Goodness-of-fit tests for the exponential and the normal distribution based on the integrated distribution function

Bernhard Klar

Institut für Mathematische Stochastik, Universität Karlsruhe,

Englerstr. 2, 76128 Karlsruhe, Germany

Bernhard.Klar@math.uni-karlsruhe.de

Abstract

This paper presents new omnibus tests for the exponential and the normal distribution which are based on the difference between the integrated distribution function $\Psi(t) = \int_t^\infty (1 - F(x))\, dx$ and its empirical counterpart. The procedures turn out to be serious competitors to classical tests for exponentiality and normality.

Key words: Goodness-of-fit test, integrated distribution function, exponential distribution, normal distribution.

1 Introduction

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed (iid) nonnegative random variables with distribution function (df) F, and consider the problem of testing whether the random sample has come from a specified parametric family of distributions. Each goodness-of-fit test which aims at detecting all departures from this hypothesized model has to use some characteristic property of the parametric class of distributions. Since the distribution of the X_i is uniquely determined by their *integrated distribution function* (idf) which, for a positive random variable X with $EX < \infty$, is defined by

$$\Psi(t) \;:=\; E(X-t)^+ \;=\; \int_t^\infty (1 - F(x))\; dx, \tag{1.1}$$

such an omnibus test may be based on the difference between the idf and its empirical counterpart.

In case the X_i are discrete valued random variables, idf-tests were proposed by Klar (1999). For common lattice models, the tests have high power with respect to competitive procedures over a large range of alternatives.

Henze and Nikitin (1998) considered tests based on the so-called integrated empirical process for testing a completely specified hypothesis and calculated local Bahadur efficiencies in the setting of shift alternatives. Since there, in contrast to (1.1), the idf is defined by integration with respect to F, it coincides with the idf as defined conventionally only in the case of uniformly distributed random variables.

It is the purpose of this paper to study tests based on the idf for two important continuous distributions, the exponential and the normal distribution. In Section 2, a test for exponentiality is considered. After the derivation of the limiting null distribution of the test statistic, the test is shown to be consistent against each alternative distribution with finite positive expectation. The finite sample properties of the test as well as of several modifications thereof are assessed by means of some simulations. The procedures turn out to have high power in comparison with classical tests for exponentiality.

Section 3 deals with testing for normality. In this case, the definition of the idf has to be adapted suitably. Again, the asymptotic distribution of the test statistic is obtained. In an extensive simulation study, powers of the new test, the Shapiro-Wilk test and the Anderson-Darling test are compared. Whereas the new test for normality compares favourably to the other tests for heavy-tailed alternatives, power against short-tailed alternatives is fairly low; in many situations, however, safeguarding against such distributions should be less important.

2 A goodness-of-fit test for exponentiality

Let $\mathcal{F} = \{F(\cdot, \vartheta) : \vartheta > 0\}$, where $F(t, \vartheta) = 1 - \exp(-\vartheta\, t)$, $t \geq 0$, denote the class of exponential distributions. The problem is to test the composite hypothesis $\mathcal{H}_0 : F \in \mathcal{F}$ against the general alternative $\mathcal{H}_1 : F \notin \mathcal{F}$. For this purpose, we propose a test statistic based on the idf (as defined in (1.1)), which, if the distribution of X is exponential with parameter ϑ, takes the form

$$\Psi(t, \vartheta) \quad = \quad e^{-\vartheta\, t} / \vartheta, \quad t \geq 0.$$

The empirical counterpart to Ψ is the empirical idf

$$\Psi_n(t) \;=\; \int_t^\infty (1 - F_n(x))\, dx \;=\; \frac{1}{n}\sum_{i=1}^n (X_i - t)\, \mathbf{1}\{X_i > t\},$$

where $\mathbf{1}$ denotes the indicator function, and $F_n(x) = n^{-1}\sum_{j=1}^n \mathbf{1}\{X_j \le x\}$ is the empirical df of $X_1, \ldots, X_n$. To perform the test, Ψ_n is compared with the estimated idf

$$\Psi(t, \hat{\vartheta}_n) \;=\; \int_t^\infty (1 - F(x, \hat{\vartheta}_n))\, dx.$$

Here, $\hat{\vartheta}_n = 1/\bar{X}_n$ with $\bar{X}_n = n^{-1}\sum_{j=1}^n X_j$ is the maximum likelihood estimator of ϑ. A convenient test statistic is

$$T_n \;=\; \hat{\vartheta}_n^3 \int_0^\infty \left(\sqrt{n}\left(\Psi_n(t) - \Psi(t, \hat{\vartheta}_n)\right)\right)^2 dt, \tag{2.2}$$

which, putting $Y_i = \hat{\vartheta}_n X_i = X_i/\bar{X}_n$ and $u = \hat{\vartheta}_n\, t$, takes the form

$$T_n \;=\; \int_0^\infty Z_n^2(u)\, du,$$

where

$$Z_n(u) \;=\; \frac{1}{\sqrt{n}}\sum_{j=1}^n \{(Y_j - u)^+ \;-\; e^{-u}\}, \quad 0 \le u < \infty. \tag{2.3}$$

Evaluating the integral yields

$$\begin{aligned} T_n \;&=\; \frac{1}{n}\sum_{i,j=1}^n \left[\int_0^{\min(Y_i,Y_j)} (Y_i - u)(Y_j - u)\, du - 2\int_0^{Y_i} (Y_i - u)e^{-u}\, du + \int_0^\infty e^{-2u}\, du\right] \\ &=\; \frac{1}{n}\sum_{i<j}\left(Y_{(i)}^2 Y_{(j)} - \frac{Y_{(i)}^3}{3}\right) + \frac{1}{3n}\sum_{i=1}^n Y_{(i)}^3 - 2\sum_{i=1}^n \left(e^{-Y_i} + Y_i - 1\right) + \frac{n}{2}, \end{aligned}$$

where $Y_{(1)} \le Y_{(2)} \le \ldots \le Y_{(n)}$ are the order statistics of $Y_1, \ldots, Y_n$. Hence, using $1/n\sum_{i=1}^n Y_i = 1$, we obtain the following alternative representation which is more suitable for computational purposes:

$$T_n \;=\; n/2 - 2\sum_{i=1}^n e^{-Y_i} - (3n)^{-1}\sum_{i=1}^n (n - i - 1)Y_{(i)}^3 + n^{-1}\sum_{i<j} Y_{(i)}^2 Y_{(j)}.$$

Note that, being a function of $Y_1, \ldots, Y_n$, the statistic T_n is scale-invariant. We thus may assume $\vartheta = 1$ without loss of generality in what follows.

To prove the weak convergence of T_n as $n \to \infty$, it is convenient to work in the Hilbert space $L_2 = L_2(\mathbb{R}_+, \mathcal{B}_+, \lambda)$ of square integrable functions on $\mathbb{R}_+$, since, in this case, T_n is a continuous functional of the process Z_n in (2.3). The inner product and the norm in L_2 are denoted by $(\cdot, \cdot)$ and $\|\cdot\|$, respectively. Neuhaus (1974) used the Hilbert space setting to prove the weak convergence of the Cramér-von Mises statistic; for sums of centered iid Hilbert space valued random variables $X_1, X_2, \ldots$, the Central Limit Theorem holds if (and only if) $\|X_1\|$ has finite variance.

To show that Z_n can be represented as the sum of iid random variables and a remainder term that is asymptotically negligible, note that the maximum likelihood estimator $\hat{\vartheta}_n$ has the representation

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta) \;=\; \frac{1}{\sqrt{n}} \sum_{j=1}^{n} l(X_j, \vartheta) \;+\; r_n, \tag{2.4}$$

where $l(x, \vartheta) = -(\vartheta^2 x - \vartheta)$ and $r_n = o_P(1)$. Thus, it follows by the Mean Value Theorem that

$$\begin{aligned}
&\sqrt{n}\left(\Psi_n(t) - \Psi(t, \hat{\vartheta}_n)\right) \\
&= \sqrt{n}\,(\Psi_n(t) - \Psi(t,1)) \;+\; \sqrt{n}\left(\Psi(t,1) - \Psi(t, \hat{\vartheta}_n)\right) \\
&= \sqrt{n}\,(\Psi_n(t) - \Psi(t,1)) \;-\; \sqrt{n}\left(\hat{\vartheta}_n - \vartheta\right) \left.\frac{\partial \Psi(t,\vartheta)}{\partial \vartheta}\right|_{\vartheta=\vartheta_n^*} \\
&= \sqrt{n}(\Psi_n(t) - \Psi(t,1)) \;-\; \left.\frac{\partial \Psi(t,\vartheta)}{\partial \vartheta}\right|_{\vartheta=1} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} l(X_j, \vartheta) \\
&\quad + \sqrt{n}\left(\hat{\vartheta}_n - \vartheta\right) \left(\left.\frac{\partial \Psi(t,\vartheta)}{\partial \vartheta}\right|_{\vartheta=1} - \left.\frac{\partial \Psi(t,\vartheta)}{\partial \vartheta}\right|_{\vartheta=\vartheta_n^*} \right) \;-\; r_n \left.\frac{\partial \Psi(t,\vartheta)}{\partial \vartheta}\right|_{\vartheta=1},
\end{aligned}$$

where ϑ_n^* is between $\hat{\vartheta}_n$ and ϑ. Hence,

$$\sqrt{n}\left(\Psi_n(t) - \Psi(t, \hat{\vartheta}_n)\right) \;=\; W_n(t) \;-\; R_n(t), \tag{2.5}$$

where

$$R_n(t) \;=\; r_n \left.\frac{\partial \Psi(t,\vartheta)}{\partial \vartheta}\right|_{\vartheta=1} + \sqrt{n}\left(\hat{\vartheta}_n - \vartheta\right) \left(\left.\frac{\partial \Psi(t,\vartheta)}{\partial \vartheta}\right|_{\vartheta=\vartheta_n^*} - \left.\frac{\partial \Psi(t,\vartheta)}{\partial \vartheta}\right|_{\vartheta=1} \right)$$

and

$$W_n(t) \;=\; \sqrt{n}(\Psi_n(t) - \Psi(t,1)) \;-\; \left.\frac{\partial \Psi(t,\vartheta)}{\partial \vartheta}\right|_{\vartheta=1} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} l(X_j, \vartheta).$$

Defining the function

$$g(t,x) \;=\; (x-t)^+ \,-\, \Psi(t,1) \,-\, l(x,1)\, \frac{\partial \Psi(t,\vartheta)}{\partial \vartheta}\bigg|_{\vartheta=1} \qquad (2.6)$$
$$\;=\; (x-t)^+ \,-\, (x-t+xt)\, e^{-t}$$

which satisfies

$$g(\cdot,x) \;\in\; L_2, \quad x \in \mathbb{R}_+,$$
$$E[\,g(t,X)\,] \;=\; \int_0^\infty g(t,x)\, e^{-x}\, dx \;=\; 0, \quad t \in \mathbb{R}_+, \qquad (2.7)$$
$$E[\,\|g(\cdot,X)\|^2\,] \;=\; E\left[\int_0^\infty g^2(t,X)\, dt\right] \;<\; \infty,$$

we obtain the representation $W_n = n^{-1/2}\sum_{j=1}^n g(\cdot,X_j)$. Since the L_2-valued random variable W_n is the sum of iid random elements and satisfies $E[(W_n,f)] = 0$ ($f \in L_2$) and $Var(\|W_n\|) < \infty$, the Hilbert space Central Limit Theorem yields the existence of a Gaussian random element W in L_2 such that

$$W_n \;\xrightarrow{\mathcal{D}}\; W \qquad (2.8)$$

(see, e.g., Araujo and Giné (1980), section 3.7). On the other hand, it is not difficult to see that $\|R_n\| = o_P(1)$. Together with (2.8) and (2.5), Theorem 4.1 in Billingsley (1968) and the continuous mapping theorem yield the following result:

2.1 Theorem *Under the hypothesis of exponentiality,*

$$Z_n \;\xrightarrow{\mathcal{D}}\; W$$

in L_2, where W is a centered Gaussian process with covariance function

$$k(s,t) \;=\; Cov(W(s),W(t)) \;=\; E[g(s,X_1)\, g(t,X_1)]$$
$$\;=\; (t-s+2)\, e^{-t} \;-\; (s+t+s\,t+2)\, e^{-(s+t)}, \quad s \le t.$$

Furthermore, $T_n \;=\; \|Z_n\|^2 \;\xrightarrow{\mathcal{D}}\; \|W\|^2$.

Let λ_j resp. φ_j, $j \ge 1$, denote the eigenvalues resp. eigenfunctions of the integral operator with kernel $k(\cdot,\cdot)$, i.e.

$$\int_0^\infty k(s,t)\, \varphi_j(t)\, dt \;=\; \lambda_j\, \varphi_j(s), \quad 0 \le s < \infty, \qquad (2.9)$$

for $j \geq 1$. The distribution of $\|W\|^2$ is that of $\sum_{j\geq 1} \lambda_j N_j^2$, where $N_1, N_2, \ldots$ is a sequence of independent unit normal variables. Differentiating (2.9) several times shows that λ_j resp. φ_j are the solutions of the equation

$$\varphi^{(4)}(t) \;+\; 2\,\varphi^{(3)}(t) \;+\; \varphi^{(2)}(t) \;\;=\;\; e^{-t}\varphi(t)/\lambda,$$

where $\varphi, \varphi' \in L_2$ and $\varphi(0) = \varphi'(0) = 0$. However, it seems that closed-form solutions of this equation do not exist. Expectation and variance of $\|W\|^2$ are given by

$$\begin{aligned} E(\,\|W\|^2\,) \;&=\; \int k(t,t)\,dt \;=\; \frac{1}{4}, \\ Var(\,\|W\|^2\,) \;&=\; 2\iint k^2(s,t)\,ds\,dt \;=\; \frac{17}{216}. \end{aligned}$$

Table 1 shows empirical critical values of T_n for several testing levels and sample sizes. Note that the distribution of T_n converges quite slowly to its asymptotic distribution. The entries in Table 1 are the 20%-trimmed means of 100 Monte-Carlo simulations, each based on 10000 replications; here, we always used $\vartheta = 1$ (recall scale invariance).

Let $\alpha \in (0,1)$, and let $z_n(\alpha)$ denote the $(1-\alpha)$-quantile of T_n under $\mathcal{H}_0$. Regarding consistency of the test which rejects the hypothesis of exponentiality if $T_n > z_n(\alpha)$, we have the following result.

2.2 Theorem *The test based on T_n is consistent against each alternative distribution P with finite positive expectation.*

Proof: Let Ψ_A and μ_A denote the idf and the expectation of the alternative distribution. Now,

$$\|\Psi_n - \Psi(\cdot,\hat\vartheta_n)\| \;\;\geq\;\; \Big|\, \|\Psi_n - \Psi(\cdot,\vartheta)\| \;-\; \|\Psi(\cdot,\vartheta) - \Psi(\cdot,\hat\vartheta_n)\|\,\Big|, \qquad (2.10)$$

where $\vartheta = 1/\mu_A$. To establish a Glivenko-Cantelli type result for Ψ_n, note that, for $\epsilon > 0$, there exists $M_\epsilon > 0$, for which $E_A(X \cdot \mathbf{1}\{X \geq M_\epsilon\}) < \epsilon$ (where E_A denotes expectation under the alternative distribution P). Now, define a finite class $\mathcal{F}_\epsilon$ consisting of the functions $g_{k,\epsilon}(x) = (x - k\epsilon)^+,\ k = 0,\ldots,[M_\epsilon/\epsilon]+1$ and $g \equiv 0$. Then, for each function $g_s(x) = (x-s)^+$ $(s \geq 0)$, there exist lower and upper approximations $g_{\epsilon,L}$ and $g_{\epsilon,U}$ in $\mathcal{F}_\epsilon$, for which

$$g_{\epsilon,L} \leq g_s \leq g_{\epsilon,U} \quad \text{and} \quad E_A[g_{\epsilon,U}(X) - g_{\epsilon,L}(X)] \leq \epsilon,$$

and Theorem II.2 in Pollard (1984) yields

$$\sup_{0\le s<\infty} |\Psi_n(s) - \Psi_A(s))| \longrightarrow 0 \quad a.s.$$

Hence, there is a common P-null set N such that for each $\omega \in N^C$ and each $s \in [0,\infty)$

$$\lim_{n\to\infty} |\Psi_n(s)(\omega) - \Psi(s,\vartheta)| = |\Psi_A(s) - \Psi(s,\vartheta)|. \tag{2.11}$$

Since, as mentioned in the introduction, a distribution is uniquely determined by the idf (see, e.g., Müller (1996)), there exists some number $t > 0$ with

$$|\Psi(t,\vartheta) - \Psi_A(t)| \ge \delta > 0. \tag{2.12}$$

On combining (2.11) and (2.12), and using Fatou's Lemma and the continuity of $\Psi(\cdot,\vartheta)$ and Ψ_A, we obtain

$$\liminf_{n\to\infty} \int_0^\infty (\Psi_n(s) - \Psi(s,\vartheta))^2 \, ds > 0 \quad a.s.$$

On the other hand,

$$\lim_{n\to\infty} \|\Psi(\cdot,\vartheta) - \Psi(\cdot,\hat{\vartheta}_n)\| = 0 \quad a.s.,$$

whence $\lim_{n\to\infty} \|\Psi_n - \Psi(\cdot,\hat{\vartheta}_n)\|^2 > 0$ $a.s.$ in view of (2.10). This implies

$$\lim_{n\to\infty} T_n = \infty \quad a.s.$$

and therefore $\lim_{n\to\infty} P(T_n \le z_n(\alpha)) = 0$. ■

In defining the test statistic, the weight function $\exp(-a\vartheta t)$ with $a > 0$ can be introduced in order to increase the power performance of the procedure. The modified statistic

$$T_{n,a} = (a\,\hat{\vartheta}_n)^3 \int_0^\infty \left(\sqrt{n}\left(\Psi_n(t) - \Psi(t,\hat{\vartheta}_n)\right)\right)^2 \exp(-a\hat{\vartheta}_n t)\, dt \tag{2.13}$$

can be rewritten as

$$T_{n,a} = a^3 \int_0^\infty (Z_n(t))^2 \, e^{-at}\, dt,$$

where the process Z_n is defined in (2.3). The weak convergence

$$T_{n,a} \xrightarrow{\mathcal{D}} a^3 \int_0^\infty W^2(t)\, e^{-at}\, dt$$

can be proved similar to Theorem 2.1 (see also Theorem 3.2). The scaling factor a^3 in (2.13) is motivated by the fact that $\lim_{a\to\infty} T_{n,a} = 2\,n$ for each fixed sample size. This result follows from an application of an Abelian theorem for Laplace transforms (see Widder (1959), p. 182, or Baringhaus et al. (1999), p. 5).

2.3 Lemma *$T_{n,a}$ has the following representation:*

$$
\begin{aligned}
T_{n,a} \;=\; & \frac{2(3a+2)\,n}{(2+a)(1+a)^2} \;-\; 2\,a^3 \sum_{i=1}^{n} \frac{\exp(-(1+a)Y_i)}{(1+a)^2} \\
& -2/n \sum_{i=1}^{n} \exp(-aY_i) + \, 2/n \sum_{i<j} (a(Y_{(j)} - Y_{(i)}) - 2) \exp(-aY_{(i)}).
\end{aligned}
$$

Proof: By definition,

$$
T_{n,a} \;=\; n\,a^3 \int_0^\infty \left(\frac{1}{n} \sum_{j=1}^{n} (Y_j - t)^+ - e^{-t} \right)^2 e^{-at}\, dt,
$$

and straightforward computation gives

$$
\begin{aligned}
T_{n,a} \;=\; & \frac{a^3\,n}{2+a} + \frac{2a^3\,n}{(1+a)^2} \;-\; 2a^3 \sum_{i=1}^{n} \frac{e^{-(1+a)Y_i} \;+\; (1+a)Y_i}{(1+a)^2} \\
& + 1/n \sum_{i=1}^{n} \left(a^2 Y_i^2 - 2aY_i + 2 - 2e^{-aY_i} \right) \\
& + 2/n \sum_{i<j} \left(a^2 Y_{(i)} Y_{(j)} - a(Y_{(i)} + Y_{(j)}) + 2 + \big(a(Y_{(j)} - Y_{(i)}) - 2 \big)\, e^{-aY_{(i)}} \right).
\end{aligned}
$$

Using $1/n \sum_{i=1}^{n} Y_i = 1$ and combining the second sum and the terms of the third sum which are symmetric in i and j, the assertion follows. ■

Table 2 shows the empirical $(1-\alpha)$-quantiles of $T_{n,a}$ for $a = 1, 5, 10, 20$ and different values of α and the sample size n. Apparently, the convergence is faster than in Table 1. The entries in Table 2 are determined in the same way as in Table 1.

We examined the dependence of the power of the test on the weight function by means of a simulation study with the following alternative distributions: Gamma, Weibull and log normal distribution with scale parameter 1 and shape parameter α (denoted by $\Gamma(\alpha)$, $W(\alpha)$ and $LN(\alpha)$), uniform distribution on (0,1) ($U(0,1)$), half normal (HN), half Cauchy (HC), χ_1^2 distribution.

The first five columns of Table 3 show the results of the test based on T_n resp. $T_{n,a}$ for $a = 1, 5, 10, 20$ (φ_0 resp. φ_a are the tests based on T_n resp. $T_{n,a}$, see below). The results indicate that the power depends heavily on a. For each of the five tests, there are alternatives for which the test is most powerful.

We therefore looked for related tests which distribute their power more evenly. To this end, we examined combinations of two or more test statistics $T_{n,a_j} (j = 1, \dots, k)$: $\mathcal{H}_0$ is rejected if at least one of the tests based on T_{n,a_j} rejects the hypothesis.

To give a formal description of the test, let $\varphi_{a,\alpha} = \mathbf{1}\{T_{n,a} > z_{n,a}(\alpha)\}$ and $\varphi_{0,\alpha} = \mathbf{1}\{T_n > z_n(\alpha)\}$ denote the level-α-tests based on $T_{n,a}$ and T_n, respectively; the tests result in 1 (rejection) or 0 (no rejection). The combined level-α-test is defined by

$$\varphi_{a_1,\dots,a_k,\alpha} \quad := \quad \max_{j=1\dots,k} \varphi_{a_j,\alpha^*},$$

where $\alpha/k \leq \alpha^* \leq \alpha$. α^* is uniquely determined by $E[\varphi_{a_1,\dots,a_k,\alpha}] = \alpha$. In practice, the quantiles have to be found empirically by a search algorithm. Using the quantiles that belong to α/k leads to a conservative test.

As a first combination, we examined $\varphi_{1,10,\alpha}$. For $n = 20$ and $\alpha = 0.1$, the quantiles are 0.108 for $T_{n,1}$ and 0.321 for $T_{n,10}$; for $\alpha = 0.05$, the corresponding values are 0.141 and 0.421. Column 6 of Table 3 shows the simulation results (in the table, the subscript α is omitted). Since the quantiles do not differ much from the $(1 - \alpha/2)$-quantiles of the original tests (see Tables 1, 2), we ran simulations with these quantiles; the results are given in column 6 of Table 3 in parenthesis. Furthermore, the results for combination $\varphi_{0,20,\alpha}$ are listed; the quantiles are 0.601 resp. 0.211 for $\alpha = 0.1$ and 0.845 resp. 0.285 for $\alpha = 0.05$. The last column shows the results for $\varphi_{0,5,20,\alpha}$; the corresponding quantiles are 0.646, 0.434 resp. 0.224 ($\alpha = 0.1$) and 0.937, 0.566 resp. 0.321 ($\alpha = 0.05$).

The conclusion is quite obvious: $\varphi_{1,10,\alpha}$ is a strong omnibus test and preferable to the other tests if nothing is known about the alternative. In comparison with $\varphi_{0,5,20,\alpha}$, the test has the further advantage that the $(1 - \alpha/2)$-quantiles can be used with little loss of power.

Besides the findings for the tests based on the idf, we provide the results of the Cramér-von Mises test W^2 and the Anderson-Darling test A^2. In comparison with $\varphi_{1,10,\alpha}$, both W^2 and A^2 behave quite similar; each of the three tests does best in about the same number of cases.

As our procedure was an omnibus one we compared it with the two procedures recommended by D'Agostino and Stephens (1986). Sometimes one has special classes of alternatives in mind. Then one should use a test directed at them. Examples of such tests are Gail-Gastwirth (1978), Lewis (1965) and Klefsjö (1983).

The results of Table 3 can also be compared with parts of Table 4 in Baringhaus and Henze (1991) and parts of Table 2 in Baringhaus and Henze (1992); in doing so, it turns out as well that the new test has high power against many alternatives.

The power of the idf tests could certainly be improved by an adaptive choice of the weight function; this remains a point of further research.

3 Testing for normality

In this section, we consider the problem of testing the composite hypothesis $\mathcal{H}_0 : F \in \mathcal{N} = \{N(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 > 0\}$ against the general alternative $\mathcal{H}_1 : F \notin \mathcal{N}$. Again, we propose a test based on the idf, which, in case of a random variable on $\mathbb{R}$ with finite expectation, is defined by

$$\psi(t) \;:=\; E(X-t)^- \;=\; \int_{-\infty}^{t} F(x)\,dx,$$

where $y^- = -\min\{y, 0\}$ (the use of the same symbol as in Section 2 should cause no confusion). Accordingly, the empirical idf is

$$\psi_n(t) \;=\; \int_{-\infty}^{t} F_n(x)\,dx \;=\; -\frac{1}{n}\sum_{i=1}^{n}(X_i - t)\,\mathbf{1}\{X_i < t\}.$$

Let Φ_{μ,σ^2} and φ_{μ,σ^2} denote the distribution function and the density of the $N(\mu, \sigma^2)$-distribution, respectively; furthermore, $\Phi = \Phi_{0,1}$ and $\varphi = \varphi_{0,1}$. If $X \sim N(\mu, \sigma^2)$, we write ψ_{μ,σ^2} instead of ψ. By differentiation, one can verify the representation

$$\psi_{\mu,\sigma^2}(x) \;=\; (x-\mu)\,\Phi_{\mu,\sigma^2}(x) \;+\; \sigma^2\,\varphi_{\mu,\sigma^2}(x), \quad x \in \mathbb{R}. \tag{3.14}$$

Since

$$\lim_{t\to\infty}\,(\psi_n(t) - \psi_{\mu,\sigma^2}(t)) \;=\; \mu - \bar{X}_n,$$

an integral statistic without weight function (as in case of the exponential distribution) cannot be used. A possible test statistic is

$$\tilde{T}_n \;=\; \frac{n}{\hat{\sigma}_n^2}\int_{-\infty}^{\infty}\left(\psi_n(t) - \psi_{\bar{X}_n,\hat{\sigma}_n^2}(t)\right)^2\;\varphi_{\bar{X}_n,\hat{\sigma}_n^2}(t)\,dt.$$

Let $Y_i = (X_i - \bar{X}_n)/\hat{\sigma}_n$, $i = 1, \ldots, n$, where $\hat{\sigma}_n^2 = n^{-1}\sum_{j=1}^{n}(X_j - \bar{X}_n)^2$ denotes the empirical variance. The substitution $u = (t - \bar{X}_n)/\hat{\sigma}_n$ and (3.14) yield

$$\tilde{T}_n \;=\; n\int_{-\infty}^{\infty}\left(\frac{1}{n}\sum_{j=1}^{n}(Y_j - u)^- \;-\; (u\,\Phi(u) + \varphi(u))\right)^2\;\varphi(u)\,du$$

$$= \int_{-\infty}^{\infty} \tilde{Z}_n^2(u)\, \varphi(u)\, du, \tag{3.15}$$

where

$$\tilde{Z}_n(u) \;=\; \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left\{ (Y_j - u)^- \;-\; (u\,\Phi(u) + \varphi(u)) \right\}, \quad u \in \mathbb{R}.$$

3.1 Lemma *$\tilde{T}_n$ has the following representation:*

$$\begin{aligned} \tilde{T}_n \;&=\; \frac{n}{3} + \frac{n\sqrt{3}}{2\,\pi} - \sum_{i=1}^{n} \left(1 - \Phi^2(Y_i) \;-\; \frac{2}{\sqrt{\pi}}\, Y_i\,(1 - \Phi(\sqrt{2}\,Y_i)) \;+\; 2\,\varphi^2(Y_i) \right) \\ &\quad + \frac{1}{n} \sum_{i,j=1}^{n} \Big((1 - \Phi(Y_i \vee Y_j))\,(1 + Y_i\,Y_j) \;-\; (Y_i \wedge Y_j)\,\varphi(Y_i \vee Y_j) \Big), \end{aligned}$$

where $x \vee y = \max(x,y)$ and $x \wedge y = \min(x,y)$.

Proof: (3.15) yields

$$\begin{aligned} \tilde{T}_n \;&=\; \frac{1}{n} \sum_{i,j=1}^{n} \int_{\max(Y_i,Y_j)}^{\infty} (u - Y_i)\,(u - Y_j)\;\varphi(u)\;du \\ &\quad - 2 \sum_{j=1}^{n} \int_{Y_j}^{\infty} (u\,\Phi(u) + \varphi(u))\,(u - Y_j)\;\varphi(u)\;du \\ &\quad + n \int_{-\infty}^{\infty} (u\,\Phi(u) + \varphi(u))^2\;\varphi(u)\;du \\ &=: \; \frac{1}{n} \sum_{i,j=1}^{n} I_1(Y_i, Y_j) \;-\; 2 \sum_{j=1}^{n} I_2(Y_j) \;+\; n\,I_3. \end{aligned} \tag{3.16}$$

Since

$$\int x\,\varphi(x)\;dx = -\varphi(x), \qquad \int x^2\,\varphi(x)\;dx = -x\,\varphi(x) + \Phi(x),$$

where $\int f(x)dx$ denotes an indefinite integral of f, we have in the case $Y_i < Y_j$

$$I_1(Y_i, Y_j) \;=\; 1 - \Phi(Y_j) \;+\; (1 - \Phi(Y_j))\;Y_i\,Y_j \;-\; Y_i\,\varphi(Y_j).$$

Furthermore,

$$\begin{aligned} \int \varphi^2(x)\;dx \;&=\; \frac{1}{2\sqrt{\pi}}\,\Phi(\sqrt{2}\,x), \\ \int x\,\varphi^2(x)\;dx \;&=\; -\frac{1}{2}\,\varphi^2(x), \\ \int x\,\Phi(x)\,\varphi(x)\;dx \;&=\; \frac{1}{2\sqrt{\pi}}\Phi(\sqrt{2}\,x) \;-\; \varphi(x)\Phi(x), \\ \int x^2\,\Phi(x)\,\varphi(x)\;dx \;&=\; \frac{1}{2}\Phi^2(x) \;-\; x\,\varphi(x)\,\Phi(x) \;-\; \frac{1}{2}\varphi^2(x), \end{aligned}$$

and hence

$$I_2(Y_j) \;=\; \frac{1}{2}\,(1-\Phi^2(Y_j)) \;-\; \frac{1}{\sqrt{\pi}}\,Y_j\,(1-\Phi(\sqrt{2}Y_j)) \;+\; \varphi^2(Y_j).$$

Using

$$\begin{aligned}\int \varphi^3(x)\,dx \;&=\; \frac{\sqrt{3}}{6\,\pi}\,\Phi(\sqrt{3}\,x),\\ \int x\,\Phi(x)\,\varphi^2(x)\,dx \;&=\; -\frac{1}{2}\varphi^2(x)\,\Phi(x) \;+\; \frac{\sqrt{3}}{12\,\pi}\Phi(\sqrt{3}\,x),\\ \int x^2\,\Phi^2(x)\,\varphi(x)\,dx \;&=\; \frac{1}{3}\Phi^3(x) \;-\; x\,\varphi(x)\,\Phi^2(x) \;-\; \varphi^2(x)\,\Phi(x) \;+\; \frac{\sqrt{3}}{6\,\pi}\,\Phi(\sqrt{3}\,x),\end{aligned}$$

gives the value $I_3 \;=\; 1/3 + \sqrt{3}/(2\,\pi)$. Plugging all results into (3.16) yields the assertion. ■

Note that, being a function of $Y_1, \ldots, Y_n$, $\tilde{T}_n$ is invariant with respect to transformations of the form $x_j \to ax_j + b$ $(a > 0)$. Consequently, we may assume $\mu = 0$ and $\sigma^2 = 1$ in the following.

Let $\vartheta = (\mu, \sigma^2)$, and write P_ϑ for the measure belonging to the df Φ_{μ,σ^2}. Similarly as in Section 2, we use the Hilbert space $\tilde{L}_2 = L_2(\mathbb{R}, \mathcal{B}, P_\vartheta)$ of functions on $\mathbb{R}$ which are square integrable with respect to P_ϑ to prove the weak convergence of $\tilde{T}_n$. If $(\cdot,\cdot)$ and $\|\cdot\|$ denote the inner product and the norm in $\tilde{L}_2$, one has $\tilde{T}_n = \|\tilde{Z}_n\|^2$; in particular, $\tilde{T}_n$ is a continuous functional of $\tilde{Z}_n$.

The estimator $\hat{\vartheta}_n = (\bar{X}_n, \hat{\sigma}_n^2)$ of ϑ has the representation (2.4), where $l(x,\vartheta) = (x-\mu,\; (x-\mu)^2 - \sigma^2)$. The function

$$\tilde{g}(t,x) \;=\; (x-t)^- \;-\; \Psi_{(0,1)}(t) \;-\; l(x,(0,1))\;\; \nabla_\vartheta \Psi(t,\vartheta)|_{\vartheta=(0,1)}\,,$$

is the analogue of g in (2.6); in view of

$$\frac{\partial \psi_\vartheta(t)}{\partial \mu} \;=\; -\Phi_{\mu,\sigma^2}(t), \qquad \frac{\partial \psi_\vartheta(t)}{\partial \sigma^2} \;=\; \frac{1}{2}\,\varphi_{\mu,\sigma^2}(t),$$

it can be written as

$$\tilde{g}(t,x) \;=\; (x-t)^- \;+\; (x-t)\,\Phi(t) \;-\; \frac{1}{2}(x^2+1)\,\varphi(t).$$

$\tilde{g}(\cdot,\cdot)$ has the following properties:

$$\begin{aligned}\tilde{g}(\cdot,x) \;&\in\; \tilde{L}_2, \quad x \in \mathbb{R},\\ E[\,\tilde{g}(t,X)\,] \;&=\; \int_{-\infty}^{\infty} \tilde{g}(t,x)\,\varphi(x)\,dx \;=\; 0, \quad t \in \mathbb{R},\\ E[\,\|\tilde{g}(\cdot,X)\|^2\,] \;&=\; E\left[\int_{-\infty}^{\infty} \tilde{g}^2(t,X)\,\varphi(t)\,dt\right] \;<\; \infty.\end{aligned}$$

3.2 Theorem *a) Let $\tilde{W}$ denote the centered Gaussian process with covariance function*

$$\begin{aligned}\tilde{k}(s,t) &= E[\tilde{g}(s,X_1)\,\tilde{g}(t,X_1)] \\ &= \Phi(s)(1+st) \,+\, t\,\varphi(s) \,-\, 3\varphi(s)\,\varphi(t)/2 \\ &\quad -t\,\Phi(t)\,\varphi(s) \,-\, s\,\Phi(s)\,\varphi(t) \,-\, \Phi(s)\,\Phi(t)(1+st)\end{aligned}$$

for $s \le t$. Then, under $\mathcal{H}_0$, $\tilde{Z}_n \xrightarrow{\mathcal{D}} \tilde{W}$ in $\tilde{L}_2$.

b) Under $\mathcal{H}_0$, $\tilde{T}_n \xrightarrow{\mathcal{D}} \|\,\tilde{W}\,\|^2$.

Proof: Defining a random variable with values in $\tilde{L}_2$ by

$$\tilde{W}_n \;=\; \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \tilde{g}(\cdot, X_j)$$

and proceeding like in Section 2 yields the weak convergence of $\tilde{Z}_n$. A series expansion of $\varphi_{\hat{\vartheta}_n}(t)$ about ϑ shows that $\tilde{T}_n \;-\; \|\,\tilde{Z}_n\,\|^2 \;=\; o_P(1)$ for $n \to \infty$. Noting that, by a) and the continuous mapping theorem, $\|\tilde{Z}_n\|^2 \xrightarrow{\mathcal{D}} \|\tilde{W}\|^2$, and using Sluzky's Theorem, the assertion follows. ■

Table 4 shows the empirical $(1-\alpha)$-quantiles for different values of the level α and the sample size n. The entries in Table 4 are the 20% trimmed means of 100 Monte-Carlo simulations, each based on 10000 replications; here, we always used $\vartheta = (0,1)$.

Let $\tilde{z}_n(\alpha)$ denote the $(1-\alpha)$-quantile of $\tilde{T}_n$ under $\mathcal{H}_0$. Similarly as in Section 2, it can be shown that the test which rejects the hypothesis of normality if $\tilde{T}_n > \tilde{z}_n(\alpha)$ is consistent against each fixed alternative distribution with positive finite variance.

To judge the power of the test based on $\tilde{T}_n$ for finite samples, we conducted a simulation study with various alternatives to normality. Random numbers are generated by routines of the IMSL-library. Besides the findings for $\tilde{T}_n$, we provide the results of the Anderson-Darling test A^2 and the Shapiro-Wilk test SW, the latter performed with the help of the IMSL routine DSPWLK. The two procedures are the omnibus tests recommended by D'Agostino and Stephens (1986). Groeneveld (1998) cites more recent studies that also indicate that the Shapiro-Wilk test has good general power properties. Certainly, other classical tests as in Gastwirth and Owens (1977) could also be considered.

Table 5 shows the results for the family of stable distributions $S(a,b)$ with characteristic function

$$\varphi(t;a,b) \;=\; \begin{cases} \exp\left(-|t|^a \exp(-ib\pi(1-|1-a|)\,\mathrm{sign}(t)/2)\right) & : \quad a \neq 1 \\ \exp\left(-|t|\,\left(1 + 2ib\log|t|\,\mathrm{sign}(t)/\pi\right)\right) & : \quad a = 1 \end{cases}$$

($-1 \leq b \leq 1, 0 < a \leq 2$). For $b = 0$, the distributions are symmetric; in particular, $S(2,0)$ is the normal distribution $N(0,2)$, and $S(1,0)$ is the Cauchy distribution.

For each of the three tests, the power increases if the stable distributions become more heavy-tailed (i.e. if the characteristic exponent a decreases) resp. if b (and thus the skewness) increases. This result is in contrast to a simulation study of Baringhaus et al. (1989), Table IV, where b had no influence on the power of any of the tests (including Shapiro-Wilk). $\tilde{T}_n$ has a slightly higher percentage of rejection than the Anderson-Darling test and the Shapiro-Wilk test for almost all values of the parameters a and b.

Table 6 and Table 7 show the power of the three tests against symmetric and skewed alternatives, respectively. All distributions are described in more detail in a simulation study of Pearson et al. (1977). S_B and S_U denote distributions from the Johnson system; SC (scale contaminated) and LC (location contaminated) are mixtures of two normal distributions with the same expectation and different variances and with different expectation and the same variance, respectively. The distributions are ordered by their kurtosis value β_2. The first seven distributions in Table 5 and the first five distributions in Table 6 are platykurtic ($\beta_2 < 3$), the others are leptokurtic ($\beta_2 > 3$). The results for the Shapiro-Wilk test differ to some extent from the results of Pearson et al. (1977); this effect may be due to the number of only 200 replications in their study. For distributions with $\beta_2 < 3$, the Shapiro-Wilk test is clearly better than the test based on $\tilde{T}_n$ and the Anderson-Darling test. For symmetric alternatives with $\beta_2 > 3$, the number of rejections is comparable for each of the three tests for sample size $n = 20$; for $n = 50$, $\tilde{T}_n$ and A^2 are often better than SW. For skewed alternatives with $\beta_2 > 3$, the tests behave very similar.

Summarizing the results of the power study, the test for normality based on the integrated distribution function compares favourably with the other tests for heavy-tailed alternatives. In comparison with the Shapiro-Wilk test, power against alternatives with $\beta_2 < 3$ is fairly low; for such distributions, the Shapiro-Wilk test was the best omnibus test in the study of Pearson et al. Presumably, use of an additional parameter in the weight function as in the case of the exponential distribution could improve the power of the idf-test against short-tailed alternatives, but in many situations safeguarding against such distributions may be of minor importance.

Acknowledgements. This work is based on a part of the author's doctoral thesis at the University of Karlsruhe, written under the supervision of Professor Norbert Henze, whose guidance is gratefully appreciated.

References

Araujo, A., and Giné, E. (1980). *The central limit theorem for real and Banach valued random variables.* New York: Wiley.

Baringhaus, L., Danschke, R., and Henze, N. (1989). Recent and classical tests for normality – a comparative study. *Commun. Statist. – Simula.*, *18*(1), 363–379.

Baringhaus, L., Gürtler, N., and Henze, N. (1999). Weighted integral test statistics and components of smooth tests of fit. *Austral. New Zeal. J. Statist.* (To appear)

Baringhaus, L., and Henze, N. (1991). A class of consistent tests for exponentiality based on the empirical Laplace transform. *Ann. Inst. Statist. Math.*, *43*, 551–564.

Baringhaus, L., and Henze, N. (1992). An adaptive omnibus test for exponentiality. *Commun. Statist. – Th. Meth.*, *21*(4), 969–978.

Billingsley, P. (1968). *Convergence of probability measures.* New York: Wiley.

D'Agostino, R., and Stephens, M. (1986). *Goodness-of-fit techniques.* New York: Marcel Dekker, Inc.

Gail, M., and Gastwirth, J. (1978). A scale-free goodness-of-fit test for the exponential distribution based on the Gini statistic. *J. Roy. Statist. Soc. B*, *40*, 350-357.

Gastwirth, J., and Owens, M. (1977). On classical tests of normality. *Biometrika*, *64*, 135-139.

Groeneveld, R. (1998). A class of quantile measures for kurtosis. *Amer. Statist.*, *51*(4), 325–329.

Henze, N., and Nikitin, Y. (1998). *A new approach to goodness–of–fit testing based on the integrated empirical process* (Preprint 98/18). Germany: University of Karlsruhe. (To appear in: J. Nonpar. Statist.)

Klar, B. (1999). Goodness-of-fit tests for discrete models based on the integrated distribution function. *Metrika*, *49*, 53-69.

Klefsjö, B. (1983). Testing exponentiality against HNBUE. *Scand. J. Statist.*, *10*, 65-75.

Lewis, P. (1965). Some results on tests for Poisson processes. *Biometrika*, *52*, 67-77.

Müller, A. (1996). Ordering of risks: A comparative study via stop–loss transforms. *Insurance: Mathematics and Economics*, *17*, 215–222.

Neuhaus, G. (1974). Asymptotic properties of the Cramér–von Mises statistic when parameters are estimated. In J. Hájek (Ed.), *Proceedings of the prague symposium on asymptotic statistics* (pp. 257–297). Praha.

Pearson, E., D'Agostino, R., and Bowman, K. (1977). Tests for departure from normality: Comparison of powers. *Biometrika*, *64*, 231-246.

Pollard, D. (1984). *Convergence of stochastic processes.* New York: Springer.

Widder, D. (1959). *The Laplace transform* (5th ed.). Princeton: Princeton University Press.

	$1-\alpha$				
n	0.5	0.9	0.95	0.975	0.99
10	0.115	0.374	0.494	0.668	1.127
20	0.124	0.452	0.623	0.863	1.397
50	0.135	0.525	0.734	0.990	1.466
100	0.141	0.557	0.775	1.028	1.438
200	0.146	0.573	0.797	1.043	1.405
500	0.150	0.584	0.805	1.040	1.376
1000	0.152	0.587	0.808	1.040	1.363
2000	0.152	0.587	0.805	1.037	1.360

Table 1: Empirical critical values of T_n

		$1-\alpha$				
	n	0.5	0.9	0.95	0.975	0.99
a=1	10	0.021	0.082	0.109	0.136	0.178
	20	0.021	0.086	0.118	0.150	0.196
	50	0.021	0.089	0.122	0.157	0.205
	100	0.021	0.089	0.124	0.160	0.207
	200	0.021	0.090	0.125	0.161	0.210
a=5	10	0.070	0.302	0.408	0.511	0.650
	20	0.068	0.303	0.416	0.531	0.690
	50	0.067	0.304	0.422	0.543	0.708
	100	0.066	0.304	0.423	0.547	0.715
	200	0.066	0.304	0.426	0.549	0.718
a=10	10	0.060	0.248	0.321	0.394	0.570
	20	0.058	0.256	0.349	0.444	0.585
	50	0.057	0.262	0.364	0.469	0.615
	100	0.057	0.263	0.366	0.472	0.619
	200	0.056	0.264	0.368	0.477	0.630
a=20	10	0.042	0.125	0.197	0.301	0.462
	20	0.039	0.165	0.217	0.294	0.444
	50	0.038	0.174	0.241	0.312	0.421
	100	0.038	0.178	0.248	0.322	0.427
	200	0.038	0.179	0.250	0.326	0.429

Table 2: Empirical critical values of $T_{n,a}$ for different values of a

distribution	φ_0	φ_1	φ_5	φ_{10}	φ_{20}	$\varphi_{1,10}$	$\varphi_{0,20}$	$\varphi_{0,5,20}$	W^2	A^2
$\Gamma(0.4)$	66	72	84	89	92	88 (87)	91	90	75	90
$\Gamma(0.6)$	30	31	41	47	53	46 (44)	52	50	32	47
$\Gamma(0.8)$	11	10	12	15	18	15 (14)	18	17	9	15
$\Gamma(1.0)$	5.0	5.0	5.0	5.0	5.0	5.0 (4.2)	5.0	5.0	4.9	5.1
$\Gamma(1.5)$	10	18	22	18	8	16 (14)	2	11	19	17
$\Gamma(2.0)$	28	45	54	47	24	44 (40)	9	35	48	46
$\Gamma(2.4)$	46	66	76	69	40	67 (62)	19	57	69	68
$\Gamma(3.0)$	70	87	93	89	63	88 (85)	38	82	89	89
$W(0.4)$	96	98	99	100	100	100 (99)	100	100	98	100
$W(0.6)$	67	70	77	80	81	81 (79)	83	82	69	81
$W(0.8)$	23	22	24	26	29	28 (26)	32	31	20	27
$W(1.4)$	21	34	36	27	12	30 (26)	6	20	35	31
$W(1.6)$	45	62	63	49	25	56 (51)	17	43	61	58
$W(2.0)$	86	94	93	83	53	92 (89)	56	83	93	92
$LN(0.6)$	54	76	95	96	83	92 (89)	28	85	89	90
$LN(0.8)$	15	23	42	47	28	35 (31)	6	27	34	34
$LN(1.0)$	19	15	11	10	5	15 (13)	15	16	15	14
$LN(1.2)$	38	33	20	14	7	30 (28)	34	33	28	27
$LN(1.5)$	68	66	59	53	44	65 (64)	66	66	61	63
$U(0,1)$	68	73	45	25	10	62 (58)	35	33	67	63
HN	14	22	18	11	5	16 (14)	3	9	21	17
HC	72	69	59	52	46	67 (66)	70	69	63	64
χ^2_1	46	50	63	70	75	69 (67)	74	73	52	71

Table 3: Empirical power of different tests for exponentiality, $\alpha = 0.05$, $n = 20$, 100000 replications

n	$1-\alpha$ 0.5	0.9	0.95	0.975	0.99
10	.0086	.0233	.0300	.0371	.0468
15	.0085	.0239	.0310	.0384	.0485
20	.0085	.0244	.0316	.0392	.0494
30	.0085	.0247	.0321	.0397	.0502
50	.0085	.0249	.0324	.0401	.0507
100	.0084	.0250	.0325	.0403	.0507
200	.0084	.0251	.0327	.0405	.0512
500	.0084	.0252	.0328	.0406	.0509
1000	.0084	.0252	.0328	.0406	.0511

Table 4: Empirical critical values of $\tilde{T}_n$

statistic	$\tilde{T}_n$			SW			A^2		
b	0.0	0.5	1.0	0.0	0.5	1.0	0.0	0.5	1.0
$a = 2.0$	5.0	5.0	5.0	5.1	5.1	5.0	4.9	5.0	4.9
$a = 1.8$	38.9	41.1	47.8	37.5	39.7	47.0	36.2	37.8	43.0
$a = 1.6$	69.1	72.6	82.1	65.8	69.6	81.3	67.4	70.3	78.3
$a = 1.4$	88.7	90.9	96.6	85.7	88.4	96.4	88.1	90.2	95.1
$a = 1.2$	97.3	98.1	99.6	95.9	97.1	99.6	97.4	98.1	99.5
$a = 1.0$	99.7	99.8	100.0	99.4	99.6	100.0	99.8	99.8	100.0

Table 5: Empirical power against stable distributions, $n = 50$, $\alpha = 0.05$, 100000 replications

statistic		$\tilde{T}_n$		SW		A^2	
distribution	β_2	$n=20$	$n=50$	$n=20$	$n=50$	$n=20$	$n=50$
$S_B(0,0.5)$	1.63	22.4	78.3	44.7	99.2	37.1	90.1
Tukey(1.5)	1.75	12.3	51.8	25.6	93.3	20.9	67.7
Beta(1,1)	1.80	10.0	42.5	20.5	88.3	17.3	58.1
$S_B(0,0.707)$	1.87	8.3	33.2	15.0	75.0	13.3	44.9
Tukey(0.7)	1.92	6.5	24.8	11.6	66.1	10.7	35.3
Tukey(3.0)	2.06	3.9	10.2	6.6	44.3	6.1	16.6
Beta(2,2)	2.14	3.9	9.3	5.4	26.6	5.8	13.2
$S_U(0,3)$	3.53	7.6	9.2	7.5	8.0	7.2	8.5
t_{10}	4.00	9.4	13.4	9.6	11.6	8.9	11.9
Logistic	4.20	11.4	17.4	11.4	14.1	10.6	15.9
$S_U(0,2)$	4.51	12.2	19.0	12.3	15.6	11.4	17.3
Tukey(10)	5.38	74.0	99.0	80.2	99.4	91.2	100.0
Laplace	6.00	26.8	52.1	25.8	41.2	27.3	54.4
SC(0.20,3)	7.54	36.2	67.5	36.8	60.8	34.8	65.3
SC(0.05,3)	7.65	18.2	32.6	19.1	32.7	17.2	28.9
SC(0.10,3)	8.33	27.1	50.0	28.2	47.7	25.5	46.2
SC(0.20,5)	11.22	70.1	96.7	70.1	95.2	70.9	96.4
SC(0.20,7)	12.84	84.6	99.5	84.8	99.3	85.4	99.4
SC(0.10,5)	16.45	53.6	84.3	54.5	83.4	52.1	82.4
SC(0.05,5)	19.96	34.8	62.0	35.7	62.2	33.6	59.1
SC(0.10,7)	21.49	66.5	93.1	67.2	92.9	65.6	92.3
SC(0.05,7)	31.40	44.2	74.3	45.0	74.8	43.1	72.1
$S_U(0,1)$	36.19	42.6	76.0	42.0	68.8	42.3	75.8
$S_U(0,0.9)$	82.08	50.5	84.5	49.8	78.6	50.9	85.1
t_4	∞	23.6	44.2	23.6	38.6	22.4	42.2
t_2	∞	53.0	85.9	52.5	81.3	52.8	85.7
t_1	∞	86.7	99.7	86.3	99.4	88.2	99.7

Table 6: Empirical power against symmetric distributions, $n = 20$ and $n = 50$, $\alpha = 0.05$, 100000 replications

statistic			$\tilde{T}_n$		SW		A^2	
distribution	$\sqrt{\beta_1}$	β_2	$n=20$	$n=50$	$n=20$	$n=50$	$n=20$	$n=50$
$S_B(0.533, 0.5)$	0.65	2.13	56.1	97.5	72.5	100.0	64.9	99.1
Beta(3,2)	-0.29	2.36	6.5	16.2	7.4	28.8	7.3	17.6
Beta(2,1)	-0.57	2.40	24.1	67.9	30.5	90.0	26.3	72.0
$S_B(1, 2)$	0.28	2.77	6.2	10.4	6.1	11.9	6.0	9.5
$S_B(1, 1)$	0.73	2.91	28.0	71.3	30.0	83.5	26.5	69.0
LC(0.20,3)	0.68	3.09	29.0	69.3	25.8	62.8	26.6	64.7
LC(0.20,5)	1.07	3.16	88.6	100.0	62.8	100.0	88.1	100.0
LC(0.20,7)	1.25	3.20	98.7	100.0	98.7	100.0	98.6	100.0
Weibull(2)	0.63	3.25	14.5	36.1	15.0	43.7	13.4	31.1
LC(0.10,3)	0.80	4.02	26.0	56.7	24.6	51.8	22.5	50.8
χ^2_{10}	0.89	4.20	23.9	56.1	24.1	59.2	21.0	48.5
LC(0.05,3)	0.68	4.35	17.9	35.1	18.2	33.3	16.0	29.7
LC(0.10,5)	1.54	5.45	75.6	97.8	76.2	98.0	72.7	97.1
$S_U(-1, 2)$	0.87	5.59	21.2	43.3	20.6	39.0	18.9	37.4
χ^2_4	1.41	6.00	50.2	91.4	52.8	95.6	46.6	89.3
LC(0.10,7)	1.96	6.40	87.7	99.5	88.0	99.5	87.4	99.4
LC(0.05,5)	1.65	7.44	52.4	83.6	54.3	85.3	49.2	80.1
χ^2_2	2.00	9.00	78.1	99.6	83.5	100.0	77.6	99.7
LC(0.05,7)	2.42	10.36	65.2	92.3	65.6	92.6	64.5	91.6
χ^2_1	2.83	15.00	95.7	100.0	98.4	100.0	97.1	100.0
Weibull(0.5)	6.62	87.72	99.5	100.0	99.9	100.0	99.8	100.0
$S_U(1, 1)$	-5.30	93.40	73.6	98.0	72.2	96.8	71.1	97.4
Lognormal(0,1,0)	6.18	113.94	91.2	100.0	93.2	100.0	90.5	100.0

Table 7: Empirical power against skewed distributions, $n = 20$ and $n = 50$, $\alpha = 0.05$, 100000 replications