_n) \to NCEX(F),$$

almost surely. Therefore, the proof of theorem is completed.

Theorem 3. The mean of the proposed test statistic T_n, under H_0 is

$$E(T_n) = \frac{\sum\limits_{i=1}^{n-1} W_i}{2(n+1)}\,,$$

where $W_i = 1 - \left(\frac{i}{n}\right)^2$ for $i = 1, ..., n-1$.

Proof. For the uniform distribution, the sample spacing $Z_i = U_{(i+1)} - U_{(i)}$ has the beta distribution with parameters 1 and n, i.e., $Z_i \sim Beta(1, n)$ and hence $E(Z_i) = (n+1)^{-1}$. Therefore,

$$E(T_n) = E\left[\frac{1}{2}\sum_{i=1}^{n-1} W_i Z_i\right] = \frac{1}{2}\sum_{i=1}^{n-1} W_i E(Z_i) = \frac{1}{2(n+1)}\sum_{i=1}^{n-1} W_i\,.$$

From the above theorem, we immediately obtain $\lim_{n\to\infty} E(T_n) = \frac{1}{3} = NCEX(U)$, where $NCEX(U)$ is the negative cumulative extropy of the uniform distribution on (0, 1).

Theorem 4. The variance of the proposed test statistic T_n, under H_0 is

$$Var(T_n) = \frac{n}{4(n+1)^2(n+2)}\sum_{i=1}^{n-1} W_i^2\,,$$

and also the second moment of T_n is

$$E(T_n^2) = \frac{1}{2(n+1)^2}\left[\left(1 - \frac{1}{n+2}\right)\sum_{i=1}^{n-1} W_i^2 + \sum\sum\nolimits_{i<j} W_i W_j\right].$$

Proof. Since $Z_i \sim Beta(1, n)$, $Var(Z_i) = \frac{n}{(n+1)^2(n+2)}$, and we have

$$Var(T_n) = Var\left[\frac{1}{2}\sum_{i=1}^{n-1} W_i Z_i\right] = \frac{1}{4}\sum_{i=1}^{n-1} W_i^2 Var(Z_i) = \frac{n}{4(n+1)^2(n+2)}\sum_{i=1}^{n-1} W_i^2\,.$$

$$\begin{aligned}
E(T_n^2) &= Var(T_n) + E^2(T_n) = \frac{n}{4(n+1)^2(n+2)}\sum_{i=1}^{n-1} W_i^2 + \frac{\left(\sum_{i=1}^{n-1} W_i\right)^2}{4(n+1)^2} \\
&= \frac{\sum_{i=1}^{n-1} W_i^2 + \left(\sum_{i=1}^{n-1} W_i\right)^2}{4(n+1)^2} - \frac{\sum_{i=1}^{n-1} W_i^2}{2(n+1)^2(n+2)} \\
&= \frac{\left[\sum_{i=1}^{n-1} W_i^2 + \sum\sum_{i<j} W_i W_j\right]}{2(n+1)^2} - \frac{\sum_{i=1}^{n-1} W_i^2}{2(n+1)^2(n+2)} \\
&= \frac{1}{2(n+1)^2}\left[\left(1 - \frac{1}{n+2}\right)\sum_{i=1}^{n-1} W_i^2 + \sum\sum\nolimits_{i<j} W_i W_j\right].
\end{aligned}$$

Remark 2. We can see that $Var(T_n)$ tends to zero as $n \to \infty$.

3. Percentage Points and Power Study

We obtain the percentage points by Monte Carlo methods. For different values of the sample size n, 100,000 samples of size n from the standard logistic distribution are generated. For each sample, the test statistic is computed. For level α, the lower(upper)-tail percentage points $T^*_{\alpha/2}(T^*_{1-\alpha/2})$ of the distribution of T_n are estimated by the $\alpha/2(1-\alpha/2)$ percentiles of the empirical distribution function of T_n based on the observed 100,000 samples. These estimates are presented in Table 1. Also, in Figure 1, we show the behavior of the critical values of the proposed test. We can see that when the sample size increases the lower and upper tails percentage points are close to each other and also percentage points tend to 1/3 when n increases.

Table 1. Percentage points of the proposed test statistic at level $\alpha = 0.05$.

n	$lower$	$upper$
10	0.2350	0.3246
20	0.2740	0.3294
30	0.2903	0.3309
40	0.2997	0.3317
50	0.3056	0.3322
75	0.3141	0.3328
100	0.3184	0.3331

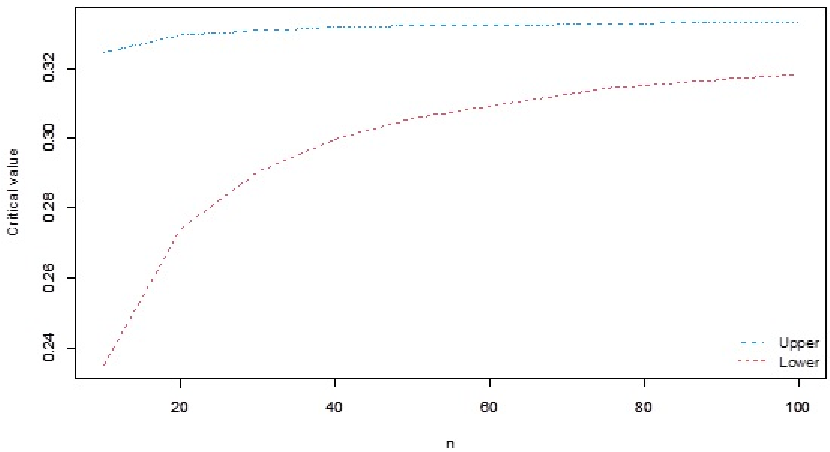

Figure 1. The lower and upper tails percentage points of the proposed test statistic for different values of sample sizes.

We evaluate the estimated type I error control using the critical values of the proposed test. We generated random samples from different logistic populations and then obtained the actual sizes of the proposed test. The results are displayed in Table 2. It is evident that the empirical percentiles given in Table 1 provides an excellent type I error control.

Table 2. Type I error control of the test for the nominal significance level $\alpha = 0.05$.

n	$logis(0, 0.5)$	$logis(0, 2)$	$logis(0, 4)$	$logis(0, 8)$
10	0.0503	0.0498	0.0487	0.0493
20	0.0483	0.0489	0.0483	0.0492
30	0.0472	0.0472	0.0497	0.0495
50	0.0494	0.0508	0.0496	0.0512

Figures 2 shows the estimated probability density functions of our test statistic with Monte Carlo samples for different sample sizes. We can see that the test statistic has closer values to 1/3 when n increases, which means that the bias of T_n with increasing decreases. We can also see that variance of T_n decreases when n increases.
The power values of the proposed test against various alternatives are computed by Monte Carlo simulations. We compare the power values of the proposed test with the existing tests. In our power comparisons, we consider the well-known tests which are applied in practice and statistical software. The test statistics of these tests are briefly

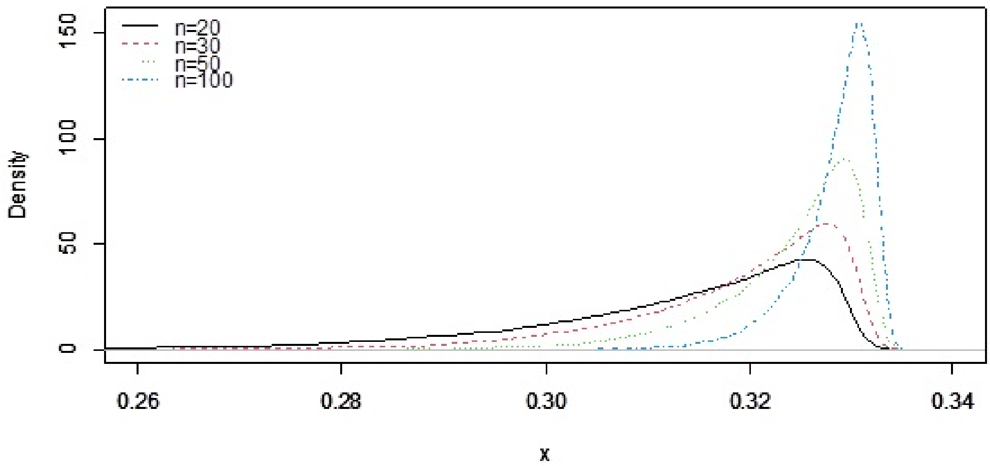

Figure 2. Empirical density functions of the test statistic generated with 100,000 samples of sizes $n = 20, 30, 50, 100$ from the standard logistic distribution.

described as follows. For more details about these tests, on can see D'Agostino and Stephens (1986).
Let $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ are the order statistics based on the random sample $X_1, ..., X_n$.

1. The Cramer-von Mises statistic (1931):

$$W^2 = \frac{1}{12n} + \sum_{i=1}^{n} \left(\frac{2i-1}{2n} - F_0(X_{(i)}; \hat{\mu}, \hat{\sigma}) \right)^2 .$$

2. The Watson statistic (1961):

$$U^2 = W^2 - n\big(\bar{P} - 0.5\big)^2,$$

where $\bar{P}$ is the mean of $F_0(X_{(i)}; \hat{\mu}, \hat{\sigma}),\ \ i = 1, ..., n.$

3. The Kolmogorov-Smirnov statistic (1933):

$$D = \max(D^+, D^-),$$

where

$$D^+ = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - F_0(X_{(i)}; \hat{\mu}, \hat{\sigma}) \right\}; \quad D^- = \max_{1 \leq i \leq n} \left\{ F_0(X_{(i)}; \hat{\mu}, \hat{\sigma}) - \frac{i-1}{n} \right\}.$$

4. The Kuiper statistic (1960):

$$V = D^+ + D^-.$$

5. The Anderson-Darling statistic (1952):

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \left\{ \log F_0(X_{(i)}; \hat{\mu}, \hat{\sigma}) + \log \left[1 - F_0(X_{(n-i+1)}; \hat{\mu}, \hat{\sigma}) \right] \right\} .$$

In the above test statistics, $F_0(x)$ is the cumulative distribution function of the logistic distribution and $(\hat{\mu}, \hat{\sigma})$ are the maximum likelihood estimates of the parameters (μ, σ).

Moreover, we consider the recent tests proposed by Alizadeh (2015) in our power comparisons. This test statistic is constructed based on the empirical likelihood ratio and it is shown that the test has a good performance as compared to the competing tests, see Alizadeh (2015). This test statistic is as follows.

$$LR = \frac{\min\limits_{1 \le m < n^{\delta}} \prod\limits_{j=1}^{n} \frac{2m}{n\left(X_{(j+m)} - X_{(j-m)}\right)}}{\prod\limits_{j=1}^{n} f_0(X_j; \hat{\mu}, \hat{\sigma})},$$

where $\delta \in (0,1)$. Here, based on the recommendation of Alizadeh (2015), we take $\delta = 0.5$. Also, f_0 denotes the logistic probability density function.
We consider various alternatives to power study of the considered tests. They are:

- the exponential distribution, $Exp(1)$;
- the Gamma distribution, $\Gamma(0.5, 1)$ and $\Gamma(2, 1)$;
- the lognormal distribution, $LN(0, 1)$;
- the Gumbel distribution, $Gu(0, 1)$;
- the Weibull distribution, $W(0.5, 1)$ and $W(2, 1)$;
- the inverse Gaussian distribution, $IG(1, 0.5)$, $IG(1, 1)$ and $IG(1, 2)$;
- the student's t distribution with 3 degrees of freedom, denoted by $t(3)$;
- the standard Cauchy distribution, denoted by $C(0, 1)$;
- the skew normal distribution, $SN(0, 1, 2)$ and $SN(0, 1, 3)$.

A simulation study is carried out to obtain the power values of all tests under the above alternatives. Under each alternative 100,000 samples of size 10, 20, 30 and 50 are generated and the test statistics are calculated. Then power of the corresponding test is computed by the frequency of the event "the statistic is in the critical region". The power estimates resulting based on a Monte Carlo study are given in Table 3, for $\alpha = 0.05$ and $n = 10, 20, 30, 50$. All computational codes are available on request from the corresponding author.
For each sample size and alternative, the bold type in these tables indicates the tests achieving the maximal power. From Table 3, it is seen that the test the proposed test T_n has a high performance and it is a powerful test (with the exception of the cases where $t(3)$, Cauchy and $SN(0, 1, 2)$ were the alternative). The power differences between the proposed test and the EDF-based tests are substantial. Also, when the sample size increases the performance of our test increases. Therefore, the proposed test T_n can be confidently recommended in practice.

Table 3. Empirical powers of the tests at significance level 5%.

$Alternative$	n	W^2	D	V	U^2	A^2	LR	T_n
$Exp(1)$	10	0.2934	0.2264	0.2950	0.2954	0.3280	0.3937	**0.4331**
	20	0.5804	0.5258	0.5940	0.5818	0.6869	0.8157	**0.8839**
	30	0.7810	0.7864	0.8188	0.7822	0.8970	0.9629	**0.9887**
	50	0.9548	0.9796	0.9772	0.9551	0.9928	0.9992	**1.0000**
$\Gamma(0.5, 1)$	10	0.5911	0.5134	0.6142	0.5948	0.6378	**0.7512**	0.7495
	20	0.9138	0.9240	0.9356	0.9150	0.9532	0.9902	**0.9939**
	30	0.9875	0.9954	0.9946	0.9876	0.9973	0.9998	**0.9999**
	50	0.9999	1.0000	1.0000	0.9999	1.0000	1.0000	**1.0000**
$\Gamma(2, 1)$	10	0.1459	0.1186	0.1406	0.1477	0.1606	0.1759	**0.2016**
	20	0.2708	0.2032	0.2436	0.2720	0.3434	0.4184	**0.5332**
	30	0.4007	0.3094	0.3749	0.4023	0.5478	0.6437	**0.7967**
	50	0.6198	0.5471	0.6074	0.6207	0.8168	0.9010	**0.9834**
$LN(0, 0.5)$	10	0.1463	0.1188	0.1371	0.1459	0.1620	0.1516	**0.2038**
	20	0.2698	0.2039	0.2379	0.2706	0.3410	0.3417	**0.4780**
	30	0.3740	0.2888	0.3368	0.3748	0.5101	0.5118	**0.7017**
	50	0.5807	0.4780	0.5419	0.5810	0.7683	0.7756	**0.9335**
$LN(0, 1)$	10	0.4623	0.3843	0.4596	0.4640	0.4999	0.5200	**0.5931**
	20	0.7910	0.7459	0.7930	0.7919	0.8577	0.8960	**0.9475**
	30	0.9235	0.9160	0.9308	0.9242	0.9689	0.9833	**0.9965**
	50	0.9936	0.9952	0.9952	0.9937	0.9990	0.9995	**1.0000**
$LN(0, 2)$	10	0.8636	0.8225	0.8732	0.8658	0.8834	**0.9316**	0.9294
	20	0.9937	0.9954	0.9955	0.9938	0.9970	0.9996	**0.9998**
	30	0.9998	1.0000	1.0000	0.9998	1.0000	1.0000	**1.0000**
	50	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	**1.0000**
$Gu(0, 1)$	10	0.0920	0.0921	0.0800	0.0867	0.1001	0.0963	**0.1100**
	20	0.1411	0.1134	0.1244	0.1409	0.1753	0.1749	**0.2403**
	30	0.1928	0.1447	0.1614	0.1926	0.2634	0.2629	**0.3730**
	50	0.2913	0.2089	0.2443	0.2916	0.4341	0.4247	**0.6170**
$W(0.5, 1)$	10	0.8106	0.7632	0.8296	0.8146	0.8412	**0.9127**	0.9080
	20	0.9908	0.9937	0.9946	0.9908	0.9962	0.9994	**0.9997**
	30	0.9996	0.9999	0.9998	0.9996	1.0000	1.0000	**1.0000**
	50	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	**1.0000**
$W(2, 1)$	10	0.0566	0.0540	0.0588	0.0578	0.0557	**0.0860**	0.0574
	20	0.0891	0.0698	0.0772	0.0896	0.0972	**0.1729**	0.1289
	30	0.1208	0.0911	0.1036	0.1214	0.1462	**0.2627**	0.2371
	50	0.1882	0.1349	0.1553	0.1884	0.2694	0.4693	**0.5301**
$IG(1, 0.5)$	10	0.5945	0.5092	0.5978	0.5962	0.6342	0.6854	**0.7275**
	20	0.9080	0.8916	0.9142	0.9088	0.9437	0.9703	**0.9861**
	30	0.9816	0.9838	0.9869	0.9818	0.9943	0.9980	**0.9996**
	50	0.9998	1.0000	0.9999	0.9998	1.0000	1.0000	**1.0000**

Table 3. Continued.

$Alternative$	n	W^2	D	V	U^2	A^2	LR	T_n
$IG(1,1)$	10	0.4005	0.3232	0.3983	0.4032	0.4388	0.4684	**0.5388**
	20	0.7321	0.6712	0.7266	0.7330	0.8129	0.8594	**0.9262**
	30	0.8892	0.8752	0.8984	0.8897	0.9526	0.9730	**0.9936**
	50	0.9882	0.9904	0.9914	0.9882	0.9986	0.9994	**1.0000**
$IG(1,2)$	10	0.2458	0.1914	0.2375	0.2479	0.2736	0.2698	**0.3374**
	20	0.4718	0.3846	0.4516	0.4722	0.5712	0.6087	**0.7432**
	30	0.6475	0.5599	0.6284	0.6481	0.7800	0.8192	**0.9294**
	50	0.8652	0.8364	0.8635	0.8656	0.9609	0.9718	**0.9976**
$t(3)$	10	0.1200	0.1091	0.1152	0.1196	**0.1272**	0.0744	0.0911
	20	0.1742	0.1477	0.1552	0.1738	**0.1900**	0.0696	0.0922
	30	0.2060	0.1737	0.1840	0.2056	**0.2275**	0.0662	0.0823
	50	0.2588	0.2178	0.2355	0.2589	**0.2924**	0.0543	0.0716
$C(0,1)$	10	0.5369	0.5034	0.5195	0.5362	**0.5389**	0.3351	0.2338
	20	0.8048	0.7638	0.7805	0.8041	**0.8070**	0.5269	0.2699
	30	0.9186	0.8898	0.9033	0.9180	**0.9216**	0.6702	0.3725
	50	0.9840	0.9760	0.9806	0.9840	**0.9844**	0.8326	0.6179
$SN(0,1,2)$	10	0.0484	0.0474	0.0501	0.0495	0.0458	**0.0624**	0.0453
	20	0.0610	0.0542	0.0560	0.0614	0.0604	**0.0921**	0.0696
	30	0.0756	0.0646	0.0664	0.0758	0.0836	**0.1205**	0.0967
	50	0.1016	0.0820	0.0865	0.1020	0.1261	**0.1821**	0.1644
$SN(0,1,3)$	10	0.0602	0.0570	0.0594	0.0612	0.0587	**0.0780**	0.0658
	20	0.0911	0.0728	0.0766	0.0920	0.1016	**0.1352**	0.1263
	30	0.1286	0.0976	0.1076	0.1293	0.1549	0.1936	**0.2022**
	50	0.1918	0.1379	0.1554	0.1924	0.2632	0.3102	**0.3623**

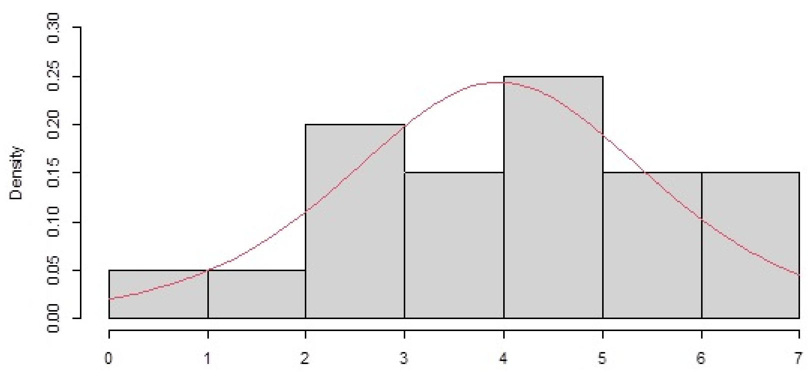

Figure 3. Histogram of data and a fitted logistic density function.

Table 4. Nickel concentrations (ppb) at four monitoring wells, USEPA (1992).

Observations $n = 20$						
58.8	331.0	587.0	1.0	14.0	3.1	262.0
64.4	942.0	56.0	39.0	85.6	8.7	151.0
10.0	19.0	27.0	637.0	81.5	21.4	

4. An Illustrative Example

Through an example, we illustrate how the proposed test can be applied to test the goodness-of-fit for the logistic distribution when the observations are available.

Example 1. We consider the data set presented by Modarress et al. (2002, p. 550). The data set consists of $n = 20$ nickel concentrations in parts per billion. These data are presented in Table 4 and histogram of them is presented in Figure 3.

Modarress et al. (2002) applied the Anderson-Darling procedure to test if the underlying distribution is log-normal, log-logistic or double exponential. They found that there is considerable uncertainty about the underlying model. Any of the three distributions are not rejected: the obtained p-values are 0.25 for the log normal and 0.15 for both the log-logistic and the double exponential.

Applying the proposed procedure to the log of this data set the following is obtained:

$$z = 6.984\ ,\ B = 3.950\,,\ D = \ 8.798\ ,\quad and\ \ E = 12.010.$$

Therefore, the approximate ML estimators of μ and σ are:

$$\hat{\mu} = 3.950 \quad and \ \ \hat{\sigma} = 1.025.$$

The value of the test statistic is $T_n = 0.32377$ and the lower and upper tail percentage points at the 5% are 0.2740 and 0.3294, respectively. Since the values of the test statistic is between the lower and upper tail percentage points, the log-logistic hypothesis is not rejected for these data at significance level of 0.05. Therefore, we can conclude that the data come from a log logistic distribution.

5. Conclusions

In this paper, we have introduced a goodness-of-fit test for the logistic distribution based on the negative cumulative extropy, and have shown that the test outperforms the EDF-goodness-of-fit tests which are commonly used in practice. The proposed test statistic is easy to compute. Consistency of the test statistic is shown and also the mean, variance and other properties of the proposed test statistic have presented.

We have carried out an extensive power comparison using Monte Carlo simulation. Through the obtained results, we have shown that the proposed test outperforms in most cases all other competitor tests. Finally, we have presented a real data set and have illustrated how the proposed test can be applied to test the goodness-of-fit for the logistic distribution when a sample is available.

Acknowledgement

The authors are grateful to anonymous referees and the associate editor for providing some useful comments on an earlier version of this manuscript.

REFERENCES

1. Abd-Elfattah, A.M. (2007), Goodness of fit tests for Logistic distribution using the stabilized probability plot, Journal of Applied Sciences Research, 3, 1650-1657.
2. Abdul Sathar, E.I. and Dhanya Nair, R. (2021), On dynamic survival extropy, Communications in Statistics - Theory and Methods, 50, 1295-1313.
3. Alizadeh, H.N. (2015), Empirical likelihood ratio-based goodness-of-fit test for the logistic distribution, Journal of Applied Statistics, 42, 1973-1983.
4. Alizadeh, H.N. (2017), Gini index-based goodness-of-fit test for the logistic distribution. Communications in Statistics - Theory and Methods, 46, 7114–7124.
5. Alizadeh, H.N. (2021a), Testing the Validity of Laplace Model Against Symmetric Models, Using Transformed Data, Statistics, Optimization & Information Computing, 10(4), 1162-1167.
6. Alizadeh, H.N. (2021b), Testing the Validity of Lindley Model Based on Informational Energy with Application to Real Medical Data, Statistics, Optimization & Information Computing, 10(2), 372-382.
7. Alizadeh, H.N. (2022), A New Goodness-of-fit test for the Logistic distribution, Sankhya B, 84, 303–319.
8. Alizadeh, H.N. (2023), Testing uniformity based on negative cumulative extropy, Communications in Statistics - Theory and Methods,52, 4998-5009.
9. Alizadeh, H.N. and Jarrahiferiz, J. (2019), On the estimation of extropy, Journal of Nonparametric Statistics, 31, 88-99.
10. Alizadeh, H.N. and Shafaei, M.N. (2023), On Testing the Adequacy of the Lindley Model and Power Study, Statistics, Optimization & Information Computing, 11(3), 719-729.
11. Alizadeh, H.N. and Shafaei, M.N. (2024), Empirical likelihood ratio-based goodness-of-fit test for the Lindley distribution, Statistics, Optimization & Information Computing.
12. Al-Shomrani, A.A., Shawky, A.I. and Arif, O.H. (2016), Log-logistic distribution for survival data analysis using MCMC. SpringerPlus 5, 1774.
13. Al-Subh, S. A.; Alodat, M. T.; Ibrahim, Kamaruzaman; and Jemain, Abdul Aziz (2012) Modified EDF goodness of fit tests for Logistic distribution under SRS and RSS, Journal of Modern Applied Statistical Methods, Vol. 11 : Iss. 2, Article 11.
14. Asadi, M. and Zohrevand, Y. (2007), On the dynamic cumulative residual entropy, Journal of Statistical Planning and Inference, 137, 1931–1941.
15. Balakrishnan, N. (1992), Handbook of the Logistic Distribution, Marcel Dekker, New York.
16. Balakrishnan, N. and Cohen, A.C. (1990), Order Statistics and Inference: Estimation Methods, Academic Press, Boston.
17. D'Agostino, R.B. and Stephens, M.A. (Eds.) (1986), Goodness-of-fit Techniques, New York: Marcel Dekker.
18. De Gruttola, V. and Lagakos, S.W. (1989), Analysis of doubly-censored survival data, with application to AIDS, Biometrics, 45, 1–12.
19. Di Crescenzo, A. and Longobardi, M. (2006), On weighted residual and past entropies, Scientiae Mathematicae Japonicae, 64, 255–266.
20. Di Crescenzo, A. and Longobardi, M. (2009a), On cumulative entropies, Journal of Statistical Planning and Inference, 139, 4072-4087.
21. Di Crescenzo, A. and Longobardi, M. (2009b), On cumulative entropies and lifetime estimations, In: Mira, J., et al. (Eds.), IWINAC 2009, Part I, LNCS 5601. Springer, Berlin, pp. 132–141. DOI: 10.22237/jmasm/1351743000
22. Drissi, N., Chonavel, T., Boucher, J.M. (2008), Generalized cumulative residual entropy for distributions with unrestricted supports, Research Letters in Signal Processing, article ID 790607, 5p.
23. Eltehiwy, M. (2023), The odd log-logistic power inverse Lindley distribution: Model, Properties and Applications, Statistics, Optimization & Information Computing, 12, 91-108.
24. Esmaeili, H., Lak, F., Alizadeh, M., and Dehghan monfared, M. esmail. (2020), The Alpha-Beta Skew Logistic Distribution: Properties and Applications, Statistics, Optimization & Information Computing, 8, 304-317.
25. Finney, D.J. (1947), The principals of biological assay, Journal of Royal Statistical Society, Series B, 9, 46–91.
26. Finney, D.J. (1952), Statistical Methods in Biological Assay, Hafner, New York.
27. Heydari, S., Afshari, M., Tahmasbi, S., and Alizadeh, M. (2021), Comparison of subspace dimension reduction methods in Logistic regression, Statistics, Optimization & Information Computing, 11, 422-444.
28. Jahanshahi, S.M.A., H. Zarei, and A.H. Khammar, (2020), On cumulative residual extropy, Probability in the Engineering and Informational Sciences, 34, 605-625.
29. Kapodistria, S. and Psarrakos, G. (2012), Some extensions of the residual lifetime and its connection to the cumulative residual entropy, Probability in the Engineering and Informational Sciences, 26, 129-146.
30. Kayal S. (2016), On generalized cumulative entropies. Probability in the Engineering and Informational Sciences, 30, 640-662.
31. Kim, M.Y., De Gruttola, V. and Lagakos, S.W. (1993), Analyzing doubly censored data with covariates, with application to AIDS, Biometrics, 49, 13–22.
32. Lad, F., Sanfilippo, G., and Agro, G. (2015), Extropy: Complementary Dual of Entropy, Statistical Science, 30, 40–58.
33. Meintanis, S.G. (2004), Goodness-of-fit tests for the Logistic distribution based on empirical transforms, Sankhyā: The Indian Journal of Statistics, 66, 306-326.
34. Modarres, R., Nayak, T. and Gastwirth, J.L. (2002), Estimation of upper quantiles under model and parameter uncertainty, Computational Statistics & Data Analysis, 39, 529–554.
35. Navarro, J. and del Aguila, Y. and Asadi M. (2010), Some new results on the cumulative residual entropy, Journal of Statistical Planning and Inference, 140, 310-322.
36. Navarro, J. and Psarrakos, G. (2017), Characterizations based on generalized cumulative residual entropy functions, Communication in Statistics- Theory and Methods, 46, 1247-1260.
37. Nikitin Y.Y. and Ragozin I.A. (2019), Goodness-of-fit tests based on a characterization of logistic distribution, Vestnik St. Petersburg University, Mathematics, 52, 169–177.
38. Nikitin Y.Y. and Ragozin I.A. (2020), Goodness-of-fit tests for the logistic location family, Journal of Applied Statistics, 47, 2610–2622.

39. Opone, F., Karakaya, K., and Ubaka, N. (2023). Statistical analysis of Covid-19 data using the odd log Logistic Kumaraswamy distribution, Statistics, Optimization & Information Computing, 12, 210-230.
40. Pearl, R. and Reed, L.J. (1920), On the rate of growth of the population of the United States since 1790 and its mathematical representation, Proceedings of the National Academy of Science, 6, 275–288.
41. Pearson E.S. and Hartley H.O., (1972), Biometrika Tables for statisticians. Cambridge University Press.
42. Plackett, R.L. (1959), The analysis of life test data, Technometrics, 1, 9–19.
43. Psarrakos, G. and Toomaj, A. (2017), On the generalized cumulative residual entropy with applications in actuarial science, Journal of Computational and Applied Mathematics, 309, 186-199.
44. Qiu, G. (2017), The extropy of order statistics and record values, Statistics & Probability Letters, 120, 52-60.
45. Qiu, G. and Jia, K. (2018a), Extropy estimators with applications in testing uniformity, Journal of Nonparametric Statistics, 30, 182-196.
46. Qiu, G. and Jia, K. (2018b), The residual extropy of order statistics, Statistics & Probability Letters, 133, 15-22.
47. Qiu, G., Wang, L. and Wang, X. (2019), On extropy properties of mixed systems, Probability in the Engineering and Informational Sciences, 33, 471–86.
48. Rao, M. (2005), More on a new concept of entropy and information. Journal of Theoretical Probability, 18, 967–981.
49. Rao, M., Chen, Y., Vemuri, B.C., Wang, F. (2004), Cumulative residual entropy: a new measure of information, IEEE Transactions on Information Theory, 50, 1220–1228.
50. Shannon, C.E. (1948), A mathematical theory of communications, Bell System Technical Journal, 27, 379–423; 623–656.
51. Tahmasebi, S. and Toomaj, A. (2022), On negative cumulative extropy with applications, Communications in Statistics - Theory and Methods, 51, 5025-5047.
52. Toomaj, A., Sunoj, S.M. and Navarro, J. (2017), Some properties of the cumulative residual entropy of coherent and mixed systems, Journal of Applied Probability, 54, 379–93.
53. USEPA, Statistical analysis of ground-water monitoring data at RCRA facilities, EPA/530-R-93-003, Office of Solid Waste, Washington, DC, 1992.
54. Verhulst, P.F. (1838), Notice sur la loi que la population poursuit dans son accroissement, Correspondance mathématique et physique, 10, 113-121.
55. Verhulst, P.F. (1845), Recherches mathematiques sur la loi d'accroissement de la population, Nouveaux Memoires de l'Academie Royale des Sciences et Belles-Lettres de Bruxelles, 18, 1–45.