A MAXIMUM ENTROPY TYPE TEST OF FIT

Sangyeol Lee\textsuperscript{1,4}, Ilia Vonta\textsuperscript{2} and Alex Karagrigoriou\textsuperscript{3}

Seoul National University, National Technical University of Athens and University of Cyprus

Abstract

In this paper, we propose a test of fit based on maximum entropy. The asymptotic distribution of the proposed test statistic is established and a corrected form for small and medium sample sizes is furnished. The performance of the test is investigated through extensive Monte Carlo simulations. Real examples are also presented and analyzed.

Keywords: Maximum entropy measure, Goodness of fit test, Reliability, Brownian bridge

1 Introduction

The maximum entropy principle (Jaynes, 1957) is a criterion for selecting a priori probabilities. For a given amount of information, the probability distribution that best describes our knowledge is the one that maximizes the Shannon entropy subject to a given evidence as constraints. That is to say, when characterizing some unknown events with a statistical model, we should always choose the one that has Maximum Entropy.

Maximum Entropy Modelling has been successfully applied to Computer Vision, Spatial Physics, Natural Language Processing and many other fields.

\textsuperscript{1}Department of Statistics, Seoul National University, Seoul, Korea. Email: sylee@stats.snu.ac.kr.
\textsuperscript{2}Department of Mathematics, National Technical University of Athens, Athens, Greece. Email: vonta@math.ntua.gr.
\textsuperscript{3}Department of Mathematics and Statistics, University of Cyprus, Nicosia, Cyprus. Email: alex@ucy.ac.cy.
\textsuperscript{4}Author to whom correspondence should be addressed.
Forte (1984) provides the Boltzmann-Shannon entropy as
\[ H(f) = -\int_{-\infty}^{\infty} f(x) \log(f(x)) \, dx \] (1)
which measures the amount of uncertainty one has about the value \( x \) of a real-valued random variable \( X \) given its probability density function \( f(x) \).

For continuous distributions, the simple definition of Shannon entropy ceases to be so useful. Instead, Jaynes (1963, 1968, 2003) gave the following formula, which is closely related to the relative entropy
\[ H_c = -\int p(x) \log \left( \frac{p(x)}{q(x)} \right) \, dx \] (2)
where \( q(x) \), which Jaynes called the “invariant measure”, is proportional to the limiting density of discrete points. In fact, \( H_c \) is equal to the negative relative entropy also known as the Kullback-Leibler divergence of \( q \) from \( p \).
The inference principle of minimizing this, due to Kullback, is known as the Principle of Minimum Discrimination Information. The Shannon entropy for a discrete random variable is not considered to be the discrete analogue of (1). Instead, Forte and Hughes (1988) proposed the function \(-\sum p_i \log(p_i/(x_i - x_{i-1}))\) as a good candidate for a discrete analogue of (1) since for a variable defined in \([a, b]\)
\[ \lim_{\max_i |x_i - x_{i-1}| \to 0} -\sum_{i=1}^{n} p_i \log(p_i/(x_i - x_{i-1})) = H(f), \] (3)
where \( p_i = P[x_{i-1} < X \leq x_i] = \int_{x_{i-1}}^{x_i} f(x) \, dx, \) \( i = 1, \ldots, n - 1 \) and \( a = x_0 < \ldots < x_n = b \).

Goodness-of-fit (gof) tests measure the degree of agreement between the distribution of an observed random sample and a theoretical statistical distribution. The problem of goodness-of-fit to any distribution on the real line, is frequently treated by partitioning the range of data in \( m \) disjoint intervals. In all cases, a test statistic is compared against a known critical value to accept or reject the hypothesis that the sample is from the postulated distribution. Over the years, numerous nonparametric gof methods including the chi-squared test and various empirical distribution function (edf) tests
(D’Agostino and Stephens, 1986), have been developed. At the same time measures of entropy, divergence and information like the ones given in (1) – (3), are quite popular in goodness of fit tests. Over the years several measures like (2) and (3), have been suggested to reflect the fact that some probability distributions are closer together than others. Many of the currently used tests, such as the likelihood ratio, the chi-squared, the score and Wald tests are defined in terms of appropriate measures. Other such tests include those based on entropy (Vasicek, 1976; Dudewicz and van der Meulen, 1981; Gokhale, 1983) and the $\phi$–family of Csiszar measures (Csiszar, 1963; Read and Cressie, 1988, Zografos et al, 1990; Pardo, 2006) and its generalizations (Mattheou and Karagrigoriou, 2010).

In this paper we propose an alternative information measure which generalizes (3) and is used for goodness of fit tests. In section 2.1 we provide the test statistic and its asymptotic distribution under the null hypothesis. In section 2.2 we provide a small sample modification and in section 3 we perform a simulation study in order to explore the capabilities of the proposed test statistic. Real examples are also presented and analyzed.

2 The Maximum Entropy Test

2.1 Development and asymptotics of the maximum entropy test

Let $Y_i, i = 1, \ldots, n$ be a random sample from a distribution with unknown cumulative distribution function $F$ and consider the following test of fit:

$$H_0 : F = F_0 \quad \text{vs.} \quad F \neq F_0.$$  

For continuous distributions, we propose the following generalization of Forte and Hughes (1988) entropy:

$$S^w(F) = - \sum_{i=1}^{m} w_i (F(s_i) - F(s_{i-1})) \log \left( \frac{F(s_i) - F(s_{i-1})}{s_i - s_{i-1}} \right),$$  \hspace{1cm} (4)

where the $w$‘s are appropriate weight functions with $0 \leq w_i \leq 1$ and $\sum_{i=1}^{m} w_i = 1$, $m$ is the number of disjoint intervals for partitioning the data range, and $-\infty < a \leq s_1 \leq \ldots \leq s_m \leq b < \infty$ are preassigned partition points. In the
case that we have two distributions, the above can be extended to the Kullback-Leibler distance. For a properly selected constant \(c\), the null hypothesis will be rejected if \(|S^w(F_n) - S^w(F_0)| \geq c\) or, even more stringently, if \(\sup_w |S^w(F_n) - S^w(F_0)| \geq c\), where \(F_n\) is the empirical distribution based on the sample, namely,

\[
F_n(x) = n^{-1} \sum_{i=1}^{n} I(Y_i \leq x).
\]

The proposed test statistic is justified by the fact that it is closely related to the entropy measure \(H(f)\) given in (1). Indeed, observe that when \(m \to \infty\) and \(\max_{1 \leq i \leq m} |s_i - s_{i-1}| \to 0\), the unweighted form of (4) takes the form

\[
S_{\text{max}}(F) = -\sum_{i=1}^{m} \left( \frac{F(s_i) - F(s_{i-1})}{s_i - s_{i-1}} \right) (s_i - s_{i-1}) \log \left( \frac{F(s_i) - F(s_{i-1})}{s_i - s_{i-1}} \right)
\approx -\sum_{i=1}^{m} f(s_i) (s_i - s_{i-1}) \log f(s_i)
\to -\int_{-\infty}^{\infty} f(x) \log f(x) dx = -E_f \log(f(X)) \equiv H(f),
\]

where \(f\) is the probability density function.

Observe that if \(F_0\) is the uniform distribution in \([0, 1]\), then \(S^w(F_0) = 0\). Note that without loss of generality we can concentrate on the Uniform distribution on \([0, 1]\) so that the hypothesis problem becomes

\[
H_0 : F = F_0 \equiv U[0, 1] \text{ vs. } H_1 : F \neq F_0 \equiv U[0, 1].
\]

Indeed, with the use of the probability integral transform we can provide the basis for testing whether the \(Y_i\)'s can reasonably be modeled as arising from any specified continuous distribution \(F_0\). Specifically, the probability integral transform can be applied to construct the equivalent set of values \(F_0(Y_i) = U_i\), and a test is then made of whether a uniform distribution is appropriate for the \(U_i\)'s. Note that this transformation is applied for the well known P-P plots and the Kolmogorov-Smirnov tests.

Remark 1. The role of weights could be vital especially before the implementation of the uniform transformation. For instance, when one deals with a heavy-tailed alternative, he/she may place more weight on the tail part, as in
the Anderson-Darling test. In our case though, the data is transformed into uniform r.v.s with \( s_i = i/m \). This approach is chosen to ease the importance of the choice of the weights. In addition, we choose to take the supremum over the weights in order to cope with any possibilities characterized by specific alternatives (see Theorem and equation (11) below). As a result, we can overcome the difficulty of choosing optimal weights irrespectively of their existence or not. In conclusion, the weight is proposed for the purpose of introducing the maximum entropy test, as a motive. The role of weight may be crucial, but a choice of specific weights is avoided by implementing into our method the uniform transformation and taking the supremum over the weights.

Although the simple null hypothesis appears frequently in practice, it is common to test the composite null hypothesis that the unknown distribution belongs to a parametric family \( \{F_{\theta}\}_{\theta \in \Theta} \), where \( \Theta \) is an open subset in \( \mathbb{R}^k \). In this case we can again consider a partition of the original sample space with \( m \) disjoint intervals. Observe though that in this case, the probability integral transformation depends on the unknown \( k \)-dimensional parameter \( \theta \). Indeed, in this case, a consistent estimator \( \hat{\theta} \) is required so that if the null hypothesis is \( H_0 : F = F_{\theta} \) then the probability integral transformation is applied for the construction of the values \( F_{\theta}(Y_i) = U_i \). In this case, the limiting distribution, given in the theorem below, can be affected by the estimation of \( \theta \). The effect, though, may diminish when \( m \) is large and \( \max_i(s_i - s_{i-1}) \) is small, as in the case of the chi-square test.

In regard to the estimating method applied for obtaining the estimator of \( \theta \), the traditional maximum likelihood estimator (MLE), under the null distribution, can be evaluated and implemented. Note though that one may alternatively consider a wider class of estimators, known as \( \Phi \)-divergence estimators. More specifically, let the partition \( \{E_i\}_{i=1,\ldots,m} \) of the original sample space. Then, the minimum \( \Phi \)-divergence estimator of \( \theta \) is any \( \hat{\theta}_{\Phi} \in \Theta \) satisfying

\[
d_{a}(\hat{\theta}_{\Phi}) = \min_{\theta \in \Theta} d_{a}(\theta) = \min_{\theta \in \Theta} \sum_{i=1}^{m} p_{\theta_{0}}(\theta) \Phi \left( \frac{\hat{p}_i}{p_{\theta_{0}}(\theta)} \right), \quad \Phi \in \Phi^a, \ a > 0 \tag{6}
\]

with

\[
p_{\theta_{0}}(\theta) = \int_{E_i} dF_{\theta}, \quad i = 1, \ldots, m,
\]
the MLE of the probability of the \( i \)th partition, and \( \Phi^* \) the class of all convex functions \( \Phi \) on \([0, \infty)\) such that \( \Phi(1) = \Phi'(1) = 0 \) and \( \Phi''(1) \neq 0 \). We also assume the conventions \( \Phi(0/0) = 0 \) and \( \Phi(u/0) = \lim_{u \to \infty} \Phi(u)/u, \ u > 0 \).

Obviously, the resulting estimator depends on the \( \Phi \)-function chosen. Observe that for \( \Phi \) having the special form

\[
\Phi_a(u) = u^{1+a} - \left(1 + \frac{1}{a}\right) u^a + \frac{1}{a}, \ a > 0
\]  

or \( \Phi_a^1(u) = \Phi_a(u)/(1 + a) \) and for \( a \to 0 \) the resulting estimator is the usual maximum likelihood estimator, for the grouped data. Note that for \( \Phi \) as in (7), the measure (6) reduces to the BHHJ measure of divergence of Basu et al. (1998) which was proposed for the development of a minimum divergence estimating method for robust parameter estimation. Basu et al. (1998) established that an ideal range for the index \( a \) in (6) is the interval \([0, 1]\) where the robust features of the estimator are better preserved.

It should be pointed out, that as long as the original data are available, it is preferable to rely on the maximum likelihood estimating procedure which provides efficient estimators which are usually simpler and easier to obtain. Furthermore, the associated tests are more powerful than those based on grouped data. It is for this reason we have chosen to use the MLE procedure for the analysis of the real examples in Section 3.3.

The theorem below provides the asymptotic distribution of the proposed test statistic.

**Theorem.** Let \( U_1, \ldots, U_n \) be a random sample from a continuous distribution with cumulative distribution function \( F \). Under \( H_0 \) given in (5), as \( n \to \infty \), we have

\[
\sqrt{n} \sup_{\{w \in W\}} |S^w(F_n)| \xrightarrow{d} \sup_{\{w \in W\}} \left| \sum_{i=1}^m w_i (B(s_i) - B(s_{i-1})) \right|,
\]

where \( B(s) \) is the Brownian bridge on \([0,1]\), \( W \) denotes the space of bounded weight functions \( w_i : [0, 1] \to [0, 1] \) with \( \sum_{i=1}^m w_i = 1 \), and \( 0 = s_0 \leq s_1 \leq \ldots \leq s_m = 1 \).
Proof. Let us consider the quantity
\[
S^w(F_n) = - \sum_{i=1}^m w_i (F_n(s_i) - F_n(s_{i-1})) \cdot \log \left( \frac{F_n(s_i) - F_n(s_{i-1})}{s_i - s_{i-1}} - 1 + 1 \right).
\] (9)

Then, by using the expansion \( \log(1 + x) \approx x \) for small \( x \) (in fact, \( |\log(1 + x) - x| \leq x^2 \) for \( |x| \leq 1/2 \)), we have
\[
S^w(F_n) \approx - \sum_{i=1}^m w_i \left( \frac{F_n(s_i) - F_n(s_{i-1})}{s_i - s_{i-1}} \right) \cdot \left[ \left( \frac{1}{n} \sum_{i=1}^n I(U_i \leq s_i) - s_i \right) - \left( \frac{1}{n} \sum_{i=1}^n I(U_i \leq s_{i-1}) - s_{i-1} \right) \right]
\times \left( \frac{1}{n} \sum_{i=1}^m w_i \left( \frac{F_n(s_i) - F_n(s_{i-1})}{s_i - s_{i-1}} \right) \left( E_n(s_i) - E_n(s_{i-1}) \right) \right)
\]
where \( E_n(s) = n^{-1/2} \sum_{i=1}^n \{I(U_i \leq s) - s\}, \ 0 \leq s \leq 1. \)

Observe that under the null hypothesis we have \( F_n(s) \to s \) a.s. as \( n \to \infty \).
Furthermore, observe that for \( n \) tending to \( \infty \),
\[
E_n(s) = \sqrt{n} (F_n(s) - s) \xrightarrow{d} B(s).
\]

Therefore, for each \( w \),
\[
\sqrt{n} \ |S^w(F_n)| \xrightarrow{d} \left| \sum_{i=1}^m w_i (B(s_i) - B(s_{i-1})) \right|.
\]

Further, it can be easily seen that
\[
\sqrt{n} \ \sup_{\{w \in W\}} |S^w(F_n)| \xrightarrow{d} \sup_{\{w \in W\}} \left| \sum_{i=1}^m w_i (B(s_i) - B(s_{i-1})) \right|
\]
(van der Vaart 1998, chap. 19). This completes the proof. \( \square \)
In order to implement our test in practice, we consider $w_{l}^{(i)}$, $l = 1, \ldots, L$, independent and identically distributed random variables from $U[0, 1]$, which are also independent from $U_{i} \sim U[0, 1]$, where $L$ is a fixed positive integer. Then, if we put $w_{li} = w_{l}^{(i)}$, we have that as $L \to \infty$,

$$
\max_{1 \leq l \leq L} \left| \sum_{i=1}^{m} w_{li} (B(s_{i}) - B(s_{i-1})) \right| \overset{d}{\to} \sup_{\{w \in W\}} \left| \sum_{i=1}^{m} w_{i} (B(s_{i}) - B(s_{i-1})) \right|. \tag{10}
$$

Subsequently, by taking $s_{i} = i/m$, $i = 1, \ldots, m$ for convenience, we can use as the maximum entropy test statistic the quantity

$$
S_{\text{max}}^{w} = \max_{1 \leq l \leq L} \left| \sum_{i=1}^{m} w_{li} \left( E_{n} \left( \frac{i}{m} \right) - E_{n} \left( \frac{i-1}{m} \right) \right) \right| \tag{11}
$$

$$
\approx \sup_{\{w \in W\}} \left| \sum_{i=1}^{m} w_{i} \left( B \left( \frac{i}{m} \right) - B \left( \frac{i-1}{m} \right) \right) \right|.
$$

**Remark 2.** The argument in (10) is true since

$$
P(\{(w_{11}, \ldots, w_{1m}) : l \geq 1\} = W) = 1;$$

otherwise, there exists an open ball $V$ in the hyperplane $W$ such that $p := P((w_{11}, \ldots, w_{1m}) \in V^{c})$ for all $l \geq 1 > 0$, which is in fact impossible since $p \leq \rho^{L}$ for any $L$ with $\rho = P((w_{11}, \ldots, w_{1m}) \in V^{c}) < 1$ (or equivalently $P((w_{11}, \ldots, w_{1m}) \in V) > 0$). This indicates

$$
\sup_{l \geq 1} \left| \sum_{i=1}^{m} w_{li} (B(s_{i}) - B(s_{i-1})) \right| = \sup_{\{w \in W\}} \left| \sum_{i=1}^{m} w_{i} (B(s_{i}) - B(s_{i-1})) \right| \text{ a.s.,}
$$

which immediately implies (10).

**Remark 3.** In regard to the choice of the value of $L$ for practical purposes, we recommend the use of $L = 1000$. We have run a number of simulations
with the uniform and a number of non-uniform distributions (Weibull, Beta, Gamma, Inverse Gaussian) with values of $L$ ranging from as low as 100 to as high as 10000. Both the size and the power of the test have been evaluated with various alternatives. The results clearly show that a value of $L$ between 500 and 2000 is sufficient for both the size and the power of the test in all cases, to be stabilized. As a result, we recommend the choice of $L = 1000$ for all practical purposes. The results presented in the simulation section are based on this choice of $L$.

2.2 Modifications for Small Sample Sizes

In this subsection, we attempt a modification of the proposed maximum entropy test so that its performance will be satisfactory even for small sample cases. Due to its asymptotic nature, as expected, the power and the size of the maximum entropy test are quite satisfactory for medium to large sample sizes but fail to perform as well for small sample sizes (results not shown). A similar situation occurs in various goodness of fit tests including the family of Cramér-von Mises (CvM) family of test statistics which includes among others, the popular Anderson Darling (AD) test (Anderson and Darling, 1954). Recall that the Cramér-von-Mises family, which is given by

$$Q = \int_{-\infty}^{\infty} [F_n(s) - F_0(s)]^2 \Psi(s) dF_0(s)$$

reduces to the AD test for $\Psi(s) = F_0(s)/(1 - F_0(s))$. The same family for appropriate functions $\Psi(\cdot)$ includes the Watson test (Watson, 1961) and the CvM test (Cramér, 1928; von Mises, 1931). Although in some cases, like the AD test, the asymptotic distribution of the test statistic can be established (see e.g. O’Reilly and Rueda, 1992) and used for large sample sizes, in most cases, there are modifications for small and medium sample sizes. All the modified goodness of fit tests including the ones corresponding to the above mentioned tests rely on the same test statistic as the original test but advocate the use of tables or formulas for critical values which depend on the parameters of the distribution under investigation and/or the sample size (see Edgeman et. al, 1988; Pavur et. al, 1992, Gunes et. al, 1997, Koutrouvelis et. al, 2010; Koutrouvelis and Karagrigoriou, 2010).

In this section we employ large Monte Carlo samples to develop modified critical values for the maximum entropy test obtained in the previous section.
for small to medium sample sizes. Our aim is the development of functional relationships to eliminate the need for extensive critical value tables. For the maximum entropy test, the number \( m \) of subintervals used plays a very crucial role in combination with the size of the sample. Naturally, for small sample sizes one should avoid using a large number of subintervals due to the possibility of zero observations in at least some of them. As a result, in our Monte Carlo study we focus on the sample size in combination with the value of \( m \), with \( L \) in (11) chosen to be equal to the recommended value of 1000 (see Remark 3). All programs are written in R. The sample sizes considered are 10, 20, 30, 50 and 100. The values of \( m \) used are 3, 4, 5, and 10 (for the choice of \( m \) see the discussion below). For each combination of \( n \) and \( m \), the following steps are performed:

- A sample of size \( n \) from \( U[0, 1] \) is selected.
- The maximum entropy test statistic given in (11) with \( L = 1000 \), is evaluated.
- The previous two steps are repeated \( M = 1000 \) times.
- The resulting 1000 statistics are ordered and the 90th, 95th, and 99th percentiles of the ordered sample which approximate the critical values for respective significance levels of 0.10, 0.05, and 0.01, are identified.

Correlation analysis using stepwise linear regression reveals strong relationships between critical values and simple functions of \( n \) and \( m \). Additional interaction terms were considered, but the marginal contributions did not warrant their inclusion. Table 1 displays the resulting regression formulas and associated adjusted coefficient of multiple determination \( R^2 \) for each significance level. For practical purposes, the critical values obtained from the asymptotic distribution given in Theorem 1 are recommended to be used in applications with \( n > 100 \) and the modified critical values from Table 1 in applications with \( n \leq 100 \).

We close this section with a discussion about the number of classes \( m \) required in tests of fit. Even for the case of the \( \chi^2 \), in spite of its wide use, it is still not clear how many partition points should be used and how the class intervals should be formed. However, it is generally recommended that researchers use equiprobable partitions (Stuart, Ord and Arnold, 1999).
The problem of determining the optimum number of classes has a long history. Mann and Wald (1942), Cochran (1952), and Dahiya and Gurland (1973) proposed similar techniques for the case of the chi-square test. More recently, Harrison (1985) proposed a generalization of the Mann and Wald method. These methods though arrive at diverse conclusions. Although Mann and Wald’s recommendation for at most 24 classes is based on asymptotic theory, they suggest that the results hold approximately for sample sizes as low as 200 and may be true for considerably smaller samples. Harrison (1985) found that at least 23 classes can ensure a power equal to 0.5. However, various numerical studies have presented empirical evidence to show that such values of $m$ are too large, resulting in loss of power in many situations. Williams (1950) indicates that the value of $m$ may be halved for practical purposes, without relevant loss of power. See also Dahiya and Gurland (1973), who suggest values of $m$ between 3 and 12 for several different alternatives in testing normality, for sample sizes of $n = 50$ and 100. Finally, Cochran suggested a fixed number of classes equal to 5. In almost all these cases, the optimum number of classes depends on various equally important factors which are the type of the alternative hypothesis under consideration, the ”distance” between the alternative hypothesis and the data, the minimum power to be achieved, the significance level of the test, and the size of the sample.

One should be always aware of the danger associated with the use of too many classes in cases where the observations are spread too thinly over the data range. Another important factor is the gain anticipated when the number of classes increases. The power of the test together with the size are the key

### Table 1: Formulas for critical values of the maximum entropy test

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Intercept</th>
<th>$n$</th>
<th>$m$</th>
<th>$\sqrt{n}$</th>
<th>$\sqrt{m}$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>90th</td>
<td>2.450566</td>
<td>0.003745</td>
<td>0.078023</td>
<td>-0.080760</td>
<td>-0.779092</td>
<td>98.71%</td>
</tr>
<tr>
<td>95th</td>
<td>2.593938</td>
<td>0.002283</td>
<td>0.001519</td>
<td>-0.234312*</td>
<td>-0.610672*</td>
<td>98.14%</td>
</tr>
<tr>
<td>99th</td>
<td>3.746811</td>
<td>0.009262</td>
<td>0.056380</td>
<td>-0.181769</td>
<td>-0.942940</td>
<td>95.61%</td>
</tr>
</tbody>
</table>

*: log(.) instead of $\sqrt{.}$
factors in deciding if the increase of the number of classes is useful. It should be
noted that even in cases, where the number of classes is allowed to increase as
the sample size increases, specific assumptions should be imposed in order to
secure satisfactory asymptotic results. One such assumption is the well known
sparseness assumption according to which the limit as \( n \to \infty \), of \( n/M \), where
\( M \) depends on \( n \), is finite. Results of this type have been reported by Holst
(1972), Dale (1986), and Read and Cressie (1988). In our analysis where we
dealt with samples with less than 100 observations, we have considered 3-20
classes and concluded that for classes at most equal to 10, the results are quite
satisfactory as compared with other tests commonly used in the literature.

3 Simulation Study and Real Examples

In this section, we report simulation results for the maximum entropy test for
small, medium, and large sample sizes. For small and medium samples we
make use of the modified critical values of Table 1, while for large samples we
use the critical values according to Theorem 1 with \( L = 1000 \). For this task,
we use samples of size 10, 20, 30, 50, 100, 300, 500, and 1000 and values of
\( m \) equal to 3, 4, 5, and 10. The Monte Carlo simulations are based on 10000
repetitions. Two real examples are presented in Section 3.3.

3.1 The Uniform Null

In order to investigate the performance of the test, we first consider the null
hypothesis

\[ H_0 : F = F_0 = \text{Beta}(1,1). \]

For alternative models we consider Beta distributions with parameters \((\alpha, \beta) = (2,5), (2,6), (0.5, 0.5), (2,2), \) and \((1,8)\). Further, we consider the mixture model:

\[ H_1 : U \sim b \cdot U[0,1/3] + (1-b) \cdot U[2/3,1], \quad b \sim \text{Binom}(1,1/3). \]  

(12)

The first part/row of Table 2 provides the size of the proposed test while
the rest provides the power of the test for the alternative models mentioned
above. The last part/row (mix) refers to the mixture model given in (12).
The results reveal that the size of the test is very close to the nominal level.
On the other hand, the power study reveals excellent discriminatory ability of the maximum entropy test against Beta (2,5), Beta (2,6) and Beta (1,8) moderate power against the mixture model (12), and, as expected, relatively poor discriminatory ability against beta distributions like Beta (0.5, 0.5) and Beta (2,2) which can be considered to be closer than the others to the null Beta(1,1) model.

3.2 Non-uniform null models

We extend now, our investigation by considering various continuous distributions and applying the probability integral transformation. In this case we have:

\[ H_0 : Y \sim F = F_0 \iff H_0 : U = F_0(Y) \sim U[0,1] \]

We choose to focus on distributions like the exponential, lognormal, Gamma, Inverse Gaussian (IG) and Weibull that frequently appear in biomedicine, engineering and reliability. For instance, the family of the two-parameter inverse Gaussian distribution is one of the basic models for describing positively skewed data which arise in a variety of fields of applied research as cardiology, hydrology, demography, linguistics, employment service, etc. Recently, Huberman et al. (1998) have argued and demonstrated the appropriateness of the inverse Gaussian family for studying the internet traffic and in particular the number of visited pages per user within an internet site. Furthermore, distributions like the Weibull are frequently encountered in survival modelling where the existence of censoring schemes makes the determination of the proper distribution an extremely challenging problem. Finally distributions like the exponential, the Gamma, the lognormal and others are very common in lifetime problems. For the null distributions we focus on Gamma, Weibull, and Inverse Gaussian distributions. Since, there are often limitations regarding the number of alternatives a test can detect (Jenssen, 2000), we have chosen to investigate the performance of the tests of fit in relation to the degree of skewness of the distribution. The shape parameter of the Gamma, the Weibull and Inverse Gaussian distributions corresponding to a skewness equal to 1.414 and 2.000 are 2 and 1, 1.259 and 1, and 4.5 and 2.25 respectively. Observe that Gamma and Weibull distributions with shape parameter equal to 1 coincide with the exponential distribution. Distributions with the same skewness val-
ues have been used as possible alternatives in each case. The scale parameter is taken to be equal to 1 in all cases, since all three distributions are scale invariant. Other skewness values considered are 0.707, 1.000 and 2.828 but due to space limitations the results are not presented here. The number of intervals used are $m = 3, 4, 5,$ and 10 and we focus on small and medium sample sizes, namely $n = 10, 20, 30, 50$ and 100. For completeness, we also present, for $n = 100,$ the best alternative test, among known tests in the literature, for the Inverse Gaussian and the Gamma cases. The best alternative tests appear to be the Anderson-Darling (AD) test, the Cumulant test of Koutrouvelis et al. (2010) and the $\Phi$-test of Vonta et. al. (2010).

Tables 3-5 clearly show the excellent performance of the proposed test in all cases. It is important to point out that the test performs well even in cases where other well known tests, fail. For instance, the Inverse Gaussian and Lognormal distributions are very often indistinguishable. In this case, the power of the test is remarkably high, making the maximum entropy test the best among all known tests for the IG distribution, including the famous AD test. Finally observe that in almost all cases the maximum entropy test is superior to the corresponding best alternative test available in the literature.

3.3 Real Data

In this section we present two real examples to show the behavior of the test in real cases. Note that in both cases the composite null hypothesis is considered. For Example 2 the parametric family used is the inverse Gaussian while for Example 1 we have considered three different parametric families namely gamma, inverse Gaussian and lognormal. These choices have been suggested in the literature since both examples have been extensively analyzed.

**Example 1.** The data given below represent active repair times (in hours) for an airborne communication transceiver.

\[ .2, .3, .5, .5, .5, .6, .6, .7, .7, .8, .8, 1.0, 1.0, 1.0, 1.1, 1.3, 1.5, 1.5, 1.5, 2.0, 2.0, 2.2, 2.5, 3.0, 3.0, 3.3, 3.3, 4.0, 4.0, 4.5, 4.7, 5.0, 5.4, 5.4, 7.0, 7.5, 8.8, 9.0, 10.3, 22.0, 24.5. \]

The data were first analyzed by von Alven (1964) who fitted successfully the Lognormal distribution. Chhikara and Folks (1977) fitted the IG distribution
and using the observed value of the Kolmogorov-Smirnov statistic they found that the fit is good (KS test statistic=0.07245267). The same conclusion is drawn by using the AD test (test statistic=0.2392647) and the Mudholkar independence characterization test (test statistic=0.2026783, Mudholkar et. al, 2001). On the other hand the Fisher’s ψ test (Fisher, 1948) for combining the p-values of independent tests as well as the usual χ² test fail at both the 5% and the 10% level (test statistic=0.3568939). Note that the test is applied to independent statistics based on skewness and kurtosis. Finally, Koutrouvelis et. al (2010) applied the Gamma distribution which was clearly rejected.

The implementation of the maximum entropy test for \( m = 3, 4, 5, \) and 10 can be used to investigate all the above conclusions. Indeed, our results verify that the lognormal and IG distributions are indistinguishable since the resulting test statistics are extremely close to each other. Further, in all cases, the p-value of the maximum entropy test is much larger than 10% so that the fit of the IG distribution is accepted at the 5% level. Finally, the test easily rejects the gamma distribution at the standard 5% level.

**Example 2.** The data represent precipitation (in inches) from Jug Bridge, Maryland.

1.01, 1.11, 1.13, 1.15, 1.16, 1.17, 1.17, 1.2, 1.52, 1.54, 1.54, 1.57, 1.64, 1.73, 1.79, 2.09, 2.09, 2.57, 2.75, 2.93, 3.19, 3.54, 3.57, 5.11, 5.62.

Folks and Chhikara (1978) found that the fit of the IG to this data is not satisfactory. The same conclusion has been drawn by O’ Reilly and Rueda (1992) who found a p-value of 0.04 by using a Monte Carlo approximation of the distribution of the AD test statistic. The modified AD test of Pavur et al. (1992) barely accepts the IG model at the 5% level and clearly rejects at the 10% level (p-value=slightly over 5%). Note though that the modified AD, CvM, and Watson test of Gunes et al. (1997) reject the IG model at the 5% level. Further, the Mudholkar independence characterization test rejects the IG distribution at the 5% and 10% levels. However, Henze and Klar (2002) found p-values between 8% and 11% using statistics based on Laplace transform.

Our maximum entropy test is extremely useful in cases like this particular one where the decisions made by various and diverse techniques are not in agreement. Indeed, since the proposed test involves a set of weights one expects
that the value of the test varies based on the weights selected. For this example, we have chosen to run the test 100 times and see the behavior of the test statistic for the different weights chosen. We have observed that the value of the test statistic in almost all instances, lies around the 5% critical level, with half of the times being below and half of the times above this critical point. As a result the null distribution is rejected at the 5% level in 50% of the cases and accepted at the same level at the remaining 50% of the cases. The p-values found between 3% and 10%, confirm the results obtained by O’Reilly and Rueda (1992), Pavur et al. (1992), and Henze and Klar (2002).

In conclusion, we have proposed a maximum entropy test and established its asymptotic distribution. Furthermore, a modified test with appropriately chosen critical values, has been proposed for small and medium sample sizes. Taking into consideration the size and the power results, the Monte Carlo simulation study for various distributions and a variety of values of $m$, has clearly shown a satisfactory performance of the maximum entropy test for small, medium and large sample sizes. Finally note the applicability of the proposed method in various classes of lifetime distributions for describing several aging criteria like the classes of increasing failure rate (IFR), the decreasing failure rate (DFR) and the new better than used (NBU) where the exponential null model is examined against all other members of the class.

References


Table 2: Uniform Null Model: Power and Size of the Maximum Entropy Test for \( n = 10, 20, 30, 50, 100, 300, 500, 1000. \)

<table>
<thead>
<tr>
<th>( m )</th>
<th>Beta</th>
<th>( 10 )</th>
<th>( 20 )</th>
<th>( 30 )</th>
<th>( 50 )</th>
<th>( 100 )</th>
<th>( 300 )</th>
<th>( 500 )</th>
<th>( 1000 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1,1</td>
<td>0.044</td>
<td>0.041</td>
<td>0.059</td>
<td>0.058</td>
<td>0.048</td>
<td>0.045</td>
<td>0.062</td>
<td>0.053</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.060</td>
<td>0.048</td>
<td>0.056</td>
<td>0.055</td>
<td>0.060</td>
<td>0.066</td>
<td>0.063</td>
<td>0.048</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.048</td>
<td>0.043</td>
<td>0.048</td>
<td>0.062</td>
<td>0.033</td>
<td>0.052</td>
<td>0.051</td>
<td>0.050</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.026</td>
<td>0.042</td>
<td>0.058</td>
<td>0.074</td>
<td>0.090</td>
<td>0.130</td>
<td>0.103</td>
<td>0.099</td>
</tr>
<tr>
<td>3</td>
<td>0.5,0.5</td>
<td>0.092</td>
<td>0.095</td>
<td>0.157</td>
<td>0.228</td>
<td>0.397</td>
<td>0.975</td>
<td>0.999</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.114</td>
<td>0.174</td>
<td>0.273</td>
<td>0.438</td>
<td>0.700</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.113</td>
<td>0.185</td>
<td>0.295</td>
<td>0.515</td>
<td>0.840</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.136</td>
<td>0.273</td>
<td>0.450</td>
<td>0.758</td>
<td>0.979</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>3</td>
<td>0.2,2</td>
<td>0.148</td>
<td>0.201</td>
<td>0.330</td>
<td>0.515</td>
<td>0.805</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.146</td>
<td>0.223</td>
<td>0.314</td>
<td>0.503</td>
<td>0.822</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.098</td>
<td>0.180</td>
<td>0.246</td>
<td>0.465</td>
<td>0.780</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.081</td>
<td>0.171</td>
<td>0.300</td>
<td>0.531</td>
<td>0.895</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>3</td>
<td>2.5</td>
<td>0.537</td>
<td>0.779</td>
<td>0.940</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.476</td>
<td>0.842</td>
<td>0.972</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.379</td>
<td>0.772</td>
<td>0.944</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.302</td>
<td>0.734</td>
<td>0.964</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>3</td>
<td>2.6</td>
<td>0.727</td>
<td>0.959</td>
<td>0.997</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.636</td>
<td>0.956</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.511</td>
<td>0.915</td>
<td>0.997</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.334</td>
<td>0.873</td>
<td>0.997</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>3</td>
<td>1.8</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.994</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.955</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>3</td>
<td>mix</td>
<td>0.540</td>
<td>0.826</td>
<td>0.974</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.348</td>
<td>0.584</td>
<td>0.798</td>
<td>0.943</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.240</td>
<td>0.521</td>
<td>0.741</td>
<td>0.955</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.185</td>
<td>0.505</td>
<td>0.824</td>
<td>0.995</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Table 3: Gamma Null Model: Power and Size of the Maximum Entropy Test for \( n = 10, 20, 30, 50, 100 \) and Best Alternative (BA) test for \( n = 100 \).

<table>
<thead>
<tr>
<th>( m )</th>
<th>( \text{Shape} )</th>
<th>( \text{MAXIMUM ENTROPY TEST} )</th>
<th>( \text{BA TEST} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Gamma</td>
<td>0.057 0.037 0.050 0.052 0.046</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>0.044 0.045 0.045 0.054 0.043</td>
<td>0.053 - 0.060 (**)</td>
</tr>
<tr>
<td>10</td>
<td>0.028 0.046 0.061 0.076 0.084</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Wei</td>
<td>0.678 0.906 0.990 0.999 1.000</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.259</td>
<td>0.537 0.883 0.983 0.999 1.000</td>
<td>1.000 (( \Phi^* ))</td>
</tr>
<tr>
<td>10</td>
<td>0.404 0.825 0.974 1.000 1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Ln</td>
<td>0.544 0.800 0.958 1.000 1.000</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.423</td>
<td>0.411 0.805 0.961 0.999 1.000</td>
<td>0.925 (R***)</td>
</tr>
<tr>
<td>10</td>
<td>0.356 0.831 0.984 1.000 1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>IG</td>
<td>0.986 1.000 1.000 1.000 1.000</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4.5</td>
<td>0.954 1.000 1.000 1.000 1.000</td>
<td>0.940 (R***)</td>
</tr>
<tr>
<td>10</td>
<td>0.866 0.999 1.000 1.000 1.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( m )</th>
<th>( \text{Shape} )</th>
<th>( \text{10} )</th>
<th>( \text{20} )</th>
<th>( \text{30} )</th>
<th>( \text{50} )</th>
<th>( \text{100} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td></td>
<td>0.057 0.037 0.048 0.053 0.046</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Gamma</td>
<td>0.059 0.057 0.050 0.057 0.051</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0.044 0.044 0.058 0.051 0.046</td>
<td>0.035-0.048(**)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.033 0.048 0.049 0.080 0.084</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Ln</td>
<td>0.278 0.396 0.670 0.958 1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.551</td>
<td>0.238 0.512 0.760 0.965 1.000</td>
<td>0.981 (R***)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.157 0.460 0.758 0.987 1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>IG</td>
<td>0.874 0.989 1.000 1.000 1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2.25</td>
<td>0.814 0.993 1.000 1.000 1.000</td>
<td>0.999 (AD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.720 0.987 1.000 1.000 1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(*) \( \Phi \)-test (Vonta et. al, 2010). (**) Cumulant test (Koutrouvelis et. al, 2010).
Table 4: Inverse Gaussian Null Model: Power and Size of the Maximum Entropy Test for \( n = 10, 20, 30, 50, 100 \) and Best Alternative (BA) test for \( n = 100 \).

<table>
<thead>
<tr>
<th>( m )</th>
<th>( \text{Shape} )</th>
<th>( \text{MAXIMUM ENTROPY TEST} )</th>
<th>( \text{BA TEST} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>IG</td>
<td>0.059</td>
<td>0.041</td>
</tr>
<tr>
<td>4</td>
<td>IG</td>
<td>0.059</td>
<td>0.049</td>
</tr>
<tr>
<td>5</td>
<td>4.5</td>
<td>0.041</td>
<td>0.045</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.030</td>
<td>0.044</td>
</tr>
<tr>
<td>3</td>
<td>Wei</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td>1.259</td>
<td>0.999</td>
<td>1.000</td>
</tr>
<tr>
<td>5</td>
<td>2.25</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>3</td>
<td>Gamma</td>
<td>0.938</td>
<td>0.998</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0.937</td>
<td>0.999</td>
</tr>
<tr>
<td>5</td>
<td>0.423</td>
<td>0.943</td>
<td>1.000</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.943</td>
<td>1.000</td>
</tr>
</tbody>
</table>

(*) \( \Phi \)-test (Vonta et. al, 2010).  (**) Cumulant test (Kourouvelis et. al, 2010).
Table 5: Weibull Null Model: Power and Size of the Maximum Entropy Test for $n = 10, 20, 30, 50, 100$ and Best Alternative (BA) test for $n = 100$.

<table>
<thead>
<tr>
<th>m</th>
<th>Shape</th>
<th>MAXIMUM ENTROPY TEST</th>
<th>BA TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>Wei</td>
<td>0.056</td>
<td>0.042</td>
</tr>
<tr>
<td>4</td>
<td>1.259</td>
<td>0.061</td>
<td>0.049</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>0.031</td>
<td>0.045</td>
</tr>
<tr>
<td>3</td>
<td>IG</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td>4.5</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>3</td>
<td>Ln</td>
<td>0.287</td>
<td>0.427</td>
</tr>
<tr>
<td>4</td>
<td>0.423</td>
<td>0.335</td>
<td>0.568</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>0.168</td>
<td>0.501</td>
</tr>
<tr>
<td>3</td>
<td>Gamma</td>
<td>0.665</td>
<td>0.896</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0.646</td>
<td>0.911</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>0.594</td>
<td>0.909</td>
</tr>
</tbody>
</table>

(†) Not Available. (*) Φ-test (Vonta et al, 2010).