Chilean Journal of Statistics
Vol. 4, No. 1, April 2013, 55–67

Goodness-of-fit
Research Paper

Powerful goodness-of-fit tests for the extreme value distribution

Mir Nabi Pirouzi Fard and Björn Holmquist

Department of Statistics, Lund University, Lund, Sweden

(Received: 10 October 2011 · Accepted in final form: 03 December 2012)

Abstract

We investigate several procedures for goodness-of-fit of the extreme value distribution. The procedures make use of recent available accurate approximations of the means and variances of order statistics from the standardized extreme value distribution, and are either modifications in estimation techniques of earlier proposed test statistics or are newly introduced test statistics based on the regression of the order statistics on their means. Properties in terms of critical values of the tests are investigated and the improved power of the tests are evaluated for a vast range of alternative distributions. Finally, we illustrate the different tests by a real data set.

Keywords: Extreme value distribution · Generalized least squares · Goodness-of-fit test · Means, variances and covariances of order statistics.

Mathematics Subject Classification: Primary 62F03 · Secondary 62E99.

1. Introduction

The extreme value distribution is widely used in the study of size effects on material strengths, the reliability of systems made up of a large number of components, in assessing the level of air pollution and in flood frequency analysis. The distribution has also an important role in modelling lifetime data. Considerable efforts have been dedicated to testing the hypothesis that data originates from an extreme value distribution. For reviews of the subject the reader is referred to D'Agostino and Stephens (1986) and Balakrishnan and Rao (1998).

Let Z have an extreme value distribution with cumulative distribution function

$$F(y) = 1 - \exp\left(-\exp\left((y-\mu)/\theta\right)\right), \qquad -\infty < y < \infty, \tag{1}$$

where the parameters $\theta > 0$ and $-\infty < \mu < \infty$. The mean and variance of this distribution, which sometimes is referred to as the Gumbel distribution, are respectively, $\mathrm{E}[Z] = \mu - \gamma\theta$ and $\mathrm{Var}[Z] = \theta^2\pi^2/6$, where $\gamma \approx 0.57721$ is Euler's constant.

*Corresponding author: Mir Nabi Pirouzi Fard. Department of Statistics, Lund University, Box 743, SE-220 07 Lund, Sweden. Email: mirnabi.pirouzifard@stat.lu.se

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

http://www.soche.cl/chjs

The many different test statistics that exist for testing for a specific location-scale family of distributions, such as the one in Equation (1), are basically based on properties of one particular distribution in that family - the standard or 'standardized' distribution - and of estimates of the best-fitting location/scale within this family. It is our belief that if both these parts are given proper attention, considerable improvements of the test procedures, in terms of power of these tests, can be achieved.

In this paper, we consider goodness-of-fit tests based either on regression, the empirical distribution function (EDF) or the stabilized probability plot. A goodness-of-fit test is a test of H_0: 'a random sample of Z-values comes from an extreme value distribution' with some unknown parameters μ and θ. In Section 2, we discuss the different test statistics for testing random samples from an extreme value distribution. We suggest modifications in the estimation procedure in the existing test statistics and also introduce new statistics for this testing purpose. In Section 3, we describe the Monte Carlo study set up to investigate the properties of these tests. The results of the power comparisons and tables of significance points are given in Section 4. In Section 5, we consider modifications needed when applied to censored data. In Section 6, we try an example with data. Finally, in Section 7, we provide some conclusions of this work.

2. Test Statistics

2.1 Estimation of location and scale parameters

Goodness-of-fit tests mostly require estimation of location and scale parameters in the tested distribution $F(y)$ which is the cumulative distribution function given in Equation (1).

Let $X_{(1)} < \cdots < X_{(n)}$ denote an ordered random sample of size n from the distribution given in Equation (1) with $\mu = 0$ and $\theta = 1$ (i.e., the standard extreme value distribution), and let

$$m_i = \mathrm{E}[X_{(i)}] \quad (i = 1, \ldots, n) \quad \text{and} \quad \sigma_{ij} = \mathrm{Cov}(X_{(i)}, X_{(j)}) \quad (i, j = 1, \ldots, n)\,.$$

Much effort have been assigned to the description of these means and covariances. Lieblein and Zelen (1956) presented the expected values, variances and covariances of the order statistics from the standard extreme value distribution for $n = 1(1)6$. Lieblein and Salzer (1957) presented a table of expected values of order statistics for $n = 1(1)10(5)25$ and the first 26 largest values for $n = 30(5)60(10)100$. White (1967, 1969), tabulated means and variances of order statistics for $n = 1(1)50(5)100$. For sample sizes $n = 1(1)15(5)30$, Balakrishnan and Chan (1992) presented tables of means, variances and covariances of the order statistics. They also in an unpublished report in 1992 at McMaster University, Hamilton, presented tables for all sample sizes up to 30.

Recently, accurate approximate expressions of the means of order statistics, were suggested by Pirouzi Fard and Holmquist (2007a)

$$m_i \approx \begin{cases} -\log(n) - \gamma, & \text{for} \quad i = 1; \\ \log\left(-\log\left(1 - \left[\frac{i-0.4866}{n+0.1840}\right]\right)\right), & \text{for} \quad i = 2, \ldots, n\,. \end{cases} \tag{2}$$

Also approximate expressions of the variances and covariances of the order statistics were given by Pirouzi Fard and Holmquist (2008) as

$$\sigma_{ij} \approx \begin{cases} \pi^2/6, & \text{for} \quad i=j=1; \\ \frac{[i-0.469]([n+0.831-i][n+0.073])^{-1}}{\log\left(\frac{n+0.831-i}{n+0.356}\right)\log\left(\frac{n+0.779-j}{n+0.356}\right)}, & \text{for} \quad 1 \le i \le j \le n, \end{cases} \tag{3}$$

where $\sigma_{ji} = \sigma_{ij}$ is the covariance of the ith and jth order statistics of standard extreme value distribution.

Define m to be the $(n \times 1)$ vector of the expected values m_i, Σ the $(n \times n)$ matrix of variances and covariances σ_{ij}.

If we let ζ be a vector of ordered random observations from Equation (1) for general μ and θ, then the elements $Z_{(i)}$ of ζ may be expressed as $Z_{(i)} = \mu + \theta X_{(i)}$, for $i = 1, \ldots, n$. Hence, $\mathrm{E}[Z_{(i)}] = \mu + \theta m_i$, in terms of the order statistics means from the standard extreme value distribution. By defining ε_i from $Z_{(i)} = \mathrm{E}[Z_{(i)}] + \varepsilon_i$, we thus have that in a regression model

$$Z_{(i)} = \mu + \theta m_i + \varepsilon_i, \qquad i = 1, \ldots, n, \tag{4}$$

the points of the n pairs $(m_i, Z_{(i)})$ should be approximately on a straight line with intercept μ on the vertical axis and slope θ. The parameters in Equation (4) can be estimated by a suitable method.

The observed values in Equation (4) are order statistics with $\mathrm{Var}[Z_{(i)}] = \theta^2 \mathrm{Var}[X_{(i)}]$ and $\mathrm{Cov}(Z_{(i)}, Z_{(j)}) = \theta^2 \mathrm{Cov}(X_{(i)}, X_{(j)})$. Here $\mathrm{Var}[X_{(i)}]$ depends on i and also $\mathrm{Cov}(X_{(i)}, X_{(j)})$ depends on i and j. Thus, ordinary least squares (OLS) estimators of μ and θ are not minimum variance estimators.

The best linear unbiased estimators of μ and θ can be obtained from the generalized least-squares (GLS) regression of the order statistic (Aitken, 1935; Lloyd, 1952) and are given by

$$\hat{\mu} = \frac{m^{\mathrm{T}}\Sigma^{-1}[m1^{\mathrm{T}} - 1m^{\mathrm{T}}]\Sigma^{-1}\zeta}{1^{\mathrm{T}}\Sigma^{-1}1m^{\mathrm{T}}\Sigma^{-1}m - (1^{\mathrm{T}}\Sigma^{-1}m)^2} \tag{5}$$

and

$$\hat{\theta} = \frac{1^{\mathrm{T}}\Sigma^{-1}[1m^{\mathrm{T}} - m1^{\mathrm{T}}]\Sigma^{-1}\zeta}{1^{\mathrm{T}}\Sigma^{-1}1m^{\mathrm{T}}\Sigma^{-1}m - (1^{\mathrm{T}}\Sigma^{-1}m)^2}, \tag{6}$$

where 1 is a n-dimensional vector of ones.

2.2 Regression tests based on coherence with linearity

Let $z_{(1)}, \ldots, z_{(n)}$ be the ordered sample of size n from the distribution function $F(y)$. The ith fitted value $\hat{z}_i$ is given by the equation

$$\hat{z}_i = \hat{\mu} + \hat{\theta} m_i, \qquad i = 1, \ldots, n, \tag{7}$$

where $\hat{\mu}$ and $\hat{\theta}$ are obtained from Equations (5) and (6).

We are interested in using residuals to test how well the data fit $\{\hat{z}_i\}$. The residuals can be expressed as $z_{(i)} - \hat{z}_i$ i.e., the differences between the observed values and the values given by the model.

We examine linearity of data from the error (or unexplained) sum of squares (ESS) divided by the total sum of squares (TSS) given by

$$T_1 = \frac{\text{ESS}}{\text{TSS}} = \frac{(\zeta - \hat{\zeta})^{\mathrm{T}}(\zeta - \hat{\zeta})}{(\zeta - \bar{z}1)^{\mathrm{T}}(\zeta - \bar{z}1)} = \frac{\sum_{i=1}^{n}(z_{(i)} - \hat{\mu} - \hat{\theta}m_i)^2}{\sum_{i=1}^{n}(z_{(i)} - \bar{z})^2},$$

where $\hat{\zeta}$ is the estimated values (by using the GLS regression) with elements given in Equation (7) and $\bar{z} = 1^{\mathrm{T}}\zeta/1^{\mathrm{T}}1$. Large values of T_1 indicate deviations from the extreme value distribution. If OLS estimates would have been used, the corresponding test statistic T_1 would have been related to that based on the squared correlation coefficient between $\{z_{(i)}\}$ and $\{m_i\}$. That statistic was studied by Kinnison (1989) using rank percentiles in the inverse of the extreme value distribution function as approximations of the expected values. Here we use the approximations given in Equations (2) and (3) in Equations (5) and (6) to estimate location and scale parameters.

Another possibility comes from defining $y_i = [z_{(i)} - \hat{\mu}]/\hat{\theta}$ for $i = 1, \ldots, n$. Since $\mathrm{E}[y_i] \approx m_i$ we let the regression estimate

$$B_n = \frac{\sum_{i=1}^{n} y_i m_i}{\sum_{i=1}^{n} m_i^2} = \frac{\sum_{i=1}^{n}(z_{(i)} - \hat{\mu})(\hat{z}_i - \hat{\mu})}{\sum_{i=1}^{n}(\hat{z}_i - \hat{\mu})^2},$$

be a test statistic with deviations from one indicating deviations from the extreme value distribution.

2.3 Goodness-of-fit tests based on the EDF

The empirical distribution function of the sample is given by $F_n(y) = k/n, z_{(k)} < y < z_{(k+1)}$, which is a step function with step size $1/n$ at the order statistics. The distance between the EDF and the hypothesized distribution, $F(y) = \hat{F}(y)$ can be considered as a way of testing for H_0. Large values of the test statistics indicate that H_0 should be rejected. In our case, the $\hat{F}(y)$ is given by

$$\hat{F}(y) = 1 - \exp\left(-\exp((y - \hat{\mu})/\hat{\theta})\right),$$

i.e., the estimated cumulative distribution function of the extreme value distribution where the parameters are estimated by GLS according to Equations (5) and (6).

One EDF test is based on the Anderson-Darling statistic, A_n^2:

$$A_n^2 = n \int_{-\infty}^{\infty} \frac{(F_n(y) - \hat{F}(y))^2}{\hat{F}(y)(1 - \hat{F}(y))} d\hat{F}(y).$$

This test statistic can be calculated by the formula

$$A_n^2 = -n - \sum_{i=1}^{n} \frac{2i-1}{n}[\log(\hat{F}(z_{(i)})) + \log(1 - \hat{F}(z_{(n+1-i)}))],$$

in terms of the ordered sample. This test is known to have good properties. The test has been studied for the extreme value distribution by Littel et al. (1979), in which case the maximum likelihood estimates were used in substitution for μ and θ. The difference here compared to standard procedure is the way the parameters are estimated. Another test statistic is

$$L_n = n^{-1/2} \sum_{i=1}^{n} \frac{\max\left\{\frac{i}{n} - \hat{F}(z_{(i)}), \hat{F}(z_{(i)}) - \frac{i-1}{n}\right\}}{(\hat{F}(z_{(i)})(1-\hat{F}(z_{(i)})))^{1/2}},$$

introduced by Liao and Shimokawa (1999), which combines the characteristics of the Kolmogorov-Smirnov, the Cramér-von Mises and the Anderson-Darling statistics and were shown to have good properties. A graphical plotting technique was used to estimate the parameters in F giving $\hat{F}$; see Liao and Shimokawa (1999, p. 29).

We here use the approximations given in Equation (2) of the means of the order statistics as the basis for replacing the plotting positions $p_i = [i-0.5]/n$ used in L_n. This modification of the test statistic is denoted L_n^+.

2.4 Goodness-of-fit test based on stabilized probability plot

Let $z_{(1)}, \ldots, z_{(n)}$ be the ordered observations in a random sample of size n from the distribution of the form $F(y) = F_0([y-\mu]/\theta)$, where μ is a location parameter and θ is a scale parameter. A probability plot is a scatter of the $z_{(i)}$'s versus a corresponding theoretical quantities $u_i = F_0^{-1}(d_i)$, where d_i is an estimate of $F_0([z_{(i)}-\mu]/\theta)$. In such plots the points should lie fairly near the line $z_{(i)} = \mu + \theta u_i$, indicating that the hypothesized distribution is a reasonable model for the data.

The interpretation of the plot can be complicated due to the existence of outliers and the non-equal variances of the plotted points. The stabilized probability plot was introduced by Michael (1983) to handle the problem, and is formed by plotting

$$s_i = [2/\pi]\sin^{-1}\left(\sqrt{F_0\left([z_{(i)}-\mu]/\theta\right)}\right) \quad \text{against} \quad r_i = [2/\pi]\sin^{-1}\left(\sqrt{[i-0.5]/n}\right), \qquad (8)$$

where, according to Michael (1983), s follows the sine distribution and all its order statistics have the same asymptotic variance. Hence by this transformation, the variance of the plotted points are approximately equal over the range of probability values.

A goodness-of-fit statistic based on the stabilized probability plot was also suggested by Michael (1983) as $D_{sp} = \max|r_i - s_i|$. Kimber (1985) used the statistic D_{sp} for testing of the extreme value distribution of maxima and applied Downton's (1966) estimates of μ and θ to obtain the critical values. Coles (1989) investigated the statistic D_{sp} for testing the extreme value distribution of minima and denoted it by D_{sp}^*. He estimated the parameters μ and θ in Equation (1), by using Blom's procedure (1958) and showed that due to the improved estimation procedure, the test statistic D_{sp}^* had higher power than Kimber's proposed test statistic. Also, in this test, we consider the estimates obtained by using Equations (5) and (6) in Equation (8), giving $\hat{s}_i$. The test statistic is then denoted $D_{sp}^+ = \max|r_i - \hat{s}_i|$.

3. The Monte Carlo Study

A simulation study was set up in order to investigate the properties of the different tests for testing extreme value distribution and in particular to investigate the power of the different test statistics for a large range of different alternative distributions. We have chosen distributions, such as Weibull, log-normal and normal that are linked to the strength data of metallic and composite materials. We also have chosen distributions such as Cauchy, logistic, Laplace, Student-t, beta and chi-square distributions that are often used in evaluating the power of test statistics for the extreme value distribution.

The critical values for all test statistics (except for L_n which were taken from Liao and Shimokawa (1999)) were for $n = 10, 20, 50$ and 100, obtained from 10^6 replicates of independent samples of size n from the standard extreme value distribution. These critical values were also verified to be invariant for other choices of the parameter values in the extreme value distribution, and thus the procedure was repeated with different sets of parameter values in order to make sure that the distributions obtained were independent of these parameter values. The critical values obtained in this manner for the different tests and for different levels and sample sizes are given in Table 3 through Table 9.

The results from the simulation study are based on 200,000 replicates in order to get three significant decimal digits in the obtained results. The results for a selection of combination of sample size and significance levels, are reported in Tables 3-9 as the fraction of rejected cases for each sample size and significance level. A more complete set of tables can be accessed from the authors upon request.

4. Power Results of the Monte Carlo Study and Percentage Points

The results of the power study show that the power of the different tests have almost the same pattern for different sample sizes. These results indicate that:

- B_n has higher power than the other tests for Weibull, log-normal and normal distributions.
- For tests on the level $\alpha = 0.01$ and sample sizes less than 30 the test statistic L_n^+ improves the power compared to L_n. For larger sample sizes this effect diminish. For tests on the level $\alpha = 0.05$ this effect can also be seen for sample size less than 20.
- A_n^2 is slightly more powerful than the other test for Cauchy and log-chi-square alternative distributions.

The power of the Anderson-Darling test, with the parameter estimates proposed here (Equations (5) and (6)) is generally of the same order as when the maximum likelihood estimates were used in substitution for μ and θ (not shown in the tables). The procedure for parameter estimation used here is simple to implement which is an advantage when used in the test statistics. Generally, D_{sp}^+ was more powerful than D_{sp}^* (not shown) for the different alternative distributions. The high performance of this improved test D_{sp}^+ and also of the T_1 test, in many occasions dominates the performance of the A_n^2 test. The quantiles of the test statistic B_n based on 10^6 replicates for $n = 3(1)20(5)60(10)100$ are given in Table 1. Similar quantiles for the statistic T_1 has been given previously (Pirouzi Fard and Holmquist, 2007b).

5. Censored Data

A nice feature of the estimation technique applied in the proposed tests is that they can easily be adapted to situations with censored samples, including both upper and lower censoring; see Castro-Kuriss (2011).

Suppose we have a sample of size n of which the k smallest and l largest observations are censored. We are thus given observations $z_{(k+1)} < z_{(k+2)} < \cdots < z_{(n-l)}$ and that k observations are less than $z_{(k+1)}$ and l observations are larger than $z_{(n-l)}$. Here of course either k or l (or both) may be 0. We may still perform the estimation of the parameters μ and θ using Equations (5) and (6) based of the standard order means $m_i, i = k+1, \ldots, n-l$ arranged in m of size $n-l-k$, variance/covariances $\sigma_{ij}, i = k+1, \ldots, n-l, j = k+1, \ldots, n-l$ arranged in Σ of size $(n-l-k) \times (n-l-k)$ and $z_{(k+1)}, \ldots, z_{(n-l)}$ arranged in ζ of size $n-l-k$.

Table 1. Quantiles of the test statistic B_n based on 10^6 replicates.

	Lower quantiles						
n	0.01	0.025	0.05	0.10	0.15	0.20	0.25
3	0.811	0.817	0.827	0.874	0.867	0.886	0.904
4	0.713	0.741	0.771	0.813	0.845	0.873	0.896
5	0.691	0.728	0.764	0.809	0.840	0.866	0.890
6	0.690	0.729	0.765	0.809	0.841	0.867	0.889
7	0.695	0.733	0.768	0.812	0.844	0.869	0.891
8	0.701	0.739	0.774	0.817	0.847	0.872	0.894
9	0.708	0.745	0.779	0.821	0.850	0.874	0.896
10	0.714	0.751	0.784	0.825	0.854	0.877	0.899
11	0.720	0.757	0.790	0.829	0.858	0.881	0.901
12	0.727	0.763	0.795	0.833	0.861	0.883	0.903
13	0.733	0.767	0.799	0.837	0.864	0.886	0.905
14	0.738	0.773	0.803	0.840	0.867	0.889	0.908
15	0.744	0.777	0.808	0.844	0.870	0.891	0.910
16	0.749	0.781	0.811	0.847	0.872	0.893	0.911
17	0.753	0.786	0.815	0.850	0.875	0.895	0.913
18	0.758	0.789	0.819	0.853	0.877	0.897	0.915
19	0.761	0.793	0.822	0.855	0.879	0.899	0.916
20	0.766	0.796	0.825	0.858	0.881	0.901	0.918
25	0.782	0.811	0.838	0.869	0.891	0.908	0.924
30	0.796	0.823	0.848	0.877	0.898	0.915	0.930
35	0.807	0.833	0.856	0.884	0.904	0.920	0.934
40	0.817	0.841	0.864	0.890	0.909	0.924	0.937
45	0.825	0.848	0.870	0.895	0.913	0.928	0.940
50	0.831	0.854	0.875	0.900	0.917	0.931	0.943
55	0.837	0.860	0.880	0.904	0.920	0.933	0.945
60	0.843	0.865	0.884	0.907	0.923	0.936	0.947
70	0.852	0.873	0.891	0.913	0.928	0.940	0.951
80	0.861	0.880	0.897	0.917	0.932	0.944	0.954
90	0.867	0.886	0.902	0.922	0.935	0.946	0.956
100	0.873	0.890	0.907	0.925	0.938	0.949	0.958

All the test statistics can then be calculated, noting that in any summation or maximization the index set should be $\{k+1,\dots,n-l\}$ instead of $\{1,\dots,n\}$. The critical values for any test statistic, should, however, be chosen based on the uncensored size $n-l-k$.

6. An Example

The following data are the logarithms of the number of million revolutions before failure for each of the 23 ball bearing in life tests: 2.884, 3.365, 3.497, 3.726, 3.741, 3.820, 3.881, 3.948, 3.950, 3.991,4.017, 4.217, 4.229, 4.229, 4.232, 4.432, 4.534, 4.591, 4.655, 4.662, 4.851, 4.852, 5.156. These data were treated by Lieblein and Zelen (1956), who assumed that the original data come from a Weibull distribution. Lawless (1982) showed that the significance level for rejecting the hypothesis that the log failure times have an extreme value distribution is over 0.25. Also, Lawless (1982) determined in this example the maximum likelihood estimates of the extreme value location and scale parameters to be $\hat{\mu} = 4.405$; $\hat{\theta} = 0.476$. With aid of Equations (5) and (6) we estimated μ and

θ, used in Equation (1), for test statistics T_1, A_n^2, D_{ps}^+ and B_n and found $\hat{\mu} = 4.4135$; $\hat{\theta} = 0.4964$. The estimates for μ and θ, using graphical plotting techniques based on plotting position $p_i = [i - 0.5]/n$ for L_n, are $\hat{\mu} = 4.3880$; $\hat{\theta} = 0.4205$. Similarly, we found the estimates for L_n^+ statistic as $\hat{\mu} = 4.3966$; $\hat{\theta} = 0.4285$ based on approximation given in Equation (2). The obtained values from test statistics T_1, A_n^2, D_{sp}^+, L_n, L_n^+, B_n are respectively $0.0590, 0.3401, 0.0838, 0.9326, 0.8796, 0.8743$. Significance levels with these statistics for $n = 23$ are tabulated in Table 2, which are obtained by 10^6 repetitions using Monte Carlo simulation. The significance level from B_n is about 0.15 and those from $T_1, A_n^2, D_{ps}^+, L_n$, L_n^+ are all greater than 0.20. There is consequently no evidence against the hypothesized extreme value distribution from any of these tests.

Table 2. Critical values of test statistics for $n = 23$ at levels $\alpha = 0.01$, 0.05, 0.10, 0.15, 0.20, 0.25 and 0.30.

α	T_1	A_n^2	D_{sp}^+	L_n	L_n^+	B_n
0.30	0.070	0.424	0.088	0.918	0.890	0.937
0.25	0.079	0.455	0.091	0.958	0.927	0.922
0.20	0.090	0.495	0.095	1.008	0.972	0.906
0.15	0.103	0.544	0.100	1.077	1.033	0.887
0.10	0.123	0.613	0.107	1.184	1.126	0.865
0.05	0.161	0.735	0.117	1.428	1.326	0.833
0.01	0.248	1.014	0.138	3.072	2.478	0.777

7. Conclusions

We have investigated several procedures for goodness-of-fit tests in connection with the extreme value distribution. We have shown that the suggested modifications of the estimation procedure of unknown parameters in existing, previously proposed, test statistics can improve the power of these tests. In addition, we have shown that the new test statistics introduced here, based on regression of order statistics on their hypothetical expected values, gives, in comparison, very powerful tests for a wide range of alternative distributions. The relative simplicity of the estimation technique has also been shown to make the tests statistics applicable in situations with censored data.

Appendix

In this section, tables of power and critical values of the different test statistics for various levels of the tests are presented. Weib(a, b) is the Weibull distribution with location parameter a and scale parameter b. The LN(a, b) is the lognormal distribution, for which the logarithms are normal with mean a and variance b, N(a, b). The symbol L$\chi^2(f)$ stands for the distribution for which the anti-log is chi-square with f degrees of freedom, $\chi^2(f)$. The $t(f)$ is the t-distribution with f degrees of freedom, U$(0, 1)$ is the uniform distribution on the unit interval, Ca(a, b) is the Cauchy distribution centred in a and scale b.

Table 3. Power comparison for sample size $n = 10$ at significance level $\alpha = 0.01$.

Alternative Distribution	T_1	A_n^2	D_{sp}^+	L_n	L_n^+	B_n
Weib(5,42)	0.011	0.009	0.010	0.012	0.012	0.011
Weib(20,25)	0.012	0.009	0.010	0.013	0.013	0.013
Weib(2,12)	0.015	0.009	0.012	0.017	0.017	0.017
Weib(2,4)	0.043	0.018	0.025	0.047	0.048	0.047
LN(0.4,0.03)	0.070	0.030	0.040	0.078	0.078	0.077
LN(0.69,0.05)	0.077	0.034	0.043	0.086	0.087	0.084
LN(2,0.1)	0.100	0.045	0.057	0.110	0.111	0.109
LN(0.65,0.27)	0.208	0.107	0.127	0.212	0.214	0.220
L$\chi^2(1)$	0.007	0.016	0.012	0.005	0.005	0.005
L$\chi^2(4)$	0.016	0.009	0.012	0.018	0.018	0.017
t(2)	0.227	0.201	0.198	0.232	0.234	0.219
t(6)	0.101	0.059	0.066	0.116	0.118	0.110
N(0,1)	0.060	0.026	0.035	0.068	0.069	0.066
Logistic(0,1)	0.089	0.048	0.056	0.104	0.106	0.100
U(0,1)	0.040	0.021	0.026	0.027	0.027	0.037
Ca(5,1)	0.429	0.472	0.457	0.357	0.359	0.338
Laplace(1,4)	0.141	0.099	0.105	0.162	0.165	0.150
$\chi^2(1)$	0.804	0.609	0.763	0.698	0.707	0.779
Beta(2,6)	0.198	0.086	0.115	0.185	0.188	0.206
Gamma(4,2)	0.254	0.130	0.157	0.247	0.250	0.264
Critical values:	0.408	0.970	0.172	6.234	3.923	0.714

Table 4. Power comparison for sample size $n = 10$ at significance level $\alpha = 0.05$.

Alternative Distribution	T_1	A_n^2	D_{sp}^+	L_n	L_n^+	B_n
Weib(5,42)	0.053	0.048	0.050	0.057	0.057	0.057
Weib(20,25)	0.054	0.047	0.051	0.061	0.061	0.062
Weib(2,12)	0.064	0.049	0.057	0.078	0.078	0.078
Weib(2,4)	0.135	0.081	0.099	0.170	0.169	0.171
LN(0.04,0.03)	0.189	0.114	0.136	0.232	0.232	0.234
LN(0.69,0.05)	0.204	0.125	0.147	0.249	0.248	0.250
LN(2,0.1)	0.244	0.152	0.175	0.291	0.291	0.294
LN(0.65,0.27)	0.403	0.266	0.299	0.447	0.447	0.455
L$\chi^2(1)$	0.048	0.067	0.053	0.026	0.027	0.027
L$\chi^2(4)$	0.065	0.050	0.057	0.079	0.079	0.079
t(2)	0.363	0.334	0.333	0.368	0.374	0.349
t(6)	0.219	0.157	0.173	0.264	0.262	0.259
N(0,1)	0.171	0.104	0.123	0.212	0.211	0.212
Logistic(0,1)	0.207	0.142	0.160	0.253	0.252	0.249
U(0,1)	0.145	0.105	0.116	0.146	0.154	0.150
Ca(5,1)	0.578	0.607	0.592	0.492	0.530	0.427
Laplace(1,4)	0.268	0.228	0.236	0.320	0.320	0.304
$\chi^2(1)$	0.927	0.818	0.898	0.914	0.918	0.918
Beta(2,6)	0.417	0.253	0.305	0.450	0.452	0.466
Gamma(4,2)	0.469	0.312	0.354	0.506	0.507	0.518
Critical values:	0.262	0.697	0.146	1.794	1.546	0.784

Table 5. Power comparison for sample size $n = 10$ at significance level $\alpha = 0.10$.

Alternative Distribution	T_1	A_n^2	D_{sp}^+	L_n	L_n^+	B_n
Weib(5,42)	0.102	0.098	0.100	0.110	0.109	0.111
Weib(20,25)	0.103	0.096	0.102	0.118	0.116	0.119
Weib(2,12)	0.117	0.100	0.111	0.145	0.141	0.147
Weib(2,4)	0.214	0.151	0.174	0.277	0.269	0.285
LN(0.4,0.03)	0.282	0.197	0.224	0.352	0.343	0.362
LN(0.69,0.05)	0.298	0.210	0.238	0.371	0.363	0.381
LN(2,0.1)	0.345	0.244	0.271	0.419	0.409	0.430
LN(0.65,0.27)	0.515	0.380	0.414	0.585	0.575	0.596
Lχ^2(1)	0.100	0.123	0.104	0.062	0.068	0.058
Lχ^2(4)	0.119	0.101	0.112	0.146	0.143	0.149
t(2)	0.446	0.420	0.420	0.459	0.467	0.428
t(6)	0.304	0.239	0.259	0.364	0.358	0.370
N(0,1)	0.257	0.181	0.207	0.327	0.318	0.335
Logistic(0,1)	0.291	0.225	0.247	0.356	0.349	0.362
U(0,1)	0.242	0.198	0.208	0.294	0.290	0.264
Ca(5,1)	0.656	0.676	0.664	0.618	0.647	0.477
Laplace(1,4)	0.353	0.319	0.330	0.412	0.410	0.406
χ^2(1)	0.961	0.891	0.940	0.965	0.963	0.958
Beta(2,6)	0.545	0.377	0.432	0.609	0.598	0.618
Gamma(4,2)	0.585	0.434	0.477	0.646	0.637	0.657
Critical values:	0.203	0.582	0.132	1.377	1.254	0.825

Table 6. Power comparison for sample size $n = 20$ at significance level $\alpha = 0.01$.

Alternative Distribution	T_1	A_n^2	D_{sp}^+	L_n	L_n^+	B_n
Weib(5,42)	0.010	0.009	0.011	0.013	0.013	0.013
Weib(20,25)	0.011	0.010	0.013	0.016	0.016	0.016
Weib(2,12)	0.016	0.011	0.018	0.025	0.025	0.025
Weib(2,4)	0.086	0.047	0.075	0.120	0.123	0.125
LN(0.4,0.03)	0.168	0.100	0.138	0.216	0.220	0.225
LN(0.69,0.05)	0.191	0.115	0.155	0.239	0.243	0.252
LN(2,0.1)	0.258	0.161	0.211	0.305	0.310	0.328
LN(0.65,0.27)	0.529	0.369	0.465	0.535	0.544	0.596
Lχ^2(1)	0.015	0.023	0.010	0.002	0.002	0.002
Lχ^2(4)	0.018	0.012	0.019	0.027	0.028	0.026
t(2)	0.434	0.470	0.415	0.456	0.461	0.405
t(6)	0.220	0.182	0.182	0.295	0.298	0.274
N(0,1)	0.136	0.081	0.113	0.183	0.186	0.188
Logistic(0,1)	0.199	0.157	0.163	0.274	0.277	0.257
U(0,1)	0.087	0.071	0.133	0.031	0.037	0.078
Ca(5,1)	0.722	0.813	0.756	0.556	0.596	0.488
Laplace(1,4)	0.281	0.302	0.253	0.378	0.383	0.334
χ^2(1)	0.996	0.976	0.999	0.979	0.983	0.994
Beta(2,6)	0.556	0.332	0.547	0.503	0.517	0.609
Gamma(4,2)	0.632	0.444	0.590	0.605	0.616	0.685
Critical values:	0.271	1.010	0.144	3.618	2.760	0.766

Table 7. Power comparison for sample size $n = 20$ at significance level $\alpha = 0.10$.

Alternative Distribution	T_1	A_n^2	D_{sp}^+	L_n	L_n^+	B_n
Weib(5,42)	0.101	0.099	0.105	0.117	0.114	0.121
Weib(20,25)	0.102	0.098	0.108	0.128	0.124	0.135
Weib(2,12)	0.123	0.110	0.129	0.175	0.168	0.186
Weib(2,4)	0.353	0.252	0.307	0.458	0.443	0.491
LN(0.4,0.03)	0.483	0.370	0.420	0.586	0.572	0.617
LN(0.69,0.05)	0.518	0.397	0.449	0.619	0.605	0.650
LN(2,0.1)	0.601	0.468	0.527	0.693	0.679	0.723
LN(0.65,0.27)	0.830	0.705	0.769	0.877	0.869	0.891
Lχ^2(1)	0.119	0.142	0.104	0.057	0.067	0.037
Lχ^2(4)	0.132	0.115	0.134	0.180	0.174	0.193
t(2)	0.663	0.686	0.657	0.696	0.706	0.588
t(6)	0.483	0.436	0.449	0.576	0.568	0.585
N(0,1)	0.432	0.327	0.376	0.538	0.524	0.568
Logistic(0,1)	0.467	0.415	0.430	0.568	0.559	0.583
U(0,1)	0.455	0.381	0.469	0.550	0.534	0.426
Ca(5,1)	0.879	0.916	0.893	0.885	0.898	0.579
Laplace(1,4)	0.538	0.570	0.542	0.635	0.633	0.611
χ^2(1)	1.000	0.998	1.000	1.000	1.000	1.000
Beta(2,6)	0.883	0.721	0.850	0.912	0.904	0.918
Gamma(4,2)	0.897	0.775	0.856	0.925	0.919	0.933
Critical values:	0.135	0.611	0.111	1.226	1.149	0.858

Table 8. Power comparison for sample size $n = 50$ at significance level $\alpha = 0.10$.

Alternative Distribution	T_1	A_n^2	D_{sp}^+	L_n	L_n^+	B_n
Weib(5,42)	0.097	0.101	0.112	0.122	0.118	0.140
Weib(20,25)	0.104	0.106	0.124	0.144	0.139	0.173
Weib(2,12)	0.157	0.144	0.181	0.235	0.225	0.290
Weib(2,4)	0.710	0.535	0.664	0.789	0.775	0.855
LN(0.4,0.03)	0.850	0.742	0.806	0.901	0.894	0.938
LN(0.69,0.05)	0.882	0.781	0.842	0.924	0.918	0.954
LN(2,0.1)	0.940	0.862	0.913	0.963	0.959	0.979
LN(0.65,0.27)	0.997	0.983	0.995	0.998	0.998	0.999
Lχ^2(1)	0.178	0.201	0.139	0.100	0.114	0.015
Lχ^2(4)	0.168	0.155	0.187	0.243	0.234	0.294
t(2)	0.917	0.954	0.921	0.947	0.951	0.764
t(6)	0.779	0.785	0.760	0.850	0.847	0.865
N(0,1)	0.789	0.673	0.743	0.856	0.847	0.905
Logistic(0,1)	0.775	0.771	0.750	0.854	0.849	0.884
U(0,1)	0.942	0.852	0.971	0.966	0.958	0.784
Ca(5,1)	0.995	0.999	0.996	0.998	0.998	0.672
Laplace(1,4)	0.823	0.906	0.843	0.905	0.906	0.874
χ^2(1)	1.000	1.000	1.000	1.000	1.000	1.000
Beta(2,6)	1.000	0.992	1.000	1.000	1.000	1.000
Gamma(4,2)	1.000	0.995	1.000	1.000	1.000	1.000
Critical values:	0.073	0.627	0.082	1.044	1.017	0.900

Table 9. Power comparison for sample size $n = 100$ at significance level $\alpha = 0.05$.

Alternative Distribution	T_1	A_n^2	D_{sp}^+	L_n	L_n^+	B_n
Weib(5,42)	0.045	0.053	0.066	0.072	0.069	0.088
Weib(20,25)	0.053	0.062	0.084	0.097	0.093	0.126
Weib(2,12)	0.123	0.118	0.169	0.218	0.209	0.287
Weib(2,4)	0.898	0.739	0.888	0.936	0.930	0.969
LN(0.4,0.03)	0.969	0.925	0.956	0.985	0.983	0.994
LN(0.69,0.05)	0.982	0.949	0.973	0.991	0.991	0.997
LN(2,0.1)	0.996	0.982	0.993	0.998	0.998	0.999
LN(0.65,0.27)	1.000	1.000	1.000	1.000	1.000	1.000
Lχ^2(1)	0.163	0.202	0.118	0.098	0.110	0.002
Lχ^2(4)	0.127	0.133	0.169	0.218	0.210	0.283
t(2)	0.984	0.997	0.987	0.995	0.996	0.838
t(6)	0.909	0.942	0.898	0.956	0.956	0.957
N(0,1)	0.937	0.876	0.917	0.967	0.964	0.985
Logistic(0,1)	0.911	0.934	0.894	0.959	0.960	0.971
U(0,1)	0.999	0.991	1.000	1.000	1.000	0.926
Ca(5,1)	1.000	1.000	1.000	1.000	1.000	0.717
Laplace(1,4)	0.939	0.990	0.952	0.984	0.984	0.965
χ^2(1)	1.000	1.000	1.000	1.000	1.000	1.000
Beta(2,6)	1.000	1.000	1.000	1.000	1.000	1.000
Gamma(4,2)	1.000	1.000	1.000	1.000	1.000	1.000
Critical values:	0.057	0.753	0.070	1.079	1.061	0.907

References

Aitken, A.C., 1935. On least squares and the linear combination of observations. Proceedings of the Royal Society of Edinburgh, A, 55, 42–48.

Blom, G., 1958. Statistical Estimates and Transformed Beta-Variables. Almqvist and Wiksell, Uppsala.

Balakrishnan, N., Chan, P.S., 1992. Order statistics from extreme value distribution I: tables of means, variances and covariances. Communications in Statistics - Simulation and Computation, 21, 1199–1217.

Balakrishnan, N., Rao, C.R., 1998. Handbook of Statistics 17. Elsevier, Amsterdam.

Castro-Kuriss, C., 2011. On a goodness-of-fit test for censored data from a location-scale distribution with applications. Chilean Journal of Statistics, 2, 115–136.

Coles, S.G., 1989. On goodness-of-fit tests for the two-parameter Weibull distribution derived from stabilized probability plot. Biometrika, 76, 593–598.

D'Agostino, R.B., Stephens, M.A., 1986. Goodness-of-fit Techniques. Marcel Dekker, NY.

Downton, F., 1966. Linear estimations of parameters in the extreme value distribution. Technometrics, 8, 3–17.

Kimber, A.C., 1985. Tests for exponential, Weibull and Gumbel distributions based on the stabilized probability plot. Biometrika, 72, 661–663.

Kinnison, R., 1989. Correlation coefficient goodness-of-fit test for the extreme-value distribution. The American Statistician, 43, 98–100.

Lawless, J.F., 1982. Statistical Models and Methods for Lifetime Data. Wiley, NY.

Liao, M., Shimokawa, T., 1999. A new goodness-of-fit test for type-I extreme-value and 2-parameter Weibull distributions with estimated parameters. Journal of Statistical Computation and Simulation, 64, 23–48.

Lieblein, J., Salzer, H.E., 1957. Table of the first moment of ranked extremes. Journal of Research of the National Bureau of Standards, 59, 203–206.

Lieblein, J., Zelen, M., 1956. Statistical investigation of the fatigue life of deep groove ball bearing. Journal of Research of the National Bureau of Standards, 57, 273–316.

Littel, R.C., McClave, J.T., Offen, W.W., 1979. Goodness of fit tests for two-parameter Weibull distribution. Communications in Statistics - Simulation and Computation, 8, 257–269.

Lloyd, E.H., 1952. Least-squares estimation of location and scale parameters using order statistics. Biometrika, 39, 88–95.

Michael, J.R., 1983. The stabilized probability plot. Biometrika, 70, 11–17.

Pirouzi Fard, M.N., Holmquist, B., 2007a. First moment approximations for order statistics from the extreme value distribution. Statistical Methodology, 4, 196–203.

Pirouzi Fard, M.N., Holmquist, B., 2007b. Goodness of fit tests for the extreme value distribution based on regression, EDF and stabilized probability plot. Working Paper. Department of Statistics, Lund University, Sweden.

Pirouzi Fard, M.N., Holmquist, B., 2008. Approximations of variances and covariances for order statistics from the standard extreme value distribution. Communications in Statistics - Simulation and Computation, 37, 1500–1506.

White, J.S., 1967. The moments of log-Weibull order statistics. Research Publication GMR-717. Research Laboratories, General Motors Corp., Warren, Michigan.

White, J.S., 1969. The moments of log-Weibull order statistics. Technometrics, 11, 373–386.